Bayesian validation of the seismic source models using the Wylfa Newydd site in the UK as a case study

Ilaria Mosca (✉ imosca@bgs.ac.uk)
British Geological Survey - Edinburgh Office
Brian Baptie
British Geological Survey - Edinburgh Office
Manuela Villani
Arup Group Ltd
Z. Lubkowski
Arup Group Ltd
T. Courtney
Formerly Horizon Nuclear Power

Research Article

Keywords: Earthquake catalogue, seismic source model, probability density function, PSHA, recurrence statistics, Bayesian inference

Posted Date: April 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-429836/v1

Bayesian validation of the seismic source models using the Wylfa Newydd site in the UK as a case study

I. Mosca, B. Baptie, M. Villani, Z. Lubkowski, and T. Courtney

Corresponding author: Ilaria Mosca, imosca@bgs.ac.uk, +447766443739

Affiliations:

Ilaria Mosca, Brian Baptie

British Geological Survey, The Lyell Centre, Research Avenue South, Edinburgh EH14 4AP, United Kingdom

Manuela Villani, Zygmunt Lubkowski

Ove Arup & Partners Int'l Ltd, 13 Fitzory St, W1T 4BQ, London, UK,

Tim J. Courtney

Formerly Horizon Nuclear Power, Downlea, Downfield, Stroud,

GL5 4HH, Gloucestershire, UK.

Abstract

In probabilistic seismic hazard assessment, the development of the seismic source characterization, especially the geometry of the seismic source models (SSMs), is controversial because it often relies on expert judgment with different interpretations of the available data from seismology, tectonics, and geology. Based on the same input datasets, different teams of experts may derive different SSMs. In this context, the verification of the models through the comparison against a set of observations is a crucial step. We present a statistical tool to compare the SSMs with the observed seismicity and rank these SSMs based on their ability to replicate the past seismicity. We simulate many synthetic catalogues derived from candidate SSMs and compare them with the observed catalogue of mainshocks using the Metropolis-Hastings Algorithm to select those that fit the observed catalogue. The candidate SSMs are then expressed by a probability density function (*pdf*) using the set of synthetic catalogues accepted by the Metropolis-Hastings Algorithm and the Bayesian inference. To help practitioners in earthquake and civil engineering understand how this tool works in practice, the proposed approach is applied to a proposed new nuclear site in the United Kingdom, Wylfa Newydd.

Keywords: Earthquake catalogue, seismic source model, probability density function, PSHA**,** recurrence statistics, Bayesian inference

1. Introduction

Probabilistic seismic hazard assessment (PSHA) forms the scientific basis for the characterization of the seismic input into seismic-resistant design codes and helps decision-makers in hazard mitigation. It aims to quantify the frequency of exceeding specific ground motion levels at a site for all possible earthquakes capable of producing damaging ground motions in a probabilistic framework accounting for the uncertainties in the input parameters (e.g. size and location of earthquakes and ground motion; Reiter 1990). The implementation of PSHA consists of two basic components: the seismic source characterization (SSC) and the ground motion characterization (GMC). The SSC describes the spatial and temporal distribution of expected earthquakes within a magnitude range in a specific region (e.g. USNRC 2012). The GMC informs us about the expected level of ground shaking for a specific site of interest considering soil conditions and building environments (e.g. Reiter 1990).

The data, upon which the SSC and GMC are defined, are often limited and therefore associated with large epistemic uncertainties. Modern PSHA studies incorporate the epistemic uncertainties in SSC and GMC using the logic tree formalism (Kulkarni et al. 1984; McGuire 2004) to capture the centre, body, and range of technically defensible interpretations (Budnitz et al. 1997; USNRC 2012). The centre of the distribution is the best estimate of the resulting interpretations, the body describes the shape of the distribution around the best estimate, and the range captures the tails of the distribution (USNRC 2012). The likelihood of fully capturing the epistemic uncertainty in the seismic hazard model is achieved by including alternative models and parameters in the logic tree where weights are assigned to each branch using expert judgment and elicitation that reflect the relative confidence in the models and parameters (Coppersmith and Bommer 2012). However, experts may interpret the observations differently and propose different weighting schemes for the input models. In this context, testing not only the hazard results but also each component of the seismic hazard model against observations is crucial to validate the model (e.g. Gerstenberger et al. 2020 and references therein). Although the role of testing and the use of Bayesian methods are increasing to capture the epistemic uncertainty in PSHA, most published studies aim to test either the GMC or the hazard estimates

against strong motion observations (e.g. Ordaz and Reyes 1999; Beauval et al. 2008; Albarello and D'Amico 2008; Stirling and Gersterberger 2010; Delavaud et al. 2012; Marzocchi and Jordan 2014), rather than the SSC (e.g. Weatherill and Burton 2009, 2010; and Musson and Winter 2012). For example, Delavaud et al. (2012) propose a strategy to combine a data-driven approach and expert judgments to build the logic tree for the GMC and to infer the range of uncertainties in the GMC. Albarello et al. (2008, 2015) use the empirical scoring test to compare the probabilistic outcomes for the Italian seismic hazard models and the empirical observations for a 25-year time window. Marzocchi and Jordan (2014, 2018) introduce the "experimental concept" to test seismic hazard estimates against a collection of observed and not yet observed, exchangeable data using error-statistical techniques, such as the p-value to quantify the statistical significance of evidence. In the last 15 years, a few studies propose to compare the SSC with the past seismicity assuming that the future seismicity follows the past. Weatherill and Burton (2009) use the K-means cluster analysis approach to define the geometry of the source model in the Aegean region based on the catalogue of recorded earthquakes. This approach is then used to attempt to construct the source model logic tree in PSHA for that region (Weatherill and Burton 2010). The approach of Musson and Winter (2012) compares synthetic catalogues generated from the SSMs with the observed catalogue in terms of the number of earthquakes and the mean magnitude of the catalogues. This approach is qualitative and does not quantify the performance of the SSM with respect to the past seismicity. Keller et al. (2014, 2019) adopt a Bayesian approach based on the Importance Sampling technique (e.g. Kass and Raftery 1995) to estimate the recurrence parameters, i.e. the seismic activity and the *b*-value, for different SSMs of France and compare the performance of the models in terms of hazard estimates. Unlike the validation of the GMC, there is no established procedure for the comparative evaluation of the SSMs with observations. Probably, this is because the length of the seismic records (up to hundreds of years) is often shorter than the average recurrence interval (from hundreds to thousands of years) of the largest earthquakes (moment magnitude Mw > 6.0) and thus the earthquake catalogue may be inadequate to represent the long-term seismicity (see Section 5).

In this work, we present and test a Bayesian approach, referred to as the Bayesian Metropolis-Hastings (BMH) approach, to validate the proposed seismic source models against the observed seismicity using statistical metrics and to rank which models better represent the past seismicity to forecast the future seismicity. The BMH approach explores the parameter space and generates synthetic catalogues that are compatible with the observed catalogue using the Metropolis-Hastings Algorithm. Then, the catalogue parameters are expressed in terms of probability density functions (*pdf*s) using the set of synthetic catalogues and the Bayesian inference. We test the BMH approach using the SSMs developed for the PSHA for a proposed nuclear site in Wylfa Newydd (North Wales, UK). The UK is an intraplate region with low levels of seismicity and the earthquake catalogue in this region is relatively short in duration (a few hundreds of years) and does not contain large (Mw > 6.0) earthquakes (e.g., Mosca et al., 2020). Facilities with such high consequence of failure require an assessment of hazard at very long return periods because seismic hazard estimates, in this case, are to be considered for low ($\leq 10^{-4}$) annual frequency of exceedance (e.g. USNRC 2012; ONR 2018) or otherwise stated for very long recurrence periods (> 10,000 years) when compared with the relatively short length of the seismic records.

2. Methodology

The SSC aims to describe the location, size, and frequency of future earthquakes through a set of parameters, such as the geometry of the seismic sources, maximum earthquake magnitude, and recurrence parameters drawn from the frequency-magnitude distribution (FMD) in each source (e.g. Budnitz et al. 1997). The study area is divided into a series of seismic sources (zones or faults). Seismic activity is considered to be spatially uniform within each seismic source zone, and earthquakes have an equal chance of occurring at any point in the source zone.

The BMH approach is a procedure to validate the seismic source models used in the SSC (Figure 1). This approach explores the full parameter space, which consists of three dimensions (i.e. the geometry of the source model, the activity rate, and the *b*-value), by generating synthetic catalogues that are

111 derived from candidate seismic source models using the Monte Carlo random sampling. Then, the
112 synthetic catalogues with a close fit to the declustered catalogue of mainshocks, are accepted by the
113 Metropolis-Hastings Algorithm and converted into probabilities using the Bayesian inference. The *pdf*
114 can be potentially used to adjust the weights in the source model logic tree that are defined by expert
115 judgments.

116 The BMH approach can be divided into six steps that are described in detail below and shown in
117 Figure 1.

118 **2.1. STEP 1: ASSEMBLING THE OBSERVED DATA**

119 The observed data used by the BMH consist of the declustered catalogue of mainshocks for the study
120 area within the completeness thresholds. It includes also the uncertainty in the epicentral location of
121 the mainshocks since this information is important in the delineation of the geometry of the SSMs.

122 The regional activity rate *a* and the *b*-value for the declustered catalogue of mainshocks are calculated
123 using the study area and the penalized maximum likelihood procedure (PMLP, Johnston et al. 1994).
124 This procedure uses the truncated Gutenberg-Richter recurrence law, different time windows of the
125 catalogue for different magnitude completeness thresholds, the correlation between *a* and *b*, and a
126 weighted prior to the *b*-value when the number of earthquakes is too small for a robust estimate of the
127 *b*-value. The results are expressed by a 5×5 matrix of possible values for *a* and *b*, *a* ± one and two
128 standard deviations, and *b* ± one and two standard deviations where the standard deviation of the
129 activity rate and *b*-value are computed from their covariance matrix (for more details, see Veneziano
130 and Van Dyke 1985; and Johnston et al. 1994). This determines 25 triplets of activity rate, *b*-value,
131 and weight. The regional estimate of the activity rate and the b-value is the most likely value in the 25
132 triplets, together with their standard deviation. The PMLP allows us to include a correction factor in
133 the activity rate calculations based on the standard error of individual earthquake magnitudes, as
134 proposed by Rhoades and Dowrick (2000).

2.2. STEP 2: SELECTION OF THE CANDIDATE SEISMIC SOURCE MODELS

Step 2 of the BMH approach defines the candidate seismic source models (SSMs) to generate synthetic catalogues. The spatial distribution of earthquakes within each source zone of the candidate source models is expected to have a uniform probability, i.e. earthquakes are equally likely anywhere in the zone. If this requirement, which can be tested using the nearest neighbour analysis (Davis 1986), is not satisfied, the source model cannot be considered as a potential candidate. Since this approach is an excellent tool to retrospectively test the SSC used in PSHA, one can consider only the source models in the SSC as candidates. Alternatively, it may be interesting to include all the published source models for the same site and potentially unpublished source models, as we do here (see Section 3).

2.3. STEP 3: GENERATION OF SYNTHETIC CATALOGUES

Assuming that the occurrence of the synthetic earthquakes follows a Poissonian process within each source zone, many synthetic catalogues are derived from the candidate source models using Monte Carlo simulations to have a stable convergence towards the final results (e.g. Ebel and Kafka 1999; Musson 2000; Assatourians and Atkinson 2013). Using the random Monte Carlo sampling, the synthetic catalogues are obtained by choosing M_{max} (i.e. the size of the largest possible earthquake in the region under investigation), the *b*-value, the activity rate, the year of occurrence, and the epicentre of the synthetic events at random within their range. We recommend fixing M_{max} to a single value corresponding to the largest maximum magnitude in the region. Although M_{max} could be also selected within a distribution, we have tested that the results will not change since M_{max} has little influence on the estimation of the FMD (Musson 2012). The sampling ranges of the recurrence parameters *a* and *b* are prior constraints on the Bayesian *pdf*s and can have a strong influence on the result. There is no standard procedure to choose the sampling range of these two parameters that must be tailored on the site under investigation by trial and error. The *a* and *b* parameters are chosen for the individual zones of the source model, and not on a regional basis. Then, the regional *a* and *b*-values for the synthetic catalogues are computed and compared with those from the observed catalogue. The sampling range

of the *b*-value can be selected within a unimodal probability distribution (e.g. Gaussian distribution) or a uniform distribution. The former is centred on the regional *b*-value within one or more standard deviations of the regional *b*-value. For the latter, the sampling range may correspond to the regional *b*-value within one or more standard deviations and all the values in the range have equal probability. Defining the sampling range for the activity rate is not straightforward and must be done by trial and error. The range of the activity rate can also be shaped within any probability distribution in principle. However, we recommend using a uniform distribution for the range of this parameter for the following reason. Unlike the *b*-value whose regional estimate represents a prior value for the individual zones, the regional activity rate is roughly the sum of the activity rates of the zones, and a prior value for *a* does not exist for each zone whose geometry is an unknown. Hence, it is not straightforward to centre the probability distribution for *a* on a specific value and a uniform distribution is a better choice not to favour or penalize some source models against others. In Section 4, we will show the results of the sensitivity analysis to set the ranges of the recurrence parameters using the testing ground of the Wylfa Newydd site.

The magnitude of the synthetic events is chosen based on the FMD that is defined by the synthetic values of M_{max}, *a* and *b.* The corresponding year of occurrence is selected within the time period from which the catalogue is assumed to be complete above the completeness magnitude (Mc) according to the completeness analysis of the earthquake catalogue. This ensures that the synthetic catalogues do not contain events of magnitude smaller than Mc. Finally, the epicentre of the synthetic earthquake is drawn within the zone using random sampling.

The procedure to generate a synthetic catalogue from a candidate SSM is the following. For each source zone of the source model, the activity rate and *b*-value are chosen within their sampling ranges using random sampling assuming that the activity rate and the *b*-value are independent. The number of synthetic events per year is computed using the Poisson distribution for each year of the catalogue's duration that is between the completeness period of the largest magnitude and the ending

of the observed catalogue. Then, Mw and the epicentral coordinates for each synthetic event are selected within their sampling range using random sampling.

2.4. STEP 4: IMPLEMENTING THE METROPOLIS-HASTINGS ALGORITHM

The Metropolis-Hastings Algorithm is a method to generate random walks (or candidate samples) from an unknown target probability distribution for which direct sampling is difficult (e.g. Metropolis et al. 1953; Tarantola 1987). For a random walk x_j, the arbitrary, conditional *pdf* $g(x_i|x_j)$ is defined for the transition $x_j \rightarrow x_i$, which depends on x_j only and not all the previous samples. The acceptance and rejection of the transitions are based on the acceptance ratio α (e.g. Mosegaard and Tarantola 2002):

$$\alpha = \min\left\{\frac{P(x_j)g(x_i|x_j)}{P(x_i)g(x_j|x_i)}, 1\right\} \quad (1)$$

where $P(x)$ is the target distribution. Assuming a uniform random number u between 0 and 1, if $u \leq \alpha$, the transition $x_j \rightarrow x_i$ is accepted; otherwise it is rejected. Then, the distribution of the accepted random walks is re-sampled to have a closer fit to the target distribution.

In the BMH approach, the target distribution is the regional estimate of the activity rate and *b*-value, together with their uncertainties, using the declustered catalogue of mainshocks and the PMLP of Johnston et al. (1994; see Subsection 2.1). The candidate samples are the regional estimates of the recurrence parameters of the synthetic catalogues for the study area computed using the PMLP. The Metropolis-Hastings Algorithm simultaneously fits the activity rate and the *b*-value in the synthetic catalogues with the target distribution assuming therefore a correlation between the recurrence parameters. Only the candidate samples that fall within the observed *pdf* are accepted by the Metropolis-Hastings Algorithm. This is described by Figure 2 where the dark and light grey histograms represent the target and synthetic distributions, respectively, before (left-hand column in Figure 2) and after applying the Metropolis-Hastings Algorithm (right-hand column in Figure 2).

Although the activity rate and the *b*-value are chosen within their sampling ranges for each source zone of the SSM, the comparison between the synthetic and observed catalogues is carried out in terms of the regional recurrence parameters for the entire study area, and not in terms of the recurrence parameters for the individual zones in the SSM. This is because there are no observations to compare the synthetic data with for each zone and the geometry of the SSM is one of the model parameters.

2.5. STEP 5: DEFINITION OF THE MISFIT FUNCTION

The misfit function (or cost function) χ^2 is a measure of the discrepancy between the observed and synthetic data. We define this function as the sum of two terms. The first term accounts for the location variability of the earthquakes in a grid cell and the second term is based on the mean magnitude in the grid cell:

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

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 cells. The standard deviation $\sigma\,(C^{obs})$ depends on the uncertainty in the epicentral location of the observed earthquakes and is defined as the square root of the absolute difference of the number of earthquakes with and without the epicentral uncertainty in the i-th grid cell:

$$\sigma(C_i^{obs}) = \sqrt{\left|\sum_{k=1}^{Neq} C_{k,i}^{obs} - \sum_{j=1}^{N'_{eq}} C_{j,i}^{obs}\right|}. \quad (3)$$

where N_{eq} is the number of mainshocks within the completeness thresholds; and $N'_{eq} = N_{eq} * 8$ is the number of mainshocks within the completeness thresholds accounting for the uncertainty in the

epicentral location. When the uncertainty in latitude and longitude for the epicentral location is taken into account, each earthquake in the observed catalogue can be located at eight different potential locations.

2.6. STEP 6: APPLYING THE BAYESIAN INFERENCE

The synthetic catalogues accepted by the Metropolis-Hastings Algorithm are converted in terms of posterior *pdf*s for the SSM using the Approximate Bayesian Computation (ABC), which is a class of computational methods in the Bayesian inference. In the Bayesian inference, the posterior *pdf* is proportional to the prior information $\rho(\boldsymbol{m})$ on the model $\boldsymbol{m}$ and the likelihood function $L(\boldsymbol{d}|\boldsymbol{m})$ of the model $\boldsymbol{m}$ on the observed data $\boldsymbol{d}$:

$$P(\boldsymbol{m}|\boldsymbol{d}) \propto \rho(\boldsymbol{m})L(\boldsymbol{d}|\boldsymbol{m}) \quad (4)$$

The ABC approximates the likelihood function by simulations and compares the posterior *pdf* of the model parameters to the observed data through its misfit function χ^2 (Tarantola 1987):

$$L(\boldsymbol{d}|\boldsymbol{m}) \propto e^{-\chi^2} \quad (5)$$

3. Applying the BMH approach to the Wylfa Newydd case

We applied the BMH approach to retrospectively test the SSC used for the site-specific PSHA of the Wylfa Newydd site, UK (53.411°N and 4.483°W; Figure 3). The PSHA for the nuclear license application of this site was developed by Arup on behalf of the operator Horizon Nuclear Power (Villani et al. 2020). In this work, we used the earthquake catalogue, the declustering, and completeness analysis as provided by Villani et al. (2020). We did not make a new analysis to decluster the earthquake catalogue and assess its completeness because this is beyond the purpose of

this paper that aims to present not a site-specific PSHA but a new tool to test which candidate source models are most appropriate to represent the future seismicity.

3.1. STEP 1

This subsection describes the data used by the BMH approach for our testing ground. We used the declustered catalogue of mainshocks that starts in July 1534 and ends on 31 December 2014. The distribution of mainshocks with magnitudes above Mw 2.1 consists of 628 earthquakes within the 300-km region from the site and the completeness thresholds in Table 1 (Figure 3). The largest earthquakes recorded in the study area occurred in 1852, 1896, 1957, and 1984 with Mw 5.0. The catalogue spans 364 years from 1650, since which the observed catalogue is assessed to be complete for events of 6.5 Mw and above, to 2014. This highlights the limitations of the earthquake catalogue that is relatively short in duration (a few hundreds of years) and does not contain large (Mw > 6.0) earthquakes.

The uncertainty in the epicentral location for historical events is defined based on four classes (Table 2) as described in Musson (1994). For each class, we define a reference value (Table 2). The location uncertainty of instrumental events has been estimated by the British Geological Survey using the HYPO71 location algorithm (Lee and Lahr 1975) since the Nineties and the HYPOINVERSE-2000 program (Klein 2002, 2003) for earthquakes after 2000. Since the location uncertainties of the historical and instrumental UK catalogues are not homogeneously and comprehensively assessed, we decided to assume a class A (i.e. 5 km) for the location uncertainty in the instrumental events.

To determine the regional recurrence parameters *a* and *b* for the seismicity of the study area, we applied the PMLP as described earlier. The minimum and maximum magnitude for the recurrence calculations were 2.1 Mw and 7.1 Mw, respectively. 7.1 Mw is the largest magnitude in the M_{max} distribution in the SSC developed by Villani et al. (2020). We did not apply the correction factor in the activity rate calculations because the magnitude uncertainty of the earthquakes is not assessed homogeneously in the catalogue for the UK (e.g. Mosca et al. 2020). EPRI (2012) and Musson (2012) recognize that care should be taken when the magnitude uncertainties are accounted for in the

recurrence statistics, especially when an earthquake catalogue contains more than one original magnitude scale to avoid counter-balancing bias in the estimation of the seismicity rates. Using the completeness thresholds in Table 1, we found that the best- fit values are N (Mw ≥ 2.1/yr) = 7.12 ± 0.51 and b =1.025 ± 0.041. Note that we provided no prior b-value for the regional estimate of this parameter.

3.2. STEP 2

We selected 14 different SSMs as potential candidates, most of which were developed for other regional PSHA for the UK or site-specific PSHA for Wylfa Newydd (Figure 4 and Table 3). SSM3 and SSM4 are based on political and administration criteria, respectively. These two SSMs were included to check how source models based on non-scientific criteria compare against source models based on seismic and tectonic information. We added a background zone for SSM2, SSM5, and SSM6 to cover the entire study area and avoid any potential bias in the simulation of the synthetic catalogues from source models that cover the entire study area.

To define the sampling range of the recurrence parameters used to generate the synthetic catalogues from the 14 SSMs, we carried out preliminary recurrence calculations for the individual zones of the 14 SSMs using the PMLP, the earthquake catalogue in Figure 3, together with its completeness analysis (Table 1), and the best regional estimate of b = 1.025 as a weighted prior for each of the zones. We found that overall the b-value varies between 0.65 and 1.43, and N (≥ 2.1Mw/yr) is between 0.01 and 5.00 considering all source zones in the 14 SSMs (see the last two columns in Table 3).

3.3. STEPS 3-4

Table 4 shows the sampling ranges of the catalogue parameters used for the simulations of the synthetic catalogues. Mmax is fixed to 7.1 Mw that is the largest maximum magnitude in the region (see Subsection 3.1). The range of the b-value is chosen to be normally distributed around the regional estimate within one standard deviation (Table 4). To set the range of the activity rate, we estimated

this parameter for each SSM as the fraction of the regional activity rate (i.e. 7.12) within four standard deviations divided by the number of zones in the SSM because the regional activity rate is roughly the sum of the activity rates of the individual source zones. The four standard deviations account for the large variability in this parameter among the zones of the SSM. Then, we grouped the ranges into four categories depending on the number of zones (see Table 4). Note that we cannot set a unimodal distribution for the sampling range of the activity rate because this parameter changes significantly from zone to zone within the same source model (see Table 3). This implies that no value of the activity rate, around which to centre the unimodal distribution, can be chosen. By trial and error, we have found that the ranges in Table 4 are the best choice to have a reasonably large range of activity rates for the synthetic events but to avoid unrealistic values that would be rejected by the Metropolis-Hastings Algorithm. Section 4 describes the influence of the range of the recurrence parameters on the results computed by the BMH approach.

We generated 300,000 synthetic catalogues for each of the 14 SSMs. This means having a total number of 4,200,000 synthetic catalogues. We tested by trial and error that 4,200,000 simulations provide a clear convergence towards stable results. The Metropolis-Hastings Algorithm accepted 1,052,929 out of 4,200,000 synthetic catalogues.

3.4. STEPS 5-6

To find the optimal grid size for the misfit function, we tested different cell sizes and plotted the accepted catalogues by the Metropolis-Hastings Algorithm in terms of the misfit function (Figure 5). We have chosen the 0.1° by 0.1° grid size because it realistically simulates the sparse seismicity in the UK.

We applied Equation 4 to convert the 1,052,929 synthetic catalogues accepted by the Metropolis-Hastings Algorithm into a 1-D posterior *pdf* for the SSM (Figure 6). The *pdf* is multimodal suggesting that more than one model fits the data equally well. The smallest probability is associated with SSM12, which consists of a single zone, and thus does not fit the catalogue of mainshocks. Also, the performance of SSM3-SSM5 is relatively poor, which gives some confidence that the method can

identify unrealistic models. Specifically, the low probability of SSM3 and SSM4 is because these models are based on non-scientific criteria and therefore are not able to replicate the past seismicity. SSM5, and to less extent SSM6, are penalized by the background zone that was added to cover the same study area as for the other SSMs (see Subsection 3.2). Although SSM2 has also a background zone, this is small and does not influence the performance of SSM2 against the observed catalogue. At the time of the publication of SSM5 and SSM6, the principal study area was considered to be a 100 km radius circle around this site to build the source model for a site-specific PSHA. Nowadays, the recommendation from IAEA is to extend the source model to 300-km from the site (e.g. IAEA 2010). SSM13 has also a relatively low probability in the posterior *pdf* probably because it was published more than 20 years ago and does not reflect the updates in the historical and instrumental seismicity in the British Isles. The other SSMs have a probability included between 0.80 and 0.92.

We analysed more carefully the *pdf* for the SSM by plotting the number of accepted catalogues in terms of SSM and the misfit function (Figure 7). There is a generally good correlation between the accepted synthetic catalogues and the misfit function for individual SSMs, i.e. the higher the number of accepted catalogues, the smaller the misfit function. An exception to this is represented by the results of SSM12. Although the Metropolis-Hastings Algorithm accepted a few catalogues for this source model, the range of the misfit function is relatively low.

4. Sensitivity analysis on the posterior *pdf*

In this section, we describe the impact of the prior distribution of the recurrence parameters, earthquake catalogue, and the size of the study area on the 1-D posterior *pdf* for the SSM. This sensitivity analysis allows us to determine whether the results from the BMH approach are strongly dependent on the parameter ranges, the earthquake catalogue, and the area under investigation. We compared the 1-D *pdfs* for the SSM obtained from the sensitivity analysis with that shown in Figure 6 (referred to as “Reference” in Figures 8 and 10).

352 **4.1. INFLUENCE OF THE SAMPLING RANGE OF THE RECURRENCE PARAMETERS**

353 To carefully assess the impact of the sampling range of the recurrence parameters, together with its
354 prior distribution, on the *pdf* for the SSM and how to set these ranges, we carried out a sensitivity
355 analysis by running the BMH approach for different combinations of sampling range for these
356 parameters (see Table 5). We considered three different ranges of the activity rate and *b*-values. In
357 Test 1, the sampling range of the recurrence parameters is chosen to be the regional values within two
358 standard deviations using a uniform distribution. In Test 2, the range for the *b*-value is that in
359 Reference (i.e. the sampling ranges in Table 4); whereas, the range for the activity rate is broader and
360 grouped into three categories, rather than four. In Test 3, the distribution of the range for the *b*-value
361 is set within a uniform distribution centred on the regional estimate within four standard deviations;
362 and the sampling ranges of the activity rate is that in "Reference".

363 The main observations from this analysis are the following (see Figure 8):

364 • The sampling range of the recurrence parameters, especially the activity rate, affects significantly
365 the *pdf* for the SSM, as shown by the number of accepted catalogues.
366 • The larger the parameter range, the lower the number of accepted catalogues from the Metropolis-
367 Hastings algorithm.
368 • The *pdf* for Test 1 is almost uniform between SSM1-SSM11 suggesting that the algorithm cannot
369 tell which source models produce synthetic catalogues that fit better the observations. This is
370 because the narrow range of the recurrence parameters excludes the extreme cases, i.e. where the
371 zones have a few or many earthquakes.
372 • In Test 2 where the range of the activity rate is large, the search tends to favour some models and
373 penalize others. For example, the range of 0.0-1.0 for the activity rate and the source model with
374 more than 10 zones penalizes SSM2 and SSM9-SSM11 and favours SSM4-SSM5 and SSM7-
375 SSM8. This is because the activity rate of the individual zones tends to be chosen such that the
376 regional synthetic activity rate for the entire study area is larger than the regional observed value.

- The result from Test 3 confirms that the influence of the *b*-value on the 1-D *pdf* is less strong than the activity rate and the results are less sensitive to the choice of the prior distribution for the *b*-value. If the uniform distribution is chosen rather than the Gaussian distribution, the number of synthetic catalogues that fit the observed catalogue of mainshocks decreases.

4.2. INFLUENCE OF THE EARTHQUAKE CATALOGUE AND STUDY AREA

We carried out two tests to check the influence of the earthquake catalogue and the size of the study area on the *pdf* for the SSM and the consistency and repeatability of the results computed by the BMH approach. In Test A, we used the instrumental catalogue for the period 1970-2014 that consists of 172 mainshocks of ≥ Mw 2.1 within 300 km from the site using the completeness thresholds in Table 1. We estimated the seismicity rate and the *b*-value for the instrumental catalogue using the PMLP. The target probability distribution for the instrumental catalogue is described by N (Mw ≥ 2.1/yr) = 7.33 ± 0.55 and b=1.050 ± 0.067. In Test B, we focused on the region within 60-km of the site because it has the strongest impact on the estimated seismic hazard (Villani et al. 2020). The entire catalogue includes 17 mainshocks of Mw ≥ 2.1 using the completeness thresholds in Table 1. The activity rate and the *b*-value for this test are N (Mw ≥ 2.1/yr) = 0.49 ±0.12 and b=0.88±0.11. We used only 13 of the SSMs because SSM3 and SSM4 are identical for this study area (Figure 9).

The results from these tests are shown in Figure 10. The trend of the *pdfs* for the SSM for Reference and Test A is very similar. Test B provides relatively similar results to the reference values except for SSM5 that gets the highest probability in the *pdf* for the SSM. The performance of this model is significantly better in Test B than in Reference because the large background zone of SSM5 is excluded from the 60-km study area (see Figure 4). A similar effect can be seen also for the large probability of SSM6; whereas, the small background zone of SSM2 does not influence the results by changing the size of the study area. As explained in Sections 3.2 and 3.4 of the Wylfa Newydd case study, the background zone was added to SSM2, SSM5, and SSM6 to cover the same study area.

5. Discussion

The goal of this article was to present a statistical tool for the objective assessment of the quality of the SSMs in the SSC against the observations using the Bayesian inference. The output of the BMH approach is expressed in terms of posterior *pdfs* for the SSMs and can be used to quantitatively compare various models with the observed seismicity in a transparent and reproducible way.

The application of the BMH approach to the Wylfa Newydd case study helps understand how to use this tool in practice. Although we provide the guidelines to apply the BMH approach, its implementation requires setting up some input elements that are specific for the site under investigation, such as the selection of the candidate SSMs, the range of the catalogue parameters to generate the synthetic catalogues and the optimal grid size to compute the misfit function. The sensitivity analysis for the Wylfa Newydd case study illustrates the influence of the input parameters on the posterior *pdf* for the SSM. The sampling ranges of the recurrence parameters must be tailored to the site under investigation by trial and error. If this is not carried out well, there is the risk to favour some models and penalize others by the Metropolis-Hastings Algorithm. The correct setting of the ranges for *a* and *b* parameters limits also the tendency of the BMH approach to favour SSMs based on past seismicity and to penalize those based on the regional geology and tectonics. This is a consequence of the fact that the Metropolis-Hastings Algorithm accepted and rejects SSMs based on the ability of the models to replicate the observed seismicity.

The retrospective validation of the SSMs developed for the PSHA of the Wylfa Newydd site shows that the source models based on political and administration criteria and the single-zone source model are not able to reproduce the past seismicity (Figure 6) suggesting that the BMH approach can identify unrealistic source models. The results also show that many source models can replicate the observed catalogues equally well (Figure 6) implying that the seismic source model is not uniquely constrained by the sparse seismic observations in the UK. The application to Wylfa Newydd, although illustrative, highlights the limitation of the tool, when the observed earthquake catalogue is short

compared to the recurrence interval of large earthquakes. The short duration of the British catalogue and the lack of the occurrence of large historical earthquakes suggest that the maximum expected earthquake is probably not included in the catalogue of past seismicity. Furthermore, the correlation between earthquakes and tectonic structures is often difficult to ascertain since there is no evidence of surface ruptures produced by known earthquakes (Baptie 2010; Mosca et al., 2020). In this application, we did not investigate the influence of the completeness thresholds for the earthquake catalogue on the posterior *pdf* for the SSM. Although the assessment of the completeness thresholds is important for the seismic hazard analysis, the scope of the paper is not to derive a PSHA at the site. However, since the completeness magnitude (Mc) values from the completeness analysis have a strong influence on the recurrence parameters, especially the *b*-value, it is likely that a change in the Mc values may result in different posterior *pdf* for the SSM.

As mentioned above, a limitation of this methodology is that it relies on the observed records that do not often include the average recurrence intervals of the potentially largest earthquakes in the region under investigation. This problem is particularly important in low strain continental environments where the rate of deformation is low and the recurrence interval of large earthquakes is of the order of several hundred to thousands of years, therefore, exceeding the relatively short sampling of an earthquake catalogue based on historical data (e.g. Stein et al 2015; Liu and Stein 2016). This implies that the occurrence of future earthquakes in regions devoid of recorded seismic activity cannot often be ruled out. Examples of events in regions with no history of seismic activity are the 1886 Charleston earthquake of magnitude between 6.9 and 7.3 Mw (Bollinger 1972), the 2017 Mw 6.5 Botswana earthquake (Gardonio et al. 2018), and the 2006 Mw 7.0 Machaze earthquake (Mozambique; Fenton and Bommer 2006). The development of the SSC must satisfy the following two criteria. Firstly, the construction of the SSC should be based on data from seismology, tectonics, geology, geophysics, and paleoseismology to account for the short length of the seismic observations compared to the seismic cycle. The use of source zones, where the seismicity is assumed to be uniform across it, allows us to reduce the likelihood of neglecting important but unknown tectonic structures that are not been associated with recorded seismicity but may generate future earthquakes. Secondly, the SSC should

reproduce well the past seismicity. In intraplate regions with low levels of seismicity and low deformation rates, such as most of the Eastern United States, Canada, and the UK, the observed catalogue is the only option to extrapolate the rates of large, rare earthquakes from the occurrence rates of small-to-moderate earthquakes (Budnitz et al. 1997; Anderson and Biasi 2016). In this context, it seems reasonable to develop a robust methodology to statistically validate the ability of the source models in the SSC to replicate the observations and therefore fully satisfy the second criterion in the development of the SSC.

6. Conclusions

This work highlights the importance of model testing and validation against the observations. The main finding of the paper is the presentation, together with its first application, of the Bayesian Metropolis-Hastings (BMH) approach to test to what degree the seismic records support or reject the proposed seismic source models in the SSC in light of the observed seismicity, i.e. one of two criteria to construct realistic source models. This is a valuable tool in regions of high seismicity with a long catalogue of seismic historical records but also in regions of moderate seismicity where the processes controlling the occurrence of earthquakes are not fully understood, the rate of deformation is low, and therefore the occurrence rate of large earthquakes is extrapolated from low-to-moderate events.

The BMH approach is applied for the evaluation of the seismic source models for the PSHA at the Wylfa Newydd site in the UK. Due to the poor correlation between British earthquakes and tectonic structures in the UK, the SSMs for the Wylfa Newydd case study did not include fault sources. Furthermore, the recurrence calculations did not include the magnitude uncertainty of the earthquakes in the catalogue because this uncertainty is not homogeneously and uniformly assessed throughout the UK catalogue (e.g. Mosca et al 2020). However, both these two features (i.e. fault sources in the source model and uncertainty in the magnitude) can be easily included in the BMH approach.

476 Future research for the SSC should focus on rigorous testing protocols for the SSC to understand
477 whether the SSMs produce forecasts that are consistent with the observations, similar to what has
478 been done for the GMC in the last decade (e.g. Delavaud et al. 2012; Douglas and Edwards 2016).
479 The use of a data-driven approach can be combined with expert judgment to adjust the weights in the
480 source model logic tree and therefore infer robust epistemic uncertainties in the SSC.

481 7. Declarations

482 This work was funded by the Innovation Flexible Funding Programme of the British Geological
483 Survey. The authors declare that they have no competing interest. The FORTRAN code for the
484 Bayesian Metropolis-Hastings approach is available upon request. The declustered catalogue of
485 mainshocks, the completeness thresholds, the seismic source models SSM7-SSM11 for the site-
486 specific PSHA of the Wylfa Newydd case study are from Villani et al (2020) listed in the references.
487 The other seismic source models used in this paper came also from published sources listed in the
488 references. The plots were made using the Generic Mapping Tools version 4.5.2
489 (www.soest.hawaii.edu/gmt).

490 8. Acknowledgements

491 We thank Horizon Nuclear Power and Arup for making the data and the reports of the seismic hazard
492 assessment for the Wylfa Newydd site available for this work. We are grateful to Margarita Segou and
493 Susanne Sargeant whose insightful comments helped improve the manuscript. We thank Merlin
494 Keller and collaborators for fruitful discussions on Monte Carlo methods and Bayesian inference.

9. References

Albarello D, D'Amico V (2008). Testing probabilistic seismic hazard estimates by comparison with observations: An example in Italy. Geophys J Int 175:1088–1094.

Albarello D, Peruzza L, Amico V (2015). A scoring test on probabilistic seismic hazard estimates in Italy. Nat Hazard Earth Sys 15:171–186.

Amante C, Eakins B (2009). ETOPO1 1Arc-Minute Global Relief Model: Procedures, data resources and analysis. TNOAA Technical Memorandum NESDIS NGDC-24.

Anderson JG, Biasi GP (2016). What is the basic assumption for probabilistic seismic hazard assessment? Seismol Res Lett 87(2A):323–326.

Assatourians K, Atkinson GM (2013). EqHaz - An open-source probabilistic seismic hazard code based on the Monte Carlo simulation approach. Seismol Res Lett 84(3):516-524.

Baptie B (2010). Seismogenesis and state of stress in the UK. Tectonophysics 482(1-4):150-159.

Beauval C, Bard P-Y, Hainzl S, Guéguen P (2008). Can strong-motion observations be used to constrain probabilistic seismic-hazard estimates? Bull Seismol Soc Am 98(2):509–520.

Bollinger GA (1972). Historical and recent seismic activity in South Carolina". Bull Seismol Soc Am 62(3):851–864.

Budnitz RJ, Apostolakis G, Boore DM, Cluff LS, Coppersmith KJ, Cornell CA, Morris PA (1997). Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on the Uncertainty and Use of Experts. Senior Seismic Hazard Analysis Committee, U.S. Nuclear Regulatory Commission, NUREG/CR-6372, Washington, DC.

Coppersmith K, Bommer J (2012). Use of the SSHAC methodology within regulated environments: cost-effective application for seismic characterization at multiple sites. Nucl Eng Des 245:233–240.

Davis JC (1986). Statistics and data analysis in geology, Wiley, New York.

Delavaud E, Scherbaum F, Kühn N, Allen T (2012). Testing the global applicability of ground motion prediction equations for active shallow crustal regions. Bull Seismol Soc Am 102(2):707-721.

Douglas J, Edwards B (2016). Recent and future developments in earthquake ground motion estimation. Earth-Sci Rev 160:203-219.

Ebel JE, Kafka AL (1999). A Monte Carlo approach to seismic hazard analysis. Bull Seismol Soc Am 89:854–866.

Fenton C, Bommer JJ (2006). The Mw 7 Machaze, Mozambique, Earthquake of 23 February 2006. Geophys Res Lett 77:426-439.

Gardonio B, Jolivet R, Calais E, Leclère H (2018). The April 2017 Mw 6.5 Botswana Earthquake: An Intraplate Event Triggered by Deep Fluids. Geophys Res Lett 45(17):8886-8896

Gerstenberger MC, Marzocchi W, Allen T, Pagani M, Adams J, Danciu L, Field EH, Fujiwara H, Luco N, Ma K-F, Meletti C, Petersen MD (2020). Probabilistic Seismic Hazard Analysis at Regional and National Scales: State of the Art and Future Challenges. Rev Geophys 58(e2019RG000653) https://doi.org/10.1029/2019RG000653.

Grünthal G, Bosse C, Sellami S, Mayer-Rosa D, Giardini D (1999). Compilation of the GSHAP regional seismic hazard for Europe, Africa, and the Middle East. Ann Geof 42(6):1215-1223.

EPRI (2012). Technical report: Central and Eastern United States Seismic Source Characterization for Nuclear Facilities. Electric Power Research Institute, Palo Alto, U.S. Nuclear Regulatory Commission, U.S. Department of Energy.

IAEA (2010). Seismic hazards in site evaluation for nuclear installations. IAEA specific safety guide series no. SSG-9, IAEA, Vienna.

Johnston AC, Coppersmith KJ, Kanter LR, Cornell CA (1994). The earthquakes of stable continental regions. EPRI Report, Electric Power Research Institute, Palo Alto.

Kass RE, and Raftery AE (1995). Bayes factors. J Am Stat Assoc 90(430) 773-795.

Keller M, Pasanisi A, Marcilhac M, Yalamas T, Secanell R, Senfaute G (2014). A Bayesian methodology applied to the estimation of earthquake recurrence parameters for seismic hazard assessment. Qual Reliab Eng Int 30:921-933.

Keller M, Mayor J, Duverger C, Senfaute G (2019). Bayesian model comparison applied to probabilistic seismic hazard assessment. Proc European Network for Business and Industrial Statistics 2019, Budapest, Hungary.

Klein FW (2002). User's Guide to HYPOINVERSE-2000, a FORTRAN Program to solve for Earthquake Locations and Magnitudes. U.S. Geological Survey Open File Report 02-171 Version 1.0, 123 pp.

Klein FW (2003). The HYPOINVERSE2000 earthquake location program. In Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (Eds) International Handbook of Earthquake and Engineering Seismology. Academic Press, San Diego, 1619-1620.

Kulkarni RB, Youngs RR, Coppersmith KJ (1984). Assessment of confidence intervals for results of seismic hazard analysis. Proc 8th World Conf Earthquake Eng, San Francisco, pp 263–270.

Lee WHK, Lahr JC (1975). HYPO71 (revised): A computer program for determining hypocenter, magnitude and first motion pattern of local earthquakes. U.S. Geological Survey Open-File Report 75-311, 116 pp.

559 Liu M, and Stein S (2016). Mid-continent earthquakes: Spatio-temporal occurrences, causes and
560 hazards. Earth-Sci Rev 162:364-386.

561 Marzocchi W, Jordan TH (2014). Testing for ontological errors in probabilistic forecasting models of
562 natural systems. P Natl A Sci USA 111(33):11,973–911,978.

563 Marzocchi W, Jordan TH (2018). Experimental concepts for testing probabilistic earthquake
564 forecasting and seismic hazard models. Geophys J Int 215(2):780–798.

565 McGuire RK (2004). Seismic Hazard and Risk Analysis, Earthquake Engineering Research Institute,
566 Oakland CA.

567 Metropolis N, Rosenbluth AW, Teller MN, Teller E (1953). Equation of state calculations by fast
568 computing machines. J Chem Phys 1:1087–1092.

569 Mosca I, Sargeant S, Baptie B, Musson RMW, Pharaoh T (2020). National seismic hazard maps for
570 the UK: 2020 update. British Geological Survey Open Report OR/20/053, United Kingdom.

571 Mosegaard K, Tarantola A (2002). Probabilistic approach to inverse problem. In Lee WHK,
572 Kanamori H, Jennings PC, and Kisslinger C (eds) International Handbook of Earthquake and
573 Engineering Seismology, Academic Press, San Diego, 81A:237-265.

574 Musson RMW (1994). A catalogue of British earthquakes, British Geological Survey Global
575 Seismology Report WL/94/04, United Kingdom.

576 Musson RMW (2000). The use of Monte Carlo simulations for seismic hazard assessment in the UK.
577 Ann Geof 43:1-9.

578 Musson RMW (2012). The effect of magnitude uncertainty on earthquake activity rates. Bull Seismol
579 Soc Am 102(6):2771–2775.

580 Musson RMW, Chadwick RA, Pharaoh TC, Henni PHO, Wild B, Carney JN (2001). Seismic hazard
581 assessment for Wylfa. British Geological Survey Technical Report CR/01/253, Edinburgh, United
582 Kingdom.

583 Musson RMW, Sargeant S (2007). Eurocode 8 seismic hazard zoning maps for the UK. British
584 Geological Survey Report CR/07/125, 62 pp.

585 Musson RMW, Winter PW (2012). Objective assessment of source models for seismic hazard studies:
586 with a worked example from UK data. Bull Earthq Eng 10:367-378.

587 ONR (2018). Nuclear Safety Technical Assessment Guide, NS-TAST-GD-013 Revision 7.

588 Ordaz M, Reyes C (1999). Earthquake hazard in Mexico City: Observations versus computations.
589 Bull Seismol Soc Am 89(5):1379–1383.

590 Reiter L (1990). Earthquake Hazard Analysis: Issues and Insights, Columbia University Press, New
591 York.

592 Rhoades DA, Dowrick DJ (2000). Effects of magnitude uncertainties on seismic hazard estimates.
593 Proc 12th World Conf Earthquake Eng, Wellington, New Zealand, Paper 1179.

594 SHWP (2001). Uniform Risk Spectra for Wylfa Power Station. Report for Magnox Electric plc,
595 Seismic Hazard Working Party, Berkeley.

596 Stein S, Liu M, Camelbeeck T, Merino M, Landgraf A, Hintersberger E, Kuebler S (2015).
597 Challenges in assessing seismic hazard in intraplate Europe. In Landgraf A, Kuebler S, Hintersberger
598 E, and Stein S (eds) Seismicity, Fault rupture and earthquake hazards in slowly deforming region.
599 Geol Soc Spec Publ 432:29-32.

600 Stirling M, Gerstenberger M (2010). Ground motion–based testing of seismic hazard models in New
601 Zealand. Bull Seismol Soc Am 100(4):1407–1414.

602 Tarantola A (1987). Inverse Problem Theory, Elsevier, Amsterdam.

603 USNRC (2012). Practical implementation guidelines for SSHAC level 3 and 4 hazard studies. U.S.
604 Nuclear Regulatory Commission, NUREG-2117, Rev. 1, Washington, DC.

605 Veneziano D, Van Dyke J (1985). Seismic parameter estimation methods. Appendix A in Seismic
606 Hazard Methodology for Nuclear Facilities in the Eastern United States (Draft 85-1). EPRI/SOG.

607 Villani M, Lubkowski Z, Free M, Musson RMW, Polidoro B, McCully R, Koskosidi A, Oakman C,
608 Courtney T, Walsh M (2020). A probabilistic seismic hazard assessment for Wylfa Newydd, a new
609 nuclear site in the United Kingdom. Bull Earthq Eng https://doi:10.1007/s10518-020-00862-8.

610 Weatherill G, Burton PW (2009). Delineation of shallow seismic source zones using K-means cluster
611 analysis, with application to the Aegean region. Geophys J Int 176:565-588.

612 Weatherill G, Burton PW (2010). An alternative approach to probabilistic seismic hazard analysis in
613 the Aegean region using Monte Carlo simulations. Tectonophysics 496:253-278.

614 Wessel P, Smith W, Scharroo R, Luis J, Wobbe F (2013). Generic Mapping Tools: Improved version
615 released. EOS Trans AGU 94:409-410.

616 Woessner J, Giardini D, the SHARE Consortium (2015). Seismic hazard estimates for the Euro-
617 Mediterranean region: A community-based probabilistic seismic hazard assessment. Bull Earthq Eng
618 13:3553-3596.

619

620 **Table 1:** Values of the completeness magnitude Mc for the earthquake catalogue.

Mc	**Completeness Period**
2.1	1999
2.5	1985
3.0	1970
3.4	1850
4.0	1750
4.5	1700
4.9	1650

621

622 **Table 2**: Classification of the uncertainty in the epicentral location of the observed historical

623 (pre-1970) and instrumental (post-1970) earthquakes.

	Uncertainty class	**Location uncertainty**	**Preferred value [km]**
Pre-1970 event	A	$\sigma < 5$ km	5 km
Pre-1970 event	B	$5 \leq \sigma < 15$ km	10 km
Pre-1970 event	C	$15 \leq \sigma < 30$ km	22.5 km
Pre-1970 event	D	$\sigma \geq 30$ km	30 km
Post-1970 event	A	$\sigma < 5$ km	5 km

624

625 **Table 3:** Name, reference, number of the source zones, variation of the recurrence parameters in the
626 source zones of the 14 candidate seismic source models (SSMs) for the study area.

SSM	Reference	Type of study	Number of zones	Variation of N(≥ 2.1Mw/yr)	Variation of the *b*-value
1	Source model developed for the Seismic Hazard Harmonization in Europe project by Woessner et al. (2015)	Regional model	13	0.18; 1.27	0.84; 1.30
2	Source model for the UK current national hazard maps developed by Musson and Sargeant (2007)	Regional model	16	0.19; 1.13	0.65; 1.43
3	Countries in the British Isles (Wales, England, North Ireland, Scotland, and Ireland) modified such that the seismicity is not clustered in each source zone using the nearest neighbour analysis	Model developed for this work	8	0.10; 2.44	0.89; 1.23
4	Regions in the British Isles (Scotland, North Ireland, Ireland, North-East, North-West, Yorkshire and the Humber, East Midlands, West Midlands, East of England, South West, South East) modified such that	Model developed for this work	12	0.07; 1.44	0.81; 1.23

	the seismicity is not clustered in each source zone using the nearest neighbour analysis				
5	Musson et al. (2001)	Site-specific model	12	0.01; 2.28	0.80; 1.33
6	SHWP (2001)	Site-specific model	7	0.07; 5.00	0.93; 1.12
7	Source model based on the observed historical and instrumental recorded seismicity and developed by Villani et al. (2020)	Site-specific model	14	0.07; 1.30	0.71; 1.32
8	Source model based on the observed historical and instrumental recorded seismicity and developed by Villani et al. (2020). The source zone in the Lleyn Peninsula was extended to the southwest to include the 3.5 Mw earthquake on 29 May 2013.	Site-specific model	14	0.07; 1.30	0.71; 1.32
9	Source model based on the potential controls of the regional geology and tectonics on the distribution of the UK seismicity and developed by Villani et al. (2020).	Site-specific model	17	0.02; 1.30	0.79; 1.32
10	Source model based on the potential controls of the regional geology and	Site-specific model	17	0.02; 1.30	0.79; 1.32

	tectonics on the distribution of the UK seismicity and developed by Villani et al. (2020). The boundary between the zones for the Anglesey and Lleyn Peninsula is shifted southern than in SSM9.				
11	Source model based on the potential controls of the regional geology and tectonics on the distribution of the UK seismicity and developed by Villani et al. (2020). The Anglesey and Lleyn Peninsula are included in a single zone.	Site-specific model	16	0.02; 1.30	0.79; 1.32
12	Single source zone, i.e. the study area	Model developed for this work	1	7.12	1.025
13	Source model developed for the Global Seismic Hazard Map project by Grünthal et al. (1999)	Regional model	7	0.36; 1.92	0.79; 1.35
14	Source model for the national seismic hazard maps for the UK developed by Mosca et al. (2020)	Regional model	14	0.06; 1.41	0.92; 1.31

627

628 **Table 4:** Sampling range of the *b*-value and the activity rate to generate the synthetic catalogues.

Parameter	**Range**
b-value	Gaussian, Mean: 1.025, Sigma: 0.041
N(≥2.1/yr) for SSMs with > 15 source zones	0.30; 0.57
N(≥2.1/yr) for SSMs 10-15 source zones	0.35; 0.70
N(≥2.1/yr) for SSMs with 4-9 source zones	0.63; 1.31
N(≥2.1/yr) for SSMs with ≤ 3 source zones	5.08; 9.16

629

630 **Table 5:** Sensitivity analysis for the sampling range of the recurrence parameters.

Parameter	Test 1	Test 2	Test 3
b-value	Uniform, 0.9; 1.2	Gaussian, Mean: 1.025, Sigma: 0.041	Uniform, 0.7; 1.3
N(≥2.1/yr)	For each SSM, (7.12 ± 2 * 0.51)/ number of zones	0.0; 1.0 for SSMs with > 10 source zones	0.30; 0.57 for SSMs with > 15 source zones
		0.4; 1.5 for SSMs with 4-9 source zones	0.35; 0.70 for SSMs with 10-15 source zones
		5.08; 9.16 for SSMs with ≤ 3 source zones	0.63; 1.31 for SSMs with 4-9 source zones
		-	5.08; 9.16 for SSMs with ≤ 3 source zones

631

632

10. List of figure captions

Figure 1 Schematic workflow of the Bayesian Metropolis-Hastings (BMH) approach.

Figure 2 Comparison of the normalized distributions of activity rates (top panel) and *b*-values (bottom panel) between synthetic and observed catalogues before and after applying the Metropolis-Hastings (M-H) Algorithm: a) N (≥ 2.1 Mw/yr) before applying the M-H Algorithm; b) N (≥ 2.1 Mw/yr) after applying the M-H Algorithm; c) *b*-value before applying the M-H; and d) *b*-value after applying the M-H Algorithm. The M-H Algorithm accepts 1,052,929 synthetic catalogues out of 4,200,000 catalogues.

Figure 3 (a) Topographic map of Europe from the global model ETOPO1 (Amante and Eakins, 2009). The black lines represent the plate boundaries and the black circle indicates the study area. (b) Distribution of the mainshocks in the study area within the completeness threshold in Table 1. Events with unknown depth are coloured in white. The black circles describe the area within 300 and 60 km from the site, respectively. The star describes the site of Wylfa Newydd.

Figure 4 14 candidate seismic source models (SSMs) for the study area. The star describes the site of Wylfa Newydd.

Figure 5 Distribution of the misfit function (Equation 2) for the 1,052,929 accepted models by the Metropolis-Hastings Algorithm and four different sizes of the cell grid.

Figure 6 1-D posterior *pdf* for the 14 candidate SSMs obtained by the Bayesian Metropolis-Hastings approach.

Figure 7 Distribution of the accepted synthetic catalogues by the BMH approach as a function of the SSMs on the left-hand side and the misfit function on the right-hand side.

Figure 8 1-D posterior *pdf* as a function of SSMs for the three tests in Table 5.

655 **Figure 9** 14 candidate seismic source models (SSMs) for the study area of 60 km within the site. The
656 star describes the site of Wylfa Newydd.

657 **Figure 10** 1-D posterior *pdf* as a function of SSMs for three cases: the entire catalogue and the 300-
658 km study area (Reference), the instrumental catalogue and the 300-km study area (Test A), and the
659 entire catalogue and the 60-km study area (Test B).

Figures

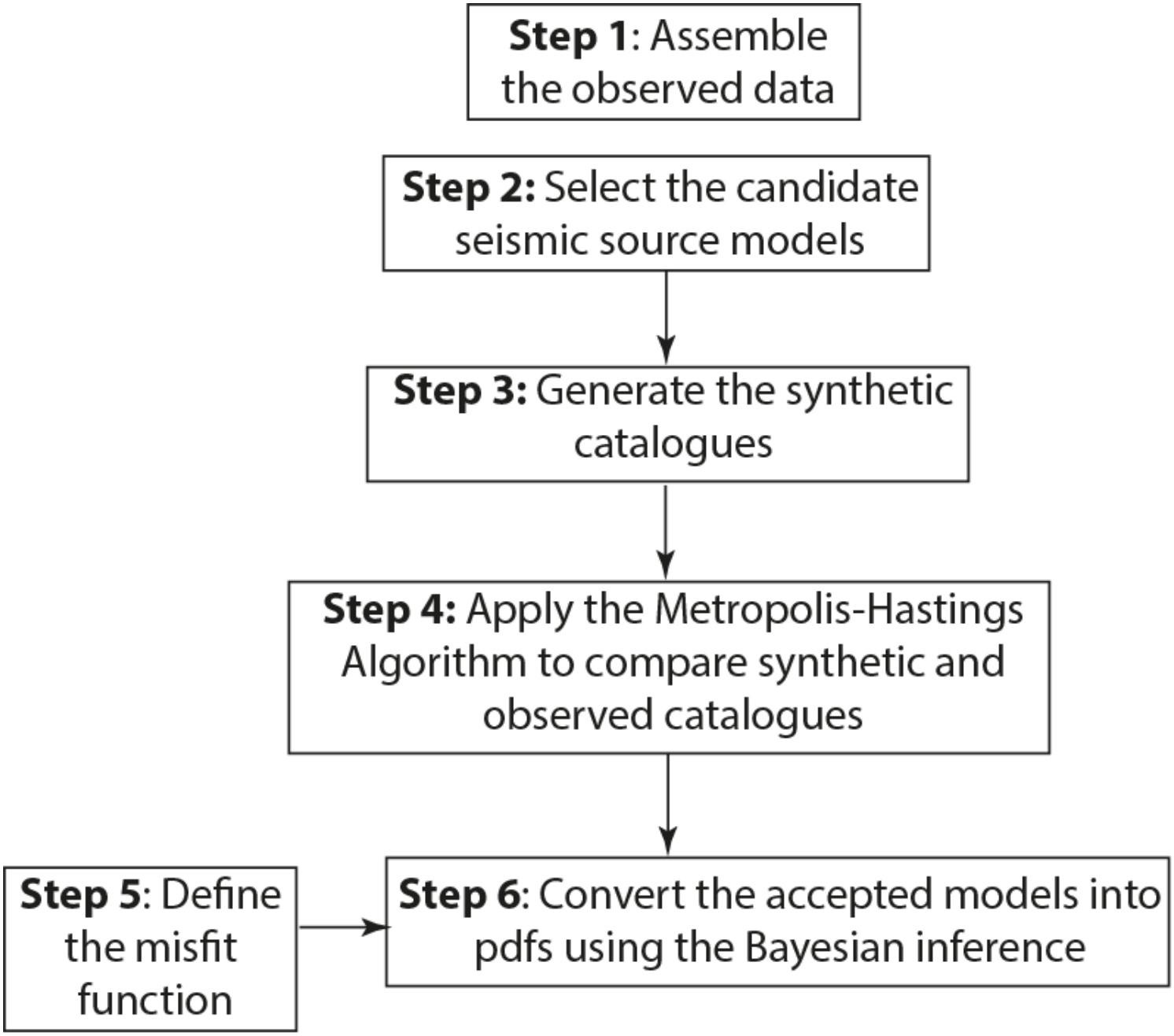

Figure 1

Schematic workflow of the Bayesian Metropolis-Hastings (BMH) approach.

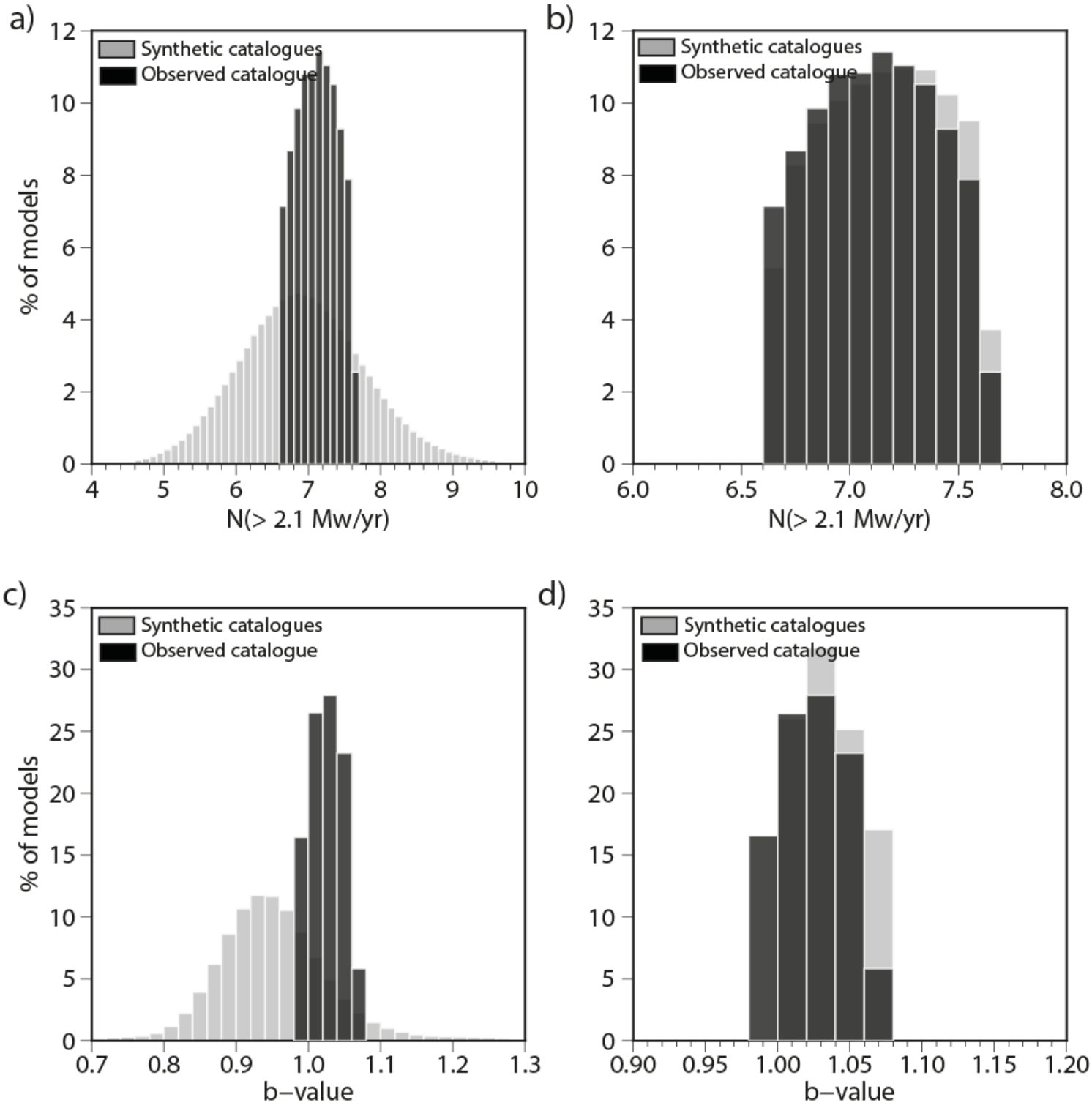

Figure 2

Comparison of the normalized distributions of activity rates (top panel) and b-values (bottom panel) between synthetic and observed catalogues before and after applying the Metropolis-Hastings (M-H) Algorithm: a) N (≥ 2.1 Mw/yr) before applying the M-H Algorithm; b) N (≥ 2.1 Mw/yr) after applying the M-H Algorithm; c) b-value before applying the M-H; and d) b-value after applying the M-H Algorithm. The M-H Algorithm accepts 1,052,929 synthetic catalogues out of 4,200,000 catalogues.

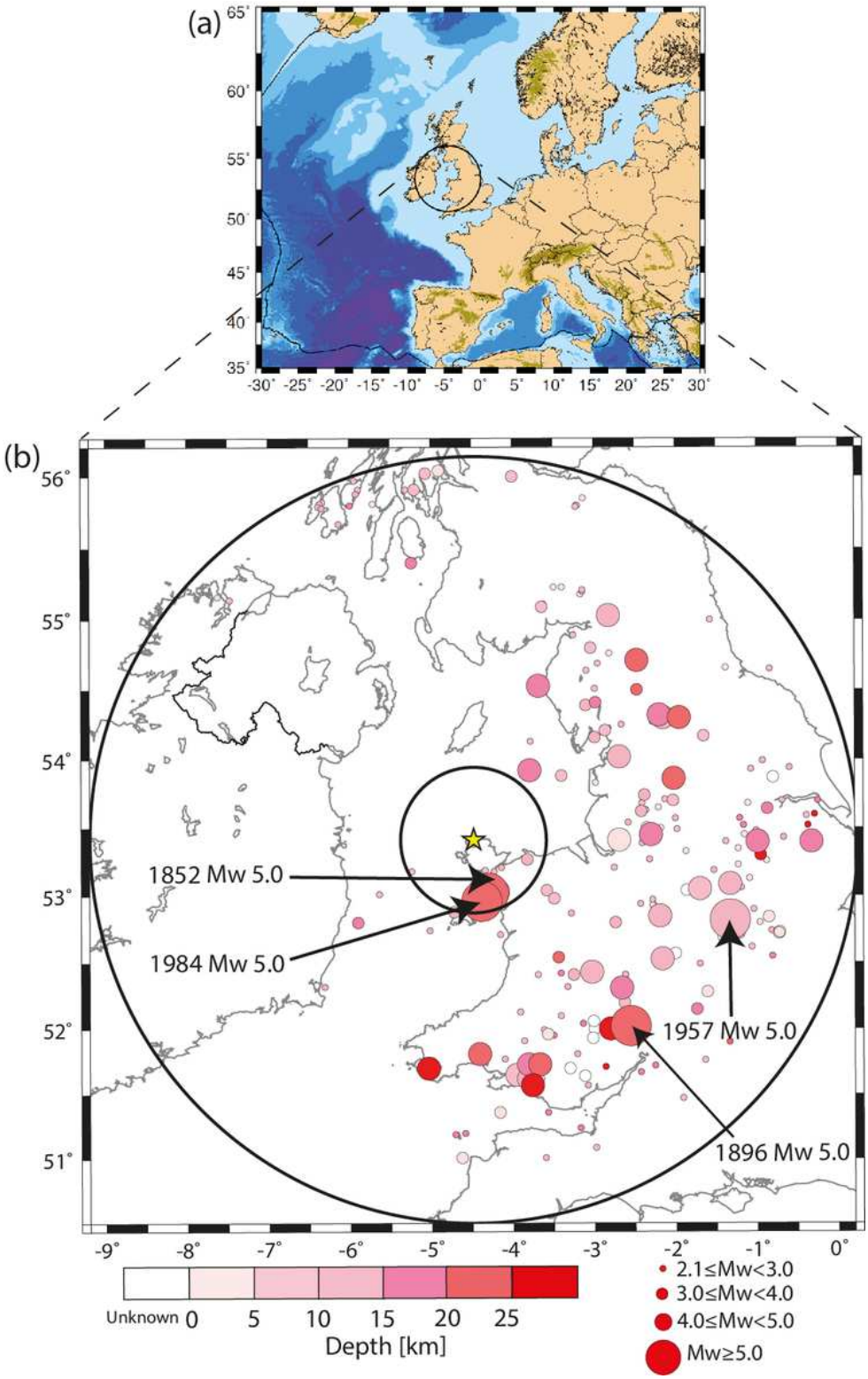

Figure 3

(a) Topographic map of Europe from the global model ETOPO1 (Amante and Eakins, 2009). The black lines represent the plate boundaries and the black circle indicates the study area. (b) Distribution of the mainshocks in the study area within the completeness threshold in Table 1. Events with unknown depth are coloured in white. The black circles describe the area within 300 and 60 km from the site, respectively. The star describes the site of Wylfa Newydd. Note: The designations employed and the presentation of

the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

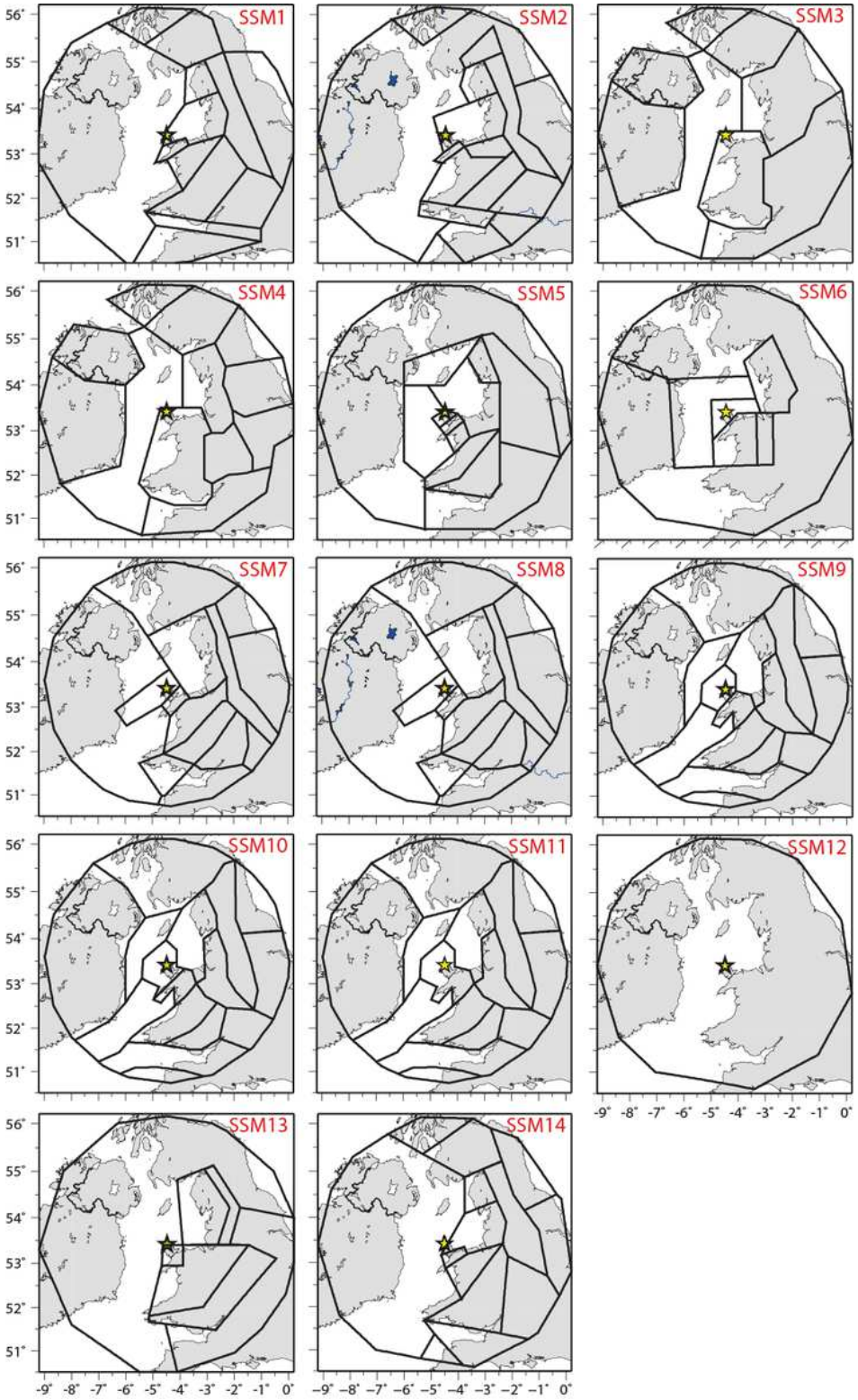

Figure 4

14 candidate seismic source models (SSMs) for the study area. The star describes the site of Wylfa Newydd. Note: The designations employed and the presentation of the material on this map do not imply

the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

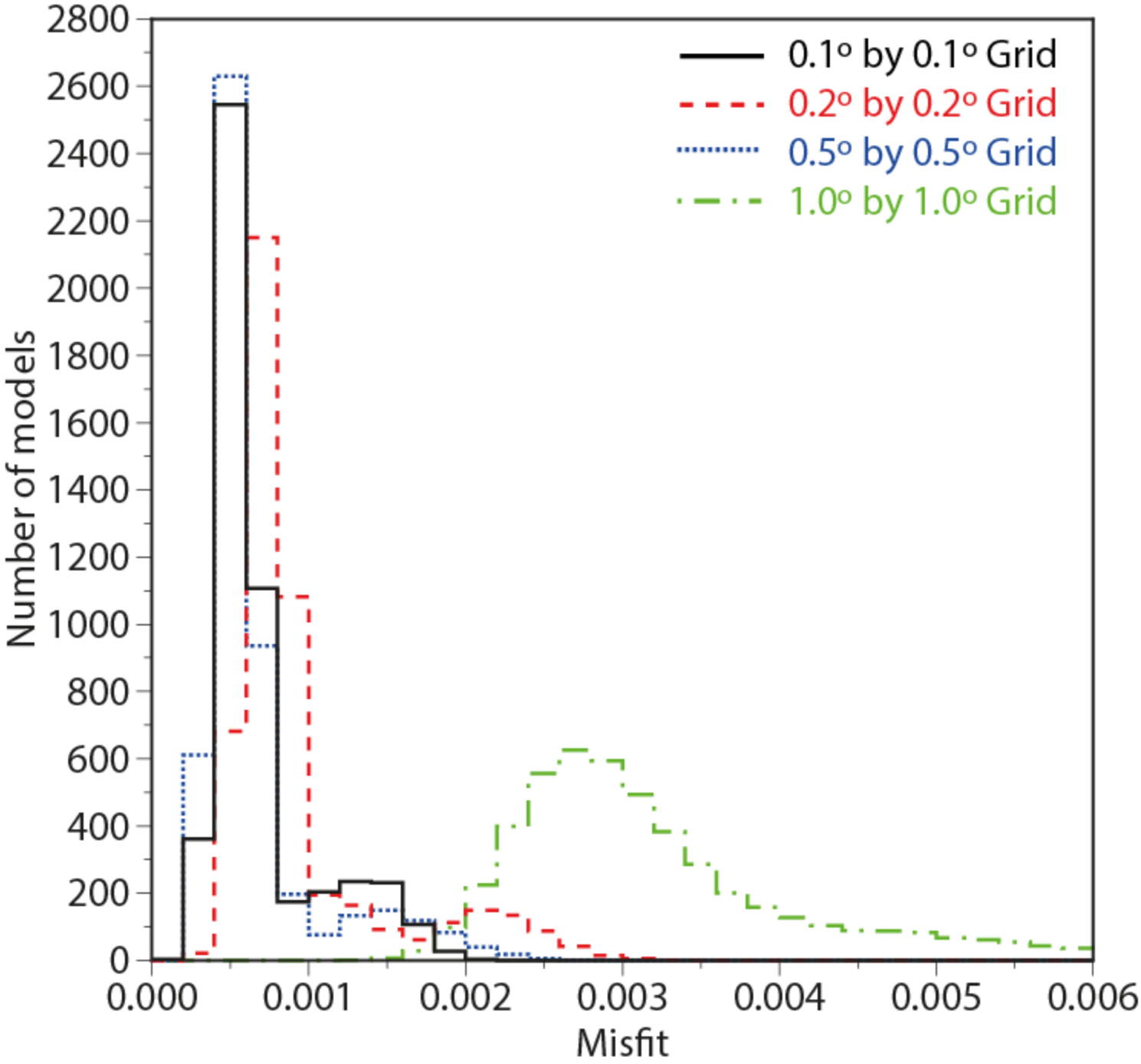

Figure 5

Distribution of the misfit function (Equation 2) for the 1,052,929 accepted models by the Metropolis-Hastings Algorithm and four different sizes of the cell grid.

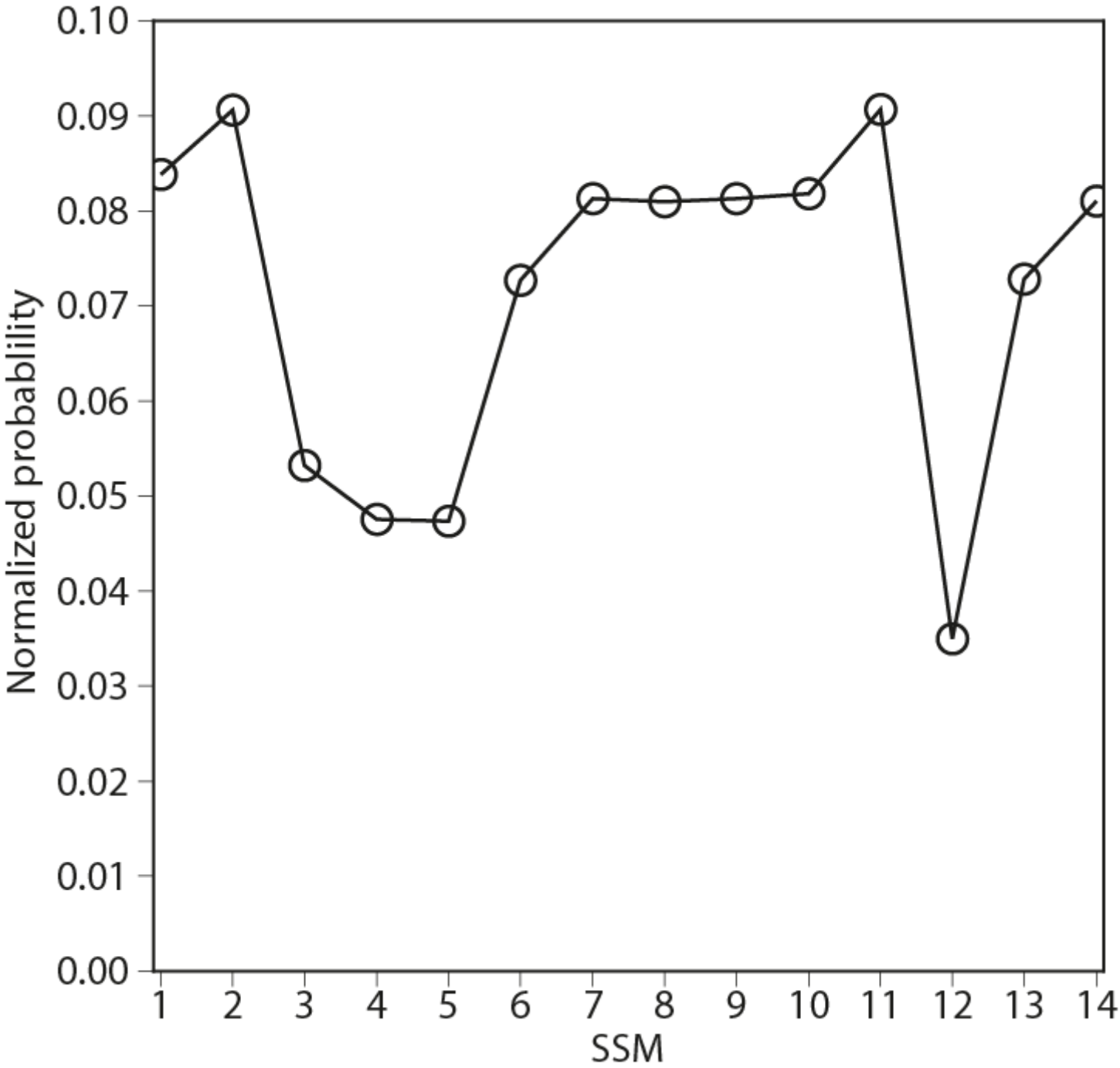

Figure 6

1-D posterior pdf for the 14 candidate SSMs obtained by the Bayesian Metropolis-Hastings approach.

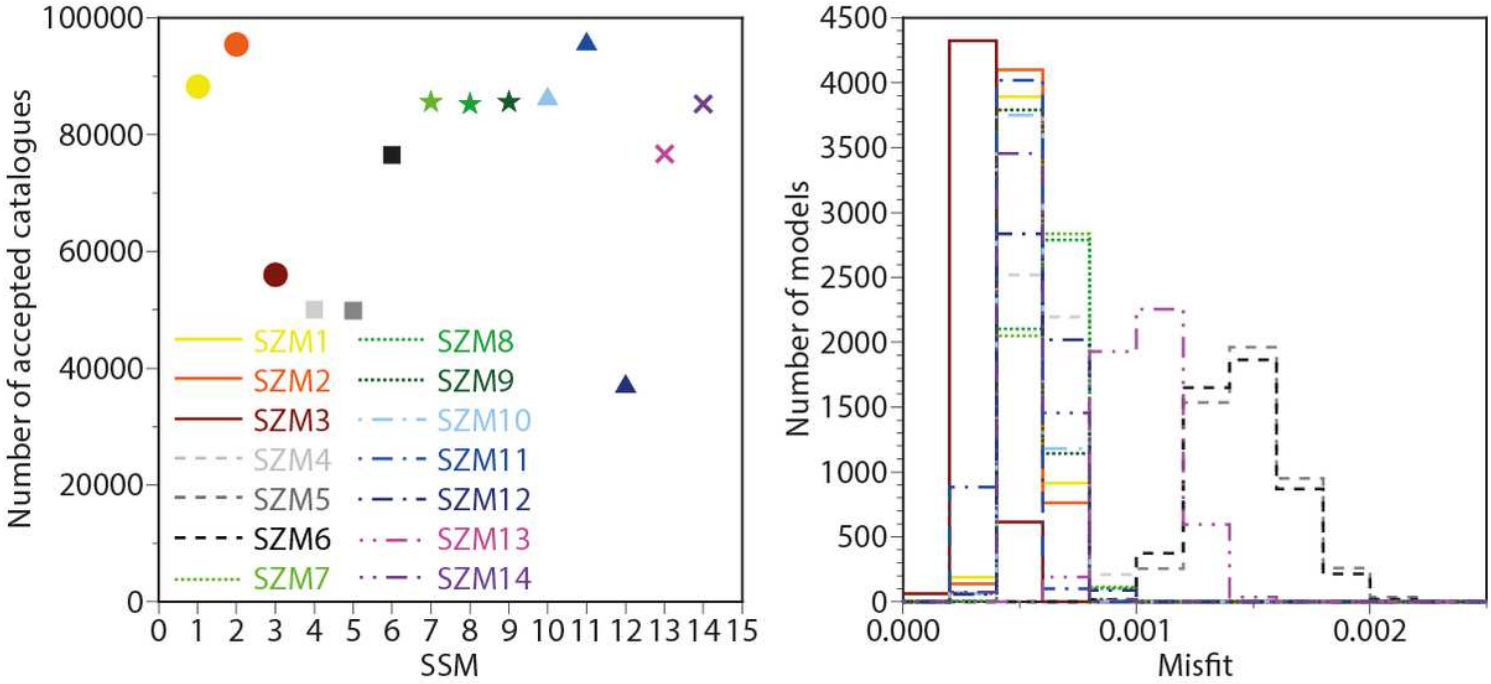

Figure 7

Distribution of the accepted synthetic catalogues by the BMH approach as a function of the SSMs on the left-hand side and the misfit function on the right-hand side.

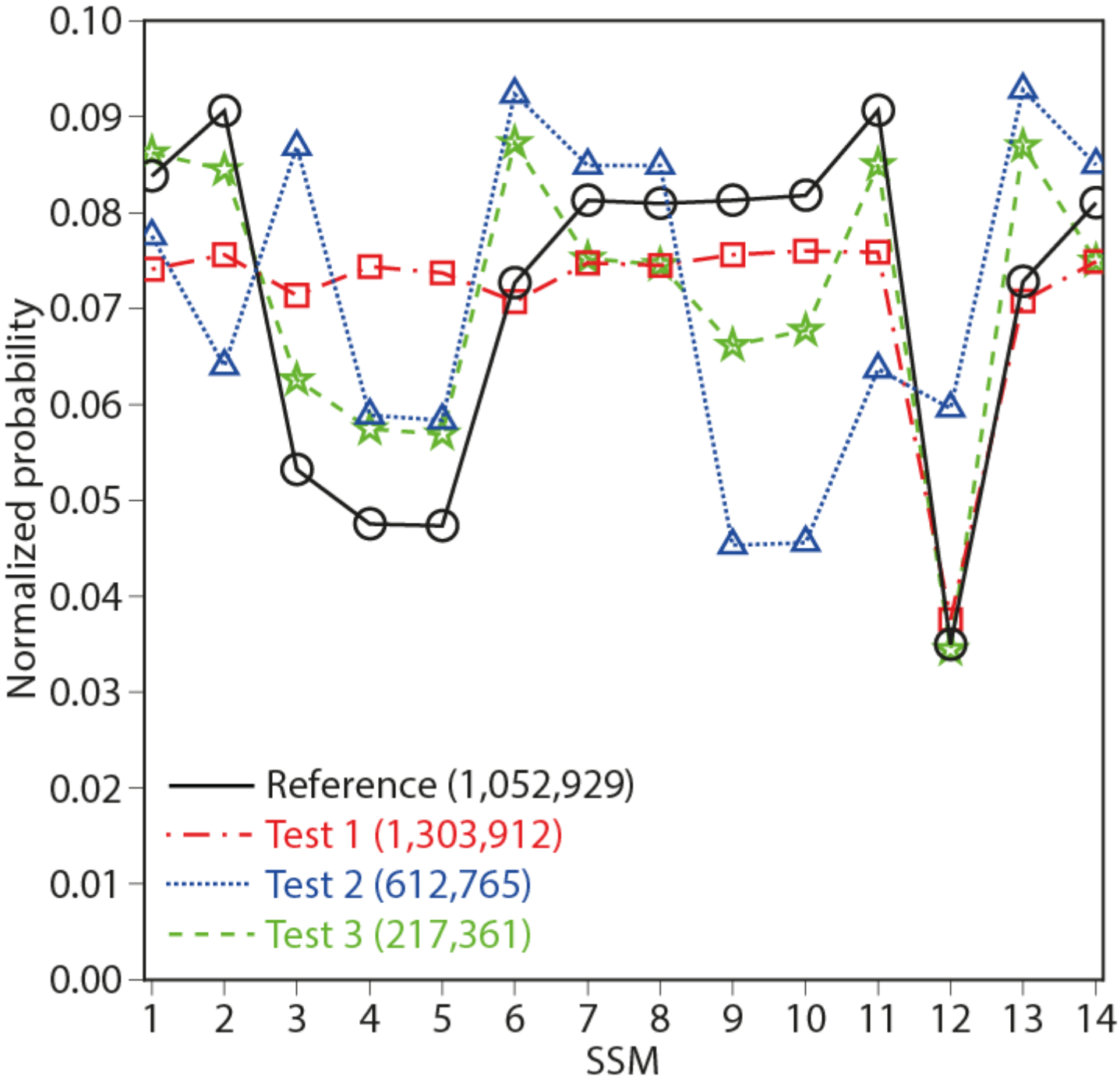

Figure 8

1-D posterior pdf as a function of SSMs for the three tests in Table 5.

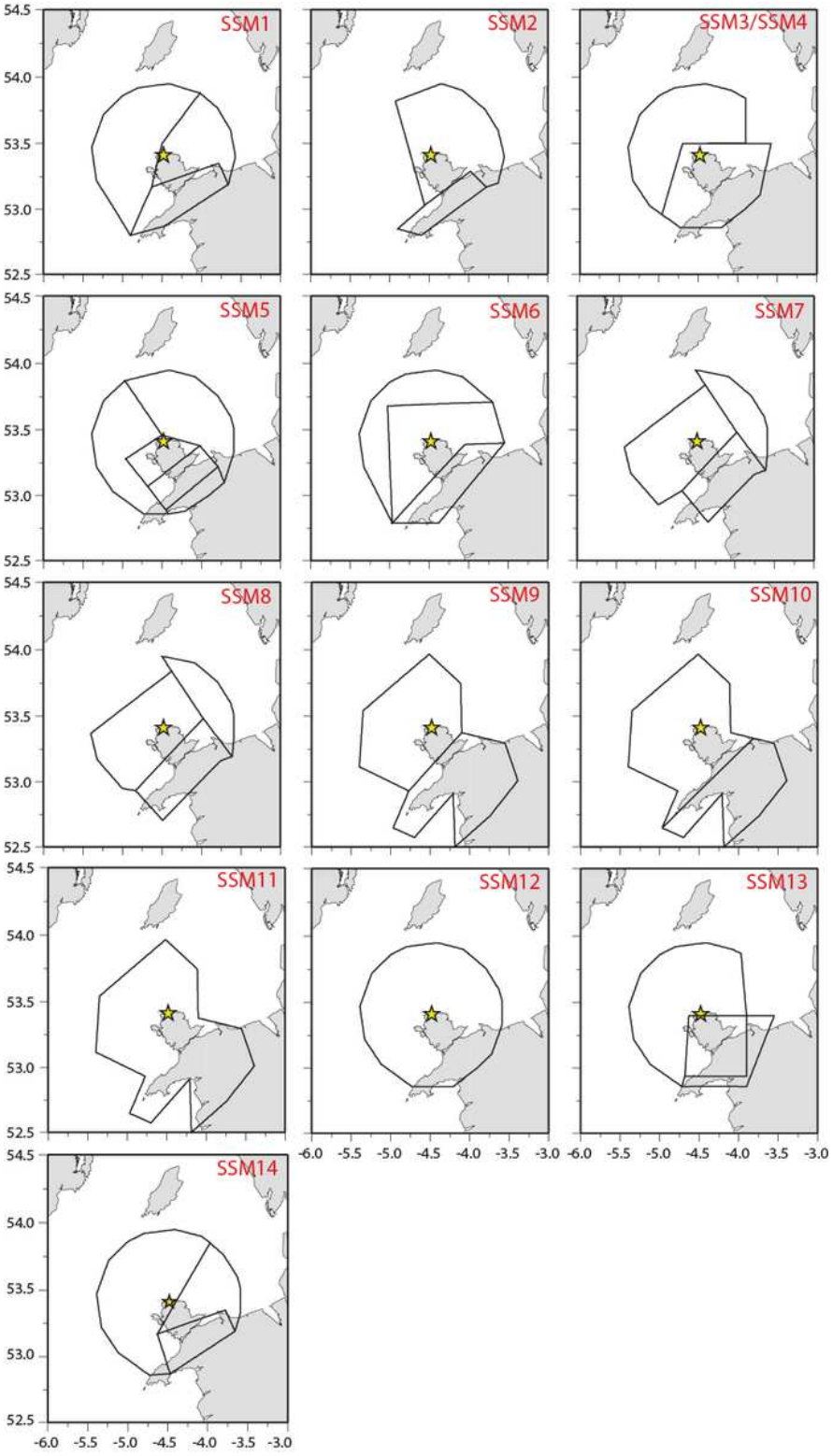

Figure 9

14 candidate seismic source models (SSMs) for the study area of 60 km within the site. The star describes the site of Wylfa Newydd. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

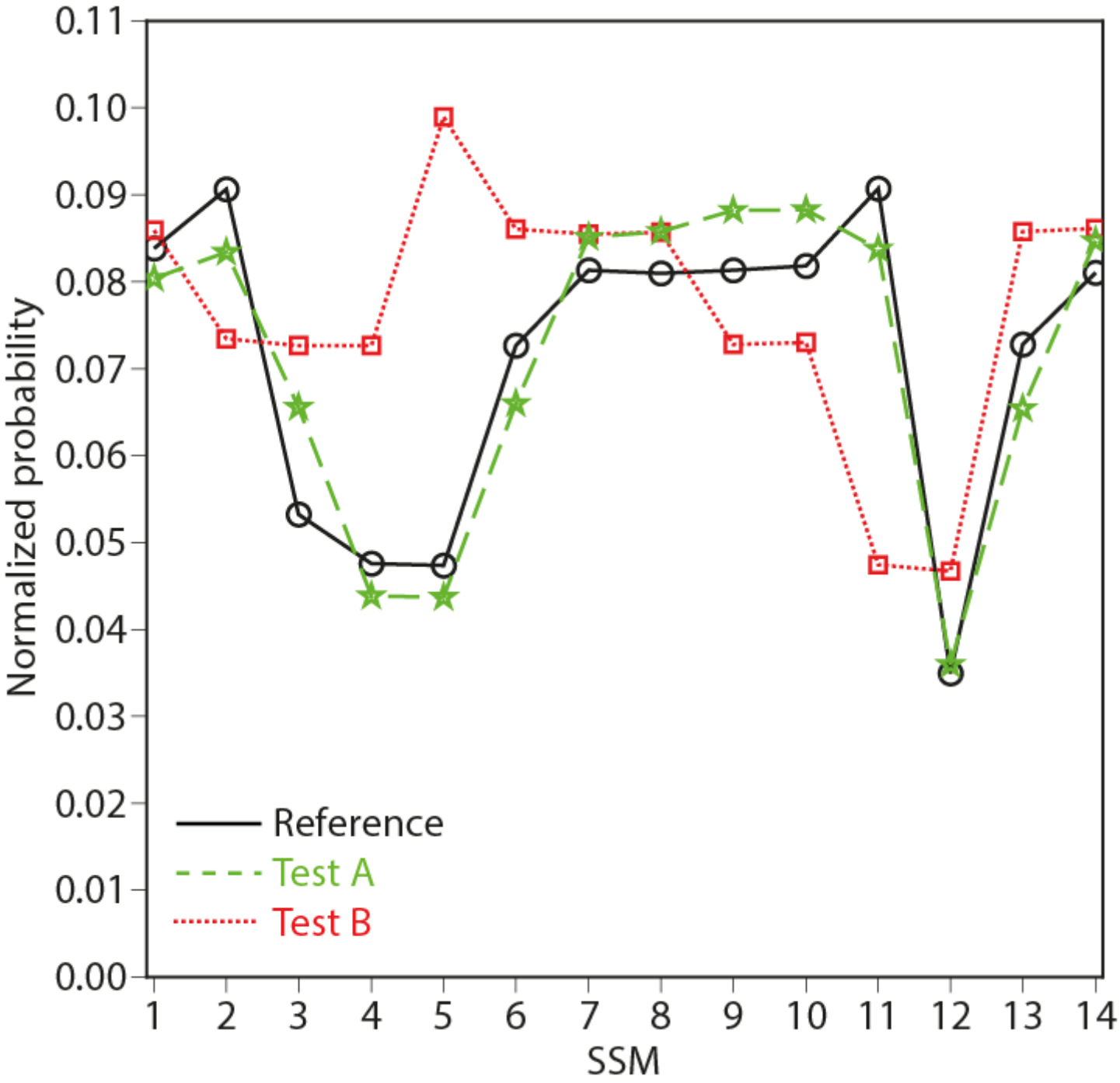

Figure 10

1-D posterior pdf as a function of SSMs for three cases the entire catalogue and the 300-km study area (Reference), the instrumental catalogue and the 300-km study area (Test A), and the entire catalogue and the 60-km study area (Test B).