Response to Reviewers Comments

We thank the associate editor, editor and two anonymous reviewers for their thoughtful and constructive comments and suggestions, which significantly help us to improve the quality of the manuscript. In this revised manuscript, we have tried our best as much as possible to address all concerns and have revised the manuscript accordingly. Below, we indicate the original comment of the respective reviewer in blue and our point-to-point response is denoted in black.

Before addressing the comments, we would like to express our sincere gratitude to the reviewers for their exceptionally informative, constructive, and detailed comments.

Reviewer #1 Evaluations:

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.

Response: We appreciated tremendously your thoughtful comments and positive review on our article. According to your nice suggestions, the concerns raised by you have been addressed in this revision as much as we can. Please see the following response and the revised manuscript for more details.

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.

Response: Per your suggestion, the following sentence has been supplemented in the end of this paragraph:

"…Particularly, a higher terrain gradient or a more unstable troposphere generally lead to a lower PBLH in ERA5 reanalysis."

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.

Response: We sincerely thank the reviewer for his or her careful reading. In the abstract, we give the full name of GLDAS (Global Land Data Assimilation System). Also, in Section 2.2, the description on GLDAS has been added. In the introduction section, a new phrase has been supplemented to illustrate the GLDAS product:

"…The biases between PBLH retrieved from the ERA5 and radiosonde could be represented by the land properties, near-surface meteorological conditions, among others, and further be optimized via a machine learning model. The GLDAS incorporates satellite- and ground-based observations and produces a global, high-resolution product regarding land states and fluxes (Rodell et al., 2004)…".

Line 154: please clarify if the interpolation is based on altitude or elevation.

Response: The interpolation is based on altitude, which has been specified in Section 2.1 as follows: "all the original soundings were evenly interpolated to the profiles with a vertical resolution of 10 m by cubic spline interpolation."

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.

Response: Thanks for your careful checks. The statement has been modified to "…Australia have a rich geographic coverage…"

Line 173: Any reference for the definition of LST?

Response: The references have been added.

Line 207: how did the authors match the stational PBLH and gridded PBLH in the comparison?

Response: The collocation procedure between radiosonde and ERA5 grid follows Guo et al, (2021), which can be descripted as follows:

(1) The grid should contain the radiosonde station.

(2) The UTC time (hour) of grid product and radiosonde stay the same.

The above descriptions have been incorporated into the revised manuscript.

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…)?

Response: All the data were used in the model training stage, and the corresponding local time of UTC was used as the input. We also tested a solution based on 24 different models that were trained by variables on different UTCs. However, the raised issue was that the boundary between different local time may has obvious gaps, due to different models used.

The above discussions have been well incorporated into the revised manuscript.

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?

Response: The study is the extension of our recent work (i.e., Guo et al., 2021), in which the bias between radiosonde and reanalysis has been revealed, especially in the daytime. Then the question that needs to be addressed is how to build a more robust and accurate seamless PBLH dataset. We believe a more robust and accurate PBLH dataset could be valuable for the study of some fields, such as climate model and air pollution.

Technical corrections:

Line 99 and 116: the definition of ERA-5 should be moved ahead.

Response: Amended as suggested.

Please keep it consistent by using either ERA5 or ERA-5 in the whole manuscript.

Response: Done as suggested.

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.

Response: As suggested, the descriptions about Table 2 in the main text has been modified to be:

"…We further perform correlation analyses between the aforementioned variables and PBLH biases between radiosonde and ERA5 reanalysis, and the statistical results are shown in Table 2…".

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.

Response: Amended as suggested.

Reviewer #2 Evaluations:

The authors have compared radiosonde-based PBL height estimates with PBL heights derived from the ERA5 reanalysis for over 300 available land stations, showing a significant bias in ERA5 PBL heights. A machine learning routine is developed to predict the ERA5 PBL height bias based on numerous input parameters, and this bias is subtracted from the ERA5 PBL height to produce a corrected dataset. This produces an immediately useful and relevant dataset that can be applied in many future studies. The work is novel, well-constructed, and succinctly explained in the paper. There are a few non-structural fixes that could improve the manuscript, but no major issues with the work, so I would only call these minor revisions.

Response: Thank you for your thoughtful comments and valuable suggestions that are crucial for us to improve the quality of our manuscript. The raised concerns have been modified as much as we can.

Notes:

Line 203: can you add some detail on what you mean by the 'second level'?

Response: As suggested, the statement has been rephrased as:

"…(1) Ri(z) in Eq. (1) exceeds 0.25, where z is the second level of radiosonde measurement…".

In equation 2, it appears that PBLH-M and PBLH-E are mis-formatted as PBLH – M and PBLH – E (where M and E are variables being subtracted) This is probably a formatting error, but is initially very confusing.

Response: The modification has been made. Thank you for providing these insights.

Line 260: I worry that randomly dividing the data can cause an issue if certain geographic regions are underrepresented in the training data. I would recommend

dividing your stations into specific regions (for example: valleys, mountains, coastal, continental, tropical, polar…) and ensuring that a subset for each region is then randomly drawn for each training/validation pool. An easier solution may be just to show that the randomly selected data already chosen for training represents these differing types of regions using a map and/or histogram.

Response: We have reflected this comment by re-assessing the distribution of data over specific regions. For a random process, the index of raw dataset (increasing from 0 to n with an increment of 1, where n is the total number of dataset) is expected to be randomly perturbed. The raw index and the randomly perturbed one are assumed to be poorly correlated. As shown in Figure A, the relationship between the raw data index in training set and the randomly perturbed index over mountain (a), plain (b), tropical island (c), and artic pole (d) regions indicate that the training dataset over different region are nearly randomly distributed. This finding indicates that arrays over different regions closes to be randomly selected.

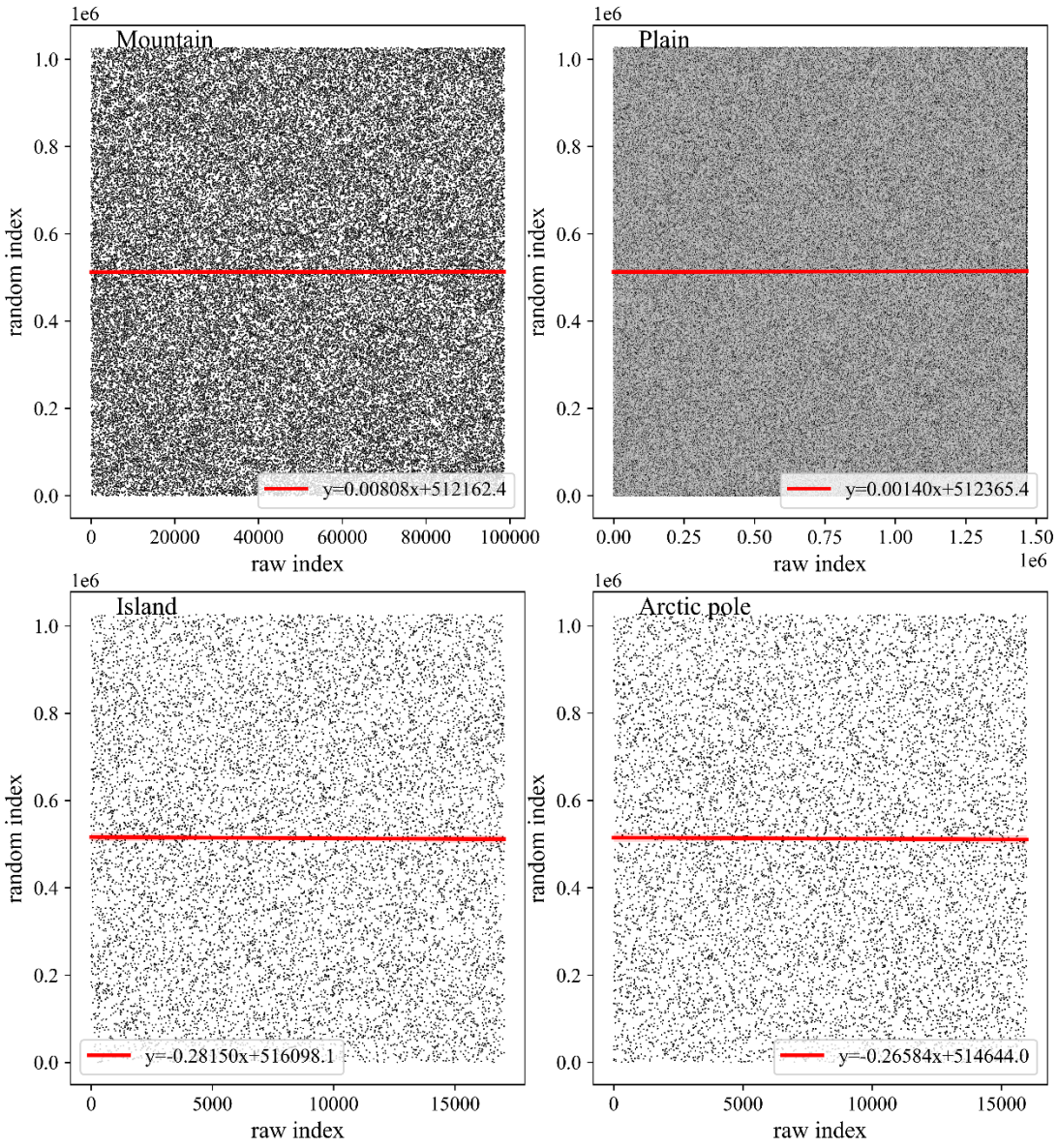

Figure A. the relationship between raw data index in training set and random index over mountain (a), plain (b), island (c), and artic pole (d) regions.

Figure 2: Dividing by calendar season for stations on both sides of the equator is not recommended, since you are lumping winter with summer, autumn with spring, etc… It would be better to combine similar seasons, so that southern hemisphere DJF is combined with northern hemisphere JJA, etc… This would better illustrate seasonal biases.

Response: We agree with you and have incorporated this suggestion throughout our paper. In the revised version, the mean bias in summer season raised from 94.15 m to

95.15 m, and from 72.46 m to 71.23 m in winter season. For the merged dataset, it is from -8.14 m to -7.88 m in summer, and it is from -0.56 m to -0.87 m in winter.

Figure 8: the panels are too small for a meaningful comparison between PBLH-R and PBLH-M (comparing the dots to the shading). I recommend making larger maps available as supporting material, or showing this comparison some other way

Response: Amended as suggested. Only 0000 UTC and 0012 UTC have been kept in the revised version. And we hope these figures in this revised manuscript could be acceptable for you.