Abstract

A regional seismic risk analysis for central-southern Malawi is conducted by focusing on the Bilila-Mtakataka Fault within the East African Rift System and by incorporating local information on population exposure and building vulnerability. The scenario-based earthquake risk assessments account for uncertainty in geometry, position, and rupture pattern of the Bilila-Mtakataka Fault as well as ground-motion variability and are based on the latest 2018 national census data. In addition, Malawi-specific seismic fragility functions, which were developed based on building surveys and laboratory tests of local construction materials, are implemented to reflect realistic seismic vulnerability of unreinforced masonry constructions in Malawi. The results from the earthquake risk assessments and sensitivity analyses based on alternative data and models highlight the importance of incorporating local information on seismic hazard characterisation, population data, and seismic vulnerability of buildings, in comparison with global data and models. For the considered case study region, individual effects of the above-mentioned model components tend to result in 20-30% or greater differences in regional seismic risk metrics, such as the affected population experiencing a certain ground shaking intensity level or the number of collapsed housing units. The improved seismic hazard-risk assessments are more effective in informing future seismic risk mitigation policies and actions.

1. Introduction

Disaster resilience is a key global goal for sustainable development, as strongly advocated by the Sendai Framework (https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030). Integrating hazard, exposure, and vulnerability and quantifying the risk to people and the built environment are essential steps to achieve this goal. Catastrophe modelling offers a viable computational framework for evaluating disaster risks of buildings and infrastructure quantitatively (Woo, 2011; Mitchell-Wallace et al., 2017). The outputs from such analyses are useful for informing disaster risk reduction policies and actions from socioeconomic and financial perspectives. Using a catastrophe model is particularly applicable to developed countries where careful financial risk management against disasters is required for insurance and reinsurance purposes. In contrast, the practice of developing and using catastrophe models is variable for less developed countries. For instance, compared to countries in the Caribbean and Central America, African countries have less experience with the use of catastrophe models and face challenges due to the low-quality and coarse spatial resolution of data and models that are necessary to characterise the hazard, exposure, and vulnerability components.

Malawi is one of the least developed countries in Sub-Saharan Africa and faces various economic and social problems, including a fast-degrading environment, rapid population growth, and a low-income volatile economy. The country is prone to multiple natural hazards, including floods, droughts, strong winds, landslides, and earthquakes (https://thinkhazard.org/en/report/152-malawi). In Malawi, the seismic risk is not negligible for three reasons. First, Malawi is located within the western branch of the East African Rift System, where large earthquakes of moment magnitude (M_w) 7+ have occurred in the past (e.g., 1910 Rukwa earthquake in Tanzania; Ambraseys and Adams, 1991). In Malawi, the 1989 Salima earthquake (Gupta and Malomo, 1995) and the 2009 Karonga sequence (Biggs et al., 2010) affected tens of thousands of people severely, causing economic losses of US$ 28 million and US$ 14 million, respectively (Chapola and Gondwe, 2016). Earthquakes originating from the major faults could affect many people and buildings (Williams et al., 2021).

Second, population growth and urbanisation are occurring at a fast pace, changing the risk profile of the country rapidly (UN-Habitat, 2010). More people are migrating into informal settlements surrounding four urban areas, i.e., Lilongwe, Blantyre, Mzuzu, and Zomba; the urban population has increased from 2.0 million in 2008 to 2.8 million in 2018 (National Statistical Office of Malawi, 2018). Third, traditional masonry structures, made of burnt/unburnt bricks and cement/mud mortar, are seismically vulnerable (Kloukinas et al., 2020). Currently, structural design provisions are applicable to larger engineered buildings (e.g., reinforced concrete) but not to residential houses (Republic of Malawi, 2019). Although guidelines for constructing safer houses, such as Bureau TNM (2016), are available, they are not yet adopted and implemented in local communities (Novelli et al., 2021).

The first milestone towards achieving improved seismic resilience in Malawi is to assess earthquake hazard and risk accurately. Due to the lack of basic information related to geological and geomorphological data, seismicity data, building data, and strength characteristics of construction materials, risk assessments involve major uncertainties. To inform risk management decisions in Malawi, integrated tools for quantitative earthquake risk assessment are essential. On a global scale, the Prompt Assessment of Global Earthquakes for Response (PAGER) system developed by Jaiswal and Wald (2008) provides rapid earthquake risk assessments upon the occurrence of major earthquakes, whereas the Global Earthquake Model (GEM, https://www.globalquakemodel.org/gem) offers earthquake data and computational platforms that are publicly available. However, the building database and seismic fragility models lack country-specific information. At a national level, Hodge et al. (2015) and Goda et al. (2016) attempted to address these problems by developing the first-generation probabilistic seismic hazard-risk models for Malawi based on fault-source-based seismic hazard models and global building collapse prediction models. Although the results provided valuable insight regarding the impact of hazard-exposure-vulnerability characteristics on seismic risk in Malawi, no local data and models were incorporated. Thus the degrees of under/overestimation of seismic hazard and risk due to the use of global data and models were unknown. In short, the current seismic hazard-risk models for Malawi need to be improved by (i) incorporating detailed fault mapping of potential seismic sources, (ii) considering recent census data as well as realistic typologies of the local building stock, and (iii) implementing seismic fragility models that reflect the local geometry and mechanical properties of buildings in Malawi. Importantly, all these improvements make the developed tools more comprehensive and will inform future seismic risk mitigation actions more effectively.

This study presents scenario-based earthquake risk assessments for central-southern Malawi using improved information on earthquake rupture sources, exposure information, local building characteristics and seismic vulnerability. The developed earthquake catastrophe model for Malawi is applied to the 2009 Karonga earthquake sequence from a retrospective perspective. Subsequently, a case study is focused on the Bilila-Mtakataka Fault (BMF; Jackson and Blenkinsop, 1997), which is approximately 110-km long and could rupture synchronously or in segments (Hodge et al., 2018; Williams et al., 2021). There is major uncertainty regarding how a future rupture of the BMF may be realised in terms of earthquake size and geometry. The regional earthquake risk assessment tool presented in this study is innovative and upgrades the previous studies by Hodge et al. (2015) and Goda et al. (2016) in major ways: (i) numerous stochastic earthquake ruptures are generated based on the recent geological/geomorphological studies of the BMF (Hodge et al., 2018; Williams et al., 2021) to account for uncertainty associated with rupture geometry and location; (ii) exposure data are based on the most recent 2018 Malawi census (National Statistical Office of Malawi, 2018); and (iii) seismic vulnerability functions are derived from new mechanical failure-mode analyses (Novelli et al., 2021) and analytical and finite-element analyses (Giordano et al., 2021) of non-engineered masonry buildings in Malawi, parameters of which were obtained from building surveys and experimental tests of local construction materials (Kloukinas et al., 2019, 2020; Voyagaki et al., 2020). The results of the analyses are presented in the form of the number of people who experience a certain level of ground shaking, e.g., peak ground acceleration (PGA) exceeding 0.2 g, and the number of buildings estimated to be at the collapse limit state. Moreover, sensitivity analyses of key hazard-exposure-vulnerability components are carried out by taking alternative models into consideration. The updated seismic hazard-risk assessments together with the sensitivity analysis results will inform the baseline regional seismic risk more accurately.

2. Study Area

Malawi is a landlocked country located on the western side of Lake Malawi, sharing its national borders with Zambia, Tanzania, and Mozambique. The population of Malawi is principally located in rural areas. The average annual population growth rate is 2.9% between 2008 and 2018. The current population and the number of households are 17.6 million and 4.0 million, respectively, with an average household size of 4.4 persons (National Statistical Office of Malawi, 2018). The population is more concentrated in the Central and Southern Regions (44% and 43%, respectively). The four urban areas in Malawi that are designated as *cities* in the census are Lilongwe (capital), Blantyre, Mzuzu, and Zomba, with populations of 989,318, 800,264, 221,272, and 105,013, respectively. The distances from the BMF to Lilongwe, Blantyre, Mzuzu, and Zomba are approximately 80 km, 110 km, 300 km, and 85 km, respectively.

Malawi is a country with moderate seismicity. The seismicity is primarily induced by the continent-scale tectonic movements of the East African Rift System, which are exhibited as a divergent boundary between the Somali and Nubia Plates (Wedmore et al., 2021). Along Lake Malawi, where regional seismicity is concentrated (**Figure 1a**), prominent border faults, such as the Livingstone and Metangula Faults (Wheeler and Karson, 1989; Flannery and Rosendahl, 1990), and intra-basin faults, such as the BMF (Jackson and Blenkinsop, 1997), form the Malawi Rift System. A notable feature of the Malawi Rift System is lower crustal seismicity, reaching depths up to 35 km (Craig et al., 2011). Faulting mechanisms of earthquakes in the Malawi Rift System tend to be of the normal type.

From the seismic hazard perspective, in central-southern Malawi, Salima suffered significant damage from the 1989 earthquake (Chapola and Gondwe, 2016), whereas Mtakataka, Golomoti, and

Balaka are located along the BMF, where the potential seismic risk could be high. The locations of the BMF and the above-mentioned towns are shown in **Figure 1b**. The BMF runs parallel to the M5 road, and its scarp is visible from the road (**Figure 1c**). This study focuses on the BMF in assessing regional earthquake hazard and risk due to its seismic potential and proximity to numerous towns and villages in central-southern Malawi. Poggi et al. (2017) conducted probabilistic seismic hazard assessments for East Africa and provided PGA estimates of 0.1 g to 0.2 g for a the return period of 475 years for central-southern Malawi. The hazard values by Poggi et al. (2017) were estimated ignoring the potential earthquake sources from the major border and intra-basin faults along Lake Malawi.

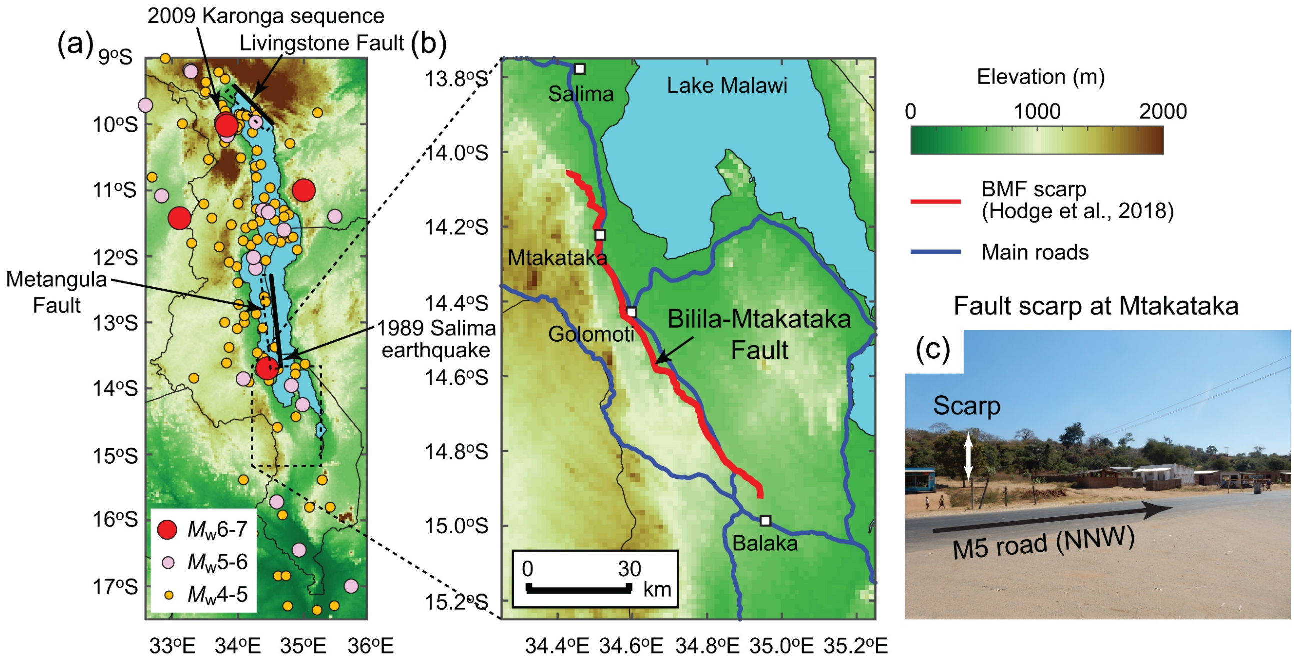


Figure 1. (a) Instrumental seismicity in Malawi based on the International Seismological Centre catalogue since 1965 (http://www.isc.ac.uk/), (b) location of the Bilila-Mtakataka Fault (BMF) based on Hodge et al. (2018), and (c) photo of the BMF scarp at Mtakataka. The background elevation data are based on ASTER-GDEM (https://asterweb.jpl.nasa.gov/gdem.asp).

3. Scenario-based Earthquake Risk Assessment

3.1 *Framework*

A general catastrophe modelling framework, where risk can be computed by convolving three key elements, i.e., hazard, exposure, and vulnerability, is adopted to perform scenario-based earthquake risk assessments for possible rupture of the BMF. The approach implemented in this study captures uncertainties associated with key elements of the analyses and is illustrated in **Figure 2**. Details of the model elements and related uncertainties are explained in **Sections 3.2** to **3.5**.

The seismic hazard module consists of an earthquake rupture model and a ground-motion model. The earthquake source model is developed based on available geological and geomorphological information (Hodge et al., 2018; Williams et al., 2021), which essentially determines the geometry of the potential earthquake rupture plane and the potential size of the rupture. Given the fault geometry, earthquake source scaling relationships (Thingbaijiam et al., 2017) are applied to estimate other earthquake source parameters, such as moment magnitude, mean slip, and width. Numerous realisations of possible earthquake rupture for a given scenario are generated via a stochastic source modelling approach (Goda, 2017). The earthquake source model is then used to simulate the spatial distribution of ground shaking in terms PGA. In simulating shaking intensities at multiple sites, ground-motion models (Akkar et al., 2014; Boore et al., 2014) are employed together with an intra-event spatial correlation model (Goda and Atkinson, 2010) and local site information based on the USGS global V_{S30} map (Wald and Allen, 2007; https://earthquake.usgs.gov/data/vs30/). The above set-up facilitates the generation of stochastic shake maps in the region of interest.

Subsequently, regional earthquake risk is evaluated by integrating stochastic shake maps with exposure information on population and housing units obtained from the 2018 census (National Statistical Office of Malawi, 2018) and with seismic fragility models of non-engineered unreinforced masonry constructions (Giordano et al., 2021; Novelli et al., 2021). The seismic fragility functions by Novelli et al. (2021) were developed based on local building surveys and experiments and the mechanical approach, Failure Mechanism Identification and Vulnerability Evaluation (FaMIVE;

D'Ayala and Speranza, 2003). On the other hand, the seismic fragility functions by Giordano et al. (2021) were developed based on the same survey-experimental data but by adopting a combination of analytical and finite-element analyses of the masonry constructions. This leads to the generation of stochastic earthquake risk maps in terms of the affected population experiencing a certain level of ground shaking and the number of collapsed housing units. These outputs are useful for evaluating the regional earthquake risk and for future earthquake risk management.

The numerical evaluations of the regional seismic risk are performed using Monte Carlo simulations. Considering the above-mentioned framework (**Figure 2**), a stochastic rupture model is first generated (**Section 3.2**), and a ground-motion field is simulated at all building locations by taking into account spatially correlated prediction variability terms (**Section 3.3**). Subsequently, a seismic vulnerability function is evaluated for individual buildings (**Sections 3.4** and **3.5**). Through the aggregation of simulated seismic damage, regional risk for a particular rupture is estimated. This process is repeated numerous times to obtain regional seismic risk assessments for the building portfolio. It is noted that the numerical procedure that is implemented in this study falls between probabilistic and deterministic seismic hazard analyses. The procedure does not account for the frequency of earthquake rupture occurrence, but it does account for ground-motion variability at building locations.

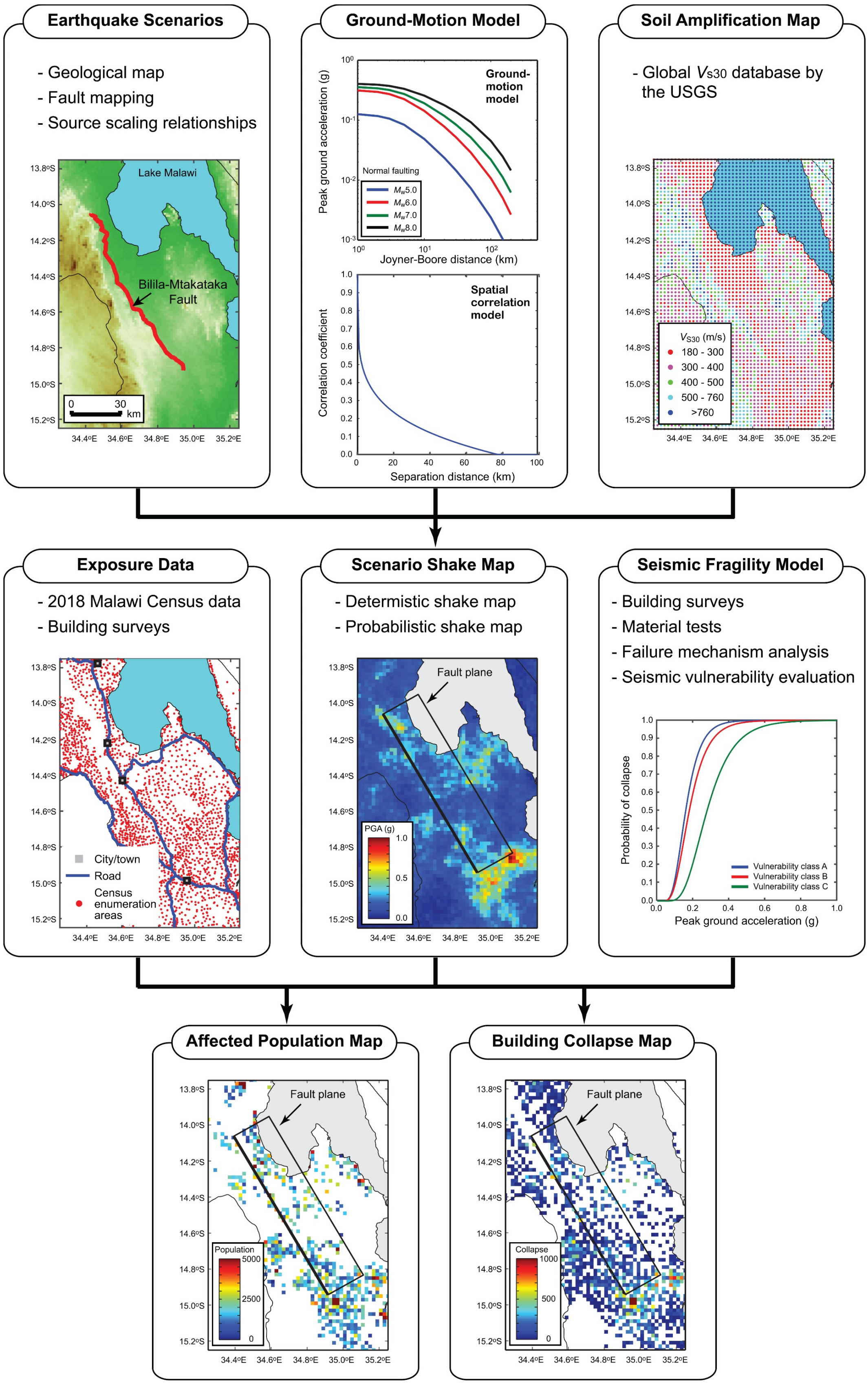


Figure 2. Scenario-based earthquake risk assessment framework.

3.2 *Earthquake rupture model for the Bilila-Mtakataka Fault*

The BMF extends from 5 km north of Balaka northward to 20 km north of Mtakataka with a scarp length of about 110 km (**Figure 1b**). Hodge et al. (2018) studied the fault scarp of the BMF based on geological field investigations, as well as analyses of high-resolution satellite images, and identified six segments along its length. The scarp height ranges between 5 m and 20 m with an average of 11 m. At places, the scarp is visible along the M5 road (**Figure 1c**) but begins to step back in a series of zig-zag patterns north of Mtakataka. On the other hand, Williams et al. (2021) re-examined the fault segmentation structure of the BMF and suggested that the six segments can be reorganised into three, i.e., northern, central, and southern segments. In determining the segmentation structure of the BMF, the reduced scarp height of the central segment may be regarded as a break in fault continuity of the BMF. Following the geological observations, both studies consider that the earthquake rupture of the BMF could occur along the entire length or discretely in segments.

In this study, the four rupture scenarios of the BMF, i.e., whole rupture and three segmented ruptures, are considered for regional earthquake risk assessments (**Table 1**). Scenario 1 spans over the entire BMF (100 km, as indicated by Williams et al. [2021]), whereas scenarios 2 to 4 are local, having fault lengths of 50 km, 15 km, and 35 km, respectively, for the northern, central, and southern segments. The total length of the BMF fault model (i.e., 100 km) is shorter than the scarp length of 110 km, as suggested by Hodge et al. (2018). We consider this difference in fault length as geometrical uncertainty of the BMF, and account for this by allowing the fault plane to extend beyond the target fault plane in stochastic source modelling. The moment magnitudes that correspond to these fault lengths are M_w7.5-7.7, M_w6.9-7.1, M_w5.9-6.1, and M_w6.6-6.8 for scenarios 1, 2, 3, and 4, respectively. In this calculation, the empirical scaling relationships by Thingbaijiam et al. (2017) are used. The BMF is oriented NNW-SSE (strike of 330°) and dips at 53° (Williams et al., 2021). The expected focal mechanism for the BMF is normal faulting. It is noted that the shorter fault lengths (or segmented rupture cases) result in more frequent occurrence but with smaller earthquake magnitudes compared to the whole rupture case. The recurrence periods listed in **Table 1** are as given by Williams et al. (2021). They assessed the recurrence periods based on the fault geometry (mainly scarp height and length), an empirical scaling relationship to estimate the average displacement when a rupture occurs in the specific fault/segment, and the slip rate of the fault/segment estimated from the regional geodetic data. Interested readers should consult Williams et al. (2021). The recurrence periods shown in **Table 1** for the four scenarios are typically longer than the return period value used for national seismic hazard maps of Malawi, which is a 475-year return period.

Table 1. Geometry of the Bilila-Mtakataka Fault (Hodge et al., 2018; Williams et al., 2021).

Rupture scenario	**Fault length[1] (km)**	**North-west corner**		**Fault width[2] (km)**	**Mean slip (m)**	M_w	**Recurrence period[3] (years)**
		Lat. (°)	**Lon. (°)**				
1: Whole fault	100	-14.9500	34.9000	44.9	2.25	7.67	3600
2: Northern segment	50	-14.5566	34.6748	28.3	0.83	7.05	1850
3: Central segment	15	-14.6746	34.7423	12.7	0.15	5.97	950
4: Southern segment	35	-14.9500	34.9000	22.3	0.50	6.73	1600

[1] The strike angle is 330° for all cases; [2] The dip angle is assumed to be 53°; [3] Estimates from Williams et al. (2021).

To account for the uncertainty of the rupture geometry, stochastic earthquake models are generated by reflecting the seismological knowledge of the earthquake rupture (Goda, 2017). The generation of the stochastic rupture planes starts with the specification of a target magnitude range (e.g., M_w7.5-7.7 for scenario 1). Subsequently, values of the length (L), width (W), mean slip (D), strike, and dip are sampled from suitable statistical distributions. The values of L, W, and D are simulated using the scaling relationships for normal faulting earthquakes (Thingbaijiam et al., 2017). On the other hand, the variability of strike and dip angles is approximated by the uniform distribution by considering the range of ±5° with respect to the representative values of 330° and 53°, respectively (Williams et al., 2021). For each scenario, 1,000 stochastic rupture models are generated; this number is sufficient to obtain stable seismic hazard and risk results (i.e., results do not fluctuate significantly, depending on the sample number) as presented in **Section 5**.

Figure 3 illustrates the generation of stochastic earthquake rupture planes for the four scenarios. To limit the spatial extent of the simulated earthquake ruptures within the seismogenic area of the BMF, an overall fault boundary is defined, the length of which coincides with the starting and ending points of the BMF (see the blue rectangles in **Figures 3a,c-e**). This bounding fault plane has a length of 110 km and a width of 45 km (including the fault section north of Mtakataka). With the dip angle of 53°, the deepest limit of the bounding fault plane reaches 37 km, which is consistent with the seismogenic thickness of the Malawi Rift System (Craig et al., 2011). Once the fault length and width are generated from the scaling relationships, the simulated rupture plane is randomly placed (or floated) within the overall bounding rupture plane (i.e., blue rectangle). The simulated parameters (i.e., L, W, and D) must result in a moment magnitude that falls within the target magnitude range, and the simulated fault plane is required to overlap at least 80% of the target rupture fault plane of the scenario of interest (see the red rectangles in **Figures 3a,c-e**). In simulating the fault plane and its position, the red rectangles (i.e., target plane) specify where the fault plane should be located, while the blue rectangles (i.e., bounding plane) serve as a place holder. Simulated fault ruptures that do not meet these criteria are discarded, and the fault parameters are resampled. After floating the simulated fault plane, the variations of the strike and dip angles are taken into account. **Figures 3a,c-e** display the simulated fault planes for scenarios 1 to 4 with grey lines, whereas **Figure 3b** shows histograms of the simulated earthquake source parameters for scenario 1. These simulated rupture planes are used for simulating ground-motion intensities in the target region.

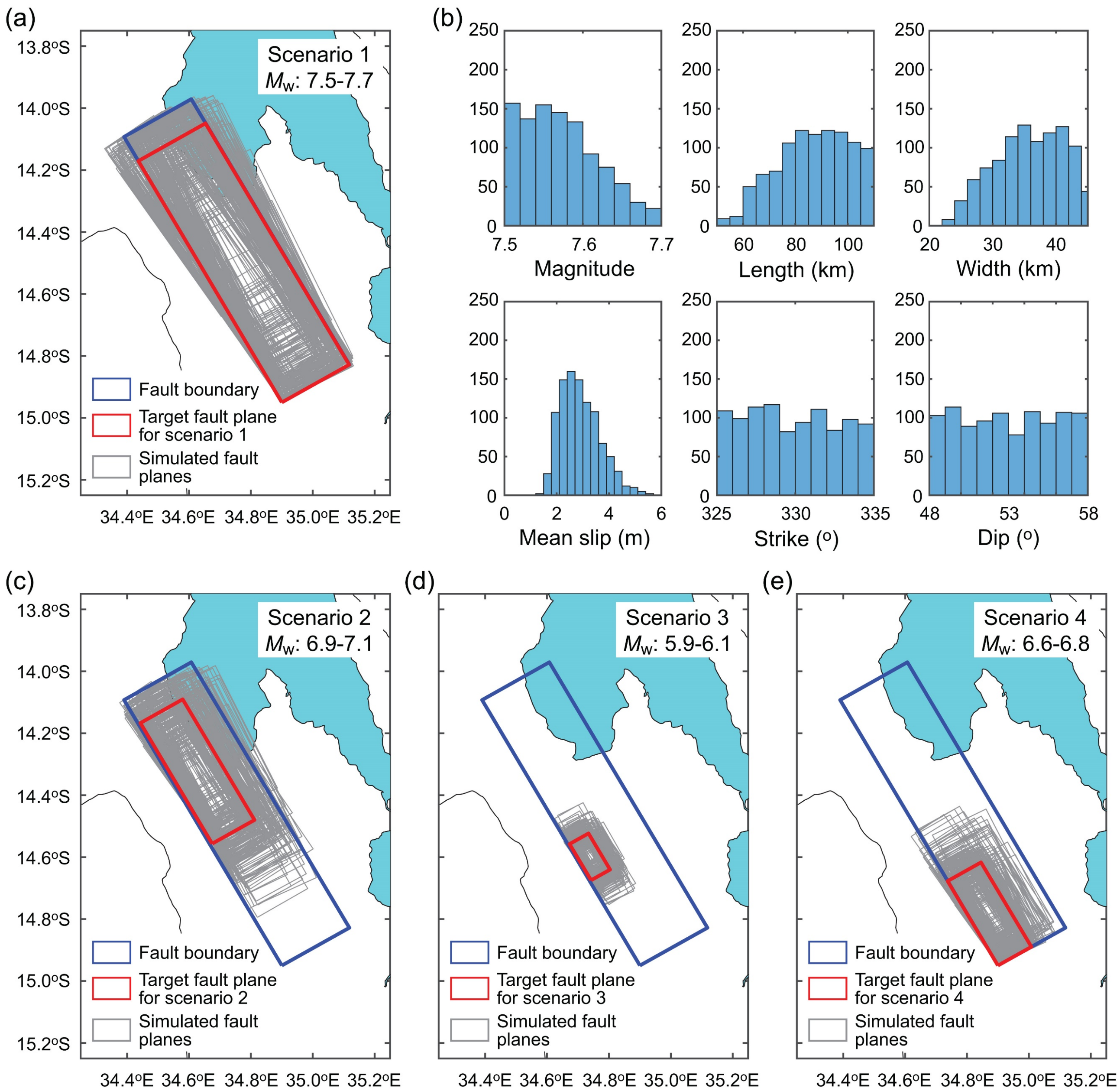


Figure 3. (a) Simulated stochastic fault planes and (b) histograms of the simulated earthquake source parameters for the BMF rupture scenario 1. Simulated stochastic fault planes for the BMF rupture scenarios 2 (b), 3 (c), and 4 (d).

3.3 *Ground-motion model*

Empirical ground-motion models are used to produce stochastic shake maps for a given earthquake scenario. A suitable set of ground-motion models, together with a spatial correlation model of intra-event variability, is required. In this study, due to the lack of regional strong motion data and region-specific ground-motion models for Sub-Saharan East Africa, two ground-motion models, i.e., Akkar et al. (2014) and Boore et al. (2014), that were developed for other seismic regions are adopted with equal weight in a logic tree. For both models, the adopted distance measure is the Joyner-Boore distance, and the focal mechanism term is of normal type. The Akkar et al. (2014) model is based on European ground-motion data, whereas the Boore et al. (2014) model is based on ground-motion data for global crustal earthquakes. Both models are applicable to shallow crustal earthquakes in the active tectonic margin; this is consistent with the previous seismic hazard studies in Malawi (e.g., Hodge et al., 2015; Poggi et al., 2017). The main ground-motion parameter that is adopted in this study is PGA, which is used in the current national seismic hazard map for Malawi (Malawi Bureau of Standards Board, 2014). PGA is preferred to spectral acceleration as the majority of buildings in the region are single-storey. The predicted PGA values for different combinations of moment magnitude and source-to-site distance are shown in **Figure 4a**. The predicted PGA values based on the Akkar et al. (2014) and Boore et al. (2014) equations are different; this is considered as epistemic uncertainty in the seismic hazard-risk assessments, and the models are incorporated as two branches of a logic tree with equal weight. In evaluating the PGA values at given locations due to an earthquake rupture, it is important to consider intra-event spatial variability of the ground-motion parameter at different locations. Empirically, it has been observed that the intra-event residuals are spatially correlated at closer locations than at more distant locations. This aspect is incorporated by simulating stochastic random fields of PGA values at different locations using the intra-event spatial correlation model of Goda and Atkinson (2010), which is shown in **Figure 4b**.

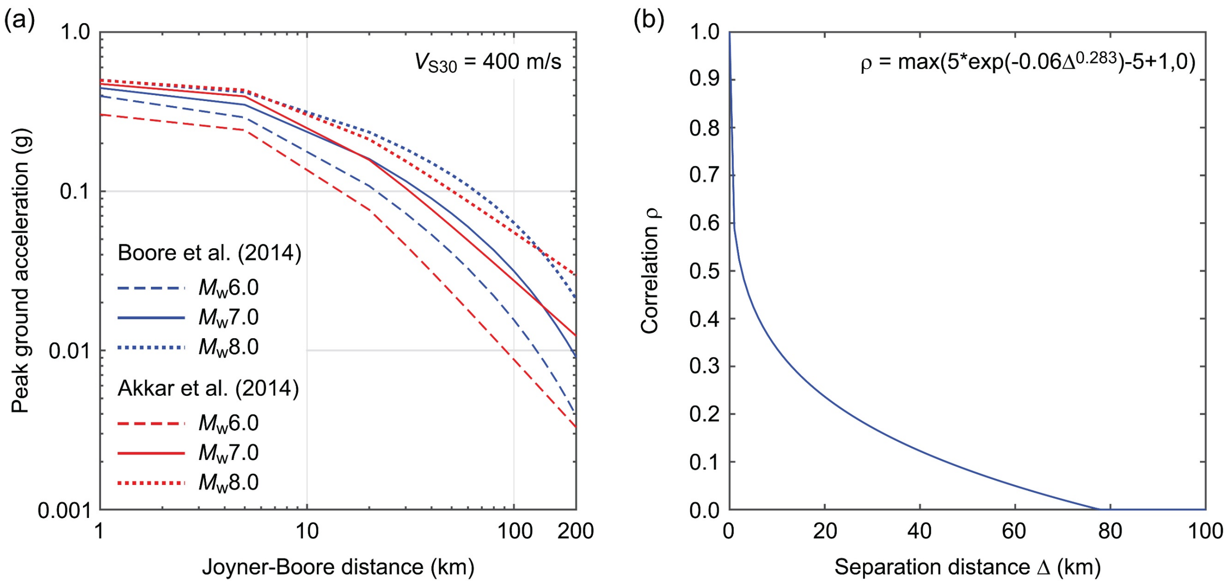


Figure 4. (a) Median predicted peak ground accelerations for a reference site condition of V_{S30} = 400 m/s for different earthquake scenarios based on Akkar et al. (2014) and Boore et al. (2014), and (b) intra-event spatial correlation model based on Goda and Atkinson (2010).

In generating scenario-based PGA shake maps, 3876 grids/cells with a grid spacing of 0.02° (circa 2 km) are set up to cover a target region surrounding the BMF (**Figure 5a**). For given earthquake rupture planes (**Figure 3**), Joyner-Boore distances (R_{jb}, shortest distance from a site to the surface projection of the rupture plane) are evaluated, and PGA values are simulated at these grid cells using a ground-motion model and the spatial correlation model. To capture local site conditions at the grid locations, the average shear-wave velocity in the upper 30 m of soil (V_{S30}) is often used in ground-motion models as a site parameter. Due to the lack of direct measurements of this site parameter in central-southern Malawi, surrogate estimates of V_{S30} based on topographical slopes (Wald and Allen, 2007) are adopted (**Figure 5b**). Although the USGS V_{S30} estimates are crude, they capture the main features of the regional topography well at the resolution of the adopted grid, as inspected by comparing **Figure 5b** with **Figure 1b**. By combining the ground-motion model and the spatial correlation model, multiple realisations of spatially correlated ground-motion fields can be generated (i.e., stochastic shake maps).

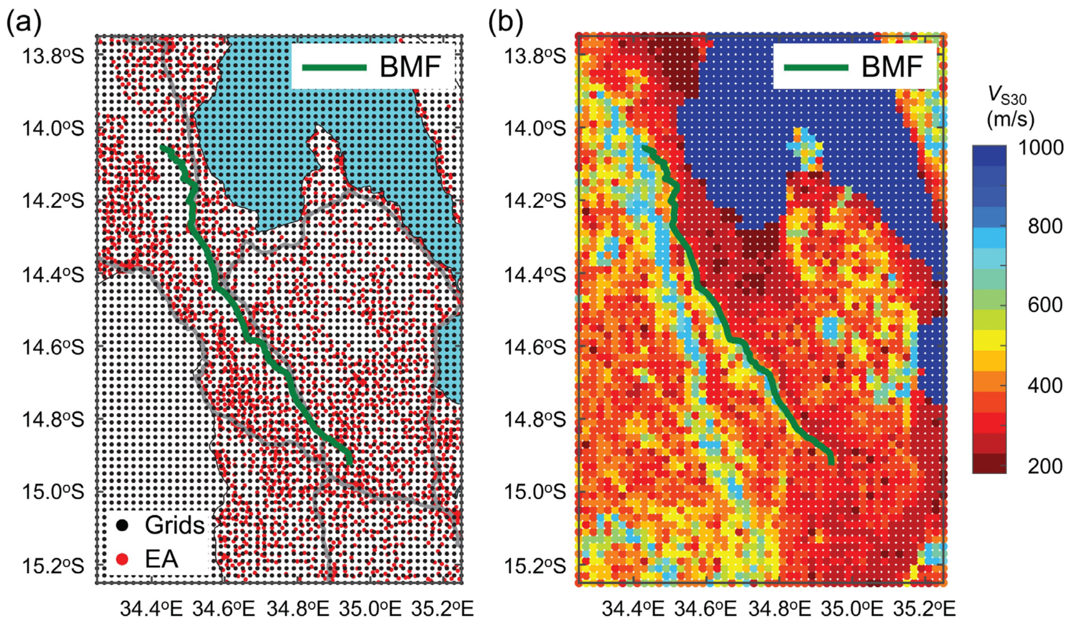


Figure 5. (a) Locations of computational grid points (black dots) and enumeration areas (EA) in the 2018 census (red dots), and (b) inferred values of average shear-wave velocity in the upper 30 m based on the USGS's V_{S30} database.

3.4 *Building characteristics and distribution in central-southern Malawi*

The Malawi census serves as the main reliable source of information on the current population and residential buildings in Malawi. The Malawi census (National Statistical Office of Malawi, 2018) classifies existing dwellings into three categories: (a) *permanent* - made of burnt clay bricks and iron sheet roofs, (b) *semi-permanent* - made of unburnt clay bricks and thatched roofs, and (c) *traditional* - made of rammed earth, daub and wattle or timber walls and lightweight thatched roofs. Out of the 4,805,431 housing units listed in the 2018 census, 41.1% are permanent, 23.0% are semi-permanent, and 35.9% are traditional. In terms of occupancy, 85% of the housing units are owner or family occupied, 12% of the housing units are rented, while the remaining 3% are institutional or other types. This indicates that 97% of the census surveyed dwellings can be considered as single-storey unreinforced constructions, which is the most dominant building type in Malawi (Kloukinas et al., 2020). Based on the definitions of the permanent, semi-permanent, and traditional types in the census, seismic vulnerabilities can be assigned to the building classifications (**Section 3.5**).

The census data are available at different administrative levels. Proportions of the building types (permanent, semi-permanent, and traditional), as well as the occupancy types, are available at the district level (of which there are 32 in Malawi). In contrast, the population and household numbers are available at the enumeration area (EA) level (of which there are 18,799). By assuming that the proportions of building types are uniformly applicable to all EAs within a district, it is possible to obtain the spatial distribution of housing units at EA level. This exposure information can be further reorganised into the computational grid set up for the scenario-based earthquake risk assessments (**Figure 5a**; note that the spatial resolutions of EAs and the grid are comparable). Consequently, there are 2,549 EAs in the target region surrounding the BMF, which accommodate 2,311,766 people (about 13% of the national population) and 623,047 housing units. Out of 623,047 housing units, 261,697 are traditional buildings (42%), 141,526 are semi-permanent buildings (22.7%), and 219,824 are permanent buildings (35.3%). **Figures 6a,b** show the histogram and spatial distribution of grid-based population data, whereas **Figures 6c,d** show the histogram and spatial distribution of grid-based housing unit data. In the grid-based housing unit maps (**Figures 6b,d** as well as **Figure 6f**), unfilled cells show that there are no people or houses based on the census's EA data. The exposure information shown in **Figures 6a-d** is used as baseline data in **Section 4**.

Recent international efforts to develop global-quality exposure data for catastrophe modelling have led to the development of Modelling Exposure Through Earth Observation Routines (METEOR; https://meteor-project.org/data), released in 2020. The METEOR database is developed based on remote sensing data with the minimum usage of country-specific data. For Malawi, the 2015 Malawi demographic and health survey data are mentioned as such country-specific data. It is noted that the total number of buildings in Malawi as indicated in the original METEOR database (3,396,386) is 29% less than the number of housing units in the 2018 census (4,805,431). Considering that the 2018 census data reflects the current situation of the building stock in Malawi more accurately, the building numbers from the METEOR database are rescaled using the total number of housing units from the 2018 census. The histogram and spatial distribution of the rescaled METEOR-based housing unit data for the region surrounding the BMF are shown in **Figures 6e,f**. Comparisons of the histograms (**Figures 6c,e**) and the building distributions (**Figures 6d,f**) indicate that the building counts in the area surrounding the BMF are still significantly underestimated by the METEOR data (after national-level rescaling), compared with the census data (i.e., 454,773 versus 623,047, a difference of 27%). This underestimation is the result of the inaccurate spatial distribution of the population and housing units in the global-quality exposure database, which distributes more buildings in urban areas. The effects of using the rescaled METEOR-based building database on the regional earthquake risk assessments are investigated in **Section 5**.

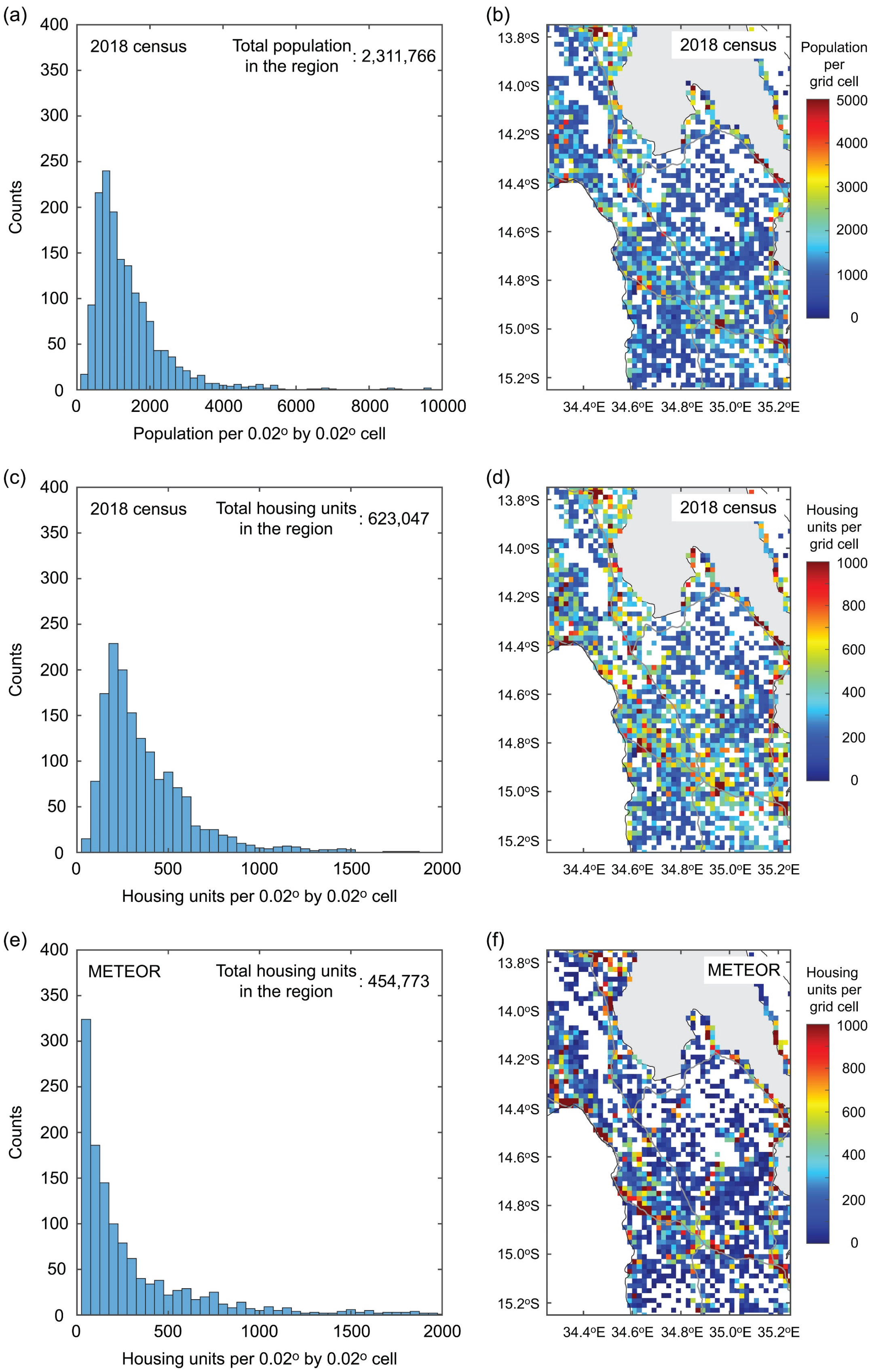


Figure 6. (a,b) Histogram and spatial distribution of the population based on the 2018 census data, (c,d) histogram and spatial distribution of housing units based on the 2018 census data, and (e,f) histogram and spatial distribution of housing units based on the METEOR exposure data.

3.5 *Seismic vulnerability of masonry buildings in central-southern Malawi*

3.5.1 *Seismic fragility functions*

Unreinforced masonry buildings in Malawi are vulnerable to seismic excitations. To investigate the seismic vulnerability of typical masonry houses in central-southern Malawi, Kloukinas et al. (2020) conducted local building surveys in Salima, Mtakataka, Golomoti, Balaka, and Mangochi. Also, a series of in-situ and laboratory tests were conducted to measure the strengths of typical construction materials in Malawi (Kloukinas et al., 2019) and to evaluate the in-plane and out-of-plane strengths of masonry wall panels (Voyagaki et al., 2020).

Moreover, Novelli et al. (2021) gathered detailed geometrical and structural features of Malawian masonry buildings and analysed the seismic vulnerability of the 646 façades using the FaMIVE method and the mechanical properties of the local materials from the experiments. This investigation resulted in the development of two sets of seismic fragility functions. The first set is based on geometric and structural features of the buildings, and all surveyed buildings were classified into seismic vulnerability classes A (poor-quality construction), B (medium-quality construction), and C (high-quality construction) (**Figure 7**; see Novelli et al. [2021] for details). The vulnerability classes can be associated with different seismic vulnerability (i.e., A being the most vulnerable and C being the least vulnerable). The second set is based on the critical failure modes of the surveyed buildings, namely out-of-plane of entire load-bearing walls (OOP), gable overturning (GABLE), overturning of vertical strips of piers or spandrels (STRIP), and in-plane failure of load-bearing walls (IP) (**Figure 8**). It is noted that the failure modes that are considered in this study do not account for foundation damage explicitly.

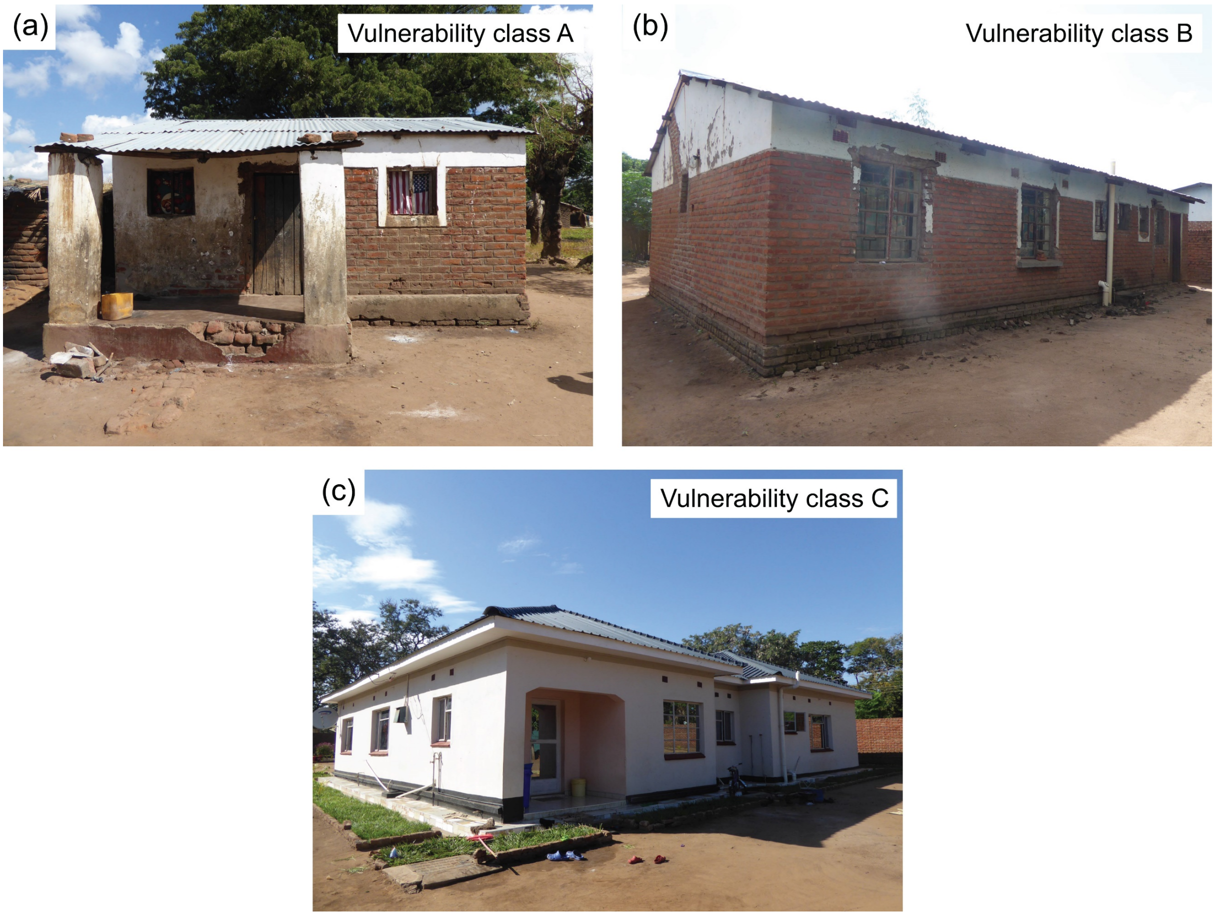


Figure 7. Photos of masonry buildings for seismic vulnerability classes A (a), B (b), and C (c). See Novelli et al. (2021).

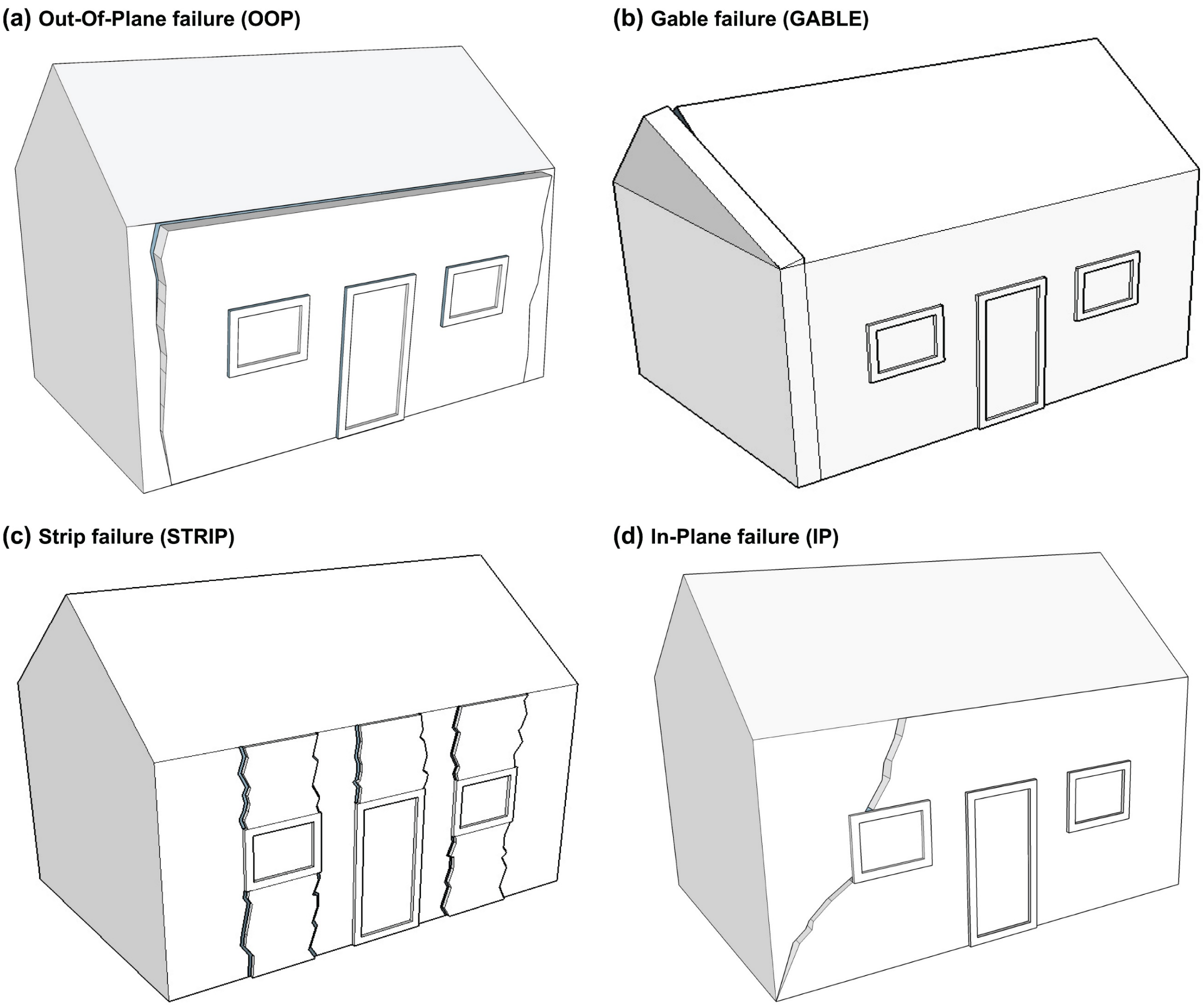


Figure 8. Failure modes that are considered in FaMIVE (Novelli et al., 2021).

In developing seismic fragility functions of Malawian masonry buildings, three different types of ultimate behaviour in terms of building capacity curves (i.e., post-yield part of a force-displacement curve) are considered: geometric instability (GI; D'Ayala and Paganoni, 2011), limited ductility (LD; Lagomarsino, 2015), and strength degradation (SD; Tomazevic, 2007). Different ultimate behaviour of the buildings characterises how buildings respond when they are damaged beyond their elastic limits. The geometric instability behaviour and the strength degradation behaviour consider three branches of a static pushover curve (with different definitions of the control points of the pushover curve), whereas the limited ductility behaviour considers two branches of a static pushover curve (without a plateau in the pushover curve). It is noted that the different ultimate behaviour (GI, LD, and SD) is applicable to vulnerability classes (A, B, and C) and failure modes (OOP, GABLE, STRIP, and IP), respectively. Therefore, there are nine applicable seismic fragility functions when buildings are classified based on the vulnerability classes, whereas there are twelve applicable functions when buildings are classified based on the failure modes. In addition, the seismic fragility functions for the geometric instability behaviour and for the strength degradation behaviour are similar; for this reason, the consideration of one behavioural type (i.e., GI or SD) is practically sufficient to capture the variability of seismic fragility functions due to modelling assumptions (Novelli et al., 2021).

Using the same experimental data (Kloukinas et al., 2019; Voyagaki et al., 2020) and survey data (Novelli et al., 2021), Giordano et al. (2021) developed different seismic fragility functions of the Malawian masonry buildings by considering two representative failure modes (i.e., OOP and IP). In developing the seismic fragility functions for the out-of-plane failure mode, analytical closed-form solutions for walls in one-way bending were applied (Giordano et al., 2020). On the other hand, the seismic fragility functions for the in-plane failure mode were derived through numerous finite-element-based simulations of Malawian masonry buildings by considering the variability of material properties that are informed by the material test results and building survey data. It is important to emphasise that the seismic fragility functions developed by Novelli et al. (2021) and Giordano et al. (2021) differ in the methodologies of developing the seismic fragility functions (and underlying assumptions) and facilitate the incorporation of epistemic (modelling) uncertainty associated with the seismic fragility analysis when both are implemented in regional seismic risk assessments.

To investigate the effects of using global seismic fragility functions in comparison to Malawi-specific seismic fragility functions (Giordano et al., 2021; Novelli et al., 2021), the PAGER functions for the building collapse limit state (Jaiswal et al., 2011) are considered. The PAGER building typologies that are applicable to Malawi are: **M2** (mud walls with horizontal wood elements), **A** (adobe block, mud mortar, straw and thatched roof), **UFB1** (unreinforced fired brick masonry with mud mortar), and **UFB4** (unreinforced fired brick masonry with cement mortar). It is noted that the PAGER seismic fragility functions are expressed in terms of Modified Mercalli Intensity (MMI). To convert the seismic intensity parameter for the fragility functions from MMI to PGA, MMI-PGA conversion equations can be used (e.g., Wald et al., 1999; Caprio et al., 2015). It is important to point out that there is significant uncertainty associated with the conversion between MMI and PGA. For instance, Caprio et al. (2015) developed global and three regional (California, Greece, and Italy) relationships. The global and Greek equations are the same, whereas the Californian and Italian equations produce higher and lower PGA, respectively, for the same MMI with respect to the global relationship (e.g., using the Californian equation results in less vulnerable seismic fragility functions, compared with the global equation). In this study, the global MMI-PGA conversion equation by Caprio et al. (2015) is implemented because the seismic fragility functions for global buildings are more consistent with the seismic fragility functions found for Malawian buildings (Novelli et al., 2021).

In this study, four sets of seismic fragility functions for building collapse are considered: FaMIVE-based models for vulnerability classes (Novelli et al., 2021), FaMIVE-based models for failure modes (Novelli et al., 2021), analytical/finite-element-based models (Giordano et al., 2021), and PAGER-based models (Jaiswal et al., 2011). The implemented functions of the four seismic fragility sets are shown in **Figure 9**. Because of the similarity of the seismic fragility functions for the geometric instability behaviour and for the strength degradation behaviour, only the models for the geometric instability are considered and are shown in **Figures 9a** and **9b**. The seismic fragility functions for the FaMIVE-based and the analytical/finite-element-based models (**Figures 9a-c**) are based on the lognormal distribution as a function of PGA, and the model parameters can be found in Novelli et al. (2021) and Giordano et al. (2021). On the other hand, the PAGER-based seismic fragility functions are computed for a range of PGA values (**Figure 9d**).

Figure 9a shows that the seismic fragility functions for typology A are more vulnerable than those for typologies B and C (**Figure 7**), and the functions for the limited ductility behaviour are more vulnerable than those for the geometric instability behaviour. **Figure 9b** shows the seismic fragility functions for different failure modes that can be ordered as OOP, GABLE, STRIP, and IP, with decreasing seismic vulnerability. By comparing **Figures 9a** and **9b** (i.e., vulnerability classes versus failure modes), the seismic fragility functions for typologies A and B are similar to those for OOP and GABLE, indicating that the overturning of entire load-bearing walls or gables is the dominant factor for building collapse. On the other hand, the seismic fragility functions for typology C are similar to those for STRIP, and the seismic fragility functions for IP have significantly higher seismic resistance than those for typologies A, B, and C. The comparison of the seismic fragility function for OOP by Giordano et al. (2021) (**Figure 9c**) with the FaMIVE-based functions (**Figures 9a** and **9b**) indicates that the former is similar to the FaMIVE-based models for vulnerability classes A and B, as well as for the failure modes OOP and GABLE. The finite-element-based seismic fragility function for IP by Giordano et al. (2021) is similar to the counterparts based on FaMIVE in terms of median, while a significant difference exists in terms of dispersion (i.e., spread of the fragility functions). The larger dispersion of the finite-element-based model can be explained by the consideration of uncertainty in more comprehensive sets of model parameters; see Giordano et al. (2021) for more details. Lastly, comparison of the PAGER-based fragility functions (**Figure 9d**) with the Malawi-specific fragility functions (**Figures 9a-c**) indicates that the former significantly underestimates the latter. The PAGER functions for **A**, **UFB1**, and **UFB4** are applicable to the clay brick masonry buildings investigated by Novelli et al. (2021) as well as Giordano et al. (2021), which reflect more realistic in-situ building geometry and local construction materials and practices in Malawi and thus better represent the seismic vulnerability of Malawian masonry buildings. The effects of considering the global PAGER-based seismic fragility models on the regional earthquake risk assessments are discussed in **Section 5**.

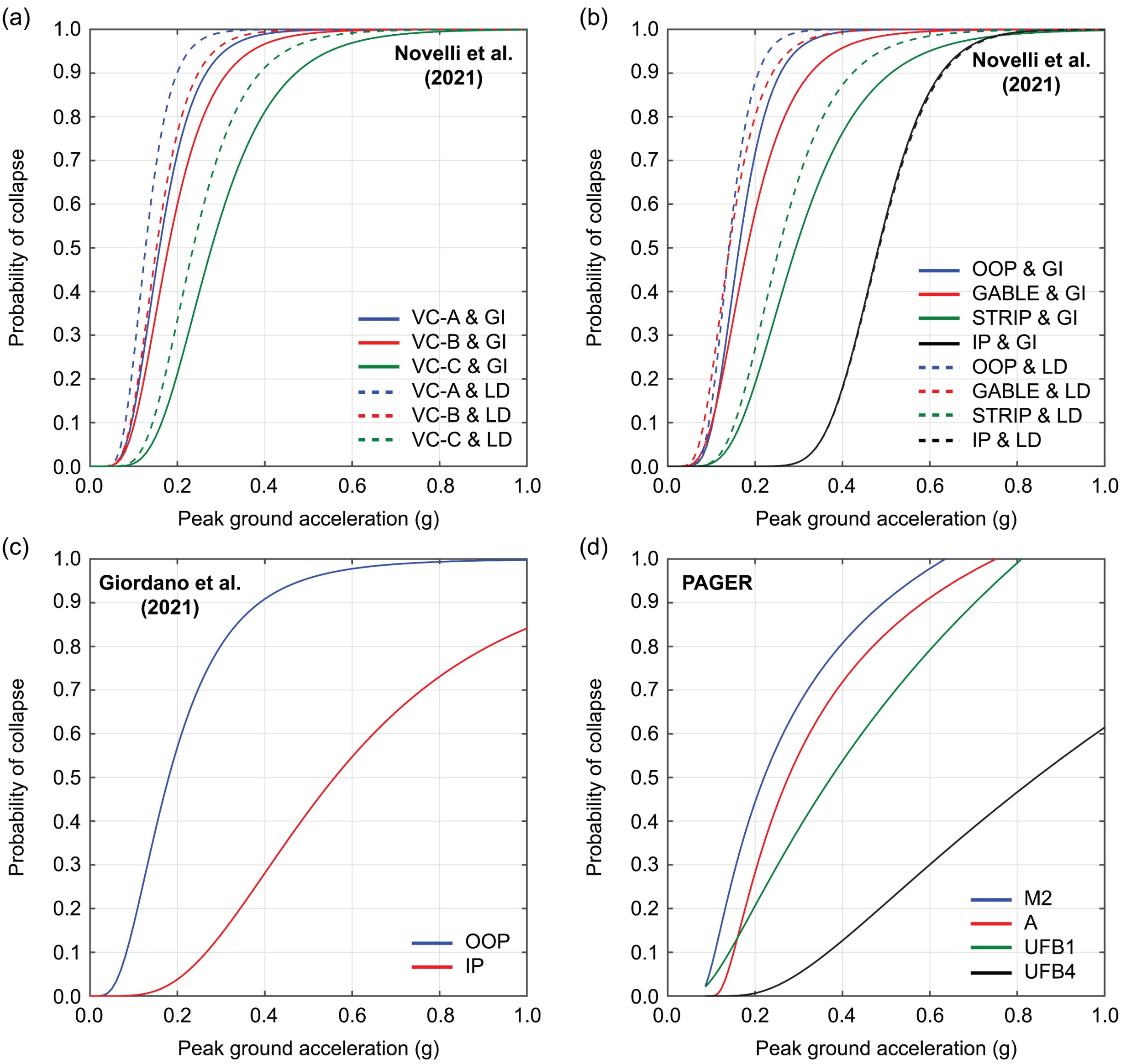


Figure 9. Comparison of seismic fragility functions for the collapse damage state: (a) FaMIVE vulnerability class (VC)-based models (Novelli et al., 2021; see **Figure 7**), (b) FaMIVE failure mode-based models (Novelli et al., 2021; see **Figure 8**), (c) Analytical/finite-element-based models (Giordano et al., 2021), and (d) PAGER-based models (Jaiswal et al., 2011). GI and LD stand for geometric instability and limited ductility, respectively. OOP and IP stand for out-of-plane and in-plane, respectively.

3.5.2 *Assignment of seismic fragility functions to census building data*

To implement the seismic fragility functions that are applicable to Malawi residential buildings in regional seismic risk assessments, it is necessary to associate the four sets of the seismic fragility models (**Figure 9**) with the census building types (i.e., traditional, semi-permanent, and permanent; **Section 3.4**). Taking into account the recent building surveys and experiments that have been conducted in Malawi, four fragility function cases are considered as follows:

- Fragility case 1 adopts the FaMIVE-based functions in terms of vulnerability classes (**Figure 9a**) for semi-permanent and permanent buildings (which are considered to be made with clay bricks), while the PAGER-**M2** function (**Figure 9d**) is considered for traditional buildings (which are considered to be made with mud walls). The semi-permanent and permanent buildings are represented by vulnerability class B and vulnerability class C, respectively, with an equal split based on the geometric instability behaviour and the limited ductility behaviour.
- Fragility case 2 considers that the FaMIVE-based functions in terms of failure modes (**Figure 9b**) are applicable to semi-permanent and permanent buildings, whereas the PAGER-**M2** function (**Figure 9d**) is suitable for traditional buildings. The semi-permanent buildings are represented by the out-of-plane failure mode (60% weight) and by the gable overturning failure mode (40% weight), with an equal split of these weights for the geometric instability behaviour and the limited ductility behaviour. On the other hand, the permanent buildings are represented by the strip failure mode (80%) and by the in-plane failure mode (20%), with an equal split of these weights for the geometric instability behaviour and the limited ductility behaviour. It is noted that the relative weights for different failure modes reflect the building survey data gathered by Novelli et al. (2021).
- Fragility case 3 adopts the analytical and finite-element-based functions (**Figure 9c**) for semi-permanent and permanent buildings, while the PAGER-**M2** function (**Figure 9d**) is applicable to traditional buildings. The semi-permanent buildings are solely represented by the out-of-plane failure mode, whereas the permanent buildings are represented by the out-of-plane and in-plane failure modes with 80% and 20% weights.
- Fragility case 4 considers that the PAGER-based functions (**Figure 9d**) are applicable. Being consistent with the above three fragility cases, traditional buildings are represented by the **M2**

class; semi-permanent buildings are represented by the **A** class; and permanent buildings are represented by the **UFB1** and **UFB4** classes with an 80%-20% split.

Overall, building proportions for different fragility functions are summarised in **Figure 10**. For interested readers, numerical values of the building proportions for different fragility cases and corresponding seismic fragility parameters are also provided as supplementary materials. The considered cases of assigning the fragility functions reflect the current building stock in Malawi and their seismic vulnerabilities based on our experience (Kloukinas et al., 2020; Novelli et al., 2021). It is noted that the assumptions made for traditional buildings may result in underestimation of seismic risks because, from numerical values of the seismic fragility functions, the global PAGER-**M2** function (**Figure 9d**) shows less vulnerability than the Malawi-specific fragility functions for adobe constructions (i.e., functions for vulnerability class A and failure modes OOP and GABLE shown in **Figures 9a-c**). One would expect that mud buildings (**M2**) are generally more vulnerable than adobe brick buildings (**A**), as shown in **Figure 9d**. The consideration of applying the PAGER-**M2** function to traditional buildings was regarded as adequate in avoiding the extrapolation of the Malawi-specific seismic fragility functions for adobe masonry buildings (Giordano et al., 2021; Novelli et al., 2021) to mud-wall masonry buildings that were not directly investigated in our previous studies. The effects of considering the country-specific versus global-quality seismic fragility functions on scenario-based earthquake risk assessments are investigated in **Section 5**. It is also important to clarify that seismic fragility functions that are shown in **Figure 9** correspond to building damage situations illustrated in **Figure 8** (i.e., reaching ultimate behaviour or collapse limit state of a building as specified in the fragility functions). This does not necessarily mean that buildings will be completely destroyed.

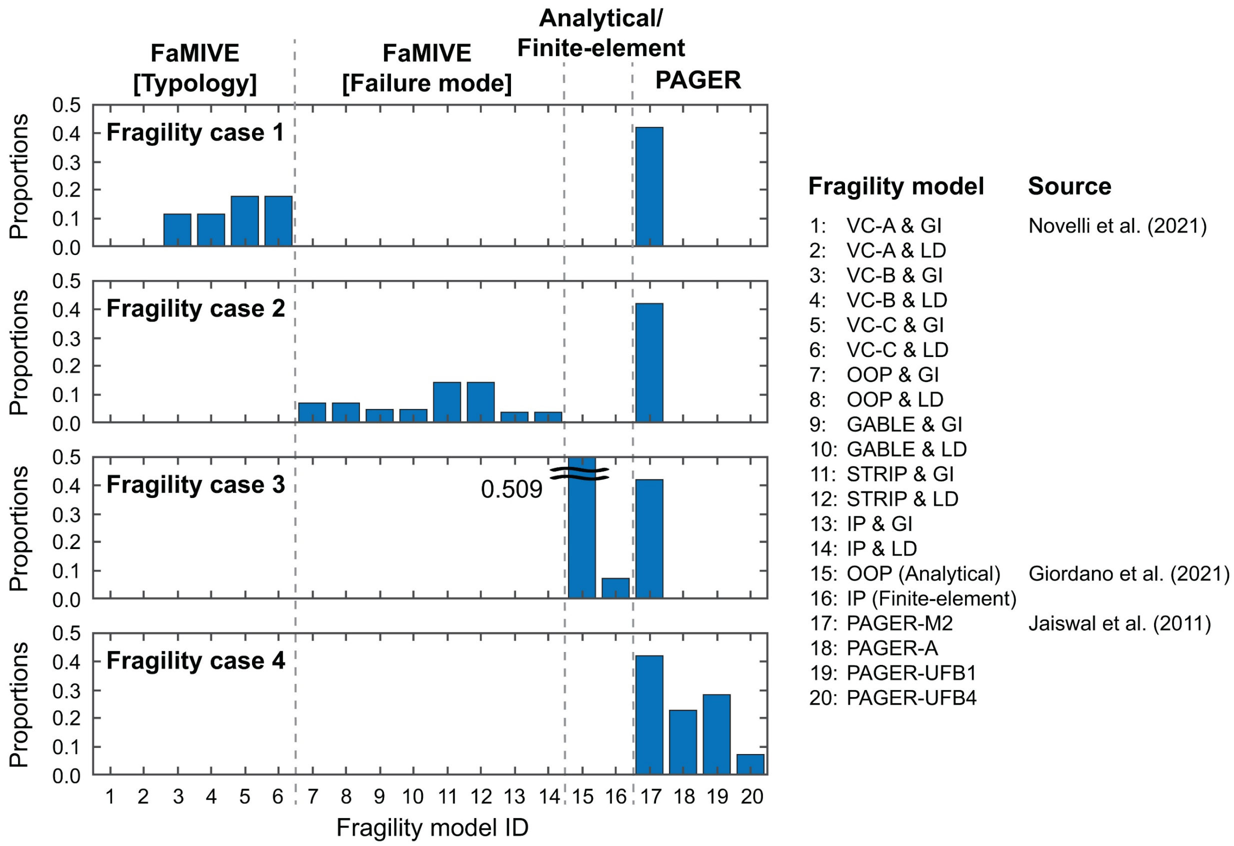


Figure 10: Building proportions of the seismic fragility cases 1 to 4.

4 Retrospective Earthquake Risk Assessment for the 2009 Karonga Sequence

A retrospective earthquake risk assessment for past major events is useful. For Malawi, two major events were the 1989 Salima earthquake in central Malawi and the 2009 Karonga earthquake sequence in northern Malawi (Chapola and Gondwe, 2016). It is important to emphasise that conducting such risk assessments and comparing them with the observed seismic damage and loss in the region are not trivial because there is significant uncertainty associated with observed earthquake shaking, if there are no recorded ground motions, and building exposure and vulnerability may change significantly over the period since these historical events. In this section, the 2009 Karonga earthquake sequence is focused upon because the earthquake event information has been relatively well documented (e.g., USGS, 2009; Biggs et al., 2010) and observed seismic damage has been reported (e.g., Chapola, 2015; Chapola and Gondwe, 2016; Kushe et al., 2017). The Salima earthquake is excluded from this retrospective risk assessment exercise because the seismic damage observations were very limited and the exposure/vulnerability data were significantly different from the current situation as described in **Section 3**.

Reported seismic damage during the 2009 Karonga sequence by different studies is variable. Kushe et al. (2017) cited two different accounts in their paper: "the Karonga earthquakes of 2009 killed four people and destroyed over 5,000 houses (source: Malawi Geological Survey)" and "2,752 houses were affected. Of these, 775 collapsed while 1,154 developed cracks (source: District Commissioner's official report)". On the other hand, Chapola and Gondwe (2016) mentioned that "the earthquakes resulted in collapse of 1,557 houses, rupture the grounds and roads, and cause liquefaction in several areas. Four people died and 300 were injured. A total of 31,220 affected by these earthquakes (source: Malawi Red Cross)." In interpreting these damage reports, it is important to recognize the definitions of 'destroyed', 'cracked', and 'affected' buildings are not identical, and are not the same as the limit states that are defined for the seismic fragility functions introduced in **Section 3.5.1**. For instance, Kushe et al. (2017) showed a photograph of an 'affected' house, for which a half of the wall was collapsed and consequently the roof was collapsed as well. Such damage is classified as 'building collapse' based on the definitions of the seismic fragility functions used in this study (**Figure 8**).

The 2009 Karonga earthquake sequence had four major events whose moment magnitudes were greater than 5.0. USGS (2009) reported that moment magnitudes of the December 6th, 8th, 12th, and 19th events were M_w5.8, 5.9, 5.4, and 6.0, respectively, whereas Biggs et al. (2010) estimated their moment magnitudes as M_w5.7, 5.8, 5.5, and 5.9, respectively. Importantly, the locations of these major earthquakes in the sequence differ significantly. To illustrate such variations, the locations of the four major events are shown in **Figure 11a**. Distances from the event locations to the centre of Karonga were small; the December 8th event was less than 10 km away. Moreover, Biggs et al. (2010) also developed an earthquake rupture model based on InSAR data; the fault plane of the December 8th event is shown in **Figure 11a**. By examining the characteristics of the major events in the Karonga sequence in relation to the proximity to the population distribution of the Karonga district, which is shown in **Figure 11b**, the December 8th event (M_w5.8-5.9) is identified as the critical event of the 2009 sequence for the retrospective seismic risk assessment.

In simulating ground shaking, the same set of ground-motion models mentioned in **Section 3.3** is employed. To demonstrate the expected spatial distribution of ground shaking intensity due to the December 8th event, median PGA values that are calculated using the Boore et al. (2004) and the Akkar et al. (2014) equations are shown in **Figures 11c** and **11d**, respectively (see **Figure 4a**). For demonstration purposes, the finite fault plane based on Biggs et al. (2010) (see **Figure 11a**) is used and variability terms in the ground-motion models are not included. It is important to observe that the Boore et al. model results in higher PGA values than the Akkar et al. (2014) model for this scenario, and PGA values in the near-source areas (less than 10 km from the rupture plane), which include the centre of Karonga (**Figure 11b**) are in the range of 0.2 to 0.3 g. With these levels of ground shaking, masonry buildings with high seismic vulnerability are expected to suffer major damage (**Figure 9**).

In carrying out a retrospective seismic risk assessment, it is essential to adjust the exposure and vulnerability information. For this purpose, the 2008 Malawi census information was consulted for the Karonga region (National Statistical Office of Malawi, 2008). The overall adjustment factor for the number of masonry buildings is calculated as 0.71 by taking the ratio of the 2008 population (256,664) and the 2018 population (359,001). This ratio is considered to be applicable to convert the total number of buildings in the Karonga region (82,920 in 2018) to the 2008 value (59,283). In addition, the proportions of permanent, semi-permanent, and traditional buildings were examined to capture the temporal changes of the building stock in the region. In 2018, the proportions of permanent, semi-permanent, and traditional buildings in the Karonga region were 0.617, 0.186, and 0.197, respectively, whereas in 2008, these proportions were 0.222, 0.325, and 0.454, respectively. The building proportion data clearly indicate the significant transitions from traditional types to permanent types over the decade. These changes are made in the retrospective seismic risk assessment. It is noted that seismic fragility functions are not altered and seismic fragility case 1 is used in the risk assessment (see **Section 3.5.2**).

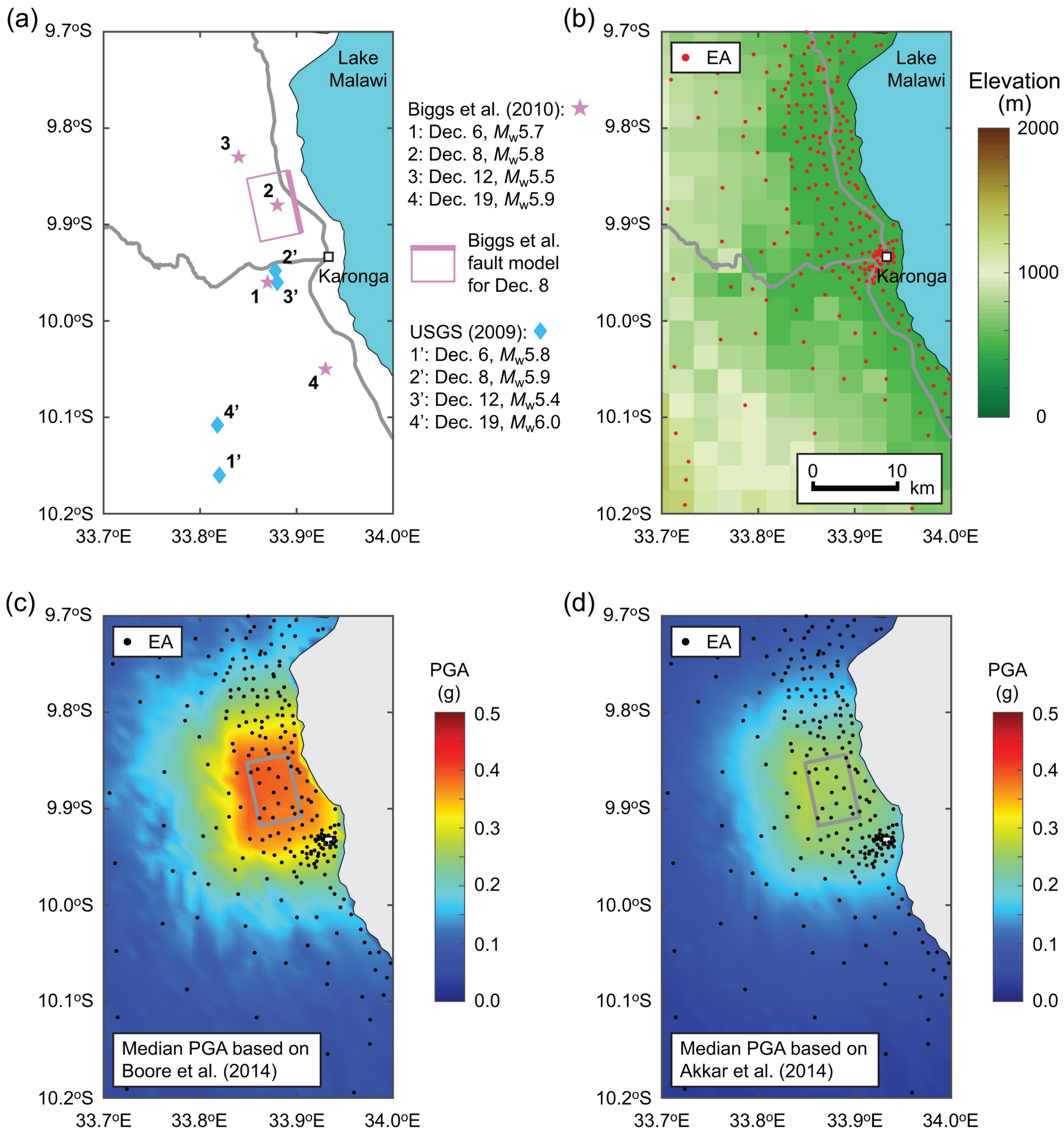


Figure 11. (a) Event information of the four major earthquakes during the 2009 Karonga earthquake sequence, (b) elevation and enumeration area data of the Karonga region, (c) median PGA shake map based on Boore et al. (2014), and (d) median PGA shake map based on Akkar et al. (2014).

The seismic risk assessment results for the December 8th, 2009 event based on three earthquake source cases are compared in **Figure 12** in terms of the cumulative distribution function of the number of collapsed housing units. The three earthquake source cases are: the finite fault plane model based on Biggs et al. (2010) and the point-source models (earthquake centroid) based on USGS (2009) and Biggs et al. (2010). When the finite fault source is considered, distances to buildings near the fault plane become smaller, consequently, ground shaking becomes more intense and results in severer damage and loss. This effect can be seen in **Figure 12**, by shifting the cumulative distribution function of the building collapse counts towards higher values compared with the point-source cases. It is also important to note that the variations of the results are significant; the main contributor of the seismic risk variations is the ground-motion variability.

From the retrospective perspective, the cumulative distribution function of the building collapse counts should be compared with the reported building collapse and damage counts in various reports and studies. This range is indicated in **Figure 12** between 2,000 and 5,000. The observed seismic damage during the 2009 sequence still falls within the predicted range of the building collapse counts but corresponds to the lower end of the distribution (i.e., 15 percentile or less). The correspondence to the lower percentile of the building collapse counts could be attributed to: (i) the actual ground motions during the 2009 sequence being lower than the adopted ground-motion models in this study predict, (ii) the seismic fragility functions derived for Malawi (**Section 3.5.1**) being biased, or (iii) some combination of (i) and (ii). Moreover, in interpreting the presented comparison, it is important to keep in mind that the seismic damage observations may not be completely consistent with what the developed seismic risk model predicts. The retrospective seismic risk assessment is useful in benchmarking the earthquake catastrophe models with respect to actual earthquake damage observations. However, it is difficult to determine which are the main causes of the discrepancy because of significant variability that is present in the seismic damage prediction and the historical event is one realisation of the possible outcomes of the considered earthquake scenario.

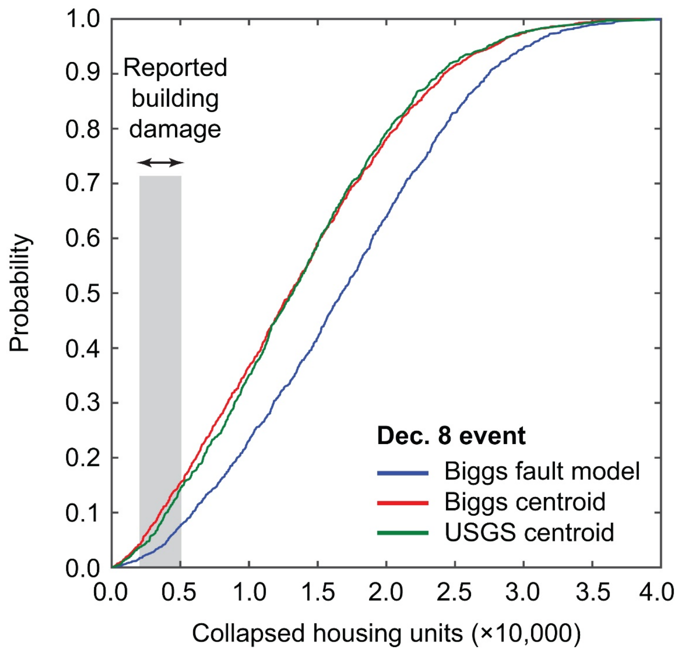


Figure 12. Comparison of cumulative distribution functions of the number of collapsed housing units based on different earthquake source information for the December 8th, 2009 event. See the earthquake event information in **Figure 11b**.

5 Earthquake Risk Assessment for the Bilila-Mtakataka Rupture Scenarios

The scenario-based earthquake risk assessments for the BMF are carried out using the earthquake catastrophe model for Malawi developed in **Section 3**. One of the main objectives of the case study is to highlight the importance of incorporating local information on exposure-vulnerability components compared with global data and models (**Section 5.1**). Moreover, by considering different rupture patterns of the BMF (i.e., synchronous versus segmented), the effects of uncertain seismic hazard scenarios on the regional earthquake risk can be quantified and visualised through integrated critical hazard-risk maps (**Section 5.2**). Such integrated hazard-risk outputs are generated by combining the PGA shake maps, the population maps where a certain shaking intensity level is exceeded (PGA = 0.2 g is adopted in this study), and the building collapse count maps. It is noted that seismic hazard values at the 475-year return period level in this region are typically between 0.1 g and 0.2 g (Hodge et al., 2015; Poggi et al., 2017), and probability of building collapse for the most vulnerable classes of Malawian buildings starts to increase rapidly at these shaking levels (**Figure 9**).

The outputs capture the uncertainties of the seismic hazard scenarios by referring to different regional risk levels. Therefore, it will be possible to define the most likely, optimistic, and pessimistic disaster scenarios for improved earthquake risk management. It is important to recognise that the earthquake hazard and risk assessments presented in this section are conditional on the occurrence of assumed earthquake ruptures. Although mean recurrence periods of these rupture cases could be assigned based on available geological and geophysical data (e.g., Table 1), these recurrence periods do not correspond to the regional seismic risk estimates presented in the following.

5.1 *Effects of local exposure-vulnerability information*

Assigning suitable seismic fragility functions to portions of the building stock of interest involves subjectivity (**Section 3.5**) because complete building-by-building surveys and inspection and determination of individual building capacities are practically infeasible. The primary objective in this section is to evaluate the effects of the exposure-vulnerability information on the regional seismic risk assessment. With this objective in mind, the whole rupture scenario of the BMF is focused upon.

Fragility cases 1 to 4, discussed in **Section 3.5**, attempt to capture the possible variations in assigning seismic fragility functions to individual buildings. **Figure 13** shows the cumulative distribution functions of the number of housing collapses in central-southern Malawi due to the 1,000 stochastic ruptures under scenario 1 (**Figures 3a,b**), considering fragility cases 1 to 4. Fragility cases 1 to 3 are based on Malawi-specific models, whereas fragility case 4 is based on global models (**Figures 9** and **10**). The census-based exposure information shown in **Figures 6a-d** is used. It can be observed from **Figure 13** that for a given fragility case, the effects of seismic hazard characterisation are significant, despite the relatively narrow range of earthquake magnitude (M_w7.5-7.7). The variability can be attributed to the uncertainty in the rupture location and geometry as well as ground-motion variability (both event-to-event and site-to-site). For instance, for fragility case 1, the number of housing unit collapses changes from 155,288 to 444,962 (a factor of 2.87 difference) between the 10^{th} and 90^{th} percentiles of the cumulative distribution function of the building collapse counts, which may be regarded as an indicative range of scenario-based hazard uncertainty (note: the median building collapse count is 298,865). This clearly demonstrates that the variability of the scenario rupture has a major influence and should not be neglected by adopting deterministic rupture scenarios alone.

Returning to our primary focus on the effects of different fragility cases on the probability distribution of the building collapse counts, the comparison for different fragility cases (**Figure 13**) indicates that fragility cases 1 to 3 lead to relatively consistent cumulative distribution functions of the building collapse counts. Fragility case 1 results in intermediate consequences among the three cases. In terms of median building collapse counts, they range from 289,387 (case 2) to 306,050 (case 3) (about 3% differences with respect to case 1). The consideration of global PAGER-based seismic fragility functions (**Figure 9d**) leads to significant underestimation of the building collapse counts, with the 10^{th}, 50^{th}, and 90^{th} percentile values being 114,944, 232,872, and 368,580, respectively. For instance, with respect to case 1, the median collapse count for case 4 is an underestimate by 22% (**Figure 13**). It is noteworthy that this underestimation is only attributed to the building portions of semi-permanent and permanent buildings (58% of the buildings considered), which are made of clay bricks. When Malawi-specific fragility functions for mud-wall buildings are properly accounted for, the differences between the Malawi-specific assessments and the global assessments are likely to be greater. The results clearly highlight the importance of considering the Malawi-specific seismic fragility functions for seismic risk assessments.

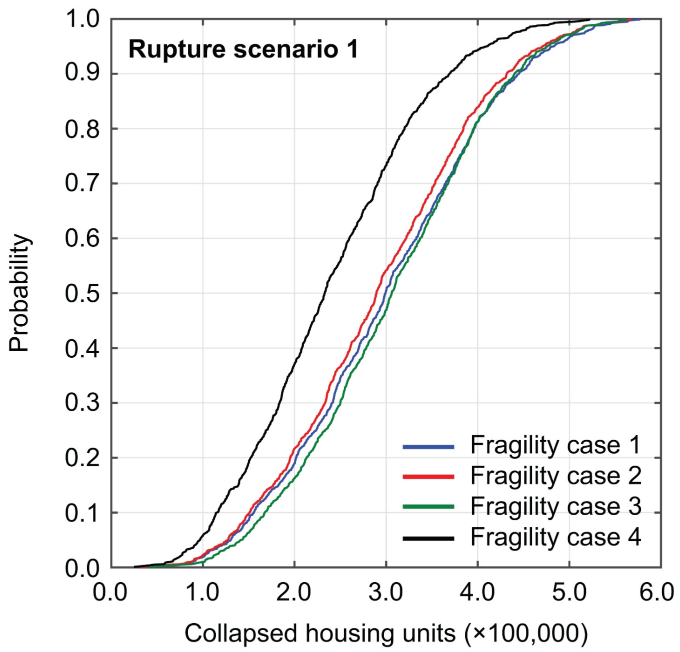


Figure 13. Comparison of cumulative distribution functions of the number of collapsed housing units for rupture scenario 1, considering fragility cases 1 to 4.

Next, the effects of considering global-quality exposure data are examined. The regional earthquake risk assessments are carried out by considering fragility case 1 but with the METEOR-based building distribution (**Figures 6e,f**). The METEOR-based building distribution is rescaled to maintain the same building number as the 2018 census at the national level; however, due to the different spatial distribution of buildings, the number of housing units within the study area is underestimated (**Section 3.4**). The cumulative distribution functions of the number of affected people experiencing PGA > 0.2 g and the number of collapsed housing units are shown in **Figure 14**. It is noted that for **Figure 14a**, the PGA threshold of 0.2 g is considered critical because many masonry buildings in Malawi may be damaged severely or fail at this level of ground shaking (**Figure 9**). The results shown in **Figure 14a** indicate that the consideration of global-quality exposure data leads to 34% underestimation of the exposed population at the specified PGA level in terms of the median (815,696 versus 1,230,783). Similar underestimation can be seen for the number of collapsed housing units (**Figure 14b)**; in terms of median, the building collapse counts are underestimated by 31% (207,654 versus 298,865). These differences are caused by the inaccurate spatial distribution of the population and buildings estimated based on the remotely sensed observations of building distribution. It is important to clarify that if the earthquake risk assessment is carried out for urban areas (e.g., Blantyre and Zomba), the opposite trends will emerge because the METEOR exposure data for Malawi distribute higher proportions of the buildings in cities and their surrounding areas than the census data.

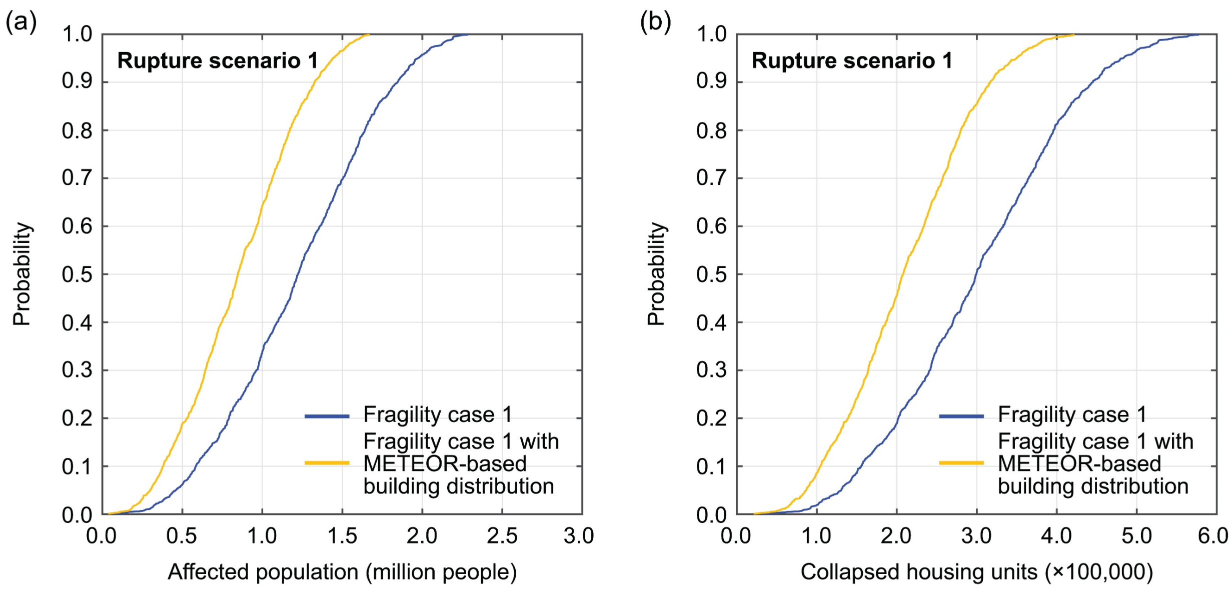


Figure 14. Comparison of cumulative distribution functions of the number of affected people experiencing PGA > 0.2 g (a) and the number of collapsed buildings (b) for rupture scenario 1 and fragility case 1 by considering the 2018 census-based spatial distribution of housing units and the METEOR-based spatial distribution of housing units (see **Figure 6c-f**).

5.2 *Integrated critical hazard-risk maps for different rupture scenarios*

The BMF could rupture in segments. In such cases, the overall regional earthquake risk is likely to be less catastrophic, but such an event is likely to occur with higher frequency; see **Table 1** for the estimates of the occurrence frequencies suggested by Williams et al. (2021) for different rupture scenarios. To quantify the regional earthquake risks for different rupture scenarios and compare them in a systematic manner, the cumulative distribution functions of the number of people who experience ground shaking in excess of 0.2 g and the number of collapsed housing units are shown in **Figure 15** for the four rupture scenarios. The census-based exposure information and fragility case 1 are considered, as the latter produces an intermediate result among the three Malawi-specific fragility cases (**Figure 13**).

The comparison of the cumulative distribution functions of the affected population and the collapsed housing units for different rupture scenarios shown in **Figure 15** indicates that the size of the event or moment magnitude has a significant influence on the overall consequences. The greater magnitude results in greater fault plane size (**Figure 3**), thus affecting a larger number of people and housing units. The curves for rupture scenarios 2 and 4 are relatively close because of similar magnitudes for these two scenarios (M_w6.9-7.1 versus M_w6.6-6.8, though there is a factor of 2.8 difference between M_w7.0 and M_w6.7 in terms of seismic moment). For the same magnitude, rupture scenario 4 tends to result in greater hazard exposure and consequences because the southern segment of the BMF is closer to Balaka (see **Figure 1b** and **Figure 6**). In terms of median estimates of the building collapse counts, rupture scenarios 1 to 4 lead to collapse counts of 298,865, 152,210, 50,023, and 142,023, respectively (i.e., a factor of 6.0 difference between scenario 1 and scenario 3). Given that the estimated frequencies of any of the BMF rupture scenarios are rare, ranging between 1 in 950 years to 1 in 3600 years (**Table 1**), the potential consequences from the BMF rupture are associated with considerable variation (as it is not possible to determine which scenario will be realised in the future). The central and local governments should be aware of such large uncertainty and the nature of low-probability high-consequence events.

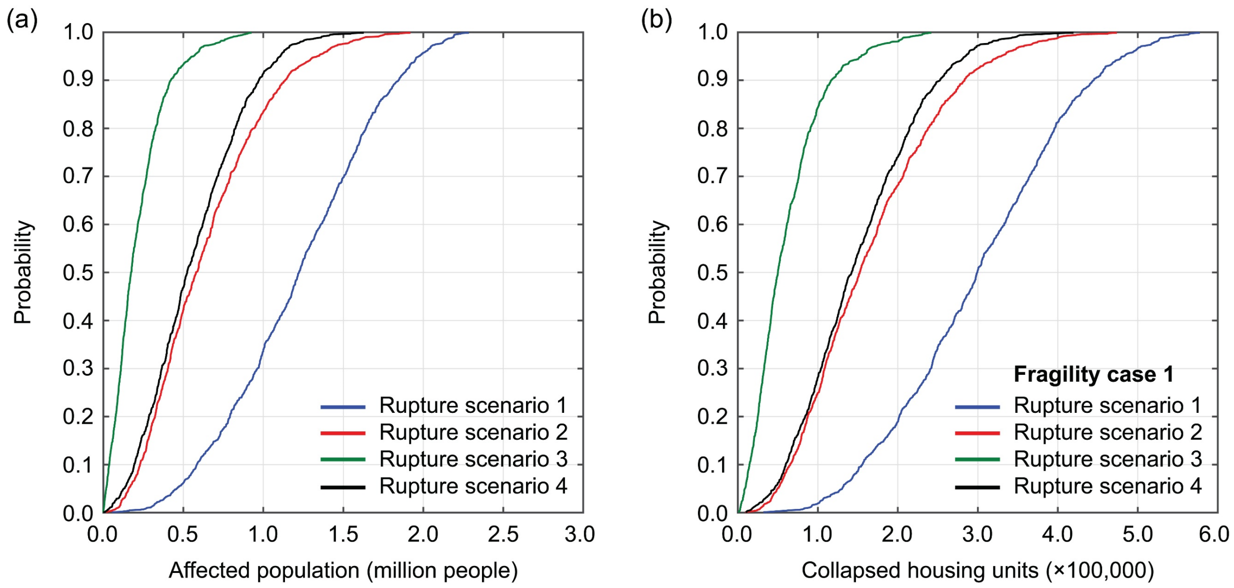


Figure 15. Comparison of cumulative distribution functions of the number of affected people experiencing PGA > 0.2 g (a) and the number of collapsed buildings (b) for rupture scenarios 1 to 4, considering fragility case 1.

To visualise the potential regional earthquake risks for different possible situations, integrated critical hazard-risk maps are useful (Goda et al., 2020). The integrated outputs combine the PGA shake maps, affected population maps, and building collapse count maps by considering three risk-based percentiles (10th, 50th, and 90th) based on the number of collapsed housing units (i.e., **Figure 15b**). Such integrated maps for the four rupture scenarios of the BMF are shown in **Figures 16** to **19**. Note that the PGA maps show values at all grid cells (i.e., coloured), whereas the affected population and the building collapse count maps show values at grid cells with non-zero population and collapsed buildings only. Taking rupture case 1, for instance (**Figure 15**), the PGA shake maps for different risk percentiles, the rupture areas tend to be greater and tend to cause more intense ground shaking at more grid locations (i.e., more yellow-to-red coloured cells exist in the PGA maps). The effects of ground shaking variability (both event-to-event and site-to-site) are significant, as discussed in **Section 5.1**. Accordingly, the highlighted grid cells for the affected people who experience PGA > 0.2 g, as well as for the housing unit collapses, expand rapidly with the risk percentile level. Overall, the combined presentations of three types of maps for three (or more) risk percentile levels offer an effective means to gain insight as to how intense and variable ground shaking can be and how seismic shaking affects the population and building stock in a region. Finally, the comparison of the integrated hazard-risk maps for the four rupture scenarios (**Figures 16 to 19**) facilitates the development of a regional view of seismic hazard and risk in relation to the probability distributions of the potential consequences in the region (i.e., **Figure 15**). Such information can be shared by different stakeholders, including government officers, professional engineers, and community leaders, to initiate more active dialogues in achieving improved seismic preparedness and resilience and to implement disaster risk reduction strategies, such as through the use of the Safer House Construction Guidelines (Bureau TNM, 2016).

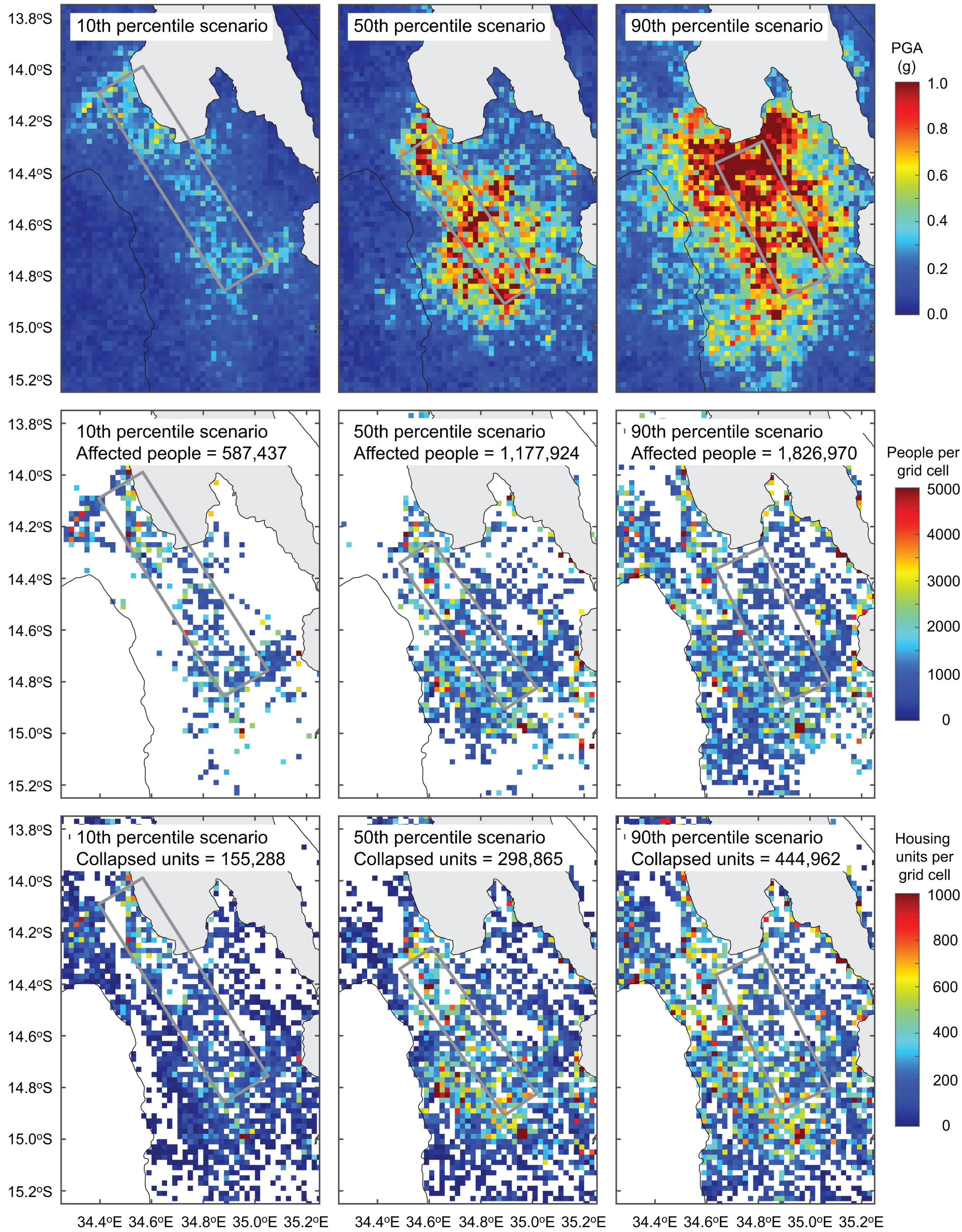


Figure 16. PGA shake map, affected population map for PGA > 0.2 g, and building collapse maps that correspond to the 10th, 50th, and 90th percentile of the building collapse risk levels for rupture scenario 1 and fragility case 1.

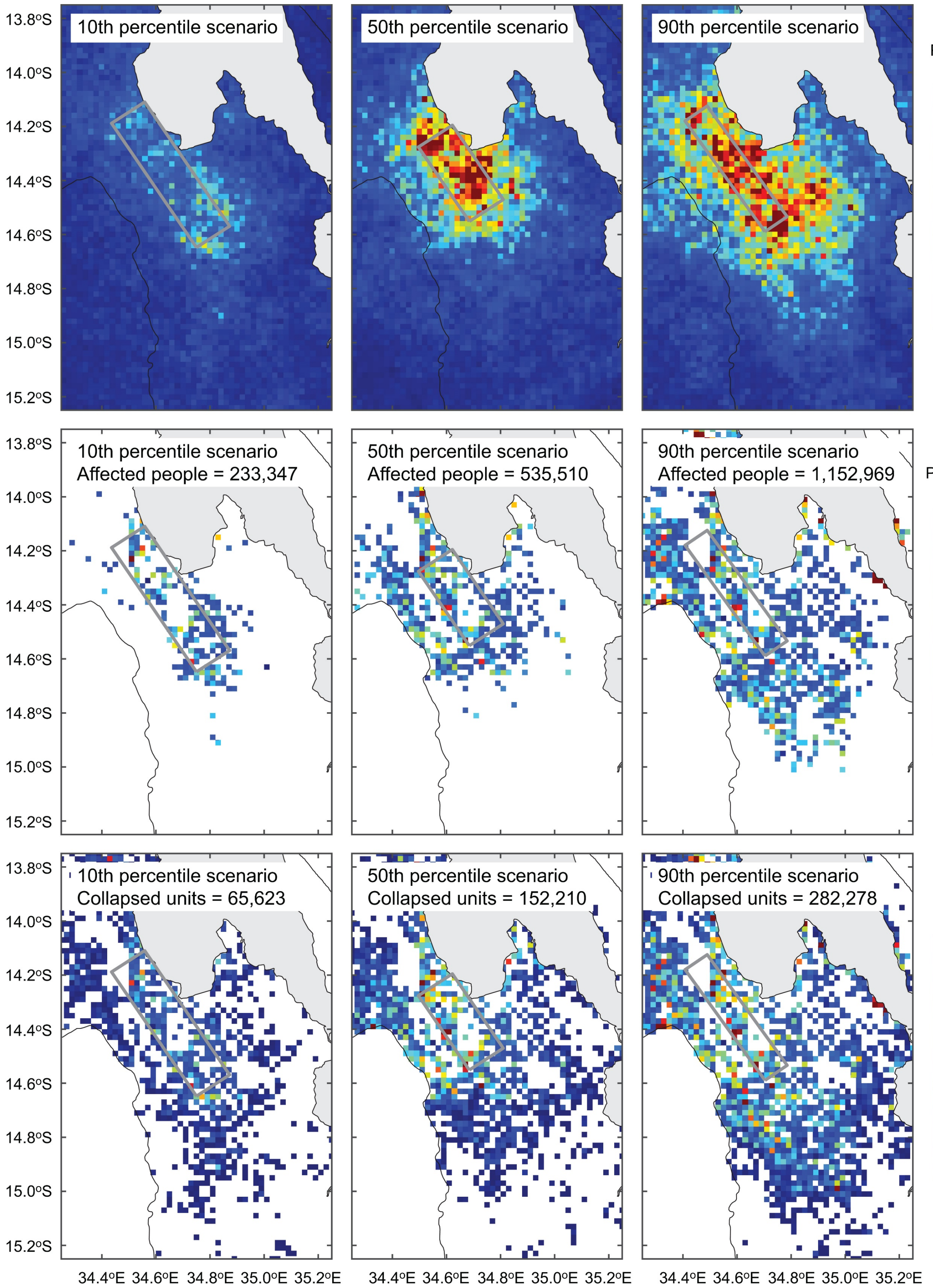
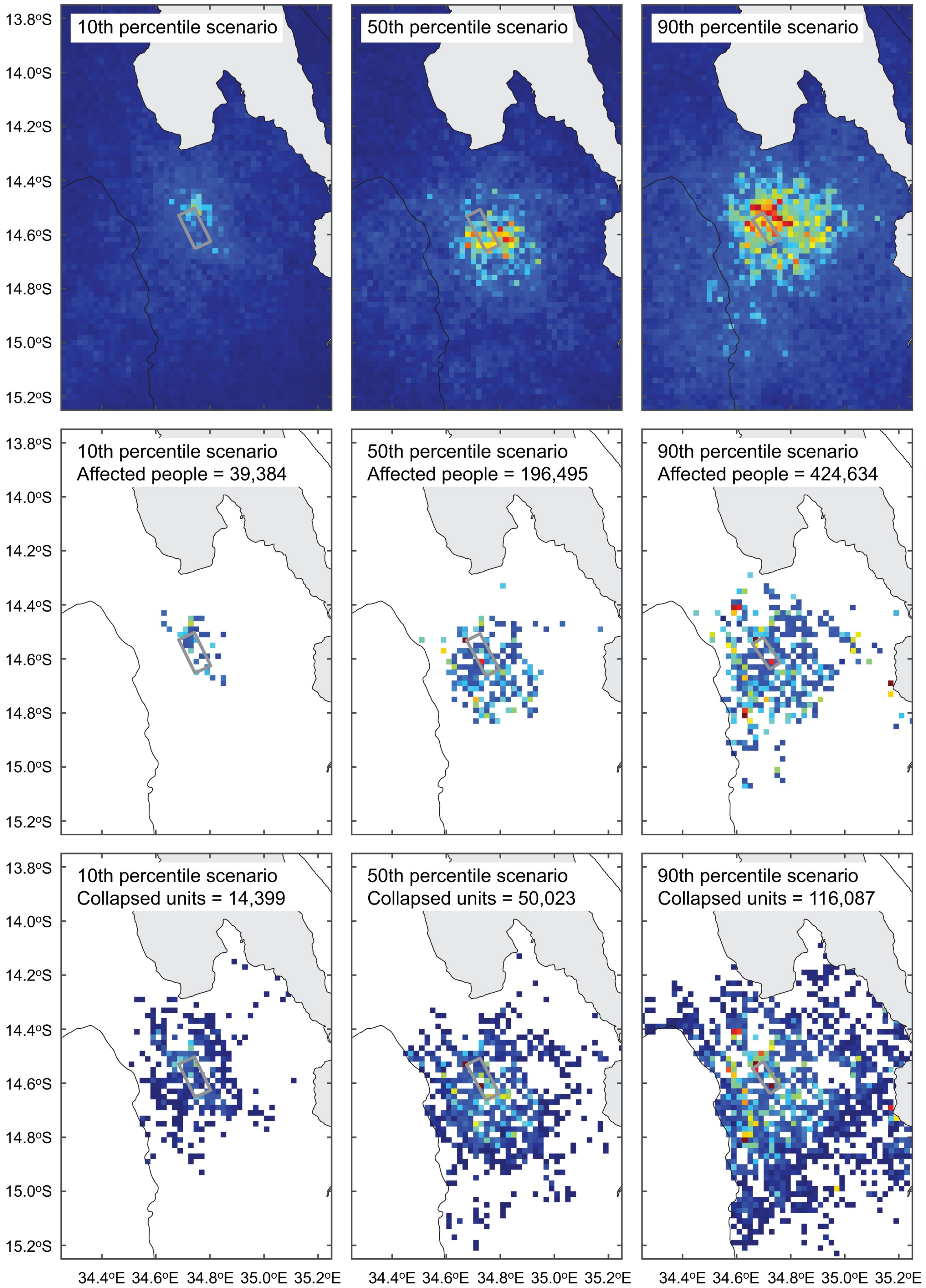


Figure 17. PGA shake map, affected population map for PGA > 0.2 g, and building collapse maps that correspond to the 10th, 50th, and 90th percentile of the building collapse risk levels for rupture scenario 2 and fragility case 1.



758 **Figure 18**. PGA shake map, affected population map for PGA > 0.2 g, and building collapse maps
759 that correspond to the 10th, 50th, and 90th percentile of the building collapse risk levels for rupture
760 scenario 3 and fragility case 1.

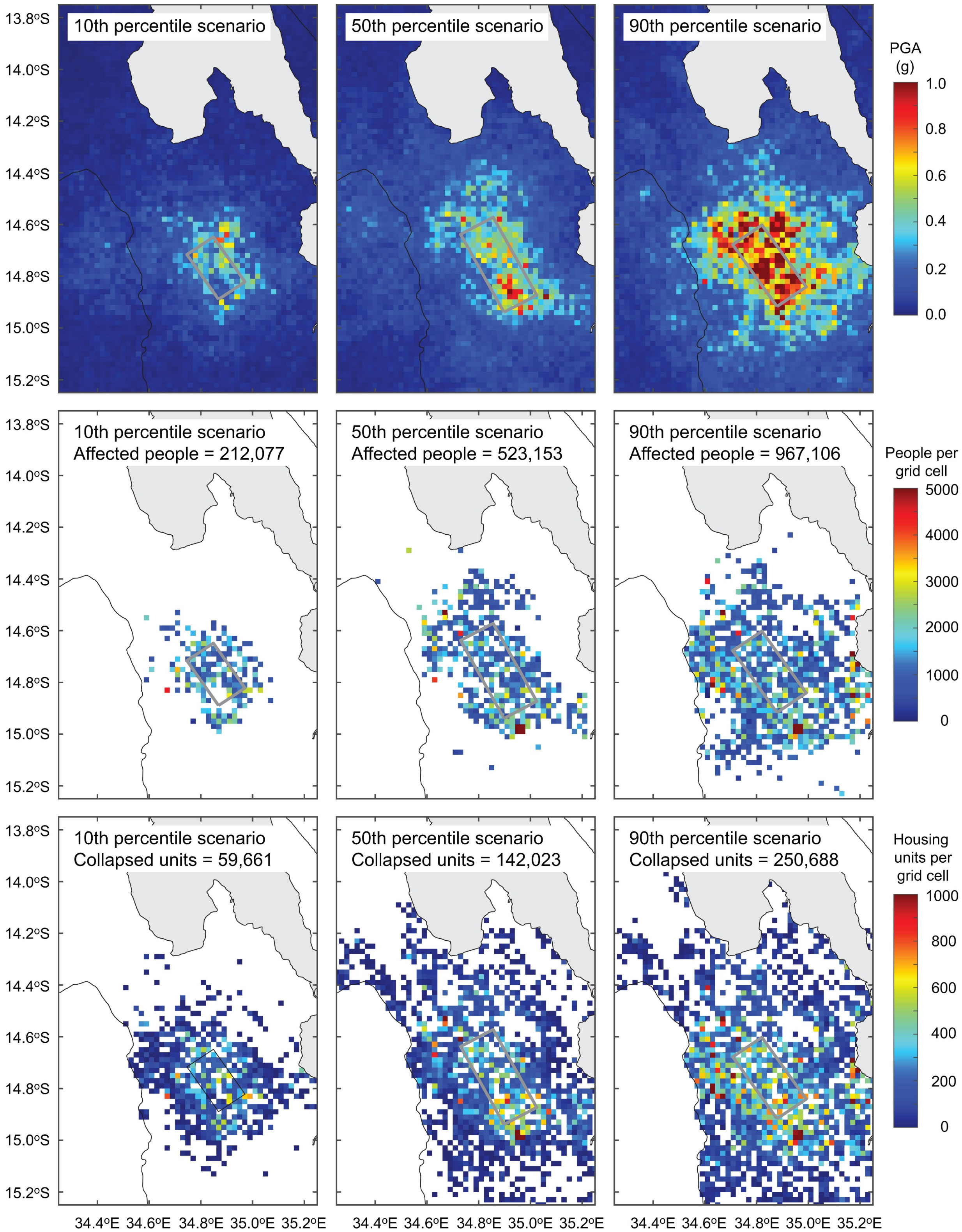


Figure 19. PGA shake map, affected population map for PGA > 0.2 g, and building collapse maps that correspond to the 10th, 50th, and 90th percentile of the building collapse risk levels for rupture scenario 4 and fragility case 1.

6. Conclusions

This study developed the regional seismic risk model for central-southern Malawi based on local information on earthquake rupture sources, exposure information, building characteristics, and seismic vulnerability, and carried out scenario-based earthquake risk assessments for the Bilila-Mtakataka Fault. The developed risk model incorporated uncertainty in earthquake ruptures via stochastic source modelling, latest Malawi exposure data from the 2018 census, building survey data and experimental tests of local construction materials, and seismic fragility functions for unreinforced masonry buildings in Malawi. The considerations of improved local information and Malawi-specific models led to more accurate evaluations of regional seismic hazard and risk and more comprehensive quantification of epistemic uncertainty associated with the model components.

The results of the earthquake risk assessments led to the following observations.

- The retrospective seismic risk assessment for the 2009 Karonga sequence (focusing upon the December 8th, 2009 event) indicates that the observed seismic damage during the 2009 sequence still falls within the predicted range of the building collapse counts but corresponds to the lower end of the distribution. This difference can be attributed to the ground-motion variability, potential bias in seismic fragility functions, uncertainty in the observed seismic damage, and combinations of the above-mentioned factors.
- The effects of seismic hazard characterisation on the regional seismic risk assessments, which can be attributed to the uncertainty in the rupture geometry and location as well as ground-motion variability, are significant. For instance, for the whole rupture case with the same overall earthquake magnitude, the number of collapsed housing units can differ by a factor of approximately 3 between the 10th and 90th percentile regional risk scenarios.
- The uncertainty associated with the rupture patterns (whole rupture versus segmented rupture) leads to different earthquake sizes and thus results in significant differences in regional seismic risk estimates. For the considered rupture cases, the median estimates of the housing unit collapses differ by a factor of approximately 6, when the whole (M_w7.6) and the

segmented (M_w6.0) cases are compared. The results are influenced not only by earthquake rupture characteristics but also by their proximity to population centres in the region.

- The consideration of local population-building data is of paramount importance. The global-quality exposure data and vulnerability models, in comparison with the local/regional data and models, can result in significant biases in the risk estimates. For the cases of the global versus Malawi-specific seismic vulnerability functions, more than 20% underestimation of the median collapse count was observed. On the other hand, when the local building data and laboratory test results are taken into account in developing seismic fragility functions, consistent risk estimates that are within approximately 3% differences were obtained. Moreover, the use of global-quality exposure data alone can lead to more than 30% underestimation of the regional seismic risk for the rural areas of central-southern Malawi. It is important to recognise that significant overestimation is possible when more urban areas of central-southern Malawi are focused upon.

The developed regional seismic risk model for central-southern Malawi has limitations. The risk assessments do not include the effects on building foundation due to geohazards (e.g., liquefaction and landslide), which are expected to be significant, especially for sites near Lake Malawi and along Shire River and its tributaries. From the methodological viewpoint, the scenario-based earthquake hazard and risk assessments do not account for the occurrence probability of the earthquake ruptures. Consequently, recurrence periods cannot be uniquely assigned to the regional seismic risk estimates that were obtained in this study. When a full PSHA model with fault sources becomes ready for use, the developed seismic risk model can be applied to conduct a comprehensive seismic risk assessment for Malawi.

In conclusions, this study highlighted the importance of incorporating local information on seismic hazard scenarios, exposure, and vulnerability. The effects of the model uncertainty on the regional earthquake risk can be quantified and visualised through probability distributions of regional risk metrics and integrated critical hazard-risk maps. Subsequently, government officers, professional engineers, and community leaders who are involved with disaster risk reduction processes can evaluate/compare available mitigation options on a uniform basis. Ultimately, they can make recommendations and decisions in shaping disaster mitigation policies and in enhancing seismic preparedness and resilience. Importantly, local situations (e.g., building typologies and demographic profiles) and local resources (e.g., material types and construction techniques) need to be reflected in choosing mitigation options which are cost effective, sustainable, and culturally acceptable in the country and the region.

Acknowledgements

The work was conducted as part of the PREPARE and SAFER-PREPARED projects funded by the Engineering and Physical Sciences Research Council (EP/P028233/1 and EP/T015462/1). K.G. is funded by the Canada Research Chair program (950-232015) and the NSERC Discovery Grant (RGPIN-2019-05898). The authors are grateful to Kingsley Chihamie Manda of the National Statistical Office of Malawi for providing the detailed 2018 census data.

Data Access Statement

All data source information for seismic hazard scenarios, exposure, and seismic fragility functions are included in the paper, except for the enumeration-level Malawi census data. The authors do not have right to make these data publicly available.

References

Akkar, S., Sandıkkaya, M.S., and Bommer, J.J. (2014). Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering*, **12**, 359–387, doi: 10.1007/s10518-013-9461-4.

Ambraseys, N.N., and Adams, R.D. (1991). Reappraisal of major African earthquakes, south of 20°N, 1900–1930. Nat Hazards, 4, 389–419, doi: 10.1007/BF00126646.

Biggs, J., Nissen, E., Craig, T., Jackson, J., and Robinson, D.P. (2010). Breaking up the hanging wall of a rift-border fault: the 2009 Karonga earthquakes, Malawi. *Geophysical Research Letters*, 37, L11305. doi: 10.1029/2010GL043179.

Boore, D.M., Stewart, J.P., Seyhan, E., and Atkinson, G.M. (2014). NGA-West 2 equations for predicting PGA, PGV, and 5%-damped PSA for shallow crustal earthquakes. *Earthquake Spectra*, 30, 1057–1085, doi: 10.1193/070113EQS184M.

Bureau TNM (2016). Safer House Construction Guidelines. Available at https://issuu.com/saferconstructionguidelines/docs/no-crocini.

Caprio, M., Tarigan, B., Worden, C.B., Wiemer, S., and Wald, D.J. (2015). Ground motion to intensity conversion equations (GMICEs): a global relationship and evaluation of regional dependency. *Bulletin of the Seismological Society of America*, 105, 1476–1490, doi: 10.1785/0120140286.

Chapola, L. (2015). The impacts of the 2009 Karonga earthquakes, Malawi. *Journal of Catholic University Malawi*, 1, 9–19.

Chapola, L., and Gondwe, J. (2016). Urban development in earthquake prone areas: lessons from 1989 Salima and 2009 Karonga earthquakes. *Journal of Catholic University Malawi Malawi*, 2, 15–26.

Craig, T. J., Jackson, J.A., Priestley, K., and Mckenzie, D. (2011). Earthquake distribution patterns in Africa: their relationship to variations in lithospheric and geological structure, and their rheological implications. *Geophysical Journal International*, 185, 403–434, doi: 10.1111/j.1365-246X.2011.04950.x.

D'Ayala, D., and Speranza, E. (2003). Definition of collapse mechanisms and seismic vulnerability of historic masonry buildings. *Earthquake Spectra*, 19, 479–509, doi: 10.1193/1.1599896.

D'Ayala, D., and Paganoni, S. (2011). Assessment and analysis of damage in L'Aquila historic city centre after 6th April 2009. *Bulletin of Earthquake Engineering*, 9, 81–104, doi: 10.1007/s10518-010-9224-4.

Flannery, J.W., and Rosendahl, B.R. (1990). The seismic stratigraphy of Lake Malawi, Africa: implications for interpreting geological processes in lacustrine rifts. *Journal of African Earth Sciences (and the Middle East)*, 10, 519–548, doi: 10.1016/0899-5362(90)90104-M.

Giordano, N., De Luca, F., and Sextos, A. (2020). Out-of-plane closed-form solution for the seismic assessment of unreinforced masonry schools in Nepal. *Engineering Structures*, 203, 109548, doi: 10.1016/j.engstruct.2019.109548.

Giordano, N., De Risi, R., Voyagaki, E., Kloukinas, P., Novelli, V., Kafodya, I., Ngoma, I., Goda, K., and Macdonald, J. (2021). Seismic fragility models for typical non-engineered URM residential buildings in Malawi. *Structures*, 32, 2266–2278, doi: 10.1016/j.istruc.2021.03.118.

Government of Malawi (2015). National Disaster Risk Management Policy. Available at: http://www.ifrc.org/docs/IDRL/43755_malawidrmpolicy2015.pdf.

Goda, K. and Atkinson, G.M. (2010). Intra-event spatial correlation of ground-motion parameters using SK-net data. *Bulletin of the Seismological Society of America*, 100, 3055–3067, doi: 10.1785/0120100031.

Goda, K., Gibson, E.D., Smith, H.R., Biggs J., and Hodge, M. (2016). Seismic risk assessment of urban and rural settlements around Lake Malawi. *Frontiers in Built Environment*, 2, 30, doi: 10.3389/fbuil.2016.00030.

Goda, K. (2017). Probabilistic characterization of seismic ground deformation due to tectonic fault movements. *Soil Dynamics and Earthquake Engineering*, 100, 316–329, doi: 10.1016/j.soildyn.2017.05.039.

Goda, K., Zhang, L., and Tesfamariam, S. (2020). Portfolio seismic loss estimation and risk-based critical scenarios for residential wooden houses in Victoria, British Columbia, Canada. *Risk Analysis*, doi: 10.1111/risa.13593.

Gupta, H.K., and Malomo, S. (1995). The Malawi earthquake of March 10, 1989: report of field survey. *Seismological Research Letters*, 66, 20–27, doi: 10.1785/gssrl.66.1.20.

Hodge, M., Biggs, J., Goda, K., and Aspinall, W.P. (2015). Assessing infrequent large earthquakes using geomorphology and geodesy in the Malawi Rift. *Natural Hazards*, 76, 1781–1806, doi: 10.1007/s11069-014-1572-y.

Hodge, M., Fagereng, A., Biggs, J., and Mdala, H.S. (2018). Controls on early-rift geometry: new perspectives from the Bilila-Mtakataka fault, Malawi. *Geophysical Research Letters*, 45, 3896–3905, doi: 10.1029/2018GL077343.

Jackson, J., and Blenkinsop, T. (1997). The Bilila-Mtakataka fault in Malawi: an active, 100-km long, normal fault segment in thick seismogenic crust. *Tectonics*, 16, 137–150, doi: 10.1029/96TC02494.

Jaiswal, K.S., and Wald, D.J. (2008). Creating a global building inventory for earthquake loss assessment and risk management (Open-File Report 2008-1160). US Geological Survey, 103 p.

Jaiswal, K.S., Wald, D.J., and D'Ayala, D. (2011). Developing empirical collapse fragility functions for global building types. *Earthquake Spectra*, 27, 775–795, doi: 10.1193/1.3606398.

Kloukinas, P., Kafodya, I., Ngoma, I., Novelli, V., Macdonald, J., and Goda, K. (2019). Strength of materials and masonry structures in Malawi. *Proceedings of the 7th International Conference on Structural Engineering, Mechanics and Computation (SEMC)*, Cape Town, South Africa, Paper 233.

Kloukinas P., Kafodya I., Ngoma I., Novelli V., Macdonald J., and Goda K. (2020). A building classification scheme of housing stock in Malawi for earthquake risk assessment. *Journal of Housing and the Built Environment*, 35, 507–537, doi: 10.1007/s10901-019-09697-5.

Kushe, J., Manda, M., Mdala, H. and Wanda, E. (2017). The earthquake/seismic risk, vulnerability and capacity profile for Karonga town. *African Journal of Environmental Science and Technology*, 11, 19–32, doi: 10.5897/AJEST2016.2217.

Lagomarsino, S. (2015). Seismic assessment of rocking masonry structures. *Bulletin of Earthquake Engineering*, 13, 97–128, doi: 10.1007/s10518-014-9609-x.

Malawi Bureau of Standards Board (2014). The structural use of masonry – Code of practice, Part 1: Unreinforced masonry walling, MS791-1.

Mitchell-Wallace, K., Jones, M., Hillier, J., and Foote, M. (2017). Natural catastrophe risk management and modelling: a practitioner's guide, Wiley-Blackwell, 536 p.

National Statistical Office (NSO) and ICF (2017). Malawi Demographic and Health Survey 2015-16. Zomba, Malawi, and Rockville, Maryland, USA. NSO and ICF.

National Statistical Office of Malawi (2008). 2008 population and housing census. Available at: http://www.nsomalawi.mw/images/stories/data_on_line/demography/census_2008/Main%20Report/Census%20Main%20Report.pdf.

National Statistical Office of Malawi (2018). 2018 population and housing census. Available at: http://www.nsomalawi.mw/index.php%3Foption%3Dcom_content%26view%3Darticle%26id%3D226:2018-malawi-population-and-housing-census%26catid%E2%80%89%3D%E2%80%898:reports%26Itemid%E2%80%89%3D%E2%80%896.

Novelli, V.I., De Risi, R., Ngoma, I., Kafodya, I., Kloukinas, P., Macdonald, J., and Goda, K. (2021). Fragility curves for non-engineered masonry buildings in developing countries derived from real data based on structural surveys and laboratory tests. *Soft Computing*, https://doi.org/10.1007/s00500-021-05603-w

Poggi, V., Durrheim, R., Tuluka, G.M., Weatherill, G., Gee, R., Pagani, M., Nyblade, A., and Delvaux, D. (2017). Assessing seismic hazard of the East African Rift: a pilot study from GEM and AfricaArray. *Bulletin of Earthquake Engineering*, 15, 4499–4529, doi: 10.1007/s10518-017-0152-4.

Republic of Malawi (2019). Malawi National Building Regulations 2019 Volume 1, 363 p.

Thingbaijam, K.K.S., Mai, P.M., and Goda, K. (2017). New empirical earthquake-source scaling laws. *Bulletin of the Seismological Society of America*, 107, 2225–2246, doi: 10.1785/0120170017.

Tomazevic, M. (2007). Damage as a measure for earthquake-resistant design of masonry structures: Slovenian experience. *Canadian Journal of Civil Engineering*, 34, 1403–1412, doi: 10.1139/L07-128.

UN-Habitat (2010). Malawi: urban housing sector profile. Available at: https://unhabitat.org/books/malawi-urban-housing-sector-profile/.

United States Geological Survey (2009). Significant Earthquakes of the World. Available at http://earthquake.usgs.gov/earthquakes/eqarchives/significant/sig_2009.php.

Voyagaki, E., Kloukinas, P., Novelli, V., De Risi, R., Kafodya, I., Ngoma, I., Goda, K., and Macdonald, J. (2020). Masonry panel testing in Malawi. *Proceedings of the 17th World Conference on Earthquake Engineering*, Sendai, Japan, Paper C003152.

Wald, D.J., and Allen, T.I. (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, 97, 1379–1395, doi: 10.1785/0120060267.

Wald, D. J., Quitoriano, V., Heaton, T. H., and Kanamori, H. (1999). Relationships between peak ground acceleration, peak ground velocity and modified Mercalli intensity in California. *Earthquake Spectra*, 15, 557–564.

Wedmore, L. N. J., Biggs, J., Floyd, M., Fagereng, A., Mdala, H., Chindandali, P., Williams, J.N., and Mphepo, F. (2021). Geodetic constraints on cratonic microplates and broad strain during rifting of thick southern African lithosphere. *Geophysical Research Letters*, 48, e2021GL093785, doi: 10.1029/2021GL093785.

Wheeler, W.H., and Karson, J.A. (1989). Structure and kinematics of the Livingstone Mountains border fault zone, Nyasa (Malawi) Rift, southwestern Tanzania. *Journal of African Earth Sciences (and the Middle East)*, 8, 393–413, doi: 10.1016/S0899-5362(89)80034-X.

Williams, J.N., Mdala, H., Fagereng, A., Wedmore, L.N.J., Biggs, J., Dulanya, Z., Chindandali, P., and Mphepo, F. (2021). A systems-based approach to parameterise seismic hazard in regions with little historical or instrumental seismicity: active fault and seismogenic source databases for southern Malawi. *Solid Earth*, 12, 187–217, doi: 10.5194/se-12-187-2021.

Woo, G. (2011). Calculating catastrophe. World Scientific, 368 p.