Geophysical Journal International

Geophys. J. Int. (2013) **195,** 1314–1322
Advance Access publication 2013 August 28

doi: 10.1093/gji/ggt310

Selecting ground-motion models developed for induced seismicity in geothermal areas

Benjamin Edwards[1] and John Douglas[2]

[1]*Swiss Seismological Service, Eidgenössische Technische Hochschule (ETH), Zürich, Sonneggstrasse 5, 8092 Zurich, Switzerland.*
E-mail: edwards@sed.ethz.ch
[2]*Seismic and Volcanic Risks Unit, Risks and Prevention Division, BRGM, Orléans, France*

Accepted 2013 July 30. Received 2013 July 30; in original form 2013 May 21

SUMMARY

We present a case study of the ranking and weighting of ground-motion prediction equations (GMPEs) for seismic hazard assessment of enhanced geothermal systems (EGSs). The study region is Cooper Basin (Australia), where a hot-fractured-rock project was established in 2002. We test the applicability of 36 GMPEs based on stochastic simulations previously proposed for use at EGSs. Each GMPE has a set of corresponding model parameters describing stress drop, regional and local (near-surface) attenuation. To select suitable GMPEs for Cooper Basin from the full set, we applied two methods. In the first, seismograms recorded on the local monitoring network were spectrally analysed to determine characteristic stress and attenuation parameters. In a second approach, residual analysis using the log-likelihood (LLH) method was used to directly compare recorded and predicted short-period response spectral accelerations. The resulting ranking was consistent with the models selected based on spectral analysis, with the advantage that a transparent weighting approach was available using the LLH method. Region-specific estimates of variability were computed, with significantly lower values observed compared to previous studies of small earthquakes. This was consistent with the limited range of stress drops and attenuation observed from the spectral analysis.

Key words: Earthquake ground motions; Earthquake source observations; Seismic monitoring and test-ban treaty verification; Seismic attenuation.

INTRODUCTION

Ground shaking from seismicity associated with stimulation and exploitation of a geothermal reservoir for heat and power production can be a significant nuisance to the local population and can, in some cases, lead to building damage. The Deep Heat Mining project (Basel, Switzerland) in 2006 triggered an M_L 3.4 (M_w 3.2) main shock and thousands of smaller shocks and led to insurance claims of more than $9 million (Giardini 2009). Two earthquakes (M_L 2.4 and 2.7) occurred in the vicinity of the Landau (Germany) geothermal power plant in 2009, which caused macroseismic intensities up to $V+$, while at another German geothermal project (Insheim) two felt tremors (M_L 2.2 and 2.4) occurred during reservoir stimulation in 2010 (Groos *et al.* 2013). Most recently, in July 2013, a geothermal project in St Gallen, Switzerland triggered a widely felt M_L 3.5 (M_w 3.4) event, which was followed by numerous smaller aftershocks. In Majer *et al.* (2012), seven steps are proposed to help assess and mitigate the seismic risk posed by geothermal systems. Step 5 of the proposal is to 'quantify the hazard from natural and induced seismic events' through either probabilistic or deterministic approaches. They suggest a two-stage approach to quantify the hazard: a baseline estimate initially established through regional seismicity, with further refinement to a site-specific hazard assessment through the analysis of induced seismicity recorded on the local monitoring network. This paper addresses the second stage.

A necessary component of any study that seeks to assess the seismic hazard (and/or risk) associated with geothermal projects is a ground-motion model that estimates measures of shaking [e.g. peak ground acceleration (PGA)] given an earthquake scenario (e.g. in terms of magnitude and source-to-site distance). Motivated by a lack of suitable models in the public literature, Douglas *et al.* (2013) derived a set of stochastic and empirical ground-motion models for application in geothermal areas. These models were based on analyses of thousands of near-source seismograms of small earthquakes, most of which were induced by geothermal activity or gas extraction, while some were natural shallow events.

Because of the considerable epistemic uncertainty in the estimation of ground motions in geothermal areas, Douglas *et al.* (2013) presented ground-motion prediction equations (GMPEs) for 36 stochastic simulation models (Boore 2003) that sought to capture this uncertainty. The ranges of the key parameters of these stochastic models (stress parameter, $\Delta\sigma$; path attenuation, Q; near-surface attenuation, κ) were defined based on the analysis of seismograms collected from numerous regions. The analysis highlighted

considerable variation in these parameters among regions and sites. When conducting a seismic hazard assessment for a given geothermal project it is not known *a priori* which of the 36 models are most applicable. Consequently Douglas *et al.* (2013) recommended, in the absence of other information by which to constrain the stochastic parameters for a given site, that a logic tree is used for seismic hazard analysis with all 36 models as branches. Subsequently, as seismograms are recorded at the geothermal project site, the applicability of some models should become evident and hence their associated branches could be assigned higher weights, while others may be downweighted or even dropped completely. There are two complementary ways in which the branch weights can be updated: direct assessment of the stochastic parameters, and comparison of ground-motion predictions and observations.

As a demonstration of the use of the stochastic simulation GMPEs and the proposed procedure for the assessment of the logic-tree weights, Douglas *et al.* (2013) present a simple analysis for Campi Flegrei (Italy), for which some seismograms from small, shallow (but natural) earthquakes were available. However, the small number of seismograms available (only 55), their limited bandwidths and low quality of the metadata meant that the logic-tree weights could not be significantly updated. The aim of this paper is to use a much larger and higher quality data set from an enhanced geothermal system (EGS), to better demonstrate the proposed selection procedure and to investigate how many records are required to significantly reduce the epistemic uncertainty in ground-motion prediction for EGSs. The data set comes from Cooper Basin (Australia), which was not considered by Douglas *et al.* (2013) when developing their ground-motion models.

The next section summarizes the 36 stochastic models developed by Douglas *et al.* (2013). They did not recommend their empirical models for application because they were derived using data from a narrow magnitude–distance range (roughly M_w 1–3 and R_{hyp} < 10 km) and a simple functional form was adopted. Because the stochastic models were based on simulated ground motions from a wider range of magnitudes and distances and a more complex functional form was fit to these simulations, they are considered more robust than the empirical equations for M_w 1–5 and R_{hyp} 1–50 km. As with empirical models, however, extrapolation of stochastic models outside their range of applicability is not recommended. Following the introduction of the models, the data set from Cooper Basin EGS is presented. The subsequent section derives estimates of the stochastic parameters from these data and investigates the impact of the number of seismograms used to estimate the parameters. Based on this analysis, a set of weights for the 36 models is proposed. A second set of weights is proposed in the following section based on comparisons between the observed response spectral accelerations and those predicted by the 36 GMPEs.

GROUND-MOTION MODELS FOR INDUCED SEISMICITY

Ground motions from small earthquakes, particularly those recorded in the near-source region, often exhibit large variability for a given magnitude and distance, the principal independent parameters for GMPEs. There are two explanations for this common observation: the first is related to the fact that metadata for small earthquakes are often poorer quality (e.g. routine automatic locations as opposed to manually reviewed locations). The second reason (exacerbated by the first) is that ground motions from small events are more sensitive to changes in hypocentral depth, while site attenuation (κ) tends to filter out, to varying degrees, the dominant high frequencies associated with smaller earthquakes (Douglas & Jousset 2011). Furthermore, it is often observed that the variability of the stress drop (or conversely, slip velocity) is significantly higher for small events than for larger events (e.g. Cotton *et al.* 2013). Whether an artefact of inversion procedures (e.g. not properly accounting for attenuation), or reality, this nevertheless reflects the greatly differing proportion of high-frequency energy observed in small earthquakes of similar size.

Analysing data from six independent regions (Basel, Soultz, Geysers, Hengill, Roswinkel and Vorendaal), Douglas *et al.* (2013) found that a significant reduction in the overall prediction uncertainty was obtained by accounting for region-specific biases. As discussed, this can be interpreted as either systematic bias in the metadata, or alternatively, region-specific 'characteristic seismicity' and recording conditions. Since magnitudes were recomputed homogeneously and hypocentral depths are generally well-constrained for sources directly below the recording network (typical in geothermal installations), Douglas *et al.* (2013) suggested that differences in source, path and site conditions were the likely cause of region-specific differences. They proposed a suite of 36 stochastic models with different source, path and site properties to cover the range observed in their data sets.

In terms of seismic hazard assessment, these 36 models can be considered to cover the epistemic uncertainty: in the case of a completely unknown site, we cannot distinguish between any of the models, and must weight them equally. In reality, of course, some information about the site of interest will always be available: for instance, if the EGS is not located on outcropping hard-rock, then we can already rule out the models with the lowest levels of site attenuation (κ). Each of the 36 models has its own associated aleatory variability, which corresponds to a region-specific model. An initial 'naïve' application of these GMPEs would be similar in terms of the resulting mean hazard to using the single empirical GMPE developed by Douglas *et al.* (2013) using data from all regions (along with the associated high aleatory variability). In terms of the hazard distribution, the obvious difference is that the empirical model leads to only one curve, while the stochastic models generate 36 individual hazard curves, representing the epistemic uncertainty. However, in the case of improved knowledge of the site's seismicity, or recording conditions, we can begin to reassess the weighting of the 36 models, reducing the epistemic uncertainty. The advantage in this case is clear, since region-specific GMPEs cannot typically be assessed due to limited recorded distance and magnitude ranges, the stochastic model approach allows us to refine the logic tree in the case of improved knowledge, which will be quickly available after the installation of a monitoring network, or even beforehand when local data already exist.

GROUND-MOTION DATA USED FOR MODEL SELECTION

The data used in this paper come from a sensitive seismic network set-up to monitor the geothermal exploitation of the reservoir at Cooper Basin (South Australia). A hot-fractured-rock project was launched at Cooper Basin in 2002 to exploit the Habanero granite reservoir at depths between 4 and 4.5 km. Various boreholes and stimulation experiments have been conducted since and triggered earthquakes have been located and characterized (Baisch *et al.* 2006). The data used here come from 2005 (data from 2003 are available but there is uncertainty over the calibration factors

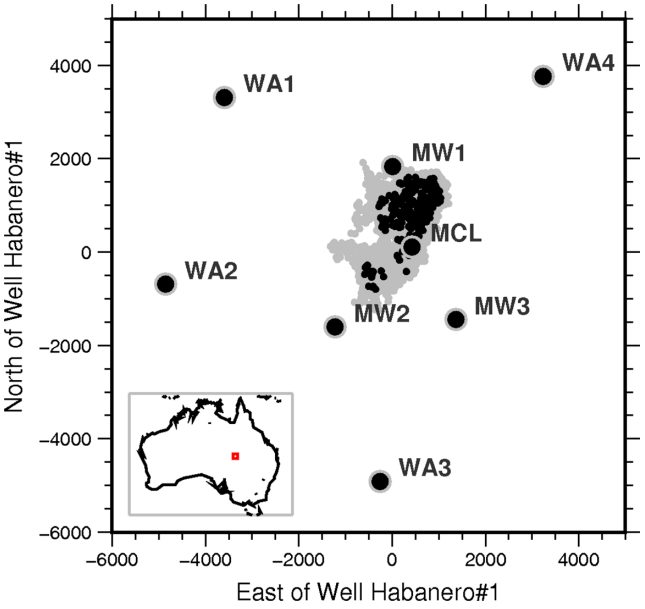

Figure 1. Map of microseismic monitoring network (McLeod #1: MCL, MW01–3 and WA01–3). Catalogued events from 2005 are indicated by grey symbols, those used in the spectral analysis for determination of κ and $\Delta\sigma$ are shown by black symbols.

of the seismometers). Data from eight stations (Stang 2011) installed by Q-Con [McLeod #1 (MCL), WA01–04 and MW01–03] are used (Fig. 1), all of which are located below the surface (all at depths of less than 357 m, except for McLeod #1 at 1.8 km). High-quality earthquake catalogues were provided by Q-Con for these data. The available records come from earthquakes with moment magnitudes between 1.7 and 3.1 (roughly following a standard Gutenberg–Richter distribution), hypocentral distances between 2.4 and 7.8 km (roughly uniformly distributed) and depths between 3.9 and 4.5 km (roughly normally distributed with a peak around 4.2 km). All records have been converted from velocity to acceleration through time-domain differentiation and application of calibration factors. Following the approach detailed in Douglas *et al.* (2013) and Edwards & Douglas (2013), all earthquakes used here have had their moment magnitudes consistently recalculated.

ESTIMATION OF STOCHASTIC PARAMETERS

As noted, the estimation of suitable stochastic model parameters allows the reduction of epistemic uncertainty related to ground-motion prediction in the study region for the magnitude and distance range of the data. For earthquakes outside the range of observations (particularly larger magnitudes), however, there still remains considerable epistemic uncertainty because it is not certain that the most appropriate stochastic parameters (or the best-fitting models) for the available data necessarily apply for such scenarios. In a first step, we look for existing parameters published in the literature. At Cooper Basin, Baisch *et al.* (2009) found an average whole-path $Q_p = 112 \pm 21$ and $\Delta\sigma = 4.7$ bars (0.47 MPa) from the analysis of over 6000 events. Assuming $Q_p \approx Q_s$ and taking the measured $V_P = 3.7\,\mathrm{km\,s^{-1}}$ and V_P:$V_S = 1.9$ from Baisch *et al.* (2009) we obtain path attenuation (t^*) of approximately 0.02 s between the event cloud and the surface. Due to the almost vertical propagation through the basin, this t^* can be completely assigned to the site-specific term, κ_0. Regional values of Q have been computed for Australian earthquakes by Allen *et al.* (2006, 2007). For southeastern Australia Allen *et al.* (2007) found

$$\log Q(f) = 3.66 - 1.44 \log f + 0.768 (\log f)^2 + 0.058 (\log f)^3 \quad (1)$$

for frequencies (f) 0.78–19.9 Hz, leading to consistently high Q (1063–6671), while for southwestern Australia Allen *et al.* (2006) found

$$Q(f) = 457 f^{0.37} \quad (2)$$

for frequencies 1.07–25.0 Hz. The study of Allen *et al.* (2007) computed a corresponding geometrical attenuation of $r_{\mathrm{hyp}}{}^{-1.3}$ in the first 90 km (where r_{hyp} is hypocentral distance) based on the decay of long-period displacement spectra. The central-east location (Fig. 1) of Cooper Basin may be slightly better described by 'southeastern Australia'. However, in light of the known trade-off between Q and geometrical decay, we may prefer to use the latter Q estimate for southwestern Australia, which corresponds to the $r_{\mathrm{hyp}}{}^{-1}$ geometrical spreading model adopted by Douglas *et al.* (2013). This frequency-dependent Q model (eq. 2) corresponds to 457 at 1 Hz and 1503 at 25 Hz, which could foreseeably be accommodated through frequency-dependent weighting of the Douglas *et al.* (2013) models developed with frequency-independent Q. Stress drop terms derived by Allen *et al.* (2006) lie between 1 and 100 bars (0.1 and 10 MPa) but show a trend that increases with magnitude, albeit weakly.

To provide further estimates of the parameters, without the numerous assumptions that may be required by using values from the literature, we here also present the average Q, $\Delta\sigma$ and κ_0 values as determined from subsets of the data used for this study. We test the impact of using 10, 25 and 50 per cent of all data to simulate the effect of a developing database for a new network installation.

Kappa estimation

We estimated site κ_0 consistent with the models of Douglas *et al.* (2013) through least-squares minimization of spectra computed over the duration of shaking. The window duration was based on 5–95 per cent of the velocity-squared integral, with spectral models fit between 10 and 100 Hz in the lin–log domain. The models were based on the Brune (1970) source with single-event common corner frequency and attenuation modelled using an exponential function (Anderson & Hough 1984); refer to Douglas *et al.* (2013) and Edwards *et al.* (2011) for an exhaustive description of the fitting procedure. At least two instruments were required to have recorded each event to include it in the processing. To estimate the impact of limited data on the choice of model weights, we simulated different stages of data collection through random sampling of the events for which we had good recordings at two or more instruments.

Since the hypocentral distance was very limited, choice of Q has a minimal impact on kappa: we chose $Q = 1200$ to be broadly consistent with the results of Allen *et al.* (2006), who found values of between 457 at 1 Hz and 1503 at 25 Hz. However, choosing the $Q = 600$ model from Douglas *et al.* (2013) would only have an impact of approximately $\Delta\kappa_0 = 0.001$s (2–4 per cent). We observe that the kappa values determined for the Cooper Basin stations were dependent on the location of the borehole. Instrument McLeod #1, located at the centre of the network (1.8 km depth) has the lowest value (0.028 s), consistent with the significant depth at which it is located. Stations MW01, MW02 and MW03 (the inner ring) lie within approximately 2.5 km of the central station (depths 109–357 m) and show the highest kappa values (0.041–0.047 s). Stations WA01, WA02 and WA03 (outer ring) lie within 5 km of the central

Table 1. Kappa (and the standard deviation $\pm\sigma$; and standard error $\pm\sigma_e$) determined for the sites of Cooper Basin, along with the impact of reduced data sets.

		100 per cent				50 per cent			25 per cent			10 per cent		
Station	Depth (m)	κ (s)	$\pm\sigma$ (s)	#	$\pm\sigma_e$ (s)	$\Delta\{\kappa_0\}$ (per cent)	$\Delta\{\pm\sigma\}$ (per cent)	$\Delta\{\#\}$ (per cent)	$\Delta\{\kappa_0\}$ (per cent)	$\Delta\{\pm\sigma\}$ (per cent)	$\Delta\{\#\}$ (per cent)	$\Delta\{\kappa_0\}$ (per cent)	$\Delta\{\pm\sigma\}$ (per cent)	$\Delta\{\#\}$ (per cent)
MCL	1783	0.0277	0.0133	113	0.0013	0.31	−2.23	−49	1.16	0.22	−76	−1.09	−13.89	−91
WA01	97.5	0.0335	0.0159	48	0.0023	−0.27	−1.47	−51	2.51	−3.36	−76	−4.20	−22.33	−90
WA02	96	0.0371	0.0136	49	0.0019	0.81	−3.17	−51	0.28	−10.13	−77	0.30	−30.00	−91
WA03	110	0.0307	0.0158	81	0.0018	0.28	−0.58	−49	1.28	−2.46	−75	1.04	−9.48	−90
WA04	97.5	0.0380	0.0149	109	0.0014	−0.05	−0.10	−50	0.26	−1.10	−76	3.34	−7.47	−90
MW01	357	0.0470	0.0128	161	0.0010	0.16	−1.80	−50	0.47	−2.81	−76	0.14	−7.29	−90
MW02	109	0.0405	0.0217	223	0.0015	0.05	−0.80	−50	−0.65	−1.09	−76	−1.55	−3.41	−90
MW03	239	0.0469	0.0134	194	0.0010	0.03	−1.09	−49	0.60	−0.90	−76	0.45	−4.16	−90

Table 2. Statistical analysis of stress parameter. All values in natural log-scale. $<\ln[\Delta\sigma]>$ is the ln-average stress parameter and $\sigma_{\ln[\Delta\sigma]}$ is the standard deviation of the individual event stress parameters; $\sigma_{<\ln[\Delta\sigma]>}$ is the standard deviation of the mean and $\sigma_{\sigma\ln[\Delta\sigma]}$ is the standard deviation of the standard deviation over 1000 randomizations.

Per cent of data	$<\ln[\Delta\sigma]>$ (bars)	$<\Delta\sigma>$ (bars)	$\sigma_{<\ln[\Delta\sigma]>}$	$\sigma_{\ln[\Delta\sigma]}$	$\sigma_{\sigma\ln[\Delta\sigma]}$
100	2.93	18.67	0.000	0.502	0.000
75	2.91	18.40	0.024	0.497	0.015
50	2.92	18.56	0.041	0.528	0.055
25	2.93	18.78	0.111	0.502	0.110
10	2.78	16.20	0.163	0.390	0.076

station and show moderate values of kappa (0.03–0.037 s), despite being located shallower than stations MW01, MW02 and MW03 (at depths of 96–110 m). The similarity of the kappa values for each ring of the network is remarkable, and may be due to similar geology for these stations or, since the earthquakes are all located near to the well-head, due to the similarity of the propagation paths to each of the stations of a given ring.

We estimated the impact of reduced data sets by bootstrapping 100 times over random subsamples of the complete data set; measuring the changes in absolute value and scatter. The impact of a reduced data set (even down to 10 per cent of the original events: corresponding to an average of 14 events) was minimal in terms of the average kappa, with changes of only a few per cent (Table 1). In the case of the standard deviation, the reduced data sets led to significant underestimation of the true uncertainty. This should not be an issue in our application, however, since we are interested in the median values; aleatory variability is independently assigned based on the work of Douglas *et al.* (2013).

Stress parameter estimation

Source corner frequencies of the spectra were re-estimated, fixing $Q = 1200$ and κ_0 as in Table 1. An inversion was performed in the log–log domain, again minimizing the least-squares misfit of the spectral model. Given the moment magnitude determined by Edwards & Douglas (2013), we can then estimate the stress parameter as

$$\Delta\sigma = M_0 \left(\frac{f_c}{0.4906\beta} \right)^3, \quad (3)$$

where $\beta = 3500\ \mathrm{m\,s^{-1}}$ and M_0 is estimated following the original formulation of the moment magnitude by Hanks & Kanamori (1979).

After selecting events with available M_w and f_c estimates a total of 95 earthquakes from Cooper Basin were assigned stress parameters. The log-average was 19 bars with a standard deviation of 0.5 (ln units; a factor of 1.65). Repeating the analysis with subselections of the events from Cooper Basin we obtain standard deviations on both the mean and standard deviation of the target value (Table 2). We see that even with only 25 per cent of the events, the mean stress parameter $<\ln[\Delta\sigma]>$ (and the variability of individual event stress parameters $\sigma_{\ln[\Delta\sigma]}$) is robust, with a variability (represented by the standard deviation) of a factor of 1.12. Reducing the data set to 10 per cent (around 9–10 events) we begin to observe larger (albeit not significant) deviations from the mean.

Based on the spectral analysis, we can choose to assign weights to the models based on expert judgement. We select models from Douglas *et al.* (2013) with $Q = 600$ and $Q = 1800$ (covering the range observed in the literature for this region), and based on the limited range of measured surface attenuation values, $\kappa = 0.04$ s (Table 1). To cover the 19 bar average seen in the spectral analyses, we then make a further selection of models with 10 and 100 bar stress parameter. This leaves four candidate models (Table 3). No preference in terms of weighting is given to GMPEs based on the two different Q models due to the uncertainty of this parameter (Table 3). The final weighting of the four selected models is then given by 0.365 for the two 10-bar and 0.135 for the two 100-bar models. These weights were chosen such that the log-average equivalent stress parameter of the weighted model was 19 bars: that is,

Table 3. Weighting scheme for the selected models based on Q and $\Delta\sigma$.

Model	Model description	Q weight	$\Delta\sigma$ weight	Total weight
19	$\Delta\sigma = 10$ bar; $Q = 600$; $\kappa = 0.04$ s	0.25	0.365	0.365
23	$\Delta\sigma = 10$ bar; $Q = 1800$; $\kappa = 0.04$ s	0.25	0.365	0.365
31	$\Delta\sigma = 100$ bar; $Q = 600$; $\kappa = 0.04$ s	0.25	0.135	0.135
35	$\Delta\sigma = 100$ bar; $Q = 1800$; $\kappa = 0.04$ s	0.25	0.135	0.135

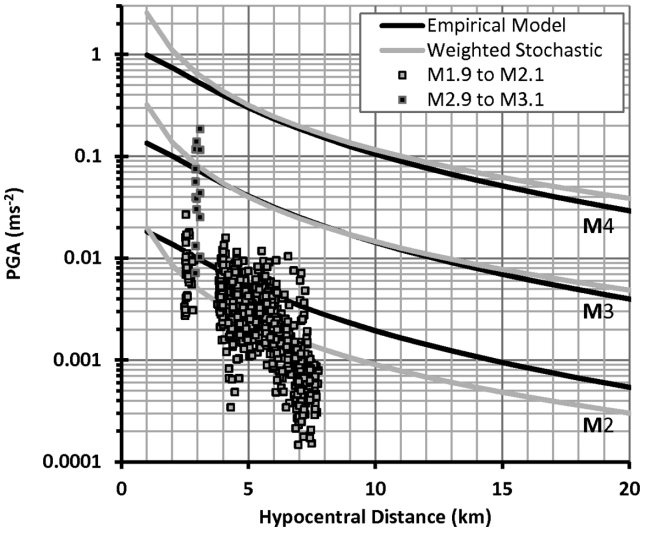

Figure 2. Comparison of median PGA predictions from the weighted stochastic GMPE (with weights based on spectral analysis) and the empirical model of Douglas *et al.* (2013).

$2 \times 0.365 \times \log(10\text{ bar}) + 2 \times 0.135 \times \log(100\text{ bar}) \approx \log(19\text{ bar})$. The resulting weighted stochastic GMPE is shown in Fig. 2 for PGA along with predictions from the purely empirical model of Douglas *et al.* (2013) and recorded data. The weighted stochastic model shows better fit to the recorded data from small events, while for the few events with $M \approx 3$, both models predict similar motions.

RESIDUAL ANALYSES

In this section, we analyse residuals computed from the 36 ground-motion models and data from Cooper Basin. 2089 pairs of horizontal time histories were available for this analysis from the eight local stations and 427 earthquakes. From these time histories, the geometric means of the pseudo-spectral accelerations (PSAs) for 5 per cent damping from each pair are computed for 0.01 s (100 Hz, assumed equal to PGA) and 0.05 s (20 Hz) natural periods. The limited bandwidth of the seismometers installed at Cooper Basin does not allow accurate PSAs to be computed for frequencies lower than about 15 Hz. Consequently, the engineering use of these data is limited to examining the response to shaking of stiff structures (e.g. low-rise masonry buildings) and non-structural elements. The observed and predicted response spectra shown by Douglas *et al.* (2013) show that PSA(0.05 s) is likely to be close to the peak PSA for the magnitude range covered by the Cooper Basin data. In this section, these two sets of PSAs are statistically compared to the 36 ground-motion models proposed by Douglas *et al.* (2013). These models consist of: the equations for the median PSAs derived using the stochastic method and the aleatory-variability models for the single-station within-event standard deviation (φ_{SS}) (e.g. Al Atik *et al.* 2010) and the zone-specific between-event standard deviation (τ_{ZS}) equal to the average of the models for this variability for the Soultz and Basel EGSs. The impact of changing the model for the aleatory variability is investigated below.

Scherbaum *et al.* (2009) and Kale & Akkar (2013) propose methods to judge the applicability of GMPEs to a given set of ground-motion data. These methods consist of the calculation of variables: log-likelihood (LLH, Scherbaum *et al.* 2009) and Euclidian distance ranking (EDR, Kale & Akkar 2013), which are both based on the differences between the natural logarithms of the observed and predicted PSAs, although the influence of the aleatory variability is different in the two cases. Because we are assuming the same aleatory variability for all tested models we could simply use the differences in mean residuals to rank the models but the use of the LLH values allows us to weight the different models in a mathematically rigorous way (see below). The lower the value of LLH and EDR, the closer the match between the observations and predictions. The large number of records available from Cooper Basin enables us to investigate the impact of the number of records available on the results of the GMPE testing. Generally, geothermal projects will have fewer records available with which to judge the applicability of the available GMPEs (especially before stimulation or early on in the stimulation phase) and hence it is useful to study whether the GMPE testing is sensitive to the number of records used.

Based on comparisons between the LLHs and EDRs (and the implied ranking of models) computed for PSA(0.01 s) and PSA(0.05 s) it was found that LLH and EDR are strongly linearly correlated for both periods, as are these values for the two structural periods. Therefore, for brevity, in the rest of this section only LLHs (and GMPE ranking and logic-tree weights derived from these values) for PSA(0.01 s) are presented. LLH is preferred to EDR as a measure of the applicability of GMPEs because of the direct link between a set of LLHs and logic-tree weights in the case of well-distributed data and similar model extrapolation behaviour (as is the case for physically based stochastic models). About a third of the seismograms from Cooper Basin required high-cut filters that removed ground motions with periods below 0.05 s (frequencies above 20 Hz), which could be affecting PSA(0.01 s). However, repeating the analyses described below with and without these seismograms showed that the influence of these band-limited records on the results is minimal. Therefore, all 2089 geometric-mean observations of PSA(0.01 s) were used, from which normalized residuals were computed. For the total variability, we use a value of 0.96 (ln units), the value computed from the estimates of φ and τ given by Douglas *et al.* (2013).

To test the stability of the LLH values with respect to the number of records being used, a bootstrap procedure was followed whereby 100 random sets of 1044 (half the total), 522 (quarter of the total) and 261 (eighth of the total) samples are selected from the 2089 available and the analysis repeated. From these results, the mean and standard deviation (from the 100 results) of the LLH of each ground-motion model were computed (Table 4). As expected, the standard deviation increases as the number of available records decreases. Surprisingly, however, the standard deviations remain low and consequently the LLHs are stable, even when only an eighth of the records are used. This suggests that even a few hundred seismograms would enable robust logic-tree weights to be computed for hazard assessments of EGS projects if it were assumed that the highest weighted models apply for magnitudes and distances outside the range covered by observations.

It is interesting to note that the best-performing models are for values of $\Delta\sigma$, Q and κ_0 similar to those previously reported for the Cooper Basin area or calculated earlier. This suggests that logic-tree weights can be preliminarily assessed based on values of these key parameters taken from the literature (if they are mutually consistent) or from seismological analyses of data, without statistically comparing the observations and predictions.

The definition of LLH allows a direct computation of logic-tree weights (Scherbaum *et al.* 2009). Such an approach is not necessarily appropriate in terms of a probabilistic seismic hazard assessment, however, because the weights do not represent the probability of a

Table 4. Mean LLHs and their standard deviations (computed using a bootstrapping procedure) for the 36 ground-motion models and various proportions of the Cooper Basin data set. Models indicated in bold are the 12 best performing models whereas those in italics are the 12 worst performing.

$\Delta\sigma$	Q	κ_0	All data	Half	Quarter	Eighth	No.
1	200	0.005	2.1007	2.0969 ± 0.0228	2.0978 ± 0.0407	2.0992 ± 0.0622	1
1	**200**	**0.02**	**1.7971**	**1.7965 ± 0.0129**	**1.7982 ± 0.0233**	**1.7988 ± 0.0335**	**2**
1	200	0.04	2.6115	2.6129 ± 0.0255	2.6157 ± 0.0464	2.6159 ± 0.0705	3
1	*200*	*0.06*	*3.7752*	*3.7780 ± 0.0384*	*3.7816 ± 0.0694*	*3.7818 ± 0.1068*	*4*
1	*600*	*0.005*	*3.7820*	*3.7755 ± 0.0438*	*3.7750 ± 0.0787*	*3.7788 ± 0.1238*	*5*
1	**600**	**0.02**	**1.7664**	**1.7648 ± 0.0147**	**1.7663 ± 0.0257**	**1.7672 ± 0.0381**	**6**
1	600	0.04	2.2444	2.2453 ± 0.0205	2.2482 ± 0.0372	2.2479 ± 0.0574	7
1	600	0.06	3.2774	3.2798 ± 0.0337	3.2836 ± 0.0609	3.2831 ± 0.0955	8
1	*1800*	*0.005*	*5.2068*	*5.1987 ± 0.0584*	*5.1963 ± 0.1035*	*5.2031 ± 0.1685*	*9*
1	**1800**	**0.02**	**1.8220**	**1.8200 ± 0.0172**	**1.8213 ± 0.0299**	**1.8224 ± 0.0453**	**10**
1	1800	0.04	2.1426	2.1432 ± 0.0190	2.1461 ± 0.0344	2.1457 ± 0.0533	11
1	1800	0.06	3.1200	3.1222 ± 0.0321	3.1261 ± 0.0580	3.1253 ± 0.0916	12
10	*200*	*0.005*	*4.5300*	*4.5227 ± 0.0477*	*4.5232 ± 0.0891*	*4.5248 ± 0.1319*	*13*
10	**200**	**0.02**	**2.0125**	**2.0091 ± 0.0223**	**2.0091 ± 0.0417**	**2.0113 ± 0.0631**	**14**
10	**200**	**0.04**	**1.7867**	**1.7857 ± 0.0125**	**1.7865 ± 0.0226**	**1.7889 ± 0.0289**	**15**
10	200	0.06	2.4931	2.4938 ± 0.0242	2.4953 ± 0.0427	2.4981 ± 0.0603	16
10	*600*	*0.005*	*8.9585*	*8.9475 ± 0.0758*	*8.9457 ± 0.1397*	*8.9512 ± 0.2162*	*17*
10	600	0.02	2.7308	2.7261 ± 0.0348	2.7252 ± 0.0636	2.7281 ± 0.0997	18
10	**600**	**0.04**	**1.7389**	**1.7372 ± 0.0136**	**1.7378 ± 0.0254**	**1.7398 ± 0.0357**	**19**
10	600	0.06	2.1503	2.1504 ± 0.0189	2.1520 ± 0.0331	2.1541 ± 0.0462	20
10	*1800*	*0.005*	*1.9403*	*11.9273 ± 0.0952*	*11.9222 ± 0.1716*	*11.9328 ± 0.2782*	*21*
10	*1800*	*0.02*	*3.1481*	*3.1427 ± 0.0406*	*3.1414 ± 0.0734*	*3.1450 ± 0.1168*	*22*
10	**1800**	**0.04**	**1.7710**	**1.7690 ± 0.0156**	**1.7695 ± 0.0290**	**1.7715 ± 0.0424**	**23**
10	1800	0.06	2.0488	2.0487 ± 0.0171	2.0503 ± 0.0300	2.0522 ± 0.0416	24
100	*200*	*0.005*	*8.3257*	*8.3160 ± 0.0695*	*8.3176 ± 0.1313*	*8.3168 ± 0.1878*	*25*
100	200	0.02	3.0079	3.0031 ± 0.0367	3.0021 ± 0.0695	3.0041 ± 0.1045	26
100	**200**	**0.04**	**1.7318**	**1.7298 ± 0.0143**	**1.7292 ± 0.0283**	**1.7324 ± 0.0398**	**27**
100	**200**	**0.06**	**2.0443**	**2.0440 ± 0.0175**	**2.0442 ± 0.0306**	**2.0482 ± 0.0384**	**28**
100	*600*	*0.005*	*6.2242*	*16.2098 ± 0.1033*	*16.2079 ± 0.1929*	*16.2123 ± 0.2911*	*29*
100	*600*	*0.02*	*4.5644*	*4.5579 ± 0.0529*	*4.5557 ± 0.0985*	*4.5585 ± 0.1523*	*30*
100	**600**	**0.04**	**1.9695**	**1.9666 ± 0.0222**	**1.9656 ± 0.0430**	**1.9685 ± 0.0650**	**31**
100	**600**	**0.06**	**1.8382**	**1.8374 ± 0.0136**	**1.8374 ± 0.0246**	**1.8408 ± 0.0298**	**32**
100	*1800*	*0.005*	*1.2923*	*21.2752 ± 0.1278*	*21.2684 ± 0.2330*	*21.2802 ± 0.3724*	*33*
100	*1800*	*0.02*	*5.4541*	*5.4468 ± 0.0607*	*5.4438 ± 0.1120*	*5.4475 ± 0.1760*	*34*
100	1800	0.04	2.1350	2.1317 ± 0.0260	2.1305 ± 0.0499	2.1333 ± 0.0766	35
100	**1800**	**0.06**	**1.7930**	**1.7920 ± 0.0130**	**1.7919 ± 0.0242**	**1.7952 ± 0.0299**	**36**

given model being correct (Delavaud *et al.* (2012). Instead, LLH-based weights represent a given model's ability to fit the observed data, favouring 'better models'. Typically, the goal of a complete logic-tree based hazard analysis is to capture not only the centre and body, but also the range of possibilities. In this sense, the LLH weights, by design, will not cover the entire range (i.e. extreme scenarios not yet recorded). To bypass this limitation, Delavaud *et al.* (2012) suggest weighting based on expert judgement, with help from LLH information. However, a fully transparent approach for logic-tree weighting still does not exist. Considering this limitation, we adopt the LLH-based logic-tree weights for this analysis. Nevertheless, for the purpose of hazard assessment in geothermal zones, we would recommend further expert elicitation to ensure that the complete range of possible models are appropriately considered in the logic tree.

Applying the equation of Scherbaum *et al.* (2009) to the values of LLH listed in Table 4 for the complete data set gives the weights summarized in Fig. 3, from which it can be seen that roughly half of the models contribute about 75 per cent of the total weight. Modifying the standard deviation associated with each model from 0.96 [the sigma proposed by Douglas *et al.* (2013), for use with the stochastic models] to 0.64 (the sigma obtained by regression analyses on the Cooper Basin data, see later) does not alter the model ranking but it slightly increases the distinction between models. Therefore, the model weights are concentrated in the best ranked models (75 per cent of the weight is contributed by roughly one-third of the models). In terms of computational efficiency, the use of 36 models in a logic tree may be problematic. In this case, it may be useful to trim the number of models from the total 36 before weighting. As discussed previously, LLH weighting tends to favour models which better predict the data. By using the LLH weighting of the full set of models it should, therefore, be possible to find and remove models that do not add any information about possible epistemic uncertainty. For instance, if the highest and second-highest weighted models predict very similar ground motions, then the second model can be removed without affecting the hazard results. Such analysis is, however, beyond the scope of this paper.

ALEATORY VARIABILITY

To assess the aleatory variability of the Cooper Basin ground-motion data, random-effects regression was performed using the functional form of model 1 of Douglas *et al.* (2013):

$$\ln\ \mathrm{PSA}(0.01\,\mathrm{s}) = a + bM_{\mathrm{w}} + c\ln\sqrt{r_{\mathrm{hyp}}^2 + h^2} + dr_{\mathrm{hyp}}, \qquad (4)$$

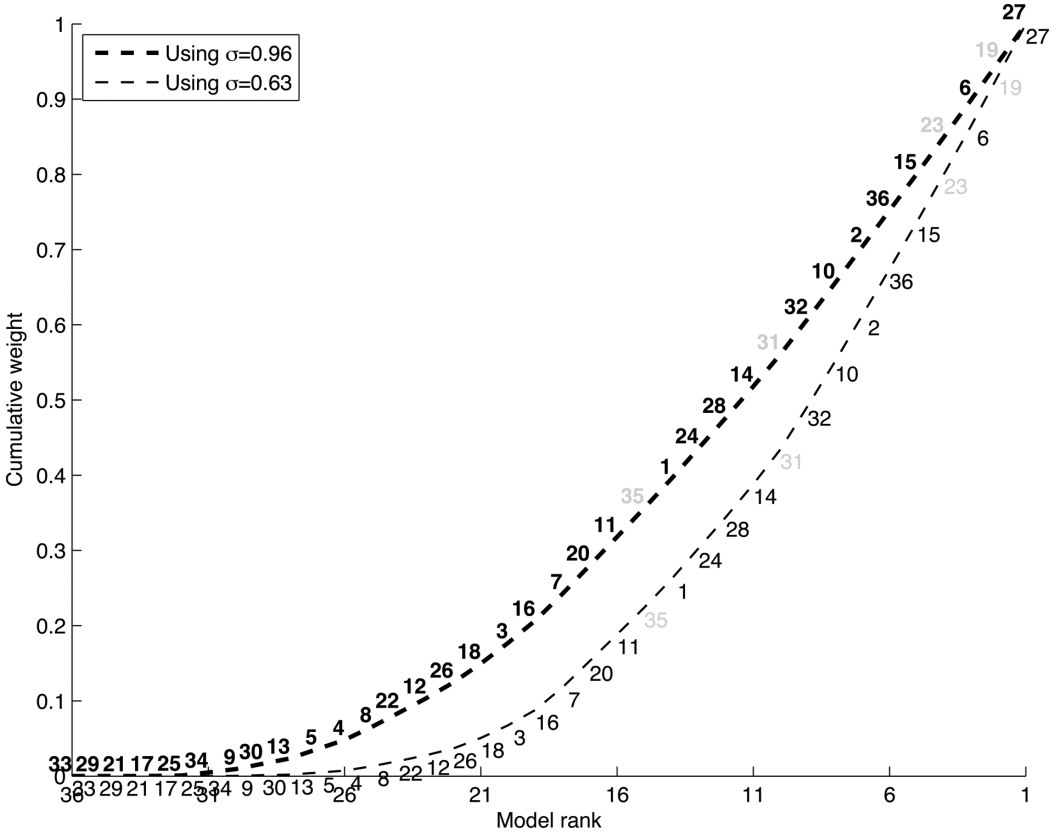

Figure 3. Ranking of models against cumulative weights for two different sigmas. See Tables 4 and 2 for the correspondence between model number and parameters of the stochastic model. Grey numbers indicate models consistent with the spectral analysis and literature (Q equal to 600–1800; $\Delta\sigma$ equal to 10–100 bars and $\kappa_0 = 0.04$ s).

where PSA(0.01 s) is in m s^{-2}, M_w is moment magnitude and a, b, c, d and h are regression coefficients. The limited distance range of the Cooper Basin data does not allow robust estimates of h (describing near-source saturation) and d (describing anelastic attenuation) to be found and, therefore, these coefficients are constrained to zero. The coefficients obtained from using the entire 2089 records are: $a = -6.899$, $b = 2.569$ and $c = -2.589$ with a between-event standard deviation (τ) of 0.099 and a within-even standard deviation (φ) of 0.627, leading to an overall standard deviation (σ) of 0.635. Comparing these coefficients to those obtained by Douglas *et al.* (2013) from regression on data from six areas indicates slightly higher magnitude dependence (coefficient b) and faster attenuation (coefficient c) for Cooper Basin data. The most notable difference, however, is the much smaller values of φ and, in particular, τ, from regression on the Cooper Basin data compared to those obtained by Douglas *et al.* (2013). The much lower value of τ can be partly explained by the use of data from a single zone (this type of τ is called τ_{ZS} by Douglas *et al.* 2013) but it appears that ground motions at Cooper Basin are much less variable than those at other EGS sites; Douglas *et al.* (2013) found for Basel $\tau_{ZS} = 0.637$ and for Soultz $\tau_{ZS} = 0.902$.

The value of τ for Cooper Basin is even lower than those associated with GMPEs derived from moderate and large earthquakes [see fig. 10(c) of Douglas *et al.* (2013), where τ for none of the considered GMPEs is lower than 0.2]. This very small τ can be related to the small variability in the stress drops of Cooper Basin: found to be 0.5 ln units, corresponding to a factor of 1.65. This is significantly lower than most other studies: for instance, Allmann & Shearer (2009) find a value of 1.46 ln units for global intraplate events; Edwards & Fäh (2013) find 1.83 and 1.43 ln units for the Swiss foreland and Alpine regions, respectively; Oth *et al.* (2010) find 1.38 ln units for Japanese earthquakes and Rietbrock *et al.* (2013) find 1.38 ln units for UK events. Cotton *et al.* (2013) showed that spectral analysis methods applied to small earthquakes often led to significantly larger stress variability than seen in larger events. They presented the required variability in stress parameter corresponding to the aleatory variability of several GMPEs and concluded that the variability should lie between 0.26 and 0.59 ln units for large events. The difference may relate to the strong regional variability in source, path and site effects for small earthquakes, either real or due to parameter trade-off. Treated independently, the data from Cooper Basin (a limited source zone) and the consistency of wave propagation may mean that the observed variability usually inherent with such small events is not apparent. One issue to consider, in this case, is whether the observed variability truly reflects the possible future variability: that is, do we account for a near-surface event outside of the seismic cloud as was the case for the largest event related to the Berlín (El Salvador) geothermal project (Bommer *et al.* 2006), or of significantly different stress drop?

Single-station φ was calculated for the eight Cooper Basin stations as $\varphi_{SS} = 0.493$, which is again smaller than that obtained by Douglas *et al.* (2013), $\varphi_{SS} = 0.576$. The relatively low value is consistent with the limited range of station kappa values (with three distinct groups corresponding to the middle, inner and outer stations). The total sigma (combining the between- and within-event variability) is, therefore, much lower than that for the empirical models of Douglas *et al.* (2013) and it is more in line with those associated with GMPEs for moderate and large earthquakes. This

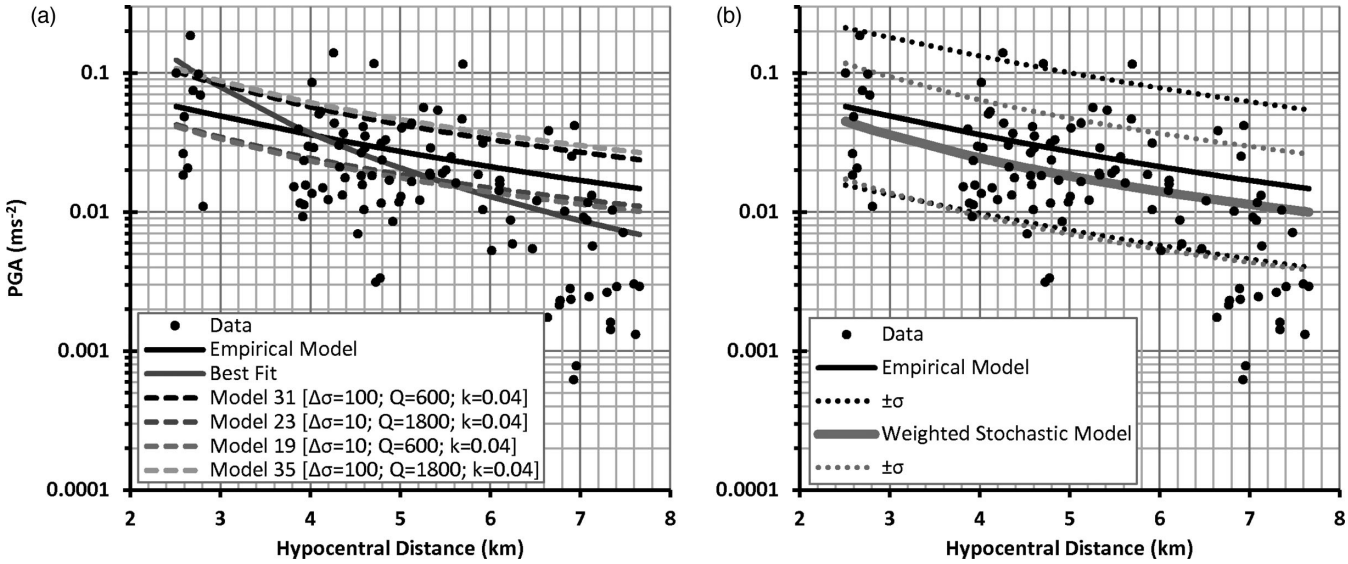

Figure 4. (a) PGA data for Cooper Basin for events with $2.5 < M < 3.1$ and the selected predictions (stochastic and empirical) from Douglas *et al.* (2013) for M 2.8 along with the best fit of the data to eq. (3). (b) The empirical model of Douglas *et al.* (2013) along with the weighted median stochastic model (from LLH testing) including their variabilities.

demonstrates that at least some of the large variability in the empirical models of Douglas *et al.* (2013) is due to mixing data from various sites when deriving these models.

REDUCTION OF EPISTEMIC UNCERTAINTY

As highlighted by Douglas *et al.* (2013) the disadvantages of applying their empirical GMPE are twofold. First, the limited magnitude–distance range means that application to rarer, but potentially damaging events, is tenuous and, secondly, the aleatory variability assigned to their equation was strongly contaminated by epistemic uncertainty from combining several regional data sets (e.g. due to differences in seismicity and attenuation). Effectively, the empirical model can be thought of as a mixture model: comprising several different sets of source and propagation behaviour, but without consideration of the increased sigma relative to a predictive relation. Nevertheless, it is not trivial to isolate such effects given limited recordings. Douglas *et al.* (2013) suggest that to reduce the uncertainty, stochastic simulation models can be used. Of course, such models are not without uncertainty outside their 'calibrated' model-space: the magnitude and distance range over which the simulation model can be tested against recorded seismograms. However, unlike empirical models, due to their physical basis, alternative models can be easily developed to cover the epistemic uncertainty outside the magnitude range available in instrumental databases. For this purpose, Douglas *et al.* (2013) provided 36 GMPEs to cover a range of simulation parameters: with various κ, Q and $\Delta\sigma$. A further benefit of testing and weighting simulation models, as performed here for Cooper Basin, is that it can help to limit the influence of epistemic uncertainty contamination related to mixing different sites.

The analysis undertaken here showed that the stochastic models can be selected based on spectral analysis, or on LLH testing. Four stochastic models performed well in LLH testing, whist also having stochastic model parameters consistent with the results of spectral analysis: Models 31, 23, 19 and 35. These models, along with the empirical model of Douglas *et al.* (2013) and the best-fitting empirical model just using the Cooper Basin data are shown in Fig. 4(a). We see that the preferred stochastic models are similar to the empirical model based on Cooper Basin data (eq. 4), while the best-fitting model for the data set shows a slightly faster decay. By producing weights following Scherbaum *et al.* (2009) the resulting weighted median stochastic model (from all 36 weighted component models) is shown in Fig. 4(b). The weighted median model leads to lower PGA at $M < 3$ (up to a factor of 1.5) than the empirical model of Douglas *et al.* (2013), while the aleatory variability is reduced by over 25 per cent.

CONCLUSIONS

In this study, we have taken an existing EGS site as a case study of the proposals by Majer *et al.* (2012) and Douglas *et al.* (2013) to characterize the expected ground motions. To simulate the realistic temporal development of a seismic database and corresponding earthquake catalogue, we randomly resampled the full databases to one-half, one-quarter and one-eighth of the full data set. Following the approach of Douglas *et al.* (2013), we then developed weighted median models to describe the subdata sets. It was found that both spectral analysis for the stochastic model parameters and residual analysis provided complementary results, with the highest weighted models from Douglas *et al.* (2013) consistent with both existing literature and values determined here. Using the LLH method, we were able to automatically assign weights using a consistent and transparent approach. The resulting models were shown to significantly reduce the epistemic uncertainty related to ground-motion prediction for EGS projects.

ACKNOWLEDGEMENTS

This study was partially funded by GEISER (Geothermal Engineering Integrating Mitigation of Induced Seismicity in Reservoirs) project funded under contract 241321 of the EC-Research Seventh Framework Programme (FP7). We thank Geodynamics Ltd (Australia) and QCon (Germany) for the data from Cooper Basin. We thank Julian Bommer and an anonymous reviewer for their detailed comments on an earlier version of this paper.

REFERENCES

Al Atik, L., Abrahamson, N., Bommer, J.J., Scherbaum, F., Cotton, F. & Kuehn, N., 2010. The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.,* **81**(5), 794–801.

Allen, T.I., Dhu, T., Cummins, P.R. & Schneider, J.F., 2006. Empirical attenuation of ground-motion spectral amplitudes in southwestern Western Australia, *Bull. seism. Soc. Am.,* **96,** 572–585.

Allen, T.I., Cummins, P.R., Dhu, T. & Schneider, J.F., 2007. Attenuation of ground-motion spectral amplitudes in southeastern Australia, *Bull. seism. Soc. Am.,* **97,** 1279–1292.

Allmann, B.P. & Shearer, P.M., 2009. Global variations of stress drop for moderate to large earthquakes, *J. geophys. Res.,* **114,** B01310, doi:10.1029/2008JB005821.

Anderson, J.G. & Hough, S.E., 1984. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. seism. Soc. Am.,* **74**(5), 1969–1993.

Baisch, S., Weidler, R., Vörös, R., Wyborn, D. & de Graaf, L., 2006. Induced seismicity during the stimulation of a geothermal HFR reservoir in the Cooper Basin, Australia, *Bull. seism. Soc. Am.,* **96**(6), 2242–2256.

Baisch, S., Vörös, R., Weidler, R. & Wyborn, D., 2009. Investigation of fault mechanisms during geothermal reservoir stimulation experiments in the Cooper Basin, Australia, *Bull. seism. Soc. Am.,* **99,** 148–158.

Bommer, J.J., Oates, S., Cepeda, J.M., Lindholm, C., Bird, J., Torres, R., Marroqun, G. & Rivas, J., 2006. Control of hazard due to seismicity induced by a hot fractured rock geothermal project, *Eng. Geol.,* **83,** 287–306.

Boore, D.M., 2003. Simulation of ground motion using the stochastic method, *Pure appl. Geophys.,* **160**(3–4), 635–676.

Brune, J.N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. geophys. Res.,* **75**(26), 4997–5009.

Cotton, F., Archuleta, R. & Causse, M., 2013. What is sigma of the stress drop? *Seismol. Res. Lett.,* **84,** 42–48.

Delavaud, E. *et al.*, 2012. Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe, *J. Seismol.,* **16,** 451–473.

Douglas, J. & Jousset, P., 2011. Modeling the difference in ground-motion magnitude-scaling in small and large earthquakes, *Seism. Res. Lett.,* **82,** 504–508.

Douglas, J. *et al.*, 2013. Predicting ground motion from induced earthquakes in geothermal areas, *Bull. seism. Soc. Am.,* **103**(3), 1875–1897.

Edwards, B. & Douglas, J., 2013. Magnitude scaling of induced earthquakes, *Geothermics,* in press.

Edwards, B. & Fäh, D., 2013. A stochastic ground-motion model for Switzerland, *Bull. seism. Soc. Am.,* **103**(1), 78–98.

Edwards, B., Fäh, D. & Giardini, D., 2011. Attenuation of seismic shear wave energy in Switzerland, *Geophys. J. Int.,* **185,** 967–984.

Giardini, D., 2009. Geothermal quake risks must be faced, *Nature,* **462,** 848–849.

Groos, J., Zeiß, J., Grund, M. & Ritter, J., 2013. Microseismicity at two geothermal power plants in Landau and Insheim in the Upper Rhine Graben, Germany, *Geophys. Res. Abstr.,* **15,** EGU2013–EGU2742.

Hanks, T.C. & Kanamori, H., 1979. A moment magnitude scale, *J. geophys. Res.,* **84**(B5), 2348–2350.

Kale, Ö. & Akkar, S., 2013. A new procedure for selecting and ranking ground-motion prediction equations (GMPEs): the Euclidean-distance based ranking (EDR) method, *Bull. seism. Soc. Am.,* **103**(2A), 1069–1084.

Majer, E., Nelson, J., Robertson-Tait, A., Savy, J. & Wong, I., 2012. Protocol for addressing induced seismicity associated with Enhanced Geothermal Systems. U.S. Department of Energy, Technical Report: DOE/EE-0662.

Oth, A., Bindi, D., Parolai, S. & DiGiacomo, D., 2010. Earthquake scaling characteristics and the scale-(in)dependence of seismic energy-to-moment ratio: insights from KiK-net data in Japan, *Geophys. Res. Lett.,* **37,** L19304, doi:10.1029/2010GL044572.

Rietbrock, A., Strasser, F. & Edwards, B., 2013. A stochastic earthquake ground-motion prediction model for the United Kingdom, *Bull. seism. Soc. Am.,* **103**(1), 57–77.

Scherbaum, F., Delavaud, E. & Riggelsen, C., 2009. Model selection in seismic hazard analysis: an information-theoretic perspective, *Bull. seism. Soc. Am.,* **99**(6), 3234–3247.

Stang, H., 2011. Documentation of the Habanero Seismic Monitoring Network. Technical report #SMR-GDY054, Geodynamics Ltd, Australia.