Response to Referees' Comments

Response to Reviewer #2:

While traditional machine learning methodologies (e.g., Random Forest) have been widely used to estimate PBLH, most studies heavily rely on specific remote sensing instruments or focuses on limited time-period or specific region of interest. More importantly, lack of enough physical explanation is another concern. To address this issue, this manuscript introduces a multi-structure deep neural network (DNN) model that is used to generate yield a robust 27-year PBLH dataset over the Southern Great Plains from 1994 to 2020. Through leveraging a variety of meteorological data, independent of remote sensing instruments, this model yielded an PBLH dataset over the SGP with robust accuracy, consistently yielding lower bias values across various conditions and datasets. Besides, the generalizability of this model to different geographic regions and climate zones are explored, exhibiting high potential and less uncertainties in terms of seasonal, diurnal variability. Overall, this manuscript is well organized with clear enough logic, I would like to offer the following suggestions for further improvement:

Response: We appreciate the reviewer's positive and comprehensive comments on our work. Following these insights, we have refined our manuscript to enhance its clarity. We have carefully considered and addressed all comments and concerns raised by the reviewer in this revision. Our detailed responses to each point are provided below.

Major Comments:

Introduction: Except for the lidar systems, the authors seem to ignore the radar wind profiler, which provides the direct measurements of turbulence in the atmosphere and thus affords the retrievals of PBLH. A variety of algorithms or methods in the literature have been proposed to accomplish this task. Therefore, the authors can argue the current literature review in this regard.

Response: Thanks for the helpful suggestion. We acknowledge the importance of radar wind profilers in measuring atmospheric turbulence and their utility in PBLH retrieval and acknowledge the relevant studies in the introduction as follows:

"In addition, wind profilers can estimate the PBLH using algorithms that analyze the signal-to-noise ratio from wind profiler data (Molod et al. 2015; Solanki et al. 2022; Liu et al. 2019; Salmun et al. 2023; Bianco and Wilczak 2002; Bianco et al. 2008; Tao et al. 2021)."

References:

Solanki, R., Guo, J., Lv, Y., Zhang, J., Wu, J., Tong, B., & Li, J. (2022). Elucidating the atmospheric boundary layer turbulence by combining UHF radar wind profiler and radiosonde measurements over urban area of Beijing. Urban Climate, 43, 101151.

Liu, B., Ma, Y., Guo, J., Gong, W., Zhang, Y., Mao, F., ... & Shi, Y. (2019). Boundary layer heights as derived from ground-based Radar wind profiler in Beijing. IEEE Transactions on Geoscience and Remote Sensing, 57(10), 8095-8104.

Molod, A., Salmun, H., and Dempsey, M., 2015: Estimating Planetary Boundary Layer Heights from NOAA Profiler Network Wind Profiler Data, J. Atmos. Ocean. Tech., 32, 1545–1561, https://doi.org/10.1175/JTECH-D-14-00155.1.

Salmun, H., Josephs, H., & Molod, A. (2023). GRWP-PBLH: Global Radar Wind Profiler Planetary Boundary Layer Height Data. Bulletin of the American Meteorological Society, 104(5), E1044-E1057.

Bianco, L., Wilczak, J. M., & White, A. B. (2008). Convective boundary layer depth estimation from wind profilers: Statistical comparison between an automated algorithm and expert estimations. Journal of Atmospheric and Oceanic Technology, 25(8), 1397-1413.

Bianco, L., and J. M. Wilczak, 2002: Convective boundary layer depth: Improved measurements by Doppler radar wind profiler using fuzzy logic methods. J. Atmos. Oceanic Technol., 19, 1745–1758, https://doi.org/10.1175/1520-0426(2002)019,1745: CBLDIM.2.0.CO;2.

Tao, C., Y. Zhang, Q. Tang, H. Ma, V. P. Ghate, S. Tang, S. Xie, and J. A. Santanello, 2021: Land–Atmosphere Coupling at the U.S. Southern Great Plains: A Comparison on Local Convective Regimes between ARM Observations, Reanalysis, and Climate Model Simulations. J. Hydrometeor., 22, 463–481, https://doi.org/10.1175/JHM-D-20-0078.1.

Line 89-102: The reason for the selection of multi-structure deep neural network (DNN) in the retrieval of PBLH lacks necessary literature support. Are there similar models constructed based on DNN? If any, how is the performance compared with other models or methods? This should be clarified and some necessary references are required to be cited here.

Response: Response: We appreciate the comment regarding the need for a clearer explanation for the choice of the deep learning model. In response, we have added a detailed discussion on the introduction as follows:

"We aim to leverage and integrate the comprehensive field observations (i.e., radiosonde and remote sensing techniques) to develop a deep learning model for direct PBLH estimation from conventional meteorological data. This strategy circumvents the limitations of relying on particular remote sensing technologies. Furthermore, our model employs an advanced deep neural network (DNN) approach (Sze et al. 2017; Schmidhuber, 2015; Nielsen, 2015; Pang et al. 2020), diverging from traditional ML methods like random forest. This deep learning model utilizes ensemble techniques, constructing arrays of various structures and using their average for the final estimation. This approach method provides particular advantages in the context of complex and nonlinear processes (Ganaie et al. 2022; Mohammed and Kora. 2023). Ensemble DNN with multi-structure designs shows very strong flexibility and robustness, so it relatively performs better and has high stability across a wide range of conditions (Xue et al. 2020; Dong et al. 2020). This facilitates the adaptability of DNN as a tool for PBLH estimation, which can be utilized under different scenarios and locations."

References:

Mohammed, A., & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University-Computer and Information Sciences, 35(2), 757-774.

Xue, W., Dai, X., & Liu, L. (2020). Remote sensing scene classification based on multi-structure deep features fusion. IEEE Access, 8, 28746-28755.

Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2020). A survey on ensemble learning. Frontiers of Computer Science, 14, 241-258.

Ganaie, M. A., Hu, M., Malik, A. K., Tanveer, M., & Suganthan, P. N. (2022). Ensemble deep learning: A review. Engineering Applications of Artificial Intelligence, 115, 105151.

Sze, V., Chen, Y.H., Yang, T.J. and Emer, J.S., 2017. Efficient processing of deep neural networks: A tutorial and survey. Proceedings of the IEEE, 105(12), pp.2295-2329.

Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural networks, 61, pp.85-117.

Nielsen, M.A., 2015. Neural networks and deep learning (Vol. 25, pp. 15-24). San Francisco, CA, USA: Determination press.
Pang, B., Nijkamp, E. and Wu, Y.N., 2020. Deep learning with tensorflow: A review. Journal of Educational and Behavioral Statistics, 45(2), pp.227-248.

Specific comments:

1. *Line 50: "it" is redundant and can be removed.*

Response: Thanks for catching this typo. We have revised it.

2. *Line 54: "climate models" -> "climate projections"*

Response: Revised as suggested.

3. *Line 83: "PBL heights using thermodynamic profiles or backscatter profiles from Lidar or Atmospheric Emitted Radiance Interferometer (AERI)" -> "PBLH using thermodynamic profiles Atmospheric Emitted Radiance Interferometer (AERI) or using backscatter profiles from Lidar".*

Response: We have rephrased the sentence as suggested.

4. *Line 87: "Moreover," -> "For example, "*

Response: The suggestion has been incorporated.

5. *Line 89: "marked progress" -> "made great progress"*

Response: Revised as suggested.

6. *Line 136: Some words are missing between "latent heat fluxes" and "the surface instruments"*

Response: We revised it as "latent heat fluxes from the surface instruments".

7. *Line 365-367: is there any supporting material for the threshold used to define low cloud (maximum cloud fraction between 0-4 km exceeding 1%)?*

Response: The ECWMF also use 1% as the threshold to identify the cloud base height. Specifically, cloud base is calculated by searching from the second lowest model level upwards, to the height of the level where cloud fraction becomes greater than 1% and condensate content greater than 1.E^{-6} kg kg^{-1} (Hersbach et al. 2023). We include this reference in the manuscript.

References:

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023): ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: https://doi.org/10.24381/cds.adbb2d47