# A global probabilistic tsunami hazard assessment from earthquake sources

GARETH DAVIES*[1], JONATHAN GRIFFIN[1], FINN LØVHOLT[2], SYLFEST GLIMSDAL[2], CARL HARBITZ[2], HONG KIE THIO[3], STEFANO LORITO[4], ROBERTO BASILI[4], JACOPO SELVA[4], ERIC GEIST[5] & MARIA ANA BAPTISTA[6,7]

[1]*Community Safety Branch, Geoscience Australia*

[2]*Norwegian Geotechnical Institute*

[3]*AECOM*

[4]*Istituto Nazionale di Geofisica e Vulcanologia*

[5]*Pacific Coast and Marine Science Centre, United States Geological Survey*

[6]*Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa*

[7]*Instituto Português do Mar e da Atmosfera*

**Correspondence: gareth.davies@ga.gov.au*

**Abstract:** Large tsunamis occur infrequently but have the capacity to cause enormous numbers of casualties, damage to the built environment and critical infrastructure, and economic losses. A sound understanding of tsunami hazard is required to underpin management of these risks, and while tsunami hazard assessments are typically conducted at regional or local scales, globally consistent assessments are required to support international disaster risk reduction efforts, and can serve as a reference for local and regional studies. This study presents a global-scale probabilistic tsunami hazard assessment (PTHA), extending previous global-scale assessments based largely on scenario analysis. Only earthquake sources are considered, as they represent about 80% of the recorded damaging tsunami events. Globally extensive estimates of tsunami run-up height are derived at various exceedance rates, and the associated uncertainties are quantified. Epistemic uncertainties in the exceedance rates of large earthquakes often lead to large uncertainties in tsunami run-up. Deviations between modelled tsunami run-up and event observations are quantified, and found to be larger than suggested in previous studies. Accounting for these deviations in PTHA is important, as it leads to a pronounced increase in predicted tsunami run-up for a given exceedance rate.

Large tsunami disasters are low-frequency events and as such may have no regional historical precedent. At any particular coastal location, relatively small, non-destructive tsunamis tend to greatly outnumber large, destructive events, while inundation and damage tend to grow rapidly with increasing tsunami flow depth (Geist & Parsons 2006; Gonzalez *et al.* 2009; Valencia *et al.* 2011). Therefore, rare tsunami events are by far the greatest contributor to tsunami risk: globally, the historical record suggests that approximately 3% of tsunamis have caused approximately 97% of tsunami fatalities, with approximately 80% of these events generated by earthquakes (Løvholt *et al.* 2014*a*; NGDC 2015). Given the absence of well-known historical precedents and our incomplete understanding of the frequency of large subduction zone earthquakes, the two most destructive tsunami in recent decades (the 2004 Andaman–Sumatra tsunami and the 2011 Tohoku tsunami) came as a surprise to many scientists, and comparable scenarios were not considered in disaster management planning (Satake & Atwater 2007; McCaffrey 2008, 2009; Synolakis 2011; Kagan & Jackson 2013; Satake 2014; Løvholt *et al.* 2014*b*; Lorito *et al.* 2015; Synolakis & Kanoglu 2015).

Management of future tsunami risk requires a sound understanding of the hazard and, in general, the historical record is too limited for a purely empirical hazard assessment (Geist & Parsons 2006; Kagan & Jackson 2013; Løvholt *et al.* 2014*a*). Instead, tsunami hazard must be assessed using a mixture of modelling and data. Two main approaches exist: the scenario approach; and

*From*: Scourse, E. M., Chapman, N. A., Tappin, D. R. & Wallis, S. R. (eds) 2018. *Tsunamis: Geology, Hazards and Risks*. Geological Society, London, Special Publications, **456**, 219–244.
First published online February 23, 2017, updated March 2, 2017, https://doi.org/10.1144/SP456.5

probabilistic tsunami hazard assessment (PTHA). In the scenario approach, a number of credible (and often presumed worst-case) tsunami scenarios are modelled to produce maps depicting a tsunami hazard metric, such as the wave run-up, inundation footprint or offshore wave height (Burbidge & Cummins 2007; Lorito *et al.* 2008; Tiberti *et al.* 2009; Harbitz *et al.* 2012; Løvholt *et al.* 2012*a*, *b*). This approach is advantageous because it is relatively simple and easily communicated. However, the results can be highly sensitive to the particular scenarios selected, and become undesirably subjective when the credible scenario parameters are highly uncertain, which is common, for example, for the moment magnitude of low return period earthquakes (Berryman *et al.* 2015). In comparison, PTHA is relatively complex and computationally intensive, as it involves the integration of tsunami hazard from a model of all possible events to determine the exceedance rate of some tsunami hazard metric (mean number of events per year above a threshold level) (Geist & Parsons 2006; Parsons & Geist 2009). Like the scenario approach, PTHA is dependent on subjective decisions. For example, it may be necessary to specify the maximum earthquake magnitude possible on a given source zone, and doing this incorrectly could result in strong over- or underestimations of the hazard (Annaka *et al.* 2007; Burbidge *et al.* 2008; Satake 2014). However, unlike the scenario approach, there are established methodologies for accounting for known uncertainties associated with these choices in PTHA (Annaka *et al.* 2007; Burbidge *et al.* 2008; Parsons & Geist 2009; Sørensen *et al.* 2012; Selva *et al.* 2016).

With few exceptions (Løvholt *et al.* 2012*a*, 2014*a*), most tsunami hazard assessments to date have been conducted at the regional, national or city level (Annaka *et al.* 2007; Burbidge *et al.* 2008; Gonzalez *et al.* 2009; Parsons & Geist 2009; Sørensen *et al.* 2012; Power 2013; Horspool *et al.* 2014; ten Brink *et al.* 2014). Analysis at these scales is better suited to supporting national-scale disaster risk management, given the typical limitations on computational and human resources for a single study, because a reduced spatial scale allows for: (1) greater detail in the modelling of tsunami sources and propagation; (2) linking local inundation maps to an exceedance rate; (3) a more in-depth assessment of predictions using historical data and palaeotsunami evidence where available; and (4) a more in-depth assessment of the associated uncertainties. However, spatial coverage limitations and methodological differences make it difficult to assess tsunami hazard globally by combining regional-scale hazard assessments, while the global impacts of tsunamis can extend far beyond the inundation zone via their economic and political consequences (UN-ISDR 2015). At the global level, the Hyogo Framework for Action was established by the United Nations to reduce disaster losses, and this has been followed by the Sendai Framework for Disaster Risk Reduction. The implementation goals of these agreements rely on a sound understanding of global disaster risk due to different natural hazards, and have motivated several previous global-scale tsunami hazard and risk assessments (Løvholt *et al.* 2012*a*, 2014*a*, *b*). Given the computational challenges of modelling tsunamis globally, the former studies were based largely on scenario approaches for tsunami hazard assessment, although a partial probabilistic treatment was employed by Løvholt *et al.* (2014*a*) in the Indian Ocean and south Pacific.

The current study extends these works with a full global probabilistic tsunami hazard assessment. Compared with previous work (Løvholt *et al.* 2012*a*, 2014*a*), the PTHA approach allows modelling events both smaller and larger than the scenarios treated previously (nominal 1/500 year exceedance rate events), and permits a more complete assessment of the associated uncertainties. The latter turn out to be large, driven in part by considerable epistemic uncertainties in the exceedance rates of high-magnitude tsunamigenic earthquakes (Berryman *et al.* 2015), and limitations in our capacity to model tsunami run-up at the global scale. Combined with appropriate exposure information, the results of the current study could also be used to support broad-scale assessments of tsunami risks (Løvholt *et al.* 2014*a*, 2015, 2016) or multi-hazard risk assessments (UN-ISDR 2015), although that step is not undertaken herein.

Given its global scale, the tsunami hazard results herein are expected to be broad-brush, and the quantified uncertainties are rather large. However, in the longer term, they could be extended with detailed analyses such as currently possible for assessments over smaller spatial scales (Gonzalez *et al.* 2009; Lorito *et al.* 2016). This would require considerable computational and human resources, necessitating a coordinated global effort which is beyond the scope of the present work.

## Methodology

### *Overview of the applied PTHA methodology*

We compute the exceedance rate (events/year) of tsunami run-up height at a globally distributed set of hazard points offshore of the coast (*c.* 100 m depth), using methods that broadly follow previous approaches (Geist & Parsons 2006; Annaka *et al.* 2007; Thio *et al.* 2007; Burbidge *et al.* 2008; Horspool *et al.* 2014). We further estimate the extra amplification that these waves undergo as they

propagate onshore, following the amplification-factor approach of Løvholt *et al.* (2012*a*). The key steps are:

- Define tsunami sources to be included in the analysis. In this study only oceanic earthquakes are considered.
- Discretize each earthquake source zone into a grid of 'unit-sources'.
- For each source zone create a synthetic earthquake catalogue, with all synthetic earthquakes consisting of uniform-slip combinations of the unit-sources. Assign a mean annual rate to each synthetic event using a Gutenberg–Richter-type model for the source zone, with uncertainties in parameter values accounted for using a logic tree.
- For each unit-source, compute the sea-surface deformation produced by 1 m of slip.
- Use the linear shallow-water equations to model the wave time series produced by the 1 m unit-source ruptures at a set of nearshore coastal points (hazard points).
- Extract the maximum tsunami wave height and period from wave time series at all hazard points for each earthquake in the catalogue. Because the sea-surface deformation and tsunami propagation models are linear, the hazard point time series for a single earthquake is efficiently computed by summing the tsunami wave time series for all included unit-sources and multiplying by the earthquake slip.
- Compute the tsunami run-up from the maximum tsunami wave height and period at the offshore hazard points.
- Combine the modelled run-up and earthquake rates for all events to compute the exceedance rate of tsunami run-up near every hazard point.

Details of the above steps are presented below.

### *Source zone definition*

As the majority of large tsunamigenic earthquakes occur on subduction zones, most earthquake sources used in this study are of this type (Fig. 1). However, we included several non-subduction zone sources in the Mediterranean, eastern Indonesia and the NE Atlantic where previous assessments indicated they are regionally significant drivers of the hazard (Sørensen *et al.* 2012; Horspool *et al.* 2014; Omira *et al.* 2015). In particular, in the NE Atlantic we include the strike-slip Gloria source, and in the Gulf of Cadiz we use a schematic source to represent the large number of thrust sources which are too small to resolve in our model (Matias *et al.* 2013; Omira *et al.* 2015).

Three-dimensional source zone fault-plane geometries are defined using the Slab 1.0 subduction interface model where available (Hayes *et al.* 2012), geometries from the SHARE project in the Mediterranean (Basili *et al.* 2013*a*, *b*), and constant dip geometries elsewhere based on Bird (2003), Burbidge *et al.* (2008), Horspool *et al.* (2014) and Berryman *et al.* (2015). The maximum seismogenic source depth is taken from the upper maximum depth suggested by Berryman *et al.* (2015). The seismogenic zone is assumed to extend to the trench (i.e. minimum depth of zero), except in the Hellenic Arc where the trench is significantly further south than the expected updip limit of the active subduction interface. Rupture to the trench is preferable for simulating tsunamis since if the minimum depth is non-zero and on the order of a few kilometres, the vertical coseismic displacement field predicted by the Okada (1985) model contains an artificial spike near the shallow rupture edge which can have a strong impact on the associated tsunami (Geist & Dmowska 1999; Goda 2015).

### *Source zone discretization with unit-sources*

Each source is discretized in a logically rectangular grid of unit-sources with dimensions approximately 100 km along-strike and 50 km downdip (Fig. 2). The use of a logically rectangular layout (with a fixed number of unit sources along-strike and a fixed number downdip) simplifies the construction of synthetic earthquake events (see the subsection on 'Synthetic earthquake event geometries' later in this section), although it means that the unit source length and width must vary to match the typically non-uniform source zone geometry (Fig. 2). Within each unit-source, a finer grid of subunit-source points records the depth and dip. These are used to retain detailed information on the non-planar source geometry when computing the unit-source seafloor deformation. The subunit-source point spacing was set to 6 km, except for unit-sources along the trench where a finer spacing (1 km) was required to ensure accurate numerical integration of the seafloor deformation. Rupture is treated as pure thrust, except on the strike-slip Gloria source zone.

### *Sea-surface deformation*

For every unit-source we compute the seafloor deformation from 1 m of earthquake slip by summing the deformation computed at each subunit-source point, with the individual deformation computations following Okada (1985) (Fig. 2), with both Lamé constants equal to $3 \times 10^{10}$ Pa. We assume that each subunit-source point represents a small rectangular region with dip and depth determined by the source depth contours, and length and width determined by the subunit-source grid spacing. Near the unit-source boundaries, these

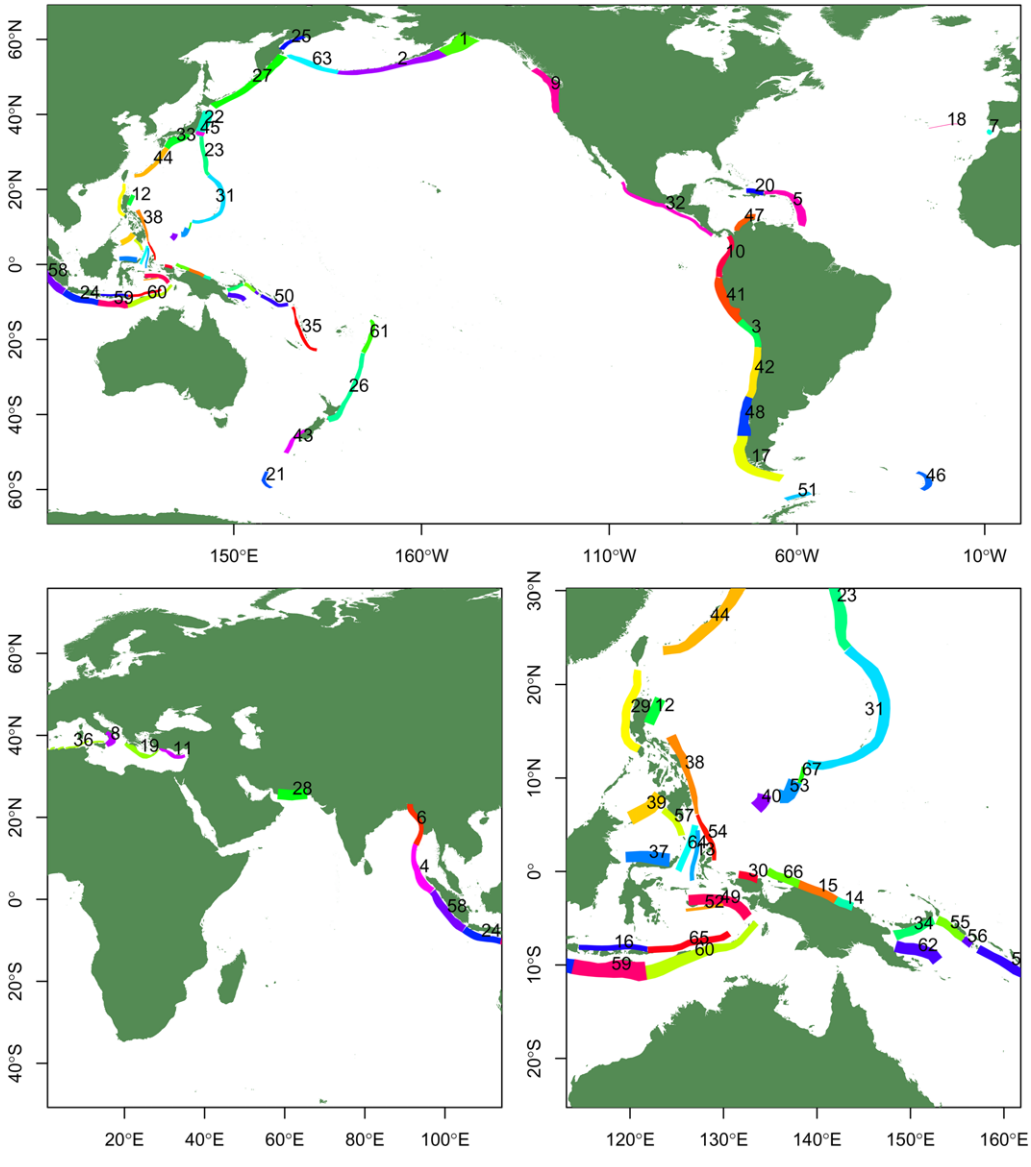


**Fig. 1.** Source zones used in the study: 1, Alaska; 2, Aleutians; 3, Altiplano; 4, Andaman; 5, Antilles; 6, Arakan; 7, Cadiz; 8, Calabrian; 9, Cascadia; 10, Columbia; 11, Cyprus; 12, Eastluzontrough; 13, Eastmolucca; 14, EastNewGuinea1; 15, EastNewGuinea2; 16, Flores; 17, Fuego; 18, Gloria; 19, Hellenic; 20, Hispaniola; 21, Hjort; 22, Honshu; 23, Izu; 24, Java; 25, Kamchatka; 26, Kermadec; 27, Kurils; 28, Makran; 29, Manila; 30, Manokwari; 31, Mariana; 32, Midamerica; 33, Nankai; 34, Newbritain; 35, Newhebrides; 36, North_african; 37, Northsulawesi; 38, Nphillip; 39, Nwsulu; 40, Palau; 41, Peru; 42, Puna; 43, Puysegur; 44, Ryukyu; 45, Sagami; 46, Sandwich; 47, Scaribbean; 48, Schile; 49, Seram; 50, Sesolomon; 51, Shetland; 52, Southseram; 53, South_yap; 54, Sphillip; 55, SouthSolomon; 56, SouthSolomonwood; 57, Sulu; 58, Sumatra; 59, Sumba; 60, Timor; 61, Tonga; 62, Trobriand; 63, Waleutians; 64, Westmolucca; 65, Wetar; 66, WestNewGuinea; 67, Yap.

rectangular subunit-source regions are reduced in size (if they extend outside the original unit-source boundary) to ensure that the subunit-source area sum equals the total unit-source interface area. The unit-source sea-surface deformation is computed by smoothing the seafloor deformation using a filter based on full potential linear wave theory (Kajiura 1963; Glimsdal *et al.* 2013). Compared

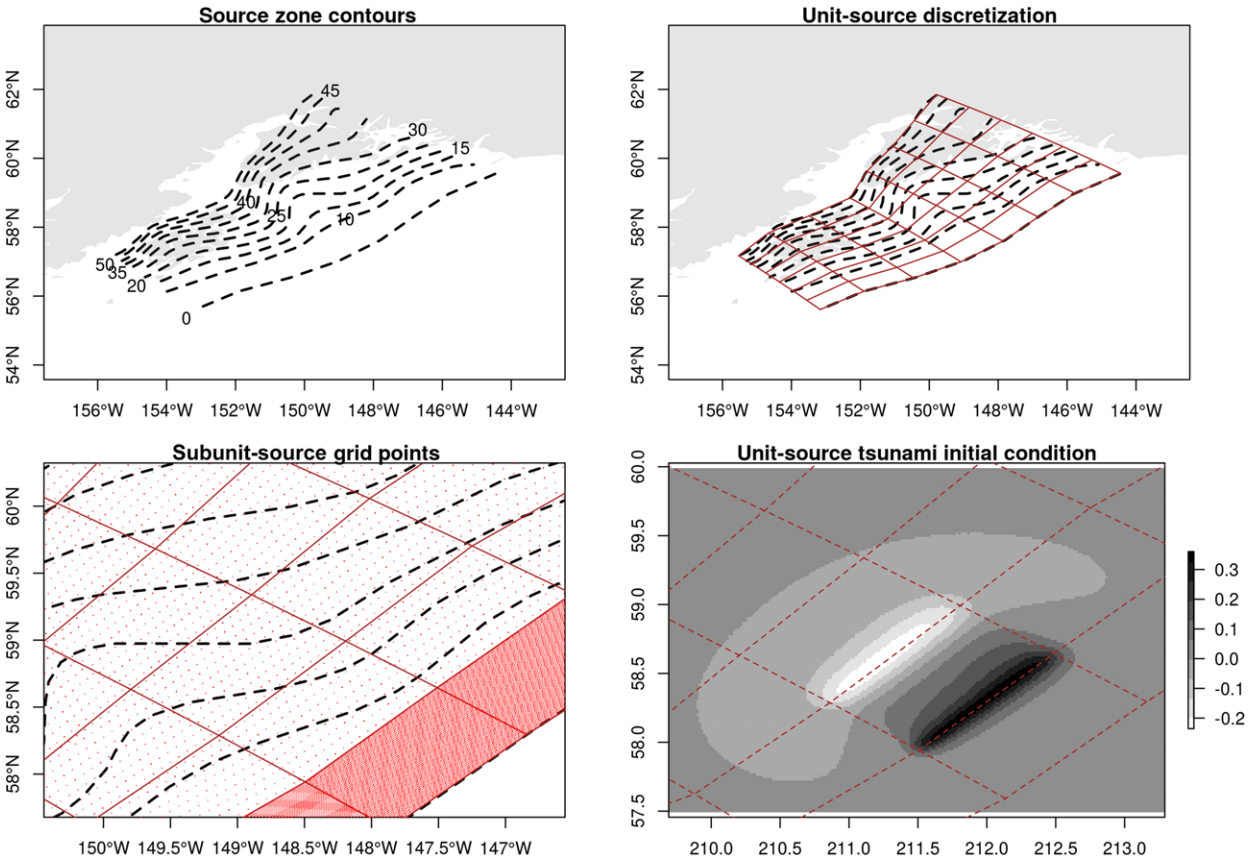


**Fig. 2.** Unit source definition. (Top left) Subduction interface contours for the Alaska source zone. Labels give the depth in kilometres. (Top right) The logically rectangular unit-source grid, overlain on the depth contours. (Bottom left) Zoom depicting the subunit-source grid points inside each unit-source. A finer spacing is used for unit-sources along the trench, to allow accurate numerical integration when computing the vertical coseismic displacement field. (Bottom right) The tsunami initial condition from one unit-source, with the scale in metres.

with directly applying the seafloor deformation to the sea surface, the filter leads to better agreement with the sea-surface deformation computed using primitive models such as the full potential equations, and Navier–Stokes for the generation and initial propagation (Saito & Furumura 2009; Saito 2013).

### *Synthetic earthquake catalogue*

On each source zone we generate a catalogue of synthetic earthquake events, all of which consist of uniform slip on a logically rectangular subset of the unit-sources (see the following subsection on 'Synthetic earthquake event geometries'). A mean annual rate is assigned to each synthetic earthquake event, such that when integrated over a source zone the exceedance rates (events/year) follow a Gutenberg–Richter-type distribution (see the subsection on 'Synthetic earthquake event rates' later in this section). The latter is forced to satisfy a seismic moment conservation principle (Bird & Kagan 2004). Parameters controlling the earthquake rates are assigned to each source zone using a range of literature-derived values (see 'Source zone parameters' later in this section), with parameter uncertainties associated with the maximum magnitude, Gutenberg–Richter $b$ value and degree of seismic coupling accounted for using a logic tree (see the subsection on 'Logic-tree weights and epistemic uncertainty quantification' later in this section). Details are provided below.

*Synthetic earthquake event geometries.* The synthetic earthquake events have moment magnitudes $m_w$ in the range:

$$m_w \in \{m_{w,min}, (m_{w,min} + 0.1), (m_{w,min} + 0.2), \ldots, (m_{w,max} - 0.1), m_{w,max}\} \quad (1)$$

where the source zone specific computation of the minimum magnitude $m_{w,min}$ and the maximum magnitude $m_{w,max}$ is described later (see the following subsection on 'Source zone parameters'). For every $m_w$ we use the scaling relationships of Strasser *et al.* (2010) to compute the desired area $A$, length $L$ and width $W$ for the synthetic earthquake events. In general, the synthetic event dimensions cannot exactly conform to $A$, $L$ and $W$ because their size is constrained by the unit-source dimensions (*c.* $100 \times 50\,km^2$: Fig. 2). However, as detailed below, for each $m_w$ we select the number of subfaults along-strike ($n_s$) and downdip ($n_d$) so that the dimensions are close to the desired values. All possible $n_s \times n_d$ logically rectangular subsets of the unit-sources are then included in the synthetic catalogue: that is, the earthquake floats through every possible position on the source zone. For the $i$-th event with magnitude $m_{w,i}$ and area $A_i$ ($m^2$), the slip $S_i$ (m) is computed as (Hanks & Kanamori 1979; Bird & Kagan 2004):

$$M_{0,i} = 10^{1.5m_{w,i}+9.05} \quad (2)$$

$$S_i = \frac{M_{0,i}}{(\mu A_i)} \quad (3)$$

where the shear modulus $\mu$ is taken as $3 \times 10^{10}$ Pa and $M_{0,i}$ is the seismic moment (N m). Note that $A_i$ will vary somewhat for events with the same magnitude because of variations in the unit-source dimensions (Fig. 2).

The detailed methods used to select the number of unit-sources along-strike ($n_s$) and downdip ($n_d$) for a given magnitude are provided here. From $m_w$ we compute $A$, $L$ and $W$ (Strasser *et al.* 2010), and check if $A$ exceeds the total source zone area, in which case the synthetic event includes all subfaults. Otherwise define $n'_s$ as $L$ divided by the average unit-source length $\bar{l}_s$ (*c.* 100 km), without rounding. If $n'_s$ exceeds the available number of subfaults along-strike, then $n_s$ is assumed to cover the full source zone length and $n_d$ is computed to best fit $A$. Otherwise define $n'_d$ as $W$ divided by the average unit-source width $\bar{l}_d$ (*c.* 50 km) without rounding. If $n'_d$ exceeds the available number of subfaults downdip, then $n_d$ is assumed to cover the full source zone width, and $n_s$ computed to best fit $A$. Otherwise, we have the typical case where both $n'_s$ and $n'_d$ can be accommodated within the source zone. Four candidate ($n_s$, $n_d$) pairs are then investigated: two pairs come from assigning $n_s$ to $n'_s$ rounded up and down (with a lower bound of 1 and upper bound from the source zone geometry), and setting $n_d$ to best match $A$; the other two pairs are produced analogously by rounding $n'_d$ and finding $n_s$ to best match $A$. From the four candidates ($n_s$, $n_d$) pairs we choose one which best matches the desired event aspect ratio: that is, which minimizes:

$$\left[\log_{10}\left(\frac{(n_s\bar{l}_s)}{(n_d\bar{l}_d)}\right) - \log_{10}\left(\frac{L}{W}\right)\right]^2 \quad (4)$$

This measures the aspect ratio error in relative terms, so that, for example, a ($n_s$, $n_d$) pair producing only 50% of the desired aspect ratio is weighted equally as one producing 200% of the desired aspect ratio.

*Source zone parameters.* To generate mean annual rates for events in the synthetic earthquake catalogue we require source zone specific parameter

values for the minimum magnitude $m_{w,min}$, maximum magnitude $m_{w,max}$, the Gutenberg–Richter $b$ value and the seismically coupled plate motion rate $c\dot{s}$ (the latter notation reflects that $c\dot{s}$ is the product of the seismic coupling rate $c$ with the plate motion rate $\dot{s}$). Below, we describe how those parameters were determined, and the approach used to quantify uncertainties. The methods used to derive the associated synthetic earthquake event rates are described in following subsections on 'Synthetic earthquake event rates' and 'Logic-tree weights and epistemic uncertainty quantification'.

The minimum magnitude $m_{w,min}$ is set to 7.5 everywhere, consistent with the area of an earthquake containing one $100 \times 50\ km^2$ unit-source (Strasser *et al.* 2010). This implies that we ignore all tsunamis generated by smaller earthquakes. While smaller earthquake can generate locally significant tsunamis (especially when combined with landslides – e.g. the 1998 Papua New Guinea tsunami: Tappin *et al.* 2008; Heidarzadeh & Satake 2015), a further reduction of $m_{w,min}$ would require using a smaller unit-source size which would increase the already heavy computational demands of the study.

Parameters $m_{w,max}$, $b$ and $c\dot{s}$ are subject to considerable epistemic uncertainty which is accounted for using a logic tree (Annaka *et al.* 2007; Thio *et al.* 2007; Horspool *et al.* 2014). Each parameter is assigned three different values (Table 1). The parameter $b$ controls the relative frequencies of small and large earthquakes, and is assigned the values 0.7, 0.95 and 1.2 on all source zones, similar to source zone specific uncertainty ranges suggested by Berryman *et al.* (2015). This range covers a number of $b$ values suggested from globally integrated analyses (Bird & Kagan 2004; Schorlemmer *et al.* 2005; Kagan 2010). The seismically coupled trench-normal plate motion rate $c\dot{s}$ is computed using three separate seismic coupling values $c$ which are based on Berryman *et al.* (2015). Where our source zones are not included in the latter study, default values of $0.5 \pm 0.2$ are used. A single average plate motion rate $\dot{s}$ is assigned to each source zone, where the trench-normal component of plate motion is used to reflect that we only model thrust earthquakes (Burbidge *et al.* 2008), except on the strike-slip Gloria source zone (Omira *et al.* 2015). Convergence angles are based on Bird (2003), and absolute plate motion rates are primarily based on Bird (2003) and Berryman *et al.* (2015); although we refer to Cummins (2007) and Socquet *et al.* (2006) for Arakan, and Horspool *et al.* (2014) for South-Seram and Timor. In the Gulf of Cadiz, we use a single schematic source zone but artificially increase $c\dot{s}$ by a factor of 4 to reflect the moment rate on numerous sources in this complex region (Matias *et al.* 2013).

For $m_{w,max}$ an overall upper limit of 9.6 is applied following Berryman *et al.* (2015). Within that constraint, the largest $m_{w,max}$ value is computed with an empirical relationship for predicting $m_w$ from rupture-area at the +1 SD limit (Strasser *et al.* 2010) assuming that the largest possible earthquake ruptures the entire source zone. This gives conservative $m_{w,max}$ values which are usually similar to those derived with the conservative $m_{w,max}$ approach of McCaffrey (2008) and Berryman *et al.* (2015), while still being consistent with our use of the Strasser *et al.* (2010) magnitude–area scaling relationships elsewhere. A second non-conservative estimate of $m_{w,max}$ is employed following Berryman *et al.* (2015), which is based on the largest earthquake thought to have occurred on each source zone. Where our source zones are not included in Berryman *et al.* (2015), we derive values from Burbidge *et al.* (2008), Storchak *et al.* (2012), Matias *et al.* (2013), Horspool *et al.* (2014) and Omira *et al.* (2015). Our third intermediate estimate of $m_{w,max}$ is derived by averaging the conservative and non-conservative estimates, following Berryman *et al.* (2015). For computational consistency with $m_{w,min} = 7.5$, we finally apply a lower limit of 7.6 to the maximum magnitudes and round all $m_{w,max}$ values to one decimal place.

*Synthetic earthquake event rates.* We assume that earthquake timings on each source zone follow a Poisson distribution: that is, the event rate is stationary and inter-event times are exponentially distributed. The mean annual rate assigned to each event in the synthetic catalogue is derived from source zone specific values for $m_{w,min}$, $m_{w,max}$, $b$ and $c\dot{s}$. Because of uncertainty in these parameters, it is ultimately necessary to integrate over multiple parameter combinations, combining their corresponding rates as explained in the following subsection on 'Logic-tree weights and epistemic uncertainty quantification'. However, this subsection presents the rate computation assuming fixed source zone parameters, which is applied to each parameter combination separately prior to their integration.

Two different parametric forms for the source zone integrated earthquake magnitude–frequency relationship are considered (giving the mean annual number of events exceeding any magnitude on the source zone): first, a Gutenberg–Richter distribution with exceedance rate function truncated between $m_{w,min}$ and $m_{w,max}$:

$$\begin{aligned} GR(m_w) &= 10^{a-bm_w} && \text{for } m_{w,min} \le m_w \le m_{w,max} \\ &= 0 && \text{for } m_w > m_{w,max} \\ &= 10^{a-bm_{w,min}} && \text{for } m_w < m_{w,min} \end{aligned} \quad (5)$$

**Table 1.** *Values for the maximum moment magnitude* $m_{w,max}$ *and trench-normal coupled slip rate* $\dot{cs}$ *(mm* $a^{-1}$*) used on all source zones (trench-parallel slip is applied on the strike-slip Gloria source zone)*

| Source name | $m^1_{w,max}$ | $m^2_{w,max}$ | $m^3_{w,max}$ | $\dot{cs}^1$ | $\dot{cs}^2$ | $\dot{cs}^3$ |
|---|---|---|---|---|---|---|
| Alaska | 9.30 | 9.30 | 9.40 | 40.00 | 45.00 | 50.00 |
| Aleutians | 9.10 | 9.30 | 9.60 | 19.00 | 29.00 | 43.00 |
| Altiplano | 8.80 | 9.10 | 9.40 | 40.00 | 46.00 | 52.00 |
| Andaman | 9.00 | 9.30 | 9.60 | 15.00 | 18.00 | 20.00 |
| Antilles | 8.00 | 8.80 | 9.60 | 3.30 | 5.50 | 7.70 |
| Arakan | 7.60 | 8.40 | 9.40 | 4.90 | 8.10 | 11.00 |
| Cadiz | 8.20 | 8.40 | 8.60 | 4.80 | 7.90 | 11.00 |
| Calabrian | 7.60 | 8.10 | 9.00 | 0.53 | 0.88 | 1.20 |
| Cascadia | 8.80 | 9.20 | 9.50 | 26.00 | 30.00 | 34.00 |
| Columbia | 8.80 | 9.20 | 9.50 | 34.00 | 39.00 | 43.00 |
| Cyprus | 7.70 | 8.30 | 9.00 | 2.00 | 3.40 | 4.80 |
| Eastluzontrough | 7.60 | 8.10 | 8.80 | 3.00 | 5.00 | 6.90 |
| Eastmollucca | 8.50 | 8.60 | 8.70 | 25.00 | 42.00 | 59.00 |
| Eenewguinea | 7.60 | 7.90 | 8.50 | 39.00 | 46.00 | 53.00 |
| Enewguinea | 7.60 | 8.20 | 8.90 | 38.00 | 44.00 | 50.00 |
| Flores | 8.10 | 8.50 | 8.90 | 8.10 | 14.00 | 19.00 |
| Fuego | 8.00 | 8.80 | 9.60 | 3.80 | 6.40 | 8.90 |
| Gloria | 8.30 | 8.60 | 8.80 | 1.20 | 2.00 | 2.80 |
| Hellenic | 8.00 | 8.60 | 9.10 | 2.00 | 6.00 | 10.00 |
| Hispaniola | 7.60 | 8.20 | 9.00 | 1.40 | 2.40 | 3.30 |
| Hjort | 7.60 | 8.00 | 8.70 | 2.60 | 4.30 | 6.00 |
| Honshu | 9.00 | 9.20 | 9.40 | 53.00 | 62.00 | 80.00 |
| Izu | 7.60 | 8.20 | 9.30 | 4.70 | 9.40 | 33.00 |
| Java | 7.80 | 8.60 | 9.40 | 5.80 | 12.00 | 41.00 |
| Kamchatka | 7.60 | 8.20 | 8.80 | 0.71 | 1.20 | 1.60 |
| Kermadec | 8.20 | 8.90 | 9.60 | 14.00 | 21.00 | 44.00 |
| Kurils | 9.00 | 9.30 | 9.60 | 56.00 | 64.00 | 72.00 |
| Makran | 8.10 | 8.80 | 9.50 | 5.40 | 9.00 | 13.00 |
| Manila | 7.80 | 8.50 | 9.20 | 4.40 | 13.00 | 61.00 |
| Manokwari | 7.80 | 8.10 | 8.50 | 11.00 | 14.00 | 17.00 |
| Mariana | 7.60 | 8.40 | 9.60 | 4.90 | 9.80 | 34.00 |
| Midamerica | 8.20 | 8.90 | 9.60 | 16.00 | 26.00 | 36.00 |
| Nankai | 8.50 | 8.80 | 9.20 | 38.00 | 43.00 | 47.00 |
| Newbritain | 8.00 | 8.40 | 8.90 | 55.00 | 64.00 | 73.00 |
| Newhebrides | 8.30 | 8.80 | 9.30 | 17.00 | 23.00 | 45.00 |
| North_african | 7.60 | 8.00 | 9.10 | 0.47 | 0.79 | 1.10 |
| Northsulawesi | 7.90 | 8.40 | 9.00 | 9.30 | 16.00 | 22.00 |
| Nphillip | 7.60 | 8.30 | 9.10 | 2.80 | 6.90 | 21.00 |
| Nwsulu | 8.00 | 8.50 | 9.00 | 2.30 | 3.80 | 5.40 |
| Palau | 7.60 | 7.80 | 8.70 | 0.37 | 0.73 | 2.60 |
| Peru | 8.80 | 9.20 | 9.60 | 43.00 | 49.00 | 55.00 |
| Puna | 8.60 | 9.10 | 9.60 | 51.00 | 59.00 | 66.00 |
| Puysegur | 7.80 | 8.40 | 9.00 | 7.40 | 10.00 | 12.00 |
| Ryukyu | 8.00 | 8.70 | 9.30 | 8.60 | 17.00 | 61.00 |
| Sagami | 8.00 | 8.20 | 8.50 | 14.00 | 16.00 | 18.00 |
| Sandwich | 7.60 | 8.20 | 9.00 | 6.20 | 12.00 | 44.00 |
| Scaribbean | 7.60 | 7.80 | 9.30 | 2.70 | 4.50 | 6.30 |
| Schile | 9.50 | 9.60 | 9.60 | 53.00 | 61.00 | 68.00 |
| Seram | 8.00 | 8.60 | 9.20 | 14.00 | 23.00 | 32.00 |
| Sesolomon | 8.10 | 8.70 | 9.20 | 29.00 | 34.00 | 39.00 |
| Shetland | 7.60 | 8.20 | 8.80 | 2.90 | 4.90 | 6.90 |
| South_yap | 7.60 | 8.00 | 8.80 | 0.48 | 0.96 | 3.40 |
| Southseram | 7.80 | 8.10 | 8.40 | 0.88 | 1.50 | 2.00 |
| Sphillip | 7.60 | 8.10 | 8.80 | 2.90 | 7.20 | 22.00 |
| Ssolomon | 8.00 | 8.40 | 8.80 | 40.00 | 46.00 | 53.00 |
| Ssolomonwood | 7.70 | 8.00 | 8.30 | 57.00 | 67.00 | 76.00 |
| Sulu | 8.20 | 8.40 | 8.70 | 2.50 | 4.20 | 5.80 |

(*Continued*)

**Table 1.** (*Continued*)

| Source name | $m^1_{w,max}$ | $m^2_{w,max}$ | $m^3_{w,max}$ | $c\dot{s}^1$ | $c\dot{s}^2$ | $c\dot{s}^3$ |
|---|---|---|---|---|---|---|
| Sumatra | 9.00 | 9.30 | 9.60 | 25.00 | 29.00 | 32.00 |
| Sumba | 8.30 | 8.80 | 9.40 | 6.70 | 13.00 | 47.00 |
| Timor | 7.60 | 8.50 | 9.40 | 1.70 | 2.80 | 3.90 |
| Tonga | 8.00 | 8.70 | 9.30 | 16.00 | 33.00 | 120.00 |
| Trobriand | 7.60 | 7.80 | 9.10 | 3.20 | 5.40 | 7.60 |
| Waleutians | 8.90 | 9.10 | 9.30 | 5.50 | 9.20 | 13.00 |
| Westmollucca | 8.50 | 8.70 | 8.90 | 1.60 | 2.70 | 3.80 |
| Wetar | 7.60 | 8.30 | 9.10 | 7.80 | 13.00 | 18.00 |
| Wnewguinea | 8.10 | 8.50 | 8.80 | 14.00 | 17.00 | 19.00 |
| Yap | 7.60 | 7.70 | 8.30 | 0.59 | 1.20 | 4.10 |

Three values are applied for each, and additionally all sources are assigned $b = 0.7$, 0.95 and 1.2, and $m_{w,min} = 7.5$.

and, secondly, a Gutenberg–Richter distribution with the density function truncated between the same range:

$$\begin{aligned} \mathrm{GR}(m_w) &= 10^{a-bm_w} - 10^{a-bm_{w,max}} \quad \text{for } m_{w,min} \le m_w \le m_{w,max} \\ &= 0 \quad \text{for } m_w > m_{w,max} \\ &= 10^{a-bm_{w,min}} - 10^{a-bm_{w,max}} \quad \text{for } m_w < m_{w,min} \end{aligned} \tag{6}$$

Here $\mathrm{GR}(m_w)$ is the exceedance rate for magnitude $m_w$ (events/year) and $a$ is a parameter controlling the overall frequency of earthquake events. Equation (5) was termed a characteristic distribution by Kagan (2002*a*, *b*) as it assigns a finite rate to events with magnitude exactly $m_{w,max}$, although it differs from the characteristic distribution of Youngs & Coppersmith (1985). Equation (6) is often termed a truncated Gutenberg–Richter distribution or a truncated Pareto distribution (Kagan 2002*a*, *b*; Kagan & Jackson 2013), although Youngs & Coppersmith (1985) refer to it as an exponential magnitude distribution. For the same source zone parameters, equation (5) implies a greater frequency of high-magnitude events than equation (6). Both appear as branches in the logic tree.

The parameter $a$ in equations (5) and (6) is adjusted so that the synthetic earthquake catalogue reproduces the seismically coupled fraction of the long-term moment rate (Bird & Kagan 2004):

$$\sum_{i \in E} \mu A_i S_i r_i = \xi \mu A_T c\dot{s}/\cos(\delta) \tag{7}$$

where the value of $a$ affects $r_i$ as explained below. In equation (7), $i$ represents the $i$-th synthetic earthquake event on the source zone among the finite set of all such synthetic events $E$, $\mu$ is the shear modulus ($3 \times 10^{10}$ Pa), $A_i$ (m$^2$) is the area of the $i$-th synthetic earthquake event with slip $S_i$ (m) and mean annual rate $r_i$, $A_T$ is the total interface area of the source zone (m$^2$), $\delta$ is the mean dip of the source zone in radians (for the strike-slip Gloria source zone, the factor $\cos(\delta)$ is ignored in equation 7 since the motion is transverse), and $\xi$ is the fraction of seismic moment released by earthquakes with $m_w \ge m_{w,min}$ on the source zone, which will be close to unity unless $m_{w,max}$ is close to $m_{w,min}$ (Kagan & Jackson 2013):

$$\xi = \frac{\left( \int_{m_{w,min}}^{m_{w,max}} (\mathrm{GR})' M_0(m_w)\, \mathrm{d}m_w \right)}{\left( \int_{-\infty}^{m_{w,max}} (\mathrm{GR})' M_0(m_w)\, \mathrm{d}m_w \right)} \tag{8}$$

Here $(\mathrm{GR})'$ is the derivative of GR (equations 5 and 6) with respect to $m_w$ (ignoring the lower magnitude bound: i.e. $m_{w,min} = -\infty$ for the derivative computation only), and $\mathrm{M}_0(m_w)$ is the seismic moment for an earthquake of magnitude $m_w$ (equation 3). The value of $\xi$ is independent of $a$ for both our modified Gutenberg–Richter relations (equations 5 and 6).

To compute $r_i$ when solving for $a$ (equation 7), note that:

$$r_i = \Pr(i|m_{w,i}) r(m_{w,i}) \tag{9}$$

where $\Pr(i|m_{w,i})$ is the conditional probability that the $i$-th earthquake event occurs given that an earthquake with the same magnitude has occurred, and $r(m_{w,i})$ is the mean annual rate of earthquakes of size $m_{w,i}$ occurring anywhere on the source zone. We assume that $\Pr(i|m_{w,i})$ is proportional to the area of the $i$-th event $A_i$ (equivalently, inversely proportional to its slip: equation 3) and to the average long-term slip rate along the trench where the $i$-th event occurs (denoted $\dot{s}_i$):

$$\Pr(i|m_{w,i}) = \frac{A_i \dot{s}_i}{\left( \sum_{j \in E(m_{w,i})} A_j \dot{s}_j \right)} \tag{10}$$

where the denominator is just a normalizing constant, and $E(m_{w,i})$ is the set of all events in the

synthetic catalogue with magnitude $m_{w,i}$. For each event, $\dot{s}_i$ is computed as the area-weighted-average of the local long-term slip rates on each unit-source which ruptures in the $i$-th event, with the unit-source long-term slip rates assigned from the nearest spatially varying trench-normal slip rates of Bird (2003) where available, and otherwise set to a constant. The use of equation (10) is further motivated below.

To model the mean annual rate of earthquakes of size $m_w$ in equation (9), we assume:

$$r(m_w) = \mathrm{GR}(m_w - 0.05) - \mathrm{GR}(m_w + 0.05) \quad (11)$$

where the factor 0.05 results from the synthetic earthquake catalogue containing events with $m_w$ in increments of 0.1. Equations (5) and (6) imply that equation (11) will contain a factor of $10^a$, which can be used to solve for $a$ using equations (7–11). Once $a$ is known, the mean annual rate of all events in the synthetic catalogue is determined directly with equations (9–11).

The use of equation (10) to assign non-constant conditional probabilities to events with the same magnitude contrasts with previous studies which employ a constant conditional probability (Horspool *et al.* 2014; Løvholt *et al.* 2014*a*; Lorito *et al.* 2016). As shown by the following idealized examples, equation (10) allows improved treatment of source zones with spatially variable width and/or long-term slip rate. First, consider a rectangular source zone where all unit-sources have the same area, but the long-term slip rate $\dot{s}_i$ varies along the trench. Equation (10) implies that an earthquake event on a fast-moving part of the source zone is proportionately more likely than an event with the same magnitude on a slower region, which we consider preferable to assuming they occur at the same rate. Secondly, consider a source zone where $\dot{s}_i$ is constant (i.e. uniform long-term slip rate) but the down-dip width is non-uniform because of variations in the source zone dip (as is typical: e.g. Fig. 2). Unit-sources then have greater area in wider parts than in narrower parts of the source zone, because of their logically rectangular layout. For a given magnitude, earthquake events with a smaller area will have higher slip $S_i$ than those with a greater area (since slip is inversely related to area: equation 3). Therefore, a constant conditional probability would imply greater long-term slip contributions from events in narrower parts of the source zone (where our synthetic earthquakes have lower area), which seems undesirable if $\dot{s}_i$ is actually constant along-strike. In contrast, equation (10) ensures that if $\dot{s}_i$ is constant, all events with the same magnitude induce the same long-term slip rate where they occur (i.e. $r_i S_i$ is constant for fixed $m_{w,i}$).

*Logic-tree weights and epistemic uncertainty quantification.* Owing to parameter uncertainties (see the earlier subsection on 'Source zone parameters'), it is necessary to compute the synthetic earthquake event rates for all combinations of $m_{w,max}$, $b$ and $c\dot{s}$ (three values each), and for the two different Gutenberg–Richter relationships (equations 5 and 6), leading to $2 \times 3^3 = 54$ different rate equations for each source zone (Fig. 3). These are combined with a weighted average to produce a mean magnitude exceedance rate equation on each source zone. The procedure for determining the weights is detailed below, and involves deriving a preliminary set of weights for each parameter combination, which are then updated based on the probability that the parameter combination would produce the observed events in the global CMT catalogue (Ekstrom *et al.* 2012).

On each source zone each parameter combination is assigned a preliminary weight equal to the product of its individual parameter weights. The individual parameter weights for each $c\dot{s}$ value are set to $1/3$; similarly, all $b$ values are given weight $1/3$; each variant of the Gutenberg–Richter model is given weight 0.5; and for $m_{w,max}$ we assign a weight of 0.45 to the upper value, 0.45 to the middle value and 0.1 to the smallest value, reflecting that it is unlikely $m_{w,max}$ has been observed on most source zones.

The weights defined with this heuristic procedure require further revision, because on some source zones particular parameter combinations lead to unrealistic earthquake rate predictions. For example, the Kermadec source zone is assigned $m_{w,max} \in \{8.2, 8.9, 9.6\}$, $b \in \{0.7, 0.95, 1.2\}$ and $c\dot{s} \in \{14, 21, 44\}$ mm a$^{-1}$. In order to conserve seismic moment, logic-tree branches with low $m_{w,max}$ and high $c\dot{s}$ need to have events with $m_w \geq m_{w,min}(=7.5)$ every few years on average (e.g. 0.42 a$^{-1}$ for the top curve in Fig. 3). In contrast, only one such event was observed in the 38 year global CMT catalogue 1976–2013 (Fig. 3: details on data selection are presented below). If the true rate of $m_w \geq 7.5$ events was 0.42, then on average we should observe about 16 events in 38 years, and we are extremely unlikely to observe only one or zero events (probability of $2 \times 10^{-6}$ for Poisson-distributed event times). Therefore, this particular branch of the logic tree seems unlikely to be correct. Although this conclusion is obvious from the absence of many events exceeding $m_{w,min}$ in the observational record, our method for assigning preliminary weights cannot detect such unrealistic logic-tree branches.

To address this issue, we update the preliminary weights based on the observed number of $m_w \geq m_{w,min}$ thrust earthquakes on the source zone. The observed events are counted using the global CMT

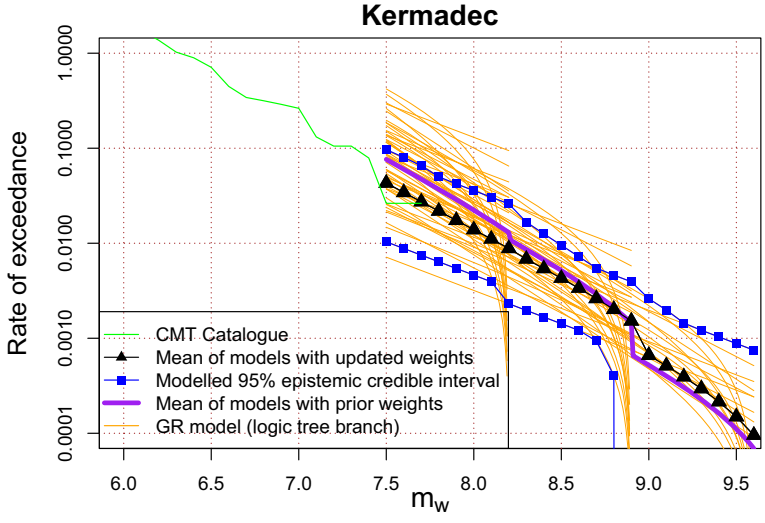


**Fig. 3.** Empirical and modelled $m_w$ exceedance rates (events/year) for the Kermadec source zone. The global CMT catalogue events include those within 0.2° of our Kermadec source zone, with depth <70 km and rake within $\pi/4$ of pure thrust. By updating the weights of the logic-tree branches based on the observed number of $m_w \geq m_{w,min}$ events, unrealistic parameter combinations (e.g. corresponding to the upper lines in this case) are given minimal weight.

catalogue (1976–2013) (Ekstrom *et al.* 2012) if they are within 0.2° of our source zones, have depth ≤70 km, $m_w \geq 7.5$ and rake within $\pi/4$ of pure thrust. The weight update is equivalent to treating the preliminary weights as prior probabilities for each parameter combination, and using Bayes' theorem to calculate posterior probabilities using the data (Kruschke 2011):

$$w_{ijkl} = \frac{w'_{ijkl} \Pr(r_e(m_w \geq 7.5)|ijkl)}{\Pr(r_e(m_w \geq 7.5))} \quad (12)$$

where $w_{ijkl}$ is the updated weight for the parameter combination with the $i$-th $m_{w,max}$ value, the $j$-th $b$ value, the $k$-th $c\dot{s}$ value and the $l$-th variant of the Gutenberg–Richter equation (equation 5 or 6); $w'_{ijkl}$ is the corresponding preliminary weight; $r_e(m_w \geq 7.5)$ is the empirically observed rate of thrust earthquakes in the global CMT catalogue; $\Pr(r_e(m_w \geq 7.5)|ijkl)$ is the probability that the empirical rate of earthquakes with $m_w \geq 7.5$ would be observed if the $ijkl$ rate equation were correct, and is computed assuming the event timings behave like a Poisson process with mean rate of $m_w \geq 7.5$ events computed from the $ijkl$ rate equation; and the denominator is a normalizing constant:

$$\Pr(r_e(m_w \geq 7.5)) = \sum_i \sum_j \sum_k \sum_l w'_{ijkl} \Pr(r_e(m_w \geq 7.5)|ijkl) \quad (13)$$

In the case of Kermadec, this leads to a mean rate curve with less frequent small earthquakes and more frequent large earthquakes, compared with the mean rate curve derived from the preliminary weights (Fig. 3). This reflects the down-weighting of unrealistic logic-tree branches which predict frequent $m_w$ events of 7.5 magnitude, as discussed above.

Although it is clearly desirable to down-weight logic-tree branches which are unlikely to be consistent with observations, in the case of Kermadec it is not visually obvious whether the 'updated-weights' mean rate curve is an improvement on the 'prior-weights' mean rate curve (Fig. 3). Because of natural variations in the observed number of earthquakes, we cannot be sure of improvements every time weight updating is applied, but improvements are expected on average when applied globally to many source zones. For illustration, consider a large number of hypothetical scenarios in which the true source zone rate curve is unknown, but corresponds to one particular branch in Figure 3 (with probability corresponding to the preliminary weights). Without weight updating, the difference between the true rate of $m_w \geq 7.5$ events and the modelled mean rate has a mean of zero and standard deviation of 0.066. However, if 38 years of data is randomly sampled from the true model and weight updating is applied, then the difference has a mean of zero and standard deviation of 0.037 (i.e. the typical error is almost halved). It is also worth noting that at the globally aggregated level, weight updating has little impact on the modelled magnitude

exceedance rate curve, with the largest change being a small reduction (*c.* 10%) in the rate of events with $m_w < 8$ which slightly improves agreement with globally aggregated data (see the subsection on 'Earthquake magnitude–frequency relationships' later in this paper). Although not pursued here, more extensive use of historical and palaeoseismic data may further improve the logic-tree weights. There is also potential to modify the approach to allow for possible non-Poisson event timings, and more fully exploit details of the earthquake magnitude observations (rather than just counting the observed events above $m_{w,min}$), although an advantage of ignoring such details is robustness to occasional large earthquakes.

The revised weights are also used to compute credible intervals representing our epistemic uncertainty: for example, the 95% credible interval is a region containing 95% of the total logic-tree weight for each $m_w$, with $\leq$2.5% of the total logic-tree weight above the upper bound and $\leq$2.5% below the lower bound (Fig. 3). Note that in this paper we use the term 'credible interval' to refer to intervals representing our epistemic uncertainty computed using the logic-tree weights, whereas the term 'confidence interval' is used for standard frequentist intervals (Kruschke 2011). The advantage of quantifying epistemic uncertainties is: (1) it highlights regions where the mean hazard is more likely to change in future hazard updates, as new knowledge may decrease epistemic uncertainties (Marzocchi *et al.* 2015); (2) the use of the weighted mean rate curve alone is insufficient testing PTHA predictions from the full logic tree (Marzocchi & Jordan 2014) and; (3) it allows the effects of uncertainty on risk assessment to be handled in an explicit manner (McGuire *et al.* 2005). A limitation of the credible interval approach used herein is that it implicitly assumes the model and logic tree cover all possible true models, which will not be correct. Although not pursued here, 'ensemble modelling' approaches have been proposed to address this issue by treating the logic-tree branches as a sample of the epistemic uncertainty which is subject to further statistical modelling (Marzocchi *et al.* 2015; Selva *et al.* 2016).

### *Wave propagation to hazard points*

For each unit-source, we model the tsunami caused by an earthquake with 1 m of slip using a numerical solution of the linear shallow-water equations in spherical coordinates on a grid with cell size $1' \times 1'$. We use the same finite-difference code as Thio *et al.* (2007), Burbidge *et al.* (2008) and Horspool *et al.* (2014), based on Satake (1995), with a fixed time step of 2 s. Global elevation values are set with the GEBCO $30'' \times 30''$ bathymetric grid. The model is evolved for 26 h of simulated time, as this is longer than required for trans-Pacific tsunami propagation from, for example, Chile to Japan (Wessel 2009), and also short enough for our simulations to complete within the time limits of our computing facility.

Because of computer storage limitations, we cannot save full time series for the unit-source tsunami model runs and so, instead, store the tsunami wave height time series at 20 s intervals at a subset of points near the coastline (termed hazard points). The hazard points are placed approximately along the 100 m depth contour at a spacing of 25 km, subject to the constraints that they are not too near ($<$1.5 arc-min) or too far ($>$22 arc-min) from the coastline. If this constraint is not met on the 100 m depth contour, then the hazard point is placed at the point closest to 100 m depth which satisfies the constraint.

Because of the linearity of the propagation model, the hazard point time series associated with any synthetic earthquake event in our catalogue can be produced by summing the time series for each contributing unit-source and multiplying by the earthquake slip (Thio *et al.* 2007). While very computationally efficient, a drawback is that we cannot model wave propagation in shallow water (*c.* 50 m). This is both because of insufficient mesh resolution to resolve the shorter tsunami wavelengths, and because non-linear processes associated with finite flow velocities (friction and momentum advection) are ignored in the linear model. Hence, our hazard points are situated offshore in depths of approximately 100 m, and other methods are required to estimate amplification and run-up in shallower water.

### *Wave amplification onshore*

Our tsunami propagation model cannot directly simulate inundation. While the usual practice in tsunami science is to use local inundation models for run-up estimation, in the current study this was not possible due to the large number of simulations, the global extent of the study, and the need for high-resolution topographic and bathymetric data. An alternative approach, which is less accurate but practically feasible, is required to estimate wave run-up (above mean sea level) from the hazard point water-surface elevation time series. Two different approaches are applied in this study.

First, we apply the amp-factor method of Løvholt *et al.* (2012*a*). This estimates the wave run-up from the modelled wave height, period and depth at the nearest hazard point, assuming that the nearshore topography can be approximated with one of seven different idealized bathymetric profile types (subjectively assigned to hazard points for this

study). For each bathymetric profile type, the amplification factors are derived by interpolation of the amplification predicted by numerical models for plane waves. Details are provided in Løvholt *et al.* (2012*a*), who found the run-up predictions compared fairly well with run-up predictions from detailed inundation models in a number of test cases.

Secondly, we estimate the tsunami run-up using Green's law to amplify the maximum water-surface elevation at the nearest hazard point, with the water depth at the coastline set to 0.5 m (Kamigaichi 2009; Sørensen *et al.* 2012; Brizuela *et al.* 2014; Horspool *et al.* 2014; Hébert & Schindelé 2015). Kamigaichi (2009) used Green's law to estimate nearshore peak tsunami wave height from the modelled water-surface elevation at offshore sites (*c.* 50 m depth), and found the estimates compared well with those from a fine mesh model for two separate earthquake–tsunamis with $m_w = 8.0$ and 6.8. Hébert & Schindelé (2015) found the approach gives a useful indication of zones of maximum run-up compared with field data for the Indian Ocean tsunami, although the errors were often of similar magnitude as the observed run-up, and no particular value of the coastline depth would consistently give the best predictions.

It is stressed that for plane non-breaking waves, the amplification factor method of Løvholt *et al.* (2012*a*) reproduces the tsunami run-up obtained from non-linear shallow-water models (Carrier & Greenspan 1958). Using Green's law for run-up estimation requires a subjective choice of the water depth at the coastline, and will produce infinite run-up heights as the coastline depth converges to zero (i.e. the shoreline). Regardless, we examine both methods here as Green's law is straightforward to apply once the coastline depth is selected, and has been employed in several past applications of PTHA.

### *Run-up exceedance rate computation*

At every hazard point, the methodology above provides the amplified tsunami run-up (see the previous subsection on 'Wave amplification onshore') and a corresponding mean annual rate for every earthquake event in the synthetic catalogue (equation 9). From this, we can compute the mean annual rate of events, with amplified wave height $h$ exceeding a threshold $h_T$ at any particular hazard point p, denoted $r_p(h \geq h_T)$:

$$r_p(h \geq h_T) = \sum_{i \in E} r_i \Pr(h \geq h_T | h_{i,p}) \qquad (14)$$

where $i$ is an event from the full synthetic earthquake catalogue $E$ with rate $r_i$, and $\Pr(h \geq h_T | h_{i,p})$ gives the probability that the real run-up exceeds $h_T$, given that the modelled run-up is $h_{i,p}$. The latter term accounts for the run-up variability associated with aleatory uncertainties (e.g. spatially variable slip) which are not directly simulated in the model, as well as model structural errors caused by, for example, errors in the subduction interface geometry and our simplified treatment of the hydrodynamics. Although the latter factors are actually epistemic uncertainties which might theoretically be treated in a logic tree, their characterization would be computationally expensive and so for simplicity herein they are lumped in with the aleatory uncertainties (Thio *et al.* 2010; Thio 2012; Horspool *et al.* 2014; Løvholt *et al.* 2016). We assume that $h$ follows a lognormal distribution with median equal to the predicted run-up at the nearest hazard point, plus some bias $\beta$. The probability of $h$ exceeding a threshold $h_T$ is:

$$\Pr(h \geq h_T | h_{i,p}) = 1 - \Phi\left(\ln(h_T) | [\ln(h_{i,p}) + \beta], \sigma\right) \qquad (15)$$

where, on the right-hand-side, $\Phi$ is the cumulative distribution function for a normal distribution with mean $[\ln(h_{i,p}) + \beta]$ and standard deviation $\sigma$, evaluated at $\ln(h_T)$ (Annaka *et al.* 2007; Thio 2012; Horspool *et al.* 2014).

Values for the logarithmic bias $\beta$ and standard deviation $\sigma$ in equation (15) were determined by comparing predictions of uniform slip scenarios in our PTHA database with observations from the NGDC/WDS global historical tsunami database (NGDC 2015) for four large earthquake–tsunami events: the 1960 Chile $m_w$ 9.5; the 1964 Alaska $m_w$ 9.2; the 2004 Andaman–Sumatra $m_w$ 9.2; and the 2011 Tohoku $m_w$ 9.0. The bias $\beta$ was estimated as the mean of $[\ln(h_{obs}) - \ln(h_p)]$, where $h_{obs}$ is the observed run-up height and $h_p$ is the predicted run-up height at the hazard point nearest to the observation. The standard deviation of the latter log-difference provided an estimate of $\sigma$.

## Results

### *Earthquake magnitude–frequency relationships*

For each source zone, magnitude–frequency curves similar to Figure 3 were derived and, in some instances, these have been compared with independently estimated exceedance rates based on longer-term historical or palaeoseismic data, or seismic moment conservation. For the Alaska source zone, the modelled $m_w = 9.2$ exceedance rate is 1/897 (1/3236, 1/538), where values in parenthesis give a credible interval covering logic-tree branches with 95% of the total weight, thus reflecting the

modelled epistemic uncertainty for the source zone. This is in reasonable agreement with 1/650–1/750 reported by Wesson *et al.* (2007) based on palaeoseismic studies. On the Honshu source zone, our modelled $m_w = 9.0$ exceedance rate is 1/486 (0, 1/209). The lower bound of 0 occurs because $m_w = 9.0$ is the smallest $m_{w,max}$ in the logic tree (corresponding to the 2011 Tohoku earthquake) with zero probability of being exceeded according to the corresponding truncated Gutenberg–Richter–frequency curves (equation 6). This estimate is comparable to values of 1/400–1/300, 1/532 and 1/880–1/260 derived using various seismic moment conservation type analyses, and 1/1300–1/400 based on tsunami deposit data (McCaffrey 2008; Uchida & Matsuzawa 2011; Kagan & Jackson 2013). For the Nankai source zone, our modelled $m_w = 8.0$ exceedance rate is 1/76 (1/136, 1/37), towards the higher end of the $1/(124 \pm 96)$ rate suggested by Burbidge *et al.* (2008) based on 11 observed events in 1500 years. On Nankai, our $m_w = 8.6$ exceedance rate is 1/284 (0, 1/137), again consistent with 1/660 years in Burbidge *et al.* (2008), noting the latter is based on only two events and thus very uncertain. In the South Chile source zone, we estimate an $m_w = 8.0$ exceedance rate of 1/48 (1/90, 1/27), and for $m_w = 8.5$ we obtain 1/131 (1/202, 1/86). These are comparable to estimates of $1/(65 \pm 41)$ and $1/(128 \pm 46)$, respectively, by Burbidge *et al.* (2008) based on the historical record since 1570. Overall, the agreement with previous studies is encouraging, and suggests the model is providing reasonable estimates of the exceedance rate of large earthquakes.

It is also of interest to compare the globally integrated magnitude–frequency curve with data. This is a challenging comparison for our model because each source zones' magnitude–frequency curve is determined separately by combining the seismic moment conservation model with the logic-tree parameters and weights. By updating the logic-tree weights with the global CMT catalogue data, we ensure some consistency between the model and data at the source zone specific level, but this process does not force the model to agree with data at the globally aggregated level. Regardless, when aggregated globally, the mean modelled exceedance rate is similar to the global CMT catalogue empirical exceedance rate (Fig. 4: CMT events include only those within 0.2° of our source zones with depth ≤70 km and a rake within $\pi/4$ radians of pure thrust). The largest discrepancy occurs for $m_w$ of approximately 8.1 where the model predicts a higher rate of events (Fig. 4). Since we expect the global CMT catalogue to be fairly complete for $m_w \geq 7.5$ (Kagan 2003), any discrepancies between the model and observed data are predominantly due to the following factors: (1) sampling variability due to the short duration (1976–2013) of the global CMT catalogue; (2) epistemic uncertainty leading to a wide range of observed rates being consistent with the model; and (3) bias in the model and inappropriate logic-tree weights.

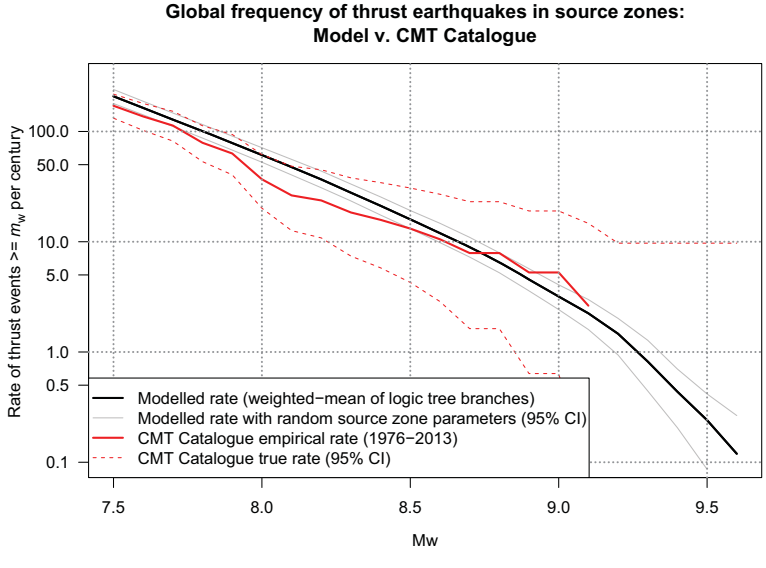


**Fig. 4.** Empirical and modelled $m_w$ exceedance rates over all source zones. Events from the global CMT catalogue were included if they had a depth ≤70 km, were within 0.2° of our source zones and had (strike–rake) within 45° of the associated source zone's updip direction.

To assess the likely impacts of sampling variability on global-scale differences between the model and data, we computed 95% confidence intervals for the true exceedance rate from the global CMT catalogue using a standard exact method (Garwood 1936). This assumes globally observed thrust earthquake events have Poisson behaviour in time. 95% confidence intervals for the true exceedance rate generally contained the modelled mean exceedance rate, although for $m_w$ of approximately 8.0 the model is close to the upper limit (Fig. 4). Based on this, we cannot confidently reject the consistency of the mean modelled rate and data at the global level. However, the uncertainties in the global rate are very large because the catalogue duration is relatively short (38 years), and there are no additional constraints imposed by models (e.g. there is no constraint from seismic moment conservation or from bounds on $m_{w,max}$ imposed in source zones with a small area). It would be possible to develop a narrower confidence interval using a longer duration database: however, this has not been pursued here since we need information on the earthquake rake to identify thrust events. To our knowledge such information is not comprehensively available in longer duration catalogues.

Even if the model is correct, epistemic uncertainties will lead to deviation between the true exceedance rate and the modelled weighted-mean rate because the latter is an average reflecting all parameter combinations in our logic tree. To assess the expected epistemic uncertainty in the globally aggregated exceedance rate, we randomly assigned each source zone to a branch of its logic tree and computed the global mean exceedance rate. The probability that each branch was assigned to its source zone was proportional to its weight $w_{ijkl}$ (equation 12). Figure 4 shows 95% credible intervals obtained by repeating the latter procedure 1000 times and computing the 2.5 and 97.5% quantiles. Importantly, this credible interval assumes that source zone parameters are independent of each other and that the model is correct, in which case the true global exceedance rate of thrust earthquakes on our source zones is likely to occur within the envelope. The substantial overlap of this credible interval and the global CMT-based confidence interval further indicates plausible consistency between the model and data at the global scale. The credible interval is much narrower than the confidence interval predicted using the 38 year global CMT catalogue (Fig. 4). In part, this reflects the information content of our source zone parameter combinations, their prior weights and the model itself: in particular, the assumptions that the earthquake rates conserve seismic moment, and that $m_{w,max}$ is constrained by the size of the source zone (using an empirical relationship of Strasser *et al.* 2010 herein). These assumptions constrain the modelled exceedance rates more strongly than does naive inference from the global CMT earthquake catalogue. The credible interval probably underestimates the true uncertainties because it neglects the possibility of correlations between the source zone parameters, which could arise, for example, if the upper $m_{w,max}$ or $c\dot{s}$ rates were actually appropriate on most source zones. Herein, we have not tried to model these correlations because it introduces additional complexity, and our conclusion that there is reasonable consistency between the model and the data would be unaffected by a broader credible interval. However, for other applications, a more careful accounting of these correlations may be required.

While the above tests do not suggest the model is inconsistent with the data, the model is nonetheless likely to be limited in a number of ways. Realistically, the model will have structural errors, and the logic-tree branches are not expected to be exhaustive of all possibilities (Field *et al.* 2014; Marzocchi *et al.* 2015). Owing to our conservative treatment of seismic coupling, no source zone has a maximum seismic coupling of less than 0.7 (Berryman *et al.* 2015). While a reasonable reflection of uncertainty at the level of an individual source zone, when globally aggregated this will lead to over-prediction of earthquake rates if many source zones have coupling values lower than 0.7. Similarly, the inclusion of a low $m_{w,max}$ value in the logic tree can lead to high rates of low $m_w$ earthquakes on some logic-tree branches (Fig. 3), potentially causing over-prediction of small events at the globally aggregated level. Although unrealistic magnitude–frequency curves are effectively downweighted by updating the logic-tree weights with the empirical rate of $m_w \geq 7.5$ thrust events in the global CMT catalogue, at the individual source zone level they may appear reasonable given the limited available data.

In summary, at the source zone specific level our tests suggest reasonable agreement of several modelled magnitude–frequency relationships with data. There is typically considerable uncertainty in the modelled rates, and in some locations this uncertainty might be reduced by integrating palaeoseismic data into the model. However, even in its current form, when aggregated globally, the modelled magnitude–frequency curve is consistent with the observed rate of thrust events in the global CMT catalogue.

### *Wave run-up from individual events*

The PTHA scenario amplified wave heights were compared against tsunami run-up observations for four large historical events in the NGDC/WDS

**Table 2.** *Summary statistics of the NGDC/WDS historical tsunami database run-up observations, compared with a corresponding event in the PTHA database with the same magnitude and source zone which best fits the data*

| Event | Amp-factor bias $\beta$ | Amp-factor $\sigma$ | Green's law bias $\beta$ | Green's law $\sigma$ |
|---|---|---|---|---|
| 1960 Chile $m_w = 9.5$ | −0.188 | 0.789 | −0.514 | 0.759 |
| 1964 Alaska $m_w = 9.2$ | −0.107 | 1.073 | −0.059 | 1.010 |
| 2004 Indian Ocean $m_w = 9.2$ | −0.0584 | 0.8715 | 0.075 | 0.847 |
| 2011 Tohoku $m_w = 9.0$ | −0.302 | 0.954 | −0.079 | 0.924 |

The amp-factor method follows Løvholt *et al.* (2012*a*) to estimate the wave run-up height. Alternatively, Green's law was applied to a 0.5 m water depth. For each event, we computed $[\ln(h_{obs}) - \ln(h_p)]$, where $h_p$ is the PTHA amplified wave height spatially closest to the observed run-up $h_{obs}$. The bias is the mean of the logarithmic difference and $\sigma$ is its standard deviation.

historical tsunami database (Table 2; Fig. 5). For each observed event, the PTHA scenario had the same $m_w$ and source zone, and where several such PTHA scenarios existed the best-fitting one was used. This allows estimation of the logarithmic bias $\beta$ and standard deviation $\sigma$ relating the predicted wave run-up height to the observations (see the earlier subsection on 'Wave amplification onshore').

Overall, the amp-factor method (Løvholt *et al.* 2012*a*) and Green's law produced a similar logarithmic bias $\beta$ and standard deviation $\sigma$ (Table 2). The bias varied substantially among the events, with the amp-factor method always over-predicting median run-up, while the Green's law approach did so in three-quarters of the cases (Table 2). Given the small overall difference in the performance of these methods, we hypothesize that the model error is due to the simplified uniform-slip representation of the tsunami source, and the fact that the complex onshore inundation, including local phenomena such as focusing and refraction, is not included in the amp-factor method, as well as the neglect of friction terms (Fritz *et al.* 2008; Lynett *et al.* 2012). The $\sigma$ values derived here are somewhat larger than the $\sigma = 0.71$ value previously reported by comparing coarse-grid uniform slip tsunami simulations with observations, and accounting for uncertainties in dip and non-uniform slip distributions (Thio *et al.* 2010; Thio 2012). Thio *et al.* (2010) stated $\sigma$ could be further reduced from 0.71 to 0.52 using a finer-grid tsunami propagation model (90 m cell size), which suggests that a substantial improvement in the hydrodynamic modelling may yield only moderate improvements in the accuracy of run-up predictions. Our results suggest that $\sigma$ values even larger than used previously may be appropriate for a 1 arc-min resolution linear shallow-water model with initial conditions based on uniform-slip earthquakes (Thio *et al.* 2010; Horspool *et al.* 2014; Løvholt *et al.* 2014*a*).

While noting similar overall performance of both amplification methods trialled in this study, hereinafter we estimate run-up heights using the amp-factor method (Løvholt *et al.* 2012*a*) because it has a more firm theoretical basis for quantifying the run-up than the Green's law approach (see above). We employ a bias $\beta = -0.164$ and $\sigma = 0.927$, corresponding to the mean and root-mean-square of the respective values in Table 2. This implies that approximately 54% of run-up heights are predicted within a factor of 2 by the best corresponding PTHA scenario.

### *Wave run-up exceedance rates*

The global pattern of 1/500 and 1/2500 exceedance rate run-up heights are shown in Figure 6. At the 1/500 exceedance rate, the model predicts high run-up around most of the Pacific Rim and eastern and southern Indonesia, with relatively low run-up heights in the Atlantic and western Indian oceans, southern Australia, and other areas which are significantly sheltered from the major subduction sources. Areas with intermediate run-up heights include many Pacific Islands, southern New Zealand, and parts of eastern and western Australia. At the 1/2500 exceedance rate, our model predicts large areas with an exceedance wave height >10 m in the Pacific and eastern Indian Oceans, as well as the Western Mediterranean near the Hellenic and Cyprus source zones, and areas adjacent the Makran source zone (Fig. 6).

The run-up heights associated with these rare events are subject to considerable epistemic uncertainty because they involve large-magnitude earthquakes. The logic-tree branches usually assign widely varying rates to large earthquakes, especially where no such events have been observed, implying that $m_{w,max}$ has a large range (Fig. 3). For example, in the Kermadec source zone, the 95% epistemic credible intervals for the rate of events with $m_w > 8.9$ always includes a lower bound of zero (Fig. 3) because our epistemic uncertainty puts significant weight on the possibility that $m_{w,max} \leq 8.9$. To quantify the impact of this on tsunami wave heights, we compute new exceedance

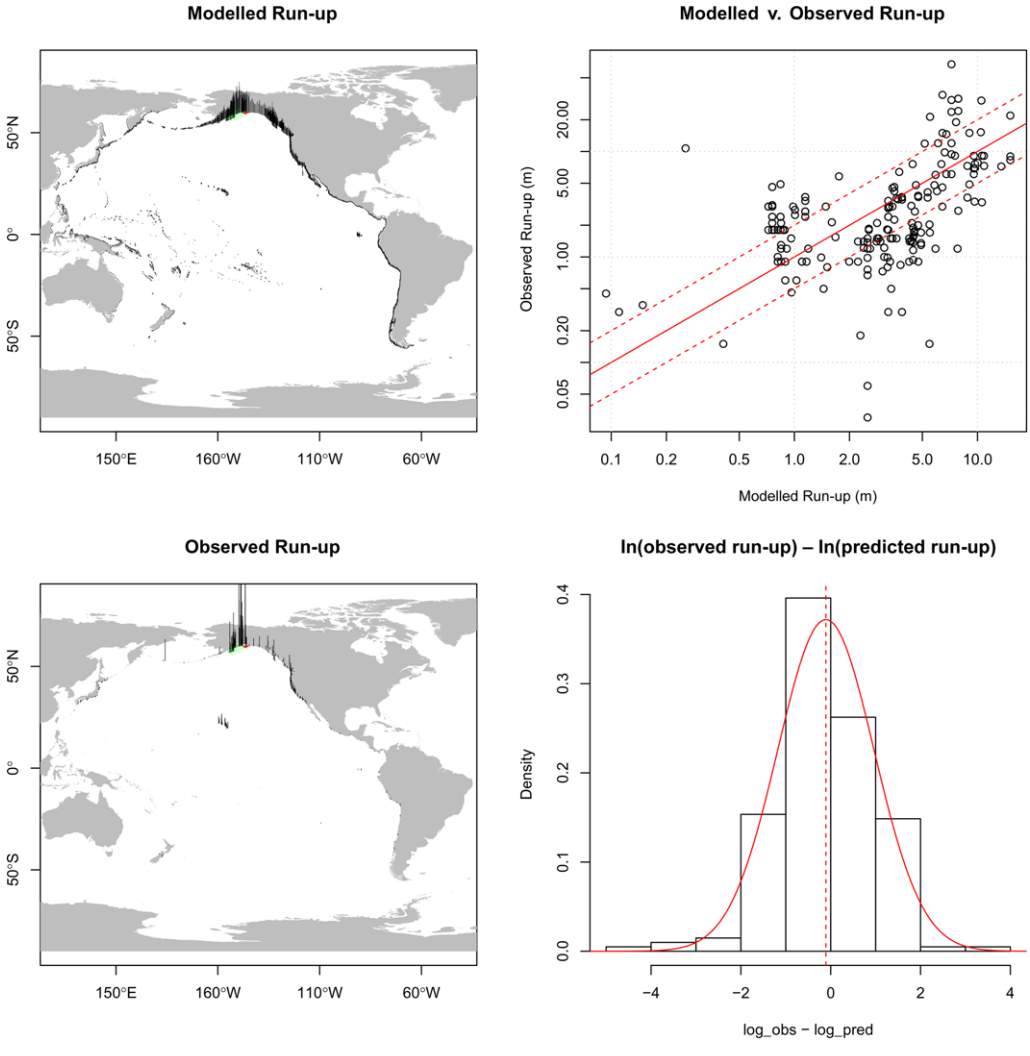


**Fig. 5.** Observations of the 1964 Alaska tsunami run-up in the NGDC/WDS historical tsunami database compared with the corresponding PTHA scenario. (Top left) Modelled wave run-up heights (>0.1 m). (Bottom left) Observed wave run-up heights (>0.1 m). (Top right) Modelled v. observed run-up heights. The solid line is $y = x$, and the dotted lines are $y = 2x$ and $y = 0.5x$. (Bottom right) Histogram of errors in log run-up and the corresponding normal distribution.

wave heights, assuming that the rates on all source zones matched, first, the lower 2.5% and, secondly, the upper 97.5% quantiles of our epistemic uncertainties (Fig. 7). For any coastal site, we can then investigate the range in exceedance wave heights predicted if all source zones which significantly affect it have 'high' or 'low' earthquake rates relative to our expectations. This simple method of uncertainty quantification is conservative because it ignores the potential for uncertainties from multiple source zones to average-out if their parameters are uncorrelated. However, at the site-specific level, we expect tsunami hazard is usually dominated by just a few source zones, which may have correlated parameters (e.g. we might consider the highest or lowest $m_{w,max}$ or $c\dot{s}$ value could be appropriate for all relevant sources). In the absence of detailed modelling of the inter-source zone parameter correlations, the conservative approach seems more reasonable for site-specific hazard estimation than does ignoring inter-source zone parameter correlations entirely.

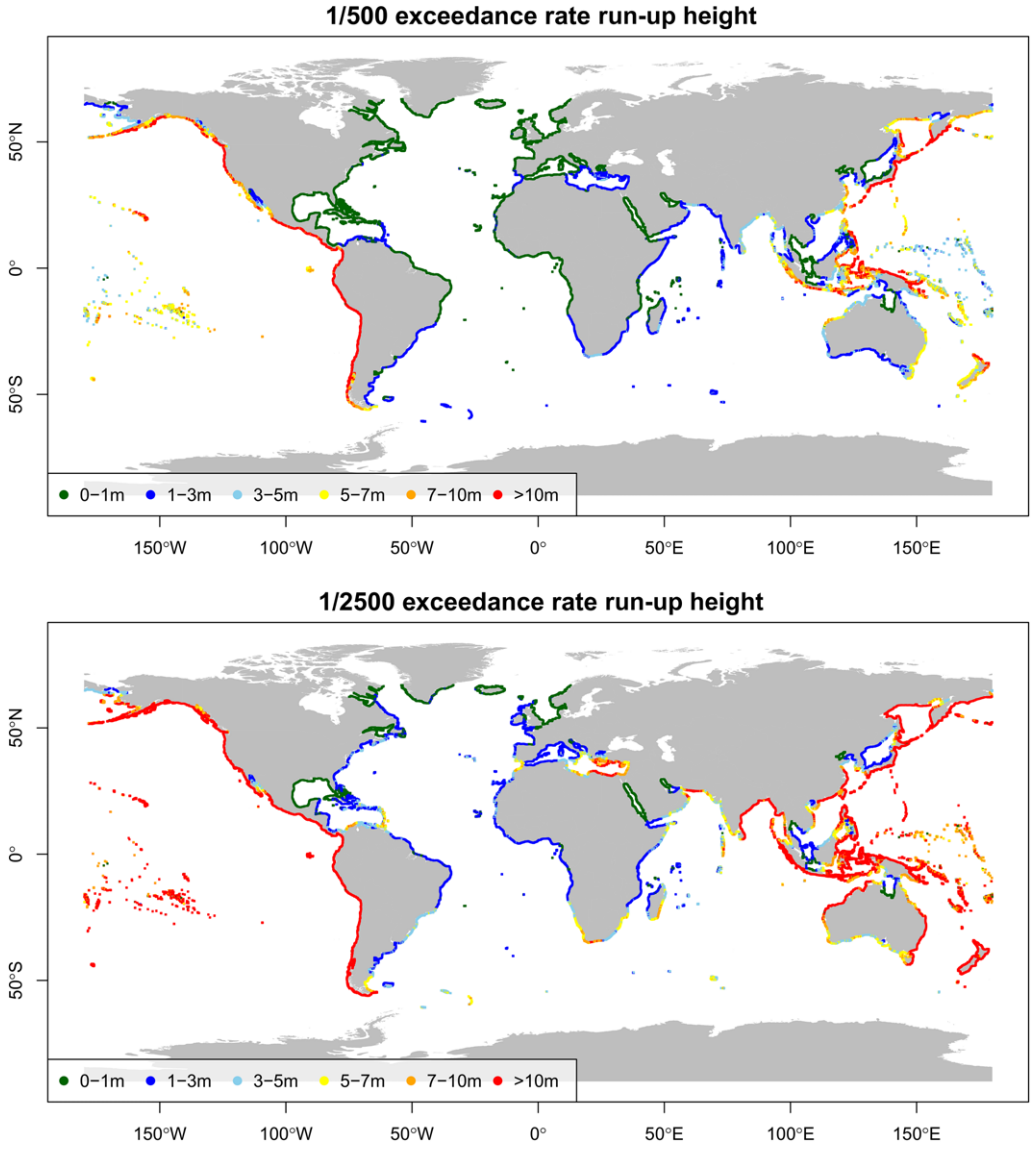


**Fig. 6.** Wave run-up heights associated with 1/500 (top) and 1/2500 (bottom) exceedance rates.

There is high epistemic uncertainty in the 1/500 exceedance rate run-up for many locations globally, with the 95% credible interval for run-up ranging from 1 to >10 m in much of Oceania and the eastern Indian Ocean (Fig. 7). This reflects exposure to source zones where $m_{w,max}$ and the seismic coupling rate have substantial epistemic uncertainty, such as Kermadec (Fig. 3). High waves are predicted even at the lower limit of the 95% credible interval near rapidly converging source zones with high seismically coupled slip rates (i.e. the South American and Japan/Kurils source zones) (Fig. 7). Other sites with high hazard even at the lower limit of the 95% credible interval include Cascadia, southern Alaska, parts of northern and eastern New Guinea, Sumatra, and the northern Moluccas, which are exposed to source zones with high seismically coupled slip rates (Berryman *et al.* 2015). Much of the Atlantic and Mediterranean coast is predicted to have relatively low tsunami hazard (<3 m run-up) for the 1/500 exceedance rate within our 95% credible

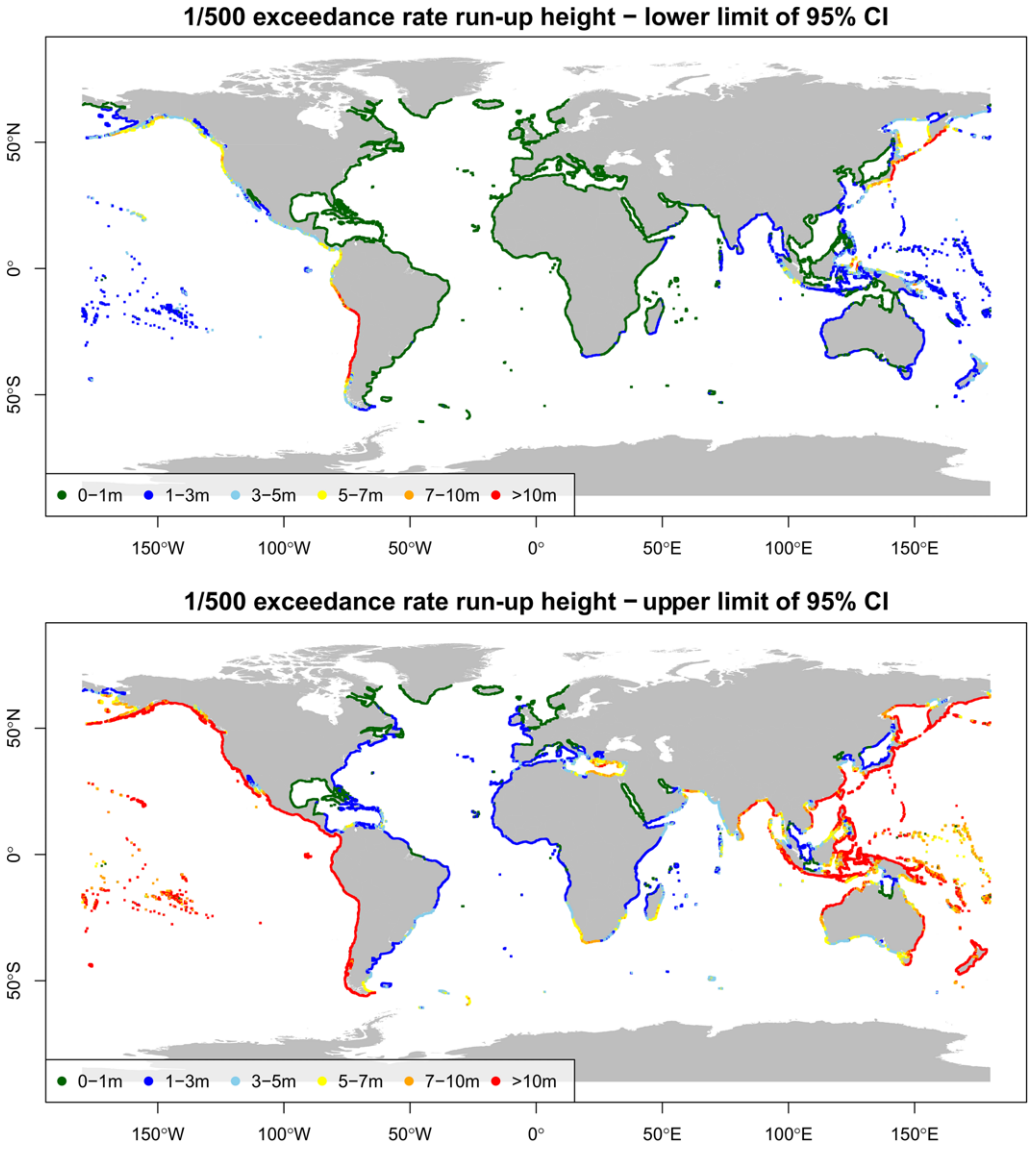


**Fig. 7.** 95% credible interval (CI) for the wave height with a 1/500 exceedance rate. Lower credible interval bound (top) and upper credible interval bound (bottom).

intervals, with exceptions close to source zones in the east and south Caribbean and the Eastern Mediterranean.

It is worth noting that in the northern Moluccas (NE Indonesia), the convergence rates suggested in Berryman *et al.* (2015) for the Halmahera source (90–100 mm $a^{-1}$, as used in this study for the Eastmolucca source zone) are significantly higher than those used by Løvholt *et al.* (2012*b*) and Horspool *et al.* (2014) (10–30 mm $a^{-1}$), and this contributes to our high hazard predictions in this area (Figs 6 & 7). Even assuming lower convergence rates, the area is considered high hazard (Løvholt *et al.* 2012*a*, *b*; Horspool *et al.* 2014). The higher convergence rates used here seem consistent with a geodetic velocity value reported in Bird (2003). However, the plate model of Bird (2003) suggests convergence rates vary rapidly in space. Given its complexity, this region warrants further analysis in future.

## Discussion and conclusions

### *Comparison with previous global analyses*

The current study builds on several previous global-scale analyses of tsunami hazard from earthquake sources (Løvholt *et al.* 2012*a*, 2014*a*) which largely estimated 1/500 annual exceedance rate run-up heights by direct simulation of tsunami from 1/500 earthquake scenarios. To demonstrate the accuracy of the simpler scenario-based method, the hazard from earthquake scenarios was compared with a full PTHA in the Indian and southern Pacific oceans in Løvholt *et al.* (2014*a*), with the PTHA results producing less conservative predictions than scenario methods in the majority of cases. The most significant advance herein is the global extension of the PTHA, which allows global estimation of wave heights for arbitrary exceedance rates with a uniform methodology (Fig. 6), and quantification of the associated epistemic uncertainties (Fig. 7).

There are substantial methodological differences between the current study and the PTHA component of Løvholt *et al.* (2014*a*). Rather than using constant dip source zone geometries, we use detailed subduction zone geometries where available (Hayes *et al.* 2012; Basili *et al.* 2013*a*, *b*), and allow for along-strike variations in the long-term convergent slip rate and downdip width to be accounted for when estimating the relative rates of thrust earthquakes with the same magnitude (equation 10). Our approach to defining the range of $m_{w,max}$ in the logic tree more closely follows Berryman *et al.* (2015) and often leads to a greater range of $m_{w,max}$ than the $m_{w,max} \pm 0.2$ approach applied previously. We also treat epistemic uncertainties in the Gutenberg–Richter $b$ parameter, which was assumed to be unity in the earlier study; and we use a simple technique to down-weight unrealistic logic-tree branches based on the observed rate of thrust events $\geq 7.5$ in the global CMT catalogue (equation 12). While our estimation of run-up from the offshore wave heights follows the same approach (Løvholt *et al.* 2012*a*), in this study uncertainties in the PTHA scenario run-up heights have been reassessed by comparison against run-up observations from four large tsunamis (Table 2). This has led to substantially larger $\sigma = 0.92$ compared to the values $\sigma = 0.52$ or 0.71 that have been used previously (Thio *et al.* 2010; Horspool *et al.* 2014). The value of $\sigma$ has a particularly strong effect on the results, with higher $\sigma$ associated with a major increase in the run-up height for a given exceedance rate (compare Figs 6 & 8). Using the lower $\sigma = 0.5$, the 1/500 run-up height reduces considerably, with values of less than 5 m now dominating over much of the central and SW Pacific Ocean, and in the Indian Ocean away from the Sunda Arc. If the uncertainties in the modelled tsunami run-up height are entirely neglected ($\sigma = 0$), then the 1/500 run-up height further reduces (Fig. 8).

Despite these methodological differences, there is broad similarity of our mean 1/500 run-up heights with those of Løvholt *et al.* (2014*a*), with relatively high hazard around much of the Pacific Rim and southern and NE Indonesia, and relatively low hazard in much of the Atlantic and western Indian oceans. The current study often predicts slightly higher run-up: for example, in the central south Pacific and eastern Australia, we often estimate 1/500 run-up heights of approximately 5–7 m, whereas corresponding values are typically $\leq 5$ m in Løvholt *et al.* (2014*a*) (Fig. 6). This is consistent with the use of a higher $\sigma$ in the current study (Figs 6 & 8). There are also regions where the current study predicts lower 1/500 tsunami run-up: for example, around northern Sumatra. One reason for this seems to be our use of Slab 1.0 to define the source zone geometry, which in this location leads to unit-sources having low dip near the shallow part of the source zone compared with the constant dip value used previously. The lower dip implies less vertical coseismic deformation for a given slip on the shallowest unit-sources. In addition, the sources used previously had a top-depth of 3 km (Horspool *et al.* 2014), compared with 0 km herein, so we expect higher slip near the trench in the former case (Goda 2015). Given the substantial methodological difference between the current study and Løvholt *et al.* (2014*a*), we are not surprised by some variations in the results: however, in general, the differences do not seem large compared with our epistemic uncertainties (Fig. 7).

### *Comparison with regional studies*

At a number of sites, the hazard results from this study have been compared with previous estimates from regional-scale tsunami hazard assessments. Overall, the comparisons reported below are encouraging, and suggest that the current PTHA methodology is providing estimates of tsunami hazard associated with major earthquake sources that are reasonably consistent with previous work. Despite this, given the global scale of this study and associated limitations (see the following subsection on 'Limitations'), we suggest that the current study be used for global-scale evaluations of tsunami hazard and risk (Løvholt *et al.* 2014*b*), and as a reference against which the results of smaller-scale studies may be compared. Detailed tsunami hazard assessments combining inundation modelling and historical or palaeotsunami data (Gonzalez *et al.* 2009; Power 2013) provide the

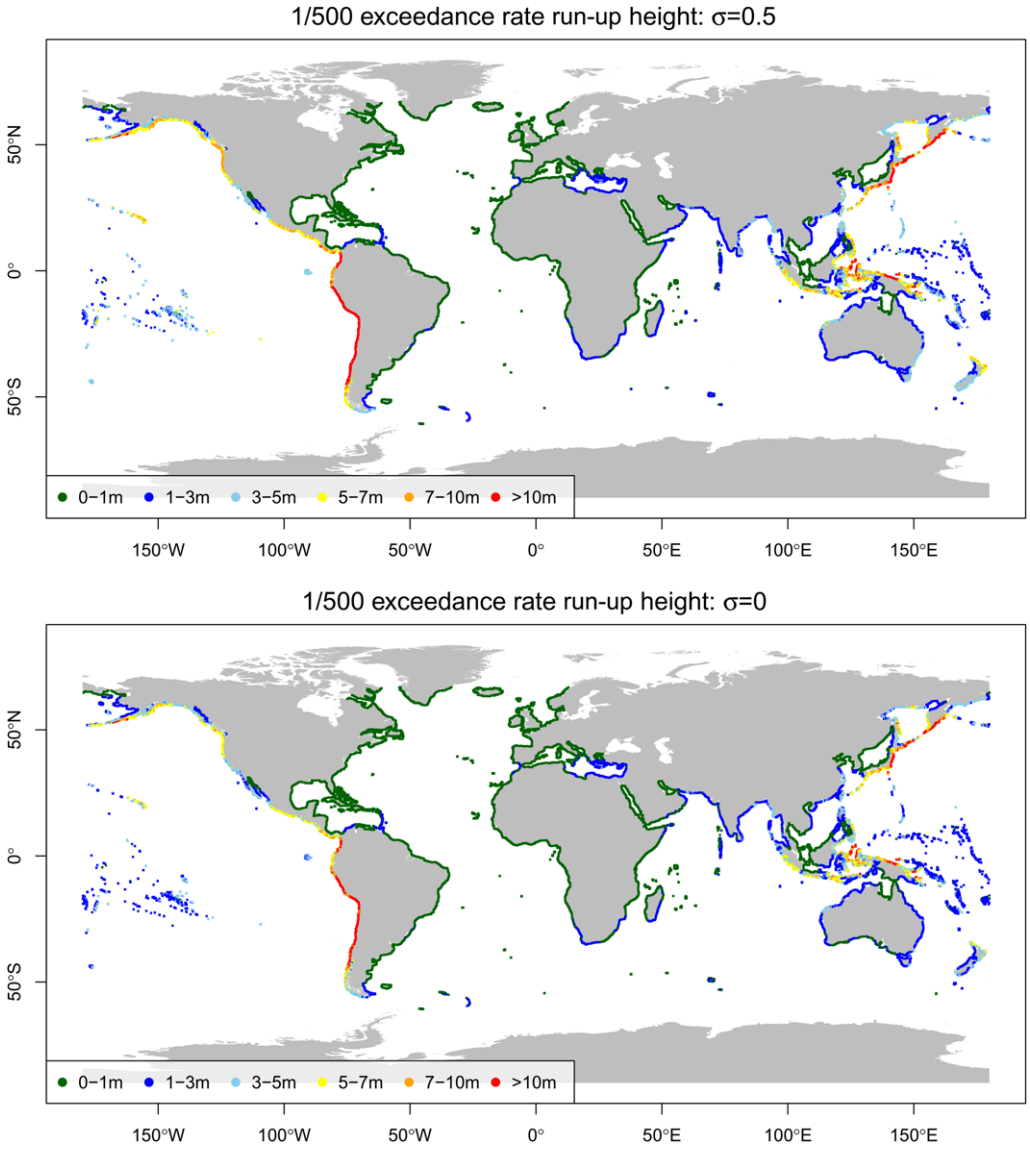


**Fig. 8.** The 1/500 exceedance rate run-up height with (top) $\sigma = 0.5$ and (bottom) $\sigma = 0$.

most appropriate information to support local-scale disaster management and planning.

In Acapulco, Mexico, Geist & Parsons (2006) analysed tsunami run-up data, and suggested 1/10 exceedance run-up of approximately 1 m, 1/100 exceedance run-up of approximately 6.5 m and 1/500 exceedance run-up of approximately 10 m. This compares quite well with the corresponding values for our nearest hazard point of 0.7, 4.2 and 10.7 m, respectively.

Along the coast of western Australia, Burbidge *et al.* (2008) modelled the exceedance rates of offshore wave height (in 100 m water depth). Tsunami hazard here is dominated by the eastern Sunda Arc. They predicted a spatial pattern in 1/500 offshore wave heights similar to that estimated herein for run-up heights (Fig. 6), with a distinct peak around latitudes of approximately 20°–22° S. In the latter region, Burbidge *et al.* (2008) modelled 1/500 wave heights of around 0.7 m in 100 m water depth (implying *c.* 3–5 m run-up height assuming amplification factors of *c.* 4–7). This is lower than the approximately 5–10 m predicted in the current study, although well within our

epistemic uncertainties (Fig. 7). While there are substantial methodological differences between the current study and Burbidge *et al.* (2008), it seems likely that the differences are in part caused by the logarithmic uncertainties in predicted wave height being employed in the current study. This factor was not treated in Burbidge *et al.* (2008) but leads to an overall increase in the run-up (Figs 6 & 8).

In the Caribbean, Parsons & Geist (2009) estimated the probability of tsunami run-up exceeding 0.5 m over a 30 year time horizon using a combination of models and historical data. Their results suggesting highest hazard in the eastern Lesser Antilles, with 0.5 m run-up exceedance rates of approximately 0.0035–0.0075 events/year. Similar 0.5 m run-up exceedance rates are predicted in the current study, with 95% of eastern Lesser Antilles sites in $\in$ [0.0032, 0.0079].

In the town of Seaside, Oregon on the NW USA coast, Gonzalez *et al.* (2009) performed a detailed local-scale PTHA, and estimated a 1/100 exceedance wave height of approximately 4 m and a 1/500 exceedance wave height of approximately 10.5 m. This compares quite well with values of 5.0 and 13.2 m, respectively, estimated at our nearest hazard point.

In the Mediterranean Sea, Sørensen *et al.* (2012) conducted a detailed PTHA, including many more earthquake sources than treated in the current study. They predicted highest 1/500 run-up (*c.* 5 m) in the Eastern Mediterranean around the Hellenic source zone, with low run-up $<$1 m in the Western Mediterranean. Similar patterns and magnitudes for 1/500 tsunami run-up are predicted in this study (Fig. 6).

In New Zealand Power (2013) compiled detailed analyses of tsunami hazard, combining modelling with historical and palaeotsunami data: 1/500 wave heights largely ranged within 4–12 m on the north coast, 4–8 m on the east and south coast, and 2–4 m on the west coast. Our model suggests comparable values of typically 7–10+ m around the north coast, 5–10 m around the east and south coast, and 3–7 m around the west coast (Fig. 6).

In Indonesia, Horspool *et al.* (2014) modelled exceedance rates for tsunami run-up height, predicting 1/500 run-up heights of 5–10+ m along the south and NE coastlines. While similar results are suggested herein, the current study suggests that the high hazard zone in eastern Indonesia may extend further southwards to include regions adjacent to the Flores and Banda Sea (Fig. 6). This is largely because our logic tree includes greater maximum $m_{w,max}$ values for source zones in this region than Horspool *et al.* (2014). However, $m_{w,max}$ values in this region are highly uncertain, leading to large epistemic uncertainties in tsunami run-up exceedance rates (Fig. 7).

In the NE Atlantic, Omira *et al.* (2015) modelled tsunami wave height exceedance rates, suggesting highest hazard on the mainland around the Gulf of Cadiz with exceedance rates of approximately 0.007 for a 1 m wave (i.e. a 50% chance of exceedance in 100 years). The current study also predicts highest mainland hazard around the Gulf of Cadiz, although our 1 m wave height exceedance rates in this region are lower (*c.* 0.002 − 0.003), which seems to be driven by lower estimates of earthquake exceedance rates on the key regional source zones in the present study.

### *Limitations*

The current study only considers tsunami hazard due to thrust earthquakes on sources identified in Figure 1 (aside from the strike-slip Gloria source zone), and we expect this to oversimplify the earthquake sources for some areas: for example, in the Mediterranean (Tiberti *et al.* 2009; Sørensen *et al.* 2012; Basili *et al.* 2013*b*). Uncertainties in the source zone schematization are not accounted for in the analysis but are expected to be significant, particularly away from major subduction zones. In the current methodology earthquakes are not permitted to cross source zones, and so our fault segmentation imposes potentially artificial limits on the range of events which are included in the model. We ignore earthquakes with $m_w < 7.5$, all non-thrust earthquake sources (except the strike-slip Gloria source zone) and most non-subduction sources, as well as tsunami from landslides, volcanoes and meteotsunami, although these are of significance in some contexts (Sørensen *et al.* 2012; Geist *et al.* 2014; Harbitz *et al.* 2014; ten Brink *et al.* 2014): for example, landslides are expected to contribute significantly to tsunami hazard in eastern Indonesia (Løvholt *et al.* 2012*b*), the Caribbean (Harbitz *et al.* 2012) and the Atlantic coast of the USA (ten Brink *et al.* 2014).

A significant source of uncertainty is related to the schematization of our earthquake sources. We assume a constant shear modulus $\mu$ which in reality will be spatially variable, with low near-trench values potentially leading to tsunami earthquakes with high slip and run-up for their magnitude (Geist & Bilek 2001; Newman *et al.* 2011). Although uncertainty in modelled run-up is accounted for using the lognormal distribution, parameters were estimated from high–magnitude earthquake–tsunami events and do not appropriately account for the tsunami hazard due to tsunami earthquakes. Uncertainties in the source zone geometry (e.g. dip, maximum depth) and the non-uniform slip distribution of real earthquakes are also not explicitly modelled,

although to some extent they are indirectly accounted for via the lognormal run-up distribution. Even for uniform slip earthquakes, small changes in the source zone geometry can have a strong impact on models of the resulting tsunami (Burbidge *et al.* 2015), and this is probably an important factor driving the large variation in wave heights predicted by different tsunami early warning systems (Greenslade *et al.* 2014). In addition, our models of earthquake rates show substantial epistemic uncertainty (e.g. Fig. 3), and our quantification of this is influenced by the subjective choice of logic-tree parameters and their weightings. On the one hand, the possible range of source zone parameters is not fully considered (since only three values are used for $m_{w,max}$, $b$ and $c\dot{s}$) suggesting the uncertainties may be underestimated. Conversely, it may be possible to reduce the uncertainties through better constraints on, for example, $m_{w,max}$ using the palaeo-record (Holschneider *et al.* 2014).

Given the global nature of this study, we have not been able to explicitly simulate tsunami run-up, but instead estimate run-up from offshore tsunami propagation models combined with amp-factors. This does not account for realistic nearshore topography which can have a large impact on tsunami run-up, and undoubtedly contributes to the large uncertainties in predicted run-up when PTHA scenarios are compared with observed events (Table 2). It may be possible to reduce these uncertainties using a higher-resolution nearshore tsunami-propagation model (Thio *et al.* 2010) which would have the advantage of concentrating higher run-up in areas where the nearshore bathymetry is particularly conducive to amplification. However, this would be computationally expensive on the global scale since the linear shallow-water equations will cease to be appropriate, and with non-linear models we cannot rely on superposition to efficiently produce many scenarios. Further, Thio *et al.* (2010) reported only moderate reductions in the logarithmic standard deviation $\sigma$ using a high-resolution hydrodynamic model, suggesting that large run-up uncertainties are predominantly driven by the representation of earthquakes, and may persist even using a more complex (and computationally expensive) treatment of tsunami inundation.

This paper is published with the permission of the CEO, Geoscience Australia. The tsunami propagation simulations in this research were undertaken on the NCI National Facility in Canberra, Australia, which is supported by the Australian Commonwealth Government.

**Correction notice:** The original version was incorrect. There was an error in the legend for the bottom part of Figure 7. 'lower' had been replaced by 'upper'.

## References

Annaka, T., Satake, K., Sakakiyama, T., Yanagisawa, K. & Shuto, N. 2007. Logic-tree approach for probabilistic tsunami hazard analysis and its applications to the Japanese coasts. *Pure and Applied Geophysics*, **164**, 577–592, https://doi.org/10.1007/s00024-006-0174-3

Basili, R., Kastelic, V. *et al.* 2013*a*. *The European Database of Seismogenic Faults (EDSF) compiled in the framework of the Project SHARE*. SHARE (Seismic Hazard Harmonization in Europe), Swiss Seismological Service, Zurich, https://doi.org/10.6092/INGV.IT-SHARE-EDSF; http://diss.rm.ingv.it/share-edsf/

Basili, R., Tiberti, M.M., Kastelic, V., Romano, F., Piatanesi, A., Selva, J. & Lorito, S. 2013*b*. Integrating geologic fault data into tsunami hazard studies. *Natural Hazards and Earth System Science*, **13**, 1025–1050, https://doi.org/10.5194/nhess-13-1025-2013

Berryman, K., Wallace, L. *et al.* 2015. *The GEM Faulted Earth Subduction Interface Characterisation Project: Version 2.0 April 2015*. GEM Technical Report GEM (Global Earthquake Model) Faulted Earth Project, available from http://www.nexus.globalquakemodel.org/gem-faulted-earth/posts

Bird, P. 2003. An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, **4**, 1–52.

Bird, P. & Kagan, Y.Y. 2004. Plate-tectonic analysis of shallow seismicity: apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling. *Bulletin of the Seismological Society of America*, **94**, 2380–2399.

Brizuela, B., Armigliato, A. & Tinti, S. 2014. Assessment of tsunami hazards for the Central American Pacific coast from southern Mexico to northern Peru. *Natural Hazards and Earth System Science*, **14**, 1889–1903, https://doi.org/10.5194/nhess-14-1889-2014

Burbidge, D. & Cummins, P. 2007. Assessing the threat to Western Australia from tsunami generated by earthquakes along the Sunda Arc. *Natural Hazards*, **43**, 319–331, https://doi.org/10.1007/s11069-007-9116-3

Burbidge, D., Cummins, P., Mleczko, R. & Thio, H. 2008. A probabilistic tsunami hazard assessment for Western Australia. *Pure and Applied Geophysics*, **165**, 2059–2088, https://doi.org/10.1007/s00024-008-0421-x

Burbidge, D., Mueller, C. & Power, W. 2015. The effect of uncertainty in earthquake fault parameters on the maximum wave height from a tsunami propagation model. *Natural Hazards and Earth System Science*, **15**, 2299–2312, https://doi.org/10.5194/nhess-15-2299-2015

Carrier, G.F. & Greenspan, H.P. 1958. Water waves of finite amplitude on a sloping beach. *Journal of Fluid Mechanics*, **4**, 97–109.

Cummins, P.R. 2007. The potential for giant tsunamigenic earthquakes in the northern Bay of Bengal. *Nature*, **449**, 75–78, https://doi.org/10.1038/nature06088

Ekstrom, G., Nettles, M. & Dziewonski, A. 2012. The global CMT project 2004, 2010: centroid-moment tensors for 13 017 earthquakes. *Physics of the Earth and*

*Planetary Interiors*, **200**, 201, 1–9, https://doi.org/10.1016/j.pepi.2012.04.002

Field, E.H., Arrowsmith, R.J. *et al.* 2014. Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3). *Bulletin of the Seismological Society of America*, **104**, 1122–1180, https://doi.org/10.1785/0120130164

Fritz, H.M., Kalligeris, N., Borrero, J.C., Broncano, P. & Ortega, E. 2008. The 15 August 2007 Peru tsunami runup observations and modeling. *Geophysical Research Letters*, **35**, 1–5, https://doi.org/10.1029/2008GL033494

Garwood, F. 1936. Fiducial limits for the Poisson distribution. *Biometrika*, **28**, 437–442, https://doi.org/10.1093/biomet/28.3-4.437

Geist, E. & Bilek, S. 2001. Effect of depth-dependent shear modulus on tsunami generation along subduction zones. *Geophysical Research Letters*, **28**, 1315–1318.

Geist, E.L. & Dmowska, R. 1999. Local tsunamis and distributed slip at the source. *Pure and Applied Geophysics*, **154**, 485–512.

Geist, E. & Parsons, T. 2006. Probabilistic analysis of tsunami hazards. *Natural Hazards*, **37**, 277–314, https://doi.org/10.1007/s11069-005-4646-z

Geist, E.L., ten Brink, U.S. & Gove, M. 2014. A framework for the probabilistic analysis of meteotsunamis. *Natural Hazards*, **74**, 123–142, https://doi.org/10.1007/s11069-014-1294-1

Glimsdal, S., Pedersen, G., Harbitz, C. & Løvholt, F. 2013. Dispersion of tsunamis: does it really matter? *Natural Hazards and Earth System Science*, **13**, 1507–1526, https://doi.org/10.5194/nhess-13-1507-2013

Goda, K. 2015. Effects of seabed surface rupture versus buried rupture on tsunami wave modeling: a case study for the 2011 Tohoku, Japan, Earthquake. *Bulletin of the Seismological Society of America*, **105**, https://doi.org/10.1785/0120150091

Gonzalez, F.I., Geist, E.L. *et al.* 2009. Probabilistic tsunami hazard assessment at Seaside, Oregon, for near- and far-field seismic sources. *Journal of Geophysical Research*, **114**, 1–19, https://doi.org/10.1029/2008JC005132

Greenslade, D.J., Annunziato, A. *et al.* 2014. An assessment of the diversity in scenario-based tsunami forecasts for the Indian Ocean. *Continental Shelf Research*, **79**, 36–45, https://doi.org/10.1016/j.csr.2013.06.001

Hanks, T. & Kanamori, H. 1979. A moment magnitude scale. *Journal of Geophysical Research*, **84**, 2348–2350.

Harbitz, C., Glimsdal, S. *et al.* 2012. Tsunami hazard in the Caribbean: regional exposure derived from credible worst case scenarios. *Continental Shelf Research*, **38**, 1–23, https://doi.org/10.1016/j.csr.2012.02.006

Harbitz, C.B., Løvholt, F. & Bungum, H. 2014. Submarine landslide tsunamis: how extreme and how likely? *Natural Hazards*, **72**, 1341–1374, https://doi.org/10.1007/s11069-013-0681-3

Hayes, G.P., Wald, D.J. & Johnson, R.L. 2012. Slab1, 0: a three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research*, **117**, https://doi.org/10.1029/2011JB008524

Hébert, H. & Schindelé, F. 2015. Tsunami impact computed from offshore modeling and coastal amplification laws: insights from the 2004 Indian Ocean Tsunami. *Pure and Applied Geophysics*, **172**, 3385–3407, https://doi.org/10.1007/s00024-015-1136-4

Heidarzadeh, M. & Satake, K. 2015. Source properties of the 1998 July 17 Papua New Guinea tsunami based on tide gauge records. *Geophysical Journal International*, **202**, 361–369.

Holschneider, M., Zöller, G., Clements, R. & Schorlemmer, D. 2014. Can we test for the maximum possible earthquake magnitude? *Journal of Geophysical Research: Solid Earth*, **119**, 2019–2028, https://doi.org/10.1002/2013JB010319

Horspool, N., Pranantyo, I. *et al.* 2014. A probabilistic tsunami hazard assessment for Indonesia. *Natural Hazards and Earth System Science*, **14**, 3105–3122, https://doi.org/10.5194/nhessd-2-3423-2014

Kagan, Y.Y. 2002*a*. Seismic moment distribution revisited: 1. Statistical results. *Geophysical Journal International*, **148**, 520–541.

Kagan, Y.Y. 2002*b*. Seismic moment distribution revisited: 2. Moment conservation principle. *Geophysical Journal International*, **149**, 731–754.

Kagan, Y.Y. 2003. Accuracy of modern global earthquake catalogs. *Physics of the Earth and Planetary Interiors*, **135**, 173–209, https://doi.org/10.1016/S0031-9201(02)00214-5

Kagan, Y.Y. 2010. Earthquake size distribution: power-law with exponent beta = 1/2. *Tectonophysics*, **490**, 103–114.

Kagan, Y.Y. & Jackson, D.D. 2013. Tohoku earthquake: a surprise? *Bulletin of the Seismological Society of America*, **103**, 1181–1194, https://doi.org/10.1785/0120120110

Kajiura, K. 1963. The leading wave of a tsunami. *Bulletin of the Earthquake Research Institute*, **41**, 535–571.

Kamigaichi, O. 2009. Tsunami forecasting and warning. *In*: Meyers, R. (ed.) *Encyclopedia of Complexity and Systems Science*. Springer, Berlin, https://doi.org/10.1007/SpringerReference_60740

Kruschke, J.K. 2011. *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*. 2nd edn. Elsevier, Amsterdam.

Lorito, S., Tiberti, M.M., Basili, R., Piatanesi, A. & Valensise, G. 2008. Earthquake-generated tsunamis in the Mediterranean Sea: scenarios of potential threats to Southern Italy. *Journal of Geophysical Research*, **113**, 1–14, https://doi.org/10.1029/2007JB004943

Lorito, S., Selva, J., Basili, R., Romano, F., Tiberti, M. & Piatanesi, A. 2015. Probabilistic hazard for seismically induced tsunamis: accuracy and feasibility of inundation maps. *Geophysical Journal International*, **200**, 574–588, https://doi.org/10.1093/gji/ggu408

Lorito, S., Romano, F. & Lay, T. 2016. Tsunamigenic earthquakes (2004–13): source processes from data inversion. *In*: Meyers, R. (ed.) *Encyclopedia of Complexity and Systems Science*. Springer Science + Business Media, New York, https://doi.org/10.1007/978-3-642-27737-5_641-1

Løvholt, F., Glimsdal, S. *et al.* 2012*a*. Tsunami hazard and exposure on the global scale. *Earth-Science Reviews*, **110**, 58–73, https://doi.org/10.1016/j.earscirev.2011.10.002

Løvholt, F., Kuhn, D., Bungum, H., Harbitz, C. & Glimsdal, S. 2012*b*. Historical tsunamis and present tsunami hazard in eastern Indonesia and southern Philippines. *Journal of Geophysical Research*, **117**, https://doi.org/10.1029/2012JB009425

Løvholt, F., Glimsdal, S., Harbitz, C., Horspool, N., Smebye, H., de Bono, A. & Nadim, F. 2014*a*. Global tsunami hazard and exposure due to large co-seismic slip. *International Journal of Disaster Risk Reduction*, **10**, 406–418, https://doi.org/10.1016/j.ijdrr.2014.04.003

Løvholt, F., Setiadi, N.J. *et al*. 2014*b*. Tsunami risk reduction – are we better prepared today than in 2004? *International Journal of Disaster Risk Reduction*, **10**, 127–142, https://doi.org/10.1016/j.ijdrr.2014.07.008

Løvholt, F., Glimsdal, S., Smebye, H., Griffin, J. & Davies, G. 2015. *Tsunami Methodology and Result Overview*. Technical Report, UN-ISDR Global Assessment Report 2015 Geoscience Australia, Canberra.

Løvholt, F., Griffin, J. & Salgado-Gálvez, M. 2016. Tsunami hazard and risk assessment at a global scale. *In*: Meyers, R. (ed.) *Encyclopedia of Complexity and Systems Science*. Springer Science + Business Media, New York, https://doi.org/10.1007/978-3-642-27737-5_642-1

Lynett, P., Weiss, R., Renteria, W., Morales, G.D.L.T., Son, S., Arcos, M.E.M. & Macinnes, B.T. 2012. Coastal impacts of the March 11th Tohoku, Japan tsunami in the Galapagos Islands. *Pure and Applied Geophysics*, **170**, 1189, https://doi.org/10.1007/s00024-012-0568-3

Marzocchi, W. & Jordan, T.H. 2014. Testing for ontological errors in probabilistic forecasting models of natural systems. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 11 973–11 978, https://doi.org/10.1073/pnas.1410183111

Marzocchi, W., Taroni, M. & Selva, J. 2015. Accounting for epistemic uncertainty in PSHA: logic tree and ensemble modeling. *Bulletin of the Seismological Society of America*, **105**, 2151–2159, https://doi.org/10.1785/0120140131

Matias, L.M., Cunha, T., Annunziato, A., Baptista, M.A. & Carrilho, F. 2013. Tsunamigenic earthquakes in the Gulf of Cadiz: fault model and recurrence. *Natural Hazards and Earth System Science*, **13**, 1–13, https://doi.org/10.5194/nhess-13-1-2013

McCaffrey, R. 2008. Global frequency of magnitude 9 earthquakes. *Geology*, **36**, 263–266, https://doi.org/10.1130/G24402A.1

McCaffrey, R. 2009. The tectonic framework of the sumatran subduction zone. *Annual Review of Earth and Planetary Sciences*, **37**, 345–366, https://doi.org/10.1146/annurev.earth.031208.100212

McGuire, R.K., Cornell, C.A. & Toro, G.R. 2005. The case for using mean seismic hazard. *Earthquake Spectra*, **21**, 879–886, https://doi.org/10.1193/1.1985447

Newman, A., Hayes, G., Wei, Y. & Convers, J. 2011. The 25 October 2010 Mentawai tsunami earthquake, from real-time discriminants, finite-fault rupture, and tsunami excitation. *Geophysical Research Letters*, **38**, 1–7, https://doi.org/10.1029/2010GL046498

NGDC. 2015. *National Geophysical Data Center/World Data Service Global Historical Tsunami Database*. National Geophysical Data Center (NGDC), Boulder, CO, https://doi.org/10.7289/V5PN93H7; https://http://www.ngdc.noaa.gov/hazard/tsu_db.shtml [last accessed 24 September 2015].

Okada, Y. 1985. Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America*, **75**, 1135–1154.

Omira, R., Baptista, M.A. & Matias, L. 2015. Probabilistic tsunami hazard in the Northeast Atlantic from near and far field tectonic sources. *Pure and Applied Geophysics*, **172**, 901–920, https://doi.org/10.1007/s00024-014-0949-x

Parsons, T. & Geist, E.L. 2009. Tsunami probability in the Caribbean Region. *Pure and Applied Geophysics*, **165**, 2089–2116, https://doi.org/10.1007/s00024-008-0416-7

Power, W. 2013. *Review of Tsunami Hazard and Risk in New Zealand 2013 Update*. GNS Science Consultancy Report, 2013/131. GNS Science, Lower Hutt, New Zealand.

Saito, T. 2013. Dynamic tsunami generation due to seabottom deformation: analytical representation based on linear potential theory. *Earth, Planets and Space*, **65**, 1411–1423, https://doi.org/10.5047/eps.2013.07.004

Saito, T. & Furumura, T. 2009. Three-dimensional simulation of tsunami generation and propagation: Application to intraplate events. *Journal of Geophysical Research*, **114**, B02307, https://doi.org/10.1029/2007JB005523

Satake, K. 1995. Linear and nonlinear computations of the 1992 nicaragua earthquake tsunami. *Pure and Applied Geophysics*, **144**, 455–470.

Satake, K. 2014. Advances in earthquake and tsunami sciences and disaster risk reduction since the 2004 Indian ocean tsunami. *Geoscience Letters*, **1**, https://doi.org/10.1186/s40562-014-0015-7

Satake, K. & Atwater, B. 2007. Long-term perspectives on giant earthquakes and tsunamis at subduction zones. *Annual Reviews of Earth and Planetary Science*, **35**, 349–374, https://doi.org/10.1146/annurev.earth.35.031306.140302

Schorlemmer, D., Wiemer, S. & Wyss, M. 2005. Variations in earthquake-size distribution across different stress regimes. *Nature*, **437**, https://doi.org/10.1038/nature04094

Selva, J., Tonini, R. *et al*. 2016. Quantification of source uncertainties in Seismic Probabilistic Tsunami Hazard Analysis (SPTHA). *Geophysical Journal International*, **205**, 1780–1803, https://doi.org/10.1093/gji/ggw107

Socquet, A., Vigny, C., Chamot-Rooke, N., Simons, W., Rangin, C. & Ambrosius, B. 2006. India and Sunda plates motion and deformation along their boundary in Myanmar determined by GPS. *Journal of Geophysical Research*, **111**, https://doi.org/10.1029/2005JB003877

Sørensen, M.B., Spada, M., Babeyko, A., Wiemer, S. & Grünthal, G. 2012. Probabilistic tsunami hazard in the Mediterranean Sea. *Journal of Geophysical Research*, **117**, https://doi.org/10.1029/2010JB008169

Storchak, D., Giacomo, D.D. *et al.* 2012. *ISC-GEM Global Instrumental Earthquake Catalogue (1900–2009)*. GEM Technical Report 2012-01. GEM (Global Earthquake Model) Faulted Earth Project.

Strasser, F., Arango, M. & Bommer, J.J. 2010. Scaling of the source dimensions of interface and intraslab subduction-zone earthquakes with moment magnitude. *Seismological Research Letters*, **81**, 941–950, https://doi.org/10.1785/gssrl.81.6.941

Synolakis, C. 2011. Tsunamis: When will we learn? *Newsweek Magazine*, http://www.newsweek.com/tsunamis-when-will-we-learn-66213

Synolakis, C. & Kanoglu, U. 2015. The Fukushima accident was preventable. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **373**, 20140379, https://doi.org/10.1098/rsta.2014.0379

Tappin, D.R., Watts, P. & Grilli, S.T. 2008. The Papua New Guinea tsunami of 17 July, 1998: anatomy of a catastrophic event. *Natural Hazards and Earth System Sciences*, **8**, 243–266, https://doi.org/10.5194/nhess-8-243-2008

ten Brink, U., Chaytor, J., Geist, E.L., Brothers, D.S. & Andrews, B.D. 2014. Assessment of tsunami hazard to the U.S. Atlantic margin. *Marine Geology*, **353**, 31–54, https://doi.org/10.1016/j.margeo.2014.02.011

Thio, H. 2012. *URS Probabilistic Tsunami Hazard System: A User Manual*. Technical Report URS Corporation, San Francisco, CA.

Thio, H.K., Somerville, P. & Ichinose, G. 2007. Probabilistic analysis of strong ground motion and tsunami hazards in Southeast Asia. *Journal of Earthquake and Tsunami*, **1**, 119–137, https://doi.org/10.1142/S1793431107000080

Thio, H., Somerville, P. & Polet, J. 2010. *Probabilistic Tsunami Hazard in California*. Technical Report Pacific Earthquake Engineering Research Center, Berkeley, CA.

Tiberti, M.M., Lorito, S., Basili, R., Kastelic, V., Piatanesi, A. & Valensise, G. 2009. Scenarios of earthquake-generated tsunamis for the Italian coast of the Adriatic Sea. *Pure and Applied Geophysics*, **165**, 2117–2142, https://doi.org/10.1007/s00024-008-0417-6

Uchida, N. & Matsuzawa, T. 2011. Coupling coefficient, hierarchical structure, and earthquake cycle for the source area of the 2011 off the Pacific coast of Tohoku earthquake inferred from small repeating earthquake data. *Earth, Planets and Space*, **63**, 675–679, https://doi.org/10.5047/eps.2011.07.006

UN-ISDR 2015. *Global Assessment Report on Disaster Risk Reduction, Making Development Sustainable: The Future of Disaster Risk Management*. Technical Report United Nations Office for Disaster Risk Reduction, Geneva.

Valencia, N., Gardi, A., Gauraz, A., Leone, F. & Guillande, R. 2011. New tsunami damage functions developed in the framework of SCHEMA project: application to European–Mediterranean coasts. *Natural Hazards and Earth System Science*, **11**, 2835–2846, https://doi.org/10.5194/nhess-11-2835-2011

Wessel, P. 2009. Analysis of observed and predicted tsunami travel times for the Pacific and Indian Oceans. *Pure and Applied Geophysics*, **166**, 301–324, https://doi.org/10.1007/s00024-008-0437-2

Wesson, R.L., Boyd, O.S., Mueller, C.S., Bufe, C.G., Frankel, A.D. & Petersen, M.D. 2007. *Revision of Time-Independent Probabilistic Seismic Hazard Maps for Alaska*. United States Geological Survey, Open-File Report, **2007-1043**.

Youngs, R.R. & Coppersmith, K.J. 1985. Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bulletin of the Seismological Society of America*, **75**, 939–964.