PSHA input model documentation for South America (SAM)

GEM Hazard Team

Version history

Table 1 summarises version history for the SAM input model, named according to the versioning system described here, and indicating which version was used in each of the global maps produced since 2018. Refer to the GEM Products Page for information on which model versions are available for various use cases. The changelog describes the changes between consecutive versions and are additive for all versions with the same model year.

The following text describes v2020.1.0.

Table 1 *– Version history for the SAM input model.*

Version	2018.1	2019.1	2022.1	2023.1	Changelog
v2016.0.0					New version of the model developed by GEM. In this version the following modifications were included to the source model: a new subduction model for Nazca subduction, and LAN subduction from CCA model; a reviewed fault model with strong implications for Colombia, Ecuador, Peru and Bolivia and updated gridded seismicity model. The gmmLT for the stable shallow crust region was modified and the one for LAN was taken from the CCA model.
v2018.0.0	X	X			Updated version of CCA-LAN subduction sources is used, which increases hazard in northeastern Venezuela.
v2018.0.1					Incorporates new faults in Colombia, Ecuador, Argentina, and Chile. Improved assignment of catalogue completeness, increased Mmax in places, homogenized the spatial pattern of the distributed seismicity and better defined the lower seismogenic depth limits. Subduction sources in the Colombia-Ecuador pacific segment were revised and a unique segment is now considered for this area.
v2020.0.0			X		Mmin extended to M4 for crustal distributed seismicity. Source ids were revised to work with disaggregation by source. Inslab source files were consolidated into a single one. gmmLT.xml updated with more recent GMPEs.
v2020.1.0				X	First version of the model.

Authors: J. Garcia-Pelaez, M. Pagani, K. Johnson, R. Styron, V. Poggi

1 Summary

The first version of the South America Model (SAM) was developed in the framework of the SARA project (SARA project v1.0, *Garcia et al., 2017*) funded by Swiss Re Foundation and benefits of the contribution of an important group of South American Institutions:

- Colombian Geological Survey, Colombia,
- National University of Colombia, Colombia,
- University of Valle, Colombia,
- Venezuelan Foundation for Seismological Research, Venezuela,
- University of Merida, Venezuela,
- National Observatory, Brazil,
- Institute of Astronomy, Geophysics and Atmospheric Sciences, Brazil,
- Federal University of Rio Grande do Norte, Brazil,
- National Polytechnic University, Ecuador,
- San Calixto Observatory, Bolivia,
- University of Chile, Chile,
- University of Concepción, Chile,
- Pontifical Catholic University of Chile, Chile,
- Research Center for Disaster Risk Management, Chile,
- Peruvian Geological Survey (INGEMMET), Peru,
- University of San Luis, Argentina,
- University of Antofagasta, Argentina,
- National Institute of Seismic Prevention, Argentina,
- National Institute of Geophysics and Volcanology, Milan, Italy.

Here, we describe an updated version of the SAM model made by GEM hazard team. This model covers the whole South American continent, with the exception of the Faulkland and the Galapagos Islands. Panama and the northeastern part of the Caribbean have been updated using newer information from the CCARA project, a GEM collaboration project funded by USAID.

2 Regional Tectonics

Seismicity in South America primarily results from the subduction of oceanic lithosphere beneath the western margin of the continent and from the eastward translation of the Caribbean plate along the continent's northern margin. In the west, the Nazca plate subducts beneath the central and northern part of the continent, while the Antarctic plate subducts beneath the southern part of the continent. This subduction occurs at the Nazca Trench, and has produced several of the largest earthquakes ever recorded, including the the 1960 *Mw* 9.4 Valdivia, Chile earthquake and the 2010 *Mw* 8.8 Maule, Chile earthquake.

This subduction has also caused substantial deformation of the upper plate as well, creating the Andes mountains. This deformation is ongoing and distributed throughout the Andean zone, and may be mostly accommodated on large, rapidly-slipping thrust faults in the Subandean Zone at the eastern rangefront of the Andes. However, oblique subduction and gravitational forces result in strike-slip and normal faulting throughout much of the Andes as well.

The Caribbean plate dips south beneath northern South America, though convergence rates are low; this plate boundary may be best characterized as strike-slip, with important left-lateral faults located on the densely-populated Caribbean coast of Colombia and Venezuela. Offshore eastern Venezuela, the Caribbean plate terminates and Caribbean-South American relative motion is accommodated on the Lesser Antilles subduction zone in the Atlantic Ocean.

To the east and south of the Andes, the South American continent may be considered stable. Nonetheless, like many stable continental regions, faults do exist–many of these date to the assembly and breakup of the supercontinent Pangea hundreds of millions of years ago. Stresses transmitted through the cold, strong South American shield produce some earthquakes in Brazil and the surrounding nations, though the rates of seismicity in central and eastern South America is a small fraction of the western and northern active margins.

3 Basic Datasets

The South American seismic model was created using a compilation of updated and harmonised databases needed for PSHA (e.g. historical and instrumental catalogue, a compilation of active fault data, a database of strong motions recordings) which were created following, to the greatest extent possible, common standards and transparent procedures. The construction of these databases was completed mainly by a number of South American scientists and engineers. In the next lines we briefly describe the main Datasets created in the framework of the SARA project and how some of them those were recently updated.

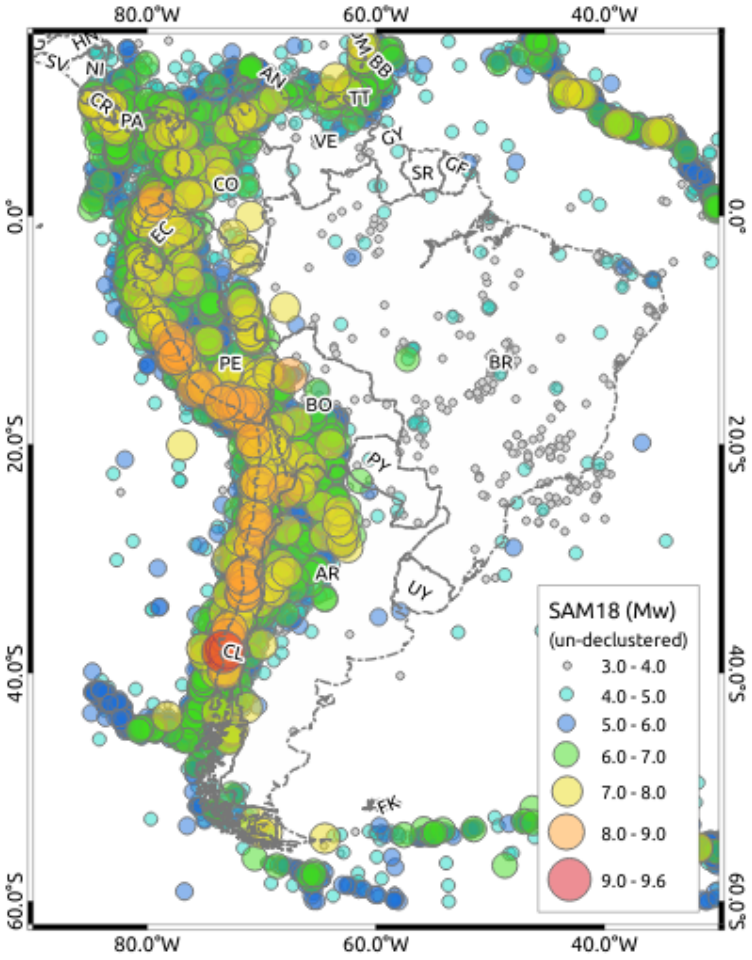
3.1 Earthquake Catalogues

An earthquake catalogue for South America was initially developed during the SARA project (SARA catalogue v1.0). This catalogue was compiled using information from a wide collection of earthquake databases and scientific publications to cover both historical/pre-instrumental and instrumental time frames, yielding a a parametric harmonized earthquake catalogue. The catalgoue has since been updated, including the following changes.

- Historical events after 1906 were replaced by instrumental recordings; now the period between 1906 and 1964 is treated the same way as the instrumental part in *Garcia et al., 2017*,
- Newer or updated information was added from the local South American agencies.

The final catalogue is presented in Figure 1, and covers the whole South America region, and part of Central America (Panama) and the Caribbean (Trinidad and Tobago). The non-declustered version consists of 106798 events from 1513 to 2017 and in the magnitude range 3.0 $\geq M_w \geq$ 9.6.

Figure 1 *– Final catalogue used to develop the SAM model.*



We also used the Global Centroid Moment Tensor (GCMT) focal mechanisms from 1976-2017 (*Dziewonski et al., 1981; Ekström et al., 2012*). This catalogue was used in the characterization of the sources as well as the definition of the geometry of whole model.

3.2 Catalogue tectonic classification and pre-processing

The classification of seismicity taking into account its possible tectonic environment is a crucial and fundamental step in the development of a seismic hazard model. Here, the catalogue was classified following Pagani et al. (2020) (see also Section 11.5), assigning earthquakes to the active shallow crust, subduction interface, subduction intraslab, and

deep earthquakes in the Bucaramunga nest. The subduction earthquakes are further subcategorized by segment (described below). The results of the classification are presented in Figure 2.

The catalogue was then declustered in order to isolate the mainshock earthquakes using the *Gardner and Knopoff (1974)* declustering algorithm and the spatial-time windows of *Uhrhammer (1985)* (see details in Section 11.5).

3.3 Fault Database

The crustal faults in the SAM PSHA are based on a fault database resulting from collaborations starting before the SARA project.

The initial compilation was completed by a group of South American geologists and based on active fault data from many published sources, including the International Lithosphere Program II-2 project (*Costa et al., 2003*), the Multinational Andean Project (*Getsinger and Hickson, 2000*) and existing national neotectonic databases and fault data collections.

The GEM hazard team collaborated with local experts from the start of the SARA project until 2020 to complete the fault database used in the current version of the PSHA input model for South America (e.g., Costa et al., 2016; Garcia et al., 2017; Yepez et al., 2017; and collaborations with between GEM and the CGS). Emphasis was given to faults that meet the following criteria:

- Slip rates equal or larger than 0.1 mm/yr (this applies to rates already available or estimated by compilers using quantitative-semiquantitative criteria)
- Evidence of Late Pleistocene tectonic activity (confirmed and suspected)
- Confirmed or suspected sources of earthquakes with $M_w \geq 5.5$.

The traces and surface projections of faults used in the model are shown in Figure 3

3.4 Ground Motion Database

The ground motion characterization is based on a database of strong ground motion recordings that were compiled during the SARA project, with contributions from scientists in Colombia, Bolivia, Brazil, Chile, Ecuador and Venezuela. This database (see Figure 4) contains about 2100 records for the different tectonic regimes present in South America (ASC: active shallow crust, SCC: stable continental crust, Subduction: interface and in-slab and Deep: Bucaramanga nest), all of them processed following the schema proposed by *Boore et al. 2012* and organized according to *Weatherill (2014)*.

3.5 Model

The SAM model was built using the GEM toolkits and the workflows described in Chapter 6.

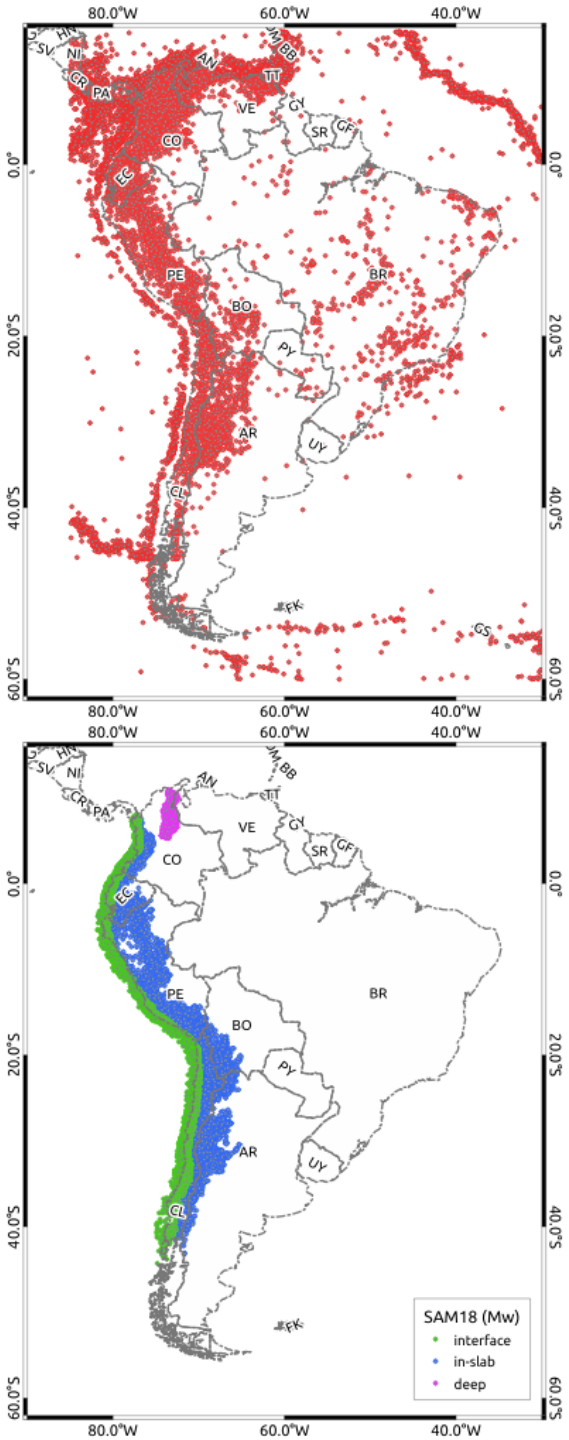


Figure 2 *– Catalogue classified by tectonic region. Top: active shallow crust. Bottom: Interface, intraslab, and deep.*

Figure 3 *– Crustal fault sources in the SAM model.*

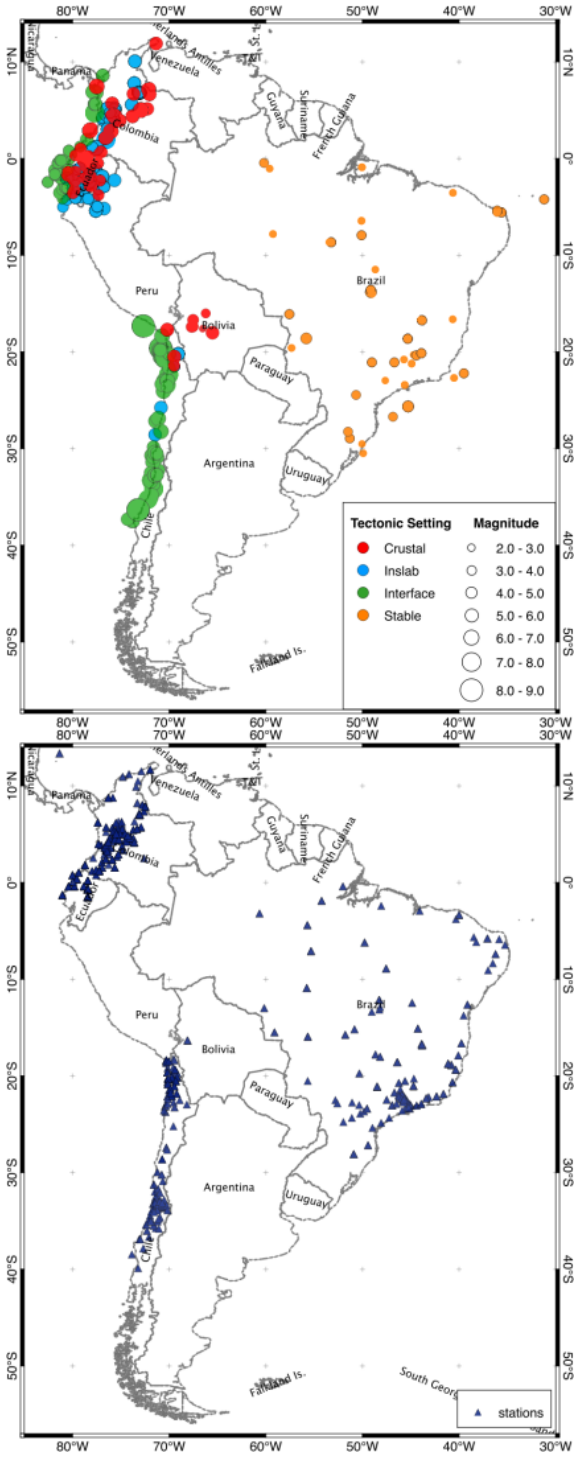


Figure 4 *– SARA Ground motion database. Left: Events classified according its tectonic context. Right: Station locations.*

4 Seismic Source Characterisation

The source characterization for the SAM model was divided into three main components:

1. the shallow crustal seismicity is modelled using a combination of **distributed seismicity** (gridded point sources with smoothed rates) and **"simple" fault sources**
2. the subduction interface seismicity, modelled as large **"complex" fault** sources with a 3D geometry,
3. the subduction intraslab and deep seismicity is modelled as **nonparametric ruptures** with pre-defined geometries.

In the following sections we briefly describe the main characteristics of each component.

4.1 Shallow seismicity: Distributed Seismicity

This component covers both, the active and stable shallow crust in South America.

Distributed seismicity (i.e. the earthquakes occurring away from faults in the fault database) is modelled using smoothed seismicity within "source zones" that correspond to regions with consistent characteristics, including tectonics, kinematics, and seismicity rates.

The catalogue was the main resource used to characterize the sources, and was used with other geologic and tectonic information to define the following parameters:

- the hypocentral depth distribution,
- the orientation and faulting styles of ruptures (focal mechanisms distribution and fault geometry characterization),
- the magnitude-frequency distribution (MFD), represented by a truncated *Gutenberg-Richter* distribution, with *b* and *a* values computed using the maximum likelihood estimator (*Wiechert, 1980*).
- the maximum magnitude, obtained by adding an increment of *0.5* to the largest observed event,
- the seismogenic thickness (following the procedure implemented to classify the seismicity),

Within each source zone, the seismicity rates are "smoothed" based on the locations of past earthquakes, using a method similar to *Frankel, 1995* (see Section 6.1). In addition, we adjusted the grid points around the fault sources to avoid double-counting of the ocurrences.

The source zone perimeters are depicted in Figure 5, and the source zone parameters are listed in Table 2.

4.2 Shallow seismicity: Crustal fault

The shallow crustal faults are modelled as simple fault sources, following the methodology described in Section 6.5.

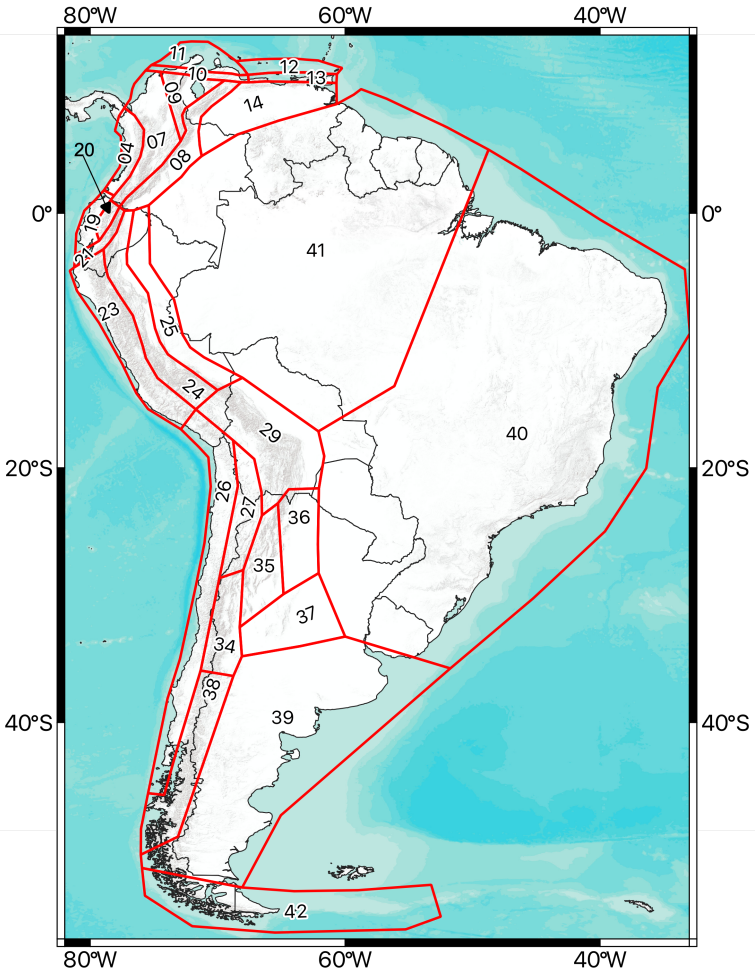


Figure 5 *– Source zone perimeters for the distributed seismicity sources.*

ID	a-value	b-value	M_{obs}	M_{max}
22	**4.993**	**1.036**	**7.45**	**7.75**
24	**4.634**	**1.023**	**7.60**	**7.90**
23	**4.735**	**0.977**	**7.50**	**7.80**
4	4.097	0.859	7.10	7.40
6	3.900	0.910	6.55	6.85
7	4.223	0.924	7.27	7.57
8	4.546	0.900	7.60	7.90
9	4.610	1.031	7.20	7.50
10	4.082	0.979	6.40	6.70
11	4.466	1.009	6.35	6.65
12	3.287	0.855	7.31	7.61
13	3.663	0.776	7.31	7.61
14	3.690	0.856	6.52	6.82
30	**3.5891**	**0.832**	**6.36**	**6.66**
32	**4.6433**	**1.141**	**7.50**	**7.80**
33	**4.0836**	**0.996**	**7.20**	**7.50**
19	4.799	1.049	6.60	6.90
20	3.805	0.922	6.32	6.62
21	5.205	1.090	7.60	7.90
23	4.402	0.910	7.28	7.58
24	4.966	0.905	7.60	7.90
25	4.421	0.940	6.71	7.01
26	3.837	0.771	7.37	7.67
27	3.737	0.865	7.06	7.36
29	3.948	0.831	6.58	6.88
32	3.481	0.712	6.95	7.25
34	4.809	0.950	7.26	7.56
35	3.992	0.813	7.48	7.78
36	3.494	0.775	7.01	7.31
37	4.443	1.019	6.21	6.51
38	4.058	0.832	7.24	7.54
39	2.846	0.685	6.69	6.99
40	4.055	1.057	5.80	6.10
41	3.560	0.817	6.34	6.64
42	3.659	0.818	7.69	7.99

Table 2 *– Source zones used to model distributed seismicity in South America. ID corresponds to Figure 5. Rows in bold are for sources that come from the model for Central America and the Carribean.*

The occurrence rates for most faults use a truncated *Gutenberg-Richter* (*GR*) model, while a few use the Youngs and Coppersmith (1985) MFD. The MFDs are constrained by fault slip

rates, using M_{min} = 6.5 and M_{max} computed from the fault dimensions. The MFD b-value corresponds to the encompassing source zone.

4.3 Subduction

4.4 Source geometry

The subduction source geometries are built using the GEM Subduction Toolkit (see Section 16.1).

We segment each subduction zone or subduction-style seismicity zone accodring to past megathrust earthquakes, current seismicity patterns, trench convergence rates and kinematics, and assistance from thorough structural and tectonic regional studies. The sources are built such that ruptures do not propagate across the defined boundaries. We include the Nazca subduction zone, the North Panama deformation belt, and the Lesser Antilles subduction zone.

Nazca Subduction zone

We test multiple segmentations, and ultimately choose the following eight-segment model, described from south to north (note that use of countries in segment geographical descriptions is approximate):

- *Segment 1 (Southern Chile):* Extends from the southernmost slab to the projection of the Juan Fernandez ridge to the trench, and is close to that proposed by *Salliard et al. (2017)*. This segment hosted both the 1960 *MW* 9.6 Valdivia and 2010 *MW* 8.8 Maule earthquakes.
- *Segment 2 (Northern Chile):* Terminates approximatively where the Iquique ridge intersects the trench. This limit agrees with *Saillard et al. (2017)* and *Villegas-Lanza et al. (2016)*, as well as interface segmentation for the 2017 USGS seismic hazard model of *Petersen et al. (2018)*.
- *Segment 3 (Southern Peru):* Terminates at the Nazca ridge, and is also used as a segment boundary by *Villegas-Lanza et al. (2016)* and *Brizzi et al. (2018)*. The ridge itself is a low-coupling region along the interface (*Brizzi et al., 2018*).
- *Segment 4 (Central Peru):* Terminates at the Mendana Fracture Zone, where the crustal age changes and seismic coupling decreases (*Villegas-Lanza et al., 2016*).
- *Segment 5 (Northern Peru):* Terminates north of the approximate intersection of an oblique strand of the Viru Fracture Zone with the trench, where seismic coupling further decreases (*Villegas-Lanza et al., 2016*).
- *Segment 6 (Southern Ecuador):* Terminates at the bouyant Carnegie Ridge intersection with the trench, close to a limit used in the highly segmented model by *Beauval et al., (2018)*.

- *Segment 7 (Northern Ecuador/Southern Colombia):* Terminates at the spreading center just north of Malpelo ridge. Approximately coincident with the northern boundary of subduction used by *Petersen et al. (2018)*.
- *Segment 8 (Northern Columbia):* Extends to the northern limit of the Nazca plate.

Segments 2-8 are assigned a locking depth of 50 km, representing the slip observed in many finite rupture models of the Finite-Source Rupture Model Database. For Segment 1 extends to 60 km, consistent with the Maule earthquake, and to yielding fault areas large enough to compute *MW* 9.6 using subduction geometry scaling relationships (Strasser, 2010; Thingbaijam and Mai, 2017; and Allen and Hayes, 2017) These locking depths are used as the cutoff depth between interface and subduction sources.

North Panama

For the North Panama Deformation Belt, we created subduction-style geometry for a single surface, following the same procedure as for the western South America Nazca subduction zone. At this convergent boundary, there is little shallow seismicity, and so we create only a slab model from six profiles spaced at 100 km (see next section).

Lesser Antilles The Lesser Antilles subdction sources was created in the framework of the CCARA project.

4.5 Interface rates

We derive a magnitude-frequency distribution (MFD) for each interface segment using a hybrid approach that combines statistics from observed seismicity with a characteristic component derived from tectonics. Source characteristics with references are summarized in Table 3. The listed characteristic magnitude (*Mchar*) is the median magnitude computed from the scaling relationship *Thingbaijam and Mai (2017)*.

Segment	***a*-Value**	***b*-Value**	**Mmax,obs**	**M*char***	**C. rate (mm/yr)**	**Coupling**
1	4.26	0.83	9.6	9.17	68	0.80
2	4.56	0.76	8.31	8.99	79	0.95
3	5.66	1.03	8.69	8.95	62	0.75
4	5.18	1.01	8.40	8.67	61	0.80
5	4.48	0.88	7.85	8.62	58	0.55
6	5.90	1.11	7.32	8.73	58	0.55
7	3.12	0.64	8.75	8.69	58	0.55
8	3.21	0.70	7.10	8.12	58	0.55

Table 3 *– Characteristics of the subduction interface sources for the Nazca subduction zone and the parameters used to derive these.* ***C. rate****: convergence rate.*

The Lesser Antilles subduction zone is discussed in the model for the Carribean and Central America.

Intraslab seismicity In our source model, segmentation boundaries from the interface extend into the slab, and are not meant to suggest barrier to rupture within the downgoing slab volume, but instead to allow spatial variability in observed seimsicity while still using non-parametric ruptures. We model the rupture geometries and rates following the standard GEM methodology for slab earthquakes, described in Section 16.1 for the range M_w 6.5-8.5. Results are in Table 4.

Subduction zone	**Segment**	***a*-Value**	***b*-Value**
Nazca	1	5.955	1.126
	2	6.733	1.115
	3	5.802	0.954
	4	4.758	0.882
	5	3.304	0.676
	6	4.568	0.828
	7	3.287	0.668

Table 4 *– Subduction intraslab parameters.*

5 Ground Motion Characterisation

The harmonized database of earthquake recordings (Figure 4), which covers the major tectonic regions found in South America (subduction interface, subduction in-slab, active shallow crustal and stable continental) was used to select suitable GMPEs for each tectonic region. This was initially completed during the SARA project.

The logic tree has since been updated. Table **??** presents the GMPEs and weights that are currently being used for South America.

Subduction Interface	**Weight**
AbrahamsonEtAl2015SInter	0.33
ZhaoEtAl2006SInter	0.33
MontalvaEtAl2016SInter	0.34
Subduction IntraSlab LAN_PRC	**Weight**
AtkinsonBoore2003SSlab	0.5
AbrahamsonEtAl2015SSlab	0.5
Active Shallow Crust	**Weight**
BindiEtAl2014Rjb	0.33
AkkarEtAlRjb2014	0.33
BooreEtAl2014	0.34
Subduction Interface CCA	**Weight**

AbrahamsonEtAl2015SInter	0.34
ZhaoEtAl2006SInter	0.33
ParkerEtAl2020SInter	0.33
Active Shallow Crust CCA	**Weight**
AbrahamsonEtAl2014	0.33
CauzziEtAl2014	0.34
AkkarEtAlRjb2014	0.33
Stable Shallow Crust	**Weight**
PezeshkEtAl2011NEHRPBC	0.33
AtkinsonBoore2006Modified2011	0.33
SilvaEtAl2002MwNSHMP2008	0.34
Subduction IntraSlab	**Weight**
MontalvaEtAl2016SSlab	0.5
AbrahamsonEtAl2015SSlab	0.5

Table 5 *– GMPEs used in the SAM model.*

5.1 References

GEM/CGS collaboration projects

Carlos C., Garcia J., Alvarado A., Audemard F., Audin L., Benavente C., Bezerra F. H., Cembrano J., González G., López M., Minaya E., Paolini M., Pérez I., Santibanez I., Arcila M., Delgado F. and Pagani M. (2016). From neotectonic data to seismogenic sources in South America: Results and lessons learned from the SARA project, Proceeding of 35th General Assembly of the European Seismological Commission, ESC2016-479, Trieste, Italy, 2016

Costa, C., H. Cisneros, M. M. Machette, R. L. Dart, (2003). A new database of Quaternary faults and folds in South America. ILP Task Group II-2 (western Hemisphere). Proceedings of the A. G. U., Fall Meeting 2003.

Garcia et al., 2017

Leonard, 2010, 2014

Yepez et al., 2017

6 Methods

The PSHA input model descried herein was among the models constructed by the GEM Secretariat, and in a systematic way that uses GEM's model-building tools. These tools helped to facilitate model construction, allowing the hazard modeler to apply commonly used methods when developing seismic hazard models. The next subsections describe some of the fundamental concepts and methods used to construct this hazard model.

6.1 Distributed Seismicity Sources

We use the term "distributed seismicity" to indicate earthquakes not clearly attributable to an individual fault source or subduction zone. To model these, we group together seismicity with common characteristics, such as focal mechanism type, strain by the same tectonic forces, rate, or 3D distribution; we then produce source models reflecting these characteristics. Here, we describe two primary source types used to model distributed seismicity.

6.2 Area Sources

Area sources consist of a statistically-determined MFD (Section 10.1) from earthquakes occuring in a volume (usually a polygon, defined by the modeler, with depth limits), with the modelled occurrence rates distributed uniformly (equal *a*- and *b*-values) over an evenly spaced grid, and paired with a hypocenter and focal mechanism. In the OpenQuake Engine, the specified hypocentral depths and focal mechanisms can be probability distributions, or singular metrics.

6.3 Smoothed Seismicity

Smoothed seismicity is modeled similarly to area sources, but rather than using a spatially-homogeneous MFD in each source, the *a*-values vary spatially based on observed seismicity.

GEM has moved away from using traditional area sources, and predominantly models distributed seismicity with an approach that combines area sources with smoothed sesimicity, incorporating methods from Frankel (1995). We define a few source zones with internally consistent tectonics (e.g., up to a few prominent focal mechanism types, reflecting the same tectonic stresses), solve for the Gutenberg-Richter *b*-value, and then smooth the occurring seismicity onto a grid of points. This approach allows us to use larger source zones (and thus more earthquakes to compute a more robust MFDs) while still capturing spatial variability in seismicity rate.

We use the declustered crustal sub-catalogue, applying the *Stepp (1971)* completeness analysis or one based on time-magnitude density plots. Then, from the earthquakes within each source zone, we compute a double truncated Gutenberg-Richter MFD from M=5 to $M_{max,obs}$ + 0.5 (bins of M 0.1), solving for *a*- and *b*-values based on *Weichert (1980)*. We

classify the earthquake probability into weighted depth bins. Lastly, we assign most-likely nodal planes based on crustal earthquake focal mechanisms within the source zone based on the GCMT catalogue.

We compute the smoothed seismicity grid by applying a Gaussian filter to the clipped, declustered catalogue for each source zone, and computing the fraction of spatial seismicity rates at each grid node. These are combined with the zone MFD to compute a grid of point-by-point earthquake occurrence rates.

In areas where we also model fault sources, we prevent double counting by dividing the magnitude occurrence bins between the two source types. If there is overlap (including a buffer around the surface projection of a fault, we cut the MFDs for distributed seismicity at *Mmax*=6.5, and use the same value as *Mmin* for fault MFDs (described in Section 6.5).

6.4 References

Frankel, A. (1995). Mapping seismic hazard in the central and eastern United States. Seismological Research Letters, 66(4), 8-21.

Stepp, J. C. (1971). "An investigation of earthquake risk in the Puget Sound area by the use of the type I distribution of largest extreme". PhD thesis. Pennsylvania State University (cited on pages 9, 25–27).

Weichert, Dieter H. "Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes." Bulletin of the Seismological Society of America70.4 (1980): 1337-1346.

6.5 Characterizing and modelling fault sources

Discrete geologic faults produce the largest earthquakes in the shallow crust. Here we describe the important characteristics of faults, and how we build fault sources for OpenQuake.

Please note that many of the hazard models developed outside of GEM may use different methods than those described here. However, the following is a description of the practices that we at GEM use for the development of our models.

7 Fault geometry and mapping

Fault geometry in map view is constrained through geologic mapping, while the geometry in cross-section view is estimated from geologic cross-section construction or based on the fault kinematics and local focal mechanisms.

In seismic hazard work, almost all faults are given as the geographic coordinates of the fault trace, with an average dip that is used to build a three dimensional representation of the fault surface.

Mapping faults for hazard work is a complicated endeavor; a more in-depth description of this process can be found at the GEM Hazard Blog.

8 Assessing fault activity

Fault activity is assessed through a variety of criteria. The first are instrumental, historical or paleoseismological evidence for earthquakes along the fault; second is strain accumulation that is rapid and localized enough to be measurable through geodetic techniques (GPS, InSAR, optical geodesy); and third is Quaternary geomorphic evidence such as fault scarps, offset streams, and so forth. If the evidence is strong in favor of activity, or a fault is thought to pose a great societal risk, then the fault will be included in the fault source model (with its appropriate uncertainty). If a fault does not display convincing evidence for activity given these criteria, it will be omitted from the fault source model.

8.1 Kinematics

The kinematics of faults, if they are not previously known from earlier studies, are inferred from the topographic and geomorphic expression of the fault, from local focal mechanisms, and from regional geodetic strain information. It is not typical that much confusion or ambiguity exists between normal, strike-slip and reverse faults, since these all have very distinct geomorphic expressions; the more confusing cases tend to be when oblique slip may be present, or when fault kinematics have changed over the millions of years of fault activity, and the topography from the previous tectonic regime is still present. It is more challenging to distinguish between left-slip and right-slip strike-slip faults if no focal mechanisms

or GPS data are available, but it is still generally possible (particularly by looking at bends or stepovers in the fault and the kinematics of faults in these regions).

8.2 Slip Rates

Fault slip rates are generally assessed through formal geologic studies of individual faults through neotectonic and paleoseismic studies, or from geodetic studies of faults or fault networks.

These are complicated and time-intensive investigations, and we at GEM do not generally do this work. Instead, we gather and evaluate the existing literature on faults in a region. There are always many more faults in an area than those that have had formal study, so we will use the rates given in the literature for the faults that have information, and then generalize that information in the context of geodetic strain rate data to infer what the slip rates may be for other structures. For example, faults or fault segments that lie along strike of faults with known slip rates are likely to have similar rates. The regional geodetic strain field provides an overall budget for slip rates within the region: if an area has 6 mm/yr of dextral shear, and the major fault in the area has a known slip rate of 3 mm/yr, then the other faults in the area cannot have dextral slip rates that add up to more than 3 mm/yr. The summed slip rate on faults may be less than the overall geodetic strain, though: some amount of strain may not be distributed on smaller structures or through continuous, plastic deformation of the crust instead of being localized on the major faults in a dataset.

8.3 Seismogenic thickness

The seismogenic thickness of a fault is the total vertical distance between the upper and lower edges of the fault that rupture in a full-length earthquake. It is thought to be a consequence of the frictional stability of the fault materials (and the encompassing crust) at the varying temperature, pressure and fluid contents through the crust. The upper limit of fault slip, the upper seismogenic depth, is usually considered to be the surface of the earth though in some instances (such as subduction zone interfaces) it may be lower. The lower limit is variable based on tectonic environment and the frictional characteristics of the fault materials.

To paint in broad brush strokes, within the continents, normal faults occupy hotter areas of the crust and rupture from (near) the surface to 10-15 km depth; the crust in reverse faulting environments is often colder and the faults will rupture from 15-25 km depth to the surface. Strike-slip faults occupy all environments, so rupture can be from the surface to 10-25 km depth.

Oceanic faults have more variability. Subduction zone interfaces can rupture to near 50 km depth, as they are very cold. Intraplate strike-slip faults can also rupture to >30 km depth, which is well into the mantle in oceanic lithosphere. Hill et al. (2015) report that the 2012 Wharton Basin earthquake east of Indonesia may have ruptured to 50 km. Oceanic spreading ridges and associated transform faults are very hot. Normal faulting does not

produce large earthquakes and the lower depth is probably ~5 km. Associated transforms are slightly cooler and faulting will extend a bit deeper.

The most sound way to assess this is to look at finite fault inversions for the largest earthquakes in a region, if these exist. Lacking this, geodetic techniques may sometimes reveal a value indicating the lower limit of fault locking, although the uncertainties are usually quite large (and underestimated). Similarly, small to microseismicity in a region can give some constraints, but be aware that small earthquakes can occur at much deeper levels in the crust than large ones, because those earthquakes can occur in unfaulted rock that exhibits stick-slip frictional behavior and brittle failure to a greater depth than mature faults with well-developed fault gouge zones and circulating fluids.

9 Building Fault Source Models

Fault source models are usually created by creating three-dimensional fault surfaces and providing information about the style, magnitudes and frequencies of earthquakes that may occur on the fault surface.

9.1 Geometry

Fault geometries are generally created as extrusions of the fault trace (or simplified trace) at a constant dip down to some limit, usually the lower boundary of the seismogenic thickness. Within OpenQuake, these are referred to as 'simple faults'.

In some instances, the geometry of a fault may change sufficiently down-dip that a more complicated representation is warranted. These are known as 'complex faults' in OpenQuake; they are represented by sets of lines of equal depth. OpenQuake then interpolates between these lines to make the fault surface. At GEM, we primarily use complex faults for subduction interfaces.

9.2 Magnitude-Frequency Distributions

The occurrence of earthquakes on a fault is parameterized through magnitude-frequency distributions (MFDs). These give the magnitudes of all the earthquakes on a fault that are to be modeled, and the frequency (or annual probability of occurrence) of earthquakes of the corresponding magnitudes.

The two most common types of MFDs are truncated Gutenberg-Richter distributions, and characteristic distributions. Other MFDs exist that may be hybrids or based on other statistical models, but these are less commonly implemented in seismic hazard analysis. At GEM, we typically use the truncated Gutenberg-Richter distribution, but many other institutions use characteristic fault sources as well. It is still scientifically unknown what the 'true' distribution is and to what degree this changes for different faults, so the choice may come down to pragmatism, familiarity, preference and tradition.

Truncated Gutenberg-Richter distributions are typical Gutenberg-Richter Distributions that are bounded (truncated) by minimum and maximum magnitudes for earthquakes, *M*min and *M*max. Within those bounds, they are parameterized by the *a* and *b* values.

*M*min and *M*max have to be chosen by the fault modeler. *M*min is usually chosen as the smallest earthquake worth modeling in a given model–lowering this value increases the computation time of the model but may not increase the accuracy of the hazard calculations; lower values are more common in smaller-scale studies. *M*max is not so easily determined. The common practice at GEM is to choose it based on the area of a fault surface and the use of an empirical magnitude-area scaling relationship such as that of Wells and Coppersmith (1984) or the more updated Leonard (2012). *M*max then represents a typical full-fault rupture. However, these scaling relationships are statistically-derived and a good amount of variation is present. If there is convincing evidence of larger *M*max on a given fault than the scaling relationship predicts, one should of course choose that larger value.

The *a* and *b* values also need to be determined for each fault. Common practice is to take the *b*-value for a broader tectonic region that encompasses the fault derived from the instrumental seismic catalog, and apply that *b*-value to every fault within the region. There are a few theoretical reasons why this should not be absolutely correct: primarily, the sum of multiple truncated Gutenberg-Richter distributions will not produce a Gutenberg-Richter distribution (in mathematical terminology, the truncated GR distribution is not Levy stable). However, it is exceedingly rare for any empirical constraints on *b*-values for individual faults to exist, so this is a pragmatic compromise.

The *a*-values are chosen so that the total moment release rate adds up to the seismic moment accumulation rate. To make this calculation, the total moment accumulation rate is calculated as the product of the fault area, the shear modulus of the rock encasing the fault, and the fault slip rate. Then, the 'aseismic coefficient', which is the fraction of this total moment accumulation rate that is not released through earthquakes, is subtracted (note that in the case of creeping faults, this moment may never physically be stored in the crust as elastic strain; nevertheless the calculation will be the same). Finally, the *a*-value is chosen so that the total amount of seismic moment released annually (on average) by all of the earthquakes on the fault equals the annual moment accumulation.

Characteristic distributions are narrow distributions that typically represent full-length rupture of a given fault. The *M*max values are chosen through fault scaling relationships or inferences from paleoseismic data. These ruptures may also occur quasi-periodically (as opposed to uniformly randomly) though this sort of time-dependence is not often used at GEM.

10 References

Hill, Emma M., et al. "The 2012 Mw 8.6 Wharton Basin sequence: A cascade of great earthquakes generated by near-orthogonal, young, oceanic mantle faults."

Journal of Geophysical Research: Solid Earth 120.5 (2015): 3723-3747. https://doi-org.www2.lib.ku.edu/10.1002/2014JB011703

Leonard, Mark. "Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release." Bulletin of the Seismological Society of America 102.6 (2012): 2797-2797. https://doi-org.www2.lib.ku.edu/10.1785/0120120249

Wells, Donald L., and Kevin J. Coppersmith. "New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement." Bulletin of the seismological Society of America 84.4 (1994): 974-1002.

11 Types of MFDs

In probabilistic seismic hazard analysis (PSHA), source models require a defined occurrence rate for earthquakes of each considered magnitude, e.g., a magnitude-frequency distribution (MFD). These rates are determined either by statistically analysing the observed seismicity over instrumental and historic time scales, or-for well characterized sources—by using the fault dimensions and slip rates to model recurrence.

Regional models built by GEM use the following common approaches to characterize seismicity rates.

11.1 Gutenberg-Richter

The Gutenberg-Richter MFD allows earthquake sources to generate earthquakes of different magnitudes. *Gutenberg and Richter (1944)* were the first to develop a functional form for the relationship between earthquake magnitude and occurrence rate, resolving a negative exponential distribution:

$$logN = a - bm \tag{1}$$

(2)

where *N* is the annual rate of earthquakes with $M > m$, *a* is the rate of all earthquakes, and *b* is the relative distribution of earthquakes among magnitudes. A higher *b*-value indicates a larger proportion of seismic moment released by small earthquakes. *a* and *b* are resolved from the available observations. Usually, *b* is close to 1.0.

11.1.1 Truncated Gutenberg-Richter

A traditional Gutenberg-Richter MFD allows for earthquakes of any magnitude, but in reality, the source in question may not be capable of producing earthquakes beyond a certain threshold. For example, fault dimensions physically limit earthquake magnitude, or the observed earthquake magnitudes saturate. To account for these constraints, a truncated MFD is used to specify a maximum magnitude (*Mmax*), simply by cutting the MFD at this magnitude. The MFD is additionally cut at a minimum magnitude ("double-truncated"), below which earthquakes are not contributing to the hazard in ways significant to engineering.

Truncated Gutenberg-Richter MFDs are commonly used in hazard models build by the GEM Secretariat. Where MFDs are produced for a source zone, such as for distributed or inslab seismicity, the upper magnitude is usually determined by adding a delta value (e.g., *MW*0.5) to *Mmax* in the earthquake catalogue or subcatalogue used to produce the MFD. This is

based on the premise that the observation period is too short to have experienced a true *Mmax* earthquake.

GEM models typically use the methodology of Weichert (1985) to compute double-truncated Gutenberg-Richter MFDs for seismic source zones, which allows for the use of different observation periods for different earthquake magnitudes (e.g., a completeness threshold).

If a seismicity distribution is not explicitly available, an MFD of this form can also be computed from a seismic moment budget using strain rates, fault dimensions, and assumed magnitude ranges and *b*-values. For models built internally by GEM, we apply this to faults with available slip rates. This methodology is described in Section 6.5.

11.2 Characteristic

Some sources do not produce earthquakes that follow the Gutenberg-Richter distribution, but instead tend to host earthquakes of nearly the same magnitude, e.g., a characteristic earthquake. In this case, an earthquake with a moderate to high magnitude occurs more frequently than would be suggested by a Gutenberg-Richter MFD. For sources of this type, the MFD reveals more frequent occurrences concentrated around the most-likely/characteristic magnitude earthquake, for example using a boxcar or Gaussian distribution (e.g., *Youngs and Coppersmith, 1985*, or *Lomnitz-Adler and Lomnitz, 1979*).

Though the *Youngs and Coppersmith* (*1985*) MFD is technically a hybrid MFD, incorporating both a characteristic component and a Gutenberg-Richter component at lower magnitudes, it is typically often categorized as a characteristic MFD. GEM uses this MFD in a few models built in-house, such as the Philippines (Section **??**) model, where sensitivity testing indicated that it produced a better fit to the regional seismicity than a double-truncated GR for crustal faults.

11.3 Hybrid types

Some subduction interface source models built by the GEM secretariat use a hybrid approach that combines the Gutenberg-Richter MFD with a characteristic MFD. The latter approach derives a double truncated Gaussian distribution to model occurrence of the maximum magnitude (*M*max) earthquake that an interface segment can theoretically support (herein called the "characteristic earthquake").

The magnitude and occurrence rate of the characteristic earthquake for an interface segment are based on the fault area (e.g., from the complex fault output by the Subduction Toolkit, see Section 16.1), the convergence rate, and a seismic coupling coefficient. We choose between three recent scaling relationships for subduction interfaces that compute magnitude from fault area: *Strasser et al.* (*2010*), *Allen and Hayes* (*2017*), and *Thingbaijam and Mai* (*2017*). We use published convergence rates and seismic coupling coefficients to determine the time needed to accumulate enough strain for the characteristic earthquake.

The coupling parameter is often challenging, in large part due to the scarcity of land and thus GPS measurements in close proximity to subduction zones. Where no other model is available, we take values from *Heuret et al.* (*2011*) or *Scholz and Campos* (*2012*), but cautiously, as many sometimes these values are suspiciously low (e.g., <0.1 where instrumentally recorded earthquakes *M*>8.0 have occurred.)

The characteristic MFD is combined with the Gutenberg-Richter MFD into a hybrid MFD by finding the intersection point of the two MFDs, and taking the Gutenberg-Richter occurrence rate below the intersection magnitude, and the characteristic rate above that magnitude.

11.4 References

Allen, T. I., & Hayes, G. P. (2017). Alternative rupture-scaling relationships for subduction interface and other offshore environments. *Bulletin of the Seismological Society of America*, 107(3), 1240-1253.

Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4), 185-188.

Heuret, A., Lallemand, S., Funiciello, F., Piromallo, C., & Faccenna, C. (2011). Physical characteristics of subduction interface type seismogenic zones revisited. *Geochemistry, Geophysics, Geosystems*, 12(1).

Lomnitz-Adler, J., & Lomnitz, C. (1979). A modified form of the Gutenberg-Richter magnitude-frequency relation. *Bulletin of the Seismological Society of America*, 69(4), 1209-1214.

Scholz, C. H., & Campos, J. (2012). The seismic coupling of subduction zones revisited. Journal of Geophysical Research: *Solid Earth*, 117(B5).

Thingbaijam, K. K. S., Martin Mai, P., & Goda, K. (2017). New empirical earthquake source-scaling laws. *Bulletin of the Seismological Society of America*, 107(5), 2225-2246.

Strasser, F. O., Arango, M. C., & Bommer, J. J. (2010). Scaling of the source dimensions of interface and intraslab subduction-zone earthquakes with moment magnitude. *Seismological Research Letters*, 81(6), 941-950.

11.5 Characterizating and processing seismic catalogues

Much of PSHA depends on the assumption that future seismicity will occur near observed past seismicity, and at rates that can be approximated by empirical or physical models. Thus, the early steps in PSHA include compiling and processing an earthquake catalogue. Beyond collecting instrumental and historic earthquake records, catalogues must be homogenized (expressed in uniform units), declustered (devoid of aftershocks and foreshocks), and filtered for completeness. The assumptions and uncertainties in the catalogue should be well understood by the modeler.

Most source types used in hazard models built by the GEM Secretariat use magnitude-frequency distributions (MFDs, Section 10.1) based on seismicity. Together with ground motion prediction equations (GMPEs), MFDs govern the computed hazard levels for time frames of interest, and so their robust calculation - and thus careful preparation of the input catalogue - is critical.

Here, we describe the ISC-GEM extended catalogue (*Weatherill et al., 2016*), which contributes the majority of earthquakes used in hazard models built internally by GEM; the workflow for combining other earthquake records with the ISC-GEM catalogue; and the remaining steps to prepare the catalogue for rate and spatial analysis. We emphasize that while most of these steps are routinely applied outside of GEM models, the following explanations only account for our own best practices.

12 The ISC-GEM catalogue

The ISC-GEM catalogue is a compilation of earthquake bulletins for seismicity occurring in the range 1900-2015. This catalogue sources records from numerous agencies to include the record deemed most accurate for each event, ensuring that no duplicates are included, and magnitudes are homogenized to *MW*. The most recent catalogue updates were completed by *Weatherill et al. (2016)* using the GEM Catalogue Toolkit, totaling 562 840 earthquakes with *MW* 2.0 to 9.6, and producing what is herein called the ISC-GEM extended catalogue. This current version is motivated by initiatives to improve regional and global scale seismicity analyses, hazard and otherwise.

Regional models developed by the GEM Secretariat use the ISC-GEM extended catalogue, augmented by data from local agencies when possible.

13 GEM Historical Earthquake Catalogue

The GEM Historical Earthquake Catalogue (*Albini et al., 2013*), includes large earthquakes (*M*>7) from before the instrumental period (1000-1903) that have been carefully reviewed to estimate a location and magnitude. The completeness of this catalogue is highly variable across the globe, and depends on how long each location has been inhabited, and the availability and quality of documentation on earthquakes occurring in this period.

14 Processing of seismicity catalogues

14.1 Catalogue homogenization

In order to use the bulletins from multiple agencies together in statistical analyses, records must be homogenized to meet the same criteria, e.g., to use the same measure of magnitude. Usually, moment magnitude (*MW*) is selected, since it does not saturate at high magnitudes. Thus, magnitudes reported in other scales must be converted. When possible, this is done using empirical relations developed for independent local datasets, but relies on global relations when too few calibration events are available.

The homogenization methodology used to build the ISC-GEM extended catalogue is described in detail in *Weatherill et al. (2016)*.

14.2 Completeness analysis

Catalogue completeness analysis accounts for the variability in instrumentation coverage throughout the catalogue duration, admitting that any catalogue is missing earthquakes beneath a magnitude threshold. This type of filtering prevents rate analysis of an incomplete catalogue - a modeling mistake that will propagate into hazard estimates. Importantly, completeness analysis must be applied to a declustered catalogue as to not confuse dependent earthquakes (such as aftershocks) with magnitude completeness.

The completeness algorithms that are applicable to *any* instrumental catalogue must depend on properties of the earthquakes, and not the stations, thus focusing on the statistics of the catalogue sample rather than the probability that a station at a known position would record an earthquake. The most common algorithmic method is by Stepp (1971), which compares the observed rate of seismicity to a predicted Poissonian rate for each magnitude, and returns a spatially constant table of time-variable magnitude thresholds. Importantly, the validity of this methodology is subject to the judgement of the user.

The Stepp (1971) is implemented in the OpenQuake Engine, and used in some steps of the modeling procedure for hazard models built by the GEM Secretariat. In other cases, we determine the completeness manually from 3D histograms that count earthquakes for magnitude-time bins, visually identifying the timings at which the occurrences rates stabilize.

14.3 Declustering

Catalogue declustering is applied in order to isolate mainshock earthquakes - that is, earthquakes that occur independently of each other - from a complete catalogue. The resulting declustered catalogue should therefore reflect the Poissonian rate at which earthquakes occur within a greater tectonic region. PSHA aims to model the hazard from sesimicity occurring at this background Poissonian rate.

Declustering algorithms identify mainshocks by comparing individual earthquakes to the "cluster" of earthquakes that occurred within a given proximity and time to that earthquake, choosing the largest for a given set of magnitude-dependent "triggering windows". The theory of declustering algorithms is described in detail in *Stiphout et. al., 2012*. The OpenQuake Hazard Modeler's Toolkit provides three different windowing options: the original implementation of Gardner and Knopoff (1974), and additionally the configurations of Uhrhammer (1986) and Gruenthal (see Stiphout et al., 2012).

In subduction zones or other complex environments, we first classify the seismicity by tectonic domain (described below), and then decluster groups of domains within which we expect seismicity to interact (i.e., interface mainshocks can trigger crustal aftershocks), and then separate the deemed mainshocks into subcatalogues based on their tectonic classification. We typically use two groups: crustal, interface, and shallow slab seismicity (that beneath the interface but with intraslab mechanisms); and deep intraslab seismicity. The declustering algorithm comparing epicentral (*not* hypocentral) proximities, and thus, declustering by groups is crucial for seismicity within slab-type volumes.

15 Classification of seismicity

The workflow used by GEM to construct seismic source models in complex tectonic regions is dependent on the use of classified seismicity, that is, the assignment of each earthquake to a tectonic domain. Separating earthquakes in this manner allows us to compute MFDs from only the seismicity occurring within a delineated domain, thus more accurately characterizing individual seismic sources or source zones. For example, in subduction zones, we separate earthquakes occurring on the interface itself from those within the downgoing slab or the overriding plate. This allows us to model the hazard from these source types using the appropriate GMPEs.

At GEM, we classify seismicity using an procedure with similar theory to *Zhao et al., (2015)* and *Garcia et al., (2012)*, which assigns earthquakes to tectonic domains defined by the modeler. In subduction zones, earthquakes are usually categorized as crustal, interface, or intraslab based on hypocentral proximity to the Moho, and the interface and slab-top complex surfaces defined by the Subduction Toolkit (Section 16.1). Where subduction zones are modeled as segmented interfaces or slabs, the domains are divided accordingly. Each tectonic domain is defined by a surface and a buffer region based on general characteristics of the corresponding cross sections. The modeler provides a tectonic hierarchy that chooses among multiple assignments for earthquakes occurring within overlapping buffers of two or more domains. Usually, we specify interface superseding intraslab, and intraslab superseding crustal. Earthquakes that do not correspond to any of the defined domains are deemed "unclassified".

The classification routine includes workarounds to correct some common misclassifications, such as to seclude dominant groups of earthquakes beneath a polygon (e.g., volcanic events); to classify large magnitude earthquakes from historic catalogues only by epicenter;

and the ability to manually classify earthquakes by their event IDs.

16 References

Albini P., Musson R.M.W., Rovida A., Locati M., Gomez Capera A.A., and Viganò D. (2014). The Global Earthquake History. Earthquake Spectra, May 2014, Vol.30, No.2, pp. 607-624. http://doi.org/10.1193/122013EQS297

Gardner, J. K., and Leon Knopoff. "Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?." Bulletin of the Seismological Society of America64.5 (1974): 1363-1367.

Stepp, J. C. (1971). "An investigation of earthquake risk in the Puget Sound area by the use of the type I distribution of largest extreme". PhD thesis. Pennsylvania State University (cited on pages 9, 25–27).

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. URL: http://www.corssa.org.

Uhrhammer, Robert A. "SEISMICITY RECORD, M> 2.5, FOR THE CENTRAL." Clement F. Shearer (1985): 199.

Weatherill, G. A., M. Pagani, and J. Garcia. "Exploring earthquake databases for the creation of magnitude-homogeneous catalogues: tools for application on a regional and global scale." Geophysical Journal International 206.3 (2016): 1652-1676.

16.1 Characterizing and modelling subduction sources

Subduction zones are plate margins where one tectonic plate 'subducts' or is thrust beneath another plate. These zones produce most of the seismicity on Earth. The zones can be complex, producing earthquakes at the interface or 'megathrust' fault between the plates, in the downgoing plate or 'slab', and in the deforming region at the margin of the upper, overriding plate. For hazard models produced by the GEM Secretariat, the plate interface and the subducting slab are characterized and modeled with subduction-specific tools we have developed alongside our modeling efforts, while the deformation within the upper plate is modeled as part of the active shallow crust.

17 Subduction interface

Among PSHA models, various source model approaches are used to model interface seismicity. Models produced by GEM use OpenQuake complex faults (surfaces with complex geometry) to account for subduction interface seismicity, and float all possible ruptures within specified magnitude limits and have a given rupture aspect ratio across the meshed surface. In some cases, we segment the surfaces along-strike to define firm barriers to rupture or capture changes in subduction characteristics. We use two predominant approaches to compute magnitude-frequency distributions (MFDs) and maximum magnitudes of the interface segments. Both use recorded instrumental (and sometimes historical) seismicity that can be attributed to the respective interface segment (classified using the methodolgy described in Section 15), fitting a Gutenberg-Richter (a negative exponential) distribution to the seismicity. One approach also includes a characteristic component, computed from the area of the interface surface, the local convergence rate, and the degree of seismic locking (a seismic coupling coefficient). MFD construction is explained in detail in Section 10.1.

18 Slab

Hazard models built by the GEM Secretariat account for intraslab seismicity using nonparametric ruptures (sources with predefined geometry) that fit within a slab volume of uniform thickness. The ruptures correspond to virtual faults within a meshed approximation of the slab volume, and forces ruptures to fit within the slab. Like the interface, the slab volume can be segmented, however here, boundaries only seldom indicate barrier to rupture (such as at a slab tear) and are more commonly used to reflect change in seismicity rate. For each slab segment, we compute a single Gutenberg-Richter MFD from the slab segment subcatalogues produced during tectonic classification (Section 15), assuming constant rates throughout each segment. Currently, moment rates are distributed uniformly among the computed ruptures, but future development will include a smoothing component.

19 The Subduction Toolkit: building the geometry of the interface surface and slab volume

Alongside the PSHA models that incorporate subduction zones, GEM has developed the Subduction Toolkit, which uses an interactive workflow to build the subdction interface and slab top geometry, an integral step in producing the subduction source model.

The subduction geometries are based on trench axes from the GEM Active Faults Database along with several geophysical datasets and models. The toolkit projects swaths of geophysical data onto cross sections along a trench axis, which are used to guide depth picking for the interface and slab upper surface. These depth profiles are then stitched together to form OpenQuake complex fault surfaces, which are used as reference frames for catalogue tectonic classification (Section 15), and for defining subduction source geometry (described above).

The data plotted on the cross sections is meant to illuminate the subsurface subduction structures and tectonic processes that contribute to seismic hazard (e.g., Figure 6). The most commonly used data include:

- hypocenters from ISC-GEM catalogue (*Weatherill et al., 2016*)
- centroid moment tensors (CMTs) from the Global CMT project (*Dziewonski et al., 1981*; *Ekstrom et al., 2012*)
- Moho depth estimates from Lithos1.0 (*Pasyanos et al., 2014*) and Crust1.0 (*Laske et al., 2013*)
- Slab depth estimates from Slab1.0 (*Hayes et al., 2011*) and Slab2.0 (*Hayes et al., 2018*)
- Shuttle Radar Topography Mission (SRTM) topography (*Farr, 2007*)
- General Bathymetric Charts of the Ocean (GEBCO) bathymetry (*Weatherall et al., 2015*)
- Volcano locations

Initially, the cross sections are automatically generated at a specified increment along the trench axis that balances data density with resolution, with azimuths perpendicular the trench. The cross section origins and azimuths can then be adjusted manually, and additional cross sections added where necessary.

The final depth profiles (or a subset) are stitched together to form an OpenQuake complex fault surface. The Toolkit allows for the full extent of the profiles to be considered in subsequent steps, or a depth range can be defined. We use these capability to separate the subduction interface from the deeper slab, and to segment the surfaces along strike (see above).

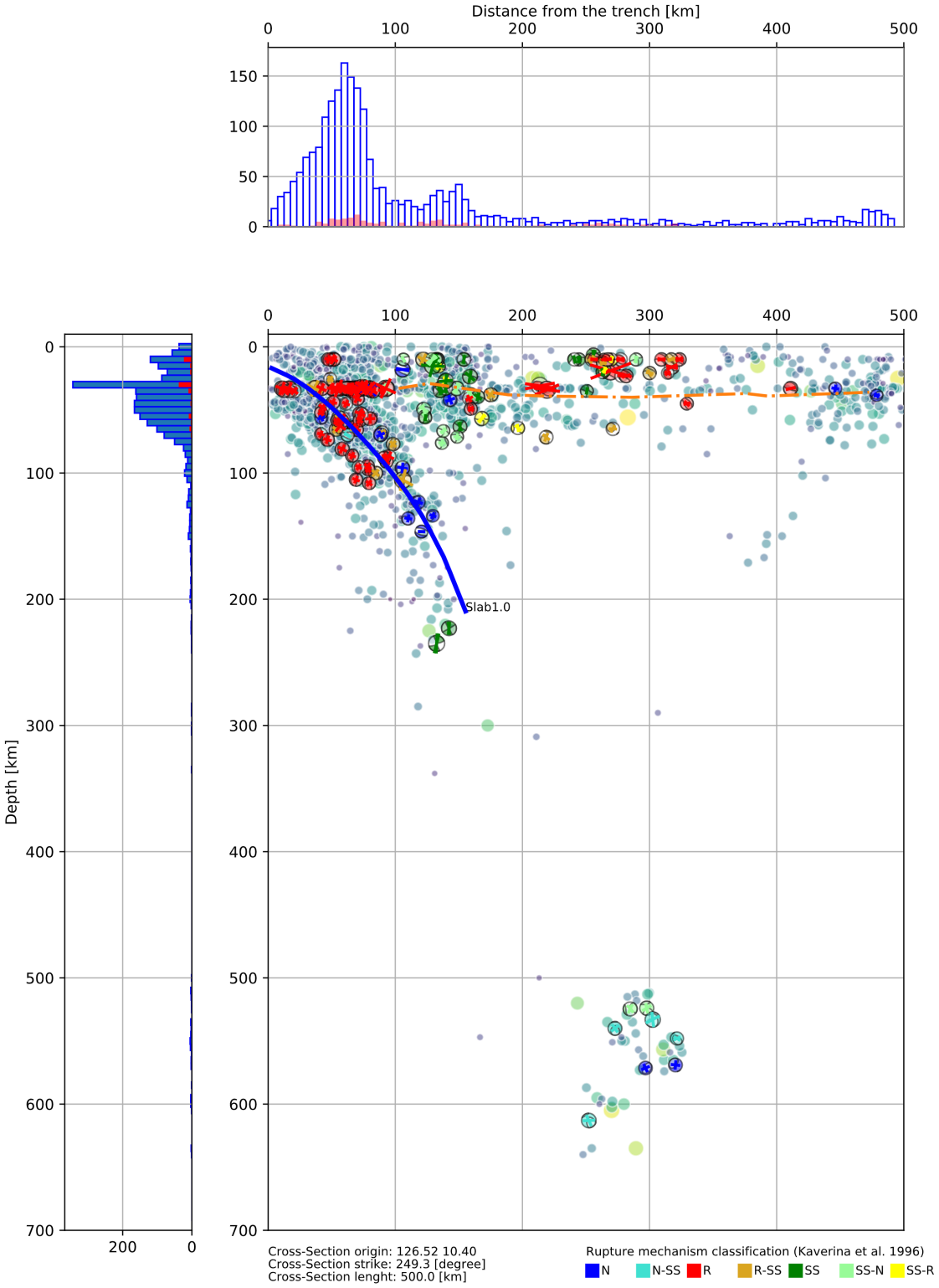


Figure 6 *– Example cross-section of a subduction zone from the Philippines*

20 References

Dziewonski, A. M., T.-A. Chou and J. H. Woodhouse, Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, J. Geophys. Res., 86, 2825-2852, 1981. doi:10.1029/JB086iB04p02825

Ekström, G., M. Nettles, and A. M. Dziewonski, The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, Phys. Earth Planet. Inter., 200-201, 1-9, 2012. doi:10.1016/j.pepi.2012.04.002

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L. and Seal, D., 2007. The shuttle radar topography mission. Reviews of geophysics, 45(2), doi:10.1029/2005RG000183.

Hayes, Gavin P., David J. Wald, and Rebecca L. Johnson. "Slab1. 0: A three-dimensional model of global subduction zone geometries." Journal of Geophysical Research: Solid Earth 117.B1 (2012).

Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. Science, 362(6410), 58-61.

Laske, Gabi, et al. "Update on CRUST1. 0—A 1-degree global model of Earth's crust." Geophys. Res. Abstr. Vol. 15. Vienna, Austria: EGU General Assembly, 2013.

Pasyanos, Michael E., et al. "LITHO1. 0: An updated crust and lithospheric model of the Earth." Journal of Geophysical Research: Solid Earth 119.3 (2014): 2153-2173.

Weatherall, P., K. M. Marks, M. Jakobsson, T. Schmitt, S. Tani, J. E. Arndt, M. Rovere, D. Chayes, V. Ferrini, and R. Wigley (2015), A new digital bathymetric model of the world's oceans, Earth and Space Science, 2, 331–345, doi:10.1002/2015EA000107.

Weatherill, G. A., M. Pagani, and J. Garcia. "Exploring earthquake databases for the creation of magnitude-homogeneous catalogues: tools for application on a regional and global scale." Geophysical Journal International 206.3 (2016): 1652-1676.