

OBJECTIVE QUANTIFICATION OF THE SEISMIC SOURCE MODEL FOR NUCLEAR SITES

Ilaria MOSCA¹, Brian BAPTIE², Manuela VILLANI³, Ziggy LUBKOWSKI⁴ & Tim COURTNEY⁵

Abstract: *The UK is a relatively low seismicity region and therefore seismic design is applied only to high consequence facilities, e.g. nuclear power plants. Knowledge of the tectonic structures beneath the UK is limited and the earthquake catalogue covers a relatively short history compared to the geological history. In this context, the definition of the seismic source model and its uncertainties is critical. In the current practice of a probabilistic seismic hazard assessment (PSHA) in the nuclear industry, expert judgement is often adopted to determine some source model parameters, e.g. location of source zones boundaries or maximum magnitude associated with the sources, which introduces some degree of subjectivity. The development of a robust seismic source zone model for PSHA and a comprehensive characterization of its uncertainties will benefit both seismic hazard assessments and seismic design in the nuclear industry. We develop statistical tools for a fully non-linear method for characterising the source zone model used for PSHA using the Metropolis-Hastings Algorithm and Bayesian inference. This approach will allow us to: 1) sample the potential models compatible with the data; 2) capture multiple sources of uncertainties in the source zone model used for the PSHA; and 3) model the key components of the source zone model jointly rather than individually. We apply this approach to the Wylfa Newydd nuclear site (Anglesey), one of the proposed sites for new nuclear power plants in the UK, and benchmark our results using the PSHA developed for this site by Arup for Horizon Nuclear Power.*

Introduction

The UK is classified as an intraplate region and as such, it has low to moderate levels of seismicity. The largest observed earthquake (5.9 moment magnitude Mw) in the UK occurred in the Dogger Bank on 7 June 1931, and the largest onshore event (5.0 Mw) is the 19 July 1984 earthquake in the Lleyn peninsula, North Wales. The historical (before 1970) earthquake catalogue for the UK includes earthquakes of magnitudes greater than 6.0, such as the event on 11 September 1275, although the epicentre and the size of these events are highly uncertain (Musson, 2015). In this context, the maximum earthquake (i.e. the largest earthquake that a seismic source is capable of generating) for the UK is estimated to be between 6.5 and 7.1 Mw in the European hazard maps developed by the EU-FP7 “Seismic Hazard Harmonization in Europe” (SHARE) project (Woessner et al., 2015) and in the new UK national hazard map (Mosca et al., 2019). Such a large earthquake is associated with a recurrence time from hundreds to thousands of years. No British earthquake recorded either historically or instrumentally has produced a surface rupture and typical fault dimensions for the largest recorded British earthquakes are of the order of 1-2 km (Baptie, 2010). Therefore, it is difficult to associate earthquakes with specific faults, particularly at depth, where the fault distributions and orientations are unclear.

Due to the low levels of seismicity in the UK, only safety critical facilities, such as nuclear power plants, are designed to withstand earthquake loading. The license application to build and operate nuclear power plants requires demonstration of adequate protection of the public and the environment from internal and external hazards, including earthquakes. Natural hazards, including the seismic hazard, are to be considered conservatively for a probability of exceedance of 10^{-4} , which corresponds to a return period of 10,000 years (ONR, 2014). The

¹ Dr, British Geological Survey, Edinburgh, United Kingdom, imosca@nerc.ac.uk

² Dr, British Geological Survey, Edinburgh, United Kingdom

³ Associate, Arup, London, United Kingdom

⁴ Associate Director, Arup, London, United Kingdom

⁵ Formally Lead Engineer for External Hazards, Horizon Nuclear Power, Gloucester, United Kingdom

quantitative assessment for this long return period is challenging in the UK because the hazard for such a return period is usually dominated by large earthquakes that are poorly constrained by the relatively short duration of earthquake catalogues, compared to the geological history (at least hundreds of thousands of years). In this context, a strong hazard study that underpins the design ground motions for a nuclear power plants must include a robust quantification of the uncertainties in every stage of the seismic hazard assessment, as stated by the Office for Nuclear Regulations (ONR): “*For external hazards, the design basis event should be derived conservatively to take into account of data and model uncertainty...*” (ONR, 2014).

Probabilistic seismic hazard assessment (PSHA) is generally used for the seismic design of critical facilities in areas of low seismicity, including the UK. It uses the widest possible amount of data, combining seismological, geological and geophysical data to produce a probabilistic description of the distribution of future shaking that may occur at a site (e.g. Reiter, 1990; McGuire, 2004). The basic elements of PSHA are: the seismic source characterization (SSC) model and the ground motion characterization (GMC) model. The SSC model refers to the model for the spatial and temporal distribution of future earthquakes of different magnitudes in a specific region. The GMC model describes the value(s) of the ground motion parameter of interest at the site from all possible earthquake scenarios.

In this work, we focus on the SSC model since the choice of models has a strong impact on the hazard. Following the SSHAC guidelines (Budnitz et al., 1997; USNRC, 2012), in the current practice of PSHA in the nuclear industry, the SSC (and the GMC) model are usually expressed as logic trees that capture the centre, body and range of the technically defensible interpretations. The likelihood of fully capturing the uncertainty in key SSC models is achieved by including alternative models and parameter values in the logic tree where weights are assigned to each branch by expert judgements that reflect the relative confidence in the models and parameters (Coppersmith and Bommer, 2012). The judgement of experts introduces some degree of subjectivity, especially in regions of low to moderate seismicity, such as the UK where the data from geology, geophysics, tectonics and seismology can be interpreted in different ways. In addition, such an approach cannot fully represent the joint uncertainties and trade-offs present in the problem.

We propose a fully non-linear methodology for quantitative assessment of the SSC model based on the Bayesian statistical analysis to model jointly the key components of the SSC model and to capture their uncertainty. This approach provides an objective way to define the source model weighting used for the PSHA. Therefore, it can retrospectively test the SSC model used for the PSHA or it can be used to develop the SSC model for the PSHA.

Below, we describe the methodology developed in this work, which we refer to as the Bayesian Metropolis-Hastings (BMH) approach. Then, we illustrate it in the case study of the Wylfa Newydd nuclear site in Anglesey, UK.

Bayesian Metropolis-Hastings approach

The BMH approach consists of a Monte Carlo methodology combined with the Bayesian inference.

A Monte Carlo approach allows us to sample many potential models compatible with the data, together with a robust estimate of their uncertainties, and to account for the non-linear and complex relationship between the model and the data. Among the Monte Carlo techniques, we have chosen the Metropolis-Hastings Algorithm because it is less computationally expensive than other Monte Carlo methods and is relatively efficient in high-dimensional model space. The Metropolis-Hastings algorithm is a Markov Chain Monte Carlo algorithm to generate samples from an unknown target probability distribution for which direct sampling is difficult (e.g. Metropolis et al., 1953; Tarantola, 1987).

Bayesian inference is a pure probabilistic procedure for information updating from observed data and translates the set of models generated by the Metropolis-Hastings Algorithm into probability density functions (*pdfs*), which can be used to infer trade-off and dependencies between the model parameters. The solution is given by the posterior *pdf* that represents all information available on the model (Tarantola, 1987). Its calculation depends on the observed

data \mathbf{d} , which are represented by the likelihood function $L(\mathbf{m}|\mathbf{d})$ of a given model \mathbf{m} , and the prior information $\rho(\mathbf{m})$:

$$P(\mathbf{m}|\mathbf{d}) \propto \rho(\mathbf{m})L(\mathbf{m}|\mathbf{d}) \quad (1)$$

The likelihood function measures the likelihood of a given model through its misfit function χ^2 (Tarantola, 1987) :

$$L(\mathbf{m}|\mathbf{d}) = \text{const} \cdot \exp(-\chi^2) \quad (2)$$

The misfit function is a measure of the discrepancy between the observed and synthetic data.

In this work, the model is the source zone model, which is a numerical representation of all possible earthquake sources, describing where and how often earthquakes occur in terms of inter-event time and magnitude-frequency in a specific region. It is usually based on information from tectonics, geodesy, geology, and past seismicity and is described by a set of parameters, including the geometry of the source zones, lower and upper bounds of the magnitude range, recurrence parameters from the Gutenberg-Richter law, and hypocentral depth. The data are represented by the observed earthquake catalogue and prior information is given by geological constraints.

The methodology consists of six steps (Figure 1) that are described in details in the next section.

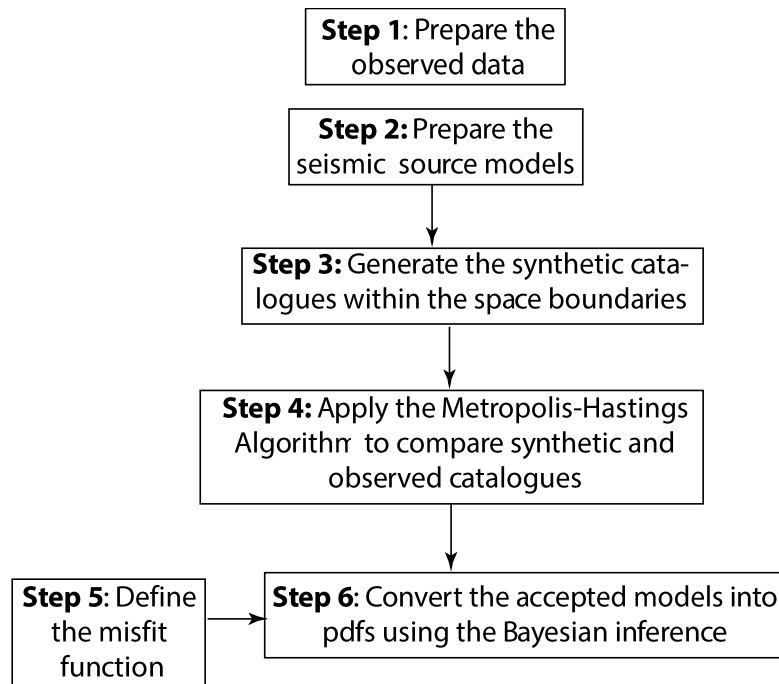


Figure 1. Workflow of the methodology.

Case study: Wylfa Newydd site

We apply the BMH approach to the Wylfa Newydd nuclear site whose PSHA was developed by Arup for Horizon Nuclear Power (Lubkowski et al., 2019). We describe the six steps of our approach for the Wylfa Newydd site in details.

Step 1: Prepare the observed data

The Wylfa Newydd site (53.411°N and 4.483°W) is situated in the Isle of Anglesey (Figure 2). The complete catalogue for the area within 300 km radius circle from the site consists of 219 mainshocks of 2.1 Mw and above (Figure 2).

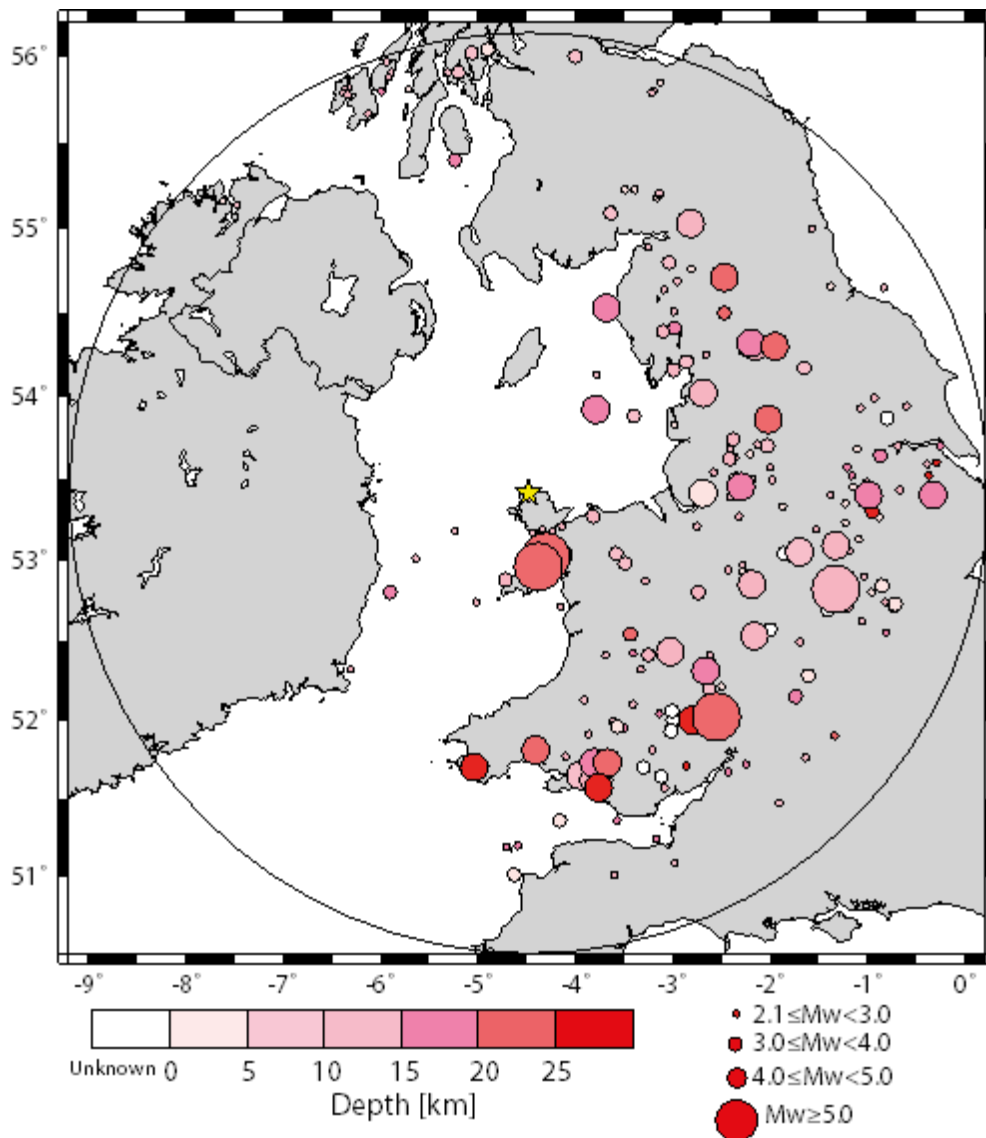


Figure 2. Distribution of the mainshocks in the study area within the completeness thresholds (Lubkowski *et al.*, 2019). Events with unknown depth are coloured in white. The black circle describe the area within 300 km radius, respectively. The yellow star describes the site of Wylfa Newydd.

Step 2: Prepare the seismic source models

We have selected 13 source zone models (SZMs) most of which were developed for regional or site-specific PSHA (Figure 3). SZM1 and SZM2 are based on the source model developed within the SHARE project by Woessner *et al.* (2015) and the model of Musson and Sargeant (2007) for the UK hazard maps, respectively. SZM7-SZM11 are the source models for the site-specific PSHA for Wylfa Newydd developed by Arup (Lubkowski *et al.*, 2019).

In order to accept a given source model as a candidate to generate synthetic catalogues, the seismicity in each source zone of that source model must be homogeneous, i.e. there must be an equal probability of occurrence of an earthquake at each point of the source zone.

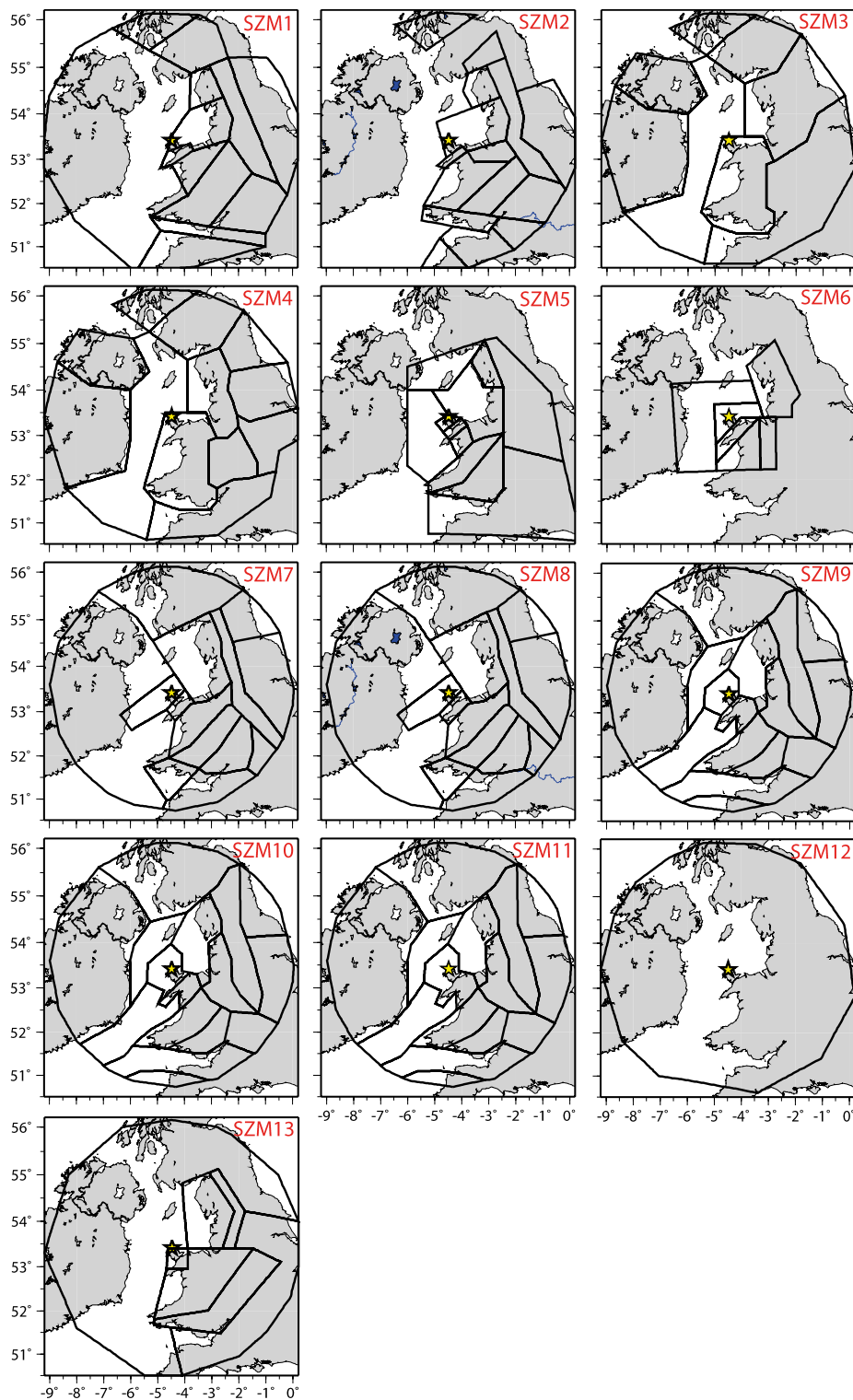


Figure 3. 13 candidate source zone models (SZMs) for the study area.

Step 3: Generate synthetic catalogues

Assuming that the occurrence of the synthetic earthquakes follows a Poisson process within each source zone, we generated 6,000,000 synthetic catalogues for the 13 candidate SZMs using Monte Carlo simulations. The synthetic catalogues are obtained by choosing the catalogue parameters, including maximum magnitude, the b -value, the activity rate, the earthquake magnitude, and the hypocentral depth, at random within their range.

Since the parameter ranges can have a strong influence on the final result, we set them within a uniform distribution so as not to strongly bias the search of the parameters in the model space

(Table 1). Furthermore, the range of the activity rate for the individual source zone of the source model depends on the number of source zones in the source model (Table 1).

| Parameter | Range |
|--|-----------|
| <i>b</i> -value | 0.7; 1.3 |
| $N(\geq 2.1/\text{yr})$ for a number of source zones > 8 | 0.0; 1.0 |
| $N(\geq 2.1/\text{yr})$ for a number of source zones between 4 and 8 | 0.5; 2.0 |
| $N(\geq 2.1/\text{yr})$ for a number of source zones ≤ 4 | 1.0; 8.0 |
| Mmax | 5.5; 7.5 |
| Mw | 2.1; Mmax |
| Hypocentral depth | 0; 30 km |

Table 1. Ranges of the catalogue parameters to generate the parameters of the synthetic catalogues.

Step 4: Apply the Metropolis-Hastings Algorithm

We apply the Metropolis-Hastings algorithm to compare the synthetic catalogues and the observed catalogue. The algorithm generates random walks sampling an initial, unknown target probability distribution. Only those walks fitting the target distribution are accepted. Then, the distribution of the accepted samples is re-sampled to have a closer fit to the target distribution. The initial target probability distribution is given by the probability distribution of the seismicity rate $N(M_w \geq 2.1/\text{yr})$ and the *b*-value computed for the entire study area. We generated 6,000,000 synthetic catalogues and only 217,603 of them fall within the observed *pdf* and therefore these ones tend to be accepted after applying the Metropolis-Hastings algorithm.

Step 5: Define the misfit function

We define the misfit function based on the number of earthquakes and the mean magnitude in a square grid:

$$\chi^2 = \frac{1}{N} \sum_i \frac{(C_i^{obs} - C_i^{th})^2}{\sigma(C^{obs})_i^2} + \frac{1}{N} \sum_i \frac{(\overline{Mw}_i^{obs} - \overline{Mw}_i^{th})^2}{\sigma(Mw^{obs})_i^2} \quad (3)$$

Where C_i^{obs} and C_i^{th} are the number of earthquakes in the *i*th grid cell for the observed and synthetic catalogue, respectively; $\sigma(C^{obs})_i$ is the standard deviation of C_i^{obs} ; \overline{Mw}_i^{obs} and \overline{Mw}_i^{th} are the mean magnitude in the *i*th grid cell for the observed and synthetic catalogue, respectively; $\sigma(Mw^{obs})_i$ is the standard deviation of the mean magnitude \overline{Mw}_i^{obs} ; and N is the number of grid squares.

We found that the optimal grid size is 0.1° by 0.1°.

Step 6: Apply the Bayesian inference

We apply the Bayesian inference to convert the set of synthetic catalogues accepted by the Metropolis-Hastings Algorithm into posterior *pdfs*, one for each parameter of the source model. Figure 4 shows the posterior *pdfs* for a set of model parameters, including the number of earthquakes, the source model and the Mmax. In the Bayesian framework, the mean is the most likely value, described by the solid lines in Figure 4, and the standard deviation is the width of the *pdf*.

Towards the full integration of the BMH approach in the PSHA

The posterior *pdfs* in Figure 4 can be used to rank the set of model parameters and to define the source model weighting in the logic tree. This provides high confidence in the logic tree for the source model because the centre, body, and range of the technically defensible interpretations for the SSC models are included in the *pdfs*.

The logic tree for the SSC model that can be inferred from the *pdf* of the SZMs in Figure 4 consists of six branches, i.e. SZM1, SZM4, SZM5, SZM6, SZM7, and SZM8, with weight 0.17, 0.18, 0.16, 0.17, 0.18, and 0.14, respectively. Our approach favours the source models based on the past seismicity and penalizes those based on tectonics because the acceptance and rejection of the synthetic catalogues by the Metropolis-Hastings Algorithm is driven by the observed earthquake catalogue. For this reason, the model SZM9-SZM11 which correspond to SM2a-SM2c in Lubkowski *et al.* (2019) are ranked less than SZM7-SZM8 (i.e. SM1a-SM1b in Lubkowski *et al.* (2019)).

We compared the hazard calculations computed using the source model logic tree inferred from Figure 4 with the PSHA developed for the Wylfa Newydd site by Arup where the source model weighting scheme consists of five branches, which here correspond to SZM7-SZM11, with weights of 0.25, 0.20, 0.25, 0.25, and 0.05, respectively (Lubkowski *et al.*, 2019). Figure 5 shows the hazard curves for the peak ground acceleration (PGA) computed for the two source model weighting schemes. We found that their agreement is excellent, suggesting that the seismic source model for the site-specific PSHA for the Wylfa Newydd site is not unique. This is because the length of the British earthquake catalogue is short and does not contain large magnitude (> 6.0 Mw) earthquakes.

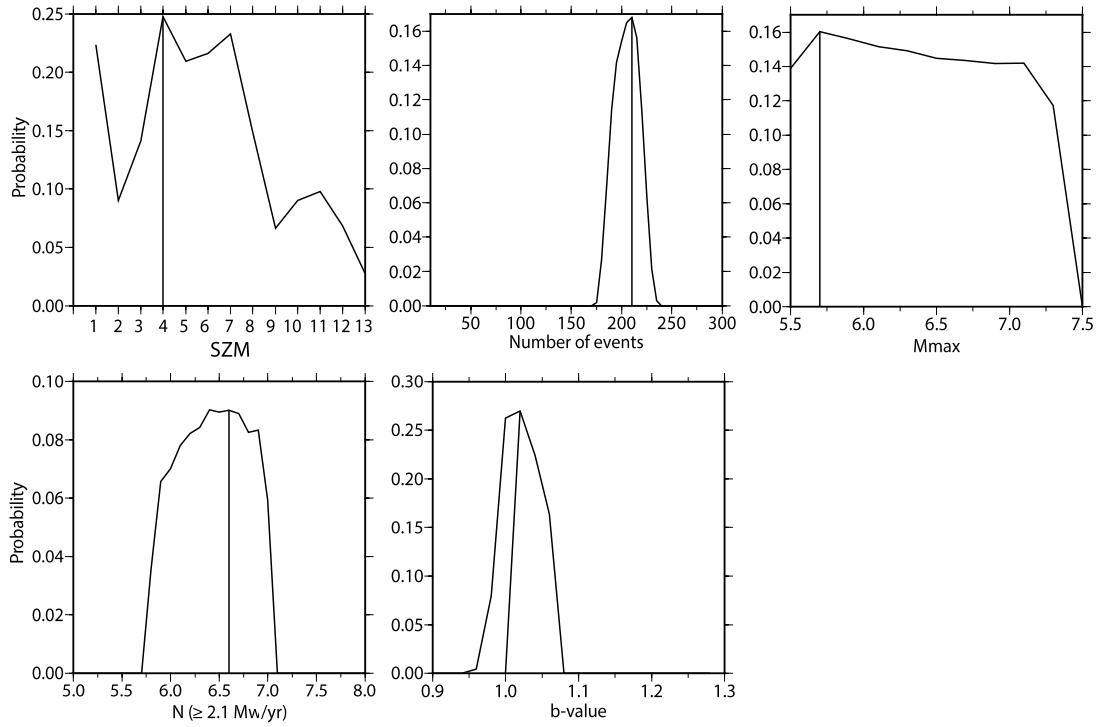


Figure 4. 1-D posterior probability density function for various model parameters. The solid line describes the most likely value in the distribution.

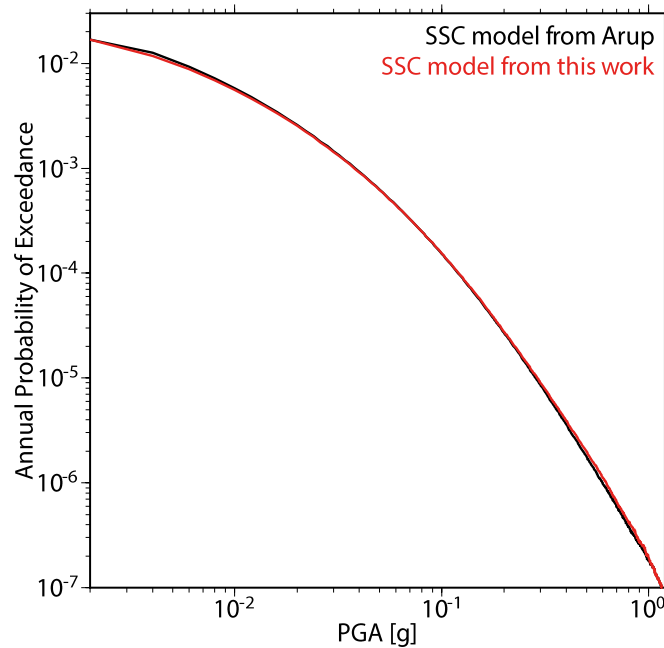


Figure 5. PGA hazard curve for the Wylfa Newydd site using the logic tree in Lubkowski *et al.* (2019) (black line) and the logic tree proposed by this study (red line).

References

- Baptie B (2010), Seismogenesis and state of stress in the UK, *Tectonophysics*, 482(1-4): 150-159.
- Budnitz RJ, G Apostolakis, DM Boore, LS Clu, KJ Coppersmith, CA Cornell, and PA Morris (1997), Recommendations for probabilistic seismic hazard analysis: Guidance on uncertainty and use of experts. *NUREG/CR-6372, Volume 1*, US Nuclear Regulatory Commission, Washington, DC.
- Coppersmith, KJ, and JJ Bommer (2012). Use of the SSHAC methodology within regulated environments: Cost-effective application for seismic characterization at multiple sites. *Nuclear Engineering and Design*, 245: 233-240.
- Lubkowski Z, Free M, Musson R, Villani M, Oakman C, Polidoro B, Koskosidi A, McCully R, Gilson B, Courtney TJ and Walsh M (2019), Seismic Hazard Assessment for the Wylfa Newydd Nuclear Power Plant, *SECED 2019 Conference*, Greenwich, United Kingdom.
- McGuire RK (2004), *Seismic Hazard and Risk Analysis*, Earthquake Engineering Research Institute, Oakland.
- Metropolis N, Rosenbluth AW, Teller MN, and Teller E (1953), Equation of state calculations fast computing machines, *Journal of Chemical Physics*, 1: 1087–1092.
- Mosca I, Sargeant S, Baptie B, Musson RWM, and Pharaoh T (2019), Revising the seismic hazard maps for the UK, *SECED 2019 Conference*, Greenwich, United Kingdom.
- Musson RWM and Sargeant S (2007), Eurocode 8 seismic hazard zoning maps for the UK. *BGS Report CR/07/125*, United Kingdom.
- Musson RWM (2015), What was the largest British earthquake?, *SECED 2015 Conference*, Cambridge, United Kingdom.
- ONR (2014). Safety Assessment Principles for nuclear facilities, Revision 0.
- Reiter L (1990), *Earthquake Hazard Analysis: Issues and Insights*, Columbia University Press, New York.
- Tarantola A (1987), *Inverse Problem Theory*, Elsevier, Amsterdam.
- USNRC (2012) Practical implementation guidelines for SSHAC level 3 and 4 hazard studies. *U.S. Nuclear Regulatory Commission, NUREG-2117, Rev. 1*, Washington, DC.
- Woessner J, Giardini D, and the SHARE Consortium (2012), Seismic hazard estimates for the Euro-Mediterranean region: A community-based probabilistic seismic hazard assessment. *Bulletin of Earthquake Engineering*, doi.org/10.1007/s10518-015-9795-1.