

Nat. Hazards Earth Syst. Sci. Discuss., author comment AC1  
<https://doi.org/10.5194/nhess-2021-306-AC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC1

Jack N. Williams et al.

---

Author comment on "Geologic and geodetic constraints on the magnitude and frequency of earthquakes along Malawi's active faults: the Malawi Seismogenic Source Model (MSSM)" by Jack N. Williams et al., Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2021-306-AC1>, 2022

---

We thank the reviewer for their constructive and positive review of our study . In the text below, we have copied their comments, and then replied to them in italicised text.

### General comments:

In the introduction, the authors give a good description of the use of fault databases in the framework of seismic hazard and risk assessment. While the results of this paper, in the shape of a seismogenic source database, are a major component of seismic hazard assessment, there are not seismic hazard results themselves. Therefore, my opinion is that the title of this paper should be modified and the words "seismic hazard" should be removed, since no hazard results are presented in the paper.

*We agree, and if invited to submit a revised version of this manuscript, we will change the title to 'Geologic and geodetic constraints on the magnitude and frequency of earthquakes along Malawi's active faults: The Malawi Seismogenic Source Database (MSSD)'*

In order to allow the results of these study to be used in a PSHA study, the way the weighting of the different source types should be done needs to be better discussed. It is not clear how the earthquake rate from the different source types should be combined. Should the weighting be done in order to fit a given MFD for the entire system? I would be useful if the authors added a section in the discussion part of the article to clarify this point. If possible, a comparison of their computed earthquake rates with the rates calculated using the earthquake catalogue could be added.

*We thank the reviewer for their insightful comments, and agree that how different source types are weighted is a significant source of uncertainty when incorporating the MSSD into probabilistic seismic hazard analysis (PSHA). In revising this manuscript, we will refer the reader to a subsequent study in which we have used the MSSD in PSHA, and have explicitly explored this uncertainty. This study is currently in review with Natural Hazards, and a pre-print of it can be found here:*

*Williams, J., Werner, M., Goda, K., Wedmore, L., De Risi, R., Biggs, J., ... & Chindandali, P. (2022). Fault-based Probabilistic Seismic Hazard Analysis in Regions with Low Strain Rates and a Thick Seismogenic Layer: A Case Study from Malawi., 27 March 2022, PREPRINT (Version 1) available at Research Square*

[<https://doi.org/10.21203/rs.3.rs-1452299/v1>]

*In summary, in this manuscript we used the MSSD to simulate stochastic earthquake event catalogs with a 2-million-year duration, and in which earthquakes occurred randomly on each MSSD source following a memoryless Poisson process. We then generated catalogs for all possible source type weighting combinations at increments of 0.1, and with the constraint that the weighting of any source type must be  $\geq 0.1$  ( $n=36$ ). We then selected the combination that, for the magnitude range 6-7.6, produced a catalog with the closest  $b$ -value to the observed  $b$ -value in Malawi as derived from instrumental seismicity (1.02; Poggi et al 2017). Following this, we selected a weighting of 0.6-0.3-0.1 for section, fault, and multifault sources respectively, which is qualitatively consistent with the inference that there must be relatively frequent small magnitude section source events to maintain a  $b$ -value  $\sim 1$ . This is discussed further in Section 4.4 and Figure 7 in Williams et al 2022.*

*Williams et al 2022, we also provide an analysis for how the moment rate ( $M_0$ ) of the synthetic earthquake catalogs generated using the MSSD compared to the instrumental catalog  $M_0$  in Malawi. This analysis concludes that although the  $M_0$  of the synthetic catalogs are 2x higher than the observed  $M_0$  in Malawi, this can be accounted for by the incomplete nature of the instrumental record in Malawi, which in turn is indicative of the low slip rate of its faults, the inference that they may be locked and/or host clustered earthquakes, and that the instrumental catalog has only a short (<60 years) duration ( see section 4.6 in Williams et al 2022).*

*We note that given that the reviewer has raised these points on this study about the MSSD, not the PSHA, we had considered whether some of the above discussion could be incorporated into this current study. However, we concluded that this would exacerbate the length of this study, and that describing the MSSD and PSHA separately remains the best strategy for presenting this work. Nevertheless, in the revised manuscript, we will surmise the above discussion with reference to this (in review) PSHA, and how it shows how different source types in the MSSD can be used in future. In either case, our PSHA shows the importance of using geologic and geodetic data to constrain the activity of faults in Malawi, and hence the importance of the MSSD's development.*

### **Specific comments :**

Line 169 – I suggest modifying the term “statistical treatment” by “exploration”

*We will make this change in the revised manuscript*

Line 218 – Simplifying the surfaces of the faults is a potentially impactful hypothesis in terms of hazard assessment. The change in the surface can both affect the moment rate estimate for the fault and the distance taken into account in the GMPEs. While it is possible that the complexity observed in the fault trace might not be present at depth, the straight line is the other end-member of the possibilities for the fault surface. Why not let the final user of the database, the hazard modeller, choose the level of simplification to be applied? Especially since modern PSHA codes can now handle rather complicated geometries.

*We agree that the way in which we simplify the fault geometries from the Malawi Active Fault Database (MAFD) into the MSSD is subjective. However, many of the attributes that are assigned to sources in the MSSD are calculated using the simplified geometries (e.g., earthquake magnitude, recurrence interval in equations 4 and 5). Hence, if a final user wishes to change the source geometry, for consistency, these attributes will also need to*

*be revised. For the MSSD to incorporate multiple interpretations of the MAFD would be impractical. In the revised manuscript we will therefore emphasise that alternative interpretations of the MAFD for seismic hazard assessment are possible. Furthermore, since the MAFD is also openly available (Williams et al 2022, <https://zenodo.org/record/5507190#.YkypWyORpB3>), final users are welcome to choose their own level of source geometry simplification, although this will require changing other source attributes. In our revisions, we will specify that the simplified source geometries in the MSSD mean they should not be used in instances where accurate fault traces at seismogenic depths are available (e.g., site-specific engineering development, assessment of surface rupture hazards).*

Line 230 – Would it be possible to add the uncertainty on the dip in the database ? This parameter can be source of large uncertainties in the hazard levels, and since the knowledge of the dip is not uniform in the system, adding the uncertainty on each fault could be useful.

*We will follow the reviewer's recommendation in the revised manuscript and add dip uncertainty as attributes in the MSSD. It should, however, be noted that our model of source geometry only considers the intermediate dip estimate.*

Line 278 – Simplifying the fault system by removing splay faults also implies to consider that the whole deformation is accommodated by the main fault. Since the metrics used in GMPEs don't usually take into account such details, the impact on the hazard would probably be minimal or within the simplification already made when using a GMPE. However, the impact of the simplification on the deformation should be commented in the text.

*We will address this comment in the revised manuscript by noting that in removing subsidiary splays, the slip rates reported in the MSSD are essentially the cumulative slip rate from both the main fault and all its associated splays.*

Line 441 – A point is missing.

*We will correct this in the revised manuscript.*

Line 455 – Some underlying assumptions behind these results should be stated here, even if there are discussed later in the article. These recurrence intervals are obtained assuming that the slip-rate is fully seismogenic. It is also assumed that each source can only host on magnitude (for one branch of the logic tree), but other magnitude frequency distributions could be possible.

*We agree and will be more explicit about the assumptions of seismic coupling and fault rupture types when describing our results in the revised manuscript.*

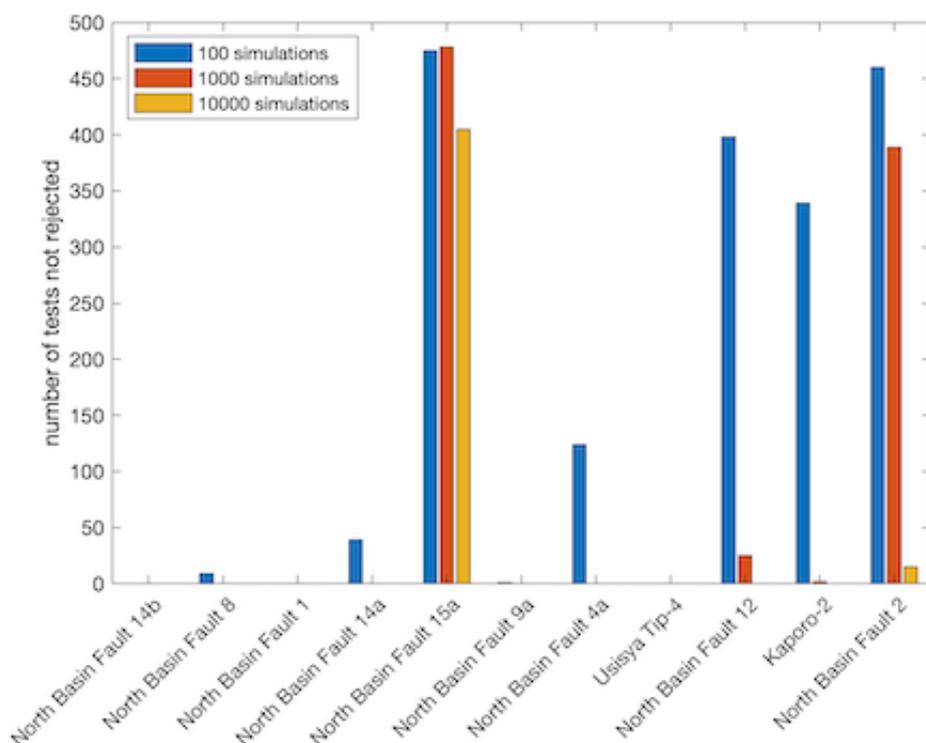
Table 3 – In this table, it is not very clear if the values are for one specific fault or for the

system as a whole. If it is for the system as a whole, can the different lines be read together? For example, is the table saying that the mean recurrence of a M6.8 earthquake is 10900 years? The legend of the table should be better detailed.

*The reviewer is correct to point out that in the initial submission, Table 3 was not described in sufficient detail. In the revised manuscript, we will clarify in the legend that the values we provide are calculated from the intermediate estimates of all MSSD sources of the given type (e.g., the section magnitudes minimum, mean, and maximum values consider the magnitudes of all section sources in the MSSD). We will also emphasise that the recurrence intervals are calculated assuming that each source only ruptures in that type.*

Line 465 - The 5% threshold is probably too severe for this type of analysis. For some fault the two distributions are very similar, and the difference are minimal, sometime affecting only the width of the distribution, but the mean values are similar. The discussion in the following paragraph is probably more useful in order to understand the difference between the slip-rate estimates.

*We acknowledge that it is surprising how many of the t-tests reject the null hypothesis that the two distributions come from probability distributions with the same mean but unequal variances. Our interpretation of this result is that since the variance of each slip rate distribution is high, many (10,000) Monte Carlo simulations must be run to achieve stable results. In other words, if less simulations are run, the result of the t-test changes each time we perform the analysis (see Figure 1 below). This large number of simulations entails that the p-value is very sensitive to even small differences in the mean between the two distributions, and hence the null hypothesis is rejected in cases when the mean values of each distribution qualitatively appear to be quite similar. In this respect, it is worth noting that z-tests may be considered more appropriate for large data sets. However, in this case, the variance of both sample distributions should be the same, which is not true for the two samples we are comparing; hence our preference to retain the two-sample t-test.*



*Figure 1: Sensitivity of T-test results to the number of Monte Carlo simulations (100, 1000, or 10000) that are performed when sampling slip rate distributions for the comparison described in Section 3.5 in the main text. Here each set of Monte Carlo simulations and the T-test is repeated 500 times. We then determine the number of tests that are accepted depending on whether 100, 1000, or 10,000 simulations are performed when sampling slip rates. Results from the T-test are stable when either 0 or 500 tests are rejected.*

*As discussed for the following comment, we will report p-values for each test in Fig 8 in the revised text. However, for 10 out of 11 cases, the p value is  $<0.01$ , and so changing the significance level will not influence our finding that in all but one case, the null hypothesis is rejected. We will incorporate the above discussion in the revised manuscript.*

Figure 8 - The authors should add indexes to these figures, so each individual fault could be identified on the map in figure 2 and in the database. Additionally, the t-test result value could be added to the figure, helping to understand the reason why one is accepted and not the others.

*We agree, and will make the necessary revisions to the figure/*