This paper basically proposed a downscaling approach to allocating building stock per province in China for better risk assessment, which well fits the scope of this journal. Given current revision status of this paper, I only have one main concern, which is about the definition of urban/township/rural. In reality, we can hardly differentiate them, in particular at pixelated level. Shanghai as we all know is highly urbanized, but still I can see many rural pixels from Figure 1, I do not believe it is the real case here. To me, 'rural' is mostly remote natural areas, I assume you intend to say 'village'.

Response: Thanks for your review. We totally understand your concern. It is true that in those highly urbanized provinces (e.g., Shanghai, Beijing, Guangzhou, Zhejiang, Jiangsu etc.), rural areas are not remote natural areas, although this is true for mountainous provinces (e.g., Sichuan, Guizhou, Yunnan, Tibet etc.). However, we consider it is also not appropriate to change "rural" to "village" in the context of this article, since villages typically have administrative boundaries. As explained in Section 2.3 of the manuscript, "The urbanity attribute of statistics in the 2010-census records is determined according to the administrative unit of the surveyed population. We also have made it clear in this section that if a residence is from a village, then the related statistics are aggregated into rural urbanity level; and if from a town, then it is township level; if from a city, it is urban level". Therefore, compared with "urban" and "township", the word "rural" only refers to those less developed/populated area within a province.

Moreover, I also see many pixels (with some built-up land) that are not assigned to any of the three grid types, but in Figure 4, buildings nearly spread all over the city of Shanghai, which confuses me a bit.

Response: Thanks for pointing this out. This difference is related to the setting of the geometry type of the visualization layer. In Figure 2 of the manuscript, in which only the Baoshan district of Shanghai is shown, the original point layer (Figure R1) has been transferred to polygon layer with grids of 0.009°×0.009°, approximate to 1km×1km resolution (Figure R2). One point in Figure R1 corresponds to one square in Figure R2. That is why some area is not assigned to any square.

While in Figure 4 of the manuscript, we show the whole Shanghai City, in which the visualization layer remains to be a point layer, only the symbology is set as square, thus it seems buildings spreads all over the whole city. But if Figure 4 is enlarged (as shown in Figure R3), there will also be gaps between grids.

Such visualization difference will not affect the calculation of exposure and seismic risk, since such spatial analyses usually need to calculate how many grids are located within specific area.

Figure R1: The original point layer of Figure 2 in the manuscript.

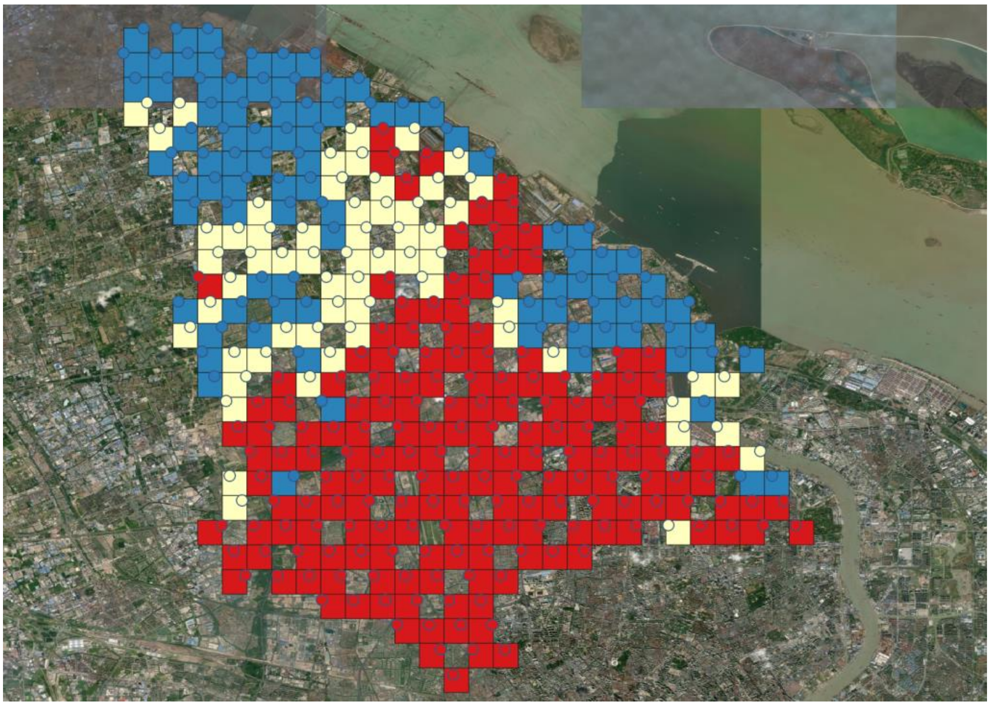

Figure R2: Converting the point layer of to polygon layer.

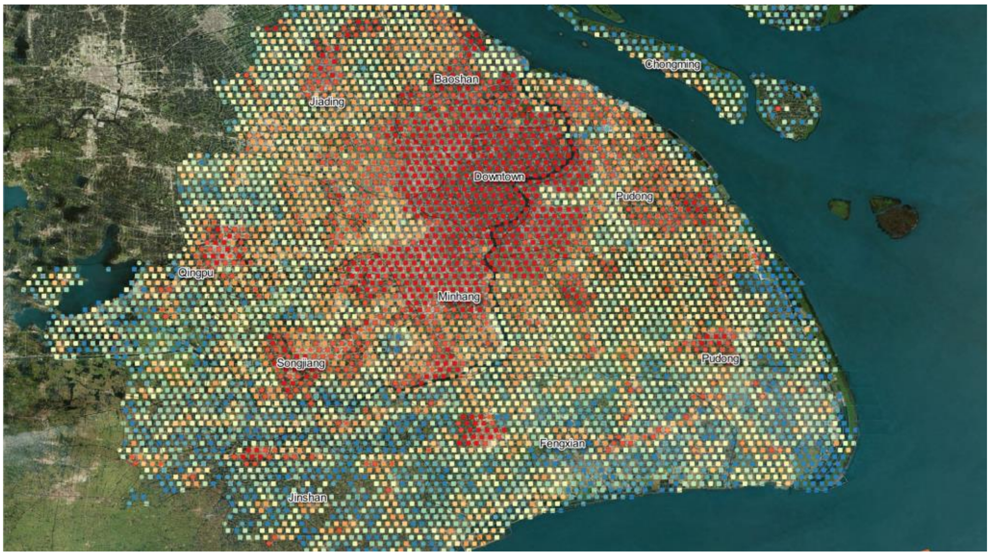

Figure R3: The enlarged version of point layer in Figure 4 of the manuscript , in which the symbology is set as square.

Minor issues:

1. Abstract part is too lengthy. The research background, method, result, and possibly implication need to be clearly stated, I suggest to remove some unnecessary details to enhance readability.

Response: Thanks for this suggestion. The abstract has been shortened by reducing the original **529** words to **344** words. The revised version of the abstract is as follows:

"To enhance the estimation accuracy of economic loss and fatality in seismic risk assessment, a high-resolution building exposure model is important. Previous studies in developing global and regional building exposure models usually use coarse administrative level (e.g., country, or sub-country level)

census data as model inputs, which cannot fully reflect the spatial heterogeneity of buildings in large countries like China. To develop a high-resolution residential building stock model for mainland China, this paper uses finer urbanity level population and building-related statistics extracted from the records in Tabulation of the 2010 Population Census of the People's Republic of China (hereafter abbreviated as the "2010-census"). In the 2010-census records, for each province, the building-related statistics are categorized into three urbanity levels (urban, township, and rural). To disaggregate these statistics into high-resolution grid level, we need to determine the urbanity attributes of grids within each province. For this purpose, the geo-coded population density profile (with 1km×1km resolution) developed in the 2015 Global Human Settlement Layer (GSHL) project is selected. Then for each province, the grids are assigned with urban/township/rural attributes according to the population density in the 2015 GHSL profile. Next, the urbanity level building-related statistics can be disaggregated into grids, and the 2015 GHSL population in each grid is used as the disaggregation weight. Based on the four structure types (steel/reinforced-concrete, mixed, brick/wood, other) and five storey classes (1, 2-3, 4-6, 7-9, ≥10) of residential buildings classified in the 2010-census records, we reclassify the residential buildings into 17 building subtypes attached with both structure type and storey class and estimate their unit construction prices. Finally, we develop a geo-coded 1km×1km resolution residential building *exposure model for 31 provinces of mainland China. In each 1km×1km grid, the floor areas of the 17 residential building subtypes and their replacement values are estimated. The model performance is evaluated to be good and its practicability in seismic risk assessment is also checked. Limitations of this paper and future improvement directions are discussed. The whole modeling process of this paper is fully reproducible, and all the modeled results are publicly accessible."*

2. How do you define high resolution? Is 1 km of high resolution? Here your modelled results are of 1 km resolution. As far as I know, even the 30-m Landsat imaginaries are claimed to be moderate resolution (see for example: https://www.montana.edu/spowell/documents/pdffiles/powell_jars.pdf). In addition, MODIS, which stands for the Moderate Resolution Imaging Spectroradiometer, is also moderate resolution, of course.

Response: Thanks for this comment. For building exposure model, the 1km*1km resolution is relatively high when compared with models at administrative level. It is true that the remote sensing datasets can be of much higher resolution, but for building exposure model development, additional attributes (e.g., the building structure type, story class, seismic design level, construction year etc.) needs to be attached with the remote sensing data to develop such a model. However, these attributes data are usually of lower resolution. That is why their final product, the building exposure model, when is of 1km*1km resolution, can be considered as a high-resolution model.

3. Two key publications on mapping buildings particularly for China are missing here: https://www.sciencedirect.com/science/article/pii/S0034425720302297

https://www-sciencedirect-com.vu-l.idm.oclc.org/science/article/pii/S016920462100150X

Response: Thanks for recommending these two publications of Li et al. (2020) and Liu et al. (2021). Based on your recommendation, we find another two related studies, namely Ji et al. (2020) and Cao and Huang (2021). We will briefly introduce four papers in Section 4 when discussing the future improvement directions for exposure model development.

The introduction of these studies will be given as follows:

"In addition, Li et al. (2020) developed the first continental-scale dataset on 3D building structure (including building footprint, height, and volume) at 1km×1km resolution for Europe, the US by using random forest models fed with remote sensing and Synthetic Aperture Radar imagery data. Liu et al. (2021) developed the urban floor area map for mainland China at 130m×130m resolution based on high spatial resolution nighttime light LUOJIA 1-01 images, a population map and a single building dataset encompassing 71 cities. Ji et al. (2020) generated the 10m×10m resolution rural settlements in the Yangtze River Delta of China by using the multi-source remote sensing

datasets with the Google Earth Engine Platform. Cao and Huang (2021) proposed a multi-spectral, multi-view, and multi-task deep network (called M3Net) for building height estimation. They estimated the building height at a spatial resolution of 2.5m×2.5m for 42 Chinese cities. Comparison with the results in Li et al. (2020) indicated that the M3Net method in Cao and Huang (2021) can better alleviate the saturation effect of high-rise building height estimation than the random forest method used in Li et al. (2020)."

References:

Cao, Y. and Huang, X.: A deep learning method for building height estimation using high-resolution multi-view imagery over urban areas: A case study of 42 Chinese cities, Remote Sensing of Environment, 264(2021), 112590, doi: 10.1016/j.rse.2021.112590, 2021.

Ji, H., Li, X., Wei, X., Liu, W., Zhang, L., and Wang, L.: Mapping 10-m Resolution Rural Settlements Using Multi-Source Remote Sensing Datasets with the Google Earth Engine Platform, Remote Sensing, 12(17), 2832, doi:10.3390/rs12172832, 2020.

Li, M., Koks, E., Taubenböck, H., and van Vliet, J.: Continental-scale mapping and analysis of 3D building structure, Remote Sensing of Environment, 245(2021), 111859, doi:10.1016/j.rse.2020.111859, 2020.

Liu, M., Ma, J., Zhou, R., Li, C., Li, D., and Hu, Y.: High-resolution mapping of mainland China’s urban floor area, Landscape and Urban Planning, 214(2021), 104187, doi:10.1016/j.landurbplan.2021.104187, 2021.