

Review nness-2022-73

The study deals with the analyses of various seismic risk scenarios for the Sabana Centro region in Colombia, located in the northern region of the capital city Bogotá, which concentrates important industrial facilities, educational facilities. It is an interesting study that follows the state-of-the-art procedures of scenario risk analyses (at least up to the computation of direct losses), for a region that hasn't been studied before, and thus is a good contribution to the scientific literature that communicates the seismic risk in the country.

Having said this, there are some comments about important issues in the study, that hopefully will help improve its clarity, coherence, and thoroughness. It must be said that after doing the review, the reviewer saw that many of the comments and limitations of the study were included in the discussion section as further developments, however there are many that should be included to make the results sound and representative of the region of study, otherwise many of the presented results could be very misleading.

Section 2:

It would be interesting to include why the Sabana Centro region is of particular interest. In previous studies of the major cities, generated GDP or % of the population in comparison with the whole country were presented as reasons for the study of a particular city or region. Additionally, it is mentioned that it concentrates many economic and industrial activities, but at the end the analysis only deals with the residential building stock.

Line 105: "The majority of the building stock of the region is comprised of one- and two-story houses" It would be good to show a reference with the numbers based on the 2018 Census for this. It is interesting that this is mentioned and still no two-story houses are considered in the analyses.

Section 3.1 Seismic hazard:

Line 171: "eighteen crustal events were selected from the catalogue to be used in this study" Does this mean that only 'historical' events included in the catalogue were included? No new possible events from the event-based tables from the PSHA model?

Table 2: Include column with the distance to the population centroid taken as reference point in the disaggregation to be able to compare this selection of scenarios with the disaggregation graphs presented in Figure 4 and their representativeness in the overall 475-years return period hazard in the region.

Given the proximity of the events based on the disaggregation, directivity effects should be considered for some of these events. As further seen in the discussion this was not considered but it would be good to mention it and not leaving it till the discussion as a further development.

Section 3.2 Exposure model for the residential building stock:

Was the replacement cost updated to 2021? 2022? In which way was this done if indeed it has been updated? If not, it should be done and explained.

Were the new inflated exposure building numbers (based on population as proxy) in any way compared to the dwellings or building numbers reported in the 2018 Census for these regions?

Table 4 only considers unreinforced masonry of 1 storey, which is known to be less vulnerable than the unreinforced masonry of 2 stories, which is actually more common in many urban areas. This typology should be included (assuming something probably based on census data or the surveys), as in the region it is very common to find 2-storey, in some cases more than single storey houses (as previously mentioned in the study also). In the current version, the study may be underestimating the losses.

Section 3.3 Physical vulnerability of residential building stock to seismic ground shaking

Chilean wood structures are known to be in better shape than those in Colombia, and they consider a different type of construction technique. The same goes to the curves used in HAZUS, which are not as representative of the local conditions and may be underestimating the risk. If they are going to be used a more thorough explanation of the limitation of using them should be included.

Table 5 needs a clarification of what each curve considers in each of the damage states. If the vulnerability model is considering a unique consequence model, there may be incompatibility between the loss ratios of the derived vulnerabilities, as each one considers each damage state in a specific different way. This is one of the main issues when combining vulnerability functions from different sources. This is particularly true given the damage results of the studies are shown considering these categories of the damage states. The reviewer saw this mentioned in the further developments of the discussion, but it should be included in the computation of the vulnerability curves here in some way, for the results to be coherent.

Given the exposure is not considering separately the 2-storey housing, there are no vulnerabilities for 2-storeys considered, even when it is more common in the urban environment than the single storey houses. This should be included.

Section 3.4. Social Vulnerability (SV)

The main criticism of the paper is in the consideration of the social vulnerability index as a percentage increase using this expression $(1+SVI)$. As stated in the study "The min-max normalization was used to standardize the SV indicators from zero to one to estimate the SVI per municipality. Higher scores indicate more socially vulnerable municipalities, and lower scores reflect less vulnerable ones. Then, the indicators were integrated by summing them with equal weight, as followed in Contreras et al. (2020c). The resulting SVI index is therefore used to adjust the percentage of economic losses with respect to the costs presented by the building inventory, i.e., multiplying them by $(1+SVI)$ (Carreño et al., 2007)." The problem with this is that there is no analysis

done on the significance of the variables included within the study and no way to know if there are variables that shouldn't be included and if anything is counted double. Additionally, considering this index as a "percentage increase" is extremely misleading. If there was a way to correlate the SVI of each variable in economic terms to the direct economic loss, then this could be done. But this is not done and there is no parametric study or anything else to validate any of the assumptions. This SVI cannot be considered a percentage unless there is backup data validating this. This has been done also in fatality modelling where the models that are presented in any publication are previously calibrated and validated with data from historic events. Moreover, considering previous events reporting post-loss amplification that include costs from the response and recovery stages in some disasters, it has been shown that numbers over 30-40% are almost non-existent (What is demand surge? Olsen and Porter 2010), while this study mention cases with increases of up to 60%. There may be a problem with the explanation of the methodology, but as it is right now it is very difficult understand how it can relate to economic losses, especially direct physical losses after an event.

(These limitations are also afterwards mentioned by the authors in the discussion, but it is a MAJOR limitation of the inclusion of the SVI methodology in the results in this study, as there is no validation or calibration of any kind for the methodology)

Table 9 numbers are misleading as a direct non-weighted average of the 18 scenarios is not probabilistically and statistically sound. It should consider the contribution of each event, otherwise the less probable events are counted in the same way as the more probable ones. In this way, as when computing AAL from a probabilistic analysis, the contributions should consider the probability of occurrence of each scenario. After saying this, it is advised not to present this table and instead present one with the analysis of each scenario done separately as in a deterministic approach, unless it is possible to demonstrate that the 18 scenarios included account for the 100% of the 475 years return period loss and a weighted average is calculated based on the contribution of each.

Section 4.4:

As stated previously, to present these analyses, it would be important to show the contribution of the 18 scenarios to the total hazard in the region (based on the disaggregation results). If not Figure 13 is misleading, considering that it says: "Within the municipalities, the mean percentage of losses is presented with respect to **the total expected losses in the region**".

"The economic losses experienced by a province due to an earthquake depends on the event's epicenter as it is depicted in Figure 12." Figure 12 does not show in any way anything regarding the epicenter. It may not be only the epicenter but also the Mw for each event the cause of the differences, so this statement is not provable from the Figure. Delete it.

Discussion:

Just until this section this is stated: "The simulations of eighteen seismic scenarios with a return period of 475 years show that half of the building stock will experience some degree of damage". How was this '475 years' return period calculated? Even when the disaggregation was done for the

'475 years' return period, how can it be confirmed that the 18 scenarios add up to the 100% contribution for the hazard for this return period? Either way this statement should be included in some way in previous sections and not only until the discussion.

Effects of SV:

The main criticism for this approach is also stated by the authors in this sentence: "First, as not all social aspects exert equal effect after an earthquake, it is necessary to develop a weighted approach to best estimate a more realistic SVI for earthquake events. A second improvement required is to devise a better way of estimating the economic impact of social vulnerability. One potential approach is to generate a database of past earthquakes with different consequences that include the economic costs." These are not needed future improvements but major limitations of the proposed approach. In this kind of analyses, as when performing a linear regression, it is important to avoid counting double and establish the significance of each variable within the analysis, if not this could be overestimating the vulnerability and losses in the region considerably. Also, using the min and max approach is very subjective as many variables as unemployment and poverty are tempered with by local organisms. This kind of indicators are good to compare and prioritize actions within regions but cannot be used in the way they are presented in this study to increase direct physical losses.

Minor modifications:

Line 118: Typo: This **GROWTH** counts for 64% of the total population of the region

Line 248: Incomplete sentence: "The number of masonry buildings represents the 88.61% of the total buildings in Sabana Centro, whereas those of concrete and wood represent 3.16% and 8.24%, respectively, with."

Line 261: Repeated to "In the absence of specific curves locally developed for the Sabana Centro province, fragility curves were selected to **to** represent these structures"