

# PSHA of the southern Pacific Islands

K.L. Johnson<sup>iD</sup>, M. Pagani<sup>iD</sup> and R.H. Styron<sup>iD</sup>

Global Earthquake Model Foundation, Pavia, Italy. E-mail: [kendra.johnson@globalquakemodel.org](mailto:kendra.johnson@globalquakemodel.org)

Accepted 2020 November 4. Received 2020 November 2; in original form 2020 November 21

## SUMMARY

The southern Pacific Islands region is highly seismically active and includes earthquakes from four major subduction systems, seafloor fracture zones and transform faults, and other sources of crustal seismicity. Since 1900, the area has experienced  $>350$  earthquakes of  $M > 7.0$ , including 11 of  $M \geq 8.0$ . Given the elevated threat of earthquakes, several probabilistic seismic hazard analyses have been published for this region or encompassed subregions; however, those that are publicly accessible do not provide complete coverage of the region using homogeneous methodologies. Here, we present a probabilistic seismic hazard model for the southern Pacific Islands that comprehensively covers the Solomon Islands in the northwest to the Tonga islands in the southeast. The seismic source model accounts for active shallow crustal seismicity with seafloor faults and gridded smoothed seismicity, subduction interfaces using faults with geometries defined based on geophysical data sets and models, and intraslab seismicity modelled by a set of ruptures that occupy the slab volume. Each source type is assigned occurrence rates based on subcatalogues classified to each respective tectonic context. Subduction interface and crustal fault occurrence rates also incorporate a tectonic component based on their respective characteristic earthquakes. We demonstrate the use of non-standard magnitude–frequency distributions to reproduce the observed occurrence rates. For subduction interface sources, we use various versions of the source model to account for epistemic uncertainty in factors impacting the maximum magnitude earthquake permissible by each source, varying the interface lower depth and segmentation as well as the magnitude scaling relationship used to compute the maximum magnitude earthquake and subsequently its occurrence rate. The ground motion characterization uses a logic tree that weights three ground motion prediction equations for each tectonic region. We compute hazard maps for 10 and 2 per cent probability of exceedance in 50 yr on rock sites, discussing the regional distribution of peak ground acceleration and spectral acceleration with a period of 1.0 s, honing in on the hazard curves and uniform hazard spectra of several capital or populous cities and drawing comparisons to other recent hazard models. The results reveal that the most hazardous landmasses are the island chains closest to subduction trenches, as well as localized areas with high rates of seismicity occurring in active shallow crust. We use seismic hazard disaggregation to demonstrate that at selected cities located above subduction zones, the PGA with 10 per cent probability of exceedance in 50 yr is controlled by  $M_w > 7.0$  subduction interface and intraslab earthquakes, while at cities far from subduction zones,  $M_w < 6.5$  crustal earthquakes contribute most. The model is used for southern Pacific Islands coverage in the Global Earthquake Model Global Hazard Mosaic.

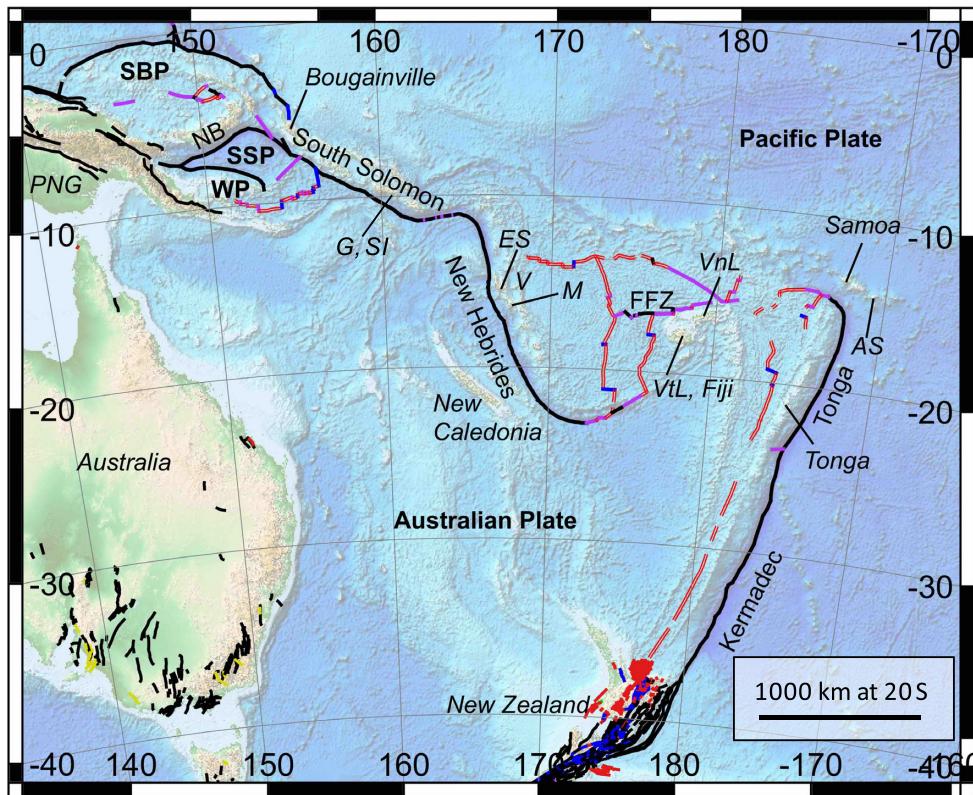
**Key words:** Earthquake Hazards; Pacific Ocean; Seismic cycle.

## 1 INTRODUCTION

The southern Pacific Islands region is tectonically complex and seismically very active, with  $\sim 350$  earthquakes  $M \geq 7.0$  occurring since 1900. Several island countries and territories are subject to this seismic hazard, the most populated being Fiji, the Solomon Islands and Vanuatu. The greatest hazard posed by these earthquakes is

triggered tsunamis, but past earthquakes have also caused shaking-related damage and fatalities (NGDC/WDS 2020).

Most of the regional seismic hazard is attributable to interface and intraslab earthquakes along the  $>6000$  km of subduction zones (Fig. 1). At the  $\sim$ north–south trending Kermadec and Tonga trenches, the Pacific plate subducts beneath the Australian plate, converging at  $\sim 80$  mm  $\text{yr}^{-1}$  in the south and increasing to



**Figure 1.** Overview of the Pacific Islands region. Solid lines show faults from the Global Active Faults Database (Styron & Pagani 2020), coloured by slip type. Red: spreading ridges. Purple: sinistral transform. Blue: dextral transform. Black: reverse and subduction thrust. Labels: bold font indicates tectonic plates, italic font geographic locations, and standard font subduction trenches. AS: American Samoa. ES: Espiritu Santo. FFZ: Fiji Fracture Zone. G: Guadacanal. M: Malekula. NB: New Britain trench. NC: New Caledonia SBP: South Bismarck Plate. SI: Solomon Islands. SSP: South Solomon Plate. V: Vanuatu. VnL: Vanuatu. VtL: Vitu Levu. WP: Woodlark Plate.

~220 mm yr<sup>-1</sup> in the north (Bird 2003). At the point of peak convergence—the northern extent of the Tonga Trench—the strike of the plate boundary changes to become approximately parallel to the plate motion. West of here, along a semi-continuous network of trenches, the Australian plate subducts beneath the Pacific plate at ~35–120 mm yr<sup>-1</sup> on the New Hebrides trench (Calmant *et al.* 2003) and ~100 mm yr<sup>-1</sup> on the South Solomon trench (Wallace *et al.* 2005). Along the westernmost Solomon trench, the Solomon Sea plate subducts beneath the Pacific plate, and furthest west, the Solomon Sea plate subducts beneath the South Bismarck plate at New Britain ~50–130 mm yr<sup>-1</sup> on the New Britain trench (Bird 2003).

In addition to subduction earthquakes, seismicity occurs in the rapidly deforming Fiji Platform due to backarc spreading and clockwise rotation along left-lateral fracture zones (Rahiman & Pettinga 2008). Some large earthquakes ( $M > 7$ ) also occur in the outer rise, and beyond here we observe widespread distributed shallow seismicity.

Here, we present a probabilistic seismic hazard analysis (PSHA) for the area spanning east–west from eastern Papua New Guinea to American Samoa and Niue, and north–south from the Solomon Islands to the southern tip of the Kermadec trench. We use the classical hazard analysis procedure, as first established by Cornell (1968), and the OpenQuake Engine (Pagani *et al.* 2014), which implements the classical approach following the formulation by Field *et al.* (2003). For sites distributed across the southern Pacific Islands, we calculate hazard curves that reveal the probability at which various ground motion levels will be exceeded during a time

period of interest. In addition to interpreting the hazard curves for a range of ground motion intensity measures types, we extract hazard maps and uniform hazard spectra from the computed hazard curves for probabilities of exceedance of 10 and 2 per cent in 50 yr on reference rock conditions. Finally, we disaggregate the hazard for large cities in the region, evaluating the significance of each tectonic region type and magnitude-distance pair at the sites of interest.

A first version of the model presented herein was included in the Global Earthquake Model (GEM) Foundation Hazard Mosaic v. 2018.1 (Pagani *et al.* 2020a). The model has since undergone internal revisions, and the seismic source characterization has been expanded to include epistemic uncertainties for several parameters used to constrain subduction sources. Here, we describe the seismic source characterization, which comprises one source model to represent crustal seismicity and multiple source models to represent subduction sources, which are weighted in a seismic source logic tree. We explain the methodology used to construct the seismic source models, and indicate the variations in the model parameters that facilitate the alternative models. The modelled epistemic uncertainties provide a first indication of the overall variability of hazard results. We plan to further improve the model's consideration of epistemic uncertainties in future versions of this model, and to complete it with a thorough sensitivity analysis.

## 2 EXISTING PSHA MODELS

Several models cover portions of the southern Pacific Islands region, including the most recent models by Suckale & Grünthal (2009),

Rong *et al.* (2010) and Petersen *et al.* (2012). However, the utility of these models is limited either because they are not publicly available, or do not homogeneously cover the region of interest. Still, these models are instructional for integrating the current understanding of crustal faults; defining source areas within the region; and comparing output, both as computed seismicity rates and ground motion exceedance values. Here, we review three recent and publicly available Pacific Islands models with overlapping coverage of the model developed herein.

Petersen *et al.* (2012), under Federal Emergency Management Agency (FEMA) directives, performed a hazard analysis for the eastern part of the region investigated here, covering the U.S. territory of American Samoa and neighbouring islands as far west as Vanuatu. They used earthquake catalogues to compute hazard from gridded seismicity and subduction zone fault models. They compare ground motions recorded in the Pacific Basin to ground motion models for subduction zone and deep earthquakes, and for shallow earthquakes, and use the relative data-model fits to define a logic tree of ground motion models. They focus on spectral accelerations at 1 Hz and 5 Hz with 2 per cent PoE in 50 yr, noting high variability across the region (1 Hz accelerations from 0.12–0.89 g and 5 Hz accelerations from 0.32–2.8 g).

Suckale & Grünthal (2009) produced a PSHA for the Vanuatu Arc, encompassing the islands that parallel the New Hebrides trench. This area overlaps with the western extent of the Petersen *et al.* (2012) model, but hones in on the single, sublinear trench. The model captures an apparent seismic gap at the approximate centre of the trench, where the buoyant d'Entrecasteaux Ridge subducts beneath the Pacific Plate, convergence rates drop abruptly, and observed seismicity rates are low. They define three different area source models for three depth slices (0–60 km, 60–120 km, and 120–200 km), and divide the seismicity into crustal, interface, and intraslab units. They use a logic tree of ground motion models to compute onshore PGA with 10 per cent PoE in 50 yr on rock ranging from 0.51–0.65 g.

The model by Rong *et al.* (2010) covers the full southern Pacific Islands region, extending westward across Papua New Guinea. Like the model developed herein, Rong *et al.* (2010) aimed to produce a model that uses the same methodology and thinking across the complete region of interest; however, the model is not publicly available. The seismic source model includes smoothed seismicity across source zones (some adopted from former studies) and characteristic subduction interface faults using tectonics and seismicity to determine recurrence parameters, as well as crustal faults based on kinematic modelling of GPS data. Using a logic tree of ground motion models, they compute PGA at rock sites with a 10 per cent PoE in 50 yr ranging from <0.05–0.85 g, with the highest reported hazard in the Solomon Islands.

Ghasemi *et al.* (2016) produced a PSHA for Papua New Guinea; this model partially overlaps with the seismic sources in the Pacific Islands model presented herein. The Ghasemi *et al.* (2016) source model uses a weighted average between an area source model and smoothed seismicity, with a fault to represent the New Britain subduction interface. The area sources are divided into two depth zones (shallow crustal and subduction), the first of which is delineated by major tectonic features/boundaries, and then refined by existing tectonic models, less significant plate boundaries and faults, and historic seismicity. The deep area sources were modelled after Slab 1.0 (Hayes *et al.* 2012). Using a logic tree of ground motion models, they compute PGA with 10 per cent PoE in 50 yr for bedrock sites, finding peak values of ~1 g.

Fiji is additionally covered by the model by Jones and Australian Geological Survey Organisation (1998); however, we exclude this from the discussion. The Jones and Australian Geological Survey Organisation (1998) model is based on a catalogue ending in 1990 and ground motion attenuation relationships from 1982, and has thus been superseded by the model by Petersen *et al.* (2012).

### 3 DATA SETS AND DERIVATIVES

Here, we describe the data sets and models we used to create the Pacific Islands seismic source model. We began with an earthquake hypocentral catalogue, focal mechanism catalogue, active faults data set, and geophysical crustal structure models, and produced a declustered catalogue with earthquakes classified to tectonic regions, as well as a detailed model of the subduction structure. These data sets and their derivatives became the basis of the seismic source characterization.

We use the magnitude-homogenized ISC-GEM extended catalogue by Weatherill *et al.* (2016), which includes earthquakes in the date range 1900–2014. Clipped to the Pacific Islands region, we keep ~110 000 earthquakes  $M_w > 2.82$  in the range 45°S to 4°N, 145°E to 160°W. We start from a complete catalogue, but use a declustered catalogue classified to tectonic regions and filtered for magnitude completeness—prepared using the Model Building Toolkit (MBTK; <https://github.com/GEMScienceTools/oq-mbtk>)—in later processing steps (Table 2). We also use the Global Centroid Moment Tensor (GCMT) focal mechanisms from 1976 to 2015 (Dziewonski *et al.* 1981; Ekström *et al.* 2012), and crustal faults with slip rates from the GEM Global Active Faults Database (GAF-DB; Styron & Pagani 2020).

From a modelling perspective, one challenge presented by the catalogues is the prevalence of hypocentres fixed to depths of 10, 33, and 35 km. We consider the impact of these depth errors throughout the source characterization.

#### 3.1 Catalogue classification

As a first step in constructing the seismic source models, we use a tectonic classification routine to separate the catalogue into crustal, interface, and intraplate earthquakes, creating a classified subcatalogue for each tectonic region of the source model. We follow the procedure by Pagani *et al.* (2020b), which classifies hypocentres based on their relative positions to reference surfaces that represent the subduction interface, subducting slab top, crustal extent, or other relevant reference frame. For each tectonic unit, the classification configuration considers user-defined buffer limits with respect to the indicated surface. A separate classification is necessary for each subduction geometry considered in the seismic source characterization; this is discussed more in Sections 3.1.1 and 4.6.

For the crustal reference surface, we use the Moho depth model of Litho1.0 (Pasayanos *et al.* 2014). The subduction surfaces are built as part of the procedure from Pagani *et al.* (2020b), discussed next, which consists of picking depth profiles to the interface from cross-sections through the subduction zones that show hypocentre positions, focal mechanisms, and guidance from the following geophysical models: Slab 1.0 (Hayes *et al.* 2012), Crust1.0 (Laske *et al.* 2013), Shuttle Radar Topography Mission (SRTM) topography (Farr *et al.* 2007), General Bathymetric Charts of the Ocean (GEBCO; GEBCO 2008) and large-scale tectonic regionalization (e.g. Chen *et al.* 2018); we note that all the datasets and models used here are openly available. The resulting profiles are connected to

**Table 1.** Fault parameters for modelled faults. ID corresponds to the labels in Supporting Information Fig. S0. GR: Gutenberg-Richter MFD. YC: Youngs & Coppersmith (1985) MFD. For GR MFDs,  $M_{\text{char}}$  is used to determine  $M_{\text{max}}$ , while for YC MFDs it constrains the characteristic component.

ID	Slip type	Slip rate ( $\text{mm yr}^{-1}$ )	Dip, dip direction	$M_{\text{char}}$	MFD type
1	Sinistral	140.4	90	7.15	GR
2	Sinistral	10.1	90	7.05	GR
3	Sinistral	13.0	90	7.25	GR
12	Sinistral	38.0	90	6.85	GR
19	Sinistral	35.2	90	7.15	GR
5	Dextral	40.2	90	7.05	GR
21	Subduction thrust	13.0	40, S	8.65	YC
22	Dextral	10.1	90	6.85	GR
23	Subduction thrust	8.5	40, SW	8.76	YC
26	Reverse	17.0	40	6.95	GR
28	Reverse	21.8	40	7.15	GR
29	Sinistral	124.6	90	7.15	GR
32	Sinistral	68.6	90	7.15	GR
33	Sinistral	39.9	90	6.55	GR
36	Sinistral	68.7	90	7.05	GR
39	Sinistral	68.8	90	7.05	GR
41	Sinistral	135.2	90	6.85	GR
43	Sinistral	263.3	90	6.65	GR

**Table 2.** Completeness thresholds used to filter the subcatalogues and define MFDs. The completeness thresholds are assigned manually from magnitude–time–density plots of the declustered subcatalogue. Due to the sparse and irregular distribution of seismic stations in the Pacific Islands region, the catalogue completeness is quite variable; this also contributes to high completeness magnitudes, even for recent years, across much of the region. Future modelling efforts may improve the completeness thresholds by incorporating observations from local and regional seismic networks.

TRT	Sources	1900	1920	1937	1940	1950	1960	1964	1980	1990	2000
Crustal	sz:1,2,3,4,5	8.0		7.0	6.0		5.5				5.0
	sz:16,12,8,9,10,17,18		7.5		6.0		5.5				
	sz:6,7,11,13,14,15		7.5		6.5		5.5	5.0			
Interface	NH1,NH4,SS1		7.5			6.5		5.5	5.0		
	NB,NH3		7.5				6.0		5.5		
	KT1,KT3,SS3,NH2,KT,NH		7.5			6.0		5.5		5.0	
	KT2,SS2,SS		7.5			6.5		5.5			
Intraslab	SS1,SS3a,SS3b,NH2,NH3,NH4,KT3,NH,SS		7.5	6.9		6.5	5.5	5.0			
	NB,SS2		7.5			6.5		5.5			
	KT1,KT2,NH1,KT		7.5	6.9		6.5		5.5	5.2		

form a surface that represents the top of the subducting plate. The surface is cut at the downdip extent of the locked interface, forming independent reference surfaces for the interface and slab-top. For subduction geometries that impose segmentation, the interface and slab surfaces are additionally cut along strike. Interface segmentation indicates our interpretation of rupture boundaries along continuous subduction zones, or significant changes in seismicity rates and tectonic characteristics. Segmentation of the slab does not necessarily imply that ruptures cannot cross the segment boundary, but instead, partially accounts for along-strike variability in the seismicity rates.

### 3.1.1 Subduction geometry

In the Pacific Islands hazard model, we include (from west to east) the New Britain, South Solomon, New Hebrides, and Kermadec-Tonga subduction zones (we discuss this last unit, which is composed of two trenches, as one continuous subduction feature because of the similar kinematics). We consider two segmentations in the seismic source characterization. In the first, herein called ‘unsegmented’ geometry, the four subduction zones are not subdivided

along strike. In the second, the ‘segmented’ geometry, the trenches are subdivided along strike according to past megathrust earthquakes, current seismicity patterns, trench convergence rates and kinematics, and assistance from thorough structural and tectonic regional studies. Here, we describe each of the subduction zones, the assumptions used to define the segmentation, and noteworthy aspects of the geometry-building process.

The ~500 km-long New Britain trench is unsegmented in both geometric interpretations. The trench axis is simple, and extends from the western trench limit to a boundary defined by Chen *et al.* (2011) as the extent of the westernmost Solomon trench super-segment. This boundary occurs at a sharp deviation in strike of the trench axis, which coincides with the western flank of a seafloor basin and a transform fault extending from the trench to New Ireland. Here, the Solomon Sea plate subducts between the South Bismarck plate at up to  $130 \text{ mm yr}^{-1}$  with  $90 \text{ mm yr}^{-1}$  of convergence.

We base the segmented geometry of the South Solomon trench, as well as convergence rates used in Section 4.3, on an arc segmentation study by Chen *et al.* (2011) which uses geographic patterns of seismicity, seafloor geomorphology, and tectonic evolution from

coral reef uplift patterns. We adopt the three super-segments, discussed here from west to east. The Bougainville segment extends from the eastern edge of the New Britain trench for  $\sim 300$  km to the western flank of the Woodlark Rise; this rise projects to a group of islands at which there is a significant uplift contrast. Here, the Solomon Sea subplate subducts beneath the Pacific plate with near-perpendicular convergence of  $\sim 100$  mm  $\text{yr}^{-1}$ . The central New Georgia segment extends eastward for  $\sim 500$  km to the Pocklington rise, subducting obliquely at  $\sim 100$  mm  $\text{yr}^{-1}$  with  $\sim 80$  mm  $\text{yr}^{-1}$  shortening (we use this value in further calculation). In this area, the subducting Australian plate can be subdivided into subplates at Woodlark Basin, across which rates vary somewhat; however, we keep these together in order to preserve a large enough declustered segment catalogue to compute robust seismicity rates. The final, easternmost segment extends for  $\sim 500$  km to just beyond Santa Ana Island, where defining a subducting plate top or interface becomes challenging. Here, the Australian plate subducts obliquely at  $\sim 90$  mm  $\text{yr}^{-1}$ , with  $\sim 70$  mm  $\text{yr}^{-1}$  of shortening. Both of these segment boundaries persist through both interface and slab; however, we additionally divide the Bougainville slab into two near-equal-width segments based on contrasting seismicity rate.

For  $\sim 250$  km eastward from the South Solomon trench, the plate boundary can still be mapped but the kinematics change and subduction of one plate does not occur. The New Hebrides subduction zone begins at the eastern extent of this, where interface and slab seismicity can be detected in the cross-sections of hypocentres and CMTs. Here, the Australian plate subducts eastward beneath the Pacific plate. In the segmented geometry for this sector, we define four segments based on Baillard *et al.* (2015). Along the westernmost segment, subduction rates are near-perpendicular at  $\sim 120$  mm  $\text{yr}^{-1}$ . This segment extends  $\sim 500$  km to a deflection in the trench strike at the d'Entrecasteaux ridge. Here, the convergence rate drops to  $\sim 90$  mm  $\text{yr}^{-1}$ , and most of the shortening transfers from the subduction interface to the backarc thrust belt (BATB)  $\sim 150$  km to the northeast; the  $\sim 300$ -km long interface segment accommodates only  $\sim 35$  mm  $\text{yr}^{-1}$ . To the east, convergence transfers back to the interface along the longest segment of the New Hebrides subduction zone ( $\sim 800$  km), with slightly arcuate geometry and convergence rates that reach  $\sim 120$  mm  $\text{yr}^{-1}$ . The eastern segment boundary occurs where the trench bends eastward, striking obliquely to the plate convergence. The final  $\sim 200$  km-long segment begins here, where the trench bends to strike obliquely. Seismicity rates are lowest here, and shortening tapers to zero. We use a representative value of 40 mm  $\text{yr}^{-1}$ .

Along the Kermadec-Tonga trench, the Pacific plate subducts westward beneath the Australian plate. In the segmented geometry, we divide this zone into three along-strike segments, similar to those defined by Bonnardot *et al.* (2007), discussing them from south to north. The southern zone extends from the southern tip of the trench system north–north–eastward for  $\sim 1200$  km to where the Louisville Seamount Chain intersects the subduction arc; this zone subducts at 60–110 mm  $\text{yr}^{-1}$ . The central zone is  $\sim 300$  km along strike, and although GPS vectors indicate convergence of  $\sim 120$  mm  $\text{yr}^{-1}$ , seismicity rates are much lower than the rest of trench system, thought to be the result of high seismic coupling between the buoyant seamounts and overriding Australian plate. The final zone, with the highest convergence rates of  $\sim 200$  mm  $\text{yr}^{-1}$ , extends north of here for  $\sim 1200$  km until the plate boundary turns westward to parallel the relative plate motion vectors.

The deep Kermadec-Tonga slab is challenging to identify simply by using the hypocentres. The slab extends to  $> 700$  km depth—well beyond the range that we include for hazard modelling—but

at intermediate depths, earthquakes also occur in a detached slab volume that has collided with the Tonga slab at  $\sim 350$ -km depth (Richards *et al.* 2011); the Slab 1.0 model includes these hypocentres in its regressions. Here, we are careful to define the slab top as the currently subducting volume, which plunges steeply, rather than flattening to near-horizontal. Seismicity in the detached slab is not considered in the source model presented here, since hypocentral depths exceed the integration distance used during classical PSHA calculations (e.g. the source-site distances are  $> 300$  km).

In the seismic source characterization, we include two downdip cut-off depths that indicate the transition from interface to intraslab seismicity. In the first, we use a uniform depth range of 9–50 km for all subduction zones (*depth\_Z1*), and in the second modify the lower depth limit to use those assigned to the full length of the respective interfaces by the Faulted Earth Project (*depth\_Z1*, Table 4, Christophersen *et al.* 2015); this is discussed more in Section 4.6. We extend the slab profiles to 300 km depth (or less, if they terminate earlier), the integration distance we use for including intraslab sources in hazard calculations, and in the case of the Kermadec-Tonga trench, a reasonable cut-off to exclude misclassification of earthquakes within the impinging detached slab.

Fig. 3 and Supporting Information Fig. S1 show the resulting slab and interface geometries, including the perimeters of the segments' surface projections. The surfaces are used for dual purposes: first as a basis for tectonic classification, and later to define the geometry of the source model faults and ruptures.

### 3.1.2 Classification settings and results

We classify the seismicity four times, producing classified catalogues for the unsegmented and segmented geometries for both depth cut-offs (*depth\_Z1* and *depth\_Z2*). For the classifications with unsegmented geometry, we include nine tectonic domains (one crustal, four interface, and four slab) while the segmented geometries include 24 tectonic domains in total (one crustal, 11 interface, and 12 slab). The domains correspond to the Moho depth model by Laske *et al.* (2013) and the subduction interfaces and slab-tops defined in Section 3.1.1.

We choose classification buffers (defined as the distance from a surface) for each tectonic domain using general characteristics of the hypocentral distribution in cross-section view, as well as geophysical and crustal characteristics. Because the Moho depth is shallow in the Pacific Islands region (mostly  $\sim 15$ -km, but ranging from  $\sim 7$  to 30 km), we set the crustal buffer limit to 30 km below Crust1.0 (Laske *et al.* 2013); this helps us to capture most hypocentres with fixed depths of 33 and 35 km. Interface surfaces are given a 20 km buffer in each direction. The slab volumes are given a 70 km thickness, which includes a 20 km buffer above the slab-top surface defined in Section 3.1.1. This thickness is an inclusive maximum, as there are few hypocentres below the slab volume, and these do not correspond to a different tectonic context considered herein. The slab-top surface is cut at 300 km, which imposes a classification lower depth of 370 km, beneath which earthquakes do not contribute significantly to the seismic hazard. We additionally classify earthquakes into a shallow slab context, which occurs directly beneath the subduction interface, but outside of its buffer region. CMTs in this proximity tend to have normal faulting mechanisms, indicating that they do not rupture the subduction interface.

**Table 3.** Parameters of crustal source zones in Fig. 2.  $N$  is the number of observations used to constrain the MFD.  $a$ -values are for the whole source zone.

Source zone	$N$	$a$ -value	$b$ -value	$M_{\max, \text{obs}}$	Mechanism type/strike/weight	Description
1	49	4.426	0.844	7.55	R/270/1.0	Fore/backarc seismicity in New Britain subduction zone
2	97	4.711	0.897	8.00	SS/130/0.7	Region of complex spreading centres and strike-slip faulting
3	42	4.912	0.949	7.65	N/270/0.5	Further backarc of South Solomon trench (Pacific Plate) where seismicity rates are lower and less dense
4	70	5.516	0.949	7.91	R/135/0.5 SS/260/0.6 R/280/0.3 N/240/0.1	Fore/backarc seismicity of New Britain and South Solomon subduction zones
5	64	5.516	0.897	7.50	N/280/1.0	Outer rise extended of New Britain trench where it merges with the South Solomon trench
6	63	4.374	0.837	6.62	N/80/1.0	Oceanic crustal region characterized by mid-ocean ridges and transform faults
7	44	3.515	0.697	7.11	SS/250/1.0	Oceanic crustal region characterized by mid-ocean ridges and transform faults
8	77	4.961	0.877	7.64	R/300/0.9	Fore/backarc of hinge between New Hebrides and South Solomon subduction zones (Pacific Plate)
9	57	5.431	0.985	7.67	R/130/1.0	New Hebrides outer rise seismicity (Australian plate)
10	89	6.051	1.062	7.70	R/345/0.45 R/280/0.45 SS/270/0.05 N/180/0.05	Fore/backarc of New Hebrides (Pacific Plate)
11	145	5.115	0.913	7.08	N/280/0.7 SS/280/0.3	North Fiji Basin; spreading ridges and transform faults
12	228	6.462	1.063	7.59	SS/260/0.9 N/210/0.1	Fiji Platform, part of Fiji Fracture zone, Lua Ridge; zone of rotation between the two subduction zones with mostly spreading ridge and transform faulting
13	257	6.793	1.195	6.64	SS/150/1.0	Zone of strike-slip seismicity that aligns with distinct lineaments
14	224	6.832	1.215	8.10	N/10/0.6 SS/100/0.25 R/190/0.15	Crustal seismicity where boundary is rotating from subduction to strike-slip
15	228	6.846	1.216	7.60	N/210/0.6 SS/60/0.4	Shallow seismicity in fore/backarc (Australian plate)
16	72	5.261	0.936	8.20	R/190/0.6 N/15/0.4	Kermadec-Tonga outer rise (Pacific Plate)
17	13	5.472	1.106	6.20	R/280/1.0	Dispersed seismicity in oceanic crust
18	10	4.187	0.897	6.99	SS/270/1.0	Dispersed seismicity in oceanic crust

Based on the configuration of the different surfaces and their buffers, some hypocentres may be assigned to more than one tectonic domain. Thus, the classification methodology by Pagani *et al.* (2020b) accounts for hypocentres that are assigned multiple classifications by using a specified hierarchy of domains. Here, we indicated that interface supersedes intraslab, and intraslab supersedes crustal, since this ordering complies with a more logical subduction zone structure (i.e., otherwise, nearly all interface earthquakes would classify as crustal). Earthquakes that may not be placed within any tectonic context are labelled ‘unclassified’ and not included in further analysis.

We inspected the classification of earthquakes with  $M_w > 7.5$  individually, and manually reassigned those we deem improperly labelled. These are mostly pre-1960 earthquakes with fixed depths, or occurring very close to the geometric boundaries.

The classifications for geometries using *depth\_Z1* yielded  $\sim 18\,000$  interface,  $\sim 16\,000$  slab and  $\sim 33\,000$  crustal earthquakes, and those using *depth\_Z2* yielded  $\sim 15\,000$  interface,  $\sim 20\,000$  slab and  $\sim 33\,000$  crustal earthquakes. Small differences in the total number of classified earthquakes are due to those classified in the shallow slab domain for geometries using *depth\_Z1* but in the slab domain for *depth\_Z2*, and consist of predominantly earthquakes with  $M < 5.5$ . Of the unclassified earthquakes,  $< 300$  have depths  $< 300$  km and  $M_w > 5.5$ , many of which are thought to occur within deeper volcanic magma chambers, and do not contribute significantly to the hazard (in this region).

### 3.2 Declustering

We decluster the catalogue following Gardner & Knopoff (1974), an approach commonly used for pre-processing earthquake catalogues in seismic hazard analysis. We test the time, space, and magnitude windowing proposed by Gardner & Knopoff (1974), and additionally, the configurations by Uhrhammer (1986) and Gruenthal (see van Stiphout *et al.* 2010), ultimately choosing the Uhrhammer (1986) parameters; we found our catalogue to be highly sensitive to the other options, resulting in such low main-shock fractions that—in some cases—it was too challenging to resolve seismicity rates.

Because seismicity interacts among adjacent tectonic contexts (i.e., interface main shocks can trigger crustal aftershocks), we decluster several domains together, and then separate the earthquakes into their respective classified, declustered catalogues. We use two groups: crustal, interface, and shallow slab seismicity; and deep slab. The declustering algorithm works in two spatial dimensions, looking only at the horizontal distance between earthquakes (e.g. comparing epicentres, not hypocentres). Thus, for steep subduction geometry such as the Tonga and Kermadec slabs, where parts of the deep slab could occupy the triggering window for large interface earthquakes, the separation into two groups is critical. While still state-of-practice in PSHA, the two-dimensionality of the selected declustering algorithm is a major limitation; however, the catalogue depth uncertainties (described at the start of Section 3) may also prohibit accurate 3D declustering, and thus we deem the grouped approach used herein to be sufficient.

**Table 4.** Characteristics of interface segments and MFD parameters for all instances of interface sources.  $a$  and  $b$  are GR parameters,  $N$  is number of observations used to constrain the MFD,  $N(M > 7)$  is the number of observations with  $M_w > 7.0$ , and Rate is the convergence rate. Convergence rates for the South Solomon Trench are from Chen *et al.* (2011); New Hebrides Segment 2 are from Baillard *et al.* (2015); and all others are from Bird (2003). Coupling coefficients for South Solomon Segments 1 and 2 are from Chen *et al.* (2011); New Hebrides Segments 2 and 3 are from Wallace *et al.* (2012); and all others are from Heuret *et al.* (2011). Many of the parameters here overlap with the ranges defined in the Faulted Earth project by Christoffersen *et al.* (2015).

Subduction zone	Segment	Lower depth (km)	$N(N(M > 7))$	$a$ -Value	$b$ -Value	$M_{\max, \text{obs}}$	$M_{\text{char}}(\text{TM/AH})$	Rate (mm yr $^{-1}$ )	Coupling
New Britain	full,z1	50	65 (7)	4.267	0.732	8.10	8.42/8.46	90	0.24
	full,z2	40	58 (5)	4.354	0.756	8.10	8.34/8.39	90	0.24
South Solomon	full,z1	50	78 (11)	4.498	0.763	8.07	8.65/8.65	90	0.50
	full,z2	40	74 (11)	4.310	0.734	8.07	8.55/8.56	90	0.50
	1,z1	50	53 (6)	3.391	0.642	7.95	8.13/8.24	70	0.60
	1,z2	40	50 (6)	3.181	0.607	7.95	8.05/8.17	70	0.60
	2,z1	50	26 (2)	5.216	0.977	8.07	8.22/8.30	80	0.50
	2,z2	40	24 (2)	4.902	0.927	8.07	8.12/8.22	80	0.50
	3,z1	50	44 (3)	4.937	0.932	8.00	8.04/8.17	100	0.26
	3,z2	40	43 (3)	5.058	0.957	8.00	7.95/8.09	100	0.26
	full,z1	50	281 (21)	5.062	0.806	8.14	8.85/9.29	90	0.50
	full,z2	31	203 (22)	4.734	0.771	8.14	8.56/8.61	90	0.50
New Hebrides	1,z1	50	10 (0)	4.411	0.972	6.35	7.58/7.80	40	0.11
	1,z2	31	10 (0)	4.567	1.015	6.35	7.35/7.62	40	0.11
	2,z1	50	159 (11)	5.096	0.860	8.14	8.49/8.51	120	0.40
	2,z2	31	128 (11)	4.954	0.851	8.14	8.29/8.35	120	0.40
	3,z1	50	26 (2)	4.250	0.797	7.60	8.12/8.22	35	0.90
	3,z2	31	17 (2)	3.718	0.736	7.60	7.94/8.08	35	0.90
	4,z1	50	78 (8)	3.942	0.709	8.02	8.25/8.32	120	0.25
	4,z2	31	47 (6)	3.381	0.649	8.02	8.07/8.18	120	0.25
	full,z1	50	851 (12)	7.319	1.145	7.97	9.29/9.54	140	0.21
	full,z2	32	761 (10)	7.222	1.136	7.97	9.13/9.38	140	0.21
Kermadec-Tonga	1,z1	50	377 (5)	7.060	1.164	7.88	8.90/9.42	110	0.21
	1,z2	32	320 (3)	7.125	1.190	7.10	8.72/8.88	110	0.21
	2,z1	50	21 (1)	6.396	1.206	7.20	8.29/8.36	120	0.12
	2,z2	32	21 (1)	6.309	1.190	7.20	8.14/8.24	120	0.12
	3,z1	50	419 (6)	6.758	1.097	7.97	8.95/9.20	200	0.04
	3,z2	32	380 (6)	6.615	1.078	7.97	8.79/9.09	200	0.04

#### 4 SEISMIC SOURCE MODEL

The primary sources of seismic hazard in the Pacific Islands region are the subduction interfaces, and their associated outer rise, forearc, backarc, and slab structures. Additional seismicity occurs around mid-ocean ridges and corresponding transform faults; on crustal faults and fracture/shear zones; distributed within shallow oceanic and continental crust; in volcanoes; and at great depths in nests thought to correspond with detached slabs.

The seismic source model presented herein includes sources of four types: subduction interfaces, subduction intraslab ruptures, active shallow crustal faults, and gridded distributed crustal seismicity. Here, we describe the source characterization. In Sections 4.1 to 4.5, we describe a single instance of the source characterization, ignoring epistemic uncertainty, in order to demonstrate the methodology used to construct seismic sources for each tectonic region type; epistemic uncertainty is addressed in Section 4.6.

##### 4.1 Crustal faults

We use crustal faults from the GEM Active Faults Database (Styron & Pagani 2020) to create fault sources. For the Pacific Islands, crustal faults with known kinematics are limited to oceanic structures with slip rates resolved by Bird (2003). The data set predominantly consists of spreading ridges and transform faults, but also includes the Fiji Fracture Zone, a transform plate boundary where the northernmost Tonga trench turns east–west, a pair of subduction thrusts between minor tectonic plates, and a few other strike-slip

faults. While other faults are known to exist, we do not have the minimum information needed to model them as seismic sources (e.g. slip rates).

We compile a set of individual faults by linking together fault segments from Bird (2003) that are continuous, and have similar orientation and slip type, choosing a representative slip rate value. For each fault, we compute the area of the fault surface from the length of the surface trace and the downdip width; the latter is constrained by the lower seismogenic depth (see below) and fault dip (if unknown, assumed from Andersonian mechanics, Anderson 1905, e.g. the optimal planar angle for failure in a given stress field). We then use the Leonard (2010) scaling relationships to compute a maximum magnitude ( $M_{\max}$ ). We keep only faults capable of generating  $M_w > 6.5$  earthquakes based on the scaling relationship, thus excluding the shortest faults, and most transform faults offsetting ridges. We also exclude faults parallel to spreading ridges, which are believed to have a very thin seismogenic coupling zone, resulting in most deformation occurring aseismically and  $M_{\max} \sim 5.8$  (Bird *et al.* 2002). In total, we keep 20 faults divided into two groups: (1) large subduction-type thrust faults that do not coincide with major plate boundaries, and (2) seafloor faults including fracture zones and ridge-offsetting transforms.

We convert each fault trace into an OpenQuake Engine simple fault source using the Fault Modeler of the MBTK (see [https://github.com/GEMScienceTools/oq-mbtk/tree/master/openquake/mbt/tools/fault\\_modeler](https://github.com/GEMScienceTools/oq-mbtk/tree/master/openquake/mbt/tools/fault_modeler)), which uses the fault's geometry (dip, seismogenic depth limits) and kinematics (rake, slip rate) to generate a magnitude frequency distribution for the source. For

most faults—all those in Group 2—we use a Gutenberg–Richter (GR) magnitude frequency distribution (MFD) truncated at  $M_{\max}$ . The faults in Group 1 are large and thus may generate higher magnitude ruptures than other crustal faults included here. However, instrumental seismicity rates of all magnitudes are low in their close proximities and these observations are not supported by the modelled GR MFD. Thus, we hypothesize that their moment release occurs mostly in larger earthquakes, and use MFDs that are modelled by a characteristic component around  $M_{\max}$ , but an exponential distribution at lower magnitudes: the Youngs & Coppersmith (1985) characteristic MFD.

For each fault, we compute the available seismic moment rate  $\dot{M}_0$  using slip rate  $\dot{s}$ ;  $M_{\max}$  computed from fault area  $A$ ; an aseismic coefficient  $c$ ; and a default shear modulus  $\mu$  of 32 GPa (typical for crustal rocks rather than oceanic crust, (e.g. Turcotte & Schubert 2014), but noting that this value trades off linearly with the aseismic coupling coefficient). We then compute the annual occurrence rates,  $n(m)$ , for each magnitude bin (0.1 units) between  $M_{\min}$  and  $M_{\max}$  by balancing them with the seismic moment rate, according to eq. (1):

$$\dot{M}_0 = \mu A \dot{s} c = \int_{M_{\min}}^{M_{\max}} n(m) M_0(m) dm, \quad (1)$$

where  $M_0$  is in units of Nm, and is computed for each magnitude  $m$  using eq. (2):

$$M_0(m) = 10^{1.5m+9.05}. \quad (2)$$

The rates for Group 2 faults, which use exponential GR MFDs, were computed using eq. (3), which follows eq. (9) of Youngs & Coppersmith (1985):

$$N(m) = N(M_{\min}) \frac{e^{-\beta(m-M_{\min})} - e^{-\beta(M_{\max}-M_{\min})}}{1 - e^{-\beta(M_{\max}-M_{\min})}}. \quad (3)$$

$N(m)$  is the cumulative rate of earthquakes with  $M > m$ ,  $\beta = b \ln(10)$ , and  $M_{\min}$  and  $M_{\max}$  are the minimum and maximum considered magnitudes, respectively.  $M_{\min}$  is the lower bound of the magnitude range across which  $M_0$  is distributed; here,  $M_{\min} = 4.0$  (use of an  $M_{\min}$  lower than this causes negligible change to the MFD rates).  $N(M_{\min})$ , the cumulative rate of earthquakes with  $m > M_{\min}$ , is determined from eq. (4):

$$N(M_{\min}) = \frac{-M_0(1.5 \ln(10) - \beta)(1 - e^{-\beta(M_{\max}-M_{\min})})}{\beta(10^{9.05+1.5M_{\min}} - 10^{9.05+1.5M_{\max}} e^{-\beta(M_{\min}-M_{\max})})}. \quad (4)$$

Eq. (4) is derived from eq. (1), where  $n(m)$  is defined by eq. (5) (eq. 8 of Youngs & Coppersmith 1985):

$$n(m) = \frac{\beta n(M_{\min}) e^{-\beta(m-M_{\min})}}{1 - e^{-\beta(M_{\min}+M_{\max})}} \quad (5)$$

The rates for Group 1 faults were computed using eq. (6), which follows Equation 16 of Youngs & Coppersmith (1985).

$$\begin{aligned} \mu A \dot{s} &= \frac{(N(m) - N(M_C)) e^{-\beta(M_{\max}-m-0.5)} M_0^{\max}}{1 - e^{-\beta(M_{\max}-m-0.5)}} \\ &\quad \left( \frac{b 10^{-c/2}}{c-b} + \frac{b e^{\beta}(1-10^{-c/2})}{c} \right) \end{aligned} \quad (6)$$

$M_0^{\max}$  and  $M_0^{\min}$  are the seismic moment in Nm of the maximum and minimum considered magnitudes, respectively, and  $M_C$  is the magnitude of the characteristic (full-fault rupture). The rate of  $M_C$  earthquakes  $N(M_C)$  is calculated using eq. (7):

$$N(M_C) = \frac{\beta(N(m) - N(M_C)) e^{-\beta(M_{\max}-m-1.5)}}{2(1 - e^{-\beta(M_{\max}-m-1.5)})}. \quad (7)$$

$c = 1.5$ , and all other parameters are as described previously.  $N(m) - N(M_C)$  accounts for the exponential (non-characteristic) component of the MFD.

For the seafloor faults in Group 2, we use a lower seismogenic depth of 10 km,  $M_{\max}$  and aseismic coefficient of 0.85, which means 85 per cent of seismic moment is released aseismically by creep (e.g. Boettcher & Jordan 2004). The Group 1 thrust faults use a seismogenic depth of 30 km and aseismic coefficient of 0.5 and  $M_{\max}$  computed from fault area using the magnitude scaling relationship by Leonard (2010) ( $M_w = \log_{10}(A) + 4.19$ , also used during the hazard calculation to produce ruptures). The aseismic coefficients are calibrated by seismicity; very little seismicity is distinctly attributable to Group 1 faults, while the Group 2 faults are somewhat more seismically productive. We use  $b$ -values from the MFDs computed for the source zones (described next) that contain the majority of the fault. The final MFDs take only the range  $M_w$  6.5 to  $M_{\max}$ , discretized by 0.1 unit magnitude; lower magnitudes are modelled by distributed seismicity (described in Section 4.2).

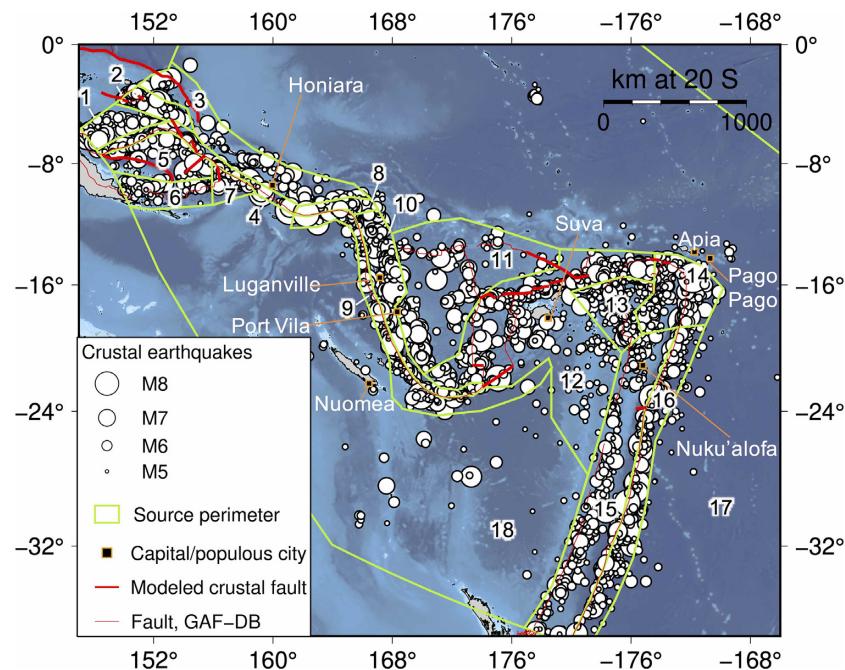
Fault source parameters are summarized in Table 1, and displayed in Fig. 2 and Supporting Information Fig. S0.

## 4.2 Distributed seismicity

We model the remaining active shallow crustal seismicity as gridded point sources for zones with common tectonic characteristics (e.g. see the descriptions in Table 3). In our approach, we define large ‘source zones’ that encompass regions with internally similar tectonics, but without considering seismicity rates, since this could create too small of areas to robustly compute earthquake statistics (e.g. Peñarubia *et al.* 2020; Rong *et al.* 2020). Instead, we account for variability in rate by distributing the modelled seismicity onto a grid with variable activity rate across each source zone. The source zones mostly correspond to trench outer rises, forearc and backarc seismicity, fracturing and rotation of the Fiji Platform, and regions of seafloor spreading.

We use the MBTK to characterize distributed crustal seismicity. For each source zone, we create a subcatalogue by clipping the declustered crustal catalogue to the perimeter of the respective source zone, and apply the completeness thresholds summarized in Table 2. Using the incremental observed seismicity rates of the corresponding subcatalogue, we compute GR  $a$ - and  $b$ -values using the algorithms defined by Weichert (1980), and use these to define a double-truncated MFD with  $M_{\min} = 5$  and  $M_{\max} = M_{\max, \text{obs}} + \Delta$  and bin width of  $M = 0.1$ . The added  $\Delta$  indicates that we cannot know whether  $M_{\max, \text{obs}}$ —the largest observed earthquake—is the largest earthquake a source can produce; we use  $\Delta = 0.5$ . We model the range of hypocentral depths using a weighted distribution of 10 km depth bins from 0–50 km noting that <10 per cent of seismicity classified as active shallow occurs in the deepest bin for most source zones. Lastly, we assign a few most-likely focal mechanisms to each source zone based on the CMTs (Ekström *et al.* 2012) of earthquakes classified as crustal within the respective source zone perimeter.

We ‘smooth’ the modelled seismicity using a method similar to that by Frankel (1995) in order to account for variations in seismicity rate within a source zone, and based on the premise that future seismicity will occur where past seismicity has been observed. Within each source zone, we smooth the seismicity onto a  $0.1^\circ$  grid of nodes using 2D smoothing kernels shaped by a Gaussian with a given standard deviation  $\sigma$ . The total seismic moment rate in the instrumental catalogue (here, the crustal catalogue clipped to the



**Figure 2.** Crustal sources. Epicentres are earthquakes classified as crustal in Section 3.1 with  $M_w \geq 5.0$ . Green lines demarcate the source zones summarized in Table 3 (NB zone 17 extends beyond the figure coordinates). Thin red lines are faults in the GAF-DB (Styron & Pagani 2020). Thick red lines are faults included in the source model presented here. Orange squares are sites (capital or large cities) for which the hazard is discussed most thoroughly. In Supporting Information Fig. S0, the faults are labelled with IDs corresponding to Table 1.

source zone) which occurred within a grid cell is computed, and the kernels are used to smooth that rate onto the surrounding grid cells. The scale of the Gaussian controls how the rate of occurrence is distributed, considering all grid cells within a specified radius  $r$ . For each node, the fraction from all other grid cells is summed to determine the portion of the total source zone seismicity that should be assigned to that cell, and then the source zone  $a$ -value is scaled accordingly. A point source is placed at the centre of the grid cell.

Here, we use a pair of smoothing kernels, giving 95 per cent weight to a Gaussian with  $\sigma = 20$  km and  $r = 50$  km, and 5 per cent weight to a Gaussian with  $\sigma = 5$  km and  $r = 20$  km. The higher-weighted kernel, which spans 100 km, is close in its reach to the width of the source zones with the highest seismicity rates, for example, forearc, backarc and outer rise sources. The observed seismicity rates decrease rapidly with distance from the subduction trench, and thus the selected smoothing radii are appropriate. However, in lower-rate source zones, the small radii will produce ‘bulls-eye’ patterns in the computed hazard. Nevertheless, we use the same smoothing kernels for all source zones included in this model, since the lower-rate zones typically encompass very little inhabited land, and have overall low hazard. However, future development of this model could consider alternative ways of calibrating the smoothing kernels.

For grid points that occur within the surface projection of crustal faults (plus a 15-km buffer), we adjust the magnitude ranges of the gridded seismicity MFDs to prevent double-counting of crustal sources. In these areas of overlap, we cap the MFDs of the distributed point sources at  $M_w = 6.5$ , and allow the faults to model all larger magnitude occurrences.

The source characterization for active shallow crustal sources uses 18 source zones, which are summarized in Table 3 and displayed in Fig. 2.

#### 4.3 Interface seismicity

We model the subduction sources following the approach by Pagani *et al.* (2020b). This methodology accounts for interface seismicity by ‘floating’ ruptures along surfaces with complex geometry defined for each interface segment (see Section 3.1.1), and models earthquake occurrence rates using a hybrid approach that combines statistics of observed seismicity with a characteristic component derived from tectonics. The statistical approach uses the same method as for crustal seismicity, and solves for a classical GR distribution (negative exponential) as in Weichert (1980) using the incremental observed seismicity rates in the subcatalogues classified to the pertinent interface segment (NB: recall that the subcatalogue is first declustered and filtered for completeness according to Table 2). The latter approach derives a double-truncated Gaussian distribution to model the earthquake that will fill the entire interface surface: a ‘characteristic earthquake’. The two approaches are combined into a hybrid MFD by taking the higher of the incremental occurrence rates for each magnitude bin, allowing for an MFD ‘shape’ that can fit observations which do not strictly comply with a GR or Gaussian distribution. Instead, this approach assumes that the rate of seismic moment released in the largest earthquakes is controlled by strain accumulation via tectonic processes (e.g. plate convergence) while the rate of lower magnitude earthquakes is controlled by exponential decay via a GR relationship. The hybrid MFD accommodates decoupling of the two controls, a primary assumption by this methodology, without imposing a pre-defined shape. In future work, we may extend the methodology to incorporate more complete moment budgeting.

We note that, compared to lower magnitude earthquakes, only a few high-magnitude observations have been recorded during the instrumental period, and thus the observed rates at magnitudes in the characteristic range have large uncertainties; however, most of the

sources do include a few earthquakes  $M_w > 7.0$  (see Table 4.) When possible, we additionally compare the resulting MFDs against other data types, such as palaeoseismic or palaeogeodetic data that can suggest long-term (Holocene to Recent) occurrence rates.

The magnitude and recurrence interval of the characteristic earthquake component for each interface segment are based on the fault area, the convergence rate, and a seismic coupling coefficient following eq. (1). The magnitude of the full-interface rupture is computed using the interface scaling relationship developed by Thingbaijam *et al.* (2017) and the segment geometry defined in Section 3.1.1. Then, the fault area, convergence rate, and coupling coefficient are used to determine the time needed to accumulate the seismic moment for that maximum magnitude earthquake following eq. (1); this is the recurrence interval of the characteristic earthquake. However, the characteristic component is not represented as a single magnitude–frequency pair, but as a Gaussian-shaped probability density function centred at the characteristic magnitude  $M_C$  and double-truncated at two standard deviations  $\sigma$  of the magnitude scaling relationship ( $\sigma = 0.15$  for Thingbaijam *et al.* 2017 and  $\sigma = 0.266$  for Allen & Hayes 2017), complying with the right-hand side of eq. (1).

The shape of the Gaussian for each MFD is derived using the Python Scipy function ‘stats.truncnorm.pdf’ (see <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.truncnorm.html>), which computes a standardized distribution for the range  $a$  to  $b$  (eqs 8–11) symmetric about  $M_w = 0$ , shifted by  $M_C$  and scaled by  $\sigma$ , and then normalized such that the Gaussian values  $G(m)$  sum to 1.0:

$$a = (M_{\min} - M_C)/\sigma \quad (8)$$

$$b = (M_{\max} - M_C)/\sigma \quad (9)$$

$$M_{\min} = M_C - 2 * \sigma \quad (10)$$

$$M_{\max} = M_C + 2 * \sigma. \quad (11)$$

Eqs (8) and (9) are rounded down and up, respectively, to the nearest MFD bin edge. Once  $G(m)$  is determined, the characteristic occurrence rates  $N(m)$  for are computed following eq. (12):

$$N(m) = \frac{M_0^{\text{TOT}}}{M_0^C} * G(m), \quad (12)$$

where  $M_0^{\text{TOT}}$  is the moment rate computed due to convergence, computed from eq. (1), and  $M_0^C$  is the moment for an earthquake of magnitude  $M_w = M_C$  following eq. (2).

The parameters used for each interface are listed in Table 4, and correspond to the segments in Fig. 3. Selecting the coupling parameter is challenging, in large part due to the scarcity of land and thus GPS measurements in close proximity to the southern Pacific Islands trenches. Where no other model is available, we take values from the most consistent subduction segment of Heuret *et al.* (2011).

Fig. 4(a) shows the MFD for South Solomon trench Segment 2. The observations for this segment demonstrate the utility of the hybrid MFD; using a GR MFD, earthquakes of  $M_w > 7.0$  (and especially  $M_w \sim 8.0$ ) would be under-predicted, while the characteristic component would fail to capture the more frequent smaller magnitude occurrences. The MFD for this segment fits reasonably well with the alternative characteristic recurrence estimate by Kuo *et al.* (2016), who used coral uplift data by Thirumalai *et al.* (2015) with their own GPS data to estimate recurrence of  $M_w \geq 8.0$  every 100 yr or more, with a preferred rate of  $M_w 8.1$  every 107 yr. Fig. 4(b) shows the MFD for New Hebrides trench Segment 3,

for which palaeoseismic constraints by Taylor *et al.* (1990) help to validate the MFD.

The MFDs for all other interface segments are shown in Supporting Information Figs S4–S7. Importantly, the characteristic component is not always dominant in the MFD. For segments where the Gaussian distribution predicts lower occurrence rates than the GR MFD for all magnitude bins, the MFD remains a GR MFD. However, the characteristic component is still instructive for constraining the maximum magnitude based on fault area.

#### 4.4 Intraslab seismicity

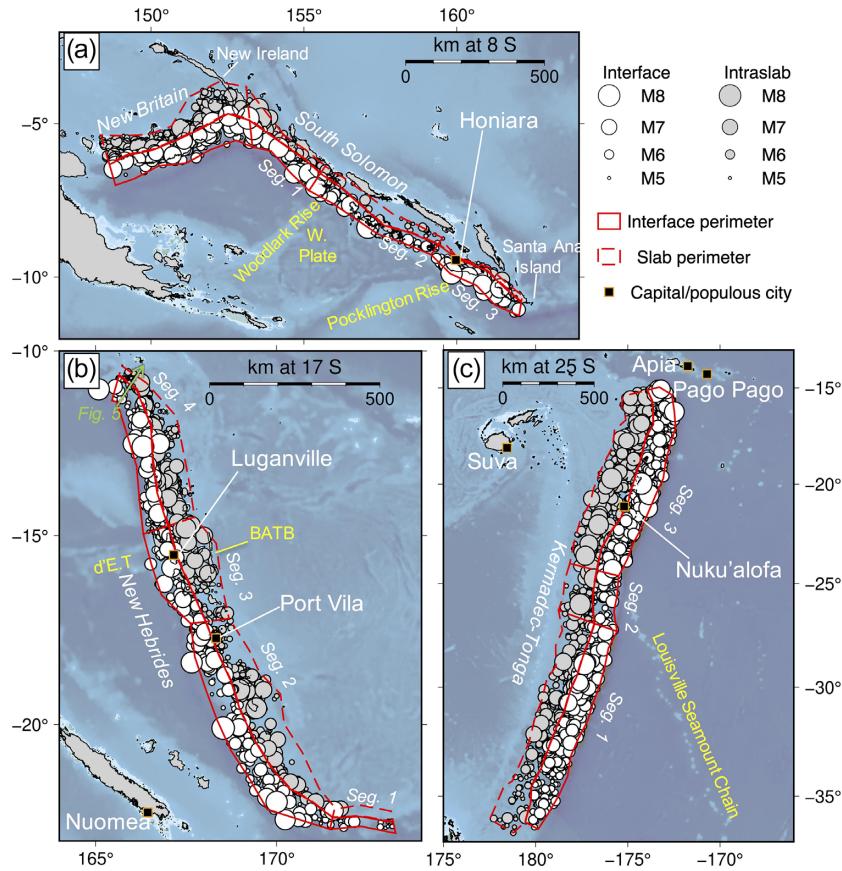
Intraslab sources are also modelled following the methodology of Pagani *et al.* (2020b). First, we use the geometry defined in Section 3.1.1 to form ‘virtual faults’ within the slab. The virtual faults are surfaces that cut the slab volume at either  $45^\circ$  and  $135^\circ$  from the slab top, and are permitted to bend to follow the slab curvature, but without leaking out of the slab volume. Next, we create a set of gridded ruptures that occur on these virtual faults, which comprise the set of possible rupture geometries for a range of magnitudes, a magnitude scaling relationship (we use that of Strasser *et al.* 2010) that could occur within the slab. Restricting the gridded ruptures to the slab volume is advantageous when paired with ground motion models that use distance metrics of  $R_{\text{rup}}$ , which consider the rupture geometry.

Once the gridded ruptures are defined, we assign each a probability of occurrence within a given investigation period (1 yr). To do this, we first define seismicity rates for the slab segments using the respective declustered subcatalogues. For each segment, we resolve a single GR MFD from the incremental observed rates (Table 5), assuming spatially constant seismicity rates, and compute the moment rate for  $M_w 6.5$  to  $M_{\max, \text{obs}} + \Delta$ , using a  $\Delta$  of 0.5 to indicate our lack of knowledge as to the true maximum magnitude earthquake that can occur. The total occurrence rate for each magnitude bin is divided among all the ruptures for that magnitude, and then converted to a probability of occurrence; the probabilities are assigned to the gridded ruptures with corresponding magnitudes.

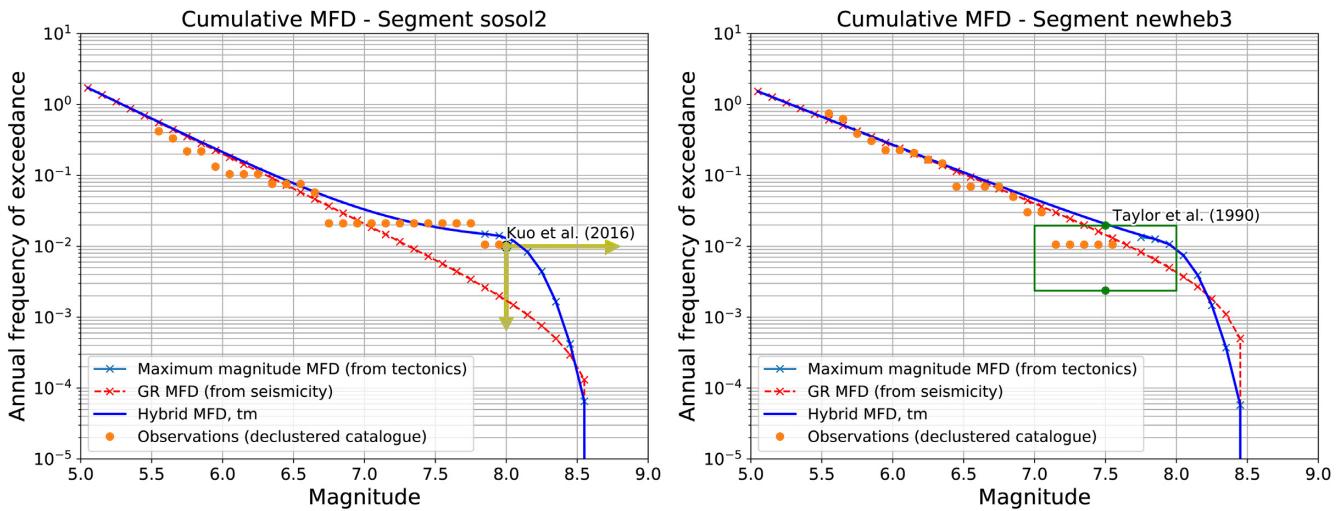
#### 4.5 Subduction classification validation and sensitivity

We investigate the effectiveness of the classification procedure, and the accuracy of our selected buffer limits and surface geometry, in order to validate our interpretation of the tectonic context. As an initial check, we observe the styles of faulting of earthquakes classified into each tectonic context by additionally classifying the GCMT catalogue of focal mechanisms (Dziewonski *et al.* 1981), and observing the most prominent focal mechanisms in the volumes we used to delimit the interfaces and subducting slabs. We compare these to the ‘expected’ mechanisms based on simplified tectonics: reverse mechanisms along the interface, and normal mechanisms within the slab.

Table 6 shows the results of this evaluation for geometries using  $\text{depth\_Z1}$ , indicating the percentage of CMTs classified to each interface segment with reverse mechanisms, and the same statistic for normal mechanisms in the subducting slab. While we achieved the expected results for the interface segments (e.g. most segments resulted in at least 80 per cent reverse mechanisms), the results were less informative for intraslab sources. The nature of subduction in the Pacific Islands region renders our proposed validation methodology more effective for interface sources than for the slabs. While the assumption that interface seismicity should manifest as reverse



**Figure 3.** Subduction sources of the Model 1 (*depth\_Z1*) segmented SSM (see Fig. 6). Epicentres refer to earthquakes classified to each of the labelled subduction interface or intraslab segments described in Section 3.1. Source perimeters are the surface projections of the surfaces defined in Section 3.1.1, which delimit the segmented interface and intraslab sources of Model 1. The unsegmented Model 1 sources use the perimeter of all segments within a given subduction system. Orange squares indicate capital or populous cities discussed in greater detail. Yellow font indicates the location of seafloor structures. BATB: Backarc thrust belt. d'E.T. d'Entrecasteaux Ridge. W. Plate: Woodlark Plate.



**Figure 4.** (a) MFD options for the South Solomon trench Segment 2, plotted with the declustered subcatalogue filtered for completeness (orange dots) and an additional recurrence estimate (yellow arrows indicating a characteristic earthquake of  $M_w \geq 8.0$  every 100 yr or longer) from Kuo *et al.* (2016), which is based on continuous GPS interpreted with the palaeogeodetic (coral uplift) data from Thirumalai *et al.* (2015). The observations cannot be fit by a GR MFD, or by the maximum magnitude MFD derived from tectonics, but have a reasonable fit to the hybrid MFD. (b) MFD options for the New Hebrides trench Segment 3, with an additional recurrence estimate range from the coral uplift analysis of Taylor *et al.* (1990). Here, the hybrid shape of the MFD is not required by the seismicity observations, but is permitted by the palaeogeodetic record. Some of the alternative MFDs for this segment, which are introduced in Section 4.6 and depicted in Supporting Information Fig. S7, do not deflect from the GR shape. (NB: The palaeo data is not used to compute the MFD; it is only used for visual comparison. The observations shown here are cumulative, but the MFDs are computed from the incremental rates of the observed seismicity.).

**Table 5.** MFD parameters for intraslab segments.  $a$  and  $b$  are GR parameters.

Subduction zone	Segment	$N$	$a$ -Value	$b$ -Value	$M_{\max, \text{obs}}$
New Britain	full,z1	65	5.168	0.896	7.5
	full,z2	64	5.000	0.867	7.5
South Solomon	full,z1	104	5.180	0.927	7.72
	full,z2	140	5.568	0.997	7.72
	1,z1	11	4.673	1.014	6.95
	1,z2	33	6.858	1.355	6.95
	2,z1	5	5.535	1.155	6.47
	2,z2	7	6.007	1.215	6.97
	3a,z1	21	4.284	0.882	6.33
	3a,z2	25	4.456	0.905	7.0
	3b,z1	56	4.281	0.804	7.7
	3b,z2	53	3.772	0.710	7.72
New Hebrides	full,z1	535	5.834	0.916	7.85
	full,z2	648	5.861	0.905	7.85
	1,z1	8	4.271	0.933	7.0
	1,z2	12	5.786	1.186	7.0
	2,z1	162	5.219	0.896	7.85
	2,z2	241	5.079	0.835	7.85
	3,z1	131	4.996	0.871	7.70
	3,z2	152	5.144	0.888	7.70
	4,z1	215	5.396	0.907	7.80
	4,z2	240	5.475	0.914	7.82
Kermadec-Tonga	full,z1	468	6.503	1.024	7.84
	full,z2	651	6.876	1.068	7.88
	1,z1	84	5.945	1.057	7.3
	1,z2	244	6.071	0.996	7.38
	2,z1	37	5.094	0.967	7.84
	2,z2	45	6.399	1.038	7.80
	3,z1	465	6.420	1.042	7.80
	3,z2	584	6.420	1.042	7.80

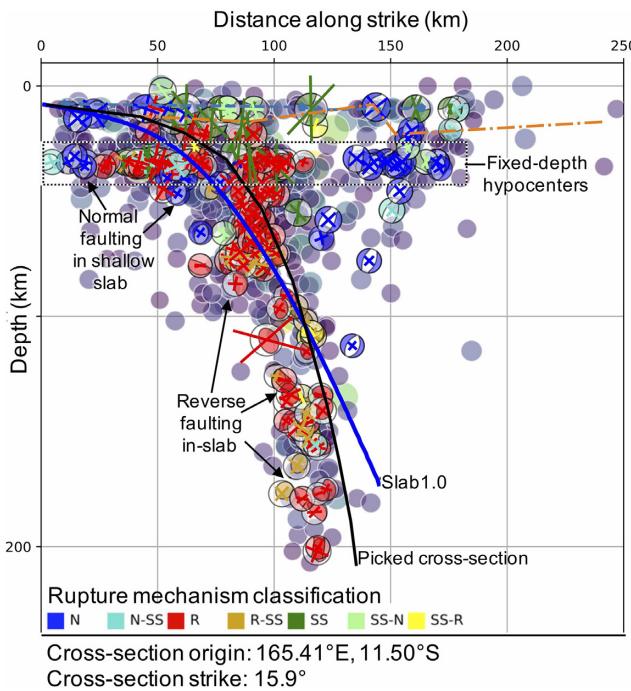
**Table 6.** Classification validation. Per cent correct is the percentage of earthquakes classified to each context that have the ‘expected’ focal mechanism. Per cent difference TMR is the relative change in the MFD total moment rate for an interface segment when the classification lower buffer is increased from 20 to 30 km. I: interface. S: intraslab

Subduction zone	Segment	Per cent correct	Per cent correct	Per cent difference
		(I)	(S)	TMR
New Britain	1	89	30	-17.4
South Solomon	1	89	0	0.3
	2	90	31	0.2
	3	92	12	4.8
New Hebrides	1	50	9	-2.0
	2	75	26	-7.6
	3	80	13	-1.2
	4	83	2	-14.4
Kermadec-Tonga	1	93	39	-1.0
	2	92	33	0.0
	3	76	58	-0.3

focal mechanisms is true based on the cross-section analysis (Section 3.1.1), the inslab seismicity tends to experience mechanisms of all types at all depths, and is often dominated by reverse mechanisms (e.g. Fig. 5). Thus, if we assume that inslab seismicity is dominated by normal faulting, this methodology cannot in itself be used to evaluate the classification and subsequent modelling sensitivity to interface-intraslab cut-off depth and choice of buffer.

A second challenge in using the CMTs to validate the classification is the dominance of focal mechanism (and hypocentres) with standardized depth assignments, the most problematic being 33 or 35 km depth, and therefore at risk of misclassification from the interface to the shallow slab or vice versa. In consideration of

this, we perform a second classification for the segmented geometry using *depth\_Z1* in which we increase the below-interface buffer to 30 km, and assess the sensitivity of the resulting interface MFDs to this classification parameter. We find that for nearly all of the interface segments, the MFD is not significantly changed when the lower buffer increases; the applied declustering eliminates many of the earthquakes from the catalogue used to determine the occurrence rates. The total annual moment rate for MFDs computed for the two classification settings changes by <2 per cent for most segments. For two segments, the moment rate counterintuitively reduces when the buffer is increased due to drastic changes to the MFD  $b$ -values that suggest the MFDs may be more sensitive to



**Figure 5.** Example of a cross-section from New Hebrides Segment 4 used in Section 3.1.1 to delineate the interface and slabtop surfaces. Focal mechanisms are coloured by the Kaverina *et al.* (1996) classification. Reverse focal mechanisms occur at all depths (0–200 km).

completeness thresholds than to the classification below-interface buffer. Importantly,  $M_{\max}$  is not affected.

#### 4.6 Source model uncertainty

An initial disaggregation analysis of the single source model used in the GEM Hazard Mosaic v. 2018.1 (Pagani *et al.* 2020a) confirmed that the subduction sources are contributing most to hazard at most of the capital and large cities, the exceptions being cities where hazard is overall low. Thus, we focus on the subduction sources to explore epistemic uncertainty of the model parameters.

In the optimal case, the logic tree for subduction sources considers the following epistemic uncertainties. For both interface and intraslab components, the source model would include a range of subduction geometries, upper and lower seismogenic locking depths (e.g. the extent of the subduction interface), and along-strike segmentations including an unsegmented model. Interface sources would additionally account for uncertainty in the MFD type (GR versus hybrid), maximum magnitude (e.g. the choice of magnitude scaling relationship), and recurrence parameters, while intraslab sources would include uncertainty in maximum magnitude (e.g. by adding different deltas to the maximum observed earthquake). Other modelling assumptions earlier in the workflow also have an impact on the resolved occurrence rates, such as the parameter choices during tectonic classification of seismicity, declustering windows or technique, and choice of completeness table; the outcomes of these choices should also be treated as uncertain.

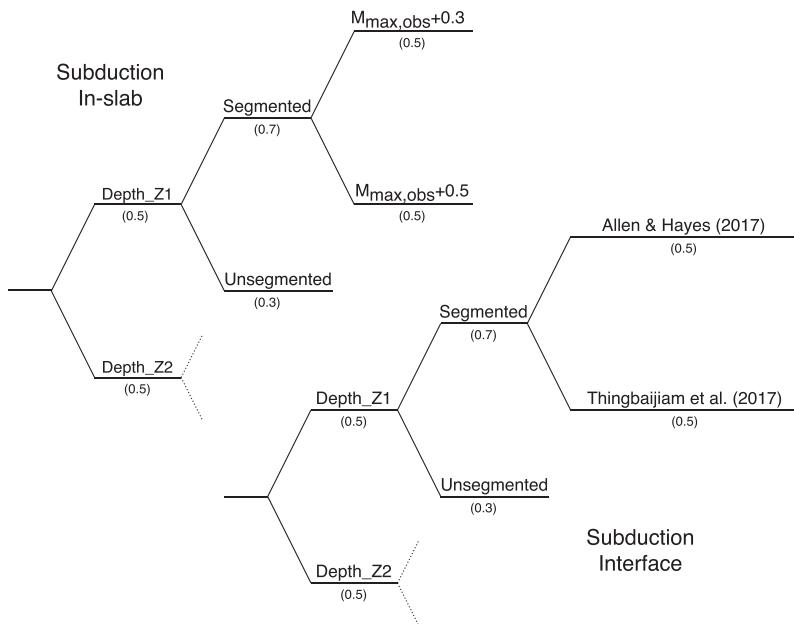
Here, we explore epistemic uncertainty by devising a source model logic tree based on the uncertain parameters that are impactful to the hazard curves at capital and large cities throughout the Pacific Islands, and for which we can propose two reasonable alternatives. The variables with two alternative hypotheses included

in the final source model logic tree are the geometric interpretations described in Section 3.1.1—two depth models used to define the boundary between subduction interface and intraslab seismicity and two versions of along-strike segmentation—as well as two values of  $M_{\max}$  for each source. The logic tree is depicted in Fig. 6.

The first division of the logic tree is by the depth used to demarcate the boundary between interface and intraslab seismicity, yielding alternative model hypotheses  $depth\_Z1$  and  $depth\_Z2$ . Because the depth models are used to construct the subduction source geometries, seismicity for the two tectonic regions must always be represented by the same depth model; otherwise, sources at depths  $z_1 < z < z_2$ , where  $z$  is the source depth, will be either doubly counted or absent (e.g. taking interfaces from model  $depth\_Z1$ , where  $z = 50$  km, but intraslab source from model  $depth\_Z2$ , where  $z = 30$  km, would result in duplicity of sources in the range 30–50 km). Thus, for the purpose of constructing the logic tree and for PSHA calculation, we treat  $depth\_Z1$  and  $depth\_Z2$  as independent source models which treat epistemic uncertainties symmetrically. Logic tree branches that diverge from  $depth\_Z1$  use the uniform 50-km interface lower limit for all included subduction zones, as discussed in the previous sections. Branches diverging from  $depth\_Z2$  use lower depth limits assigned to the full length of the interfaces by the Faulted Earth Project (Table 3; Christoffersen *et al.* 2015). We weight the two depth models equally at 0.5.

For each depth model, two along-strike segmentations are considered: *Segmented*, which uses the boundaries described in Section 3.1.1; and *Unsegmented*, which does not segment the subduction zones. In the *Unsegmented* model, we use a single coupling coefficient per interface, using a representative value for each from the suite of values in the segmented model (see Table 3). One significant reason to include the *Unsegmented* model is to permit ruptures to occur anywhere on the interface or within the slab, as opposed to being constrained by the segment boundaries. A second reason is to allow for rare instances in which a full subduction interface ruptures together, and to exceed or approach the possibility of  $M_w$  9 on each interface included in the model, consistent with the hypothesis that most or all subduction zones can generate  $M_w$  9+ events (e.g. McCaffrey 2008). Because of the workflow we use to produce interface MFDs (e.g. in the tectonic component,  $M_{\max}$  is determined using magnitude scaling relationships), the *Unsegmented* interface inherently introduces an alternative value of  $M_{\max}$ . This is also true for portions of each slab, where the  $M_{\max, \text{obs}}$  increases from that of the segment to that of the full slab. We assign a higher weight (0.7) to the *Segmented* geometry, which—due to its discretization—has more deliberate selection of convergence rates and seismic coupling, whereas the *Unsegmented* geometry uses a single representative value for the full length of the subduction zone. Additionally, the highest magnitudes in the *Unsegmented* model of each subduction interface are not constrained by observations, and so we treat these as particularly rare, regardless of their modelled occurrence rates. Lastly, the *Segmented* version of the intraslab sources permits some control on the spatial distribution of seismicity, since the current methodology does not yet allow for smoothed rupture probabilities. Segmentation is applied separately to the interface and intraslab sources, despite using the same bounds, so this consideration adds four branch sets to the logic tree.

For each segmentation hypothesis, we assign two different values of  $M_{\max}$  ( $M_{\max}1$  and  $M_{\max}2$ ) to the intraslab sources. Here,  $M_{\max}$  is based on statistics, using the convention of  $M_{\max, \text{obs}} + \Delta$  as in the initial implementation in Section 4.4. The hypotheses  $M_{\max}1$  and



**Figure 6.** The logic tree used to incorporate a range of subduction zone source characterizations into the model herein. The first level of division is downdip segmentation, or in other words, the depth used to divide the subduction sources into interface and intraslab, and includes two alternative hypotheses: *depth\_Z1*, which is uniformly 50 km, and *depth\_Z2*, which uses the depths from the Faulted Earth Project (Christophersen *et al.* 2015). The second level of division is the along-strike segmentation. The segmented model corresponds to that described in Section 3.1.1, while in the unsegmented model, each of the major subduction zones is characterized as a single structure. The third and final level of division is  $M_{\max}$ , which is applied to both interface and intraslab sources, but using different methodologies. For the interface sources, the two instances of  $M_{\max}$  are determined using magnitude scaling relationships, while for intraslab sources, different values of  $\Delta$  are added to the  $M_{\max, \text{obs}}$ . Sources across the first level of division are isolated from each other, because allowing these branches to interact would lead to either duplication or absence of known sources; however, all other uncertainties interact, yielding 24 realizations of the source model.

$M_{\max,2}$  are equally weighted, using  $\Delta = 0.3$  and  $\Delta = 0.5$  respectively. In rare cases, ruptures of  $M_{\max}$  cannot fit within the bounds of the slab according to the defined slab volume and selected rupture aspects ratios; in these cases, the magnitude of the largest permitted rupture is used as  $M_{\max}$ .

For the interface sources within each segmentation hypothesis, we include two alternative hypotheses that use different magnitude-scaling relations (MSR): MSR<sub>AH</sub> and MSR<sub>TM</sub>, which use the magnitude-area interface MSR of Allen & Hayes (2017) and Thingbaijam *et al.* (2017), respectively. The two scaling relationships provide different values of  $M_{\max}$ , but also different  $\sigma$  values about the median  $M_{\max}$  (0.266 and 0.15 respectively); this causes the characteristic components of the hybrid MFDs to shift and to stretch or compress relative to each other. Because of the higher  $\sigma$  in the scaling relationship of Allen & Hayes (2017), the Gaussian for most subduction segments is stretched such that the characteristic component does not impact the shape of the MFD (see Supporting Information Figs S4–S7). Allen & Hayes (2017) provide two interface MSRs: a preferable bilinear equation, and—for completeness—a secondary linear equation, both of which are said to be valid up to  $M_w$  9.5. As a default, we use the preferred bilinear equation, with one exception. For the unsegmented Kermadec-Tonga interface extending to 50 km depth, the bilinear equation predicts a median magnitude that greatly exceeds what we consider to be realistic, and which is outside of the legitimate range specified by the authors ( $M_w > 10.6$ ); we instead use the linear equation, which yields  $M_w$  9.5. In all cases, the MSR used to determine  $M_{\max}$  is also used during the classical PSHA calculations.

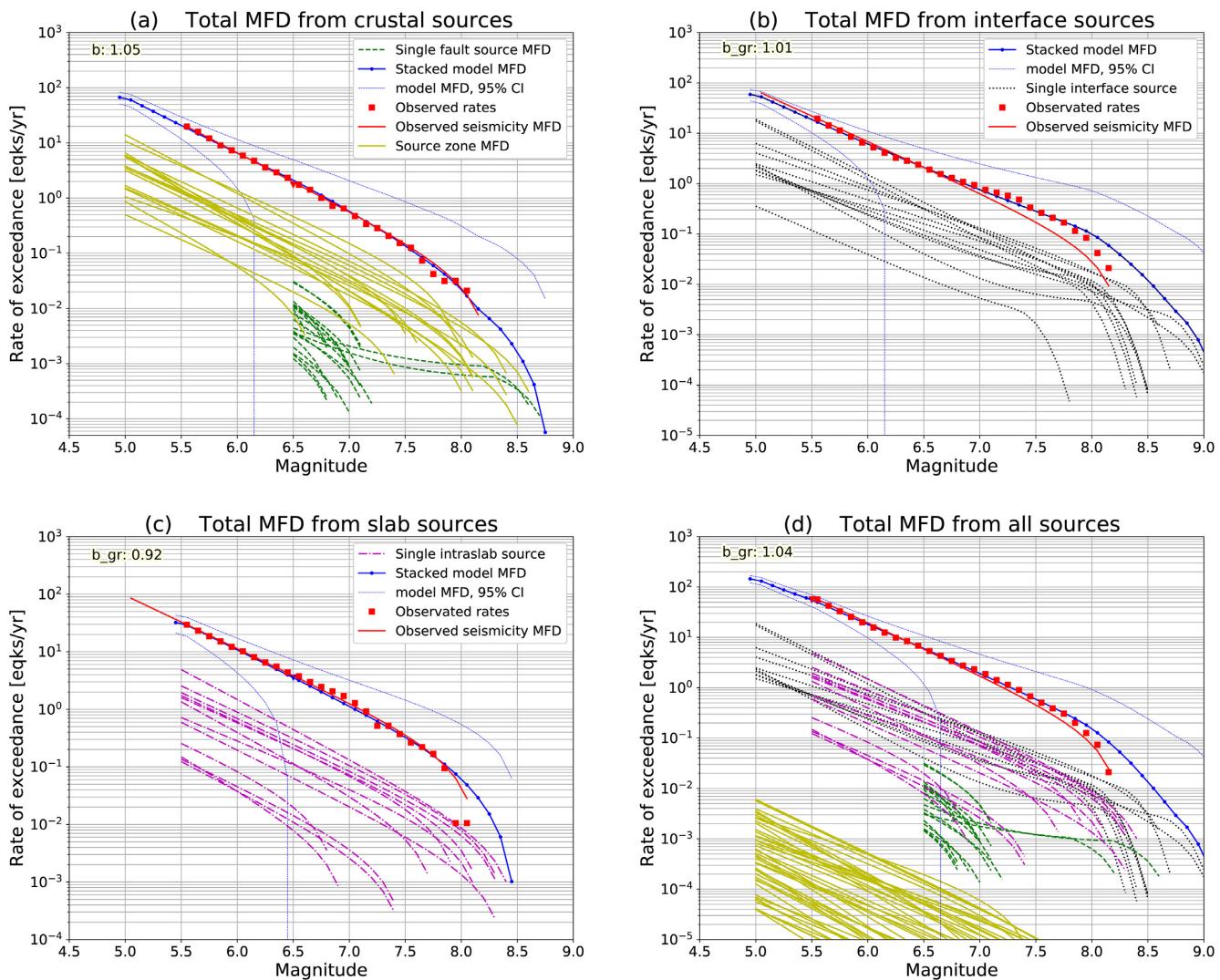
The source characterization for each end branch in general follows the workflow described in the previous sections. However, the unique depth model and segmentation combinations require their

own set of subcatalogues, as noted in Sections 3.1 and 3.2. All resulting MFDs for each branch and each subduction zone or segment are included in Supporting Information Figs S4–S11.

#### 4.6.1 Source model sanity check

A seismic source model should include all sources that pose a shaking hazard to the region of interest as evidenced by earthquake catalogues, as well as the geologic and palaeoseismic records. Here, we provide an initial ‘sanity check’ of the source model developed herein by comparing the modelled sources to the catalogue observations for each tectonic context individually, and for the full model, demonstrating on the branch developed in Section 4 (Fig. 7). In each instance, the catalogue (or subcatalogue) of observations used for comparison is declustered and filtered for completeness using the most conservative completeness pairings of the relevant tectonic contexts. For each context, and the whole region taken together, the observations fit the model MFD to well within the 95 per cent confidence interval. Here, we discuss the possible explanations for data-model misfits observed in Fig. 7.

Fig. 7(a) shows the total MFD from modelled crustal sources (e.g. gridded seismicity and crustal faults) compared to the MFD computed from the observed seismicity. The modelled seismicity captures the observations reasonably well at all magnitudes, with the largest mismatch at  $M_w > 7.5$ : the least frequent and thus most uncertain occurrences. The close match at  $M_w \sim 6.5$  was partly accomplished by our choice to use Youngs and Coppersmith MFDs for the two large thrust faults (Section 4.1); using GR MFDs for all faults resulted in a higher rate of modelled than observed earthquakes at this magnitude. The largest crustal earthquakes admitted by the model are attributable to the two largest faults. These faults



**Figure 7.** Total cumulative MFD for (a) crustal, (b) interface, (c) slab and (d) all sources for Model 1, segmented. Solid yellow lines show the MFDs of selected point sources from all source zones. Dashed green lines show the modelled crustal faults. Dash-dot magenta lines show the MFD for each intraslab segment. Dotted black lines show the MFDs of interface segments. The blue line is the sum of all individual source MFDs for the respective tectonic context. Thin dotted blue lines are the 95 per cent confidence interval for the summed MFDs, calculated as the normal approximation of a Poisson distribution ( $\lambda \pm 1.96\sqrt{\lambda}$ , where  $\lambda$  is the mean rate). Red dots show the observed cumulative rates of earthquakes classified to the respective tectonic contexts included in each frame (declustered and filtered for completeness), and red lines are the GR MFDs fit to these occurrences. We note that in all cases, the initiation of GR MFD tapering is dependent on the maximum magnitude assigned to the truncated MFD.

are actually characterized as subduction interfaces by Bird (2003), and thus their seismogenic depth limits and dips yield large fault areas that permit earthquakes of  $M_w > 8.5$ .

Fig. 7(c) shows the total MFD of modelled intraslab sources compared to the MFD fit to observations. Here, the summed MFD provides a good fit to the observed seismicity, modelling a slightly lower rate of earthquakes in the range of  $M_w$  6.5–7.0 than is observed, while predicting a higher rate than is observed for the largest observed magnitude ( $M_w$  7.8). Here, it is less informative to draw a comparison to the MFD formed by a single GR, since the observations comprise the numerous slabs included in the model, which do not all interact and thus cannot be expected to comply well with a single MFD. This is addressed more in Section 4.6.2.

Fig. 7(b) shows the total MFD of modelled interface sources compared to the MFD fit to observations. Here, the summed MFD from the approach used in Section 4.3 yields a small improvement

to the fit to observations compared to the singular GR MFD. The observations reveal an upward deflection from the exponential shape of a GR MFD at  $M_w \sim 6.5$ , and thus cannot realistically be captured by only GR MFDs. The modelled MFD predicts a higher rate of occurrences with  $M_w > 7.8$ , allowing a  $M_w \sim 8$  or larger earthquake on one of the subduction interfaces in the southern Pacific Islands every 10 yr. One explanation for this overprediction is that we use a catalogue that covers only a short time-frame respective to the life of the subduction systems, and thus the observations do not capture the true average recurrence of the most infrequent earthquakes; one occurrence of a  $M_w$  8.5 or larger earthquake during the instrumental period would change the trend of the observed occurrences. Alternatively, any of our modelling assumptions could be flawed. For example, the interface segmentation model, coupling coefficients, or parameters and scaling relationships used to compute  $M_{max}$  for each segment could be incorrect, our delineation of

the subduction interfaces and classification buffers could have led to misassignment of larger magnitude events, or the completeness thresholds applied to the catalogue could be too optimistic. Lastly, like for intraslab sources, the summed interface MFD is somewhat misrepresentative, since the many interface segments represent seismicity rates of different subduction zones, and thus together may not comply with a single MFD. The characteristics of the summed hybrid MFD in Fig. 7(b) persist in the singular interface segments; that is, while occurrences of most magnitudes are generally fit well by GR MFDs, the highest magnitudes are sometimes better fit by a tectonic (characteristic) component.

Finally, Fig. 7(d) joins the separate tectonic contexts into a singular total MFD for the model, comparing it to the full classified catalogue. Like the subduction contexts, the total model MFD predicts a higher rate of occurrences with  $M_w > 7.5$ , modelling slightly more than one  $M_w 8.0$  or larger earthquake per decade. The full catalogue—not filtered by completeness—includes 11 earthquakes of  $M_w > 8$  from 1900–1914, yielding an annual rate of 0.096 that complies with the rate depicted in Fig. 7. Thus, we deem the modelled rates acceptable, and attribute the mismatch with the observations to the same factors as for interface sources.

#### 4.6.2 Likelihood-based tests

The ‘sanity check’ in Section 4.6.1 qualitatively demonstrates an overall good fit of the source model developed in Section 4 to the observations of the ISC-GEM catalogue, both for individual tectonic region types and for the model as a whole; however, it does not validate the source model quantitatively, or confirm that all the source models included in the seismic source characterization can indeed reproduce the observations. Furthermore, the sanity checks do not evaluate the spatial distribution of seismicity produced by the source models.

We quantitatively evaluate the performances of the source models developed herein using statistical tests originally developed by the Regional Earthquake Likelihood Models (RELM) working group and adapted for use by the Collaboratory for the Study of Earthquake Predictability (CSEP) testing centres (e.g. Schorlemmer *et al.* 2007, 2010; Zechar *et al.* 2010). For each end branch of the source model logic tree, we test retrospectively against the observations used to constrain the models, executing three tests: the N-, M-, and S-tests, which together evaluate the models’ performance in terms of number of events generated, and the distribution of those events in magnitude and space.

All end branches of the source model pass the three tests. Linear regression of the parameters used in each end branch reveals that the most impactful parameter on the test scores is the first parameter of the logic tree: the depth cut-off between the subduction interface and intraslab sources. The branches that use *depth\_Z1* (a uniform value of 50 km; see Section 4 of the main text) perform better on the M- and S-tests than the branches using *depth\_Z2*.

More information on the testing procedure as well as interpretation of the test results is included in the Supporting Information.

## 5 GROUND MOTION CHARACTERIZATION

Due to the scarcity of land in the Pacific Islands region, station coverage is sparse, and regional ground motion prediction equations (GMPEs) have not been developed. Furthermore, residual analysis is challenging, as there are very few strong motion records available

with source-site distances  $<300$  km (the source-site distance cut-off for most tectonic region types used herein, and the limit of the data used to develop many of the GMPEs). In order to overcome this challenge, Petersen *et al.* (2012) supplemented Pacific Islands strong motion data with recordings from around the Pacific Basin, finding that strong-motion recordings for earthquakes occurring in the western Pacific (ranging north–south from the Aleutian Islands to Tonga, and as far west as Taiwan) were best predicted by Zhao *et al.* (2006).

Another approach to choosing GMPEs when limited data are available is to use selection criteria such as those of Bommer *et al.* (2010), eliminating GMPEs in the greater database until only a small number remain. These criteria include consideration of the intended tectonic context; applicability of the data used to derive the GMPE to the source model, e.g. appropriate magnitude and distance ranges; details of the regression; and whether or not the GMPE has undergone peer review. Ghasemi *et al.* (2016), combined the results of the residual analysis by Petersen *et al.* (2012) with these selection criteria to develop a ground motion model logic tree for PSHA of Papua New Guinea. They use Zhao *et al.* (2006), Boore & Atkinson (2008) and Chiou & Youngs (2008) weighted at 0.4, 0.3 and 0.3, respectively for active shallow crustal sources, and Youngs *et al.* (1997), Atkinson & Boore (2003) and Zhao *et al.* (2006) weighted at 0.3, 0.3 and 0.4, respectively, for both types of subduction sources. The higher weight assigned to Zhao *et al.* (2006) is based on the results of Petersen *et al.* (2012).

We initially adopt the logic tree of Ghasemi *et al.* (2016), since the Papua New Guinea model has overlapping coverage with the model presented herein. Together, the GMPEs for each tectonic region type encompass a range of possible values for each magnitude and distance combination, which partly compensates for our inability to constrain the ground motions using recorded data. In the case of the crustal sources, Zhao *et al.* (2006) tends to predict the highest ground motions for the magnitudes and distances considered. For subduction sources, it predicts the median of the three GMPEs included in the logic tree for most magnitudes and distance pairs, but becomes the maximum for the highest magnitudes ( $M_w > 8.0$ ).

For the model presented herein, we update the adopted logic based on the following, and with focus on subduction interface and intraslab (NB: these sources are revealed by disaggregation to contribute most to the hazard in capital and large cities in the Pacific Islands, as discussed in Section 7.3). First, we consider that new subduction GMPEs meant to be globally applicable have emerged since the model of Ghasemi *et al.* (2016) was developed. In particular, Abrahamson *et al.* (2016) provided subduction interface and intraslab GMPEs that use more recent and more widespread data than prior models. The Abrahamson *et al.* (2016) interface GMPE has been checked against recordings from the 2010  $M_w 8.8$  Maule and 2011  $M_w 9.1$  Tohoku earthquakes (these earthquakes occurred after the database of strong motion data were compiled). Secondly, at the highest magnitudes included in the source model herein, the GMPE of Zhao *et al.* (2006) diverges from other globally applicable subduction GMPEs, predicting exceedingly high values (e.g. Supporting Information Figs S2 and S3). This is particularly consequential for the unsegmented subduction interfaces, and specifically the Kermadec-Tonga interface, for which  $M_{max} = 9.5$  when computed using the MSR of Allen and Hayes (2017). Thus, for subduction tectonic region types, we replace Zhao *et al.* (2006) with Abrahamson *et al.* (2016), and equalize the logic tree weights. The final ground motion model logic tree is summarized in Table 7.

**Table 7.** Ground motion prediction equations by tectonic region and weight. When multiple versions are available, the original form of the relevant tectonic region type was used.

Tectonic region	GMPE	weight
Active shallow crust	Zhao et al. (2006)	0.4
	Boore & Atkinson (2008)	0.3
	Chiou & Youngs (2008)	0.3
Subduction interface	Youngs et al. (1997)	0.33
	Atkinson & Boore (2003)	0.33
	Abrahamson et al. (2016)	0.34
Subduction intraslab	Youngs et al. (1997)	0.33
	Atkinson & Boore (2003)	0.33
	Abrahamson et al. (2016)	0.34

## 6 RESULTS

We use the OpenQuake Engine (Pagani *et al.* 2014) to compute the hazard across the southern Pacific Islands, focusing on return periods of 475 and 2475 yr, corresponding to a 10 per cent and 2 per cent probability of exceedance (PoE) in 50 yr. Here, we discuss the map distribution of mean peak ground acceleration (PGA) and spectral acceleration (SA) for a period ( $T$ ) of 1.0 s, as well as PGA hazard curves and uniform hazard spectra (UHS) for several capital and large cities. Throughout, hazard is computed for reference rock site conditions estimated by shear wave velocity in the upper 30 meters of crust of 800 m/s.

We first observe hazard across the full region. Fig. 8 shows PGA and SA 1.0 s with 10 and 2 per cent PoE in 50 yr. The subduction sources—and especially the interfaces—are most evident within the hazard pattern, with the highest hazard along the New Hebrides and South Solomon trenches. The Fiji Fracture Zone and some other areas with seafloor faults have more subtle concentrations of hazard. The Gaussian smoothing used to model distributed crustal sources is evident as ‘spots’ away from the more significant sources, indicating the locations of past seismicity. Onshore, PGA values for 10 and 2 per cent PoE in 50 yr reach 1.23 and 2.26 g, respectively, on the Solomon Islands island of Guadalcanal. Here, SA 1.0 s reaches 0.79 and 1.41 g for the same PoEs. Notably, the onshore areas with the highest PGA (e.g. >1.0 g for 10 per cent PoE in 50 yr) are confined to a narrow band of land within ~80 km of the nearest subduction trench.

Fig. 9 shows PGA with 10 per cent and 2 per cent PoE in 50 yr for the Solomon Islands. PGA ranges from 0.33–1.23 g and 0.62–2.26 g at 10 per cent and 2 per cent PoE in 50 yr, respectively. In general, hazard contours are approximately parallel to the Solomon trench, and PGA decreases with distance. In the capital city of Honiara, PGA is 0.95 and 1.68 g for 10 and 2 per cent PoE, respectively (Fig. 12).

Fig. 10 shows the PGA with 10 and 2 per cent PoE in 50 yr for New Caledonia and Vanuatu. In New Caledonia, low hazard persists for both return periods across most of the land area. On the main island, PGA with 10 and 2 per cent in 50 yr reaches 0.15 and 0.30 g; however, on the smaller, Loyalty islands—closer to the New Hebrides trench—these values increase to 0.38 and 0.61 g. Noumea, the capital city, is along the southwest coast of the island, and far from any well-defined active tectonic structure (crustal Source Zone 18, Table 3), with PGA at 10 and 2 per cent PoE in 50 yr of only 0.11 and 0.20 g, rendering it one of the least seismically hazardous locations within the southern Pacific Islands.

The islands of Vanuatu sit on the overriding plate of the subduction zone, and are overall closer than New Caledonia to the trench itself, and thus experience consistently higher hazard. Here, PGA

with 10 and 2 per cent PoE in 50 yr ranges from 0.72 to 1.13 g and 1.17 to 2.00 g, respectively. The hazard pattern is generally consistent with that of the Solomon trench in that the PGA contours trend parallel to the trench, increasing trenchward; however, the hazard gradient in the dip direction is much lower in Vanuatu, and so hazard is higher for most of the land area. The capital city, Port Vila, on the island of Malekula has PGA of 0.83 and 1.37 g for 10 and 2 per cent PoE in 50 yr, while Luganville—a populous city on the island of Espiritu Santo to the northwest—is slightly higher at 0.91 and 1.49 g, respectively.

The higher PGA values in Vanuatu compared to the Solomon Islands reflect the different geometry and convergence properties of the two subduction zones; the New Hebrides interface and slab extend farther from the trench than those of the South Solomon Trench, and have higher seismic productivity. Segment 2 of the New Hebrides subduction zone—along which many of the Vanuatu islands are located—has a higher convergence rate than the full length of the Solomon trench. Additionally, the crustal sources that encompass most of the Vanuatu Islands (smoothed seismicity within Source Zone 10) have more frequent occurrences than those covering the Solomon Islands (Source Zone 3). The highest concentration of earthquakes in Source Zone 10 overlaps with the surface projection of New Hebrides intraslab Segment 3, corresponding to where plate convergence is accommodated by crustal faults in the Pacific plate.

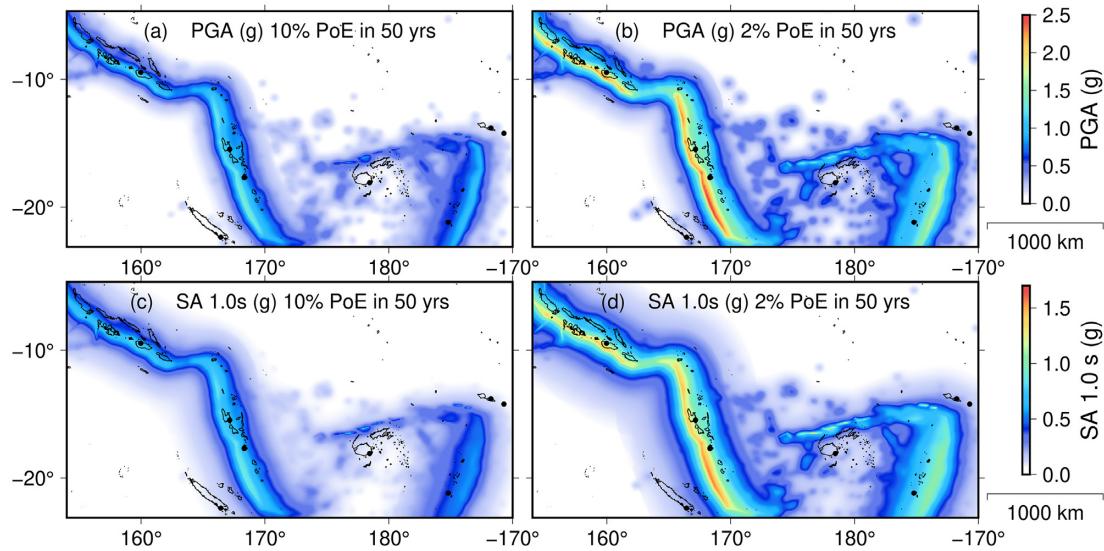
Fig. 11 shows mean PGA with 10 and 2 per cent PoE in 50 yr for the Fiji islands. PGA values range from 0.15–0.56 g for 2 per cent PoE in 50 yr and 0.09–0.34 g for 10 per cent PoE in 50 yr and are generally low considering their central location within the tectonic complexity of the southern Pacific Islands. The capital city of Suva has a PGA of 0.14 and 0.28 g for 10 and 2 per cent PoE in 50 yr.

## 7 DISCUSSION

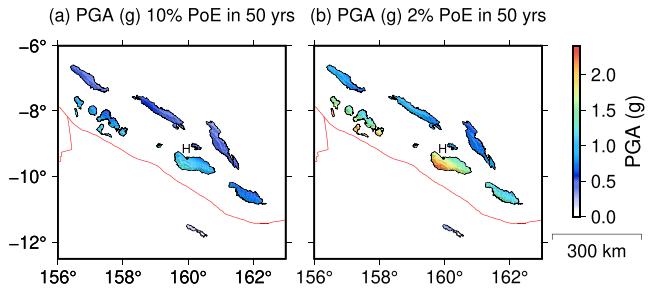
### 7.1 Impact of epistemic uncertainties

The seismic source models included in the seismic source characterization contribute to the range of hazard results for each city in Fig. 12. Here, we evaluate the impact of the alternative hypotheses considered in the logic tree by comparing hazard curves computed for two source model logic tree end branches that differ by only on parameter; for example, we compare the change due to use of *depth\_Z1* versus *depth\_Z2* for a fixed segmentation and means of computing  $M_{\max}$ , repeating this test for each branch set in the logic tree structure (e.g. each uncertainty; Supporting Information Figs S15–S19). For each test, the same ground motion models are used: Boore & Atkinson (2008) for crustal sources, Abrahamson *et al.* (2016) for interface sources, and Youngs *et al.* (1997) for intraslab sources, all of which use the site-to-rupture distance metric. This evaluation provides an initial look at the sensitivity of the hazard results to source model parameters, but is not considered a comprehensive sensitivity analysis.

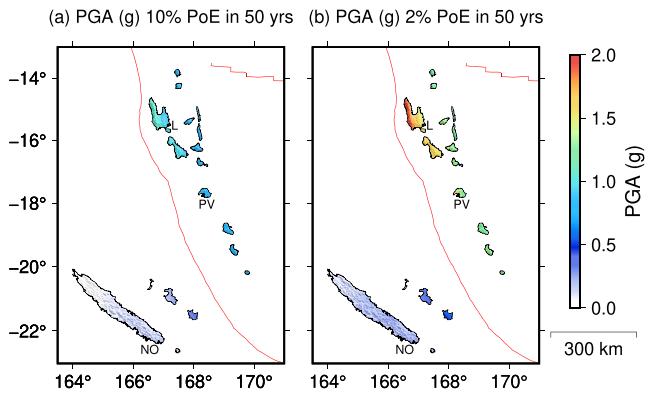
We find that the impact of each epistemic uncertainty on the hazard results varies by city, and depends on factors such as the position of the city relative to the proposed subduction segment boundaries and interface-intraslab contacts of the closest subduction zone, and the characteristics of the respective subduction zone. For example, cities near the New Hebrides subduction zone have higher hazard due to the unsegmented interfaces, those cut using *depth\_Z2*, and those using the Allen & Hayes (2017) magnitude scaling relationship with respect to their counterparts, while the opposite is true



**Figure 8.** Mean hazard for the full region (on and offshore), showing (a) PGA (g) with 10 per cent and (b) 2 per cent PoE in 50 yr, and (c) SA ( $T=1.0\text{s}$ ) (g) with 10 per cent and (d) 2 per cent PoE in 50 yr on rock. Scales are for a reference latitude of  $16^\circ\text{S}$ .

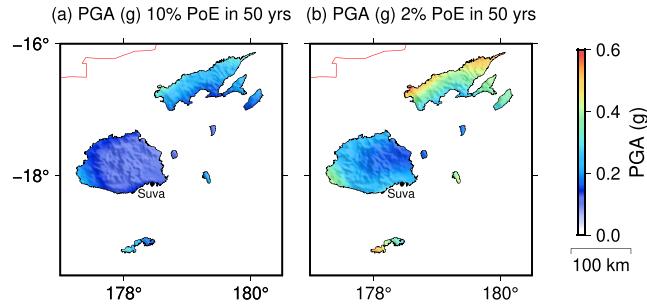


**Figure 9.** Mean PGA for the Solomon Islands, (a) 10 per cent and (b) 2 per cent probability of exceedance in 50 yr on rock. Red lines show faults of the GAF-DB (Styron & Pagani 2020). Islands with area  $< 100 \text{ km}^2$  are not plotted. Scale corresponds to  $10^\circ\text{ S}$ . H: Honiara.



**Figure 10.** Mean PGA for Vanuatu and New Caledonia, (a) 10 per cent and (b) 2 per cent probability of exceedance in 50 yr on rock. Red lines show faults of the GAF-DB (Styron & Pagani 2020). Islands with area  $< 100 \text{ km}^2$  are not plotted. Scale corresponds to  $18^\circ\text{ S}$ . L: Luganville. PV: Port Vila. NO: Noumea.

for Nuku’alofa, which is situated above the Kermadec-Tonga slab. The use of a segmented versus unsegmented slab and the intraslab maximum magnitude have only a small impact on the hazard curves for most cities included in this analysis, with the highest impact on

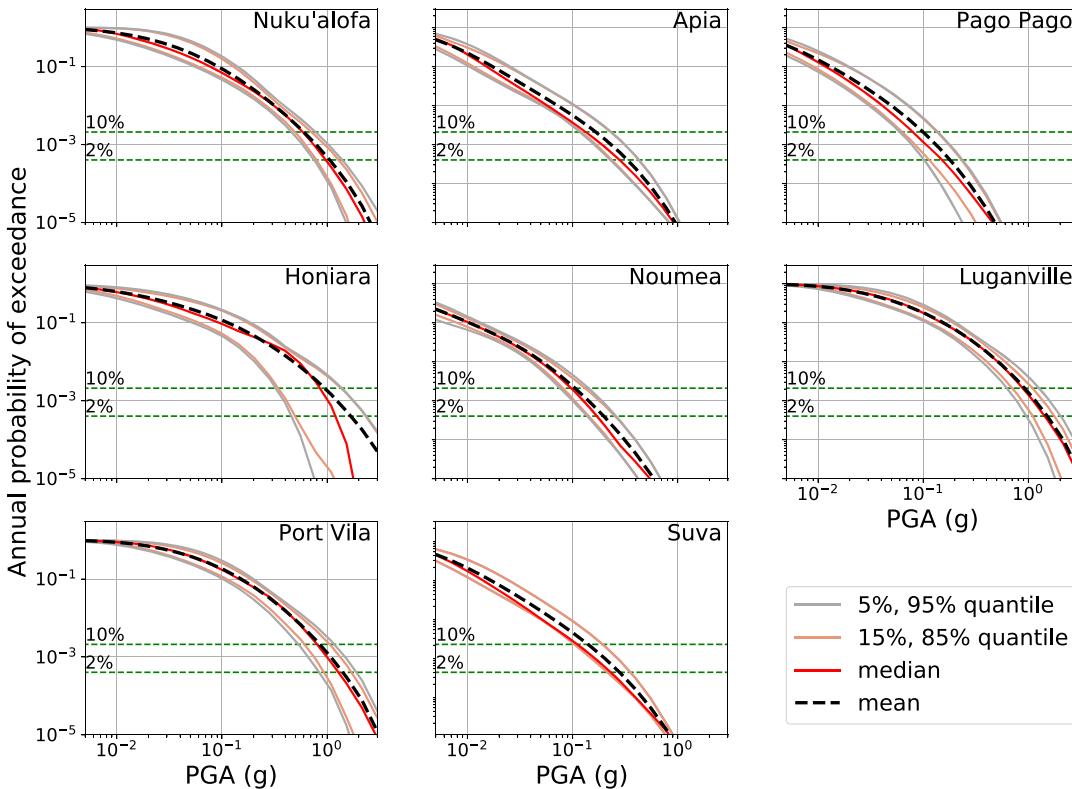


**Figure 11.** Mean PGA for Fiji, (a) 10 per cent and (b) 2 per cent probability of exceedance in 50 yr on rock. Red lines show faults of the GAF-DB (Styron & Pagani 2020). Islands with area  $< 100 \text{ km}^2$  are not plotted. Scale corresponds to  $18^\circ\text{ S}$ .

cities situated within the mapped perimeter of the slab sources (e.g. Nuku’alofa, Port Vila; see Fig. 3).

## 7.2 Comparison to former models and other regions

The earlier model that overlaps most with the one presented herein is by Petersen *et al.* (2012), which covers the southern Pacific Islands with a western limit of Vanuatu. The general characteristics of the hazard patterns between the two models are similar, with the highest hazard concentrating along subduction zones and near the Fiji Fracture Zone. In Fig. 13, we compare intensity measure levels at significant sites as computed by the present model versus Petersen *et al.* (2012). In Nuku’alofa, the Petersen *et al.* (2012) value at SA ( $T = 0.2 \text{ s}$ ) with a 2 per cent PoE in 50 yr exceeds the mean hazard computed here by  $\sim 0.5 \text{ g}$ , and is also outside the 0.95 quantile (Supporting Information Fig. S12), but at higher spectral periods (e.g.  $T = 1.0 \text{ s}$ ), the absolute discrepancy decreases. In Port Vila, mean SA for  $T = 0.2$  and  $T = 1.0 \text{ s}$  values from Petersen *et al.* (2012) are close to the mean and median quantile from the present model, respectively. Away from the subduction zones in Suva and Pago Pago, the model herein produces higher values of hazard and in some cases a notable relative difference from the Petersen *et al.* (2012) results; however, the values are similar in absolute terms,



**Figure 12.** Hazard curves showing the annual probability of exceedance for several capital or populous cities within the southern Pacific Islands. The dashed black lines show the hazard curve for mean PGA—the predominant result discussed throughout the text. The other curves are PGA quantiles: bright red shows median PGA (quantile 0.5), faint red shows quantiles 0.15 and 0.85, and grey shows quantiles 0.05 and 0.95.

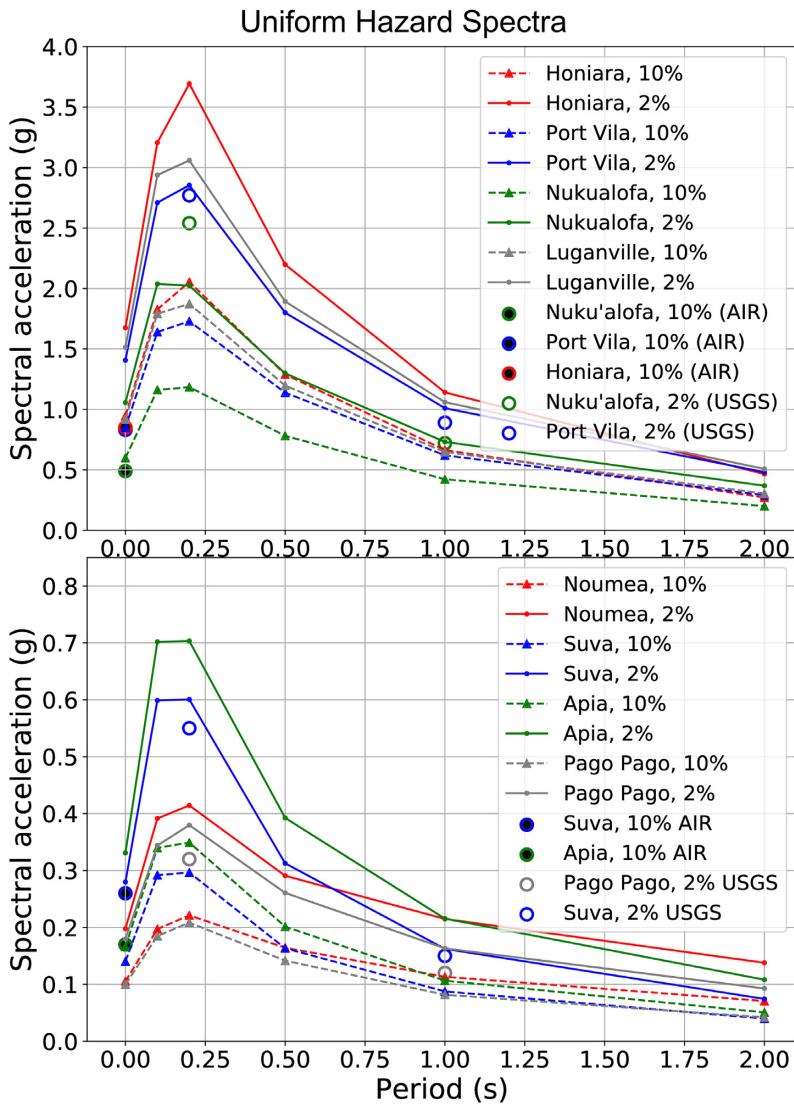
differing by  $<0.1$  g (Fig. 13), and the Petersen *et al.* (2012) values always fall within the range of values computed by the present model (Supporting Information Fig. S12).

We suspect that the smaller discrepancies between the two sets of results merely exemplify a difference in modelling choices, such as completeness thresholds and declustering windowing in the initial catalogue preparation, and the means of classifying seismicity to the different tectonic regions. However, in the case of Nuku'alofa, we suggest that the difference in selected GMPEs contributes strongly. The Petersen *et al.* (2012) model solely uses Zhao *et al.* (2006) for subduction sources, accounting for epistemic uncertainty with vertical shifts. On the other hand, we excluded Zhao *et al.* (2006) due to its increasing divergence from the other models at high magnitudes, e.g.  $M \geq 9.0$ , included in both models, and in particular on the Kermadec-Tonga trench.

The model by Suckale & Grünthal (2009) covers the islands of Vanuatu. The hazard patterns computed from the model herein are similar to the Suckale & Grünthal (2009) results for PGA with a 10 per cent PoE in 50 yr on rock; however, the values from the present model are again higher. Suckale & Grünthal (2009) report onshore values in the range 0.51 to 0.66 g, while we compute values of 0.72 to 1.13 g, ranges that do not overlap. The reasons for this discrepancy are likely similar to those for Petersen *et al.* (2012): the two models are based on different catalogues and catalogue pre-processing assumptions; use different methods to classify earthquakes to the tectonic regions; and use different GMPEs. Additionally, the Suckale & Grünthal (2009) model uses only distributed seismicity sources, whereas we include faults to model the subduction sources.

The model presented herein covers approximately the same region as Rong *et al.* (2010). Their study presents hazard computed for different soil conditions, and so the results are generally difficult to compare; however, they do include PGA on rock with 10 per cent PoE in 50 yr for a few capital cities. For capital cities near subduction zones (Nuku'alofa, Port Vila, and Honiaria) as well as in Apia, computed hazard values are within the range of values computed here (and within  $\sim 0.1$  g of the mean). In Suva, the difference is slightly larger (Fig. 13b), and Rong *et al.* (2010) compute a higher value ( $\sim 0.26$  g versus 0.14 g) which exceeds the 0.95 quantile value computed here. A possible explanation is that Rong *et al.* (2010) used a different database of faults and alternative methodologies to constrain their rates; their source model includes a crustal fault to the southeast of Fiji which is absent from the present model. They also use a different group of GMPEs for crustal sources.

The discrepancies in mean PGA values, and in particular the increase in hazard, between the present model and the former ones discussed in this section are mostly attributable to changes in modelling approaches and GMPEs. Thus, we additionally compare the highest hazard values computed here to those for other global subduction zones in order to assess whether the Pacific Islands hazard, as represented by this model, is uniquely high. We identify other areas of the hazard map computed from the GEM Hazard Mosaic of the Global Earthquake Model (Pagani *et al.* 2020a) where the mean PGA on rock with 10 per cent PoE in 50 yr exceeds 1.0 g. In addition to the Pacific Islands, the GHM exceeds 1.0 g on the small Indonesia islands of Pagai-utara and Pagai-seletan; over small lengths of coastal Honshu and Hokkaido in Japan (areas that protrude trenchward with respect to the coastlines); central to northern



**Figure 13.** Uniform hazard spectra on rock for several capital or populous cities within the southern Pacific Islands. (a) Shows cities with overall higher hazard values and (b) for lower hazard sites. Points show site values reported by other studies. AIR: Rong *et al.* (2010, fig. 5). USGS: Petersen *et al.* (2012).

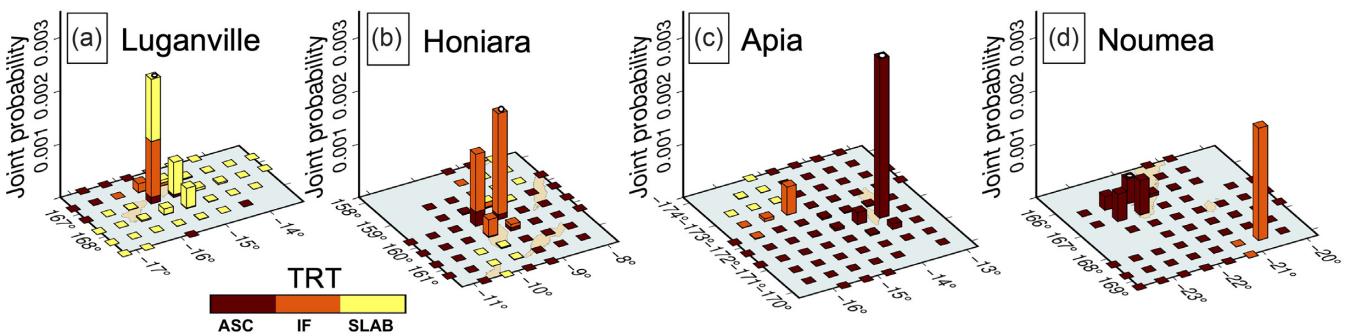
coastal Chile; and New Britain of Papua New Guinea (also covered by the model presented herein). We note that some of these areas have experienced  $M_w > 9.0$  earthquakes during the instrumental period, while others—like the Pacific Islands—encompass land area that is particularly close to the subduction trench (e.g. <100 km); however, a more thorough evaluation of these PGA global maxima is beyond the scope of this study.

### 7.3 Disaggregation

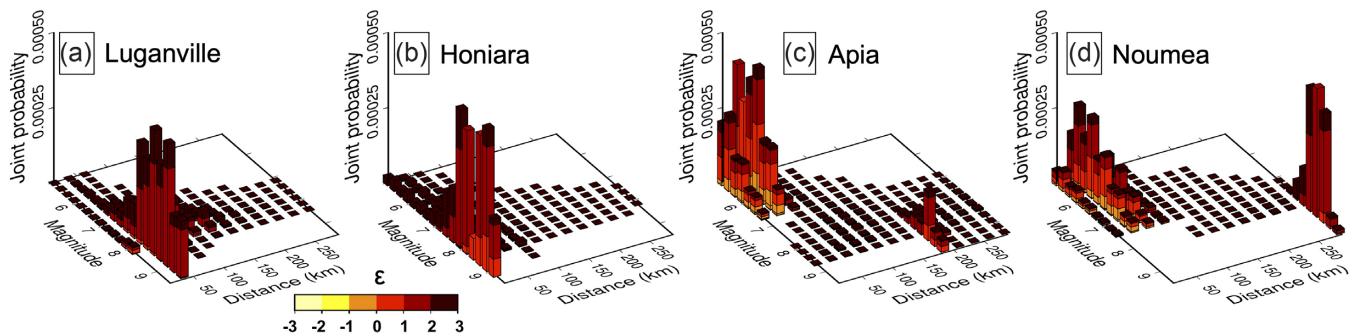
Seismic hazard disaggregation is a way to analyse how the different components in a hazard model contribute to the ground-shaking values computed for a given return period (e.g. Bazzurro & Cornell 1999; Pagani & Marcellini 2007). Here, we disaggregate the PGA with 10 percent PoE in 50 yr on rock for several cities (Figs 14 and 15); all of the text in this section refers to this intensity metric and return period. We focus primarily on capital or populous cities, selecting Luganville, Vanuatu; Apia, Samoa; Honiara, Solomon Islands; and Noumea, New Caledonia. Two of the selected cities are close to local hazard peaks (Luganville and Honiara) and two are

farther away from the most productive seismic sources (Apia and Noumea). We examine disaggregation for two parameter configurations: (1) tectonic region type by position (latitude and longitude) and (2) magnitude and distance, including the range of ground motion uncertainties ( $\epsilon$ ).

Luganville is on the southeastern shore of Espiritu Santo—the largest of the Vanuatu islands—and close in map-view to the defined interface-slab limit for both *depth\_Z1* and *depth\_Z2* of New Hebrides trench Segment 3. The computed PGA is 0.92 g. The predominant tectonic contexts contributing to ground-shaking hazard in Luganville are intraslab ruptures, especially those just to the north and east of the city (e.g. <0.5° away, Fig. 14a), and interface earthquakes that rupture directly beneath the city and to its west. Despite the high rate of upper-plate seismicity near Luganville, crustal earthquakes have a very minor role. Fig. 15(a) reveals that most ground-shaking hazard is due to  $M_w$  7.0 and larger earthquakes 25–150 km from the site. The sources that contribute most are within 50 km of Luganville with  $M_w$  7.5–8.5, corresponding to both interface and intraslab sources on New Hebrides Segment 3, as well as the unsegmented versions of both tectonic regions; the MFDs



**Figure 14.** Disaggregation of the mean PGA computed at 10 per cent PoE on rock for four large or capital cities in southern Pacific Islands nations: (a) Luganville at  $167.16^{\circ}\text{E}$ ,  $15.52^{\circ}\text{S}$ , PGA = 0.91g, (b) Honiara at  $159.97^{\circ}\text{E}$ ,  $9.45^{\circ}\text{S}$ , PGA = 0.95g, (c) Apia at  $171.75^{\circ}\text{W}$ ,  $13.85^{\circ}\text{S}$ , PGA = 0.20g and (d) Noumea at  $166.44^{\circ}\text{E}$ ,  $22.27^{\circ}\text{S}$ , PGA = 0.11g. Disaggregation is by source latitude ( $0.5^{\circ}$  bin width), longitude ( $0.5^{\circ}$  bin width) and tectonic region type. The white dot indicates the spatial bin that includes the city of interest. The colour bar indicates crustal (ASC), interface (IF) and intraslab (SLAB) sources.



**Figure 15.** Disaggregation of the mean PGA computed at 10 per cent PoE on rock for four capital cities in the southern Pacific Islands: (a) Luganville at  $167.16^{\circ}\text{E}$ ,  $15.52^{\circ}\text{S}$ , PGA = 0.91g, (b) Honiara at  $159.97^{\circ}\text{E}$ ,  $9.45^{\circ}\text{S}$ , PGA = 0.95g, (c) Apia at  $171.75^{\circ}\text{W}$ ,  $13.85^{\circ}\text{S}$ , PGA = 0.20g and (d) Noumea at  $166.44^{\circ}\text{E}$ ,  $22.27^{\circ}\text{S}$ , PGA = 0.11g. Disaggregation is by source magnitude, distance and  $\epsilon$  defined by standard deviation of the GMPEs.

for both segments and both tectonic region types predict  $M_w > 7$  earthquakes on the order of tens of yr.

Like Luganville, Honiara is located along an island arc that parallels a subduction zone; however, the different subduction geometry and productivity of the South Solomon trench compared to the New Hebrides trench results in a different balance of source contributions. Fig. 14(b) shows that the ground-shaking hazard in Honiara (PGA of 0.95 g) is almost entirely controlled by subduction interface seismicity. Although Honiara is close to the surface projection of the interface-slab boundary, seismicity rates within the South Solomon slab Segments 1 and 2 are lower than those of the New Hebrides trench, and the slab itself is less expansive (Fig. 3). By magnitude, the most hazardous sources to Honiara are predominantly  $M_w > 7$  earthquakes close to the city, and especially those within 25 km (Fig. 15b). The magnitude bins with the highest contributions are  $M_w \sim 8\text{--}9$ , which have return periods as low as  $\sim 100$  yr ( $M_w 8$ , Solomon interface Segment 1; Supporting Information Fig. S4).

Apia, Samoa is located within a region of dispersed crustal seismicity beyond the outer rise of the Kermadec-Tonga subduction zone (Source Zone 17, see Table 3 and Fig. 2). The disaggregation results in Figs 14(c) and 15(c) reveal that the highest contributors to the hazard in Apia are nearby crustal sources ( $<50$  km away) with  $M_w < 7$ , and peaking at  $M_w < 6$ . This is logical considering the lower mean PGA (0.17 g), and the lower-magnitude earthquakes that occurs in nearby source zones more closely associated with subduction processes. However, interface earthquakes at 175 km distance are also contributing non-negligibly, and may have an even higher contribution to the ground-shaking hazard at longer return periods.

Noumea, New Caledonia is also located within a source zone of dispersed crustal seismicity (Source Zone 18, see Table 3 and Fig. 2). Here, the PGA value is low (0.11 g) due to a lack of proximal, high-magnitude seismicity relative to other sites, and therefore the disaggregation reveals a relatively higher contribution by more distant sources. Figs 14(d) and 15(d) show that while  $M_w < 7$  crustal sources within 100 km of Noumea contribute most to the ground-shaking hazard (and especially those at distances  $<50$  km), subduction sources at  $>250$  km distances with  $M_w > 8$ —namely, the New Hebrides Segment 2 interface, and instances in which the New Hebrides interface is unsegmented—also contribute. At this site, the difference in source prominence is also visible in the UHS when compared to sites with similar mean PGA values (Fig. 13). The intensity measure levels for Noumea with both 10 per cent and 2 per cent PoE in 50 yr at periods  $<1.0$  s are lower than for the other compared cities, but for higher periods, Noumea has the highest spectral accelerations.

#### 7.4 Model uncertainty

The model presented herein is meant to include all foreseeable earthquake occurrences within the area of interest that pose a hazard, as well as the range of possible ground motions that these sources could cause. Here, we discuss the epistemic uncertainties that are included, and shortcomings to address in future versions.

The ground motion model uses a set of GMPEs for each tectonic context to capture epistemic uncertainty, accounting for the range of realistic possibilities. As stated, GMPE selection for the Pacific

Islands region is difficult due to a scarcity of seismic records within the source-site distance range considered by the hazard model presented here (e.g. 0–300 km for most source types), and especially for short distances where the strongest shaking will occur (e.g. Petersen *et al.* 2012; Ghasemi *et al.* 2016). The logic tree defined by Ghasemi *et al.* (2016) and the modified version used here capture a range of ground motions that cannot be ruled out by the limited available records. However, as more strong-motion data becomes available for the Pacific Islands, residual analyses should be repeated to test the more recent GMPEs, including the updates to the NGA models. We note that while Bommer *et al.* (2010) recommends using the most recent versions of any GMPE, use of the 2014 NGA models will not significantly change the computed ground motions at most sites for the PoEs considered here; crustal sources are seldom the most significant contributor to hazard, and the ruptures that occur on the modelled crustal faults tend to have low impact to sites onshore for the return periods investigated here.

In the seismic source model, we focused treatment of epistemic uncertainty on the subduction sources, taking into account the downdip and along-strike source segmentation, and the maximum magnitude permitted by the sources. Because of the methodology used to constrain the rates of interface sources, including maximum magnitude as an uncertainty changes the *shape* of MFDs in addition to the upper magnitude truncation; this aspect partially accommodates for the variability in occurrence rates that would arise from using alternative coupling coefficients and convergent rates, but future versions of this model should consider these parameters which control the tectonic component of the MFD independently as more research becomes available. Furthermore, epistemic uncertainty could be included in the characterization of active shallow crustal sources.

## 8 CONCLUSION

Herein, we have presented a PSHA model for the southern Pacific Islands region. The seismic source model uses a combination of smoothed seismicity, 3D faults, and non-parametric ruptures to capture the complex tectonics of the region, including four subduction zones and a range of seafloor faulting mechanisms. The ground motion model uses weighted GMPEs for each tectonic region type. Here, we have discussed the hazard results for reference rock conditions across a ~10-km grid for PoEs of interest to engineering.

The model results and hazard disaggregation show that the highest hazard is due to subduction sources, and islands within the surface projections of the interface and intraslab volumes show the highest values of PGA on rock with 10 per cent PoE in 50 yr, exceeding 1.0 g in some places. For the return periods evaluated here, the PGA hazard is often due to nearby sources at larger magnitudes ( $M_w > 7.0$ ). However, at sites with lower hazard (e.g. Samoa, American Samoa, Fiji, and New Caledonia), smaller magnitudes are more significant. At longer spectral periods (e.g. 2.0 s), the significance of subduction interface sources increases; importantly, these interface sources additionally pose the threat of secondary hazards, namely tsunamis, to the Southern Pacific Islands.

This hazard model was originally created to provide homogeneous coverage of the Southern Pacific Islands—ranging from the Solomon Islands to Tonga—in the GEM Hazard Mosaic (Pagani *et al.* 2020a). Although spatially homogeneous, we acknowledge that the degree of investigation into different model components was variable. The seismic source model received more focus than

the ground motion model in this work. The source model uses rigorous catalogue classification, carefully defined source geometries, and occurrence rates based on both seismicity and tectonics, aiming for a realistic source characterization both spatially and temporally, and accounting for the widely varied tectonics. In particular, the model makes use of a subduction source modelling procedure by Pagani *et al.* (2020b) that defines finite ruptures for intraslab earthquakes, incorporates a characteristic component into MFDs that considers convergence rate and seismic coupling, and accounts for some of the epistemic uncertainties in interface segmentation and the maximum magnitude of subduction earthquakes. In general, the source model captures the observed seismicity reasonably well, including the rare palaeoseismic and palaeogeodetic rates that sample longer observation periods. The ground motion model was adopted from a model with coverage of a neighbouring region (Ghasemi *et al.* 2016, with the influence of Petersen *et al.* 2012), but modified based on the availability of new GMMs, and should be a focus of future modelling efforts.

## ACKNOWLEDGEMENTS

We are grateful to Kelvin Berryman, who did a preliminary review of this model, and to our colleagues Robin Gee and Thomas Chartier for engaging in insightful discussions during the revision process. We also thank Matteo Taroni and an anonymous reviewer for their constructive reviews, which greatly improved the PSHA model we developed, as well as the publication. We also acknowledge the support provided by public and private participants of the GEM Foundation ([www.globalquakemodel.org](http://www.globalquakemodel.org)). The views and opinions expressed in this paper are those of the authors and do not necessarily reflect the official policy or position of the GEM Foundation. Maps were created using QGIS and the Generic Mapping Tools software (Wessel *et al.* 2013).

## REFERENCES

- Abrahamson, N., Gregor, N. & Addo, K., 2016. BC Hydro ground motion prediction equations for subduction earthquakes, *Earthq. Spectra*, **32**(1), 23–44.
- Allen, T.I. & Hayes, G.P., 2017. Alternative rupture-scaling relationships for subduction interface and other offshore environments, *Bull. seism. Soc. Am.*, **107**(3), 1240–1253.
- Anderson, E.M., 1905. The dynamics of faulting, *Trans. Edinburgh Geol. Soc.*, **8**(3), 387–402.
- Atkinson, G.M. & Boore, D.M., 2003. Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions, *Bull. seism. Soc. Am.*, **93**(4), 1703–1729.
- Baillard, C., Crawford, W.C., Ballu, V., Régnier, M., Pelletier, B. & Garaebiti, E., 2015. Seismicity and shallow slab geometry in the central Vanuatu subduction zone, *J. geophys. Res.*, **120**(8), 5606–5623.
- Bazzurro, P. & Cornell, C.A., 1999. Disaggregation of seismic hazard, *Bull. seism. Soc. Am.*, **89**(2), 501–520.
- Bird, P., 2003. An updated digital model of plate boundaries, *Geochim. Geophys. Geosyst.*, **4**(3), doi:10.1029/2001GC000252.
- Bird, P., Kagan, Y.Y., Jackson, D.D., Stein, S. & Freymueller, J., 2002. Plate tectonics and earthquake potential of spreading ridges and oceanic transform faults, *Plate Boundary Zones*, AGU Monograph, Vol. 30, pp. 203–218, eds Stein, S. & Freymueller, J.T., Geodyn. Ser., doi:10.1029/GD030p0203.
- Boettcher, M.S. & Jordan, T.H., 2004. Earthquake scaling relations for mid-ocean ridge transform faults, *J. geophys. Res.*, **109**(B12302), 21, doi:10.1029/2004JB003110.

- Bommer, J.J., Douglas, J., Scherbaum, F., Cotton, F., Bungum, H. & Fäh, D., 2010. On the selection of ground-motion prediction equations for seismic hazard analysis, *Seismol. Res. Lett.*, **81**(5), 783–793.
- Bonnardot, M.-A., Régnier, M., Ruellan, E., Christova, C. & Tric, E., 2007. Seismicity and state of stress within the overriding plate of the Tonga-Kermadec subduction zone, *Tectonics*, **26**(5), doi:10.1029/2006TC002044.
- Boore, D.M. & Atkinson, G.M., 2008. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthq. Spectra*, **24**(1), 99–138.
- Calmant, S., Pelletier, B., Lebellegard, P., Bevis, M., Taylor, F.W. & Phillips, D.A., 2003. New insights on the tectonics along the New Hebrides subduction zone based on GPS results, *J. geophys. Res.*, **108**(B6), doi:10.1029/2001JB000644.
- Chen, M.-C., Frohlich, C., Taylor, F.W., Burr, G. & Van Ufford, A.Q., 2011. Arc segmentation and seismicity in the Solomon Islands arc, SW Pacific, *Tectonophysics*, **507**(1–4), 47–69.
- Chen, Y.-S., Weatherill, G., Pagani, M. & Cotton, F., 2018. A transparent and data-driven global tectonic regionalization model for seismic hazard assessment, *Geophys. J. Int.*, **213**(2), 1263–1280.
- Chiou, B.S.-J. & Youngs, R.R., 2008. An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthq. Spectra*, **24**, 173–215.
- Christophersen, A., et al., 2015. Development of the Global Earthquake Model's neotectonic fault database, *Nat. Hazards*, **79**(1), 111–135.
- Cornell, C.A., 1968. Engineering seismic risk analysis, *Bull. seism. Soc. Am.*, **58**, 1583–1606.
- Dziewonski, A., Chou, T.-A. & Woodhouse, J., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.*, **86**(B4), 2825–2852.
- Ekström, G., Nettles, M. & Dziewoński, A., 2012. The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.*, **200**, 1–9.
- Farr, T.G. et al., 2007. The shuttle radar topography mission, *Rev. Geophys.*, **45**(2), 1–33.
- Field, E.H., Jordan, T.H. & Cornell, C.A., 2003. OpenSHA—a developing community-modeling environment for seismic hazard analysis, *Seismol. Res. Lett.*, **74**, 406–419.
- Frankel, A., 1995. Mapping seismic hazard in the central and eastern United States, *Seismol. Res. Lett.*, **66**(4), 8–21.
- Gardner, J.K. & Knopoff, L., 1974. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, *Bull. seism. Soc. Am.*, **64**(5), 1363–1367.
- GEBCO, 2008. *The GEBCO-08 Grid*, International Hydrographic Organization and the Intergovernmental Oceanographic Commission of UNESCO, p. 19.
- Ghasemi, H., McKee, C., Leonard, M., Cummins, P., Moihoi, M., Spiro, S., Taranu, F. & Buri, E., 2016. Probabilistic seismic hazard map of Papua New Guinea, *Nat. Hazards*, **81**(2), 1003–1025.
- Hayes, G.P., Wald, D.J. & Johnson, R.L., 2012. Slab1.0: a three-dimensional model of global subduction zone geometries, *J. geophys. Res.*, **117**(B1), doi:10.1029/2011JB008524.
- Heuret, A., Lallemand, S., Funiciello, F., Piromallo, C. & Faccenna, C., 2011. Physical characteristics of subduction interface type seismogenic zones revisited, *Geochem. Geophys. Geosyst.*, **12**(1), Q01004, doi:10.1029/2010GC003230.
- Jones, T., Australian Geological Survey Organisation, 1998. *Probabilistic Earthquake Hazard Assessment for Fiji*, Australian Geological Survey Organisation. OCLC: 40620647.
- Kaverina, A., Lander, A. & Prozorov, A., 1996. Global creepex distribution and its relation to earthquake-source geometry and tectonic origin, *Geophys. J. Int.*, **125**(1), 249–265.
- Kuo, Y.-T., et al., 2016. Characteristics on fault coupling along the Solomon megathrust based on GPS observations from 2011 to 2014, *Geophys. Res. Lett.*, **43**(16), 8519–8526.
- Laske, G., Masters, G., Ma, Z. & Pasquano, M., 2013. Update on Crust1.0—a 1-degree global model of Earth's crust, in *EGU General Assembly*, EGU2013-2658, Vienna, Austria.
- Leonard, M., 2010. Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release, *Bull. seism. Soc. Am.*, **100**(5A), 1971–1988.
- McCaffrey, R., 2008. Global frequency of magnitude 9 earthquakes, *Geol.*, **36**(3), 263–266.
- NGDC/WDS, 2020. Significant earthquake database, Technical report, National Geophysical Data Center, NOAA, doi:10.7289/V5TD9V7K (last accessed: 2020 May 5).
- Pagani, M. & Marcellini, A., 2007. Seismic-hazard disaggregation: a fully probabilistic methodology, *Bull. seism. Soc. Am.*, **97**(5), 1688–1701.
- Pagani, M., et al., 2014. OpenQuake Engine: an open hazard (and risk) software for the Global Earthquake Model, *Seismol. Res. Lett.*, **85**(3), 692–702.
- Pagani, M., et al., 2020a. The 2018 version of the Global Earthquake Model: hazard component, *Earthq. Spectra*, **36**, 226–251.
- Pagani, M., Johnson, K. & Palaez, J.G., 2020b. Modelling subduction sources for probabilistic seismic hazard analysis, *Geol. Soc. Lond. Spec. Publ.*, **501**, doi:10.1144/SP501-2019-120.
- Pasquano, M.E., Masters, T.G., Laske, G. & Ma, Z., 2014. Litho1.0: an updated crust and lithospheric model of the earth, *J. geophys. Res.*, **119**(3), 2153–2173.
- Peñarubia, H.C., et al., 2020. Probabilistic seismic hazard analysis model for the Philippines, *Earthq. Spectra*, **36**, 44–68.
- Petersen, M.D., Harmsen, S.C., Rukstales, K.S., Mueller, C.S., McNamara, D.E., Luco, N. & Walling, M., 2012. *Seismic Hazard of American Samoa and Neighboring South Pacific Islands—methods, Data, Parameters, and Results*, US Department of the Interior, US Geological Survey.
- Rahiman, T.I. & Pettinga, J.R., 2008. Analysis of lineaments and their relationship to Neogene fracturing, se Viti Levu, Fiji, *Bull. geol. Soc. Am.*, **120**(11–12), 1544–1555.
- Richards, S., Holm, R. & Barber, G., 2011. When slabs collide: a tectonic assessment of deep earthquakes in the Tonga-Vanuatu region, *Geology*, **39**(8), 787–790.
- Rong, Y., Mahdyiar, M., Shen-Tu, B., Shabestari, K. & Guin, J., 2010. Probabilistic seismic hazard assessment for south Pacific Islands, in *Proceedings of 9th US National Conference on Earthquake Engineering (9NCEE)*, Toronto.
- Rong, Y., Xu, X., Cheng, J., Chen, G., Magistrale, H. & Shen, Z.-K., 2020. A probabilistic seismic hazard model for mainland China, *Earthq. Spectra*, **36**, 181–209.
- Schorlemmer, D., Gerstenberger, M., Wiemer, S., Jackson, D. & Rhoades, D., 2007. Earthquake likelihood model testing, *Seismol. Res. Lett.*, **78**(1), 17–29.
- Schorlemmer, D., et al., 2010. First results of the regional earthquake likelihood models experiment, in *Seismogenesis and Earthquake Forecasting: The Frank Evison Volume II*, pp. 5–22, Springer.
- 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, *Seismol. Res. Lett.*, **81**, 941–950.
- Styron, R. & Pagani, M., 2020. The GEM Global Active Faults Database (GAF-DB), *Earthq. Spectra*, **36**(1\_suppl), 160–180, doi.org/10.1177/8755293020944182, in press.
- Suckale, J. & Grünthal, G., 2009. Probabilistic seismic hazard model for Vanuatu, *Bull. seism. Soc. Am.*, **99**(4), 2108–2126.
- Taylor, F.W., Edwards, R.L., Wasserburg, G. & Frohlich, C., 1990. Seismic recurrence intervals and timing of aseismic subduction inferred from emerged corals and reefs of the central Vanuatu (New Hebrides) frontal arc, *J. geophys. Res.*, **95**(B1), 393–408.
- Thingbaijam K, K.S., Mai, P. & Goda, K., 2017. New empirical earthquake source-scaling laws, *Bull. seism. Soc. Am.*, **107**(5), 2225–2946.
- Thirumalai, K., et al., 2015. Variable Holocene deformation above a shallow subduction zone extremely close to the trench, *Nat. Commun.*, **6**(1), 1–6.

- Turcotte, D. & Schubert, G., 2014. *Geodynamics*, Cambridge Univ. Press.
- Uhrhammer, R., 1986. Characteristics of Northern and Central California seismicity, *Earthq. Notes*, **57**(1), 21.
- van Stiphout, T., Zhuang, J. & Marsan, D., 2010. Theme V—basic techniques/models for analysing seismicity, Technical report, Community Online Resource for Statistical Seismicity Analysis.
- Wallace, L.M., McCaffrey, R., Beavan, J. & Ellis, S., 2005. Rapid microplate rotations and backarc rifting at the transition between collision and subduction, *Geol.*, **33**(11), 857–860.
- Wallace, L.M., Fagereng, Å. & Ellis, S., 2012. Upper plate tectonic stress state may influence interseismic coupling on subduction megathrusts, *Geol.*, **40**(10), 895–898.
- Weatherill, G.A., Pagani, M. & Garcia, J., 2016. Exploring earthquake databases for the creation of magnitude-homogeneous catalogues: tools for application on a regional and global scale, *Geophys. J. Int.*, **206**(3), 1652–1676.
- Weichert, D.H., 1980. Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, *Bull. seism. Soc. Am.*, **70**(4), 1337–1346.
- Wessel, P., Smith, W.H., Scharroo, R., Luis, J. & Wobbe, F., 2013. Generic mapping tools: improved version released, *EOS, Trans. Am. geophys. Un.*, **94**(45), 409–410.
- Youngs, R., Chiou, S., Silva, W. & Humphrey, J.R., 1997. Strong ground motion attenuation relationships for subduction zone earthquakes, *Seismol. Res. Lett.*, **68**, 58–73.
- Youngs, R.R. & Coppersmith, K.J., 1985. Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates, *Bull. seism. Soc. Am.*, **75**, 939–964.
- Zechar, J.D., Gerstenberger, M.C. & Rhoades, D.A., 2010. Likelihood-based tests for evaluating space–rate–magnitude earthquake forecasts, *Bull. seism. Soc. Am.*, **100**(3), 1184–1195.
- Zhao, J.X., et al., 2006. Attenuation relations of strong ground motion in Japan using site classification based on predominant period, *Bull. seism. Soc. Am.*, **96**, 898–913.

## SUPPORTING INFORMATION

Supplementary data are available at [GJI](#) online.

### **supplement.zip**

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.