Fault-based probabilistic seismic hazard analysis in regions with low strain rates and a thick seismogenic layer: a case study from Malawi


1School of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK. E-mail: jack Williams@otago.ac.nz
2School of Earth and Environmental Sciences, Cardiff University, Cardiff CF10 3AT, UK
3Department of Geology, University of Otago, Dunedin NZ 9054, New Zealand
4Department of Earth Sciences, Western University, Ontario N6A 5B7, Canada
5Department of Civil Engineering, University of Bristol, Bristol BS8 1TR, UK
6Malawi Geological Survey Department, P.O. Box 9, Mzuzu, Malawi
7Geography and Earth Sciences Department, University of Malawi, P.O Box 280, Zomba, Malawi
8Malawi Geological Survey Department, P.O. Box 27, Zomba, Malawi

Accepted 2023 February 7. Received 2023 February 7; in original form 2022 May 30

SUMMARY

Historical and instrumental earthquake catalogs in low strain rate regions are not necessarily indicative of the long-term spatio-temporal distribution of seismicity. This implies that probabilistic seismic hazard analysis (PSHA) should also consider geologic and geodetic data through fault-based seismogenic sources. However, it is not always clear how on-fault magnitude-frequency distributions (MFDs) should be described and, if the seismogenic layer is especially thick, how fault sources should be extrapolated down-dip. We explore these issues in the context of a new PSHA for Malawi, where regional extensional rates are 0.5–2 mm yr\(^{-1}\), and the seismogenic layer is 30–40-km thick, the instrumental catalog is ~60 yr long and fault-based sources were recently collated in the Malawi Seismogenic Source Model. Furthermore, Malawi is one of several countries along the East African Rift where exposure to seismic hazard is growing, but PSHA does not typically consider fault sources. We use stochastic event catalogs to explore different fault source down-dip extents and MFDs. Our PSHA indicates that hazard levels are highest for a Gutenberg–Richter on-fault MFD, even at low probabilities of exceedance (2 per cent in 50 yr), whilst seismic hazard levels are also sensitive to how relatively short (<50 km) fault sources are extrapolated down-dip. For sites close to fault sources (<40 km), seismic hazard levels are doubled compared to previous instrumental-seismicity based PSHA in Malawi. Cumulatively, these results highlight the need for careful fault source modelling in PSHA of low strain rate regions and the need for new fault-based PSHA elsewhere in the East Africa Rift.

Key words: Earthquake hazards; Statistical seismology; Continental tectonics: extensional.

1 INTRODUCTION

In low strain rate regions, long-term deformation rates may not be adequately captured by historical and instrumental earthquake catalogs (Lombardi & Marzocchi 2007; Clark et al. 2012; Stevens & Avouac 2021; Iturrieta et al. 2022). Therefore, probabilistic seismic hazard analysis (PSHA) that uses these catalogs to develop areal or smoothed-seismicity sources (e.g. Zahran et al. 2015; Goitom et al. 2017; Poggi et al. 2017) will be limited by the data available to constrain future earthquake rates. Incorporating geodetic and geologic data in PSHA through fault-based seismogenic sources can partly address this challenge. However, careful treatment of fault sources is required in low strain rate regions as estimated seismic hazard levels are highly sensitive to assumptions about their geometry and segmentation (Hodge et al. 2015; DuRoss et al. 2016; Gómez-Novelli et al. 2020; Valentini et al. 2020; Visini et al. 2020; Goda & Sharipov 2021). A promising approach for representing fault segmentation in PSHA is inversion techniques that constrain the rate of all plausible ruptures in an interconnected fault system (Field et al. 2014, 2021; Geist & Parsons 2018; Chartier et al. 2019; Gerstenberger et al. 2022). However, various constraints on earthquake rates from geologic, geodetic and seismologic data must be satisfied to apply these techniques and this will be especially challenging in low strain rate regions where such data is scarce and...
earthquake rates may be non-stationary (Cox et al. 2012; Hodge et al. 2015; Vallage & Bollinger 2020; Stevens & Avouac 2021; Iturrieta et al. 2022).

The use of fault-based sources in PSHA also requires extrapolating geologic and geodetic constraints on fault deformation through the Earth’s crust. In particular, assumptions about a fault’s down-dip geometry and extent can influence its seismic moment rate, interaction with neighbouring faults and source-to-site distance in ground motion calculations. In most continental regions, the seismogenic layer is 10–20-km thick (e.g. Jackson et al. 2021) and so only a few studies have explicitly considered this uncertainty in PSHA (Wu et al. 2017; Ellis et al. 2021). However, assumptions for how faults are extrapolated down-dip in PSHA of regions where the seismogenic layer is much thicker (20–40 km), such as in the East African and Baikal rifts (Nyblade & Langston 1995; Déverchère et al. 2001; Lavayssiére et al. 2019; Craig & Jackson 2021), have not been assessed.

In this study, we present a new fault-based PSHA for Malawi, which is located within the Western Branch of the East African Rift (EAR). The seismogenic layer in Malawi is estimated to be ~30–40-km thick (Jackson & Blenkinsop 1993; Nyblade & Langston 1995; Ebinger et al. 2019; Stevens et al. 2021) and though geologically derived EAR extension rates are low (0.5–2 mm yr⁻¹; Stamps et al. 2021; Wedmore et al. 2021), they are an order of magnitude higher than inferred from the last 50 yr of instrumentally recorded seismicity (Hodge et al. 2015; Ebinger et al. 2019). Hence, it provides an ideal case study to investigate PSHA in a low strain rate region with a thick seismogenic layer. Furthermore, constraints on the seismogenic potential of Malawi’s active faults have improved with the collection of new geologic and geodetic data (Hodge et al. 2019; Scholz et al. 2020; Shillington et al. 2020; Wedmore et al. 2020a, b, 2021; Kolawole et al. 2021; Williams et al. 2022c), which have been synthesised into fault sources for PSHA in the Malawi Seismogenic Source Model (MSSM; Williams et al. 2022b).

Malawi is also one of several countries along the EAR where seismic risk is being exacerbated by population growth, rapid urbanisation and the development of seismically vulnerable building stock (Goda et al. 2016; Meghraoui et al. 2016; Poggi et al. 2017; World-Bank 2019; Kloukinas et al. 2020; Giordano et al. 2021). However, PSHA in the EAR still mainly considers the relatively short historical (150–600 yr) and instrumental (< 60 yr) record of seismicity alone (e.g. Midzi et al. 1999; Bwambale et al. 2016; Delvaux et al. 2017; Goitom et al. 2017; Poggi et al. 2017; Tuluka et al. 2020; Msabi & Ferdinand 2021). Here, we incorporate the MSSM into a PSHA for Malawi through stochastic event catalogs (Musson 1999; Atkinson & Goda 2013) to provide critical inputs for assessing Malawi’s increasing seismic risk Goda et al. (2016, 2022) and to examine how uncertainty in the thd down-dip extent and segmentation of fault sources influences PSHA in low strain rate regions.

2 BACKGROUND TO SEISMIC HAZARD ASSESSMENT IN MALAWI

2.1 The seismotectonic setting of Malawi

Malawi’s national borders are closely aligned to a 900-km long section of the EAR’s Western Branch, with earthquakes of moment magnitude (Mw) > 4.5 and active faults > 50-km long documented throughout (Fig. 1; Dixey 1926; Ebinger et al. 1987; Specht & Rosendahl 1989; Chapola & Kaphwiyo 1992; Poggi et al. 2017; Wedmore et al. 2020a; Williams et al. 2022c). In central and northern Malawi, the EAR has mostly been flooded by Lake Malawi, whilst in southern Malawi the rift is onshore and at its southern end has intersected and reactivated Karoo (i.e. Triassic-Jurassic) age faults (Dulanya 2017; Wedmore et al. 2020b; Kolawole et al. 2021; Williams et al. 2021a; Dulanya et al. 2022). Only negligible amounts of melt have been inferred from geophysical observations of Malawi’s crust (Ninju et al. 2019; Accardo et al. 2020; Hopper et al. 2020) and so rift extension is primarily accommodated by normal fault earthquakes (Biggs et al. 2010; Hodge et al. 2015; Ebinger et al. 2019; Williams et al. 2022b).

The most comprehensive instrumental record of earthquakes in Malawi is the Sub-Saharan Africa Global Earthquake Model (SSA-GEM) catalog (Fig. 1a; Poggi et al. 2017). This was mainly developed from the International Seismological Centre (ISC) catalog, which in Malawi is complete since 1965 for events Mw > 4.5 (Hodge et al. 2015). Within this record, two events stand out: the 1989 Mw 6.3 Salima Earthquake and the 2009 Karonga earthquake sequence. The former was assigned a VIII on the Modified Mercalli Intensity Scale (MMI; Gupta & Malomo 1995) and its 32 ± 5 km focal depth is typical of Malawi’s 30–40-km thick seismogenic layer (Jackson & Blenkinsop 1993; Nyblade & Langston 1995; Ebinger et al. 2019; Craig & Jackson 2021). In contrast, the Karonga earthquake sequence primarily consisted of four shallow (focal depths 5–10 km) Mw 5.5–5.9 events over 13 d and resulted in a 9–18-km long surface rupture along the previously unrecognized St Mary Fault (Biggs et al. 2010; Hamiel et al. 2012; Machyekye et al. 2015; Kolawole et al. 2018b). There are records of M 5–7 earthquakes in and around Malawi in the early 20th century; however, the lack of instrumentation in East Africa during this period mean the locations and magnitudes assigned to these events have large uncertainties (Ambraseys 1991; Poggi et al. 2017; Wedmore et al. 2022). Focal mechanism stress inversions indicate a normal fault stress state in Malawi with an ENE-WSW trending minimum principal compressive stress (Delvaux & Barth 2010; Ebinger et al. 2019; Williams et al. 2019).

2.2 Previous seismic hazard assessment in Malawi

In a PSHA for Malawi that considered areal sources developed from the SSA-GEM catalog, Poggi et al. (2017) found that for peak ground acceleration (PGA) there is a 10 per cent probability of exceeding (PoE) 0.10–0.15 g in 50 yr. Hazard levels were relatively uniform across Malawi in this study as it is largely covered by a single ~380 000 km² areal source zone that extends from Mozambique to southern Tanzania. However, geodetic models indicate that EAR extension rates increase from south to north Malawi (Saria et al. 2014; Stamps et al. 2021; Wedmore et al. 2021) and there are also geologic observations of across-rift variations in fault activity (Accardo et al. 2018; Shillington et al. 2020; Wedmore et al. 2020a). These spatial variations in deformation were incorporated into a PSHA by Hodge et al. (2015), who developed seven fault-based seismogenic sources in Malawi using previously mapped rift-bounding faults (Flannery & Rosendahl 1990) and geodetically derived regional extension rates (Stamps et al. 2008). Inclusion of these sources into PSHA resulted in higher hazard levels adjacent to these faults [10 per cent PoE 0.15-0.25 g in 50 yr vs 10 per cent PoE 0.10–15 g in 50 yr in Poggi et al. (2017)], with these increases greatest for cases that inferred relatively frequent, moderate magnitude ruptures of short discrete fault segments, rather than rarer, larger magnitude earthquakes spanning entire faults.
The PSHA presented by Hodge et al. (2015) was later incorporated into a quantitative seismic risk assessment (Goda et al. 2016), which highlighted that Malawi was at serious risk of building collapse during moderate-large magnitude earthquake (>10 000 buildings collapse for MMI > 6.5 shaking). The high exposure to moderate ground motions in Malawi is indicative of its seismically vulnerable building stock (Giordano et al. 2021; Novelli et al. 2021). Currently, no provisions are made for seismic loading in the official masonry construction code of practice in Malawi (MS791-1:2014; Malawi Bureau of Standards Board 2014). Earthquakes are qualitatively acknowledged for informal buildings through the Safer House Construction Guidelines (Cassani et al. 2016); however, considerable challenges exist in applying and enforcing these guidelines (Ngoma et al. 2019; Novelli et al. 2021).

Since 2015, high resolution digital elevation models (Hodge et al. 2019, 2020; Wedmore et al. 2020, a), aeromagnetic and gravity data (Kolawole et al. 2018a, 2021; Chisenga et al. 2019) and a new generation of seismic reflection data in Lake Malawi (Shillington et al. 2016, 2020; Scholz et al. 2020) have led to significant advances in the identification and mapping of active faults in Malawi. These data sets were combined into the Malawi Active Fault Database (MAFD), a geospatial database for 113 faults that are inferred to be active in Malawi and neighbouring areas of Mozambique and Tanzania (Williams et al. 2021b, 2022c). In addition, new constraints on fault slip rates in Malawi have been provided from new geodetic data (Stamps et al. 2018; Wedmore et al. 2021) and the fault offsets of a 75 ka reflector in seismic reflection data in Lake Malawi (Shillington et al. 2020). These new data were combined into the MSSM, a database that provides slip rate, earthquake magnitudes

Figure 1. Map of Malawi in the context of (a) the Malawi Active Fault Database (MAFD; Williams et al. 2022c) and the Sub-Saharan Africa Global Earthquake Model (SSA-GEM) catalog Poggi et al. (2017) and (b) regional geological terranes Fullgraf et al. (2017). BMF; Bilila-Mtakataka Fault.
and recurrence interval estimates of faults included in the MAFD (Williams et al. 2022a, b).

3 PSHA WORKFLOW

Here we combine the MSSM with previously defined EAR areal sources (Poggi et al. 2017) to develop a new PSHA for Malawi. We incorporate the MSSM (v1.2) by considering both the earthquake magnitude and recurrence interval assigned to each MSSM source (the ‘Direct MSSM’ approach) and a moment rate balancing approach that explicitly explores different hypotheses for the down-dip extent of the MSSM sources and whether they exhibit G–R or characteristic earthquake magnitude-frequency-distribution (MFD; the ‘Adapted MSSM’ approach; Youngs & Coppersmith 1985; Convertito et al. 2006). These different MSSM interpretations are realised in five stochastic event catalogs, with off-fault events considered using an areal source based catalog. The PSHA is then formulated by evaluating these five catalogs using four ground motion models (GMMs; Atkinson & Adams 2013; Akkar et al. 2014; Boore et al. 2014; Chiou & Youngs 2014). Hence, for a given site, PoE and spectral period, we calculate 20 ground motion parameters. Following an ensemble approach, we use the mean and distribution of these values to describe seismic hazard and its uncertainty (Marzocchi et al. 2015; Meletti et al. 2021).

Our analysis is performed for a rectangular region that bounds Malawi with a grid spacing of 0.2° (Fig. 1). For each grid point, we consider two values for the average shear wave velocity to 30 m depth (V30): (1) a reference site condition (V30 = 760 m s\(^{-1}\)) and (2) the value derived from the USGS V30 database (Wald & Allen 2007). In addition, we performed site specific PSHA for the three largest cities in Malawi: Lilongwe, Blantyre and Mzuzu (Fig. 1a). We describe the earthquake sources and stochastic event catalogs further in Section 4, the GMMs in Section 5.1 and seismic hazard calculations in Section 5.2. PSHA results are presented in Section 6. A summary of our PSHA workflow is shown in Fig. 2. Abbreviations and symbols are listed in Table 1.

4 SOURCE MODELS

4.1 The Direct approach to the MSSM

The MSSM (v1.2) is a geospatial database of 140 geometrically defined section sources, 108 fault sources and 27 multifault sources that were identified from the 113 faults contained within the MAFD (Fig. 3; Williams et al. 2021b, 2022c). The number of faults in the MAFD and fault sources in the MSSM are not the same due to the requirement that sources are ≥5-km long and that faults that splay in map view are considered to represent different sources. A full description of the MSSM is provided by Williams et al. (2022b) and so we only briefly summarise the parameters of interest for PSHA here. An earthquake magnitude (\(m_s\)) and single event displacement (\(\bar{D}_s\)) estimate is assigned to each MSSM source using the Leonard (2010) magnitude–area scaling relationships for interplate dip-slip faults. For these calculations, source width (\(W_s\)) is derived through

\[
W_s = \begin{cases} 
c_1 L_s^{\frac{1}{2}}, & \text{if } c_1 L_s^{\frac{1}{2}} < \frac{z}{\sin \delta} \\
\frac{z}{\sin \delta}, & \text{if } c_1 L_s^{\frac{1}{2}} \geq \frac{z}{\sin \delta} 
\end{cases}
\]

where \(c_1\) is an empirically derived parameter and equals 17.5 metres\(^2\) for interplate dip-slip faults (Leonard 2010), \(L_s\) and \(\delta\) are source length and dip and \(z\) is the seismogenic layer thickness. Previous studies have estimated that \(z\) is 30-40 km in Malawi (Ebinger et al. 2019; Stevens et al. 2021) and we apply an intermediate estimate of 35 km here. If two sources intersect down-dip, the shorter source is assumed to be truncated by the longer one and its \(W_s\) accordingly revised (Scholz & Condreras 1998).

For some sources in Lake Malawi, slip rates (\(S_s\)) have been derived from offsets across a 75 ka reflector in seismic reflection surveys (Shillington et al. 2020). Elsewhere, \(S_s\) are estimated using a systems-based approach that incorporates geodetically derived regional extension rates (Wedmore et al. 2021; Williams et al. 2021a, 2022b). The recurrence interval (\(R_s\)) for a full source earthquake with magnitude \(m_s\) is then derived by combining the slip rate (\(S_s\)) and displacement (\(D_s\)) through the relationship: \(R_s = D_s / S_s\) (Wallace 1970).

4.2 The Adapted approach to the MSSM

The discrete section, fault and multifault sources in the MSSM are not an exhaustive list of potential earthquake ruptures in Malawi; in reality earthquakes can ‘float’ within a larger fault network (Visini et al. 2020; Field et al. 2021). In addition, eq. (1) suggests that a source’s down-dip extent is dependent on its length (‘length-limited’); however, we cannot exclude the possibility that unless intersected by another source, all the MSSM sources propagate through Malawi’s 35-km thick seismogenic layer (‘layer-limited’). This uncertainty is raised further in Malawi by: (1) possible lateral variations in the lower crust’s composition and strain rate that can locally modulate whether the down-dip extent of faults is seismic or aseismic (Fagereg 2013; Hellebrekers et al. 2019; Wedmore et al. 2020a) and (2) intrarift faults in Malawi may accommodate upper-crustal flexural extensional strains that are induced from bending in the hanging-wall of large displacement (>5 km) border faults (Turcotte & Schubert 1982; Billings & Kattenhorn 2005; Kolawole et al. 2018a; Shillington et al. 2020; Williams et al. 2022b). Another possibility is that large earthquakes propagate below the seismogenic layer (‘dynamic overshoot’; Shaw 2013; Ellis et al. 2021). We do not explicitly consider dynamic overshoot in this PSHA, though its implications are discussed in Williams et al. (2022b).

We explore the uncertainty on length- or layer-limited source down-dip extents using an ‘adapted’ approach to the MSSM. We do this by calculating a source’s seismic moment release rate (\(M_0\)) through

\[
\dot{M}_0 = \mu S_s A_s
\]

where \(\mu\) is crustal rigidity and is taken as 33 GPa for consistency with the Leonard (2010) scaling relationships and \(A_s\) is source area and is accordingly adjusted for length- and layer-limited width cases (Fig. A7). We then combine the source’s \(\dot{M}_0\) with a \(b\)-value to develop continuous recurrence models that follow a G–R or characteristic MFD and allow ruptures to float anywhere within the fault plane (Figs 4 and 5; Youngs & Coppersmith 1985; Convertito et al. 2006; Goda & Sharipov 2021). For each source, MFD and width case, nine recurrence models are generated to incorporate uncertainty in the \(b\)-value and the source’s largest magnitude event (\(M_{\text{max}}\); Figs 2 and A1). The equations that allow us to balance the \(\dot{M}_0\) through source-specific magnitude probability distribution functions and earthquake rates and that were derived by Youngs &
Coppersmith (1985) and Convertito et al. (2006), are provided in Appendix A2.

Since the Adapted MSSM source models are continuous across a range of magnitudes, the division of the MSSM into discrete section, fault and multifault sources is not necessary. In other words, section source seismicity is already incorporated into the MFD of the larger fault or multifault source that they belong to (Fig. 4). Likewise, if a MSSM fault source is a constituent of a multifault source. In the Adapted MSSM approach we therefore only consider 79 sources; all multifault sources and fault sources that are not part of a multifault source.

4.3 Areal sources

We use the areal source zones developed for East Africa by Poggi et al. (2017) to incorporate: (1) earthquakes on unknown faults in Malawi, (2) earthquakes on faults included in the MAFD but not in the MSSM due to their short length (<5 km) and (3) earthquakes in regions adjacent (< 200 km) to Malawi, where no fault-based earthquake sources have been developed (Fig. 6). These areal source zones are defined by a truncated exponential G–R relationship that is fitted to the seismicity recorded in each area in the SSA-GEM catalog (Table 2; Poggi et al. 2017).

There are regions to the southwest and east of Malawi that are not covered by the Poggi et al. (2017) sources. However, since their seismic hazard is non-zero, we define areal sources for these regions by adjusting global rates of stable craton seismicity to their respective areas (Fig. 6, Table 2; Fenton et al. 2006). Strictly speaking these regions do not meet the criteria of ‘stable cratons’ set by Fenton et al. (2006) as they are within 200 km of passive margins and/or regions of Phanerozoic deformation. However, given the lack of recorded earthquakes in these regions, these estimates remain the best constraint on their seismicity.

Areal sources in this PSHA provide additional challenges. First, areal sources within Malawi require an upper magnitude bound ($M_{\text{Max}}$) that is indicative of the largest earthquake that could occur on an unmapped fault. Here, we set $M_{\text{Max}}$ to 7.0, which is equivalent to a 40-km long fault for the Leonard (2010) scaling. This is guided by: (1) comparisons to other national-scale seismic hazard models where $M_{\text{Max}}$ ranges between 6.5 and 7.5 (Stirling et al. 2012; Field et al. 2014; Woessner et al. 2015; Wang et al. 2016) and (2) the inference from its thick seismogenic layer and lack of chronostratigraphic data that there are relatively large unmapped active faults in Malawi (Williams et al. 2022c).

Secondly, high earthquake location uncertainties in Malawi (5–20 km; Jackson & Blenkinsop 1993; Gaherty et al. 2019) mean it is not possible to filter out ‘on-fault’ earthquakes from the
Table 1. List of acronyms and symbols used in this study.

<table>
<thead>
<tr>
<th>Acronym/symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acronyms</strong></td>
<td></td>
</tr>
<tr>
<td>CoV</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>EAR</td>
<td>East African Rift</td>
</tr>
<tr>
<td>G–R</td>
<td>Gutenberg–Richter</td>
</tr>
<tr>
<td>GMM</td>
<td>Ground Motion Model</td>
</tr>
<tr>
<td>GSRM</td>
<td>Global Strain Rate Model</td>
</tr>
<tr>
<td>MAFD</td>
<td>Malawi Active Fault Database</td>
</tr>
<tr>
<td>MMI</td>
<td>Modified Mercalli Intensity Scale</td>
</tr>
<tr>
<td>MPS19</td>
<td>Italian Seismic Hazard Model (Modello di Pericolosità Sismica)</td>
</tr>
<tr>
<td>MSSM</td>
<td>Malawi Seismogenic Source Model</td>
</tr>
<tr>
<td>MFD</td>
<td>Magnitude–frequency distribution</td>
</tr>
<tr>
<td>PGA</td>
<td>Peak ground acceleration</td>
</tr>
<tr>
<td>PoE</td>
<td>Probability of exceedance</td>
</tr>
<tr>
<td>PSHA</td>
<td>Probabilistic seismic hazard analysis</td>
</tr>
<tr>
<td>SA</td>
<td>Spectral Acceleration</td>
</tr>
<tr>
<td>SHIFT</td>
<td>Seismic hazard inferred from tectonics</td>
</tr>
<tr>
<td>SSA-GEM</td>
<td>Sub-Sahara African Global Earthquake Model catalog</td>
</tr>
<tr>
<td>SSA-GSRM</td>
<td>Sub-Saharan African Geodetic Strain Rate Model</td>
</tr>
<tr>
<td>$V_{SS0}$</td>
<td>Average shear wave velocity to 30 m depth</td>
</tr>
<tr>
<td><strong>Symbols</strong></td>
<td></td>
</tr>
<tr>
<td>$\alpha_C$</td>
<td>Activity rate for a characteristic MFD</td>
</tr>
<tr>
<td>$\alpha_{NC}$</td>
<td>Activity rate for the non-characteristic magnitude range in a characteristic MFD</td>
</tr>
<tr>
<td>$\alpha_{GR}$</td>
<td>Activity rate for a G–R MFD</td>
</tr>
<tr>
<td>$\beta$</td>
<td>The product of the $b$-value and ln10</td>
</tr>
<tr>
<td>$\delta_s$</td>
<td>Source dip</td>
</tr>
<tr>
<td>$\Delta m_1$</td>
<td>Magnitude range in characteristic MFD where recurrence rate is less than characteristic portion</td>
</tr>
<tr>
<td>$\Delta m_2$</td>
<td>Magnitude range for characteristic events</td>
</tr>
<tr>
<td>$\lambda_s$</td>
<td>Annual occurrence rate for source events</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Crustal rigidity</td>
</tr>
<tr>
<td>$A_s$</td>
<td>Source area</td>
</tr>
<tr>
<td>$c_1$ &amp; $c_2$</td>
<td>Empirical constants from Leonard (2010)</td>
</tr>
<tr>
<td>$D_s$</td>
<td>Source single event displacement</td>
</tr>
<tr>
<td>$f_M(m)$</td>
<td>Source probability density function for magnitude $m$</td>
</tr>
<tr>
<td>$L_s$</td>
<td>Source length</td>
</tr>
<tr>
<td>$m$</td>
<td>Earthquake magnitude</td>
</tr>
<tr>
<td>$m_s$</td>
<td>Source earthquake magnitude estimate</td>
</tr>
<tr>
<td>$M_{Max}$</td>
<td>Maximum expected earthquake magnitude</td>
</tr>
<tr>
<td>$M_{Max}$</td>
<td>Seismic moment for $M_{Max}$</td>
</tr>
<tr>
<td>$M_{Min}$</td>
<td>Minimum earthquake magnitude considered for the PSHA</td>
</tr>
<tr>
<td>$M_0$</td>
<td>Seismic moment release rate</td>
</tr>
<tr>
<td>$\dot{M}_0$</td>
<td>Moment magnitude</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of event catalog–GMM combinations</td>
</tr>
<tr>
<td>$R_t$</td>
<td>Source recurrence interval</td>
</tr>
<tr>
<td>$S_s$</td>
<td>Source slip rate</td>
</tr>
<tr>
<td>$t_e$</td>
<td>Time to source’s next event in stochastic catalog</td>
</tr>
<tr>
<td>$v_{COM}$</td>
<td>Rate of ground motion exceedance</td>
</tr>
<tr>
<td>$W_s$</td>
<td>Source width</td>
</tr>
<tr>
<td>$z$</td>
<td>Seismogenic layer thickness</td>
</tr>
</tbody>
</table>

SSA-GEM catalog when estimating the areal source’s G–R parameters (with the exception of the 2009 Karonga earthquakes). An additive combination of the areal and MSSM sources could therefore lead to double-counting of seismicity across the magnitude range where these sources overlap (i.e. $M_W$ 4.5–7.0). Combining these sources in this way, does not however, overestimate the moment release rate in Malawi (as constrained by geodesy) nor does it imply an unusually high proportion of off-fault seismicity (Section 4.6.2, Appendix A4). We therefore retain this relatively simple approach for combining areal and MSSM sources and acknowledge that future PSHA in Malawi should more critically examine how these sources are incorporated (Section 7.2).

### 4.4 Stochastic event catalog generation

To perform the PSHA, we generated two million 1-yr long simulations for each of the five MSSM cases we wish to consider: the Direct-MSSM approach (Section 4.1) and four catalogs in the Adapted MSSM approach to cumulatively explore whether MSSM sources exhibit G–R or characteristic seismicity and if their...
down-dip extrapolation is length- or layer-limited (Section 4.2, Figs 2, 4 and 7). To represent our additive combination of the MSSM and areal sources (Section 4.3), each MSSM-based catalog is combined with an equivalent length areal source catalog.

The identical length of each MSSM-based catalog is chosen to reflect that we have no constraints on what the ‘true’ source width or MFD is in Malawi and so we equally weight each hypothesis (Frankel et al. 2000; Goda & Sharipov 2021). Another interpretation is that these catalogs can be merged into one ‘Combined Catalog’ that consists of 10 million 1-yr catalog simulations of MSSM and areal source seismicity (Fig. 7). Compared to conventional PSHA, in which ground motions are derived from integrating over all sources, distances and magnitudes, these different stochastic catalogs provide an intuitive and flexible approach to explore alternative on-fault MFD and down-dip extents for a large number of sources (Musson 1999; Atkinson & Goda 2013).
Earthquake occurrence is modelled in the catalogs using a memory-less Poisson process and so each event is independent of time for a given source (e.g. Zhuang et al. 2012; Pace et al. 2016). In the Direct-MSSM catalog, a source’s annual occurrence rate ($\lambda_s$) is taken from the inverse of its recurrence interval, so that the time to its next event ($t_s$) is

$$t_s = -\ln(1-u)/\lambda_s$$  \hspace{1cm} (3)

where $u$ is a sample from the standard uniform distribution. The event magnitude is then sampled from a random normal variable centred around the source’s Leonard (2010) magnitude–area scaling with a standard deviation of 0.1 and truncated $\pm 0.2$ magnitude units.

Since section, fault and multifault source types are mutually exclusive, weightings must be assigned to describe their relative likelihood in the Direct-MSSM catalog. We therefore generated catalogs for all possible source type weighting combinations at intervals of 0.1 with the limitation that the weighting of any source type is $\geq 0.1$ (Fig. 8a). We then combined each catalog with the areal source catalog and searched for the weighting combination that produced a catalog with the closest $b$-value to the regional estimate (1.02; Hodge et al. 2015; Poggi et al. 2017). From this test, we selected a weighting for section, fault and multifault sources of 0.1–0.3–0.6.
of events is sampled from eq. (A2) (Appendix A2). (Fig. 8). Further details of this weighting procedure are given in Appendix A1.

Within each simulation cycle of the MSSM-Adapted catalogs, we randomly sample one of the nine recurrence models that is generated for each source width-MFD case (Fig. A1, Table A2) given the weightings in Fig. 2. Then from the catalogs with a G–R MFD, \( \lambda_M \) equals \( \alpha_{GR} \) as defined in eq. (A4) and event magnitudes are taken from the probability density function in eq. (A1). For a characteristic MFD, \( \lambda_c \) is defined by \( \alpha_C \) (eq. A5) and the magnitude of events is sampled from eq. (A2) (Appendix A2).

For the areal source catalog, \( \lambda_s \) and the magnitude distribution for each source is defined by its G–R relationship (Table 2). Events occur randomly anywhere within each source; however, events >200 km from the region assessed in the PSHA are subsequently removed (Fig. 6).

For the purposes of calculating ground motions, all events in the MSSM-based catalogs are presumed to be normal faulting earthquakes. This is consistent with their moderate fault dip, the regional stress state (Delvaux & Barth 2010; Ebinger et al. 2019; Williams et al. 2019), Late Quaternary fault slickensides (Wedmore et al. 2020a) and the offset of sediments under Lake Malawi (Accardo et al. 2018; Shillington et al. 2020). However, of the 63 focal mechanisms that were resolved during a 2-yr deployment of seismometers in northern Malawi, seven were strike-slip (Ebinger et al. 2019) and it has been proposed that some historical events elsewhere in the EAR were strike-slip (Ayele & Kulhanek 2000). To recognize this, 10 per cent of events in the Areal Source Catalog are randomly assigned to be strike-slip when applying the GMM (Section 5.1).

### 4.5 Rupture geometry

We define the rupture geometry of events in the MSSM-based catalogs using the MSSM geometric model (Fig. 3d; Williams et al. 2022b). This model consists of 2D planes in 3D space; however, for the purpose of source-to-site calculations, we convert it to a set of grid points at intervals of \( 1 \times 1 \times 0.6 \) km in the \( x \times y \times z \) direction respectively. In the Direct-MSSM catalog, the source geometry defines the lateral extent of their events. In addition, we allow the depth interval of smaller section or fault sources to randomly float within a larger fault or multifault plane.

In the length-limited Adapted-MSSM catalogs, we use the same geometric source model as the Direct-MSSM catalog. However, in the layer-limited catalogs, we revise this model so that all sources are extrapolated to a depth of 35 km unless intersected by another source. The length and width of each event in the Adapted-MSSM catalogs are calculated from the Leonard (2010) scaling relationships between source length, width and magnitude and then floating these dimensions randomly within the larger source plane (Fig. 5). In this approach, section boundaries or fault tips (for multifault sources) are not considered rupture barriers unlike in the Direct-MSSM Catalog.

In cases where MSSM sources intersect and it is not possible to fit a rupture onto the cut-off plane given the Leonard (2010) length-width scaling (eq. 1), an area that matches the event’s magnitude is randomly fitted onto the plane instead. To save computational resources, events of \( M_w<5.4 \) in the Adapted-MSSM catalog are treated as point sources and are randomly located on the source plane. Areal source catalog events are also treated as point sources and their depth is randomly sampled from a normal distribution truncated between 5–35 km and with a mean and standard deviation of 20 and 5 km, respectively.

### 4.6 Stochastic event catalog validation

We test the output of the stochastic event catalogs in two ways: (1) internal tests to investigate if the sources’ moment rate (\( M_0 \)) and MFD shape match those from their input recurrence models and (2) external tests to determine if the total \( M_0 \) of these catalogs is consistent with independent constraints for the \( M_0 \) in Malawi from instrumental seismicity (Poggi et al. 2017) and geodesy (Kreemer et al. 2014; Stamps et al. 2018). Further details of these tests are provided in Appendix A.

#### 4.6.1 Internal tests

For the Direct- and Adapted-MSSM internal tests, there is generally a good correlation between the \( M_0 \) of MSSM source’s calculated using their slip rate and area (eq. (2) and in the stochastic event catalogs (Fig. 9). We consider the Direct MSSM’s catalog \( M_0 (1.52 \times 10^{18} \text{Nm yr}^{-1}) \) an acceptable fit to the target \( \dot{M}_0 \), as defined by weighing each source by rupture type and then summing up their \( \dot{M}_0 \) (1.48 \times 10^{18} \text{Nm yr}^{-1}; Table A3).

For the MSSM adapted sources, the \( \dot{M}_0 \) of the entire catalog is 2–10 per cent lower than calculated from just combining the
conclude that the MSSM and areal sources are sufficiently well represented in the stochastic event catalogs that we can used them to formulate our PSHA (Section 5.2) and hence they can replace the need to evaluate these sources using numerical integration instead (i.e. ‘conventional’ PSHA).

4.6.2 External tests

For the external tests, we first calculate the $M_0$ of the areal sources developed in Malawi by Poggi et al. (2017) using their original $M_{Max}$ estimate (7.9). This analysis is therefore representative of the $M_0$ in Malawi from an extrapolated instrumental record (Appendix A3). From this approach we derive a $M_0$ of $\sim 7.45 \times 10^{17}$ Nm yr$^{-1}$ (Table A4), which is just ~40 per cent of the $M_0$ from the Combined Catalog ($\sim 1.89 \times 10^{18}$ Nm yr$^{-1}$; Fig. 7).

Part of this discrepancy reflects differences in the $M_{Max}$ estimates for Malawi between Poggi et al. (2017) and the MSSM-based Combined Catalog ($M_{Max}$ 7.9 vs 8.1). If the MSSM-based $M_{Max}$ estimate is applied to the areal sources, the total $M_0$ increases to $9.3 \times 10^{17}$ Nm yr$^{-1}$. Increasing $M_{Max}$ therefore reduces the $M_0$ discrepancy, but cannot account for it alone. Instead, we consider three other possibilities: (1) the geodetic $M_0$ from which the slip-rates in the MSSM are mainly derived from (Wedmore et al. 2021) is released aseismically, (2) the geodetic $M_0$ is an overestimate, or (3) extrapolating the instrumental record underestimates long-term earthquake activity in Malawi.

With regards to the first scenario, the nucleation of earthquakes throughout Malawi’s seismogenic layer (Ebinger et al. 2019; Stevens et al. 2021) and energetic slowly decaying aftershock sequences imply overall highly coupled faults (Ben-Zion 2008; Gahtery et al. 2019). Some degree of aseismic deformation in Malawi may still occur (Ebinger et al. 2019), with shallow (depths <5–10 km) aseismic afterslip observed following a $M_W$ 5.2 earthquake near Karonga in 2014 (Zheng et al. 2020) and the 2006 $M_W$ 7.0 Machaze earthquake in nearby Mozambique (Copley et al. 2012; Lloyd et al. 2019).

For the second scenario, we compare the Combined Catalog to independent estimates of the geodetic $M_0$ in Malawi from: (1) the Global Strain Rate Model (GSRM v2.1; Kreemer et al. 2014) and (2) the Sub-Saharan African Geodetic Strain Rate Model (SSA-GSRM v1.0; Stamps et al. 2018). Following the approach used in the seismic hazard inferred from tectonics (SHIFT) model (Bird & Liu 2007; Bird & Kreemer 2015), the $M_0$ of the GSRM v2.1 and SSA-GSRM v1.0 models are $9.5 \times 10^{17}$ Nm yr$^{-1}$ and...
Figure 8. (a) MFD for the Direct MSSM-Areal combined Catalog in which 36 possible weighting combinations of section, fault and multifault ruptures are explored, with the MFD curve of the selected weighting combination highlighted. (b) MFD for the Direct MSSM-Area catalog for optimal set of rupture weightings, defined by a catalog with a \( b \)-value of 1.02. The MFD for individual source types is also shown.

Figure 9. Comparison of the analytical moment rate \( \dot{M}_0 \) of sources in the MSSM and their \( \dot{M}_0 \) in (a) the Direct MSSM catalog and (b) two of the Adapted MSSM simulated catalogs. In (a) the Direct MSSM source analytical \( \dot{M}_0 \) is weighted by source type as indicated in Section 4.4.

3.5 \times 10^{18} \text{ Nm yr}^{-1} (Appendix A5). The Combined Catalog \( \dot{M}_0 \) (1.89 \times 10^{18} \text{ Nm yr}^{-1}) is intermediate between these estimates and so suggests the Wedmore et al. (2021) geodetic model does not imply anomalously high extension rates in Malawi. Furthermore, the MSSM slip rates that are derived from the 75 ka seismic reflector offsets in Lake Malawi are consistent with geodetically derived regional extensional rates (Shillington et al. 2020; Williams et al. 2022b).

The Combined Catalog \( \dot{M}_0 \) is also not necessarily identical to the input geodetic \( \dot{M}_0 \) since: (1) it incorporates events that accommodate hanging-wall flexural extension along intrarift faults in northern and central Malawi and this deformation will not be captured by large-scale geodetic models (Muirhead et al. 2016; Shillington et al. 2020; Williams et al. 2022b), (2) areal source events are independent of the input geodetic \( \dot{M}_0 \) and (3) not all of the geodetic \( \dot{M}_0 \) is converted to seismic \( \dot{M}_0 \) due to the obliquity of MSSM sources to the regional extension direction (Williams et al. 2021a) and/or because they do not extend across the full width of the seismogenic layer (Section 4.2 and Fig. 5). We calculate that if all MSSM sources were optimally oriented to the regional extension direction and extended to the base of the seismogenic layer, the total \( \dot{M}_0 \) (1.88 \times 10^{18} \text{ Nm yr}^{-1}, Appendix A4) is nearly identical to the Combined Catalog (1.89 \times 10^{18} \text{ Nm yr}^{-1}). Therefore, a physical interpretation of the areal source source events is that they accommodate the deformation required to prevent space problems that would otherwise arise from normal fault obliquity and narrow fault widths in Malawi.

To investigate the third scenario that the SSA-GEM catalog underestimates long term earthquake activity in Malawi, we divided the Combined Catalog into 50-yr increments and compared these samples’ \( \dot{M}_0 \) and MFD shape to the non-declustered 50-yr long (1965–2015) SSA-GEM catalog (Poggi et al. 2017). In this way,
we can investigate how likely it is that the observed seismicity in Malawi would have been simulated in our event catalogs. We find that 12 per cent of the 50 yr Combined Catalog have a $M_0$ equal to or less than the SSA-GEM catalog $M_0$ (Fig. 10a). The MFD of this 50 yr period of instrumental seismicity is also within the variability of seismicity within the Combined Catalog 50 yr samples (Fig. 10b). Hence, if the Combined Catalog is representative of seismicity in Malawi, then the observed $M_0$ between 1965 and 2015 is relatively, but not inconceivably, low.

We note too that the G–R stochastic catalog 50-yr samples cannot replicate the SSA-GEM $M_0$ (Fig. 10). This result could be used to argue against applying G–R recurrence models to MSSM sources (Wesnousky et al. 1983; Ishibe & Shimazaki 2012). However, the event catalogs are generated using a time-independent Poisson approach (eq. 3) and it is plausible that clustered seismicity such as triggered events and/or long aftershock sequences, as occurred during the 2009 Karonga earthquakes, allow a MSSM source’s MFD to align with a G–R relationship (Page & Felzer 2015; Stirling & Gerstenberger 2018; Wang et al. 2021). We suggest that if a non-Poisson approach was used to generate the MSSM-catalogs, then more G–R catalog 50 yr samples would have a $M_0$ that is comparable to the SSA-GEM catalog, even though the long-term deformation rates would not change.

In summary, there are many challenges in reconciling Malawi’s observed seismic $M_0$ and the MSSM-Areal Combined $M_0$. We propose that this reflects an incomplete instrumental earthquake record, which in turn is indicative of the limited duration, poor instrumental coverage, clustered seismicity, low regional extension rates and locked faults (Ambraseys & Adams 1991; Biggs et al. 2010; Hodge et al. 2015; Stevens et al. 2021). This is further highlighted by the large uncertainty in how the catalog may be extrapolated to larger magnitudes (Fig. 10b; Tinti & Mulargia 1987). Our analysis does not consider uncertainty within the MSSM itself, or with applying the Leonard (2010) scaling relationships to faults in Malawi (Section 7.2). Nevertheless, the Combined Catalog satisfies constraints on the distribution of across-rift regional extensional strain (Shillington et al. 2020; Wedmore et al. 2020a), the regional b-value (Poggi et al. 2017) and the source specific $M_0$ (Fig. 9). Furthermore, we have explored alternative hypotheses for uncertainty in on-fault MFDs and source down dip-extents. We therefore propose that it represents the best source model for PSHA currently available for Malawi.

5 SEISMIC HAZARD CALCULATIONS

5.1 Ground motion model

In the absence of strong ground motion data in East Africa (Midzi et al. 1999; Hodge et al. 2015; Poggi et al. 2017), we apply GMM from other similar tectonic terranes. In selecting the GMMs, we consider variations between 1D seismic velocity models in northern and southern Malawi in the crust’s top 5 km (Fig. A8;Ebinger et al. 2019; Stevens et al. 2021) and that large magnitude earthquakes in Malawi have generated remarkably little fracturing in the surrounding crust (Wedmore et al. 2020b; Carpenter et al. 2022; Williams et al. 2022d). This led us to apply and equally weight three well-tested active crust GMMs (Akkar et al. 2014; Boore et al. 2014; Chiou & Youngs 2014) and one stable crust GMM (eastern crustal GMM from Atkinson & Adams 2013). Some of these GMMs were previously applied in East Africa by Poggi et al. (2017) and the ratio of active to stable crust GMMs is same as this study.

Cumulatively, these GMMs allow us to explore various source to site measurements: closest horizontal distance to rupture’s surface projection (Joyner-Boore distance, $R_{JB}$), closest distance to rupture plane ($R_{epi}$), epicentral distance ($R_{ep}$) and hypocentral distance ($R_{hypo}$). In the instances that $R_{ep}$ and $R_{hypo}$ are applied to the MSSM fault-based events, distances are measured to a point randomly sampled within the simulated rupture’s geometry. $R_{hypo} < 10$ km are not considered by the Atkinson & Adams (2013) GMM, so in these instances, ground motions are calculated with $R_{hypo}$ fixed to 10 km.

To incorporate aleatory uncertainty, the respective sigma model for each GMM is applied to the calculated median ground motion. For the site-specific PSHA in the three largest cities in Malawi (Lilongwe, Blantryre and Mzuzu, Fig. 1), $V_{S30}$ is set to reference values of 300 and 760 m s$^{-1}$ and a range of spectral accelerations (SA) between 0-3 s are considered. For the PSHA maps, which are developed from 756 sites across Malawi in a $0.2 \times 0.2$ latitude and longitude grid, we consider PGA only and both a reference $V_{S30}$ value (760 m s$^{-1}$) and the site-specific value derived from the USGS $V_{S30}$ database (Fig. A6; Wald & Allen 2007). Ground motion calculations do not incorporate differences in lake floor elevation and substrata for the 59 sites in our grid under Lake Malawi. Because of this uncertainty, these sites are not explicitly shown in the hazard maps. However, to facilitate our understanding for how the MSSM sources influence hazard, they are retained in the hazard map comparisons.

5.2 Sensitivity analysis

We calculate the seismic hazard and its uncertainty by following the ensemble modelling framework used in the latest Italian seismic hazard model [Modello di Pericolosita Sismica (MPS19); Marzocchi et al. 2015; Meletti et al. 2021]. In this approach, N seismic hazard curves are generated for each site, where N is the number of source model-GMM combinations. For a given PoE and SA, N hazard values can therefore be sampled and fitted to a continuous distribution (i.e. ‘horizontal dissections’ of the curves), where the central value represents the seismic hazard estimate and the dispersion mimics the epistemic uncertainty (Marzocchi et al. 2015). In this study, $N = 20$ given that we consider four GMMs and five interpretations of the MSSM in the stochastic event catalogs (Figs 2 and 7). For each catalog–GMM combination, the annual probability (or rate) at which a specific ground motion intensity is exceeded ($v_{GM \geq gm}$) is calculated given the catalog’s two million yr length (Section 4.4).

For the site-specific PSHA, the 20 ground motion intensity values at a given PoE and SA are described by a beta distribution, as this provides good fits to unimodal distributions bounded between 0 and 1 (Marzocchi et al. 2015). The spatial distribution of seismic hazard uncertainty is of greater interest for the PSHA maps and so is described by: (1) the interquartile range of the 20 seismic hazard values calculated at each site and (2) their Coefficient of Variation (CoV). The former describes the spatial distribution of the absolute uncertainty, whilst the latter is indicative of the uncertainty once normalized by the hazard level (Meletti et al. 2021). This analysis provides only a minimum bound on hazard uncertainty as we do not consider the uncertainty in the MSSM slip rate and recurrence interval estimates, the areal sources, or the nine recurrence models explored in the Adapted MSSM catalogs (Figs 2 and A1). Stochastic event catalog generation and seismic hazard calculations were performed using bespoke codes written in MATLAB and...
Figure 10. (a) Comparison of the moment rate ($M_0$) in 50 yr samples of the stochastic event catalogs and the observed $M_0$ from the SSA-GEM catalog between 1965-2015 for the assessed region shown in Fig. 6 (Poggi et al. 2017). Results for the Combined Catalog samples are shown as a kernel distribution ($n = 200000$) and are also resolved into the five different interpretations of the MSSM ($n = 40000$). The percent by each catalog indicates the proportion of 50 yr samples that had a $M_0 \leq$ SSA-GEM catalog. For context, the $M_0$ over the assessed region from the the Global Strain Model (GSRM v.2.1; Kreemer et al. 2014) and Sub-Saharan African Geodetic Strain Model (SSA-GSRM v.1.0; Stamps et al. 2018) are also shown (see also Appendix A5). (b) Comparison of the Combined Catalog and SSA-GEM catalog MFD, with the MFD for 10 000 random 50 yr long samples of the Combined Catalog, colored by MSSM catalog also plotted. We also fit and extrapolate a G–R relationship and its associated uncertainties, to the 75 events in the SSA-GEM catalog in the assessed region with $M_W > 4.5$ following Tinti & Mulargia (1987). The SSA-GEM catalog has not been declustered in this analysis.


6 PSHA RESULTS

6.1 Site specific PSHA

The seismic hazard of the three selected sites (Lilongwe, Blantyre and Mzuzu; Fig. 1) shows considerable diversity. The mean hazard is lowest in Lilongwe (10 per cent PoE 0.11 g in 50 yr for PGA and $V_{ss3}$ of 760 m s$^{-1}$). Lilongwe is ~55 km from the nearest MSSM source and so local (<40 km) $M_W$ 5–6 events in the areal source model present the main source of hazard (Figs 11–13). The MSSM sources do, however, become important contributors to hazard in Lilongwe at low PoE and longer (>1 s) vibration periods (Fig. 12d). Hazard levels are higher in Blantyre and Mzuzu (10 per cent PoE 0.15–0.2 g in 50 yr). This reflects that both sites are < 20 km from MSSM sources, which dominate their hazard (Figs 11–13).

Since the hazard at high PoE and short vibration periods is dominated by areal source events in Lilongwe, seismic hazard uncertainty is mainly driven by the GMM selection (Fig. 11a). Conversely, in Blantyre and Mzuzu, both the source model (i.e. the event catalogs) and GMM selection contribute to uncertainty. In particular, high hazard levels are derived in Blantyre for the combinations that consider the Atkinson & Adams (2013) GMM and G–R on-fault MFD (Fig. 11e). However, in Mzuzu, the highest hazard levels are found for the Atkinson & Adams (2013) GMM regardless of the on-fault MFD (Fig. 11f). In all cases, the uncertainty in how the MSSM sources propagate through Malawi’s 35-km thick seismogenic layer (i.e. length- or layer- limited width) do not significantly influence hazard estimates (Fig. 11).

6.2 Malawi seismic hazard maps

We first assess the relative contribution of areal and MSSM sources to seismic hazard in Malawi through maps that consider these sources separately. Figs 14(a) and (d) indicates that the hazard from areal sources is generally spatially uniform in Malawi, with a broad zone of relatively high hazard following the the relatively high $M_0$ Rukwa–Malawi source zone (Tables 2 and A4, Figs 6 and A3; Poggi et al. 2017), which broadly corresponds to the EAR’s trajectory through Malawi (Fig. 1).

By contrast, the fault-based MSSM sources imply a more complex spatial pattern of seismic hazard, with localized regions of relatively high hazard (10 per cent PoE ~0.2–0.3 g in 50 yr) adjacent to rift-bounding ‘border’ faults (Figs 14b and d). This reflects that border faults are thought to accommodate 50–90 per cent of extension in low strain magma-poor continental rifts such as in Malawi (Agostini et al. 2011; Accardo et al. 2018; Muirhead et al. 2019; Shillington et al. 2020; Wright et al. 2020; Wedmore et al. 2020a) and so these faults are assigned relatively high slip rates (0.5–2 mm yr$^{-1}$) in the MSSM. These maps indicate hazard levels are highest in regions surrounding the northern end of Lake Malawi, where rift extension rates are relatively high (Wedmore et al. 2021) and intrarift faults also accommodate local hanging-wall flexural extension (Shillington et al. 2020; Williams et al. 2022b).

As observed in the site-specific PSHA, G–R recurrence models imply higher hazard levels than the Direct MSSM or characteristic approach at high PoE (Fig. 15d and e). The effect on hazard levels for different source down-dip extents are smaller (<0.1 g for 10 per cent PoE in 50 yr hazard levels) and localized to regions with relatively short (<50 km) MSSM sources such as around Karonga and Malawi’s southern tip (Fig. 15f). The uncertainty in the MSSM sources MFD mean that the seismic hazard of regions close to faults have a higher CoV and interquartile range than regions where
the hazard is dominated by just the single areal source model we consider. By comparison, multiple areal and smoothed seismicity sources are considered in the Italian seismic hazard model and this means that regions far from faults have the highest CoV in hazard levels (Meletti et al. 2021). Regions peripheral to Malawi also have relatively high CoV (Fig. 15c), but this likely reflects their very low seismicity rate (Fig. 6) and not the underlying uncertainty in our analysis.

To quantify how the MSSM-Areal Combined map compares to seismic hazard maps previously developed for Malawi by Poggi et al. (2017) and Hodge et al. (2015), we find the closest sites within each map’s respective grids and then subtract these hazard estimates from the MSSM-Areal Combined estimate using the reference $V_{S30}$ value (760 m s$^{-1}$). This is the same $V_{S30}$ condition used by Hodge et al. (2015), whereas Poggi et al. (2017) considered a slightly lower $V_{S30}$ estimate (600 m s$^{-1}$). Poggi et al. (2017) also truncated the GMM sampling at ±3 standard deviations ($\epsilon$) whilst our sampling was unbounded (Section 5.1). Nevertheless, this different GMM sampling approach is unlikely to influence the hazard map comparisons, given that analysis of the disaggregation plots (Fig. 13) indicates that only ~1–2 per cent of the probability mass were simulated when $\epsilon > 3$ (Fig. 13). We compare our hazard map to Hodge et al. (2015) for a 500 yr return period (equivalent to ~9 per cent PoE in 50 yr), which was the minimum return period considered in their study.

For sites <40 km from the MSSM sources, our seismic hazard estimates for 10 per cent PoE in 50 yr level are up to 0.3 g higher than in the map from Poggi et al. (2017), with increases highest around the relatively high slip-rate border faults. The median difference between the Poggi et al. (2017) and MSSM-Areal combined maps is, however, only 0.01 g (Fig. A12a). This reflects that most sites in Malawi are sufficiently far from active faults (> 40 km), for areal sources to be the main contributor to hazard at high PoE. Indeed, at these sites > 40 km from active faults, near-identical hazard levels are expected because we incorporate off-fault seismicity using the areal sources developed by Poggi et al. (2017).

The MSSM-Areal Combined map indicates higher hazard levels than in Hodge et al. (2015), particularly at sites where new fault sources have been included (increases of 0.2–0.3 g; Fig. 16d–f). Locally, the MSSM-Areal combined map indicates lower hazard levels around the Bandawe and Mbamba faults (Fig. 16d). These faults were included as sources by Hodge et al. (2015) but not in the MSSM, as new seismic reflection data indicates that these are inactive faults (McCartney & Scholz 2016; Accardo et al. 2018; Scholz et al. 2020). Differences in GMM selection and fault and areal source modelling may have also affected comparisons between these two maps and we discuss this further in Section 7.1.
7 DISCUSSION

7.1 Implications for seismic hazard and risk in Malawi

The higher moment rate ($M_0$) implied by the MSSM-based catalogs leads to elevated seismic hazard estimates in Malawi compared to previous instrumental-seismicity based PSHA (Fig. 16; Poggi et al. 2017). However, this increase is mainly observed for sites within 40 km of a MSSM source (Fig. 16a–c). This result demonstrates the importance of fault-based sources for understanding both the magnitude and spatial distribution of seismic hazard in Malawi.

Compared to the first generation of fault-based PSHA in Malawi (Hodge et al. 2015), the MSSM-Areal combined map indicates higher seismic hazard levels in the EAR valley (Fig. 16d–f). This reflects the incorporation of new fault sources in the MSSM (107 vs 7) and in particular intrarift faults, which have been highlighted as overlooked sources of seismic hazard in Malawi (Biggs et al. 2010; Shillington et al. 2020; Wedmore et al. 2020a). However, the Livingstone Fault is broadly coincident between the two maps and even though its slip rate estimate is lower in the MSSM (2.0 vs...
3.8 mm yr$^{-1}$), the MSSM-Areal combined map implies higher seismic hazard adjacent to it (Fig. 16d–f). This reflects a combination of: (1) our exploration of G–R on-fault MFD, which leads to higher hazard levels than the characteristic-only on-fault MFD considered by Hodge et al. (2015), (2) addition of nearby intrarift fault sources, (3) inclusion of a stable crust GMM and our increased estimates for (4) the base of the seismogenic layer (35 vs 30 km) and (5) the maximum background event magnitude (MW 7 vs MW 6.25–6.75).

For global context, with a 10 per cent PoE $\sim 0.2–0.4$ g (PGA) in 50 yr close to MSSM sources and 10 per cent PoE $\sim 0.10–15$ g PGA 50 yr in regions peripheral to these sources (Fig. 14), the seismic hazard in Malawi can be considered comparable to other regions with low slip rate normal faults, such as Italy (Meletti et al. 2021) and the Basin and Range Province in the USA (Petersen et al. 2015). Regional extension rates in Malawi are slightly lower than in these regions (0.5–1.5 mm yr$^{-1}$ vs $\sim$3 mm yr$^{-1}$; D’Agostino et al. 2011; Hammond et al. 2014). However, this may be compensated for by Malawi’s relatively thick seismogenic layer, which allows normal faults to reach larger lengths and widths (100–150-km long, 40-km wide) than in Italy (<60-km long, 15–20-km wide; Basili et al. 2008; Valentini et al. 2017) and although the Basin and Range Province’s Wasatch Fault is 370-km long, its seismogenic width is

---

**Figure 14.** PGA seismic hazard maps for Malawi for (a)–(c) 10 per cent PoE in 50 yr and (d)–(f) 2 per cent PoE in 50 yr for reference $V_{S30}$ value (760 m s$^{-1}$). Figure is arranged so each column represents a different catalog. Red lines depict the MSSM sources (Williams et al. 2022b). For equivalent maps for the slope-based USGS $V_{S30}$ values (Wald & Allen 2007), see Fig. A11.
assumed to be 20 km (Valentini et al. 2020) and paleoseismic data suggest most ruptures are 20–100-km long (DuRoss et al. 2016). In turn, Malawi’s wider faults will have a disproportionately high $M_0$ and can host larger magnitude earthquakes (Jackson & Blenkinsop 1997; Hodge et al. 2020).

To fully explore and quantify the implications of this study for seismic risk in Malawi, results should be combined with seismic vulnerability and exposure assessments (Goda et al. 2016, 2022; Ngoma et al. 2019; Kloukinas et al. 2020; Giordano et al. 2021; Novelli et al. 2021). Nevertheless, some implications of this PSHA to seismic risk are apparent. For example, given the low quality and high turnover of building stock in Malawi (Giordano et al. 2021), the results that are of most practical importance are for high PoE and in these instances the MSSM sources affect hazard levels only at long vibration periods and/or sites close (<40 km) to active faults (Fig. 12). Hence, off-fault areal sources are still important contributors to seismic risk in Malawi and future work should consider improving its seismic network so that future PSHA can use a more finely resolved off-fault source model.

Instances where the hazard estimates at low PoE are of importance in Malawi include hydro-electric dams in the Shire River valley in southern Malawi, which generates 80 per cent of Malawi’s electricity (Taulo et al. 2015). The development of geothermal resources in Malawi, whose locations are inherently controlled by active faults (Dulanya et al. 2010; Gondwe 2015; Dávalos-Elizondo et al. 2021), should also consider local seismic hazard. PSHA only considers ground motions and other secondary seismic hazards in Malawi, such as liquefaction, fault displacement, landslides and seiches (Williams et al. 2022c), are not considered here.

Figure 15. Sensitivity analysis for seismic hazard maps. (a) Combined 10 per cent PoE in 50 yr seismic hazard map as shown in Fig. 14(c), but with adjusted color axis. (b) Interquartile range and (c) Coefficient of Variation (CoV) of the 20 seismic hazard values calculated from each event catalog–GMM combination. (d)–(f) Maps showing how source modelling affects hazard uncertainty. (d) Difference in seismic hazard values for a G–R and MSSM Direct on-fault MFD, for the length-limited case. (e) Same as (d) but comparison is between a G–R and characteristic MFD. (f) Difference in seismic hazard values for layer- and length-limited source down dip extents for a G–R on-fault MFD. All maps are for 10 per cent PoE in 50 yr hazard level, PGA and the generic $V_{S30}$ value (760 m s$^{-1}$). Comparison in (d)–(f) are for maps generated with the Boore et al. (2014) GMM. The area shown under Lake Malawi (blue outline) is included for the purpose of comparing hazard maps only.
7.2 Using fault-based sources for PSHA in Malawi and other regions with low strain rates and a thick seismogenic layer

Fault-based MSSM sources are incorporated into our PSHA of Malawi using stochastic event catalogs. Cumulatively, these catalogs explore five realisations of the MSSM (Fig. 2) with different on-fault MFD and down-dip extension of fault sources through Malawi’s seismogenic layer (i.e. ‘length’ or ‘layer’ limited faults).

We consider alternative down-dip extents as it is not clear how sources should be extrapolated through Malawi’s ~35-km thick seismogenic layer (see Section 4.2). The influence of this uncertainty is only significant in regions in Malawi with relatively short (<50 km) sources (Fig. 15f). Longer sources, which also tend to have the highest slip rates, are not sensitive to this uncertainty as they are expected to extend throughout the seismogenic layer (Fig. A7).
Three on-fault MFD are considered for MSSM sources during the PSHA: G–R, characteristic, or the ‘Direct-MSSM’ approach where sources may rupture in geometrically defined sections, faults, or multifault ruptures (Fig. 4). We find that for a 10 per cent PoE in 50 yr level, a G–R MFD implies higher hazard levels (Figs 11 and 15). This result can be understood in terms of the relatively frequent moderate magnitude seismicity (\(M_W > 6.8\); Fig. 7) that is inherent to G–R MFD and which in a low strain rate region is considerably more likely within a 50–500 yr time-frame than large magnitude characteristic events (Valentini et al. 2017; Goda & Sharipov 2021). Although the differences in hazard levels between G–R, characteristic and Direct MSSM MFD are reduced at lower PoE, they can still be significant for sites that are close to many long (<40 km) low slip rate (0.05–1 mm yr\(^{-1}\)) faults (e.g. Blantyre, Fig. 11).

In future, the on-fault MFD could be constrained by inversion-based source models (e.g. Field et al. 2021); however, there are currently considerable challenges to defining a regional MFD target for Malawi when developing these models (Section 4.6.2). In the meantime, we consider it prudent that multiple on-fault MFD cases should be explored during PSHA of regions like Malawi, with low strain rates and instrumental and historical records much shorter than the earthquake recurrence intervals of individual faults. Although a smaller source of uncertainty, our results also imply that alternative cases could be considered for source down-dip extrapolation in regions with an abnormally thick seismogenic layer (>20 km) and short (<50 km) faults.

The stochastic event catalogs used in this PSHA are simulated on the assumption that earthquakes in Malawi: (1) follow the Leonard (2010) scaling relations, (2) ruptures do not propagate below the presumed base of the seismogenic layer at 35 km and (3) earthquake inter-event times can be described by a time-independent Poisson process (Section 4.4). With so few well-instrumented \(M_W > 7\) continental normal fault earthquakes it is difficult to critically examine the first two assumptions (see also; Williams et al. 2022b). For the latter point, the 2009 Karonga earthquake sequence imply that fault interaction through static stress changes can lead to clustered non-Poisson seismicity in Malawi (Biggs et al. 2010; Fagereng 2013; Gaherty et al. 2019); indeed this is a widespread observation in low strain rate regions (e.g. Beanland & Berryman 1989; Wedmore et al. 2017; Griffin et al. 2020; Stevens & Avouac 2021; Itrurrieta et al. 2022). In these cases, earthquake interevent times are more appropriately modelled using two-parameter time-dependent distributions such as Weibull or Brownian Passage Time (Matthews et al. 2002; Zöller & Hainzl 2007; Cowie 2012). Seismic hazard assessment in Malawi should also recognize that previous large magnitude (\(M_W > 7\)) earthquakes in the EAR were followed by long and damaging aftershock sequences (Ambraseys 1991; Ambroseys & Adams 1991; Gaulon et al. 1992; Lloyd et al. 2019). Future seismic hazard assessment in Malawi should consider these challenges when the appropriate paleoseismic and seismic records become available.

7.3 Application of GMM in Malawi

An ongoing challenge with PSHA in Malawi and elsewhere in the EAR, is the lack of geotechnical (i.e. \(V_{50}\) measurements) and strong motion data (Midzi et al. 1999; Hodge et al. 2015; Poggi et al. 2017). This raises uncertainties when applying slope-based proxies for \(V_{50}\) (Figs A6 and A11) and means our use of global GMMs implicitly assumes that the ground motion behaviour in Malawi will be similar to other regions (i.e. the ‘ergodic assumption’; Anderson & Brune 1999). This could be addressed in the future by considering whether regional weak ground motion data in Malawi can be used to adjust GMMs (Yenier & Atkinson 2015). An additional problem is that EAR seismicity is characterised by deep moderate-large magnitude (\(M_W > 6\)) normal fault earthquakes and it is difficult to calibrate GMMs for these events as so few of them have been recorded (Akkar et al. 2014; Boore et al. 2014).

The incorporation of a stable crust GMM into PSHA in Malawi raises further challenges. In particular, the near-field (<10 km) motions associated with events in stable crust are poorly understood. This could be addressed though incorporation of the Next Generation Attenuation East (NGA-East) GMM (Goulet et al. 2018) in East Africa and this would also allow \(R_{100}\) to be considered in a stable crust GMMS. However, the NGA-East was developed for a reference \(V_{50}\) condition of 3000 m s\(^{-1}\) and challenges remain in adapting the site amplification factors for lower \(V_{50}\) values (Kolaj et al. 2019) that are likely in Malawi (200–800 m s\(^{-1}\), Fig. A6).

7.4 Implications for seismic hazard elsewhere in the EAR

Our comparisons of PSHA maps for Malawi (Fig. 16a–c) highlight the limitations of using short historical and instrumental catalogs to assess seismic hazard in regions with low strain rates (Section 4.6.2). These challenges apply elsewhere in the EAR, where despite abundant evidence for late Quaternary faulting (e.g. Vittori et al. 1997; Lerdal & Talbot 2002; Wanke 2005; Kervyn et al. 2006; Zielke & Strecker 2009; Fontijn et al. 2010; Nicholas et al. 2016; Delvaux et al. 2012, 2017; Muirhead et al. 2016; Daly et al. 2020; Wedmore et al. 2022), no fault-based PSHA has been attempted outside Malawi. This partly reflects the lack of chronoseismographic data needed to estimate fault slip rates in the EAR; however, this can be addressed to an extent by incorporating regional geodetic data using the MSSM systems-based approach (Williams et al. 2021a). Given the increasing levels of seismic risk (Goda et al. 2016; Meghraoui et al. 2016; Poggi et al. 2017; World-Bank 2019), we suggest there is a clear need to develop new fault-based PSHA maps elsewhere in the EAR.

8 CONCLUSIONS

We use the MSSM to develop a new fault-based PSHA in Malawi. We find that for sites close (<40 km) to these fault-based sources, seismic levels are higher than indicated by previous instrumental-based PSHA (Poggi et al. 2017). This replicates the findings of the first generation of fault-based PSHA in Malawi (Hodge et al. 2015). However, the incorporation of more fault sources in the MSSM (107 vs 7), source modelling and GMM selection, leads to a more complex seismic hazard pattern than Hodge et al. (2015). These results should motivate the development of more fault-based PSHA elsewhere in the EAR, as previous seismic hazard assessment has used the instrumental record of seismicity alone and in some situations, this may underestimate hazard levels.

The stochastic event catalogs we use to incorporate the MSSM into PSHA explore alternative hypotheses for on-fault MFD and the down-dip extension of fault-based sources through Malawi’s 35-km thick seismogenic layer. We find that seismic hazard levels are only sensitive to the down-dip extension of relatively short sources (differences of ~0.1 g for 10 per cent PoE in 50 yr for regions with <50-km long sources), whilst the assumed MFD can influence the hazard estimates for all sources. In particular, we find that compared
to a characteristic MFD or ‘direct’ implementation of the MSSM, 
a G–R MFD increases seismic hazard levels by up to 0.2 g for 
sites close to low slip rate sources (<1 mm yr⁻¹) and high PoE 
(10 per cent PoE in 50 yr).

Our new PSHA is also useful for highlighting sources of un-
certainty that present key targets for future research in Malawi. In 
particular, (1) the uncertainty in fault slip rates, (2) the applicability 
of a Poisson model for earthquake recurrence, (3) refining off-fault 
and (4) the lack of local fault scaling and strong ground motion data to assess the applicability of global empirical scaling 
relations in Malawi. Nevertheless, we suggest our incorporation of 
a rich active fault and geodetic data set makes this the most 
robust assessment of seismic hazard currently available for Malawi and 
presents a framework for assessing seismic hazard in other low 
strain rate regions.

AUTHORS’ CONTRIBUTIONS

Conceptualization: All authors. Methodology: All authors. Software: JW, KG and RD. Investigation: All authors. Writing – original draft preparation: JW. Writing – reviewing and editing: All authors. Data Curation: JW and LW. Funding acquisition: JB, AF, KG and MW.

ACKNOWLEDGMENTS

This study is funded by the EPSRC-Global Challenges Research Fund PREPARE (EP/P028233/1) and SAFER-PREPARED (part of the ‘Innovative data services for aquaculture, seismic resilience and drought adaptation in East Africa’ grant; EP/T015462/1) projects. Seismic hazard calculations were performed using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/. Color maps in Figs 5, 14, 15, 16, A5 and A11 were produced following Crameri et al. (2020). We thank Vitor Silva (Global Earthquake Model Foundation) for providing the seismic hazard values from Poggi et al. (2017). This study benefited from constructive and detailed comments from Bruno Pace, Francesco Visini and an anonymous reviewer.

DATA AVAILABILITY


MATLAB codes for the generation of the MSSM sources and the probabilistic seismic hazard analysis that is described in this study are available at: https://github.com/jack-williams1/Malawi_PSHA and version 1.0 is archived at https://doi.org/10.5281/zenodo.7265780.

REFERENCES


leech river valley fault and devil’s mountain fault system, *Sustainability*, 13(3), 1–36.


Williams, J.N. et al., 2021b. Malawi Active Fault Database (v1.0) [Data set], doi:10.5281/ZENODO.5507189.


Williams, J.N. et al., 2022b. Geologic and geodetic constraints on the magnitude and frequency of earthquakes along Malawi’s active faults: the Malawi Seismogenic Source Model (MSSM), Natural Hazards Earth Syst., 22(11), 3637–3639.


Woessner, J. et al., 2015. The 2013 European Seismic Hazard Model: key different weighings of section, fault and multifault Malawi Seismogenic Source Model (MSSM) source types can be fitted to the regional b-value in the Direct MSSM catalog, (2) if the Adapted MSSM catalogs replicate the seismicity implied by the Youngs & Coppersmith (1985) recurrence models (Appendix A2), (3) in Appendix A3, if the moment rate (M0) of the areal sources in the catalog match the M0 derived analytically from their G–R relation (Poggi et al. 2017), (4) an examination of combining MSSM and areal source seismicity in Malawi (Appendix A4) and (5) in Appendix A5, if the Combined Matched catalogs independent estimates of the geodetic M0 in Malawi (Kreemer et al. 2014; Stamps et al. 2018). Results from this testing are also presented in Section 4.6 in the main article.

### A1 Weighting sources in the Direct MSSM catalog

For a MSSM fault source that is divided along-strike into smaller section sources and/or combined with closely spaced faults into a multifault source, the frequency of one type of source event will impact the frequency of other source types (Williams et al. 2022b); simply assuming that events along these different source types are independent of each other will lead to double counting of the fault’s seismicity. As is common in PSHA, combining these different MSSM source types can be achieved through a logic tree approach, with weightings assigned to each branch to describe their relative likelihood.

Section sources imply relatively frequent moderate magnitude seismicity, whilst multifault sources are indicative of rarer larger magnitude events (Williams et al. 2022b). Therefore, the gradient

APPENDIX A: STOCHASTIC EVENT CATALOG TESTS

Herein, we provide further descriptions of the internal and external tests performed on the stochastic catalog that were used in our probabilistic seismic hazard analysis (PSHA) of Malawi: (1) how different weightings of section, fault and multifault Malawi Seismogenic Source Model (MSSM) sources can be fitted to the regional b-value in the Direct MSSM catalog, (2) if the Adapted MSSM catalogs replicate the seismicity implied by the Youngs & Coppersmith (1985) recurrence models (Appendix A2), (3) in Appendix A3, if the moment rate (M0) of the areal sources in the catalog match the M0 derived analytically from their G–R relation (Poggi et al. 2017), (4) an examination of combining MSSM and areal source seismicity in Malawi (Appendix A4) and (5) in Appendix A5, if the Combined Matched catalogs independent estimates of the geodetic M0 in Malawi (Kreemer et al. 2014; Stamps et al. 2018). Results from this testing are also presented in Section 4.6 in the main article.
of the the Direct MSSM stochastic event catalogs MFD (i.e. b-value) will be sensitive to the weighting of these sources types. A convenient way of weighting these sources is therefore to find the catalog, which when combined with areal source events, has a b-value closest to the regional estimate (1.02; Hodge et al. 2015; Poggi et al. 2017).

To perform this analysis, we generate two million 1-yr long event catalogs for all possible source type weighting combinations at intervals of 0.1 with the limitation that the weighting of any source type \( \geq 0.1 \) \((n = 36, \text{Table A1})\). In cases where a fault source is not divided into section sources and/or combined into a multifault source types, its weighting is accordingly re-adjusted.

### Table A1. The b-value that is derived from generating stochastic event catalogs that consider all 36 weighting combinations of fault, section and multifault MSSM sources. The target b-value (1.02). See also Fig. 8.

<table>
<thead>
<tr>
<th>Fault</th>
<th>Section</th>
<th>Multifault</th>
<th>b-value</th>
<th>Moment rate (Nm yr(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.8</td>
<td>0.98</td>
<td>1.44E-18</td>
</tr>
<tr>
<td>0.1</td>
<td>0.2</td>
<td>0.7</td>
<td>0.97</td>
<td>1.32E-18</td>
</tr>
<tr>
<td>0.1</td>
<td>0.3</td>
<td>0.6</td>
<td>0.96</td>
<td>1.27E-18</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.5</td>
<td>0.95</td>
<td>1.23E-18</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5</td>
<td>0.4</td>
<td>0.95</td>
<td>1.16E-18</td>
</tr>
<tr>
<td>0.1</td>
<td>0.6</td>
<td>0.3</td>
<td>0.94</td>
<td>1.09E-18</td>
</tr>
<tr>
<td>0.1</td>
<td>0.7</td>
<td>0.2</td>
<td>0.94</td>
<td>1.07E-18</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>0.1</td>
<td>0.93</td>
<td>1.03E-18</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1</td>
<td>0.7</td>
<td>1.00</td>
<td>1.50E-18</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>0.6</td>
<td>0.97</td>
<td>1.43E-18</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
<td>0.96</td>
<td>1.32E-18</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
<td>0.96</td>
<td>1.26E-18</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5</td>
<td>0.3</td>
<td>0.95</td>
<td>1.22E-18</td>
</tr>
<tr>
<td>0.2</td>
<td>0.6</td>
<td>0.2</td>
<td>0.94</td>
<td>1.18E-18</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1</td>
<td>0.7</td>
<td>1.00</td>
<td>1.52E-18</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2</td>
<td>0.6</td>
<td>0.98</td>
<td>1.48E-18</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
<td>0.5</td>
<td>0.96</td>
<td>1.44E-18</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.95</td>
<td>1.40E-18</td>
</tr>
<tr>
<td>0.3</td>
<td>0.5</td>
<td>0.3</td>
<td>0.95</td>
<td>1.36E-18</td>
</tr>
<tr>
<td>0.3</td>
<td>0.6</td>
<td>0.2</td>
<td>0.95</td>
<td>1.32E-18</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
<td>0.5</td>
<td>1.00</td>
<td>1.55E-18</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>0.4</td>
<td>0.98</td>
<td>1.45E-18</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.96</td>
<td>1.37E-18</td>
</tr>
<tr>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.95</td>
<td>1.31E-18</td>
</tr>
<tr>
<td>0.4</td>
<td>0.5</td>
<td>0.1</td>
<td>0.94</td>
<td>1.30E-18</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.4</td>
<td>0.99</td>
<td>1.53E-18</td>
</tr>
<tr>
<td>0.5</td>
<td>0.2</td>
<td>0.3</td>
<td>0.97</td>
<td>1.46E-18</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.96</td>
<td>1.38E-18</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.1</td>
<td>0.95</td>
<td>1.31E-18</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1</td>
<td>0.3</td>
<td>0.98</td>
<td>1.56E-18</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2</td>
<td>0.2</td>
<td>0.97</td>
<td>1.47E-18</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3</td>
<td>0.1</td>
<td>0.96</td>
<td>1.41E-18</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1</td>
<td>0.2</td>
<td>0.98</td>
<td>1.52E-18</td>
</tr>
<tr>
<td>0.7</td>
<td>0.2</td>
<td>0.1</td>
<td>0.97</td>
<td>1.46E-18</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1</td>
<td>0.1</td>
<td>0.98</td>
<td>1.55E-18</td>
</tr>
</tbody>
</table>

Direct MSSM catalogs are then combined with the areal source catalog for events in the assessed region (Fig. 6).

We use the maximum-likelihood indicator to derive the b-value of each catalog (Aki 1965) over the magnitude range Mc 4.5–7.6. At greater magnitudes, the MFD is non-G–R (Fig. 8). From this, we find the optimal source type weighting is 0.3–0.1–0.6 for fault, section and multifault MSSM sources respectively (b-value = 0.998; Fig. 8, Table A1). No catalog can directly replicate the estimated b-value of 1.02 in Malawi (Table A1). However, the b-value for the optimal source type weighting is within the b-value measurement uncertainty (–0.95–.105; Hodge et al. 2015). The low weighting assigned to MSSM section sources in this approach reflects that much of the moderate magnitude seismicity they produce is instead incorporated into the areal source events; hence to maintain a b-value \( \sim 1 \), the MSSM sources should be dominated by larger fault and multifault events.

This MSSM source type weighting analysis does not explicitly consider the earthquake rates produced by each weighting combination \( (M_0, \text{Table A1}) \). These variations are a consequence of the Leonard (2010) scaling relationships, which disproportionately increase the \( M_0 \) of longer (and wider) sources. Since the \( M_0 \) inferred from recorded seismicity in Malawi is poorly constrained (Section 4.6.2; Hodge et al. 2015; Ebinger et al. 2019), we did not consider this a useful constraint when selecting the source type weighting combination.

### A2 Adapted MSSM sources

In the Adapted MSSM approach, we distribute a source’s \( M_0 \) (eq. 2) across either a G–R or characteristic MFD (Section 4.2, Fig. 4). In this section, we first provide the equations given by Youngs & Coppersmith (1985) and Convertito et al. (2006) that were used to develop these sources and then examine if the stochastic event catalogs reproduce the seismicity anticipated by these recurrence models.

For a G–R distribution, the probability density function \( f_M(m) \) for magnitude \( m \) is given by

\[
f_M(m) = \frac{\beta e^{(-\beta m - M_{min})}}{1 - e^{(-\beta (M_{Max} - M_{Min}))}}, \quad \text{for} \quad M_{min} \leq m \leq M_{Max} \quad (A1)
\]

where \( \beta = b \ln(10) \) and \( M_{min} \) and \( M_{Max} \) describe the range of event magnitudes that are assessed for each source. For a characteristic recurrence model, \( f_M(m) \) is given by

\[
f_M(m) = \begin{cases} \frac{\beta e^{(-\beta m - M_{min})}}{(1 + C e^{-\beta (M_{Max} - M_{Min}) - \Delta m_1} - e^{-\beta \Delta m_2})}, & \text{for} \quad M_{min} \leq m \leq M_{c} \\ \frac{\beta e^{(-\beta m - M_{Max} - \Delta m_1 - \Delta m_2)}}{(1 + C e^{-\beta (M_{Max} - M_{Min}) - \Delta m_1} - e^{-\beta \Delta m_2})}, & \text{for} \quad M_{c} \leq m \leq M_{Max} \end{cases} \quad (A2)
\]

where \( \Delta m_1 \) is the magnitude range across which the G–R portion of the source’s MFD that has a recurrence rate lower than the characteristic portion, \( \Delta m_2 \) is the magnitude range over which characteristic earthquakes occur and which is bounded by the minimum characteristic earthquake magnitude (\( M_c \)) and \( M_{Max} \) (Fig. 4a; Youngs & Coppersmith 1985; Convertito et al. 2006). The constant \( C \) is

\[
C = \frac{\beta}{1 - e^{-\beta (M_{Max} - M_{Min})}} \Delta m_2 \quad (A3)
\]

The annual frequency, or ‘activity rate,’ for events with \( m \geq M_{min} \) for a G–R magnitude frequency relationship \( \alpha_{GR} \) is

\[
\alpha_{GR} = \frac{\hat{M}_0 (c - b)(1 - e^{-\beta (M_{Max} - M_{Min})})}{b M_0^\beta e^{-\beta (M_{Max} - M_{Min})}} \quad (A4)
\]
Figure A1. Nine MFD for the Chingale Step Fault for (a) characteristic and (b) G–R type seismicity in the MSSM-adapted approach (Youngs & Coppersmith 1985; Convertito et al. 2006). Each MFD considers a different $b$-value and $M_{\text{Max}}$ combination and is assigned a weighting as described in Fig. 2. (c) Comparison of these MFD $M_0$ for a given fault width, $b$-value and $M_{\text{Max}}$ case. The target $M_0$ indicates the $M_0$ as calculated from eq. (2). Results for the length-limited case and are also shown in Table A1.

Table A2. Comparisons of the Chingale Step Fault $M_0$ that are produced by a characteristic and G–R models in the Adapted MSSM approach. In this approach, nine recurrence models are generated for each MFD, which cumulatively explore uncertainty in the $b$-value and $M_{\text{Max}}$. These are randomly sampled in the event catalogs using the weightings in Fig. 2. Analysis for the length-limited case. $M_0$ ratio is the ratio of the recurrence model’s $M_0$ to the Chingale Step Fault’s $M_0$ as calculated from eq. (2) ($3.47 \times 10^{15}$ Nm yr$^{-1}$). Results also shown in Fig. A2.

<table>
<thead>
<tr>
<th>$b$-value</th>
<th>$M_{\text{Max}}$ shift</th>
<th>Weighting</th>
<th>Characteristic $M_0$ (Nm yr$^{-1}$)</th>
<th>Char $M_0$ ratio</th>
<th>G–R $M_0$ (Nm yr$^{-1}$)</th>
<th>G–R $M_0$ ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.12</td>
<td>-0.15</td>
<td>0.048</td>
<td>$3.39 \times 10^{15}$</td>
<td>97.7 per cent</td>
<td>$3.14 \times 10^{15}$</td>
<td>90.5 per cent</td>
</tr>
<tr>
<td>1.12</td>
<td>0</td>
<td>0.096</td>
<td>$3.40 \times 10^{15}$</td>
<td>98.9 per cent</td>
<td>$3.18 \times 10^{15}$</td>
<td>91.7 per cent</td>
</tr>
<tr>
<td>1.12</td>
<td>+0.15</td>
<td>0.016</td>
<td>$3.40 \times 10^{15}$</td>
<td>98.0 per cent</td>
<td>$3.21 \times 10^{15}$</td>
<td>92.5 per cent</td>
</tr>
<tr>
<td>1.02</td>
<td>-0.15</td>
<td>0.204</td>
<td>$3.39 \times 10^{15}$</td>
<td>98.0 per cent</td>
<td>$3.30 \times 10^{15}$</td>
<td>94.6 per cent</td>
</tr>
<tr>
<td>1.02</td>
<td>0</td>
<td>0.408</td>
<td>$3.41 \times 10^{15}$</td>
<td>98.3 per cent</td>
<td>$3.32 \times 10^{15}$</td>
<td>95.1 per cent</td>
</tr>
<tr>
<td>1.02</td>
<td>+0.15</td>
<td>0.068</td>
<td>$3.41 \times 10^{15}$</td>
<td>98.3 per cent</td>
<td>$3.35 \times 10^{15}$</td>
<td>95.7 per cent</td>
</tr>
<tr>
<td>0.92</td>
<td>-0.15</td>
<td>0.048</td>
<td>$3.39 \times 10^{15}$</td>
<td>98.3 per cent</td>
<td>$3.37 \times 10^{15}$</td>
<td>97.1 per cent</td>
</tr>
<tr>
<td>0.92</td>
<td>0</td>
<td>0.096</td>
<td>$3.41 \times 10^{15}$</td>
<td>98.3 per cent</td>
<td>$3.38 \times 10^{15}$</td>
<td>97.4 per cent</td>
</tr>
<tr>
<td>0.92</td>
<td>+0.15</td>
<td>0.016</td>
<td>$3.41 \times 10^{15}$</td>
<td>98.3 per cent</td>
<td>$3.38 \times 10^{15}$</td>
<td>97.4 per cent</td>
</tr>
</tbody>
</table>

Table A3. Comparison of total $M_0$ from all MSSM Sources as derived by their slip rate and area (Target $M_0$) and as replicated in the stochastic event catalogs (Catalog $M_0$). For the Adapted MSSM sources, a comparison is to the total $M_0$ derived from eqs (A1) and (A7) (Source model $M_0$).

<table>
<thead>
<tr>
<th>Catalog</th>
<th>Target $M_0$ (Nm yr$^{-1}$)</th>
<th>Source Models $M_0$ (Nm yr$^{-1}$)</th>
<th>Catalog $M_0$ (Nm yr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct MSSM</td>
<td>$1.48 \times 10^{18}$</td>
<td>N/A</td>
<td>$1.52 \times 10^{18}$</td>
</tr>
<tr>
<td>Length-limited, char</td>
<td>$1.61 \times 10^{18}$</td>
<td>$1.58 \times 10^{18}$</td>
<td>$1.63 \times 10^{18}$</td>
</tr>
<tr>
<td>Length-limited, G–R</td>
<td>$1.61 \times 10^{18}$</td>
<td>$1.53 \times 10^{18}$</td>
<td>$1.48 \times 10^{18}$</td>
</tr>
<tr>
<td>Layer-limited, char</td>
<td>$1.80 \times 10^{18}$</td>
<td>$1.78 \times 10^{18}$</td>
<td>$1.78 \times 10^{18}$</td>
</tr>
<tr>
<td>Layer-limited, G–R</td>
<td>$1.80 \times 10^{18}$</td>
<td>$1.71 \times 10^{18}$</td>
<td>$1.65 \times 10^{18}$</td>
</tr>
</tbody>
</table>

Table A4. G–R relationships and moment rate ($M_0$) for areal source zones located within the region assessed during PSHA (Fig. 6).

<table>
<thead>
<tr>
<th>Source ID (Poggi et al. 2017)</th>
<th>Source zone</th>
<th>a-value</th>
<th>b-value</th>
<th>$M_{\text{Max}}$</th>
<th>Discretized $M_0$ (Nm yr$^{-1}$)</th>
<th>Catalog $M_0$ (Nm yr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Tanganynika</td>
<td>2.9</td>
<td>1.02</td>
<td>7.9</td>
<td>$1.2 \times 10^{16}$</td>
<td>$1.3 \times 10^{16}$</td>
</tr>
<tr>
<td>9</td>
<td>Rukwa–Malawi</td>
<td>4.7</td>
<td>1.02</td>
<td>7.9</td>
<td>$7.3 \times 10^{17}$</td>
<td>$7.4 \times 10^{17}$</td>
</tr>
<tr>
<td>13</td>
<td>Kariba-Okavango</td>
<td>2.8</td>
<td>0.99</td>
<td>6.9</td>
<td>$4.3 \times 10^{15}$</td>
<td>$4.5 \times 10^{15}$</td>
</tr>
<tr>
<td>20</td>
<td>Rovuma Basin</td>
<td>2.6</td>
<td>1.02</td>
<td>6.9</td>
<td>$1.8 \times 10^{15}$</td>
<td>$1.8 \times 10^{15}$</td>
</tr>
<tr>
<td>N/A</td>
<td>Nyanga</td>
<td>0.4</td>
<td>0.8</td>
<td>7.0</td>
<td>$2.6 \times 10^{15}$</td>
<td>$3.0 \times 10^{14}$</td>
</tr>
<tr>
<td>N/A</td>
<td>Northeast Mozambique</td>
<td>0.1</td>
<td>0.8</td>
<td>7.0</td>
<td>$1.3 \times 10^{15}$</td>
<td>$1.7 \times 10^{14}$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$7.45 \times 10^{17}$</td>
<td>$7.62 \times 10^{17}$</td>
</tr>
</tbody>
</table>

In this analysis, the $a$-value has been scaled from Table 2 given the overlap between the source zone and assessed region and events $M_W > 7$ have been removed from the event catalogs.

where $M_{\text{Max}}$ is the seismic moment for $M_{\text{Max}}$ and $c$ is the parameter from the relation $\log M_0 = cm + d$ and equals 1.5 (Hanks & Kanamori 1979). For characteristic earthquakes, the activity rate
Figure A2. The (a) median and (b) weighted average of the nine MFD curves for the Chingale Step Fault (slip rate $\sim 0.04$ mm yr$^{-1}$) derived from all possible variations in the $b$-value and $M_{\text{max}}$ in the adapted moment rate balancing approach (see also Fig. 2 and A1). These are then plotted with respect to the MFD of the Chingale Step Fault in the stochastic event catalogs. In (a) a comparison is also made for how the occurrence rate of the Chingale Step Fault in the Direct MSSM approach is replicated in the catalogs. (c) and (d) Equivalent to (a) and (b), but for the Livingstone Fault (slip rate 2.1 mm yr$^{-1}$).

Figure A3. Discretized and event catalog MFD for the six areal sources from Poggi et al. (2017) that lie within the region assessed for PSHA (Fig. 6).

$\alpha_C$ is

$$\alpha_C = \alpha_{NC} \frac{\beta \Delta m_2 e^{-\beta(M_{\text{Max}}-M_{\text{Min}})} e^{\beta(M_{\text{Max}}-M_{\text{Min}})}}{1 - e^{-\beta(M_{\text{Max}}-M_{\text{Min}})}}$$

(A5)

where $\alpha_{NC}$ represents the activity rate of the non-characteristic magnitude range (i.e. for $M_{\text{Min}} \leq m \leq M_c$, Fig. 4a) and is given by

$$\alpha_{NC} = \frac{M_0(1 - e^{-\beta(M_{\text{Max}}-M_{\text{Min}})})}{K M^0_{\text{Max}} e^{-\beta(M_{\text{Max}}-M_{\text{Min}})}}$$

(A6)

where the constant $K$ is defined by

$$K = h e^{\beta \Delta m_2} + e^{\beta \Delta m_2} (1 - 10^{-c \Delta m_2})$$

(A7)

$\beta$ in eqs (A1)–(A7) is taken from the regional $b$-value in Malawi (1.02; Poggi et al. 2017). The Leonard (2010) area-magnitude scaling relationships are used to derive $M_{\text{Max}}$ for each source and hence it is equivalent to the magnitude estimate $m_s$ used in the Direct MSSM approach (Section 4.1). The uncertainty in these parameters is explored by converting them to discrete variables, with the $b$-value shifted by $\pm 0.1$ with weightings of 0.16 for lower and upper cases and $M_{\text{Max}}$ shifted $\pm 0.15$ with weightings of 0.3 and 0.1 for lower and upper cases, respectively (Fig. 2 and A1, Table A2). These weightings are based on expert opinion and follow that of Goda & Sharipov (2021). Nine source-specific recurrence models
To consider whether the stochastic event catalog $M_0$ of the six areal sources that lie within the region assessed during the PSHA (Fig. 6) matches their $M_0$ as derived from their $a$- and $b$-value, we first adjust the $a$-value so that it is consistent with the size of the overlap between the assessed region and the areal source (Table A4). We then discretize this G–R relationship into magnitude bins of 0.01 and calculate the $M_0$ of each bin and for consistency with the event catalogs (Section 4.3), by assuming the magnitude probability distribution follows a truncated exponential relationship (Cosentino et al. 1977). In this $M_0$ comparison, events in the areal source catalog with $m>M_{\text{Max}}$ (i.e. 7.0) were not removed as they are for the PSHA (Section 4.3).

For most source zones, the catalog’s $M_0$ and MFD matches that expected from discretizing their G–R relationship (Table A4 and Fig. A3). Discrepancies do exist for the Nyanga and Northeast Mozambique sources. However, as these sources have very low rates of seismicity (annual probability of $M_W>5$ events is <0.001), they are only minor contributors to Malawi’s seismic hazard. We therefore conclude that the two million simulations in the areal source catalogs are of sufficient duration to characterise off-fault seismicity for the PSHA. Notably, the $M_0$ implied by these areal sources for Malawi ($7.45 \times 10^{17}$ Nm yr$^{-1}$, Table A4) is lower than that derived in in the MSSM-based catalogs ($1.89 \times 10^{18}$ Nm yr$^{-1}$, Fig. 7). We discuss this further in Section 4.6.2.

A4 Combining MSSM and areal sources in stochastic event catalogs

The historical and instrumental record of seismicity in Malawi comprises events on known mapped active faults that are included in the MSSM and on hitherto unknown faults (i.e. ‘off-fault’ or ‘background’ earthquakes). In theory, distinguishing whether these earthquakes are on- or off-fault can be determined by resolving whether their locations fall within km-scale 3D buffer zones around known faults (Powers & Field 2013; Gerstenberger et al. 2022). However, this type of analysis is challenging in Malawi due to high location uncertainties (5–20 km; Jackson & Blenkinsop 1993; Gaherty et al. 2019), such that the 2009 Karonga earthquakes are the only events in the SSA-GEM catalog that can be confidently placed on a known active fault (Biggs et al. 2010; Macheyeki et al. 2015). In using the SSA-GEM catalog to fit G–R relationships to areal sources (Table 2; Poggi et al. 2017) and estimating that the largest areal source event, $M_{\text{Max}}$, is $M_W$ 7, our PSHA implies all SSA-GEM catalog events $M_W<7$ are off-fault. This is not true and hence raises the risk that between $M_W$ 4.5–7, we are double-counting recorded seismicity as both areal source and fault-based source events. To investigate how this simplification impacts our PSHA, we examine: (1) if the areal source seismicity is still consistent with the SSA-GEM catalog after removing the 2009 Karonga earthquakes and (2) the physical basis of areal source events in the context of the MSSM’s development.

For the first test, we consider all earthquakes in the SSA-GEM catalog between 1965–2015 that fall within the Rukwa–Malawi

A3 Areal source catalog validation

![Figure A4. MFD for all events in the Rukwa–Malawi areal source zone between 1965–2015 in the SSA-GEM catalog (‘Unfiltered’) and the MFD after removing the four largest events in the 2009 Karonga earthquake sequence (‘Filtered’). Also shown are the G–R distributions that can be fitted these filtered and unfiltered catalogs (Tinti & Mulargia 1987) and the Rukwa–Malawi source zone stochastic event catalog.](https://academic.oup.com/gji/article-lookup/doi/10.1093/gji/ggab193)
source zone of Poggi et al. (2017), which is the principal areal source in our PSHA (Section 6.2). We then plot the MFD: (1) for all events in this area in this period, (2) all events after removing the four principal \( M_W \geq 5.5 \) earthquakes in the 2009 Karonga sequence (Biggs et al. 2010), (3) the G–R relationships that can be fitted to the filtered and unfiltered catalogs (Tinti & Mulargia 1987) and (4) the stochastic event catalog for this source zone (Fig. A4).

This analysis indicates that the Rukwa–Malawi areal source stochastic event catalog actually corresponds most closely to the SSA-GEM catalog once it has been fitted for the Karonga earthquakes (Fig. A4). This result can be explained by: (1) when deriving the \( b \)-value for areal sources in East Africa, (Poggi et al. 2017) grouped some sources together and so the \( b \)-value assigned to the Rukwa–Malawi source is relatively independent of whether the Karonga earthquakes are included, (2) some events in this earthquake sequence may have been removed during catalog declustering prior to fitting G–R relationships to the sources (Poggi et al. 2017). In either case, the areal sources used in our PSHA for Malawi are not impacted by the (incorrect) inclusion of the 2009 Karonga earthquakes when defining their G–R parameters. Nevertheless, there are likely other ‘on-fault’ earthquakes in the SSA-GEM catalog that we cannot resolve due to their large location uncertainties and so the possibility remains that combining the areal sources in Table 2 with the MSSM sources in the stochastic event catalogs leads to an over-estimate of the \( \dot{M}_0 \) in Malawi. We discuss this further below.

The MSSM assigns most fault slip rates by partitioning regional geodetically derived extension rates across all known mapped faults (Williams et al. 2022b). This systems-based approach is therefore subtly different from traditional methods of obtaining slip rates in...
which each fault is considered individually using on-fault paleoseismic or geomorphic constraints. Significantly, it is implicit in the systems-based approach that not all of the geodetically derived extension rate is converted to fault slip, as a correction is made for the obliquity of faults to the regional extension direction (Williams et al. 2021a). In this context, areal source events could represent the seismic moment that is lost from projecting the fault dip direction through the regional extension azimuth.

To quantify this, we define the obliquity factor (OF) of MSSM source \( i \), as the ratio of the source’s \( M_0 \) relative to its \( M_0 \) if it was optimally oriented to the regional extension direction,

\[
OF_i = \cos(\theta_i - \phi)
\]

(A8)

where \( \theta_i \) is the source’s strike and \( \phi \) is the regional extension it is projected through. In some of the MSSM-based catalogs, a source’s down-dip is not necessarily extrapolated through the full width of the seismogenic layer, which also implies that the geodetic \( M_0 \) is not all converted to seismic \( M_0 \) (Section 4.2). To account for these two effects, we calculated the total \( M_0 \) from all MSSM fault sources on the basis that they were all optimally oriented to the regional extension direction (i.e. \( OF = 1 \)) and that they extend through the full width of Malawi’s 35-km thick seismogenic layer (Ebinger et al. 2019; Stevens et al. 2021). The latter correction was not made for the MSSM sources whose down-dip extrapolation implies they intersect with another fault (Section 4.5) and neither correction was made for MSSM sources whose slip rate was estimated independently from offset of the 75 ka seismic reflection (Shillington et al. 2020; Williams et al. 2022b).

The total \( M_0 \) of all MSSM fault sources, given the corrections for fault obliquity and down-dip extent is \( 1.88 \times 10^{18} \) N m yr\(^{-1}\). By comparison, the combined catalog \( M_0 \) is \( 1.89 \times 10^{18} \) N m yr\(^{-1}\) and so implies that the inclusion of areal sources developed from the SSA-GEM catalog (Poggi et al. 2017) essentially compensates for the geodetic \( M_0 \) that is lost due to normal fault obliquity and limited down-dip extent. Hence, although the Combined Catalog represents an oversimplified combination of areal and MSSM sources, this does not lead to an overestimate the \( M_0 \) in Malawi compared to that inferred from geodesy and optimally oriented faults. This comparison does not, however, consider the possibility of aseismic moment release in Malawi (Section 4.6.2; Zheng et al. 2020).

As a final comparison, of the Combined Catalog’s \( 1.88 \times 10^{18} \) N m yr\(^{-1}\) \( M_0 \), areal sources in Malawi contribute \( 2.8 \times 10^{17} \) N m yr\(^{-1}\) and MSSM sources contribute \( 1.61 \times 10^{18} \) N m yr\(^{-1}\). Our PSHA therefore implies that \( \sim 15 \) per cent of seismicity in Malawi is ‘off-fault.’ Caution should be applied when comparing this value to other seismic hazard models due to differences in how on- and off-fault sources are developed. Nevertheless, this proportion of off-fault seismicity is lower than estimated in other seismic hazard models (20–50 per cent; Field et al. 2014; Johnson et al. 2022).

A5 Analysis of independent geodetic models of Malawi

To derive independent estimates of the geodetic \( M_0 \) in Malawi, we consider the Global Strain Rate Model (GSRM v.2.1; Kreemer et al. 2014) and the Sub-Saharan African Geodetic Strain Rate Model (SSA-GSRM v.1.0; Stamps et al. 2018). Both models are thus distinct from the geodetic constraints used to generate the MSSM-based stochastic event catalogs (Wedmore et al. 2021). We first divide the assessed region shown in Fig. 6 into a grid with intervals of 0.1° × 0.1° longitude and latitude. These grid sizes do not necessarily reflect the true spatial resolution of these geodetic models; however, the spatial variations in strain rate in regions with few stations are minimized (Stamps et al. 2018). For both models, within each grid \( i \), we first calculate: (1) the second invariant of strain: \( (\epsilon_{11}^2 + \epsilon_{22}^2 + \epsilon_{33}^2)/\mu \), where \( \epsilon_{11}, \epsilon_{22}, \) and \( \epsilon_{33} \) are the two principal strain rates in the horizontal plane of grid \( i \) and (2) the strain rate style: \( (\epsilon_{11} + \epsilon_{22})/\max(\abs{\epsilon_{11}}, \abs{\epsilon_{22}}) \) where a positive style indicates extension and vice versa (Kreemer et al. 2014). We then calculate the moment rate of each grid \( M_0(i) \) through

\[
M_0(i) = A_i z \mu \left\{ \begin{array}{ll}
\frac{1}{\mu} \epsilon_{11}, & \text{if } \epsilon_{22} < 0 \\
\frac{1}{\mu} \epsilon_{22}, & \text{if } \epsilon_{22} \geq 0
\end{array} \right.
\]

(A9)

where \( A_i \) is the area of each grid, \( z \) is the thickness of the seismogenic crust (35 km; Ebinger et al. 2019; Stevens et al. 2021), \( \mu \) represents.
Figure A7. Comparison for the length-width scaling for the length- and layer-limited Adapted MSSM sources. Length-limited source widths follow eq. (1), in which they follow the Leonard (2010) scaling up to a width at which they will exceed the 35-km thick seismogenic layer in Malawi (dashed line indicates scaling for a 53° dipping fault). Layer-limited sources are extrapolated down-dip to a depth of 35 km (Section 4.2). Exceptions to these scalings occur when the down-dip extent of MSSM sources are presumed to intersect and cut each other off (Williams et al. 2022b). The layer-limited source with a width >45 km is the St Mary Fault, which has a relatively low dip (45°; Biggs et al. 2010; Kolawole et al. 2018a).

The shear modulus (33 GPa for consistency with; Leonard 2010), \( \theta \) is fault dip (53°, section 4.5) and the three principal strain rates of each grid (\( \dot{\epsilon}_{11} \leq \dot{\epsilon}_{22} \leq \dot{\epsilon}_{33} \)) are derived by invoking that the vertical strain rate (\( \dot{\epsilon}_{rr} \)) is a principal strain rate and that to maintain incompressibility, \( \dot{\epsilon}_{11} + \dot{\epsilon}_{22} + \dot{\epsilon}_{rr} = 0 \) (Bird & Liu 2007; Bird & Kreemer 2015). Eq. (A9) is therefore similar to the \( \dot{M}_0 \) calculation in the SHIFT model (Bird & Liu 2007; Bird & Kreemer 2015); however, we consider the seismic coupling factor (\( c \)) = 1 (Section 4.6.2) and the assumed 53° dip of faults mean they do not satisfy the criteria that \( 1/sin(\theta) = 2 \). To derive the total geodetic \( \dot{M}_0 \) across the assessed region, we sum the \( \dot{M}_{0(i)} \) from each 0.1° × 0.1° grid.

Malawi can be considered as a region of low magnitude extensional deformation in both of the assessed geodetic models (Fig. A5), which is consistent with the model developed by Wedmore et al. (2021) and observations from seismicity (Ebinger et al. 2019; Willams et al. 2019; Stevens et al. 2021). The SSA-GSRM v.1.0 implies greater spatial variability in the magnitude and style of strain in Malawi than the GSRM v.2.1 model. This likely reflects the more comprehensive suite of geodetic data used to develop the SSA-GSRM v.1.0 and although it indicates strike-slip and even contraction, in regions that have experienced normal fault earthquakes (Biggs et al. 2010; Ebinger et al. 2019), such discrepancies may be reconciled by local strain rotations at the scale of individual faults (Twiss & Unruh 1998; Philippon et al. 2015; Williams et al. 2019). The MSSM based Combined Catalog \( \dot{M}_0 \) (1.89 \( \times \) 10\(^{18} \) Nm yr\(^{-1} \); Fig. 7) is approximately intermediate between the total estimates of \( \dot{M}_0 \) derived from these geodetic models (8.2 \( \times \) 10\(^{17} \) and 3.5 \( \times \) 10\(^{18} \) Nm yr\(^{-1} \) respectively; Fig. A5). This is discussed further in Section 4.6.2.
Figure A8. 1D seismic velocity models previously derived in northern (Ebinger et al. 2019) and southern (Stevens et al. 2021) Malawi from short-term seismic deployments. For context the velocity model for the generic rock site from Boore (2016) is also shown for depths 0–8 km.
Figure A9. Empirical cumulative distribution functions for the 20 ground motions intensity values derived for each site for PGA and for (a)–(c) 10 per cent PoE in 50 yr and (d)–(f) 2 per cent PoE in 50 yr. For comparison, the cumulative distribution functions for the beta distribution that is fitted to these 20 values in Fig. 11 is also shown. In addition, we randomly take 20 samples from these beta distribution and then perform a two sample Kolmogrov-Smirnov test for the null hypothesis that these random samples come from the same continuous distribution as the 20 calculated ground motion values. We repeat this test 1000 times and the number of times where the null hypothesis is not rejected (at a 5 per cent significance level) is reported in the title of each plot.
Figure A10. (a)–(c) Seismic hazard curves as shown in Fig. 11 but for a 3 s spectral acceleration. In addition, we show the mean value and Beta distribution that may be fitted to these values for (d)–(f) 10 per cent PoE in 50 yr and (g)–(i) 2 per cent PoE in 50 yr. Line colors represent different event catalogs and symbols represent different GMMs. For $V_{30}$ value of 760 m s$^{-1}$. 

<SEismic hazard curves as shown in Fig. 11 but for a 3 s spectral acceleration. In addition, we show the mean value and Beta distribution that may be fitted to these values for (d)–(f) 10 per cent PoE in 50 yr and (g)–(i) 2 per cent PoE in 50 yr. Line colors represent different event catalogs and symbols represent different GMMs. For $V_{30}$ value of 760 m s$^{-1}$.>
**Figure A11.** Equivalent to Fig. 14 with PGA seismic hazard maps for Malawi, but using the USGS $V_{S30}$ values (Fig. A6; Wald & Allen 2007). Maps (a)–(c) are for 10 per cent PoE in 50 yr and (d)–(f) 2 per cent PoE in 50 yr. Figure is arranged so each column represents a different catalog. Red lines depict the MSSM fault sources (Williams et al. 2022b).
Figure A12. Histogram for the differences in seismic hazard levels in Malawi between the maps presented in this study and from (a) Poggi et al. (2017) for 10 per cent PoE in 50 yr and (b) from the mixed rupture catalog of Hodge et al. (2015) for a 500 yr return period (see also Fig. 16). Histogram considers the difference in hazard levels for each point in a 0.2° × 0.2° latitude and longitude grid across Malawi (n = 756).