Reviewer #3 (Response to Reviewer)

The study by Che et al. represents a significant advancement in the field of urban geography and Earth observations. The development of the 3D-GloBFP dataset is a groundbreaking achievement that fills a critical gap in the availability of global, high-resolution, and accurate building height information. The methodology employed is innovative and rigorous, resulting in a dataset with exceptional performance and reliability. The implications and applications of the 3D-GloBFP dataset are vast, spanning from climate modeling to sustainable development policies. Overall, this study deserves high praise for its contributions to the scientific community and beyond. However, to further strengthen the research, I would suggest addressing the following minor issues:

Response: thank you very much for your positive comments. We have carefully revised our manuscript based on your comments. We provided a detailed response to your comments below.

Comment #1: The division of the 33 regions mentioned in the paper is not particularly clear and requires a brief elaboration or a reference to the specific figure where they are illustrated.

Response: thank you for your suggestion. We included Figure 3, which shows the distribution of subregions and the number of training samples within each:

"We divided the globe into 33 regions and developed the building height estimation model for each region, considering the non-uniform spatial distribution of samples and the heterogeneous building heights. Firstly, we divided the globe into 13 regions based on geographic spatial distance and regional development levels to ensure that each region has enough samples to train effective models. For instance, the Central and West Asian countries were considered as a single region for model training and estimation with 40040 training samples. However, given China's complex urban 3D structure and significant building heterogeneity (Wu et al., 2023), we further divided China into 21 regions. We built a separate height regression model for each region to ensure the effectiveness of the height estimation. For instance, considering the inadequacy of samples in Northwest China, we considered the provinces in Northwest as a single region with 8050 training samples for model training. Additionally, we considered the Beijing-Tianjin-Hebei, Yangtze River Delta, and Peral River Delta urban agglomerations as three separate regions due to the comparable economic levels and population size." (page 8-9, line 186-195)

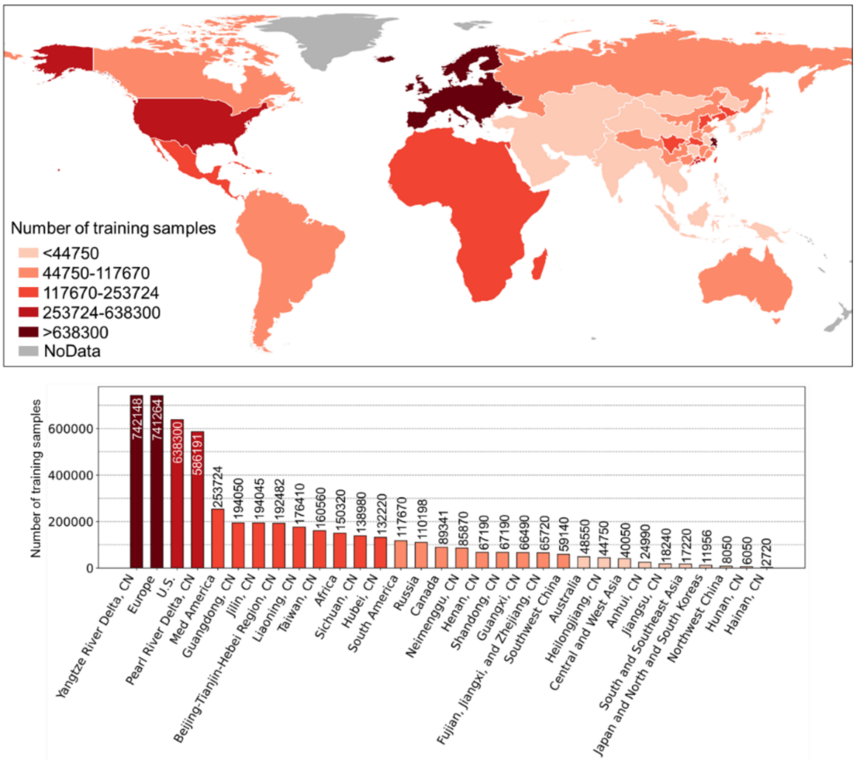

__Figure 3__. Distribution of subregions.

Reference:

Wu, W.-B., Ma, J., Banzhaf, E., Meadows, M. E., Yu, Z.-W., Guo, F.-X., Sengupta, D., Cai, X.-X., and Zhao, B.: A first Chinese building height estimate at 10 m resolution (CNBH-10 m) using multi-source earth observations and machine learning, Remote Sensing of Environment, 291, 113578, https://doi.org/10.1016/j.rse.2023.113578, 2023.

Comment #2: The first sentence of several paragraphs in the result section introduces the methodology, but it is recommended to revise them to summarize the findings of the current paragraph instead. You may place the corresponding methodology in the method section, or write it after the figure caption.

Response: thank you for your comments. We checked the first sentences of all the paragraphs in the result section and summarized the findings at the beginning of paragraphs, including:

"Our 3D-GloBFP showed the most similar numerical distribution patterns to the reference heights across the United States, China, and Europe (Fig. 7)." (page 16, line 343-344)

"In the respective regional comparisons, first, we found that our 3D-GloBFP outperforms other building-scale height datasets in the US." (page 17, line 354-355)

"Second, our 3D-GloBFP is similar to the reference height in terms of distribution and spatial patterns in China." (page 18, line 379)

The description of methodology was reorganized in method section.

Comment #3: While most of the results presented in this paper appropriately utilize the present simple tense, there are some instances where the past tense has been used inappropriately, see for example, line 149.

Response: thank you for the comment. We checked the tense throughout the paper. We used the simple past tense to describe the completed process and present the research findings. And the simple present tense is used to describe the contents of figures and tables. The language has been polished by professional editing. The revised sentences in the article include:

"The validation results with interpreted heights from Google Earth Street Views indicated the estimated results are consistent with the reference heights in the metropolitans of countries around the world, particularly for those landmark buildings." (page 12, line 273)

"Our estimated building heights provided more details of urban morphology and show more accurate results compared to the other four existing global datasets." (page 14, line 289)

"The building volume and area of representative cities varied significantly across different regions worldwide." (page 22, line 438)

Comment #4: Expanding the discussion of challenges and future work would provide valuable insights into the dataset's limitations and potential for growth.

Response: thank you for your suggestion. First, improving the accuracy of building height estimations across various scenarios remains a key challenge. For instance, in densely built urban cores with more high-rise buildings, the complex structures and dense spatial patterns can lead to an underestimation of high-rise buildings. Furthermore, generating three-dimensional building footprints with temporal information is also a challenge. The dataset reflects the building structure for only a single year, whereas building heights may vary over time due to new construction and the demolition of existing buildings. In future work, we plan to use LiDAR datasets and filed survey data to improve the accuracy of building height estimation, especially for high-rise buildings. Also, we aim to develop multi-temporal building datasets to capture the dynamic changes in three-dimensional urban landscapes.

We have revised the manuscript to include a more detailed examination of the challenges and the future work:

"The resolution of coarse-resolution remote sensing dataset (e.g., DSM with a 30 m resolution and nighttime light with a 463.83 m resolution) make it difficult to capture the heterogeneity features of super tall buildings, especially in densely built urban cores. Moreover, height and material of high-rise buildings, as well as the side-looking scene illumination Sentinel sensor, can cause complex multipath effects, complicating radar signal propagation, and ultimately affecting the accuracy of height estimations (Frantz et al., 2021; Stilla et al., 2003)." (page 11, line 260-264)

"Second, the current version of 3D-GloBFP shows relatively lower accuracy in areas with limited building height samples (i.e., suburb of South America). Integrating additional data (i.e., ground survey data and LiDAR datasets) to create more representative samples can enhance the accuracy of building height estimation. Additionally, the current version of 3D-GloBFP represents building height of a single year (i.e., 2020), as the model inputs (i.e., multi-source datasets) were collected around that time. This temporal limitation restricts the dataset's ability to reflect changes over time. We are also committed to producing 3D building datasets with temporal information to capture the dynamic changes of urban landscape." (page 23-24, line 462-468)

Comment #5: Identifying gaps in current knowledge, discussing opportunities for integrating additional data sources, and outlining plans for updating and maintaining the dataset over time would demonstrate the authors' commitment to ongoing improvement and research.

Response: thank you for your insightful comments. The limitations of the dataset include the accuracy in areas with limited samples and the coverage in some countries. First, areas with fewer samples, primarily undeveloped areas or cities with inadequate data show lower precision. Second, spatial coverage of three-dimensional building height data is incomplete in some regions due to the lack of detailed building boundary extraction, particularly in rural areas of certain countries. To address these issues, we will integrate additional data sources, including ground survey data, LiDAR, and other publicly available remote sensing data to enhance the accuracy of building height estimation models. Additionally, we will regularly update the dataset as more building boundary datasets become available.

We have summarized the gaps and plans for integrating datasets and updating the proposed dataset to enhance the manuscript:

"4.6 Limitations and future work

While this study provides valuable insights, several limitations must be acknowledged. First, the coverage is limited in certain regions, leading to tiled spatial gaps within some countries. These gaps are due to the limited coverage of Microsoft building footprints at the time of data creation. As more building footprint datasets become available, we will continue to update and enhance 3D-GloBFP using comprehensive open-source data. Second, the current version of 3D-GloBFP shows relatively lower accuracy in areas with limited building height samples (i.e., suburb of South America). Integrating additional data (i.e., ground survey data and LiDAR datasets) to create more representative samples can enhance the accuracy of building height estimation." (page 23, line 458-464)

By addressing these points, the authors can enhance the readability and presentation of their work, thereby ensuring that the 3D-GloBFP dataset becomes an invaluable asset to the scientific community.