Dear author/s,

Thank you very much for your efforts. I have included my notes and suggestions below.

We sincerely thank the reviewer for their time, effort, and valuable feedback. We have carefully considered all comments and provide detailed, point-by-point responses in the attached document, where our replies are indicated in blue. The proposed revisions to the manuscript are also included in the document and are highlighted in *blue italics.*

Abstract

The abstract says that multi-hazard approaches are not widely used in real-world situations, but it should explain what that means. It is suggested to make clear that older models often get risk wrong because they treat hazards like liquefaction or tsunamis as if they happen separately. The effect seen in the Napier case study should be measured, for example, by showing which hazard caused the most damage, or by noting that the model used data from 30,000 homes to produce its detailed results. Use stronger action words like “enable high-resolution risk pricing” or "permit evidence-based land-use planning." This will make the research more useful for people outside of academic circles.

With regard to this specific point, we agree that making these points in the abstract would be beneficial. We propose to do this by adding the following last paragraph to the revised abstract:

Results include damage state and damage ratio metrics for individual and combined perils, allowing attribution of losses across hazards. Earthquake shaking and liquefaction emerge as the dominant drivers of risk, followed by tsunami, lateral spreading, and landslides. These findings demonstrate the importance of capturing interdependent hazards in earthquake risk analysis. The framework provides decision makers, urban planners, and the (re)insurance sector with actionable metrics to guide resilience investments, refine underwriting, and minimize losses from cascading hazard events.

Introduction

Please clearly define what you mean by internal consistency in this study. Does it mean consistency over time, across locations, or from different sources? Please point out the main problem with how samples are chosen. Say if the framework uses any methods to reduce differences in results or not.

We were referring to consistency across the different cascading earthquake perils caused by the triggering event. We propose to make this clearer in the revised paper by adding the following lines to Section 1.3:

The key advantage, and novelty, of the proposed framework lies in the generation of hazard intensities that are internally consistent across all perils within each simulated earthquake event. Here, internal consistency refers to the conditional dependence of all cascading hazards on a common set of event-specific ground motion parameters and spatial footprint, ensuring coherence across perils and locations for a given realisation. In contrast to approaches that sample hazards independently, this framework derives secondary perils (e.g., liquefaction, landslides, lateral spreading) directly from the primary hazard fields (ground shaking and tsunami), which serve as inputs to both hazard-specific models and subsequent fragility assessments.

As a result, all sampled hazard intensities and damage states follow a correlated cascade, where dependencies are preserved throughout the modelling chain. For example, damage arising from landslides and lateral spreading is jointly conditioned on the same underlying ground motion intensity (e.g., PGA), rather than being treated as independent realisations. This approach reduces artificial variability introduced by independent sampling and produces more physically consistent event scenarios. The process is repeated across all stochastic event sets (SES), enabling the construction of long-term, statistically robust hazard and damage distributions that capture both frequency and interdependence among hazards.

Methodology

It is important to explain how the relationship between different intensity measures (IM) is managed. Giving a clear description of how these IM are combined would help.

The IM's are not combined, but their impacts are. To address this comment, we propose the following revisions to Section 2.3 to make this clearer. Specifically, to lines: 191-210 and 226-232:

2.3 Framework Workflow

Each earthquake in the catalogue is analysed using an MCS-based approach, which simulates damage from each individual peril and the combined effects of cascading hazards. Across all modules, hazard intensities are derived consistently from the same simulated earthquake event. *Ground shaking intensities (e.g. PGA) modelled act as inputs for ground-failure-related hazards such as liquefaction, lateral spreading, and landslides, while tsunami inundation is modelled directly from the earthquake source parameters. A detailed description of the hazard-specific models and their parameterisation is provided in the Supplementary Material.*

The overall procedure, detailed in steps 1–9 and illustrated in Figure 2, is as follows:

1) **Event Selection:** A single earthquake event is selected from the stochastic catalogue (see Sect. 3.1).
2) **Shaking Intensity Modeling:** GMPEs are used to generate a ground shaking footprint (e.g., PGA) for the event at all exposure locations.
3) ***Tsunami Inundation Modeling:*** *The earthquake ruptures are used to model tsunami generation and inundation depths, which are mapped to exposed assets (Sect. 3.3).*

4) ***Liquefaction Severity Modeling:*** *Liquefaction Severity Number (LSN, Tonkin & Taylor, 2013) is calculated using local geotechnical conditions and PGA derived from the ground shaking modelling in Step 2. LSN is sampled from a Poisson distribution within predefined susceptibility zones (Sect. 3.2.3), for the modelled PGA and earthquake magnitude.*
5) ***Lateral Spreading Modeling:*** *Lateral Displacement Index (LDI; Zhang et al., 2004) is calculated per site using geotechnical data, slope angle, distance to a free face, and PGA derived from Step 2. The resulting displacements are assigned to exposed buildings (Sect. 3.2.4).*
6) ***Landslide Modeling:*** *Landslide probabilities (EILP) are calculated from an earthquake-induced landslide susceptibility model to identify potential landslide source areas (Massey et al., 2021b, 2022). This model uses PGA derived from the ground shaking field in Step 2 as input at each source location. Landslide runout modelling identifies potential debris-inundation areas should a landslide occur (Brideau et al., 2020, 2021; Massey et al., 2021a). The resulting probabilities are assigned to exposed buildings (Sect. 3.2.5)*

Across all modules, hazard intensities are derived consistently from the same simulated earthquake event. Ground shaking intensities (e.g., PGA) act as inputs for ground-failure-related hazards such as liquefaction, lateral spreading, and landslides, while tsunami inundation is modelled directly from the earthquake source parameters.

For each event in the catalogue, multiple Monte Carlo simulations (MCS) are performed to sample hazard intensities from their respective uncertainty distributions. Damage states are then sampled from the corresponding damage state probabilities, resulting in a range of possible damage scenarios for each event. This probabilistic sampling approach enables the representation of uncertainty in both hazard and impact, generating multiple realisations of physically consistent damage outcomes. A detailed description of the hazard-specific models and their parameterization is provided in the Supplementary Material.

The number of MCS´s required to achieve stable mean damage estimates for each event in the catalogue can be determined via convergence analysis. Due to the inclusion of multiple hazards and uncertainty sources, the minimum number of scenarios (N) needed for convergence is typically higher than in single-peril models.

The current method (shown in Figure 2) calculates damage separately and then combines it. In real life, shaking can weaken a building and make it less able to withstand a tsunami. The paper should discuss this complex damage process, or clearly state whether it uses a simple "max-damage" or "union" rule and what its limits are.

Thank you for this comment. The combination of damage states is handled through the damage state harmonization procedure, which maps damage from individual perils into a unified damage scale based on the fragility model definitions. This approach does not account for sequential weakening effects (e.g., earthquake damage increasing vulnerability to tsunami loads) except for structures which experience both significant liquefaction and ground shaking (as described in Section 3.2.7 in the main manuscript). We have clarified this limitation in the manuscript and added a reference to Petrone et al. (2020), which presents fragility functions for sequential earthquake–tsunami loading.

Lines 364 to 369:

In multi-hazard risk assessment, particularly when evaluating cascading impacts from cascading hazards, a key challenge arises from the need to integrate damage estimates derived from different fragility models, each with their own damage state definitions and granularity. This process, known as damage state harmonization, involves mapping and aligning the distinct damage classifications used for different perils into a unified framework. While this approach enables the combination of damage from multiple hazards, it does not explicitly account for sequential weakening effects, where damage caused by one hazard (e.g., earthquake shaking) may increase vulnerability to subsequent hazards (e.g., tsunami loading). Such path-dependent damage processes are therefore not represented in the current implementation, although they have been explored in previous studies (e.g., Petrone et al., 2020).

Crescenzo Petrone, Tiziana Rossetto, Marco Baiguera, Camilo De la Barra Bustamante, Ioanna Ioannou, Fragility functions for a reinforced concrete structure subjected to earthquake and tsunami in sequence, Engineering Structures, Volume 205, 2020, 110120, ISSN 0141-0296, https://doi.org/10.1016/j.engstruct.2019.110120.

The paper says more simulations are needed for results to settle than in single-hazard models. Please give a number or measure for this. In a multi-hazard model, the average damage might settle faster than the highest possible loss.

Thank you for this comment. We agree with it. Convergence is currently analysed and discussed in Section 3.2. We will add a sentence there to clarify that the concatenation of uncertainties across cascading perils generally requires more simulations than single-hazard models.

Lines 391-395

A convergence analysis was conducted to assess the statistical stability of mean damage ratios (MDR) derived from the stochastic seismic simulations. Convergence refers to the point at which the cumulative mean MDR stabilizes with minimal variation as the number of simulations increases—a principle rooted in Monte Carlo–based uncertainty quantification (Burt and Garman, 1971; Ata, 2007). In the multi-hazard context, this analysis is particularly important due to the concatenation of uncertainties across cascading perils, which generally requires a larger number of simulations compared to single-hazard models. Two convergence tests were performed.

The tsunami model uses a set of 330 pre-made scenarios based on simple slip patterns. There might be a difference between the detailed slip used for shaking and the simpler slip used for the tsunami scenarios. To keep things clear, the paper should discuss how the seafloor changes in the specific simulation compared to those in the scenario library.

The slip distributions used in these models to generate the tsunamis were generated stochastically as described in Appendix A2.The slip in these models varies across the rupture plane. The ground motions were calculated using GMPEs, as described in Appendix A1. GMPEs don’t specifically take the slip distribution into account. The effect of variations in the slip is accounted for in the uncertainty parameters in the GMPE. So it is the other way around. The tsunami models are the ones based on detailed slip distributions, not the ones used to calculate the shaking. This means that the seafloor changes in the tsunami simulation are identical to those in the scenario library. We refer the reviewer to both Section 3.2 and to the details of the method as presented in the Appendix for further information about this point.

The landslide hazard is shown as just one movement value. Landslides in this setting can occur immediately during an earthquake or later due to rain on weakened slopes. The method should make clear whether it models only landslides that occur during the earthquake.

We agree with the reviewer that different triggering mechanisms for landslides can exist. In this study, however, all landslides are earthquake-induced and triggered by ground shaking (and are therefore purely co-seismic). Other triggering mechanisms (e.g., slope weakening) and the associated landslides are not considered. We will clarify this point at the first mention of the landslides modelled in the manuscript.

The model shows liquefaction and lateral spreading as separate parts in the diagram, but these often happen together. It should be clear whether the model considers them together or treats them as separate ways buildings can be damaged.

Thank you for this comment. In the framework, liquefaction and lateral spreading are treated as separate hazards. While both are triggered by ground shaking and depend on geological (e.g., material type, age, cementation, and compaction) and geotechnical properties (e.g., soil strength and water saturation), lateral spreading additionally requires specific conditions such as ground slope or proximity to a free face.

As a result, the two processes are modelled independently, although they share a common triggering mechanism through the same ground shaking field. The resulting damage states are then combined through the damage state harmonization procedure described above.

Results

The results show total damage numbers but often do not compare the Average Annual Loss for each hazard, either on its own or when combined.

Please add a Peril Contribution Plot. This would let readers see when hazards like tsunamis start to cause more damage than ground shaking.

Thank you for this suggestion. We believe that Figures 8a and 8b already provide a clear representation of the contribution from each hazard. Figure 8a presents damage ratios by return period relative to the total building exposure, allowing comparison of the contribution of each hazard to overall losses. Figure 8b provides a complementary view, showing damage ratios relative only to buildings exposed to each respective hazard. We propose adding the text below with the specific AADR values per peril as well. Please see the paragraph below.

We would be happy to include additional plots or further clarification if required.

Revised text:
The estimated average annual damage ratios (AADR) for the buildings exposed to each hazard are 0.00568 for ground shaking, 0.00644 for shaking damage and induced liquefaction, 0.00107 for landslides, 0.00392 for lateral spreading, and 0.00029–0.00031 for tsunami inundation.

The results are based on the max-damage rule. Please test how sensitive the results are to this rule. If the model used a different way to combine damage, like adding up the chances of damage instead of just taking the worst case, how much would the total loss change?

We thank the reviewer for this valuable comment. To assess the sensitivity of the results to the damage combination rule, we tested alternative approaches for combining damage across hazards:

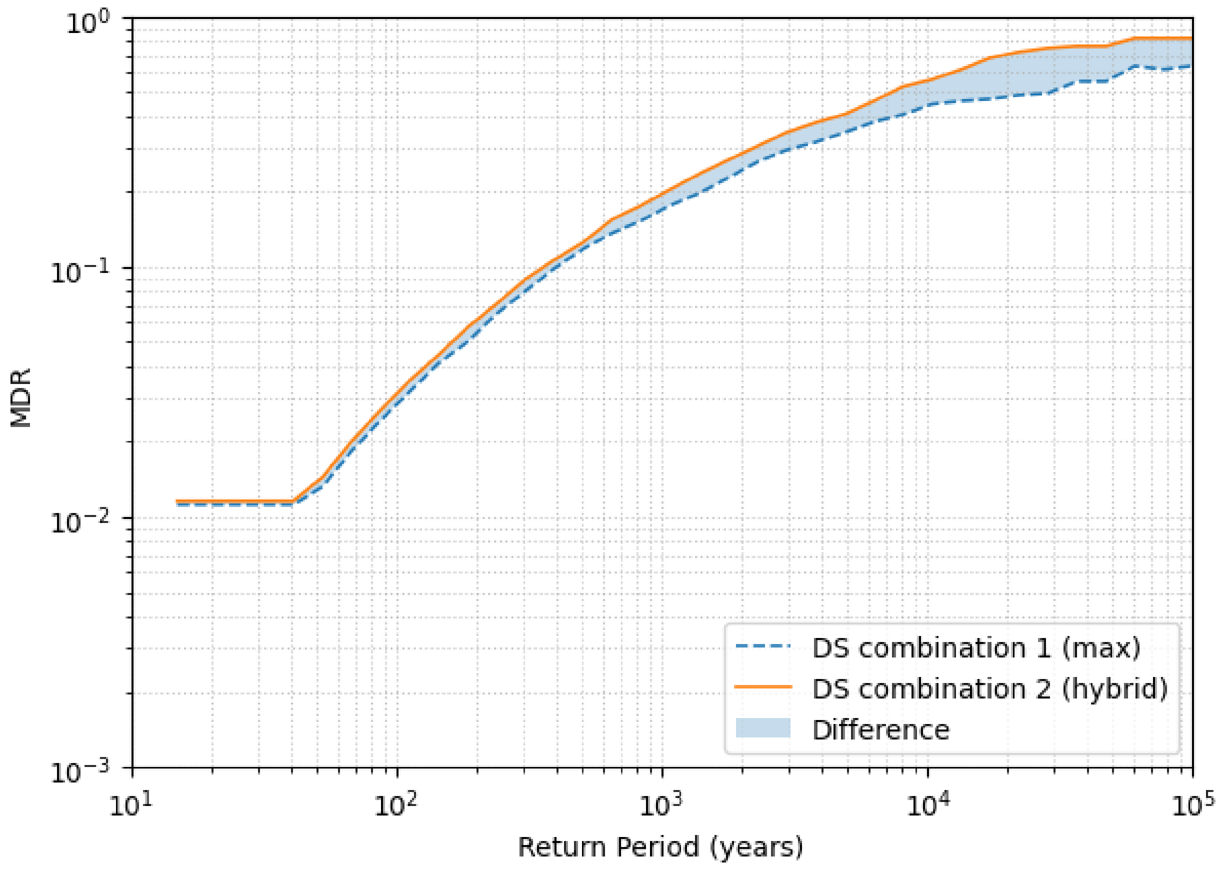
1- Maximum damage state (baseline approach): for each simulation, the highest damage state across all perils is selected, and the corresponding damage ratio is applied.

2- Hybrid additive approach: damage ratios are combined additively for lower damage states (DS1–DS3), while for higher damage states (>DS3), the maximum damage state is retained.

3- Mean damage ratio across perils: the average damage ratio across all hazards was also explored.

The mean-based approach was discarded, as it led to a systematic underestimation of losses by averaging damage across hazards, while in practice losses are driven by the most severe damage affecting each asset.

The comparison between the maximum and hybrid approaches is shown in Figure 4.2. Differences are negligible for return periods below approximately 150 years. Differences become observable above ~150 years, coinciding with the increasing contribution of landslides (see Figure 4.1), and become more pronounced beyond ~500 years due to the contribution of tsunami losses.



Please note this is not included in the manuscript, but can be included in the Supplementary material if required.

Discussion

To keep the calculations manageable, the tsunami model library was limited to 330 scenarios. During the simulation, the model picks the closest scenario from this set. The discussion should clearly measure the error caused by this choice. It is also important to discuss whether this approach might miss the most extreme risks, especially for events that do not fit neatly into pre-made scenarios.

An exact quantification of the effect of this approximation isn't possible without simulating a lot more than 330 events, making the problem once again intractable. Before simulating the tsunami, we checked that the 10 scenarios were diverse and covered a reasonable range of slip patterns (for example we made sure that all were not shallow or deep). Given the wide spread in the tsunami losses in the larger earthquakes, we feel we have successfully picked a diverse set of slip distributions so the average estimates of the loss should be reasonably accurate. We agree that the extreme values of loss (high or low) are uncertain, but this is likely to remain the case even if significantly more models were run.

The uncertainty in the extreme values in the loss is why extreme losses were not presented or discussed in the paper.

We note that all the perils have the problem that the MCS method may not capture the most extreme hazard intensity for that peril to various degrees. The way we handle this is to run the MCS process for long enough to ensure convergence in the average loss values as described in Section 3.2. Based on the convergence testing, we feel confident that the MCS simulation has been run long enough that average or median impact metrics should be robust for tsunami and all the other perils considered in this paper.

The framework uses a max-damage rule to combine damage from different hazards. This is a cautious simplification. The discussion should talk about how damage can depend on the order of events, for example, how shaking can make buildings more likely to collapse in a tsunami. Saying that future versions could include more detailed damage rules would give a clear plan for future research.

Thank you for this comment. We believe that we have addressed this point in a previous response (please see page 7)

The paper says that more simulations are needed for results to settle in multi-hazard models than in single-hazard ones. The discussion would be stronger with more detail on the trade-off between how quickly the model runs and how stable the results are.

Thank you for this comment. We believe that we have addressed this point in a previous response (please see page 4)

Please discuss the potential issues with using US-based HAZUS damage estimates for New Zealand buildings. The discussion should also suggest using local damage data in the future to improve the average damage estimates.

Thank you for this comment. Due to the limited availability of fragility models for New Zealand residential buildings, we adopted HAZUS 2.1 functions, supported by similarities with local timber building typologies. We acknowledge that using US-based models introduces uncertainty due to differences in construction practices and design standards. This limitation has been clarified in the manuscript, and we note that locally derived fragility functions would improve damage estimates in future work.

Revised text:
While HAZUS-based fragility functions provide a consistent and well-documented basis for damage estimation, their application to New Zealand building stock introduces uncertainty due to differences in construction practices, materials, and design standards. Although similarities exist for certain building typologies (e.g., timber structures), these models may not fully capture local vulnerability characteristics. The use of locally derived fragility functions, based on New Zealand-specific empirical or analytical data, would improve the representation of damage and reduce uncertainty in future applications.

The current model does not include other hazards that can occur at the same time (such as tides or storm surges) or long-term changes (such as sea-level rise). Talking about how rising sea levels from climate change would affect tsunami flooding in this model would make it more useful for long-term city planning, not just for earthquake risk.

Thank you for this comment. We will add a statement in the manuscript noting that these factors may influence tsunami inundation and that incorporating sea-level rise scenarios would improve the framework for long-term planning applications.

Revised text:
The current framework does not account for concurrent coastal processes (e.g., tides or storm surge) or long-term changes such as sea-level rise. These factors may influence tsunami inundation patterns and associated damage, particularly in low-lying coastal areas. Incorporating sea-level rise scenarios could modify inundation extents and depths, potentially increasing exposure and losses. Extending the framework to include these effects would improve its applicability for long-term risk assessment and climate-informed urban planning.

Conclusion

The paper says that risk increases as more hazards are included. Please give a simple scientific summary about how much the risk increases. For example, adding more hazards raised the average yearly loss compared to a model with only shaking. This gives other researchers a number to compare with different locations or building types.
Thank you for this suggestion. We believe that this point has been partially addressed in a previous response, where we reported the percentage contribution to the AADR (see page 6).

The results show that shaking and liquefaction are the main hazards in Napier. Please clearly say when this order of importance is true.

Thank you for this comment. The dominance of ground shaking and liquefaction observed in our results applies at the scale of the overall building portfolio considered in this study, reflecting the widespread exposure and local geological and geotechnical conditions in Napier. This order of importance is therefore context-dependent and may differ at smaller spatial scales or in areas with different exposure or site characteristics.

We would be happy to clarify this point in the discussion section if the reviewer considers it appropriate.

In the conclusion, talk about how much consistency was reached. While the hazards come from the same source, the ways buildings are damaged mostly stay separate.

We believe that these points have been addressed in our responses to the previous comments. We will ensure that this aspect is clearly reflected in the conclusions section of the manuscript.

Thank you again for your work,

Best regards.

We thank the reviewer again for their valuable comments and suggestions.