**Comment on nhess-2022-59**
Anonymous Referee #3


Review of Muhammad et al

This paper presents a methodology for time-dependent probabilistic tsunami hazard analysis with stochastic earthquake rupture modelling, using the Mentawai region of the Sunda Subduction Zone as a case study. This is a novel and ambitious approach, and it is exciting to see the efforts made by the authors. In my view, the complexity of the model does however pose some challenges, and I think there are a number of points that require further justification and/or consideration of the choices made in the model. I expect this will require some effort to revise the model.

Major comments

Justification of the choice of a time-dependent approach. A number of recent studies of global paleoeearthquake records (Williams et al 2019; Griffin et al 2020; Moernaut 2020) have, to varying degrees, provided empirical support for weakly quasiperiodic earthquake recurrence as a general model, which can be used to justify the use of renewal models for hazard assessment. That said, the Mentawai record of Philibosian et al (2017) looks to be more random than quasiperiodic in the analysis presented by Griffin et al (2020), although perhaps a different result might be obtained using the segmentation model presented.
The posterior BPT parameter estimates given for each segment are also relevant – some give values of alpha ~1 (segments 2, 3 and 4), implying random recurrence (i.e. Poisson), while others are ~0.6 (segments 1, 5 and 6), implying moderately quasiperiodic recurrence. So, I think some comment needs to be made here that:

- At a global scale there is empirical support for weakly quasiperiodic earthquake recurrence as a general model (see Griffin et al 2020);
- Excluding the hypothesis at the individual fault level is difficult, particularly for short records (Williams et al 2019; Griffin et al 2020)
- The data from Philibosian et al (2017) is somewhat equivocal about whether earthquake recurrence here is truly time-dependent, and the Poisson hypothesis cannot be confidently excluded using these data. But the global studies mentioned above suggest it is not unreasonable to assume time-dependence as a hypothesis.

The discussion section of the paper could then discuss the implications of this assumption in light of the different values of alpha obtained for each segment.

In estimating parameters for the BPT distribution, the authors use the data to estimate the prior distribution of mu, before then using the same data to calculate the posterior probability distribution of mu. This is incorrect. I would suggest using an uninformative prior (e.g. as used by Fitzenz et al 2010). An alternative approach could be to use an informative prior for mu based on the slip rate (e.g. as determined from geodesy), but this may become complex (e.g. due to having to estimate coupling of the fault). The 450 year long record is short for accurately estimate model parameters. This is, of course, what a Bayesian approach should be helping with, but needs more care about the choice of priors.

I am also concerned that fitting the model parameters to each segment individually is problematic. Later you consider multi-segment ruptures, and it is not clear how all this fits together. Do the recurrence statistics obtained from the sum of all synthetic ruptures across all segments match the recurrence statistics from the sum of all historic/paleo ruptures in your data? Checking this could be a good test for your model.
Also related to parameter estimation, some of the posterior histograms seem a bit spiky; does this improve if the number of samples is increased beyond 10,000?

Spatio-temporal completeness of the paleo record compared with the instrumental record is an issue that I think could lead to biases in the parameter estimates. It is very unlikely that events similar to the Mw 7.8 2010 Mentawai event would be visible in the coral record; this event occurred near the trench and caused <4 cm subsidence on the Mentawai Islands as measured with GPS (Hill et al 2012). Related to the above, the Mmin of 7.6 (L129), while reasonable from a tsunami hazard assessment perspective, would mean that you are modelling events that are unlikely to be present in the paleoeartquake record. I am unsure of how the frequency of these events could be determined in the time-dependent approach. Therefore it seems likely in your current approach that smaller events are missed in the paleoeartquake record, therefore affecting the recurrence model parameters.

The 1D rupture segmentation is a problem for tsunami hazard assessment, as the resulting tsunami size depends so significantly on the depth of rupture. Compare the 2007 Bengkulu earthquakes (Mw 8.4 and 7.9), that were down-dip of the trench and did not generate a significant tsunami, with the 2010 Mentawai earthquakes (Mw7.8), which occurred near the trench and did generate a significant tsunami. It is not clear whether such events are discriminated by the stochastic modelling approach with 1D segmentation – it seems they probably aren’t, but I may not be understanding correctly. A related problem is low-rigidity near the trench and its tsunamigenic potential, as in the 2010 Mentawai tsunami? How might the assumption of constant (and relatively high) rigidity (L309-310) bias your tsunami hazard results?

The maximum magnitude of 9.0 seems too low, which seems related to the segmentation model. If the potential for ruptures connecting with other segments of the Sunda Subduction Zone is considered, then larger Mmax values are justified. Significantly larger Mmax’s were used in Horspool et al (2014). Even if the paleoeartquake record for the past 450 years suggests events haven’t exceeded Mw 9.0, we also don’t expect these magnitude events to occur all that often. So allow for the possibility that they are missing from the record.
Some area of the coast of Padang show zero probability of inundation (Figure 17), while in others the potential inundation extent extends quite a way inland. This raises some significant concerns for me about the quality of the inundation modelling and/or the elevation data used, given how low-lying the coast is in this area. If only SRTM data was used, this could significantly underestimate inundation extent (see Griffin et al 2015, Figure 8). Are buildings included in the elevation model?

Detailed comments

L15: Suggest change ‘A total of >’ to ‘More than’

L18: Forecast periods begin in what year?

L136: Choice of BPT is fine, but hasn’t really been justified here. Why is this chosen over lognormal, Weibull or Gamma? Some of your justification seems to be presented later in Section 2.2.

L174: Several thousand years

L185: Perhaps rephrase as ‘reflects the expectations of elastic rebound theory’, or similar.
L192: Should probably cite others who've used Bayesian approaches to fitting time-dependent models to earthquake records, in particular Rhoades et al (1994) and Fitzenz et al (2010).

L197 and Table 1: These should not be referred to as tsunamigenic. For half of them we have no information on whether a tsunami was generated; coseismic deformation on the Mentawai Islands observed in coral paleogeodetic records suggests they probably were, but we don't actually know.

L324. Please give a link or citation for DEM5 and Bathy5.

L332: Might be a typo here – Griffin et al (2016) used a Manning’s roughness of 0.036 as a conservative minimum for land (grassland; for the Mentawai Islands). For the urban context here, 0.06 may be reasonable, e.g. Griffin et al (2015) suggested a Manning’s roughness of 0.08 for the city of Padang. See also Kaiser et al (2011) for a discussion of choice of Mannings n.

Ling 501-502: The time-independent model has too low an Mmax (9.0) to be considered worst-case. See earlier comments about choice of Mmax.
Table 1: Change Shieh to Sieh.

Figure 3, and also in the text. I do not think the term ‘occurred’ scenarios is the best terminology. These are modelled scenarios that have not actually occurred.

References:


