Bull Earthquake Eng (2016) [x:x–x]
DOI [xxx/xxx]

ORIGINAL RESEARCH PAPER

Probabilistic Seismic Hazard Assessment of Sub-Saharan Africa

Valerio Poggi[1] • Raymond Durrheim[2] • Georges Mavonga Tuluka[3] • Graeme Weatherill[1] • Julio Garcia[1] • Marco Pagani[1] • Andrew Nyblade[4] • Damien Delvaux[5]

Received: [D M Y] / Accepted: [D M Y] / Published online: [D M Y]

Abstract The East African Rift System (EARS) is the major active tectonic feature of the Sub-Saharan Africa (SSA) region. Although the seismicity level of such a divergent plate boundary can be described as moderate, several damaging earthquakes have been reported in historical times, and the seismic risk is exacerbated by the high vulnerability of the local buildings and structures. Formulation and enforcement of national seismic codes is therefore an essential future risk mitigation strategy. Nonetheless, a reliable risk assessment cannot be done without the calibration of an updated seismic hazard model for the region. A major limitation affecting the assessment of seismic hazard in Sub-Saharan Africa is the lack of basic information needed to construct source and ground motion models. The historical earthquake record is sparse, with significant variation in completeness over time across different regions. The instrumental catalogue is complete down to sufficient magnitude only for a relatively short time span. In addition, mapping of seimogenically active faults is still an on-going task, and few faults in the region are sufficiently constrained as to allow them to be directly represented within the seismic hazard model. Recent studies have identified major seismogenic lineaments, but there is substantial lack of kinematic information for intermediate-to-small scale tectonic features, information that is essential for the proper calibration of earthquake recurrence models.

✉ Valerio Poggi
e-mail: valerio.poggi@globalquakemodel.org

1 Global Earthquake Model Foundation (GEM), Pavia, Italy
2 University of the Witwatersrand, Johannesburg, South Africa
3 Goma Volcanic Observatory, DR Congo
4 Penn State University, USA
5 Royal Museum for central Africa, Tervuren, Belgium

In this study, we propose a new probabilistic seismic hazard model for the Sub-Saharan Africa region. The model has been calibrated on the most recent and up to date information available from scientific literature, global bulletins and local earthquake catalogues, such as those from the partner *AfricaArray* project. In this report we describe in detail all working assumptions, main processing steps, data analyses and interpretations used for the model setup.

Keywords Probabilistic seismic hazard analysis · GMPEs · Uncertainty analysis · Earthquake engineering · Logic-tree

1 Introduction

Earthquakes pose a significant risk in many regions of the Sub-Saharan Africa (SSA), more particularly along the tectonically active East African Rift System (EARS). Further away from this rift system, the remainder of SSA is largely considered a stable intra-plate region characterized by a relatively low rate of seismicity. Nonetheless, several large earthquakes have been reported in historical times. While most earthquakes occur near plate boundaries (inter-plate seismicity) along the EARS, it must be noted that a damaging earthquake can occur anywhere, especially as cities grow and many buildings are constructed without taking potential ground shaking into account. Even a moderate-sized events can prove disastrous should it occur near a city with many vulnerable buildings, as happened when an Mw 5.7 earthquake struck Agadir, Morocco in 1960, causing some 15,000 deaths.

Damaging earthquakes with M > 6 occur almost annually in the East African Rift, and five M > 7 earthquakes have occurred in eastern Africa since 1900. The largest known event in the region is the 13 December 1910 Ms 7.4 Rukwa (Tanzania) event that badly cracked all European-style houses in towns on the eastern shore of Lake Tanganyika (Midzi and Manzunzu, 2014; Ambraseys, 1991a; Ambrasys and Adams, 1991). A Ms 6.9 earthquake that occurred on 6 January 1928 in the Subakia Valley (part of the Kenya Rift, some 200 km northwest of Nairobi) produced a 38 km long surface rupture with a maximum throw of 2.4m and destroyed, or damaged beyond repair, all European-style houses within 15 km of the rupture, fortunately without causing casualties (Ambraseys, 1991b).

During the last decade there have been several other events that have caused loss of life (Durrheim, in press). On 5 December 2005 an Mw 6.8 event caused several deaths and damaged school buildings and hundreds of dwellings in the Democratic Republic of Congo (DRC) and western Tanzania. The 22 February 2006 Mozambican Mw 7 earthquake was one of the largest ever recorded in southern Africa, producing a surface rupture with a displacement of more than 1 m. Shaking was felt as far away as Zimbabwe and South Africa. Four people were killed, 27 injured, and at least 160 buildings damaged. On 3 February 2008 an Mw 5.9 earthquake struck the Lake Kivu region of the DRC and neighbouring Rwanda. The event was located approximately 20 km north of Bukavu (population 250,000) and can be regarded as a “near miss”. A second earthquake followed the main shock 3 and 1⁄2 hours later. Numerous buildings collapsed or suffered significant structural damage, trapping many people under rubble. At least 40 people died and more than 400 were injured.

While these events caused relatively small losses, the population of the region has increased enormously over the last century and increasingly urbanized; trends that are expected to continue well into the mid 21st century. Building methods have changed from wattle and daub or timber with grass roofs, which have a large inherent resistance to earthquake shaking, to European-style unreinforced masonry constructions, which are far more vulnerable to shaking (references). The occurrence of similar events close to a town would likely cause serious human and economic losses today.

The mitigation of earthquake risk in Africa requires coordinated action on several fronts. Firstly, seismic hazard assessments should be improved by maintaining and expanding seismic monitoring networks, supplementing historical and paleoseismic catalogues, and mapping active faults and the near-surface. Secondly, building codes should be formulated and enforced, and vulnerable existing buildings and infrastructure reinforced to prevent serious damage or collapse when subjected to strong shaking. Lastly, disaster management agencies, emergency first responders, and the general public should be trained to act effectively and sensibly during an earthquake, and equipped to deal with the aftermath. National efforts to assess seismic hazard are reviewed by Worku (2014).

The Global Earthquake Model (GEM) Foundation was launched in 2009 with the vision of promoting the collaborative development of tools and models for earthquake hazard and risk assessment. In this paper, we illustrate part of the activities completed within USAID-funded pilot project, where we seek to gain knowledge and build capacity to mitigate and reduce seismic risk in regions affected by earthquakes in sub-Saharan Africa associated with the East African Rift System by combining the expertise, technologies and infrastructure developed by partner *AfricaArray* project. Within this framework, a new regional probabilistic seismic hazard model based on distributed seismicity is proposed and discussed.

2 Tectonic of the East African Rift System

The African continent is a palimpsest recording a lengthy tectonic history, and the EARS is superimposed on structures formed during earlier tectonic episodes. On a broad scale, much of it can be explained by plate tectonics and the Wilson cycle, for example to amalgamation and dispersal of Gondwana. However, there are other phenomena, such as the rise of the African superswell, that are not well understood. The East African Rift System (EARS) stretches quasi-continuously from the Afar depression in northern Ethiopia to the Southwest Indian Ocean Ridge (SWIR) at the junction with the Antarctic plate. It is the southern branch of three rifts that radiate from a triple junction. The north-western rift lies along the axis of the Red Sea, while the north-eastern rift bisects the Gulf of Aden reaching an oceanic triple junction where meets the Indian Ocean Ridge.

The EARS includes the world’s youngest continental flood basalt province (Ethiopia) and is superimposed on a broad region of high topographic elevation (the 1000 m high eastern and southern African plateaus). This high elevation region and its offshore extension in the south-eastern Atlantic define the “African Superswell” (Nyblade and Robinson, 1994), which lies on average 500 m higher than the global topographic mean. The analysis of long-wavelength gravity and topographic relief over Africa shows that more than half of this anomalous topography is dynamically supported by convective mantle upwelling associated with a large, slow shear wave seismic velocity mantle anomaly, the African superplume (Lithgow-Bertelloni and Silveri, 1998; Ritsema et al., 1998; Gurnis et al., 2000).

The initiation of Cenozoic rifting is estimated to start in the mid-Tertiary with the onset of volcanism in the Turkana Rift (Furman et al., 2006) followed by uplift and flood basalts in Ethiopia (Pik et al., 2003). The process was followed by extension in the Main Ethiopian Rift and the western and eastern (Kenya) branches (Roberts et al., 2012), and further south in the Malawi Rift (Lyons et al., 2011).

3 Methodology

Seismic Hazard is evaluated for the regions surrounding the EARS by developing a probabilistic model based on distributed seismicity. The choice of a distributed seismicity model was mostly driven by current data, including local earthquake catalogues, faults, focal mechanisms and strain information. Unfortunately, available data appeared not sufficient for the implementation of alternative probabilistic models.

For a given site, the distributed seismicity approach determines the probabilities of exceeding, at least once in a given time span, a set of ground motion levels of engineering interest generated by a number of seismically and tectonically homogenous earthquake source zones. In its simplest representation, each source is considered independent from others and the earthquake rupture process within zones is assumed to follow a *Poisson* process. More comprehensive descriptions of Probabilistic Seismic Hazard Analysis (PSHA) can be found for example in Field (2003), McGuire (2004) and USNRC (2012). Calculation of seismic hazard is made through the use of the *OpenQuake* engine, an open source seismic hazard and risk calculation software developed, maintained and distributed by the Global Earthquake Model.

4 The SSA-GEM Earthquake Catalogue

The starting point for PSHA analysis is the definition of the seismicity characteristics, in terms of both the long-term recurrence as well as the seismotectonic properties (e.g. style of faulting, depth distribution etc.), for the study area. This can be done in multiple ways, but the basic - and probably the most common - approach is in the use of an earthquake catalogue. For the purposes of constraining earthquake recurrence, it is critical to identify which portions of the catalogue can be considered - as much as possible – to be a complete record of all earthquake events indirectly reported (the historical and macroseismic component) or directly recorded (the instrumental component) on a specific area and over a certain time span.

If several catalogues are available for a given study area, information (location solutions, reported time, intensity scale) can be quite heterogeneous and some objective criteria for selection, merging and homogenisation are needed. This is usually the case when neighbouring agencies are reporting same events but with different magnitude types. Same issue is affecting source solutions, for instance when different earthquake phases, processing algorithm or base model assumptions (e.g. earth velocity structure) are used.

GEM has recently developed a set of open-source tools that helps scientists going though the catalogue harmonisation process. In this study we make use of these tools (aka GEM Catalogue Toolkit, Weatherill et al. 2016) to produce a state-of-art earthquake catalogue for Sub-Saharan Africa with homogenous magnitude representation (Mw). Such catalogue (hereinafter SSA-GEM) is obtained by augmenting available global catalogues (e.g. ISC-Reviewed, ISC-GEM, GCMT) with information from local agencies and regional projects, particularly from the *AfricaArray* framework (missing reference). In the following we describe in detail the necessary steps, main assumptions and choices we faced to set up the SSA-GEM catalogue.

4.1 Source Data

4.1.1 ISC Reviewed Bulletin

The manually reviewed bulletin from the International Seismological Centre (ISC) was used as one of the primary sources of information for the earthquake catalogue. The ISC bulletin covers a period ranging from the beginning of the 20th century to present day. In our geographic area selection (-40° to 20° North, 10° to 60° East) it spans the period 1904-2013, for a total of 26322 events from 89 international and national (local) agencies. Magnitude scale representation is however not homogenous and varies between agencies and considered time period.

4.1.2 ISC-GEM Catalogue

The ISC-GEM global instrumental earthquake catalogue (Storchak et al., 2013, 2015) is a refined version of the ISC bulletin, which improves the accuracy on magnitude and location solutions for large global events (Mw > 5.5) in the period 1900-2012. Events reported in the ISC-GEM catalogue are considered as reference events, which have priority over other estimates from global bulletins. Intensity is homogeneously represented by using moment magnitude (Mw) from globally calibrated magnitude conversion relations. ISC-GEM catalogue is presently in its version 3, which is the one used in this study. 285 events (out of 24375) are available for the selected study region.

4.1.3 Harvard/GCMT Bulletin

The Global Centroid Moment Tensor catalogue (GCMT, Ekström et al., 2012) is a collection of moment tensor solutions for earthquakes with Mw > 5. The catalogue covers the period 1976 to present, with a total of more than 40000 global events, 614 of which are of interest for this study. Note that within ISC bulletin, the Global Centroid Moment Tensor catalogue is indicated with two separated agency labels, HRVD and GCMT, indicating the migration of the project from Harvard (Harvard CMT Project) to the Lamont-Doherty Earth Observatory (LDEO) of the Columbia University in 2006. Moment tensor solutions from the GCMT are considered as reference for the calibration of magnitude conversion relations used in this study.

4.1.4 GEM Historical Earthquake Archive

The GEM Historical Earthquake Catalogue (GEH) is a global collection of reviewed historical records consisting of 825 events (M > 7) covering the period

1000-1903 (pre instrumental period). Eight earthquakes from the GEH catalogue are relevant for the study region. This is likely due to the lack of historical records in sun-Saharan and central Africa, and poses the problem of completeness of the regional earthquake record for large magnitudes, which may consequently bias the calibration of annual occurrence rates for these events.

4.1.5 AfricaArray regional earthquake catalogues

We extended the Earthquake record by integration of three local catalogues developed within the frame of the *AfricaArray* project (reference). These catalogues are the result of regional earthquake monitoring performed with temporary and permanent seismic network installations. The available *AfricaArray* catalogues are:

I. The Tanzanian Broadband Seismic Experiment (TZB), with 2218 events covering the period 1994-1995 and Ms magnitude between 1.43 and 4.42;
II. The Ethiopian Plateau Catalogue (ETP), with 253 events covering the period 2001-2002 and with Ms magnitude between 1.75 ad 4.05;
III. The AfricaArray Eastern Africa Seismic experiment (AAE), with 1023 events in the period 2009-2011 and Ms magnitude range 1.28-4.04.

Although these catalogues extend the record to very low magnitudes, their primary application within the present hazard study was for the local definition of seismicity distribution patterns in order to elucidate potentially seismogenic structures within the rift system and the surrounding regions. Subsequently, these are used to improve the design of a new area source model for Sub-Saharan Africa.

4.2 Location Solution

In many applications, preference for earthquake location solution should be given to local agencies, while solutions from global agencies and teleseismic events should be alternatively used in those cases where local agencies are not available on the territory (e.g. not yet established) or where large solution uncertainty exists, e.g. due to insufficient station coverage. For the case of Sub-Saharan Africa, although solutions from several local agencies are made available through ISC bulletin, there is general lack information regarding network operation (particularly before 1980) and metadata - including quality of the solutions - which makes the use of their locations often questionable. Nonetheless, events recorded teleseismically are unlikely to be affected by changes in station location or operation over time, with a consequent decreased bias in the solution error for different periods of the catalogue. For these reasons, solutions from global agencies have been preferred, while restricting the use of local agencies only to those cases where no other information was made available.

By mapping the activity period of the different seismological agencies over time, we identify five main time intervals with a different scheme of agency prioritisation (see summary **Table 1**).

Table 1 Prioritisation of agencies for preferred location solution. Selection is done differently for separated time periods, accounting for network operation and reliability of the estimate. We refer to ISC website (http://www.isc.ac.uk/iscbulletin/agencies, last access August 2016) for acronyms not otherwise described in the article.

Period	Agency Priority List
1000-1900	GEH
1901-1959	ISC-GEM, ISC, ISS, GUTE, GEH
1960-1964	ISC-GEM, EHB, ISC, ISS, GEH
1965-1980	EHB, ISC, NEIC, IDC, GCMT, HRVD, GCMT-NDK, BUL, PRE, LSZ, TAN, CNG, GEH
1981-2015	EHB, ISC, NEIC, IDC, GCMT, HRVD, GCMT-NDK, AAE, ETP, TZB, PRE, LSZ, NAI, TAN, CNG, EAF, GEH

4.3 Magnitude Homogenisation

An unbiased seismicity analysis requires that the seismic record is represented homogeneously in terms of the magnitude scale, to avoid inconsistencies due to the different processing schemes used within different scales and the manifestation of saturation effects. Among the several scales that can possibly be used as reference, the most natural choice is moment magnitude (Mw), due to its direct connection to earthquake size and energy, and the absence of a saturation at high magnitudes. However, events with a native (directly obtained from data) estimate of Mw are limited, and very often a conversion from other scales is necessary.

Calibration of regional conversion rules from local datasets is generally advisable; however, it can be limited by availability of events with multiple magnitude scale representations. Alternatively, a two- (or three-) step conversion with an intermediate dummy intensity measure (IM) of larger availability can be used, with the drawback of the progressive accumulation of uncertainty at each conversion step. If no calibration data are available at all, globally calibrated conversion rules can still be applied.

Table 2 List of magnitude agencies and corresponding Mw conversion rules. Agencies are sorted according to decreasing priority for the catalogue harmonisation.

Agency	Mw Conversion Rule	Rage	Reference
ISC	$0.616M_S+2.369$	$M_S < 6$	Weatherill et al. 2016
	$0.994M_S+0.1$	$M_S > 6$	
	$1.084m_b-0.142$	$m_b < 6.5$	
NEIC	$0.723M_S+1.798$	$M_S < 6.5$	
	$1.005M_S-0.026$	$M_S > 6.5$	
	$1.159m_b-0.659$	$m_b < 6.5$	
PRE	M_L	$M_L < 6$	Assumed 1:1 scaling and arbitrary uncertainty (0.3)
BUL	M_{blg}	$M_{blg} < 6$	
TZB, ETP, AAE	$1.02+0.47M_L+0.05M_L^2$	$M_L < 5$	Edwards et al., 2010
PAS	$0.616(M_S-0.2)+2.369$	$M_S > 6$	ISC-M_S corrected (as suggested in Engdahl and Villasenor, 2002)
	$0.994(M_S-0.2)+0.1$	$M_S < 6$	

For the definition of ad-hoc magnitude conversion rules, we used in this study the functionalities offered by the GEM catalogue toolkit (Weatherill et al., 2016), which allows for the exploration and statistical analysis of local, regional and global datasets to build statistical regression models for the IM conversion. In the SSA region, unfortunately, we experienced a substantial lack of calibration data to implement local Mw conversion rules and in several cases we had to rely on globally calibrated relations (see Table 2)

4.4 Duplicate Findings

When merging different earthquake catalogues, in the absence of a defined one issue is the identification of duplicate events. To face this problem, events falling within a window of prescribed spatial and temporal width are assumed representing the same earthquake. Best results have been obtained with a window of 0.5° and 120s. These values appear sufficient to capture relative uncertainty in earthquake solution between agencies, which is particularly relevant for teleseismic events. The use of larger values had led to erroneous results, by misinterpreting aftershocks sequences. After catalogue merging, previously defined priority rules for magnitude and location agency selection are applied and the final catalogue released (see **Fig. 1**)

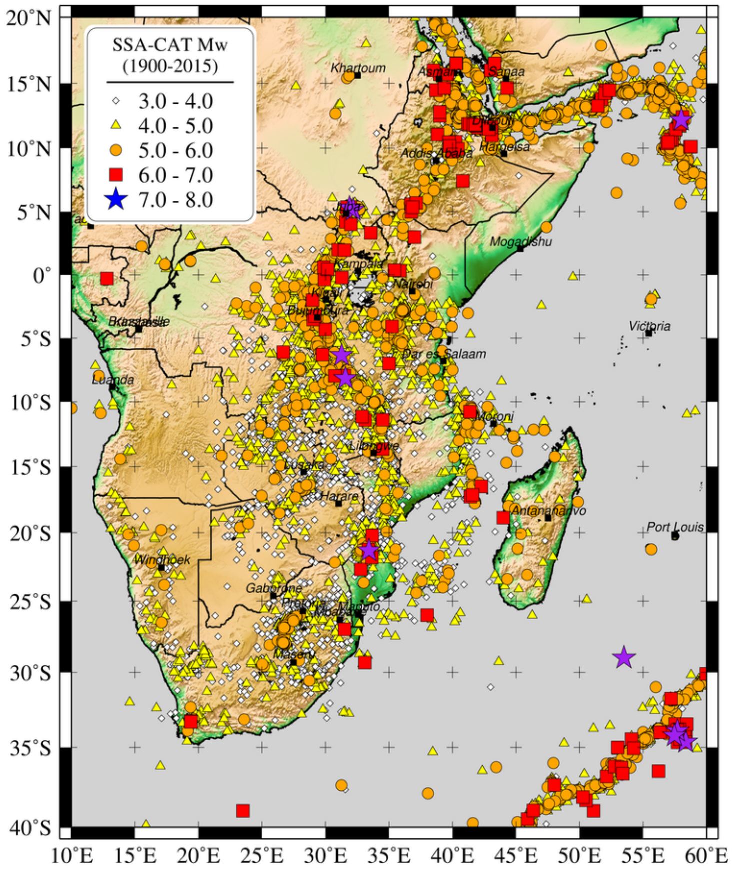
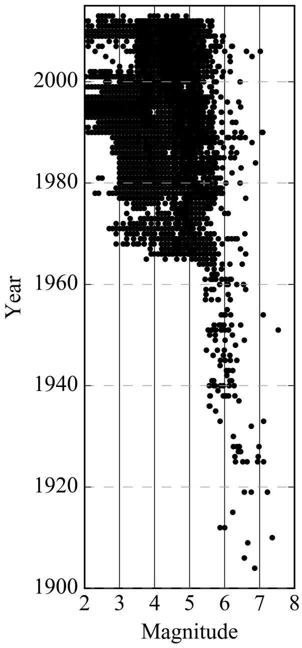


Fig. 1 Left - Distribution of events from the homogenised SSA earthquake catalogue. Bottom: Magnitude over time distribution.



4.5 Catalogue Declustering

A widespread assumption in standard PSHA is that earthquake occurrence rates are independent of the observation time and that their probability distribution is that of a Poisson process. However, earthquake catalogues are naturally affected by the presence of groups of correlated events (clusters), such as fore-/after-shocks and seismic swarms, which are highly dependent in space and time.

In order to estimate poissonian seismicity rates, those dependent events have to be removed by filtering the catalogue prior to the calibration of any occurrence relationship. Such procedure is called catalogue declustering and several algorithms have been proposed in literature to face this issue (see van Stiphout et al., 2012 for a review). Among others, one of the most popular is from Gardner and Knopoff (1974), due to its conceptual and computational simplicity. The algorithm isolates and removes dependent events from a sorted catalogue by virtue of a fixed time-distance window centred on each (assumed) earthquake main shock and proportional to its magnitude. Although several window variants exist (see Uhrhammer, 1986 or van Stiphout et al., 2012), we use the original magnitude-scaling relation of Gardner and Knopoff (1974). Declustered GEM-SSA catalogue consists of 7259 events out of the original 29803 in the magnitude range $3 \le Mw \le 7.53$.

5 Seismic Source Zonation

The proposed seismic hazard model for Sub-Saharan Africa is based on distributed seismicity, and requires the discretisation of the study area in source zones of supposedly uniform temporal and spatial earthquake occurrence. Such approach is commonly used when observed seismicity cannot be reliably linked to any known (or inferred) geologic structure, which is often the case of low seismicity regions. Main advantage of using area source zones (ASZ) stays in their flexibility as to how to define the properties of seismogenesis within a region, although selection criteria might be highly subjective and often debatable between experts.

For the development of area source model we followed a mixed approach, which accounts for both observed seismicity and the geological/tectonic characteristics of the study region. Such approach closely descends from the methodology advocated by Vilanova (2014), which consists in the definition of a set of objective criteria for ASZ boundary delineation. Seismicity constraints have been obtained from the analysis (completeness, occurrence rates) of the ad-hoc developed SSA-GEM Earthquake catalogue and it will be discussed in more detail in the next section. Aside, tectonic information was derived mostly from scientific literature and by integration of available datasets.

The current area source model consists of a total of 19 zones distributed over 6 main tectonic groups (**Table 3**, **Fig. 2**), which we assume of homogenous rheological and mechanical behaviour with respect to the underlying crustal geology. The definition of these groups is essential for the regional calibration of b-values. Within four zones (7, 11, 12, 13), we further define sub-regions of larger observed seismicity. We assume these layers (marked with .1) to inherit all the basic seismotectonic features of the containing (background) zone, but with occurrence rates readjusted to match and explain the irregular spatial distribution of local seismicity. In the following we describe in detail the main seismotectonic characteristics of each group.

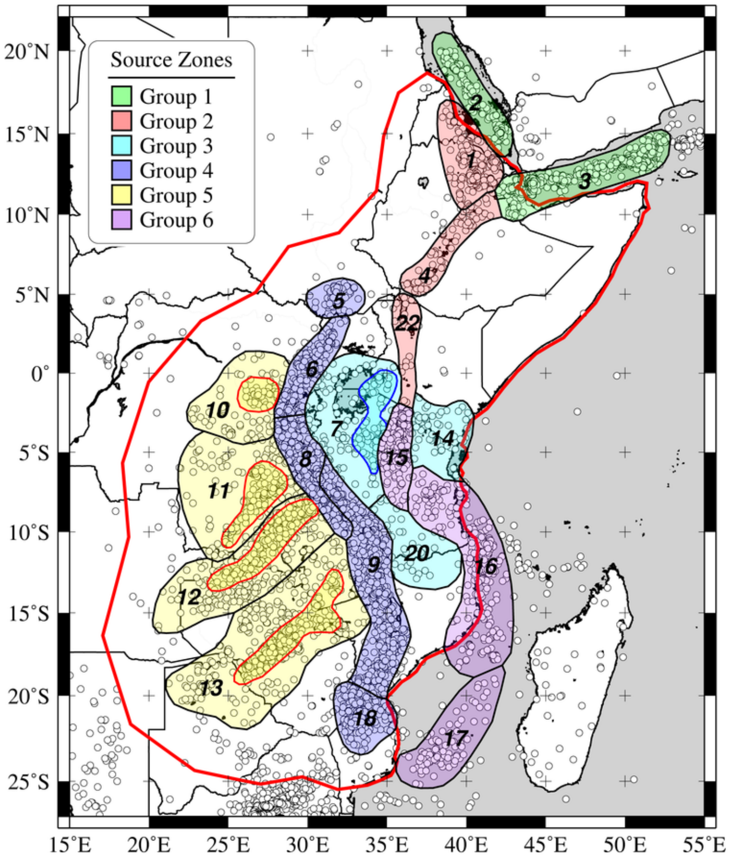


Fig. 2 Source zonation model used in this study (see **Table 3** for details). Area sources belonging to same tectonic group are represented with unique colour. Calculation area is marked with red solid line.

5.1 Group 1 and 2 - Horn of Africa

The Afar Rift triple junction is a key point in the Arabia, Nubian and Somalia plate tectonics, because it represents the point of accommodation of three concurring extensional regimes, which are the Red Sea and Gulf of Aden spreading ridges to the North and the Ethiopian rift system to the South. The whole area is interested by a significant seismic activity and several large earthquakes have been observed in historical and modern times. Surface geology and focal mechanism of earthquakes show that the whole region is dominated by normal faulting (Shudofsky, 1985; Kebede and Kulhanek, 1991), with a minor although not negligible strike slip component.

We formally separated the Red Sea and Gulf of Aden source zones (group 1) from the inland zones of the triple junction's southern branch (Afar, Ethiopian plateau and Ethiopian rift valley; group 2), which did not evolved (yet) in oceanic spread. The rationale behind this choice stays in the likely different seismic attenuation behaviour of the two neighbouring regions. However, this hypothesis has to be confirmed by analysis of local seismic recordings. The Main Ethiopian Rift, in particular, is a single-extensional rift basin between Nubia and Somalia extending from the Afar triple junction (Wolfenden et al., 2004; Keir et al., 2009) to the Lake Turkana depression in northern Kenya. Few earthquakes focal mechanisms exist for the Ethiopian Rift and most of them show ESE-WNW orientation and normal fault (Casey et al., 2006, Damien et al., 2010).

Table 3. Source zones of the current SSA model separated by tectonic group.

Group	Source	Name
1	2	South Read Sea
	3	Gulf of Aden
2	1	Afar Depression - Eritrea
	4	Main Ethiopian Rift
	22	North Kenya - Lake Turkana
3	7	Lake Victoria
	14	South Kenia
	20	Rowuma Basin
4	5	South Sudan
	6	Western Rift - Lake Kivu
	8	Western Rift - Tanganika
	9	Malawi - Nyasa Rift
	18	South Mozambique
5	10	Walikale and Masisi
	11	Luama rift
	12	Mweru - Katanga - Upemba
	13	Kariba - Okavango
6	15	Eastern Rift
	16	Davie Ridge
	17	Mozambique channel

5.2 Group 3 - African Microplates

South of Lake Turkana, seismic and tectonic activity delineate two branches, the Eastern and Western Rifts, which bound a relatively unfaulted, scarcely seismic domain centered on a 2.5–3 Ga old assemblage of metamorphic and granitic terranes (Tanzanian craton). This domain has remained undisturbed tectonically since the Archean (e.g., Chesley et al., 1999), except for minor seismicity under Lake Victoria, and it was interpreted by Hartnady (2002) as the present-day Victoria microplate. Seismic, xenolith and gravity data show that the 150–200 km thick lithosphere of the Tanzanian craton is colder and stronger than surrounding orogenic belts (Wendlandt and Morgan, 1982; Boyd and Gurney, 1986; Green et al.,1991; Ritsema et al., 1998; Weeraratne et al., 2003), which might lead to a less attenuating seismic propagation behaviour of the region. The low seismicity belt extends toward south of the Tanzanian craton and partially interests north Mozambique. This region, however, has been formally separated from the Victoria microplate as the independent tectonic domain of the Rowuma microplate (missing reference).

Although seismicity for these microplate regions is comparably very low with respect to neighbouring rifts zones, it still has to be represented into hazard model. This is done trough implementation of low-rates background area sources. Additionally to Victoria and Rowuma microplates, source group 3 partially extends to Nubian and Somalian plates to include the seismic clusters of South Sudan (to north-west) and Kenya (to the east). These regions have definitely higher seismic productivity, but they reasonably share a similar tectonic setup, which makes them suitable for the calibration of a common *b*-value.

5.3 Group 4 - Western Rift System

This group contains four sources, which cover segments of the Western branch of the EAR, showing the highest rates of seismicity along the whole rift system. It includes the Albertine Rift (which contains the Albertine Graben, Semliki Basin and Ruwenzori Mountains), the Lake Kivu Basin including the Virunga volcanic area, Lake Tanganyika and Malawi. The present-day fault kinematics as evidenced by the focal mechanism of events in the Albertine Rifts is normal faulting under NW-SE extension. Focal mechanisms in the Lake Kivu area display also a normal faulting with general trend N-S as opposed to the NE-SW of the Albertine-Ruwenzori segment. The Lake Tanganyika occupies the central part of the Western Branch. The stress inversion for North Tanganyika gives an ESE-WNW under Normal Faulting regime with a slight strike slip component. The southern part belongs to the TRM (Tanganyika-Rukwa-Malawi) rift segment, along which the kinematic model of Chorowicz (2005) infer dextral strike slip movements under NW-SE extension. The Rukwa Rift forms the central part of the TRM segment. In addition, the Ufipa Plateau between the Rukwa and South Tanganyika depression is affected by the 160 km-long Kanda active normal fault that might have generated the 1910 Ms 7.4 earthquake, which is the strongest ever recorded in the east Africa Rift (Vittori et al., 1997, Delvaux, 2010).

Most of the seismicity of the EAR is concentrated in the magma-poor Western Rift, which initiated around 25 Ma simultaneously with the Eastern branch (Roberts et al., 2012). The Western branch is characterized by low-volume volcanic activity, large ($M > 6.5$) magnitude earthquakes, and hypocenters at depths up to 30–40 km (Yang and Chen, 2010; Craig et al., 2011). From Lake Albert to southern Rukwa, the width of the Western branch does not extend more than 40–70 km, with large volcanic centers coincident with the basin segmentation (Virunga, South-Kivu, and Rungwe). The Western Rift connects southward with the Malawi Rift via the reactivated Mesozoic Rukwa Rift (Delvaux et al., 2012). The Malawi Rift itself shares similarities with the Tanganyika basin, with long and well-defined normal faults (e.g., Livingstone escarpment) and limited volcanism. The 2009 Karonga earthquake swarm, with 4 $Mw > 5.5$ events (Biggs et al., 2010), however, showed that additional hanging wall normal faults participate in present-day extension. Recent coring in Lake Malawi indicates that modern rift initiation may be as young as early to middle Pliocene, considerably younger than most prior estimates (Lyons et al., 2011).

Craig et al. (2011) found that seismicity (i.e. centroid depths) is confined in the uppermost 11 km. The seismicity observed within these areas shows very few events before 1960, probably because of the lack of seismic stations. The maximum magnitude observed corresponds to 7.3 (M_W); this was generated by the 1910 earthquake occurred in Rukwa. The accuracy of the focal depth is generally poor owing to the sparse station spacing. However, micro seismic studies indicate that earthquakes are generally between depths of 10 to 20 km in the Western Rift Valley (Zana, 1977; Zana and Hamaguchi, 1978).

5.4 Group 5 - Continental Africa

Masisi is located at the northwest of Lake Kivu. A study of earthquake focal mechanisms by Tanaka et al. (1980) showed that the direction of the fault traces in that area is SE-NW, and the average focal mechanism is normal faulting with the tension axis perpendicular to the strike of the fault traces. The last strong earthquake occurred in the Masisi area on 29 April 1995 (M_b 5.1) (Mavonga, 2007).

The most prominent seismotectonic features in this region are the Upemba and Moero or Mweru Rifts. The Upemba Rift is characterized by a NE-SW striking fault extending along its eastern side (Studt et al., 1908). The Upemba Rift may extend northward to the Kabalo area, which experienced an earthquake with magnitude Mw 6.5 on 11 September 1992. Detailed investigation has revealed that the main geological features in the Kabalo area trend in the NNE-SSW direction, similar to those found in the Upemba Rift (Zana et al., 2004).

5.5 Group 6 - Eastern Rift System

The Eastern Rift branch is characterized by a broad zone of shallow (5–15 km) and smaller magnitude seismicity, but voluminous volcanism (e.g., Dawson, 1992; Yang and Chen, 2010; Craig et al., 2011). The Eastern Rift includes the ~25Ma Turkana Rift, which reactivates part of an Eocene-Oligocene rift system (George et al., 1998; Pik et al., 2006). South of Lake Turkana, rifting and volcanism initiated at about 25 Ma (Furman et al., 2006; McDougall and Brown, 2009) with active eruptive centers along its length and moderate seismic activity. The seismically active southernmost part of the Eastern Rift, < 5 Myr old in the Natron basin, experienced in 2007 a discrete strain accommodation event rarely observed in a continental rift, with slow slip on a normal fault followed by a dike intrusion (Calais et al., 2008; Biggs et al., 2009).

South of the Natron basin, the Eastern branch of the EAR splits into the Pangani, Manyara, and Eyasi Rifts at an apparent triple junction (North Tanzanian Divergence, NTD) (Le Gall et al., 2004, 2008). The continuation of the Eastern branch south of the NTD appears more prominent along the Manyara Rift (Macheyeki et al.,2008), which may therefore form the eastern boundary of the Victoria plate. The aseismic plateau between the Manyara and Pangani Rifts has been interpreted as a microplate (Masai block), separate from Victoria and Somalia (Dawson, 1992; Le Gall et al., 2008). Farther south, the Manyara and Pangani Rifts connect with the Usangu basin to the southwest and with the Kerimbas Rift to the east. The presence of 17–19 Ma phonolites intruding the basin sediments [Rasskazov et al., 2003] indicates that the Usangu basin likely initiated in the early stage of rift development. The Usangu basin shows moderate seismicity and connects to the south with the Malawi Rift, while the Kerimbas Rift is continuous offshore with the Davie Ridge, a narrow, NS trending, zone of seismicity with purely east-west extensional focal mechanisms (Mougenot et al., 1986; Grimison and Chen, 1988). The southward continuation of the Davie Ridge is unclear, but it may connect with the Quathlamba Seismic Axis, a linear cluster of seismicity between Madagascar and southern Mozambique (Hartnady, 1990; Hartnady et al., 1992).

South of the Malawi Rift, active deformation extends along the seismically active Urema graben and further south along the Chissenga seismic zone and the Urrongas protorift swell (Hartnady, 2006), where the Mw 7.0 Machaze, Mozambique, earthquake of 23 February 2006 occurred (Fenton and Bommer, 2006; Yang and Chen, 2008).

From the study of focal mechanism, Delvaux & Barth (2010), deduce that an E-W extension occurs along the Davie Ridge. This ridge is considered to be the southward continuation of the Eastern Branch of the EARS.

The Western Branch of the EARS continue south of the Mbeya triple junction by the Malawi Rift and by more weakly expressed asymetric structures along the coastal region of Central Mozambique. The Mbeya lies at the triple junction between the Somalia, Victoria and Rovuma plates (Ebinger et al., 1989, Delvaux and Hanon, 1993). It contains the Rungwe volcanic province and links the NW trending South Rukwa and North Malawi rift basins with the NE trending Usangu basin. In February 2006, a Mw7.0 earthquake occurred in the Coastal region of Central Mozambique and generated a surface fault rupture observed over 15 km, with a possible overall extension of 30 km with a vertical separation from 0.4 to 2.05m and a component of left-lateral displacement of maximum 0.7m (Fenton and Boommer, 2006).

South of the Manyara-Dodoma Rift un Central Tanzania, active extensional deformation associated in the Eastern Branch of the EARS seems to jump laterally to the coastal region and the Indian Ocean. The coastal region of Kenya and Tanzania display homogeneous extension in an ENE-WSW orientation and a pure normal faulting regime. Between Mozambique and Madagascar, the Mozambique Channel is known for its seismicity associated mainly to the Davie Ridge.

6 Calibration of the Seismic Source Model

6.1 Source Depth Distribution

A model for source depth distribution was calibrated based on the available information from SSA catalogue. Unfortunately, not all reported events included an estimation of hypocentral depth solution. In few cases, moreover, such estimate was considered not reliable because of the large uncertainty (generally at depths larger than 40Km) or because being explicitly assigned a-priory (e.g. fixed solution depths of 5, 10, 15 and 33 km,). These events have been removed from the analysis. Nonetheless, a sufficient number of samples were available to perform a reasonable statistic (**Fig. 3**).

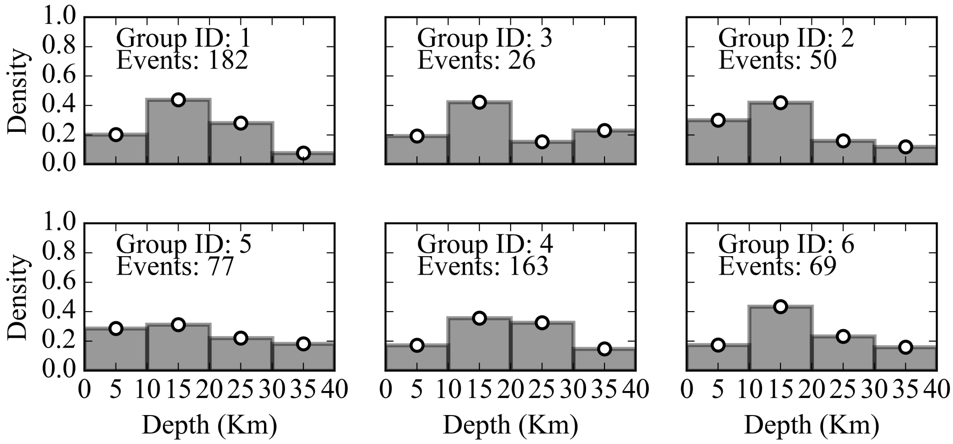


Fig. 3 Distribution of the hypocentral depth solutions of earthquake events falling into the main six source groups.

6.2 Source Mechanism

Geometry of the source is fully described by focal mechanism parameters strike, dip and rake. While strike and dip uniquely describe fault orientation, rake is used to further specify the style of displacement (normal, thrust, strike-slip or oblique). Such source parameters can be estimated directly by the analysis of fault-plane solutions from moment tensor inversion of earthquake recordings, or indirectly by the analysis of local and regional stress regimes and existing geological structures. We based our considerations on the geological and seismological literature available for the area (e.g references).

Tectonic regime of the study region is mostly extensional, although a minor but not negligible transform component is also relevant in many areas. Normal faulting style was modelled by imposing a standard (constant) dip angle of 60° and rake of -90°, while adding where necessary strike-slip component by allowing oblique strike on fault plane. Since in most cases precise information on average slip direction was not available, either left lateral (-45°) and right-lateral (-135°) rake components were allowed with equal probability.

The overall strike distribution was calibrated by performing statistic analysis on the outcropping fault structures available from the database of Macgregor (2015). To do this, mapped fault lineaments were split into segments of fixed length (1 km), in order to weight segments of different length proportionately, but also to avoid issues related to arbitrary segmentation of main faults. Segment statistic was then used to constrain average strike orientation in each zone (e.g. Fig. 4). In few cases, bimodal (and even more complex) distributions were found, which are likely due to a mixed tectonic regime.

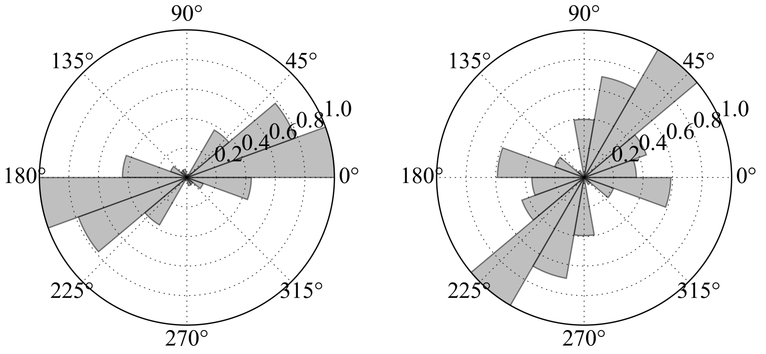


Fig. 4 Distribution of fault orientation (strike) for two example source zones (4 and 7). Input information is from the fault database of Macgregor (2015).

6.3 Seismicity Analysis

6.3.1 Magnitude-Frequency Distribution

Seismicity of each area source zone is assumed to follow a double truncated Gutenberg-Richter magnitude occurrence relation (or magnitude-frequency distribution, MFD). Lower truncation is arbitrarily assigned to M_W 4.5 for all zones. Upper truncation is defined as the magnitude of the largest earthquake assumed possible (or, rather, plausible) for an area. A different maximum magnitude (M_{MAX}) estimate is derived independently for each source group as the largest observed event plus an arbitrary - although quite conservative - increment of 0.5 magnitude units.

b-values have been calibrated for the whole catalogue and independently for each source group. Occurrence rates (a-values) have been conversely calculated for each source zone separately by imposing the previously calibrated b-values. This strategy was necessary given the limited amount of data available for the study area, and particularly for those zones of quite limited extension.

In addition to using standard and well-established approaches (e.g. Weichert's maximum likelihood), we tested an alternative strategy based on direct inversion of incremental earthquake occurrences. Seismicity parameters (*a*- and *b*-values) are obtained by minimizing the residuals between observed rates in discrete magnitude bins and a theoretical truncated MFD model (e.g. **Fig.** 5a). Such strategy is advantageous in that target observations are independent and the results is therefore not affected by discontinuous earthquake records, as for the case of uncertain completeness of reported large magnitudes. Moreover, a variety of *a priori* constraints (e.g. fixed b-value or maximum magnitude) can easily be included in the analysis.

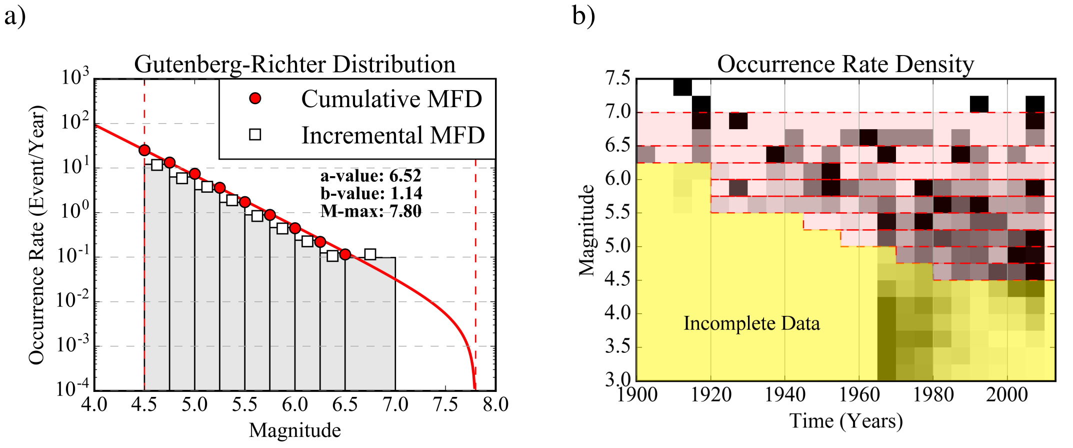


Fig. 5 a) Gutenberg-Richter magnitude occurrence relation of the declustered SSA-CAT. Red solid line and grey histogram are the fitted relation, while symbols represents observer rates (cumulative and incremental) for discrete magnitude bins. b) Corresponding catalogue completeness: on the background (with normalized scale) is the distribution of annual rates computed for discrete time windows (5 years), while in red is the time completeness table. Magnitude bins are discretized as in a).

6.3.2 Completeness Analysis

Earthquake catalogue completeness (Rydelek and Sacks (1989); Woessner and Wiemer (2005)) is evaluated for different temporal periods and magnitude ranges by integration of two complementary procedures. First, results from the unsupervised Stepp (1997) algorithm are evaluated, using the implementation available within the GEM's Hazard Modeller Toolkit (HMTK) (Weatherill, 2014). This method, however, proved to be unstable, giving potentially erroneous results in the case of sparse and irregular data coverage, as it is unfortunately the case for Sub-Saharan Africa. As subsequent manual refinement, therefore, we manually adjusted the completeness estimates by iterative comparison of the corresponding magnitude-frequency distribution. Such procedure is performed in a first stage for the whole SSA catalogue (e.g. **Fig.** 5b) and then for each source zone group.

6.3.3 Earthquake Rate Balancing

In order to avoid duplicate counts of events on overlapping zones (e.g. 12 and 12.1), redistribution of seismic rates is necessary. Background events have to be removed from the rates computed for the topmost overlapping layer, so that join calculation of the occurrence rates for the two zones will keep the total balance unmodified. First, unit-area background rate is obtained by counting the occurrences in the background region not falling also into the overlapping layer. This can be done by simple subtraction of the total events observed in the two zones. The background rate is then removed from the occurrence of the overlapping zone after rescaling by local area extension. For simplicity, we limited this procedure to just one single overlapping zone, but such strategy can nonetheless be extended to the use of several layers, each delimited by contouring the average density level of events over the area. This approach would be an intermediate approach between standard distributed and smoothed seismicity models.

Table 4 Calibrated seismicity parameters for each source zones, divided by tectonic group. Sources marked with .1 are representing overlapping layers within a background zone.

Group	Source	a-value	b-value	M-max
1	2	4.83	1.02	7.2
	3	5.38		
2	1	4.48	0.95	7.5
	4	4.18		
	22	3.70		
3	7	4.00	1.02	6.9
	7.1	4.23		
	14	4.34		
	20	3.31		
4	5	4.22	1.02	7.9
	6	4.89		
	8	4.84		
	9	4.93		
	18	4.40		
5	10	3.90	0.99	6.9
	10.1	3.92		
	11	3.51		
	11.1	3.93		
	12	4.05		
	12.1	4.13		
	13	4.08		
	13.1	3.99		
6	15	5.31	1.16	7.4
	16	5.45		
	17	4.77		

7 Logic Tree Implementation

While aleatory (or random) component of the model uncertainty is generally taken into account by describing the probability distribution of model parameters, the epistemic component, which is related to the available level of knowledge and/or the adopted initial assumptions and simplification, can be quantified by using a logic-tree strategy. In a logic-tree approach, different interpretations of the model components are considered concurrently. Statistic analysis is performed a-posteriori on the weighted outcome of each model realisation (or logic-tree branches). *OpenQuake* allows the use of different branching levels, each of those representing a separate contribution to uncertainty. A multilevel strategy assures the full exploration of the model variability by computation of all possible permutations of those model parameters affected by epistemic uncertainty. We applied this strategy to account for the difference between existing ground motion prediction models and for the variability of source parameters not directly constrained by available data.

Table 5 Weighting scheme used for the GMPE logic tree. Source zones sharing the same weights are grouped into four main categories (A-D).

Group ID	Source ID	CY	AK	AB	PZ
A	1, 2, 3, 4, 17	0.5	0.5	0	0
B	5, 6, 8, 9, 1, 8, 22	0.375	0.375	0.125	0.125
C	15	0.25	0.25	0.25	0.25
D	7, 10, 11, 12, 13, 14, 16, 20	0.125	0.125	0.375	0.375

7.1 Ground Motion Prediction Equations

The optimum strategy for the selection of most representative Ground Motion Prediction Equations (GMPE) is the direct comparison of empirical ground motion estimates with observed earthquake recordings in a sufficiently representative range of magnitudes and distances. The GEM Ground Motion Toolkit (GMTK) offers a set of simple functionalities to pursue this goal (Weatherill, 2014b). Unfortunately, Sub-Saharan Africa is affected by a severe lack of data availability. The use of AfricaArray networks did not contribute significantly, due to too low magnitudes covered, lack of recordings in the near to intermediate distance range (<50km). For these reasons, we alternatively had to rely for GMPE selection on a simpler - but less accurate - selection criteria, based on direct evaluation and comparison of GMPE features, such a tectonic context of validity, type and quality of data used for calibration and suitability of the functional form (Cotton et al. 2006).

In a first round, sixteen GMPEs from worldwide were selected as possible candidates, covering four different tectonic contexts: active shallow crust (ASC), stable continental crust (SCC), cratons (CRT) and volcanic areas (VLC). However, ground motion prediction equations from CRT and VLC settings have been trimmed off nearly immediately, because of the questionable applicability to the investigated area and the lack of available data to perform ad-hoc seismicity analysis. This last issue is particularly critical in case of volcano-related seismicity, which is nonetheless a possibly significant contribution to seismic hazard at specific sites. Once more data will be made available, it is advisable that this component will be progressively integrated into the model.

In a second attempt, GMPEs for ASC and SCC have then been assigned to different area source groups. While we used ASC GMPEs for areas involving plate boundary segmentation, SCC GMPEs where used to model ground motion in all intra-plate areas. The rationale behind this choice is the evolution of the African rifting. Given the relative young age of the process, it might be expected extra-rift regions to be less exposed to asthenospheric upwelling, and therefore to preserve a mechanical behaviour and a seismicity footprint typical of stable continental areas. However, after some sensitivity test calculation, we found that using a sharp separation between regions of different tectonic setting would have led to unjustifiable large differences in the computed ground motion across certain zone boundaries. In order to minimize such effect, while retaining the assumption of diversity in crustal attenuation and stress-drop, we proceeded with an alternative approach.

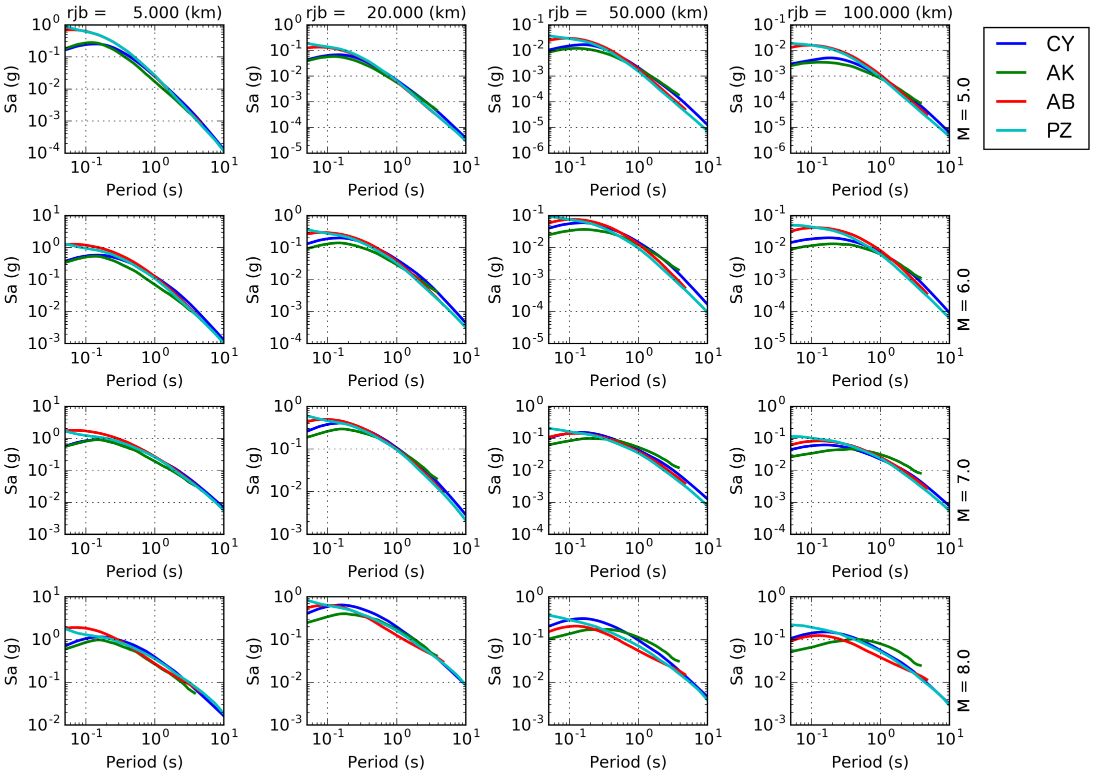


Fig. 6 Comparison between response spectra predicted by the four selected GMPEs as a function of Mw magnitude (along columns) and Joyner-Boore distance (along rows).

The current logic-tree model was restricted to the use of four GMPEs, respectively two for active shallow crust (CY - Chiou and Youngs, 2014; AK - Akkar et al., 2014) and two for stable continental conditions (AB - Atkinson and Boore, 2006; PZ - Pezeshk et al. 2011). We assigned then all of the selected GMPEs to each source zone, but allowing the corresponding logic-tree weight to vary in agreement with the likelihood for each specific tectonic type. Assignment of weights was agreed on the basis of the direct judgement of local seismotectonic conditions from a pool of experts within the region. The full list of weights is summarised in **Table 5**. Zones sharing the same weighting scheme have then been clustered into four main groups (named A to D) to reduce the total number of end-branches into the logic tree implementation. In Fig. 6 it is presented a comparison of the response spectra from the selected GMPE and for a range of magnitude and distance values.

7.2 Source Model Uncertainty

The source model logic tree has currently a master branch that includes the area source zonation previously described. On top of that, additional branching levels have been implemented to describe the epistemic variability of the assumed maximum magnitude of each zone (e.g. **Fig. 7**). Given the poor constraints available for its calibration, maximum magnitude is assumed to have a relative possible error of ±0.2, assigned empirically with a certain level of conservatism. The higher weight (0.5) is assigned to the original unmodified magnitude estimate, while edge values (±0.2) have a lower probability of 0.25 each.

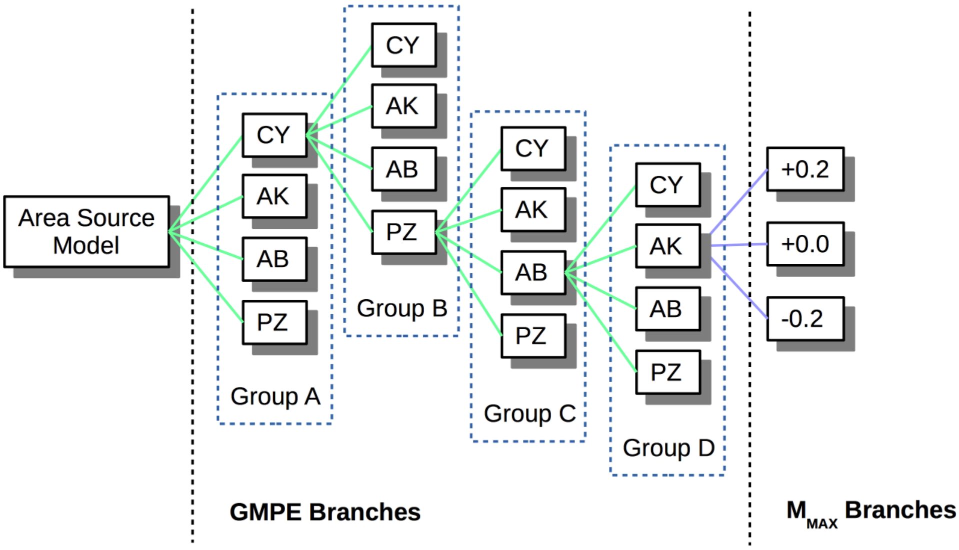


Fig. 7. Schematic representation of branch permutation for the current logic-tree implementation.

8 PSHA Results

8.1 *OpenQuake* Settings

Hazard computations have been performed using OpenQuake engine (Version 2.0) through the available classical calculator for distributed seismicity (see OpenQuake Reference Manual for details on available calculators). Investigation area consists of a mesh of 79109 sites spaced at approximately 10 km. Such area includes all earthquake source zones described in chapter 3, plus a buffer region of not less than about 100 km. For each site of the mesh, free rock conditions are assumed, with a fixed 30-metre averaged shear-wave velocity (Vs_{30}) reference of 600 m/s (corresponding to stiff soil transition in EC8 and NERHP classification).

Target ground motion intensity for calculation is 5% damped response spectral acceleration (in g), estimated for probabilities of exceedance of 10% and 2% within an investigation time of 50 years. This corresponds respectively to return periods of about 2474 and 474 years. Due to the substantial lack of historical records for proper calibration of the large magnitude rates, we avoid using longer return periods.

According to the possibilities of the selected GMPEs, spectral acceleration has been computed at PGA and for the response spectral periods of 0.05s, 0.1s, 0.2s, 0.5s, 1s and 2s. Ground motion integration has been conservatively truncated at 3 sigma of the predictions. Output of the calculation are mean and quantile (0.15, 0.5 and 0.85) hazard curves at each site, together with Uniform Hazard Spectra (UHS) and hazard maps, which are described in the next sections.

8.2 Calculation Outputs

Hazard calculations have been performed for each site of the investigation grid. For the sake of conciseness, however, in the following we illustrate hazard results for four selected African capitals, which are considered of significant for risk analysis:

- Addis Ababa (Ethiopia);
- Kampala (Uganda);
- Nairobi (Kenya);
- Bujumbura (Burundi).

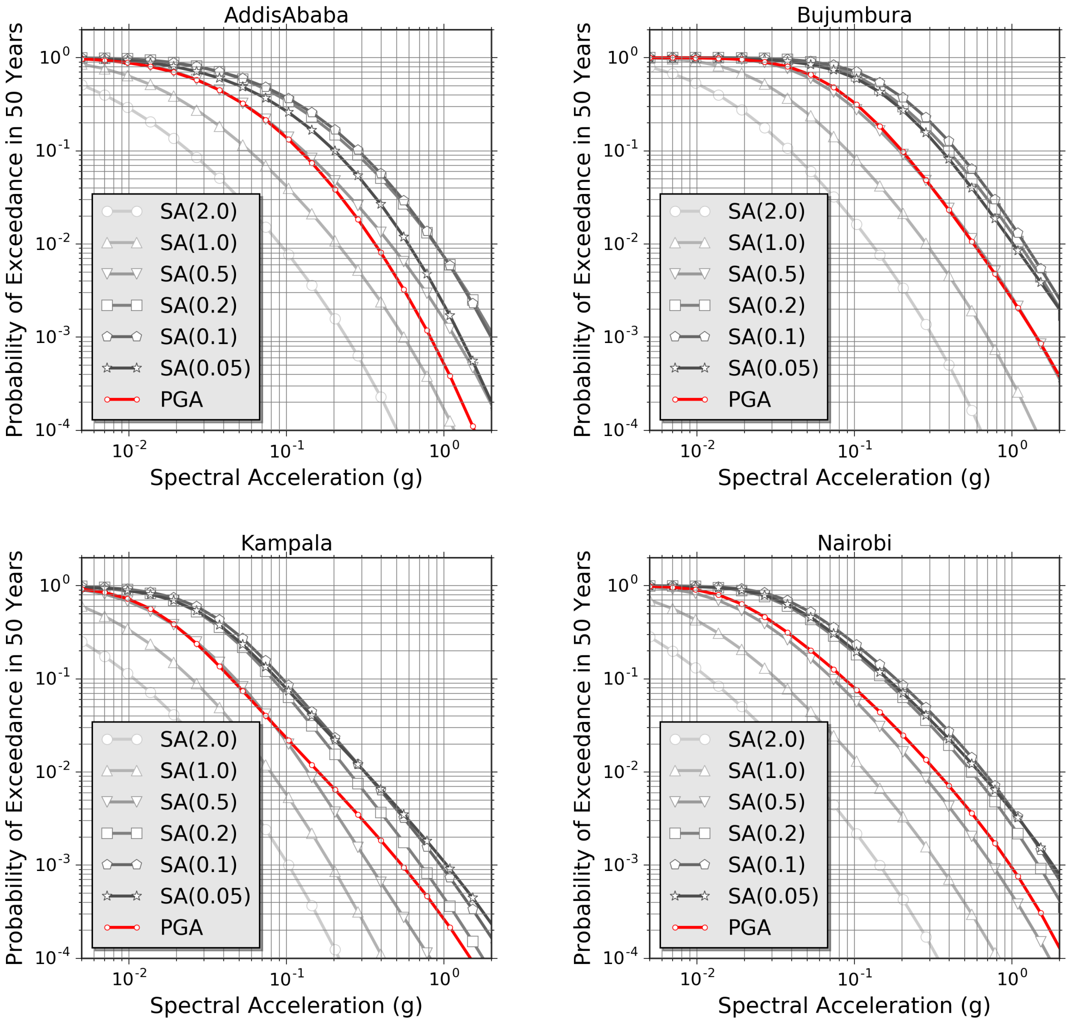


Fig. 8 Mean hazard curves computed for a range of spectral periods, including PGA (in red), at the four example African locations.

8.2.1 Earthquake Hazard Curves

Hazard curves are calculated for fixed acceleration ordinates between 0.005g and 2.13g and separately for each prescribed spectral period (including PGA). Acceleration corresponding to the target probability of exceedance(s) is subsequently extracted from the curves by linear interpolation. In **Fig. 8** are presented mean hazard curves for different spectral periods at the four example locations. It has to be noted the unusual behaviour of the hazard curves in Kampala at long periods; this is likely due to concurrence of GMPE for shallow crust and stable continental conditions, which is affecting differently the different probabilities.

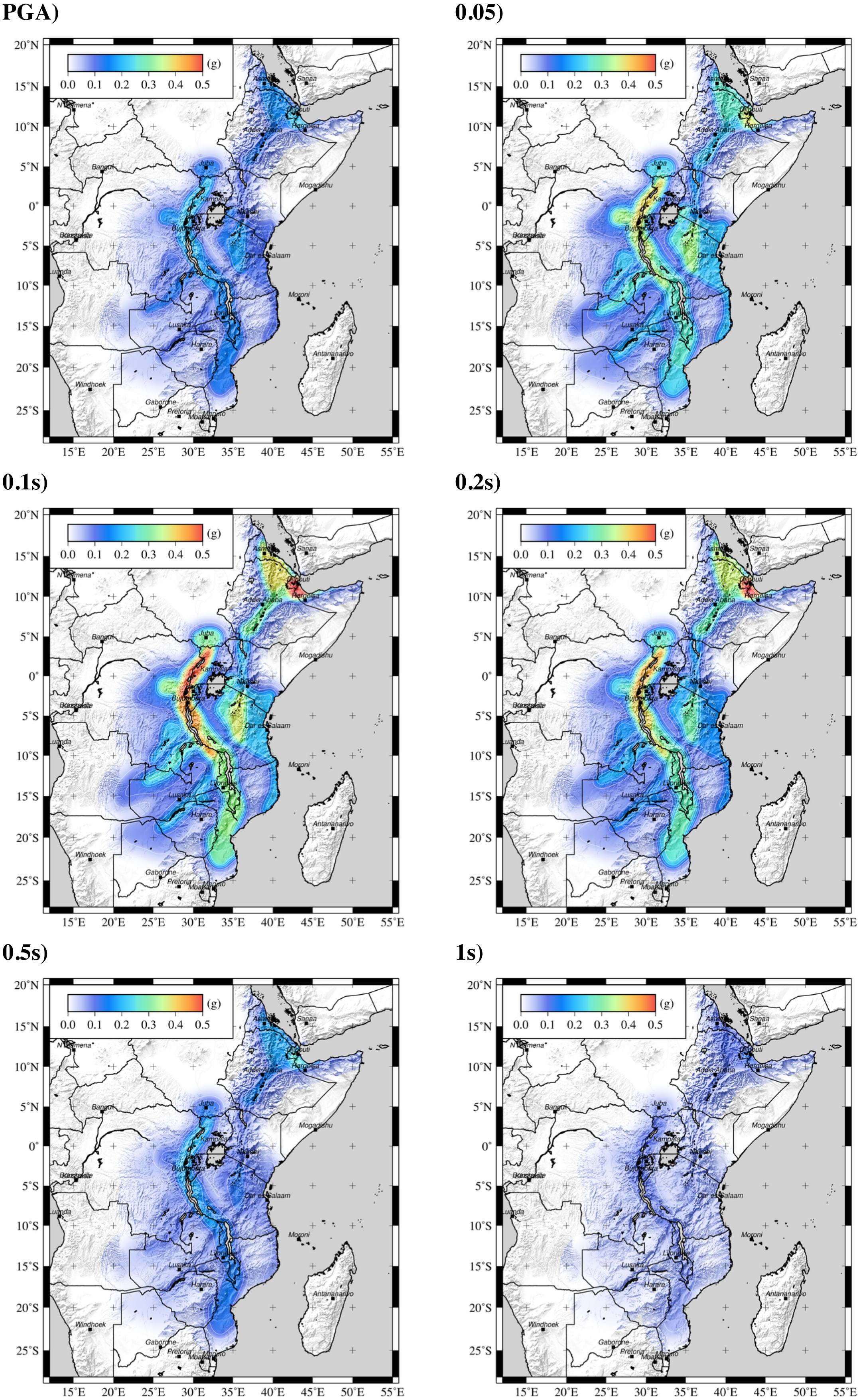
8.2.2 Earthquake Hazard Maps

A series of hazard maps have been produced for different return periods and POEs. Largest accelerations are found for periods of 0.1s and 0.2s along the eastern branch of the EARS (0.51g), particularly for source zones 6 and 8. Moderate acceleration (less than 0.35g) is experienced in the Afar region (zone 1) in north Ethiopia. South Ethiopia (zone 4) presents levels (0.24g at 0.2s) that are comparable to western EARS (zone 15) and the side seismic belts of Zambia (zone 12). Remaining portions of the rift are affected by an overall lower hazard, with accelerations generally lower than 0.2g.

8.2.3 Uniform Hazard Spectra

Uniform Hazard Spectra (UHS) are computed by collecting ground motion for a given probability of exceedance over a spectrum of different response periods. This representation is useful to highlight those periods where larger spectral acceleration is expected. It is however important to stress that UHS cannot be directly used to model local scenarios (e.g. for the selection of a reference earthquake), as the different spectral ordinates might be (and likely are) linked to different controlling events. For that purpose, a disaggregation procedure is best suited.

In **Fig. 10** mean and quantile UHS are shown for the four selected African capitals. It is evident a considerable contribution to hazard for periods between 0.1s and 0.2s. This is not surprising as it is related to residual effect of local soil conditions, which are likely to affect ground motion predicted in this frequency band (5-10Hz). Such period range is also significant from engineering perspective, as it match the resonance response of common buildings in urban environments.


PGA)

0.05)

0.1s)

0.2s)

0.5s)

1s)

Fig. 9 Map of spectral acceleration (g) for 10% probability of exceedance in 50 years.

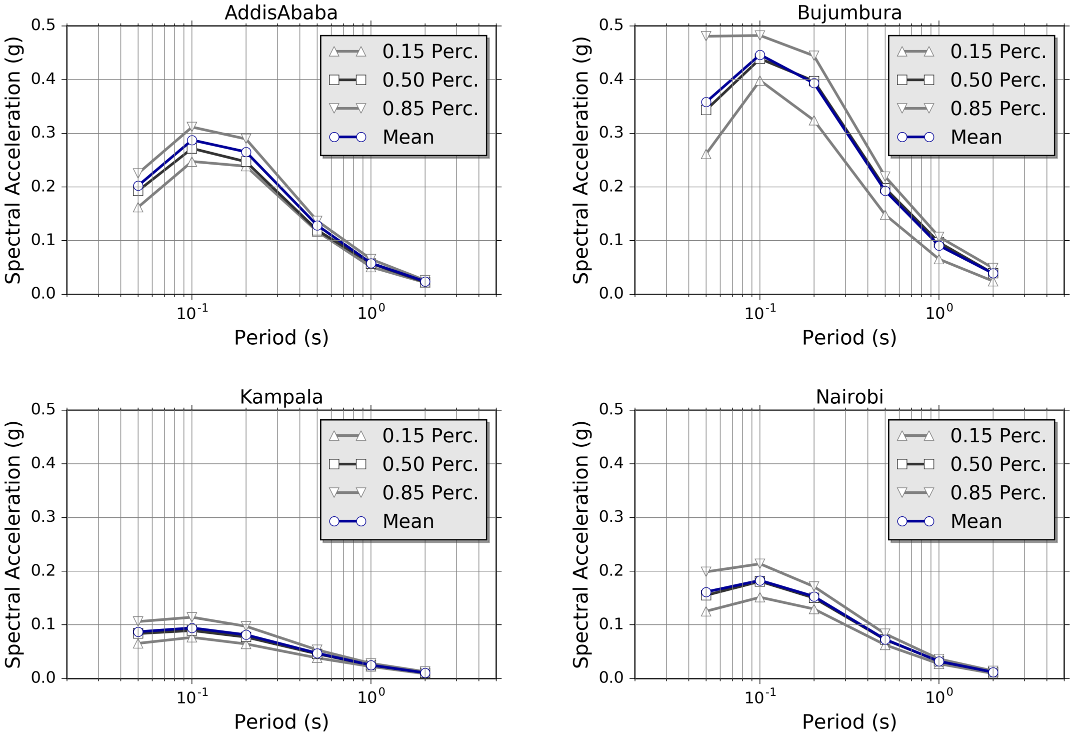


Fig. 10 Mean and quantile Uniform Hazard Spectra computed for the four selected African capitals.

8.2.4 PSHA Disaggregation

9 Discussion and Conclusions

Text

10 Data and Resources

OpenQuake input files for hazard computation are made openly available though GEM platform.

Acknowledgments

This study is made possible by the generous support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of GEM and do not necessarily reflect the views of USAID or the United States Government.

References

Storchak, D.A., D. Di Giacomo, I. Bondár, E. R. Engdahl, J. Harris, W.H.K. Lee, A. Villaseñor and P. Bormann, 2013. Public Release of the ISC-GEM Global Instrumental Earthquake Catalogue (1900-2009). Seism. Res. Lett., 84, 5, 810-815.

Storchak, D.A., D. Di Giacomo, E.R. Engdahl, J. Harris, I. Bondár, W.H.K. Lee, P. Bormann and A. Villaseñor, 2015. The ISC-GEM Global Instrumental Earthquake Catalogue (1900-2009): Introduction, Phys. Earth Planet. Int., 239, 48-63.

Stiphout, T. van, J. Zhuang, and D. Marsan (2012). Theme V - Models and Techniques for Analysing Seismicity. Technical report. Community Online Resource for Statistical Seismicity Analysis. doi:10.5078/corssa-52382934.

Weatherill, G. A. (2014) OpenQuake Hazard Modeller’s Toolkit - User Guide. Global Earthquake Model (GEM). Technical Report