Reviewer #2 (Response to Reviewer)

This paper presents the world's first three-dimensional building footprint dataset, 3D-GloBFP, which integrates multi-source remote sensing data and various reference building height data. By employing machine learning methods, it generates high-precision global building height data. This dataset holds significant importance across multiple application domains, including urban planning, environmental monitoring, disaster management, and energy consumption analysis. The research demonstrates substantial promise and value, providing a crucial foundation for the acquisition and application of global 3D building data. I believe that 3D-GloBFP is an indispensable foundational dataset for urban research. During the review process, I identified several areas that require further clarification and improvement.

Response: thank you very much for your constructive comments. We have carefully revised our manuscript and provided further responses for your further review.

Comment #1: In Section 3.2.1 "Division of Subregions," the information of training and testing samples (e.g., the total amount) for each sub-region should be explicitly provided. Clearly specifying the selection criteria and distribution of these samples will help readers better understand the process of model training and validation.

Response: thank you for pointing out this issue. We enhanced the description of our training and testing datasets, including adding the information of distribution and the number of samples in each subregion:

"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)

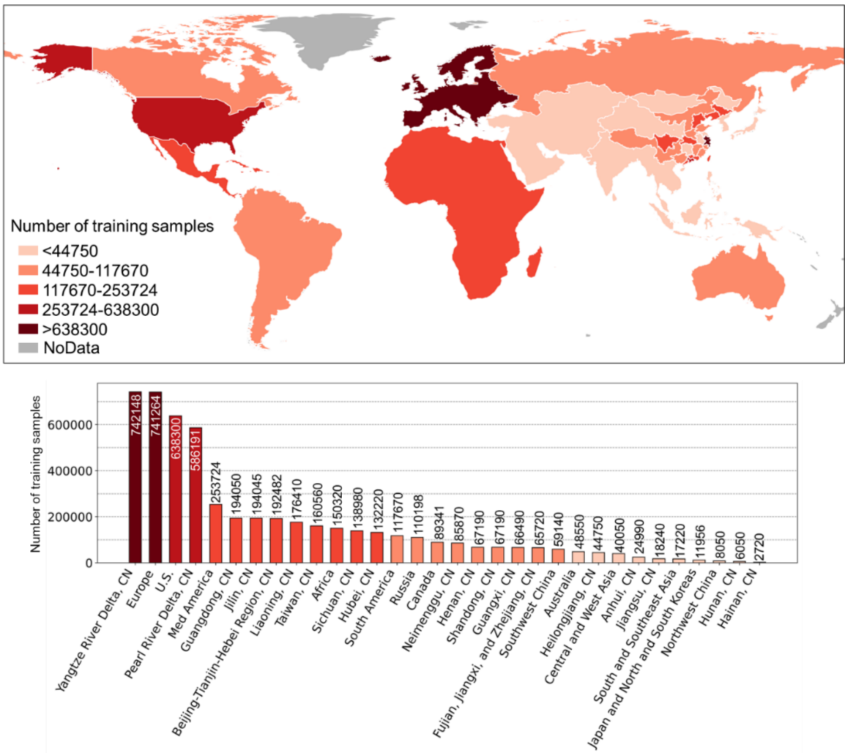

***Figure 3**. Distribution of subregions.*

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: I noticed that there are some missing tiles in Ghana and incomplete regions in Guangdong, please ensure that the dataset is complete globally.

Response: thank you for raising this concern. We completed the building footprints with height attributes of our datasets. The updated results of Ghana and Guangdong are showed in Fig R3 and Fig R4.

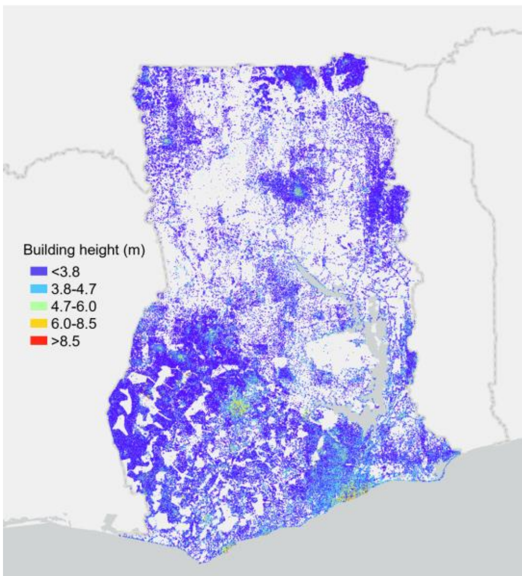

Figure R3. 3D-GloBFP in Ghana.

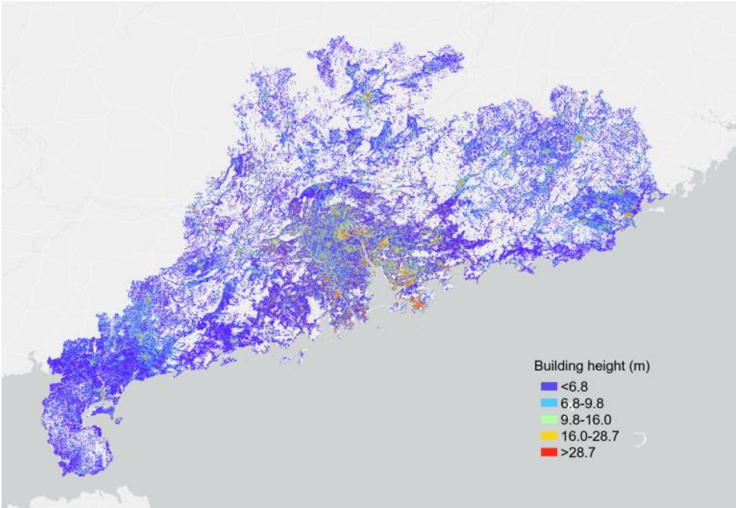

Figure R4. 3D-GloBFP in Guangdong.

Comment #3: In China, reference building heights used as training samples were mostly concentrated in city centers. Please clarify the accuracy of the model's estimates for building heights in different urban areas (e.g., urban fringe). This is crucial for validating the model's applicability in diverse environments.

Response: thank you for the great question. We evaluated the model performance across different urban regions, including city center, suburban area, and urban fringe. The results show that the estimated results are consistent with reference data in various city environments.

To assess the model performance in different regions, we divided the cities into several rings based on road network, and calculated the R^2 and RMSE between estimated and reference height. It is worthy to note that these reference samples were not used for model training. We selected Shanghai, Guangzhou-Foshan, and Chengdu as samples cities due to their heterogeneity of building heights from city center to fringe areas.

The model shows relatively high reliability in estimating building heights across different rings(**Fig. R5**). In Shanghai, the model performs well in the Middle Ring (R^2 = 0.88, RMSE = 8.91 m), indicating the model can capture the height distributions in city center. The R^2 of suburban and outer ring areas is 0.90 with an RMSE of approximately 13.7 m. Although the RMSE is relatively higher in suburban ring and outer areas, the estimated results generally agree well with reference height in different regions. In Guangzhou-Foshan metropolitan area, the model's accuracy of the areas outside outer ring (R^2 = 0.73, RMSE = 9.51 m) is slightly higher than in the Middle Ring (R^2 = 0.66, RMSE = 14.02 m). The results indicate that the model performs well across different regions of Guangzhou-Foshan, with more accurate estimates in outer ring and outside areas. In Chengdu, the model's performance in the Middle Ring (R^2 = 0.40, RMSE = 9.41 m) is slightly lower than in the outside outer ring area (R^2 = 0.74, RMSE = 7.77 m), suggesting higher accuracy in the more peripheral areas. Buildings in city centers have more complex functions and higher height heterogeneity compared to suburbs, which may make it challenging to accurately estimate building heights accurately. Overall, these results demonstrate that the model performs effectively across different urban regions, providing supports for the model's application in various urban environments.

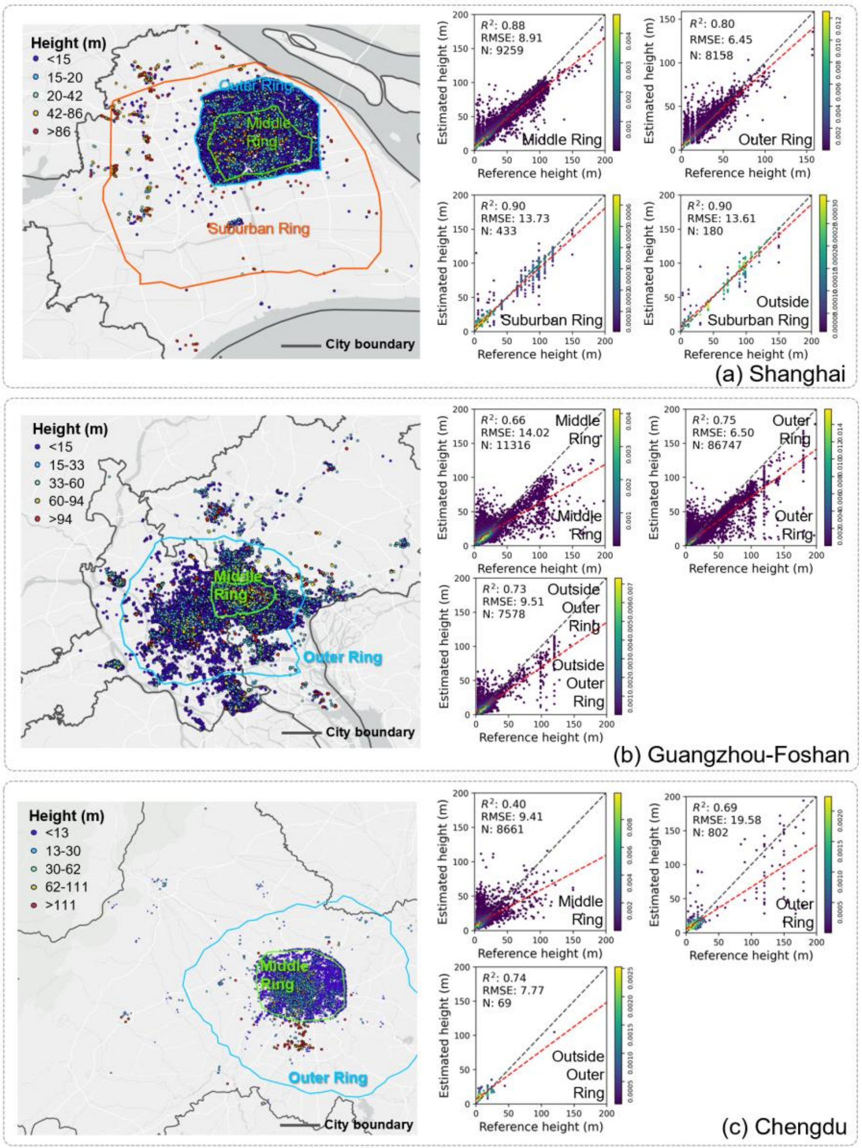

Figure R5. Model performance from the urban core to fringe areas. (a) Shanghai, (b) Guangzhou-Foshan, (c) Chengdu.

Comment #4: For high-rise buildings, especially super tall buildings, what are the potential reasons for height underestimation? It is recommended to include a detailed error analysis and explanation.

Response: thank you for your question. We discussed the reason about the underestimation of high-rise buildings in Section 4.1 as:

"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 11-12, line 260-265)

Reference:

Frantz, D., Schug, F., Okujeni, A., Navacchi, C., Wagner, W., van der Linden, S., and Hostert, P.: National-scale *mapping of building height using Sentinel-1 and Sentinel-2 time series, Remote Sensing of Environment, 252, 112128-112128, https://doi.org/10.1016/j.rse.2020.112128, 2021.*

Stilla, U., Soergel, U., and Thoennessen, U.: Potential and limits of InSAR data for building reconstruction in built-up areas, ISPRS Journal of Photogrammetry and Remote Sensing, 58, 113-123, https://doi.org/10.1016/S0924-2716(03)00021-2, 2003.