Comment on “Residential building stock modeling for mainland China targeted for seismic risk assessment” by Xin et al.

Comment 1: The authors present an interesting approach to achieve a nation-wide model for the building stock to be used in seismic risk assessment. Based on various statistical data and information derived from secondary sources and remotely-sensed data, they present a method to derive at a geo-coded 1km×1km resolution residential building exposure model for 31 provinces of mainland China. Moreover, based on a sensitivity analysis for one case study, the authors present possible sources of uncertainty in the results, and show how these may be decreased during future research efforts.

Response: Thank you very much for your time and efforts in reviewing this manuscript.

Comment 2: Overall, the paper is timely, and well-structured. The overall point of criticism is that the sole use of statistical data to derive at a “real world” building stock model neglects region-specific and very local impacts on the distribution and quality of assets, which may lead to systematic over- and underestimation in certain areas of the country.

Response: It is true that regional specific and very local impacts are not considered in this modeling process. This is mainly limited by the difficulty in collecting such detailed data for specific regions, which are usually proprietary and not publicly accessible.

Comment 3: Nevertheless, I strongly believe that the method is worth being published so that the international research community can further refine the method and decrease inherent uncertainties.

Response: Thanks for this affirmation. Our detailed responses to your comments are as follows.

Some minor comments:

Comment 4: Line 152/153: something is missing here. Should be re-formulated.

Response: Thanks for pointing this out. The complete reason that the township/street level population generated by using the multi-variate regression method in Fu et al. (2014a) tends to overpredict the population density in a sparsely populated area and underpredict the population density in a densely populated area is enriched as follows:

"The reasons for such discrepancies are that: (1) The population density developed for each land use type by using the multi-variate regression method is the average population density, thus the over/under prediction of the actual population density in certain areas is inevitable; (2) When applying the multi-variate regression method, no additional supplementary data (e.g., road density, nighttime light) is employed to adjust the development level difference in different regions, because the development level is much higher than the average in places such as the downtown area of metropolitan cities like Shenzhen and Guangzhou."

We will add the above explanation to the revised manuscript.

Comment 5: Line 193/194: could be better formulated.

Response: Thanks for pointing this out. The old expression "*The census for the year 2020 is just initiated and normally it takes around two years to publish the final surveyed data. Therefore, the current latest census data are for the year 2010*" will be rephrased as "*Detailed statistics for the year 2020 are not publicly accessible yet. Therefore, census data for the year 2010 will be used to elaborate the modeling process*".

Comment 6: Line 270/271: please elaborate a bit more why the spatial coverage is limited.

Response: Thanks for pointing this out. The limited spatial coverage of PopGrid China is related to its development method, namely the multi-variate regression method (Fu et al., 2014a). In this method, it was assumed that the spatial distribution of population is limited within the six land use types recognized from the Landsat TM images, namely cultivated land, forest land, grass land, rural residential land, urban residential land, industrial and transportation land. However, in actual cases, population distribute more widely and are beyond these land use types.

Comment 7: Line 469: in times of almost unlimited computing capacity, this should not be an issue. In contrast, applying the same unit price over the entire area of (mainland) China is a major source of uncertainty of the method, which should be addressed in more detail in the respective section 4.

Response: Thanks for pointing this out. In Line 467-469, we emphasized that "*There are significant differences across the country in terms of economic development level, geographic climatic diversity, and standardization in building construction. Therefore, it is mainly for computational convenience that this paper applies the same unit construction price for all the provinces and all the urbanity levels.*". The "computation convenience" here does not refer to the limitation in computing capacity of hardware. Instead, we mean it is convenient to only use a uniform price list to preliminarily calculate the replacement value of residential buildings in each grid, since it is quite difficult to compile a complete and accurate building construction price list for different provinces and urbanities. On one hand, this is because different documents include different items in calculating their unit construction prices. For example, some only consider the cost to build the main structure, while others also consider the cost of supporting facility and landscaping. On the other hand, different areas have different seismic fortification requirement, which will also alter the unit construction price for the same type of building.

Actually, before compiling the unit construction price in Table 4 of the manuscript, we have consulted cost-engineers from the real-estate industry. Their internal documents on cost control indicate that for the same type of residential building, the unit construction price of the main structure in different regions in China is limited to 300 RMB. Also, according to Li et al. (2021), the average unit construction price of multi-story reinforced concrete in urban areas of Tibet reaches 3200 RMB/m^2, which is quite comparable to the unit construction price for the same type of building in those more developed coastal areas.

Therefore, we prefer to only provide a reference set of unit construction price in Table 4 and avoid to over-manipulate it. As you point out, this will cause major uncertainty and we totally agree. However, for the sake of model users, this simplicity and transparency will make it more convenient for them to adjust the reference price to their targeted study area by simply multiplying some rectification factors.

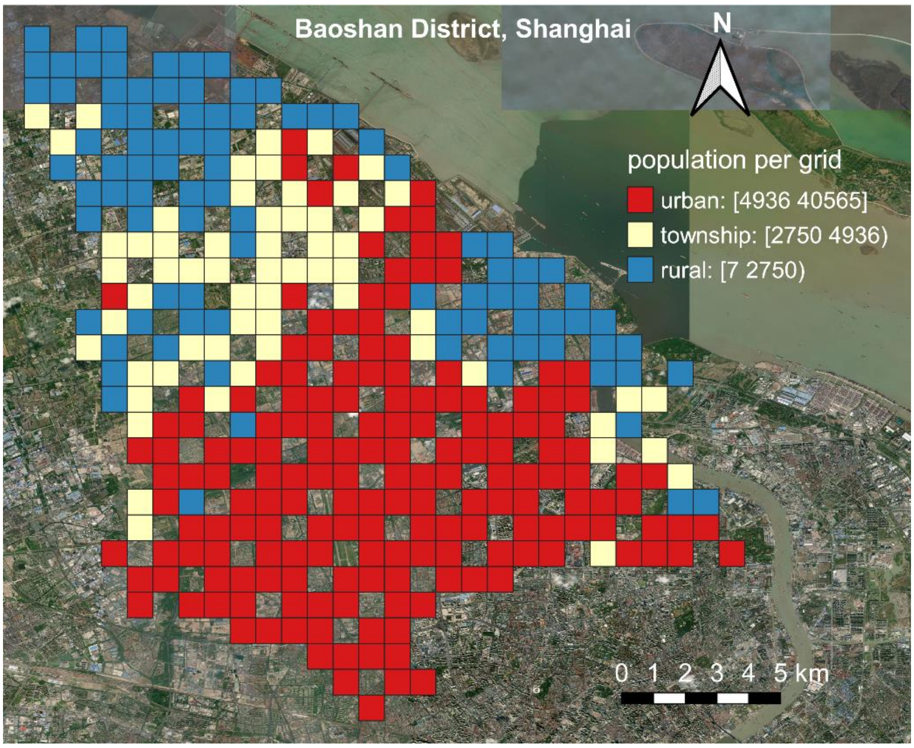
We will add the above explanation to the limitation discussion in Section 4.

Comment 8: Section 3.1.2: Here it is not clear to me what the key message is; obviously, higher buildings will have a higher density of floor area and, thus, a higher population density. My suggestion is to elaborate this a bit more, or to delete this particular section.

Response: Thanks for this comment. As implied from the title of this subsection, this short section serves as an example to demonstrate the modeled floor area in Shanghai, which is to help potential readers to conduct direct comparison with other reports or modeling results. For example, they can directly check whether the grids that have the highest modeled floor area are within those most prosperous regions in Shanghai, as explained in Line 474-476 of the manuscript.

Comment 9: Figure 1: technically, the classes are not clearly distinguishable (what if a grid has exactly 4936 or 2750 inhabitants?), please adjust.

Response: Thanks for pointing this out. We think you refer to Figure 2. The modified population ranges for urban/township/rural urbanity levels in Figure 2 are as follows.



Comment 10: Figure 4 (and related section in the main text body): from my understanding it would be more explanatory how well your method is suitable for application if you would compute the differences between the modelled floor area per km^2 and the 3d view provided in inlet (c), also in terms of uncertainty

quantification. Please also consider the similar issue of classes given already for Figure 1 (and check all the other Figures, also in Figure 9 this is wrong).

Response: We are afraid there is some misunderstanding here. Indeed, the population data in Figure 2 (not Figure 1) and Figure 4(c) are from the 2015 GHSL developed by the European Union. And this population density profile is the base for us to divide the grids in each province into urban/township/rural levels. After this, we further disaggregate 2010-census statistics for urban/township/rural levels into corresponding grids.

Therefore, Figure 2 is to demonstrate how we assign urban/township/rural attributes to grids according to 2015 GHSL population density. Figure 4(a) is the example demonstrating one of our modeled products, namely the floor area in each grid. Figure 4(c) is plotted on the base of Figure 4(a), and its height is determined by the 2015 GHSL population in these grids, but the floor area is the same as that in Figure 4(a). By plotting Figure 4(c), we mainly want to show the location of those most densely populated grids.

Figure 9 is quite different from Figure 2 or Figure 4, because it is an application of the modeled results. It gives the seismic loss ratios calculated based our modeled building floor area, replacement value as well as an empirical vulnerability curve and the intensity map of the 2008 Ms8.0 Wenchuan earthquake.

Comment 11: Given these constraints I recommend revisions before the manuscript may become acceptable for publication.

Response: We deeply appreciate your generosity in spending time on reviewing this manuscript, which requires a lot of patience and efforts. We hope our responses have solved your concerns on this work. If not, we would like to make further explanation.

References mentioned in the responses:

Fu, J., Jiang, D. and Huang, Y.: Populationgrid_China, Acta Geographica Sinica, 69(Supplement), 41–44, doi: 10.11821/dlxb2014S006, 2014a (in Chinese).

Li, C., Li, Z., Lyu, H., and Gao, M.: Probabilistic seismic risk assessment for the Eastern Himalayas, China, Earthquake Spectra, doi: 10.1177/8755293021999056, 2021.