

**Manuscript title: A new regionally consistent exposure database for Central Asia: population and residential buildings**

This manuscript proposes a regional exposure database featuring the population and residential building stock for the countries of Central Asia. This new dataset has a regional scale, with the appropriate resolution to be used in multi-hazard and risk assessment. Moreover, the authors provide estimates for exposure for the year 2080 that can support current and future risk mitigation strategies. To achieve this, the authors use high resolution population and building datasets to generate gridded exposure datasets.

The introduction clearly states the research goals and main novelties. The methodological aspects of the research are very interesting. The authors rely on local data and expert knowledge to provide reliable building characterization and replacement costs. The manuscript presents relevant information for other risk scientists. This information is clearly laid out in maps, figures, and tables, making the paper useful, appealing, and easy to read. The exposure layers for 2080 are a significant addition to the research outputs. These rely on future scenarios of population and urbanization to propose different possibilities for future exposure in Central Asia. The results section is brief, clear, and very well written. I praise the authors for presenting the results of their work so well.

The manuscript contains a couple of typos, which can be easily corrected, but it is also missing several references, some of which are essential to the research. I strongly suggest thorough proof-read by the authors. Regarding the methodological aspects of the research, there are only a couple of points that need to be explicit in the manuscript. One is regarding the SSPs, specifically the motivation for the scenarios chosen and the uncertainties that they account for in this research (i.e., it seems that future population and urbanization are being accounted for, but not future sustainability, resilience, or vulnerability of the residential building stock). The other point that needs to be clearly addressed is the expected reduction in the future number of buildings, which is most likely due to the fact the future abandoned or unoccupied buildings are not being accounted for in the 2080 exposure dataset.

Based on this assessment, I believe the manuscript would be an excellent contribution to the scientific community, and that its main contents are up to the standards of NHESS. I also want to extend my congratulations to the authors for their hard work. My recommendation is to **accept the manuscript, subject to minor revisions**. Please find my comments line to line below, which I hope can help the authors greatly improve the quality of the final publication.

**Introduction**

Lines 30-35: Reference Yu et al. 2019 is missing in the references section. The authors need to include it.

Lines 45-50: Reference UNECE 2017 is missing in the references section. The authors need to include it.

Lines 55-60: The authors mention that the only regional-scale exposure dataset of residential buildings available at the time (April 2023) is provided by Pittore et al., (2020). I believe the Global Earthquake Model Foundation has provided an updated exposure model for Central Asia in terms of population, building counts and replacement costs (See Yepes-Estrada et al. 2023 and [https://github.com/gem/global\\_exposure\\_model](https://github.com/gem/global_exposure_model)). The authors should consider mentioning the update as well.

Lines 75-80: The authors state that the first regional-scale exposure dataset for Central Asia was developed by Pittore et al., (2020) using ground-based and remote sensing datasets. The Global Earthquake Model Foundation also has updated exposure datasets covering Central Asia using a bottom-up approach (Yepes-Estrada et al. 2023). Herein, however, the authors maintain that they are proposing the 'first regionally-consistent exposure database for Central Asia'. Were the first efforts and methodologies to map the residential exposure in the region inconsistent? How so? There are novelties in the new dataset the authors propose, but this assertion needs further explanation to be included in the manuscript.

## **Methodology**

Lines 120-125: The authors mention that comparing Facebook data versus national census data results in differences in the total population that exceed 20% in 7 oblasts. What is the total number of oblasts under study? What is the difference between the total population estimates from the census and Facebook? Would it be possible to elaborate? The use of Facebook data is an interesting approach and I think it would be interesting for other researchers to see how it compares to national databases.

Lines 135-140: The Global Earthquake Model building taxonomy by Brzev et al., 2013 has been updated by Silva et al., 2022 (<https://doi.org/10.1007/s13753-022-00400-x>).

Lines 160-165: The authors used the Facebook population data and adjusted it to respect local estimates of the population. Then the buildings were taken from Pittore et al. (2020). Both datasets end up distributed in variable-resolution grids. These disaggregation techniques can result in an inconsistent number of occupants per dwelling and occupants per building in each cell, especially if datasets developed independently are being combined. Do the time stamps of the population and building datasets agree with each other? (i.e. the population of 2020 and the number of buildings by 2020). Was the final number of occupants per dwelling and building consistent with the subnational local estimates after the distribution? It would be a great addition to the methodology section to briefly mention the checks performed after the disaggregation methods.

## **Development of exposure layers for 2080**

Lines 210-220: HAZUS is mentioned throughout the manuscript, but I could not find the reference in the reference section (HAZUS 2021? FEMA 2021? Inventory technical manual?). This is the third reference that is either missing from the manuscript or difficult to find in the corresponding section. There is also a typo: 'manual (2021).g dife Costs'. I strongly recommend the authors to proofread the manuscript carefully.

Lines 235-245: The SSPs chosen to develop these layers were SSP1, SSP4 and SSP5. This is a very interesting choice. However, the authors mention that these scenarios envisage different development drivers, but provide no further motivation for selecting these. For example, why was the SSP2 not included? The 'middle of the road' considers less strong deviations from the current fertility trajectories, hence it is the likely scenario in terms of the future population and urbanization. Could the authors elaborate on what motivated this choice? Was supporting risk management strategies a part of this motivation? How so?

Lines 245-265: It is unclear how the SSPs are used beyond the scenarios of future population and urbanization. The uncertainty regarding the future population is taken from the SSPs. The different future population scenarios are used to infer a future number of buildings, and the urbanization layers inform location and density. However, there seems to be no uncertainty considered in the future characterization of the buildings, which influences the final vulnerability and sustainability of the future exposure datasets.

This would be a very difficult task, and the authors rightly rely on expert judgement alone to propose a single set of rules for future building characterization. There is no mention of this set of rules changing depending on the SSPs. If this is the case, the authors need to be explicitly clear in the manuscript that the SSPs only inform the population figures and building allocation. This is important given the choice of SSPs by the authors. For example, the SSP1 (sustainability with low emissions) and the SSP4 (high inequality with high emissions) are clearly opposite scenarios. For example, a reader could understand that sustainability and a more resilient built environment were at the core of creating the future building characterization for an SSP1 scenario, but it doesn't seem to be the case. I think this section really needs more clarity on this point.

## **Results**

The results section is brief, clear, and very well written. I praise the authors for presenting the results of their work so well. I would even recommend the authors include a figure like Figure 3 for all the sub typologies but only if it is possible within their time constraints.

The only comment here is regarding Table 5. The building variation is mostly negative in the scenarios, given the population prospects of Central Asia. However, it was not explained in the methodology why the building count is decreasing. I believe that it is because the dataset is meant to include occupied structures only and abandoned or unoccupied structures are not being considered. This needs to be explicit in the methodology.

This is important because a sustained population decrease does not always lead to a sustained decrease in the exposed number of structures. For example, despite a steady decline in population, the number of total dwellings exposed in Albania rose by more than 25% between 2001 and 2011 ([www.instat.gov.al](http://www.instat.gov.al)). In other words, residential exposure in Albania (new dwellings become new buildings) is increasing despite the decreasing population, as there are other factors beyond population driving changes in residential exposure. Moreover, less population might lead to fewer occupied structures, but abandoned or unoccupied structures are still at risk of flooding, shaking, and causing debris, which will incur an economic loss for society. If the datasets the authors proposed account only for occupied structures, that is perfectly reasonable, but the limitations of this approach should be mentioned as well.

## **Discussion**

Lines 345-350: The authors mention that the final data is provided in GED4ALL format. Is it open access already? Consider using 'will be provided'.

## **References**

Please include the missing references