Anonymous Referee #2

Received and published: 20 June 2020

Comment 1: The Introduction starts with some common sentences on earthquake loss and related action on the UN level, however, this paragraph is not very suitable to introduce the topic of the manuscript to the potential readers. Therefore, I kindly recommend rewriting.

Response: Thank you for this suggestion. We will reorganize this part in the revised manuscript by using materials more closely related to our focus.

Comment 2: Moreover, the statement of world-wide earthquake loss should go clearly beyond the referred two studies of one of the co-authors of this manuscript.

Response: Thank you. We will add other references to this statement.

Comment 3: One more detail: For me it is not clear why in a paper from 2011 earthquake loss can be given in 2016 values, please clarify.

Response: This is because the work in Daniell et al. (2017) is based on the work of Daniell et al. (2011). To avoid confusion, we will delete the latter and only keep the former reference.

Comment 4: Then the authors address the need to have information on the building stock level when risk assessment should be undertaken. They state that in cases where such information is not available, obtaining necessary information is not practicable. I strongly disagree with this statement, throughout the relevant literature there are many different methods presented of how to do so. This needs thorough revision.

Response: Thank you for this comment. We will replace the description of "obtaining necessary information is not practicable" by "a series of ancillary data based on remote sensing technologies will be resorted to". And a succinct summarization of related methods will be added.

Comment 5: Moreover, the authors elaborate on a method to compile such information by taking census data in consideration. How the building stock value is correlating to statistical information on population density (lines 49-57)? How did the authors generally treat the MAUP issue when using data bond to administrative borders?

Response: A detailed description of the data used and the modelling process we adopted are provided in the "Data Sources and Methodology" section. In short, the building stock value (here we use the replacement value) is derived by multiplying the residential building floor area and the unit construction price we compiled for the building types used in this work. And the residential building floor area is derived directly from the statistical information provided in the 2010-census of China. Since the 2010-cenus data are categorized at urbanity level, to disaggregate the census records into grid level, the 2015 Global Human Settlement population density profile is used to derive the apportion weight of each grid.

If MAUP refers to the "modifiable areal unit problem", then we did not consider the MAUP problem during the modeling process. According to the explanation of MAUP from the internet: "MAUP is a statistical biasing effect when samples in a given area are used to represent information such as density in a given area. The area defined by an analyst is often arbitrary, thus measurement such as density could be deceptive because that density measure could have widely different results based on shape and scale chosen for analysis." However, in this work the provincial boundaries are not arbitrarily defined and the 2010-census data to be disaggregated is also derived from survey population within each province administratively. Therefore, all the boundaries are clearly defined. What we need to do is to distribute the census data into grid level for each province. Therefore, we do not think the MAUP issue will affect our modelled results and to be honest, we have not seen discussions on this MAUP issue in studies of this kind even after an extended literature review.

Comment 6: Further down the text body, the authors correctly state that "to better cope with this spatial mismatch between natural hazards [spatial occurrence] and administrative boundaries, building stock models should be geocoded with relatively high resolution and be disaggregated from more detailed census data". The last statement means that from a methodological point of view, such an assessment will not guide us to precise results that can be used as a proxy for the building stock. So somehow, the introductory section is unclear with respect to what the authors would like to show us in their study.
Response: Thank you for pointing this out. You are right that geocoded resolution referred to in this study (1km*1km) is still not specified enough to provide precise results that can be used as a proxy for the building stock. The purpose of this work is to develop grid-level residential building stock model for seismic risk assessment in mainland China and this is achieved by disaggregating 2010-census into 1km*1km grid with the 2015 GHS population density profile as the bridge. We will add more method description in the introduction part to avoid possible confusion and misunderstanding.

Comment 7: Finally, the research gap is not properly defined, nor is the niche to be filled by this work easily accessible to potential readers. In the method section the authors explain how they aggregated information on the building characteristics to information on the population density, both at a final resolution of 1 km grid cells. In this respect it remains unclear how the other building types were excluded from the grids, as information on e.g. building design and material in the statistics ("Long Table") are also related to other building types, right? -> Needs clarification.
Response: Thank you for this comment. No building type is excluded in the modelling process. The 2010-census recorded number of families living in buildings classified by four structure type (brick-wood, mixed, steel-RC, other) and by five story classes (1, 2-3, 4-6, 7-9, ≥10) is regrouped into 17 building sub-types attached with structure type and story class. Based on these building sub-types, the unit construction price for each of them will be easier to be compiled.

Comment 8: Further on, the authors present different methods of how to merge different types of information such as the amount of buildings of different height or different construction type to these grid cells, resulting in a certain spatial probability for the different data. It remains open, however, how this information was finally be checked against the real world situation, and as such it remains open. How e.g. information on population was distributed or allocated to different building categories?

Response: For each urbanity level of each province, number of families living in different residential building types (classified by structure type or story) are given in the 2010-census. The average population per family and the average floor area per capita are also given in the 2010-census. Therefore, the overall population living in each building type and the total floor area of different building types (classified by structure type or story) can be derived for each urbanity of each province. This is the way how information on population and building is related. We now need to distribute the overall building floor area for the whole urbanity into grid level, based on the apportion weight derived from the 2015 GHS population density profile.

Comment 9: Occupation rate and building values were then allocated to the different building sub-categories, and spatially distributed over the grid cells. Results of values per grid cell where then compared to (A) a study published by Wu et al. (2014) on the net capital stock, (B) more detailed information available on the residential floor area for the Shanghai district, and (C) an empirical earthquake vulnerability study published by one of the co-authors of this manuscript, linking vulnerability to reported loss. The authors conclude that the results from the present manuscript (in terms of what? Potential value of buildings? Potential loss resulting from an earthquake scenario?) are in line with results from other studies, a statement which cannot be supported by the referee evaluating the information provided by the authors. In the present form, the results of the study are not validated, they are only opposed to other studies on building values (in case A), to the area used for computation of values (in case B) and to vulnerability, linking the newly generated building values to an empirical vulnerability function and comparing the results to some loss reports available (in case C).
Response: We are afraid we cannot agree with the conclusions in the comment here. We believe you also agree that there are different levels of validation tests, from a general comparison of an overall value, to an exact match of data distribution pattern, or even an building-by-building comparison. Ideally, the best validation practice will be comparing the modelled results with in-situ site surveys, which will be undoubtedly expensive and extremely time consuming for we to do this for the whole mainland China, also we cannot afford to do this. It is 35 years ago that the Chinese government conducted its last country-wide building survey in 1985. However, considering the main application of the model developed in this study is for seismic risk assessment, an exact validation may be not necessary. Firstly, when damaging earthquake occurs, it usually affect at least several counties or even provinces. What we want to do is develop a relatively high-resolution building stock model based on the open census data. When distributing the census data into grid level, to better reflect the spatial heterogeneity of population and building distribution, we use the 2015 GHS population density profile as the bridge. And we believe it is better than distribute census data into an administrative unit evenly.

When it comes to the validation of our modelled building stock replacement value and floor area, a reasonable test is to check whether the model can be used to better reflect the seismic loss distribution pattern or to compare modelled results with outputs derived by using different methods at coarser level. And we do not think this is a "self validates self" thing, since when disaggregating the urbanity level census data into grid level, this process is not related to or affected by prefecture or county/district boundary. And we think if we reaggregate the grid-level data into county/district/prefecture level and compare them with other independent sources of validation data. If they are still consistent, then it can be regarded as a positive signal indicating the reliability of our modelled results. The three validation tests in this work are performed for different purposes.

(A) The comparison with Wu et al. (2014) is to first make sure that the replacement value of residential building stock we modelled is of reasonable magnitude. Since our model and that of Wu et al. (2014) are different in method, model target and the treatment of depreciation issue, but the ratio between our residential building replacement value and the net capital stock value in Wu et al. (2014) is within the range of 0.31~0.65. Recently, for each province we also checked the average ratio between newly constructed residential building value and the value of all newly constructed buildings in the past five years. This ratio is within the range of 0.38~0.75. Through this comparison, we consider that the replacement value of residential building stock we estimate for each province is reasonable.

(B) We then compare our modelled residential building floor area at district level with the statistical results of Shanghai only, since such data in other provinces are not available. Since the disaggregation of the urbanity level census data into grid level is not limited by or related to county/district boundary, when our modelled results are reaggregated into county/district level and have high correlation coefficient (0.91 without adjustment) with county-level statistical records, we believe this consistency is suggesting the reliability of our modelled residential building floor area for Shanghai.

Recently, after a broader check of databases related to building statistical data, we found that although in other provinces there are no such records of accumulated residential building floor area like that given in Shanghai 2015 statistical yearbook, there are provincial 2010-cenus data available (although not with open access, but can be bought), in which the residential building floor area can be directly derived for each county/district. We will allocate these data and further validate our modelled residential floor area for all counties/districts of all the 31 provinces. The comparison will be added to the revised manuscript.

(C) In applying our modelled residential building replacement value to loss estimation by using the intensity map of Wenchuan Ms8.0 earthquake, we are trying to check the availability of our modelled results in risk analysis, since the model is targeted for seismic risk analysis.

We do not think that the empirical loss curve derived by Daniell (2014) is used in this process is a problem. First of all, this empirical loss curve was derived long before the preparation of this study. Secondly, the derivation of this empirical curve is based on extensive collection of damage and loss records from journals, books, reports, conference proceedings and even newspapers. Therefore, only because the author of this empirical curve is also a co-author of this work should not decrease the reliability of his previous research findings. In Daniell (2014), he also developed empirical loss curves for other countries and regions through careful selection and rectification of historical records and it turns out that his curve performs better than previous studies in predicting seismic loss, since he considered the change in vulnerability of exposure with time. That is why in the study, the empirical loss curve in Daniell (2014) for mainland China is used.

Since our estimated loss is around 144-288 billion RMB (in 2015 current price), while the median loss estimated from post-earthquake investigation is around 212.32-247.25 billion RMB (in 2008 current price) for residential buildings in Sichuan province, we think this can be taken as a consistent result. Of course.

Comment 10: With respect to the latter, further questions arise with respect to different construction types and their individual structural vulnerability concerning earthquakes, this needs careful interpretation and more information on the comparison performed.
Response: In case C, the empirical loss curve of Daniell (2014) gives the relation between macroseismic intensity and loss ratio (the ratio between building repairment cost and replacement cost). This curve is called as "empirical" because it is regressed from the loss ratio and intensity pairs of historical damaging earthquakes. And the loss ratio is the derived from the loss and exposed value considering all building types. Therefore, there is no further classification of buildings when applying this empirical curve in rapid loss estimation, only the seismic intensity map and the sum of the replacement value of all buildings in each grid are needed.

It is worth to note that in this work, the replacement value not the repairment value is used in the loss estimation, which means the depreciation of buildings are not considered. The difference in building vulnerability will not affect the estimated replacement value, but will affect the repairment value greatly.

Comment 11: As such, the added value of the material presented here is not clear to me. Statements such as "Therefore, the estimated loss range, based on the buildings stock model developed in this study and the empirical loss function developed in Daniell (2014), is quite compatible with that given in previous studies. This compatibility further validates the robustness of our residential building stock model" (lines 392ff.) may be right but cannot be proven from what the authors have shown.
Response: We fully understand your concern, since the loss estimation process involves uncertainties from both the estimation of replacement value of residential buildings and the uncertainty in development of the empirical loss curve. That is why the estimated loss is given in range form, not a specific value. The reliability of the modelled replacement value and floor area have been checked in case A and case B. In case C, we want to further check the prospect of applying the modelled results in seismic risk assessment and our estimated loss of 144-288 billion RMB (in 2015 current price) is approximate to loss estimated from post-earthquake investigation, which is 212.32-247.25 billion RMB (in 2008 current price). We think this compatibility indicates the positive prospect of applying our model to seismic risk assessment.

Comment 12: Finally, information given in the conclusion is by far too generalized, instead, the overall limitation and the specific ones, such as those arising from the use of aggregated statistical data and the underlying factors, should be discussed in detail.
Response: Thank you for your patience to read this manuscript through. We will strengthen

discussion on limitations of this work.

Comment 13: Hence, I cannot recommend publication of the work in its present form, the manuscript needs major revisions in terms of (1) problem statement, including extended literature review, (2) explanation of methods, (3) compilation of results, and – most important –(4) discussion, including discussion on limitations and uncertainties, so that (5) sound conclusions will become possible.
Response: Thank you for your careful review. We will take these comments seriously and modify our manuscript accordingly in the revised version.

References:

Daniell, J.: Development of socio-economic fragility functions for use in worldwide rapid earthquake loss estimation procedures, Ph.D. Thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany., 2014.

Daniell, J. E., Khazai, B., Wenzel, F. and Vervaeck, A.: The CATDAT damaging earthquakes database, Natural hazards and earth system sciences, 11(8), 2235–2251, doi:https://doi.org/10.5194/nhess-11-2235-2011, 2011.

Daniell, J. E., Schaefer, A. M. and Wenzel, F.: Losses Associated with Secondary Effects in Earthquakes, Front. Built Environ., 3, doi:10.3389/fbuil.2017.00030, 2017.

Wu, J., Li, N. and SHI, P.: Benchmark wealth capital stock estimations across China's 344 prefectures: 1978 to 2012, China Economic Review, 31, 288–302, doi:10.1016/j.chieco.2014.10.008, 2014.