RESEARCH

Regionally adjusted stochastic earthquake ground motion models, associated variabilities and epistemic uncertainties

**Jaleena Sunny · Marco de Angelis ·
Benjamin Edwards**

Abstract An optimisation-based calibration technique, using the area metric, is applied to determine the input parameters of a stochastic earthquake-waveform simulation method. The calibration algorithm updates a model prior, specifically an estimate of a region's seismological (source, path and site) parameters, typically developed using waveform data, or existing models, from a wide range of sources. In the absence of calibration, this can result in overestimates of a target region's ground motion variability, and in some cases, introduce biases. The proposed method simultaneously attains optimum estimates of median, range and distribution (uncertainty) of these seismological parameters, and resultant ground motions, for a specific target region, through calibration of physically constrained parametric models to local ground motion data. We apply the method to Italy, a region of moderate seismicity, using response spectra recorded in the European Strong Motion (ESM) dataset. As a prior, we utilise independent seismological models developed using strong motion data across a wider European context. The calibration obtains values of each seismological parameter considered (such as, but not limited to, quality factor, geometrical spreading and stress drop), to develop a suite of optimal parameters for locally adjusted stochastic ground motion simulation. We consider both the epistemic and aleatory variability associated with the calibration process. We were able to reduce the area metric (misfit) value by up to 88% for the simulations using updated parameters, compared to the initial prior. This framework for the calibration and updating of seismological models can help achieve robust and transparent regionally adjusted estimates of ground motion, and to consider epistemic uncertainty through correlated parameters.

Keywords Stochastic modelling · Calibration · Area metric · Epistemic uncertainty · Aleatory uncertainty

Highlights
• The article introduces a stochastic model calibration algorithm to obtain optimised input parameters for the given data.
• The algorithm significantly improves the accuracy and reliability of the models.
• The article provides a set of models rather than a single model considering the epistemic uncertainty in the process..

J. Sunny (✉) · B. Edwards
Jane Herdman Building, School of Environmental Sciences, University of Liverpool, Liverpool L69 3GP, UK
jaleenasunny@gmail.com

M. de Angelis
Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, UK

1 Introduction

The modelling of earthquake ground motion is considered a key component within the framework of probabilistic seismic hazard analysis (PSHA). Subse-

quent assessments, such as risk analyses, therefore also rely strongly on the appropriate modelling of earthquake ground motion, both in terms of expected shaking and its variability. Often, ground motion models (GMMs), developed for predicting ground motion intensities, are derived from data recorded in seismically active regions. The resulting GMMs, often in the form of empirical ground motion prediction equations (GMPEs), are considered robust. However, there are many regions that have insufficient recorded data in the magnitude-distance range required to derive native models without extrapolation. To overcome this lack of data requires the adaption of GMMs developed for data-rich regions (the host) to the seismological environment of the data-poor region. Such adjustments are computed using various methodologies, such as simulation-based approaches (Atkinson and Boore 1995, 2006), the hybrid empirical approach (Campbell 2003; Pezeshk et al. 2011) and the referenced empirical approach (Atkinson 2008).

Ground motion simulation-based approaches have received particular attention in recent years due to increases in computational power, and due to the attractiveness of physically based (thus, arguably more justified) methodologies for direct simulation or model adjustment. Different simulation techniques have been developed to estimate strong ground motion for seismic hazard analysis and modelling future earthquakes. These techniques include the stochastic simulation approach (Hanks and McGuire 1981; Boore 1983, 2003; Beresenev and Atkinson 1997) which is characterised by separating the ground motion model into the source, path and site components of a band-limited stochastic waveform; the composite source modelling technique (Boatwright 1982; Anderson 2015; Atkinson et al. 2009) which uses the fact that an earthquake source is composed by several sub-events; the empirical Green's function technique (Somerville et al. 1991) and many other advanced rupture simulation and wave-propagation techniques (Graves and Pitarka 2010; Olsen and Takedatsu 2015; Mai et al. 2010). Stochastic simulations are widely used to obtain acceleration time series and for the development of ground motion prediction equations (Boore 1983; Atkinson and Boore 1995; Atkinson and Silva 2000; Atkinson and Boore 2006). In applications such as PSHA, simulations from these stochastic processes are required to be reliable and accurate for application in a given (target) region (Atkinson and Boore, 1995), or even at

a specific site. Here we focus on the stochastic simulations using SMSIM — 'FORTRAN Programs for Simulating Ground Motions from Earthquakes' introduced by Boore (1983). This method is based on the work of McGuire and Hanks (1980), which explains the Fourier amplitude spectrum (FAS) of a ground acceleration as the combination of the source, path, and site effects with a random phase spectrum. The physical considerations within this approach are mainly defined through the parametric spectrum for the source model, such as the ω-square model, or various alternatives, and the form of signal attenuation and signal dispersion with distance (Boore 2003).

For reliable simulations, SMSIM, or similar simulation methods, require region-specific models of source, path, and site parameters. These parameters are derived from the region-specific waveform datasets through various techniques, such as spectral decomposition, attenuation tomography, and inversions of Fourier spectra (Bindi and Kotha 2020; Rietbrock et al. 2013; Edwards et al. 2019). Even when considering data-sparse regions, such approaches can be used to develop input for simulation tools, such as SMSIM. However, one of the biggest challenges in this case is to justify magnitude-scaling behaviour, particularly of the source properties: i.e. how representative are the source parameters of small earthquakes for larger ones, for which we have no data? A benefit of the proposed calibration approach is that the more robust parameter scaling (such as magnitude dependence) inherent in a seismological prior developed using data from seismically active regions can remain (unless data indicates otherwise) in the locally re-calibrated models. Furthermore, with the regular integration of new waveform data into an existing dataset, the repeated computation of seismological parameters (such as stress drop) becomes a cumbersome process. Being able to rapidly re-calibrate seismological models (without having to return to the analysis of individual records) for use in simulation techniques therefore clearly provides an advantage, particularly in dynamic systems such as induced seismicity.

In the following, we introduce a data-driven re-calibration algorithm to refine stochastic ground motion simulations, taking advantage of existing seismological models as a prior. The algorithm is based on a validation metric known as the stochastic area metric (AM), which happens to coincide with the Wasserstein distance in the univariate case, as elucidated in de Ange-

lis et al. (2021). This metric was recently used by the authors for ranking empirical ground motion models (Sunny et al. 2021). The software to efficiently calculate the metric has been made available by the authors, see the 'Data and resources' section. The calibration algorithm runs a large number of stochastic ground-motion simulations over varying combinations of simulation parameters, computes the area metric against the empirical data for each combination, and outputs the combination that yields the minimum distance. As a significant advantage over existing approaches, we also explicitly account for the epistemic uncertainty by generating sets of seismological parameter combinations, and corresponding ground motions, which fall within a defined confidence band of the observed data. Effectively, we, therefore, consider the inferential uncertainty in the dataset (or data subset). We consider both aleatory and epistemic uncertainty involved in the calibration process. The aleatory uncertainty (inherent randomness) is explicitly incorporated in the simulation model through a sigma parameter as commonly used in empirical ground motion models, while the epistemic uncertainty (reducible by acquiring more data or with better models) is accounted for by determining not only an 'optimal', but a suite of simulation parameter combinations. Epistemic uncertainty can arise due to many factors. In this study, it arises from (1) limited sample size within a specific model space (e.g. large magnitude, short distance records), and (2) model uncertainty because of the idealised assumptions made by the physics-based solver. The variability observed between simulated and observed data is considered aleatory uncertainty, and thus irreducible (as defined by sigma in empirical ground motion models).

In this study, we use 5% damped response spectral accelerations (PSA) from Italy, along with associated metadata, from the European Strong Motion (ESM) database. We use the duration model of Afshari and Stewart (2016), and the spectral decomposition results for strong motion waveform data in Europe from Bindi and Kotha (2020) as prior for the seismological parameters. The resulting simulations, using this prior, are initially analysed and compared with the observed peak ground acceleration (PGA) data. Simulation parameters are subsequently calibrated to improve the performance of the model. We then extend the model calibration to ground motions at response spectral periods of 0.1, 0.5, 0.8, and 1 s. We finally explore the

epistemic uncertainty of the calibration procedure, and simulation-based GMMs in general. Through consideration of correlated simulation parameters, we present a period-by-period iterative approach that determines a suite of models valid across the analysed period range. The resulting suite of GMMs can be used to account for epistemic uncertainty in applications such as PSHA.

2 Data and methodology

The European Strong Motion (ESM) dataset is used for this study. We have evaluated the flat file supplied by them for this study (Lanzano et al. 2018). Its choice is based on the consideration that it is the most up-to-date dataset available for Europe. It has been developed in the framework of the European project Network of European Research Infrastructures for Earthquake Risk Assessment and Mitigation (NERA) by Luzi et al. (2016). The ESM dataset comprises data recordings from 1077 stations, significantly expanding on its predecessor, the RESORCE dataset (Akkar et al. 2014), which included only 150 stations. We have only considered recordings with given Vs30 values (either in situ or estimated) — i.e. 17,218 recordings in total. Whenever an in situ measured Vs30 was unavailable, we adopted the Vs30 values from the slope proxy according to Wald and Allen (2007). The geometric mean of horizontal PGA and PSA values are used throughout this study. Here, we consider geopolitical boundaries as the basis for the calibration of stochastic models, and specifically, we opted for the country Italy (country code: IT) owing to its moderate seismicity. This offers the possibility to benchmark predictions against moderate and large events relevant to engineering applications. Our IT region dataset consists of 11,341 recordings within a magnitude range of 3.5 to 6.0 at epicentral distances below 552 km. The magnitude range used here is truncated at M_w 6, since SMSIM relies on a point-source assumption, with simple geometric approximations to account for larger ruptures (Atkinson et al. 2009). However, the method can be extended for application to larger events using a suitable finite fault stochastic method, for instance. In a preliminary step, we simulate the PGA values corresponding to the metadata (M, R, depth, dip and focal mechanism) of recordings (Fig. 1) using SMSIM and the seismological model prior, as outlined above.
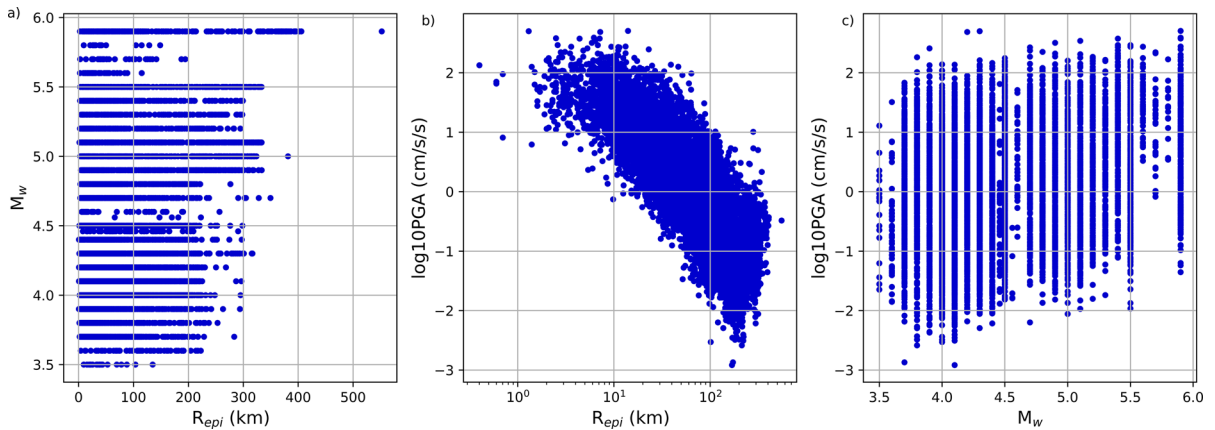
Fig. 1 The distribution of the Italian data used in this study (before down-sampling). (a) Epicentral distance vs M_w. (b) Epicentral distance vs PGA. (c) M_w vs PGA values

2.1 Stochastic simulations using SMSIM

SMSIM (**S**tochastic **M**ethod **Sim**ulation) is a set of FORTRAN programs based on the stochastic method to simulate ground motions from given earthquake parameters (Boore 2003). Its essence is to limit the frequency-band of stochastic time series to match the amplitude spectra, on average, to a target Fourier amplitude spectrum (FAS). In this study, we apply the random vibration theory (Vanmarcke and Lai 1980) simplification of the method in order to rapidly determine peak oscillator response values (i.e. PSA) at different periods. The a priori source, path and site parameters used for the stochastic modelling of this regional dataset are adopted from Bindi and Kotha (2020). Bindi and Kotha (2020) determined seismological parameters for the wider European region using waveform data from the ESM database. They proposed a suite of parameters that describe the FAS of these waveforms, with three region-specific adaptations for attenuation. We utilise the attenuation model applicable to the Italian region for our prior. Duration parameters are taken from Afshari and Stewart (2016), who provide values of significant duration parameters for 5–75% (D75), 5–95% (D95), and 20–80% (D80) of the normalised cumulative Arias intensity. In the case of stochastic simulations, 2(D80) is the most appropriate duration, according to Boore and Thompson (2014); Kolli and Bora (2021), and is therefore adopted for our simulations in SMSIM.

Bindi and Kotha (2020) expressed Brune stress drop ($\Delta\sigma$) for low and high magnitudes separately: 30 bars

for magnitudes below 5.17 and 60 bars for magnitudes above 5.17. We considered the stress drop distribution given by Bindi and Kotha (2020) and Razafindrakoto et al. (2021) adopting a magnitude-dependent stress drop (in Pa) model given by:

$$\log_{10}(\Delta\sigma) = \max[6.34, 6.02 + 0.09266 M_w] \Rightarrow M_w < 5.17 \tag{1}$$

$$\log_{10}(\Delta\sigma) = \min[6.79, 6.5 + 0.35(M_w - 5.17)] \Rightarrow M_w >= 5.17 \tag{2}$$

Baseline (uncalibrated) simulations, using the seismological prior (Table 1), are analysed using the full dataset (Fig. 1), and the residuals (observed–simulated, in $\log_{10}$ scale) are plotted in Fig. 2. A normal distribution is used as a prior model for simplicity. Any other bell-shaped distribution with unbounded support could have been used for the same purpose. The residuals from instruments at rock sites are highlighted in order to understand the effect of site amplification, as we did not directly consider site amplification within the simulations (Fig. 2, *top*). When the rock sites are analysed, the residuals are reasonably well distributed with respect to the zero horizontal line (i.e. are unbiased). However, ground motions at soil sites are all under-predicted, as expected. In order to account for site amplification in the empirical data, we post-process computations from SMSIM using the site amplification component of the Boore et al. (2014) GMPE, taking a reference velocity of 760 m/s and the site-specific Vs_{30}. Although the application of the amplification model has resulted in

Table 1 Brief summary of seismological prior model parameters (as used in SMSIM input)

Parameter	Units	Value	Description	Calibrated and prior used
rho	g/cm3	2.7	Source component: Density of the medium	No
beta	km/s	3.2	Source component: Velocity of the medium	No
prtitn	–	0.707	Source component: Partition factor	No
radpat	–	0.55	Source component: Radiation pattern	No
fs	–	2	Source component: free surface factor	No
Source number	–	1	Source component: Single corner frequency	No
r_ref	km	10	Path component: Reference distance	No
nsegs rlow(i) a_s	–	3 rlow(1) = 10 a_s(1) = -1.35 rlow(2) = 71 a_s(2) = -0.577 rlow(3) = 119 a_s(3) = -1.53	Path component: Geometrical spreading	Yes a_s(1) = $N(-1.35, 0.3)$ a_s(2) = $N(-0.577, 0.1)$ a_s(3) = $N(-1.53, 0.3)$
Qr1 s	–	250.4 0.29	Path component: Quality factor	Yes Qr1 = $N(250.4, 50)$ s = $N(0.29, 0.06)$
nknots rdur(i) dur(i) Slope of last segment	–	3 rdur(1) = 10 dur(1) = 0.0646 rdur(2) = 50 dur(2) = 0.0865 0.0373	Path component: path duration	No
kappa	s	0.02	Site diminution factor	Yes $N(0.02, 0.01)$

Nomenclature is as used in SMSIM. The $N(location, scale)$ provides information on the prior normal distribution of each parameter used in this study

some improvement in residuals, there has not been a significant change (Fig. 2, *bottom*). The mean value of all residuals, which before including site effects was 0.298, reduced to 0.226 using this amplification model.

2.2 Validation and calibration of simulations using the area metric

In order to assess the quality of simulations from SMSIM, and as a basis for subsequent model calibration, we use the area metric. This metric can be used for analysing the absolute shift between the observed data and the corresponding simulations. The area metric defines the area between probability distributions, quantifying how dissimilar they are. This can be considered as a measure of misfit between the marginal distribution of the data and the marginal distribution of the model (Sunny et al. 2021). The AM considers the

observed data and the simulated data as two cumulative distribution functions and computes the area between them. The larger the AM value, the more distant the model and data become from each other.

Due to the computational expense of SMSIM in simulating all available recordings (11,341), we downsample the full dataset during model calibration. The aim with this is to appropriately reflect the full dataset in terms of correspondence of intensity to the metadata, while avoiding data redundancy and associated computational expense of using the dataset in its entirety. The data are, therefore, uniformly sampled from different bins of the magnitude-distance distribution, which we consider to be the primary driver of ground motion intensity. The bins are selected in such a way that the magnitude range ($M_w < 6$) is divided into 5 equally spaced bins and uniform 50 km bins are selected over the whole distance range. 15 recordings are then randomly sampled from each bin, creating the new down-
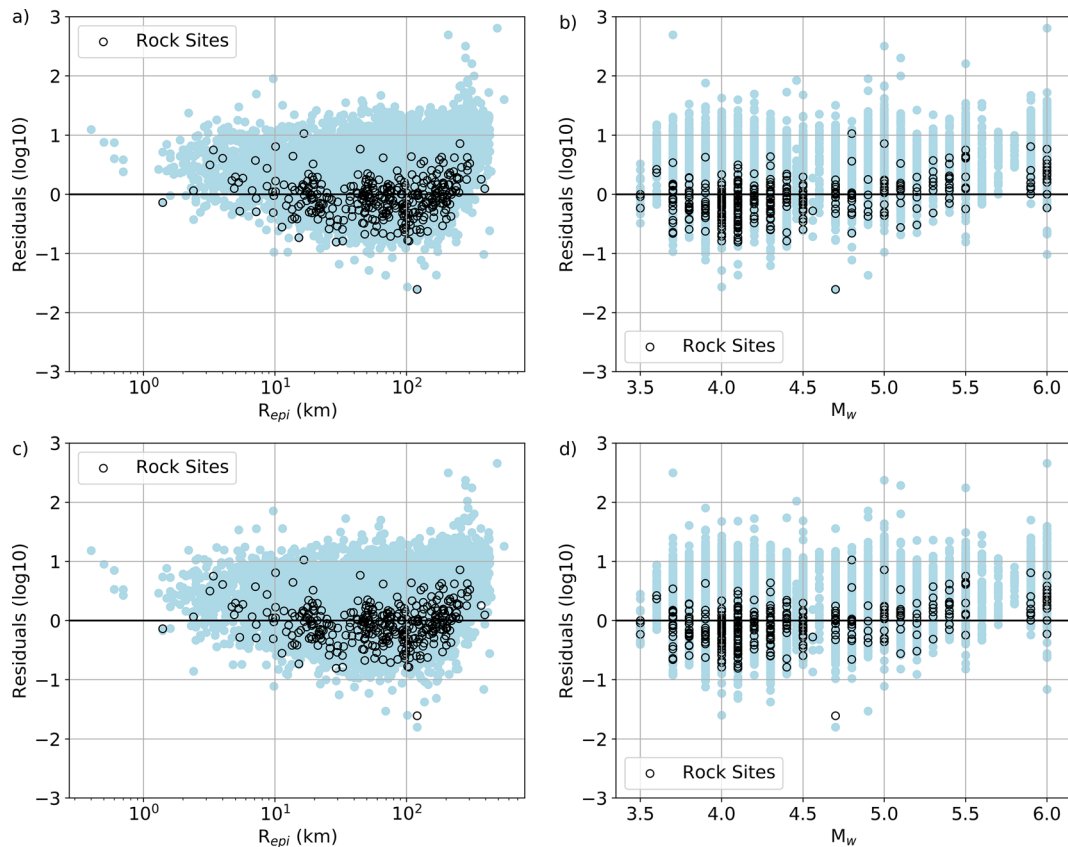
Fig. 2 The distribution of the residuals (PGA) before (top panel) and after (bottom panel) considering the site amplification. The empty dots denote the data recorded only from rock sites. (left panel) Epicentral distance vs residuals; (right panel) M_w vs residuals

sampled dataset for the calibration. The maximum number of available recordings is selected from the bins at extreme ranges of magnitude and distance (i.e. where fewer than 15 recordings are present). This results in a total of 505 recordings after down-sampling (4.7% of the entire database). The distribution of the down-sampled data can be inferred from Fig. 3, and can be directly compared to the full dataset in Fig. 1. The time difference in the simulation process reduces from about 20 min for the full dataset, to 1 min for the down-sampled dataset.

The area metric is calculated for the prior model (Table 1), after adding the site corrections, to define the baseline fit and understand the behaviour of subsequent simulations. For the given data subset, AM = 0.446. When the data- and simulation-specific cumulative distributions are compared, there is a difference in the simulations towards the tails (Fig. 4), which arises due to the inherent randomness of the process, i.e. the aleatory

uncertainty, which is as yet not considered. The aleatory uncertainty in our simulations is introduced by considering a sigma value (Strasser et al. 2009), which is commonly used to quantify the variability of ground motion in GMMs. The GMM output is therefore described in terms of a median (simulated) value and a logarithmic standard deviation (here, base-10), sigma (σ) (e.g. Strasser et al. (2009)) as shown in Eq. 3:

$$\log_{10}(Y_{obs}) = \log_{10}(Y_{pred}) + \log_{10}[S(Vs_{ref} = 760 m/s, V_{S30})] + N(0, \sigma) \quad (3)$$

where Y_{obs} are the observed data, Y_{pred} are the median simulation values (without site amplification) of PSA, and S are site terms from Boore et al. (2014) for a given V_{S30}. σ defines the aleatory variability associated with the ground motion prediction and is added onto the simulated values. We use a value of $\sigma = 0.34$, which is inside the common range of sigma used in ground-motion models. The performance of sim-
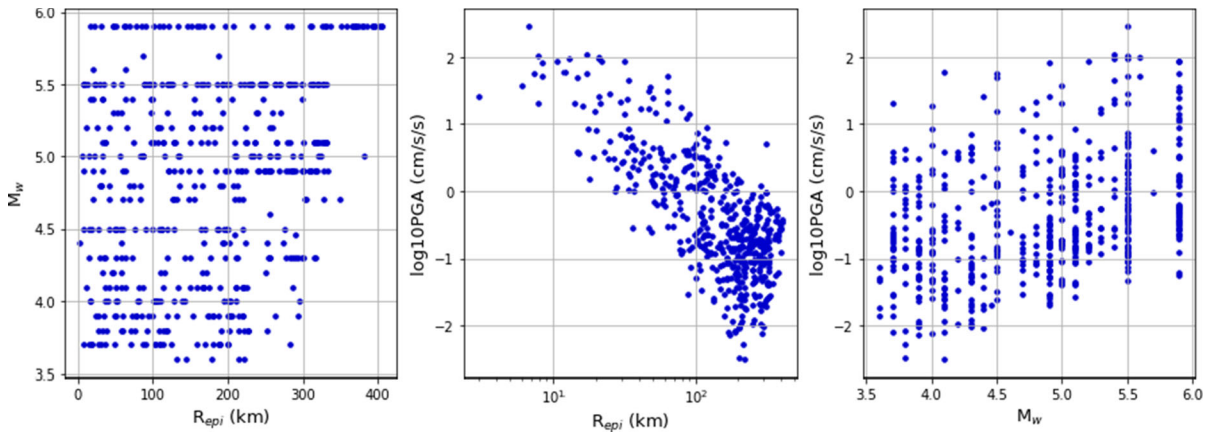
Fig. 3 The distribution of the Italian data after down-sampling. (a) Epicentral distance vs M_w. (b) Epicentral distance vs PGA. (c) M_w vs PGA values

ulations towards the tail of the distributions significantly increased after introducing the sigma component (Fig. 4). The area metric also increased slightly from 0.446 to 0.447. The final setup for the calibration process (Eq. 3) therefore considers the prior model to include: the attenuation model of Bindi and Kotha (2020), a magnitude-dependent stress drop model (Eqs. 1 and 2), site amplification from Boore et al. (2014), and a generic sigma component. The naive implementation of the models directly from the literature results in a 0.226 shift in the mean value of residuals, as shown in Fig. 2, which corresponds to a reasonable level of under-prediction. One source of difference, for instance, could be an incompatibility between the various model components, such as the

spectral model and RVT duration as they have been derived independently.

We analyse the performance of SMSIM simulations using random perturbations of the prior. For each model perturbation, we determine the AM (e.g. Fig. 4) and check whether a minimum AM value has been obtained. The calibration process starts by generating independent and identically random samples (iid) of the major parameters in SMSIM with the median value taken from the prior — here, parameters for Italy. The data-generating mechanism is a normal distribution (with location and scale defined in Table 1 as $N(location, scale)$) and produces variables with high probability density around the prior values. In other applications, the user should consider the most appro-

Fig. 4 The area metric plot before (a) and after (b) considering aleatory variability (the sigma component)
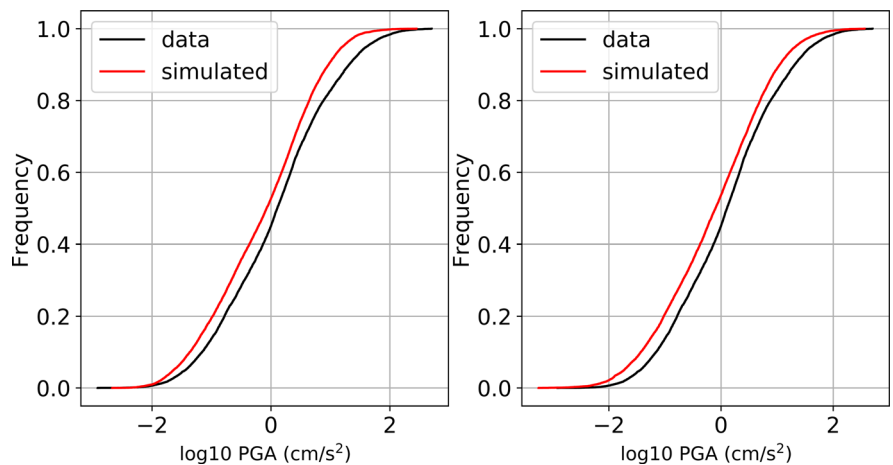
priate sampling mechanism. The calibration process accounts for the epistemic uncertainty which can arise due to many factors such as selection of models, trade-off between parameters, significant digits, missing data or sparse data. In this study, we, therefore, consider both epistemic and aleatory uncertainty as two different entities (recall that the aleatory component was introduced through considering sigma, σ). Aleatory uncertainty here comes from the observed dataset values such as magnitude and epicentral distance and also from the randomness of the model, while epistemic uncertainty comes from the seismological parameters given as the input to the model.

Initially, 1000 uncorrelated sets of simulation parameters are randomly generated. In this first instance, each variable in a given parameter combination is independently sampled from its normal distribution (Table 1). All corresponding simulations are then analysed using each combination of seismological input parameters and the AM values are calculated. Path- and site-specific parameters are considered for calibration in this study, as these are most sensitive to regional variability. Specifically, we aim to refine the frequency-dependent quality factor, geometrical spreading model and site damping (κ_0). We also calibrate the aleatory ground-motion variability (sigma) along with the other parameters. The combinations of parameters providing minimum AM, with some tolerance around this minimum, provide statistically improved simulations compared to those determined using median parameters prior. If we are not able to obtain a reduction in AM from the first suite of simulations, the trial distribution is widened and the process is repeated. In this instance, we began with a standard deviation of approximately 20% on each parameter (Table 1), which proved sufficient. If calibration cannot be achieved at this range, it would be possible to increase the standard deviation to widen the search.

Due to the high number of trial parameter combinations, all simulations performed during the calibration stage are calculated for and compared to, the down-sampled dataset. This significantly increases the speed of the execution of the process ($\sim$ 20 times faster). This is discussed in the Data and Methodology section. Once the process is completed using the down-sampled data and the final parameters are determined, we use this combination of parameters to simulate ground motion for the complete dataset and validate our results. No differences in the residual trends were observed between the down-sampled and full datasets. The algorithm is given in the flowchart provided in Fig. 5. Because of the randomness involved in the process (random selection of input seismological parameters), the algorithm does not provide a unique solution after the completion of the calibration processes. In the following, we, therefore, extend the approach to explore the epistemic uncertainty of the resulting GMM.

Rather than simply obtain the minimum misfit (or optimum) model the approach outlined above allows us to explore the epistemic uncertainty of our model. To further understand the interaction of various parameter combinations and uncertainties, we, therefore, build a confidence band around the observed data (Fig. 6), with the aim to attain simulations that fit into this confidence band, along with their corresponding input parameters. In order to build confidence bands, we calculate the empirical cumulative distribution function (ECDF) of the data, i.e. the distribution function associated with the empirical measure of a sample. We then consider all the simulations that fit into the banded ECDF, rather than aiming for the single ECDF that uniquely explains the dataset. While the target is no longer the minimal AM, these simulations are nevertheless those that provide minimum AM values from all trials. We calculated the confidence band of the ECDF using the Dvoretzky-Kiefer-Wolfowitz-Massart (DKW) inequality. The DKW inequality bounds how close an empirically determined distribution function will be to the distribution function from which the empirical samples are drawn. This can be used as a method for generating cumulative distribution function (CDF)-based confidence bands. The interval that contains the true CDF of n values, $F(x)$, with probability 1-α is specified using DKW inequality as

$$F_n(x) - \varepsilon \leq F(x) \leq F_n(x) + \varepsilon \quad \text{where } \varepsilon = \sqrt{\frac{\ln \frac{2}{\alpha}}{2n}}$$
(4)

A suite of parameters is therefore determined, which provides simulations that fit within the ECDF confidence band. This provides us with a quantitative assessment of how the seismological parameters vary within the confidence band of the data. In this way, we account for epistemic uncertainty related to data confidence (and availability). We have defined models to fit within the confidence band where a maximum of 10% of recordings lie outside the defined bands (i.e. 10% tol-
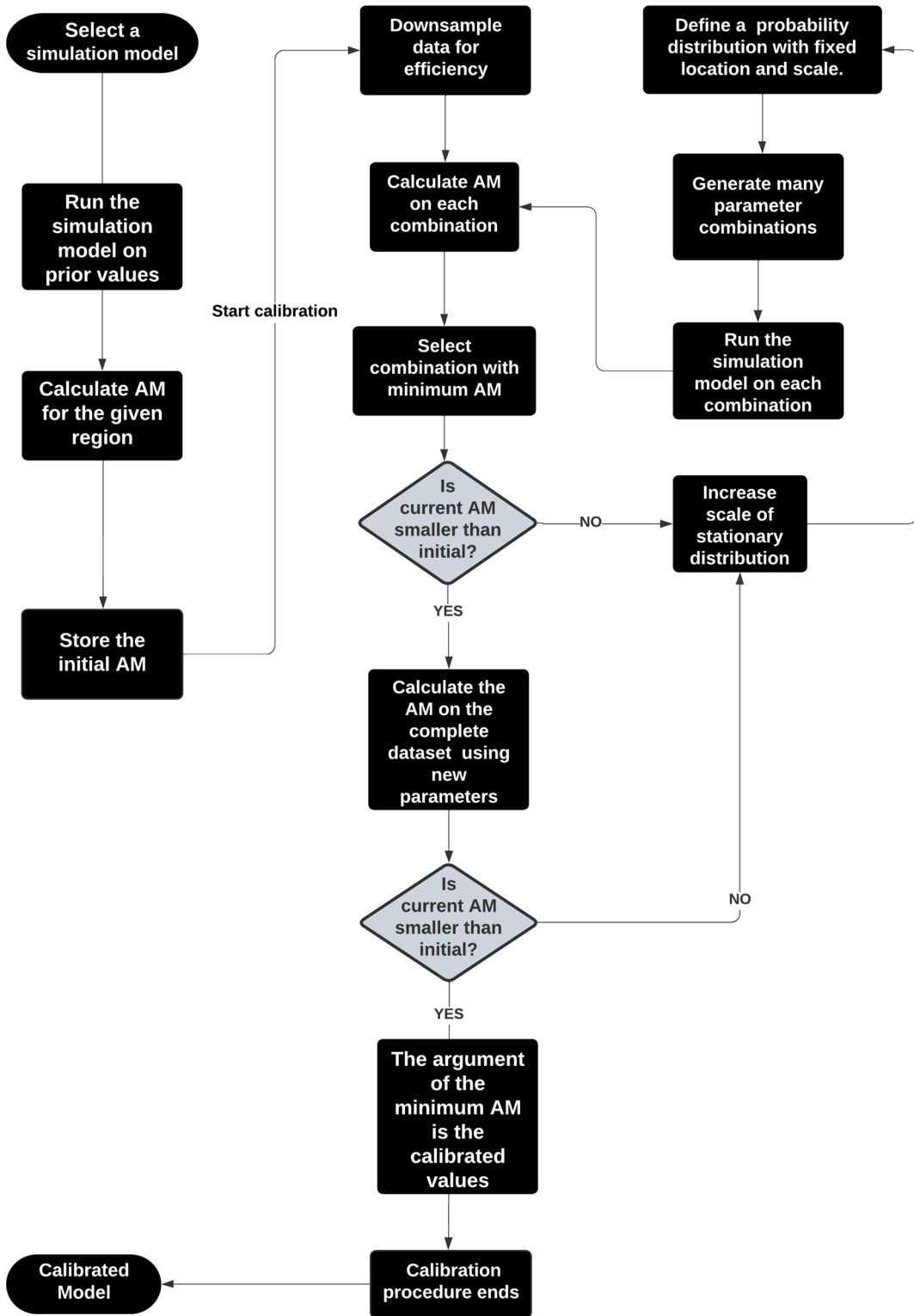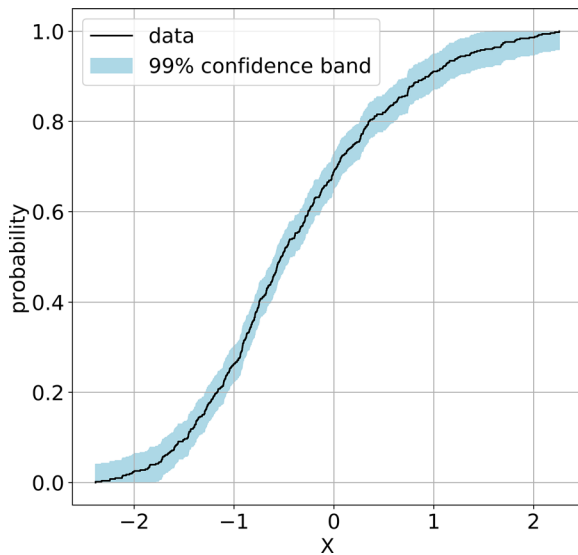
Fig. 5 Flowchart describing the algorithm for calibration of SMSIM parameters

Fig. 6 The confidence band using DKW inequality. Blue region represents the 99% confidence band of the given data (black curve)

erance). This is chosen because it is difficult to obtain the parameter combinations which provide simulations that are 100% inside the chosen bands.

Several ECDF confidence bands are created. The confidence band will be wider at high confidence levels and narrower at lower confidence levels. The simulation models that are compatible with each confidence interval are then determined. The number of parameter combinations consistent with wider confidence bands is larger, reflecting our increasing inability to define model parameters at higher levels of data confidence. An example of the 99% confidence band created using DKW inequality is shown in Fig. 6. The blue region

in the figure represents the target confidence band for which we aim to define the combination of seismological parameters. We also analysed the optimum parameter combination for periods of 0.1, 0.5, 0.8, and 1 s.

3 Results of model calibration

The calibration procedure leads to modest changes to the prior (Table 2). After calibration, all simulations perform better than those using the initial (uncalibrated) combination of parameters. We first analysed the residual plots (observed value–simulated value, in log_{10}) before and after the re-calibration of SMSIM parameters to obtain an understanding of how the simulations changed with the minimum AM set of simulation parameters. The residual plots are computed and plotted with respect to epicentral distance (R_{epi}) and moment magnitude (M_w). We initially performed residual analysis for the down-sampled data (Figs. 7 and 8) and then moved to the complete dataset (Fig. 9) to confirm whether the calibrated combination of parameters is valid.

The residual distribution before re-calibration showed significant under-prediction (Fig. 7). Most of the residuals are distributed well above the zero horizontal line and the mean value of residuals before optimising the parameters is 0.446. The residual distribution after re-calibration is more symmetrically distributed around the zero horizontal line and the under-prediction is less significant. The mean value of residuals changed to −0.01 after the calibration procedure. The standard deviation of residuals, after considering the sigma component, changed slightly from 0.49 to 0.42 after implementing the algorithm on the down-sampled data.

Table 2 The initial parameters and the updated (minimum AM) parameters with the difference in each parameter

Parameter	SMSIM parameter	Initial	Calibrated	Difference	Location	Scale
$Q0$	Qr1	250.40	249.36	1.04	261.93	26.13
$alpha$	s	0.29	0.33	0.04	0.31	0.07
$gamma$	a_s(1)	−1.35	−0.84	0.51	−0.89	0.15
$gamma1$	a_s(2)	−0.57	−0.40	−0.17	−0.51	0.14
$gamma2$	a_s(3)	−1.53	−1.52	0.01	−1.32	0.32
$kappa$	kappa	0.02	0.01	−0.01	0.018	0.008
$sigma$	–	0.33	0.33	0.00	0.33	0.01

Note that at this point the calibration is performed only on (and therefore valid for) PGA. The last two columns provide the location and scale of the distribution of each parameter
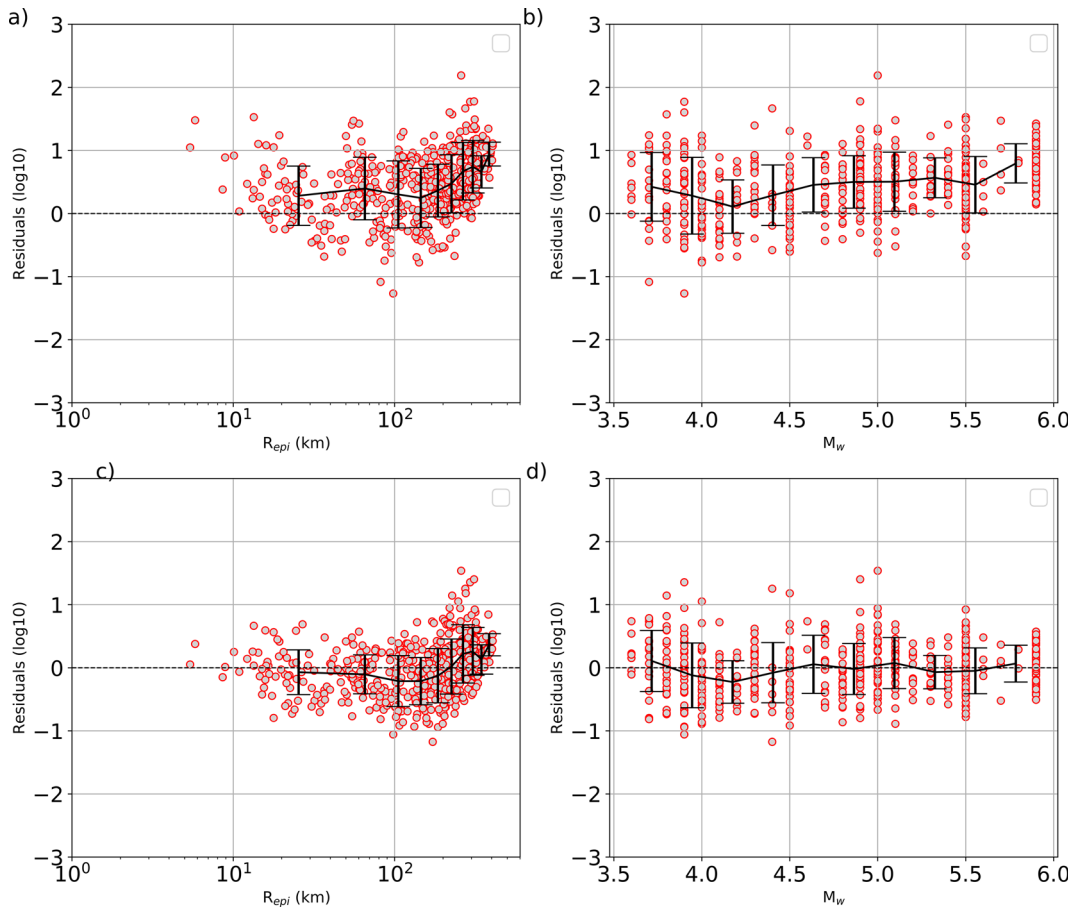
Fig. 7 The distribution of the residuals (down-sampled data) before (top) and after (bottom) optimising the parameters. The top panel shows the plots before calibration and the bottom panel

indicates the plots after calibration. (Left) Epicentral distance vs residuals; (right) M_w vs residuals

While the algorithm attempts to reduce the area in between the data and the simulation CDFs, the distribution of the residuals will also converge, becoming more symmetrical to the zero horizontal, as the area metric is minimised. The change in AM values before and after calibration is shown in Fig. 8. The newly calibrated parameters, the old parameters and their difference are given in Table 2. The mean and standard deviation of residuals within separate bins of distance and magnitude are shown in Fig. 7. The standard deviation values clearly decrease within each bin analysed and the mean value of each bin shifts towards the zero line, consistent with the average shift over the entire dataset. The AM plots are generated before and after the re-calibration of the down-sampled data (Fig. 8). After the re-calibration process, the simulated PGA values

moved closer to the empirical (target) dataset. Specifically, the AM value decreased from 0.447 to 0.0674.

Once we have the updated seismological model parameters, developed using the down-sampled data, we simulate ground motion records for the complete ESM dataset in order to validate our results. As noted earlier, sigma is also optimised in this study alongside the seismological model. The simulation residuals, using the updated model and complete dataset, are symmetrically distributed around the zero horizontal line as shown in Fig. 9. The sample mean and sample standard deviation of residuals in different bins are also analysed and plotted, most of the bins have sample mean values near zero and have reduced sample standard deviations. The AM value is calculated and it is reduced to 0.0516 from 0.226 after and before the
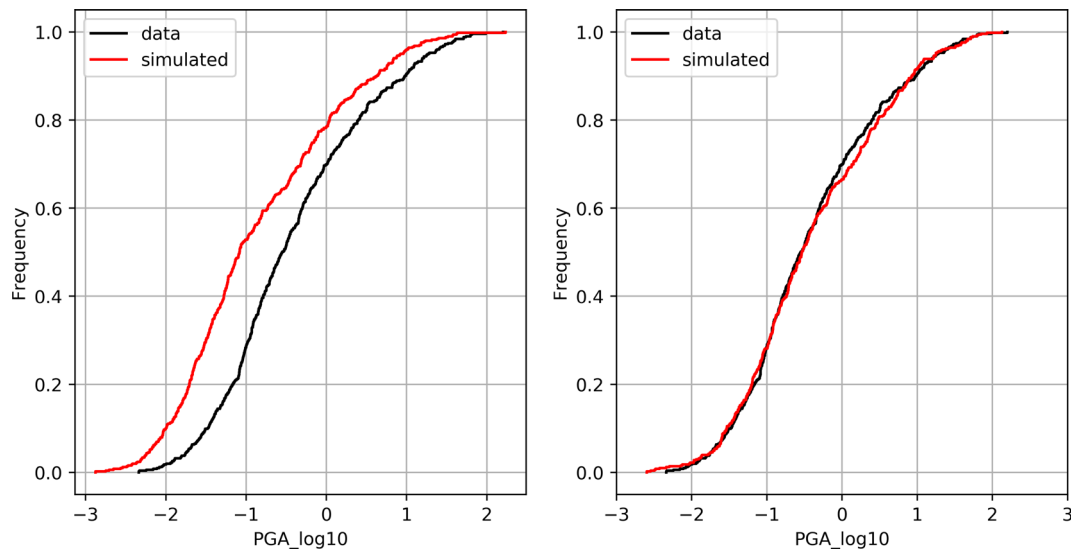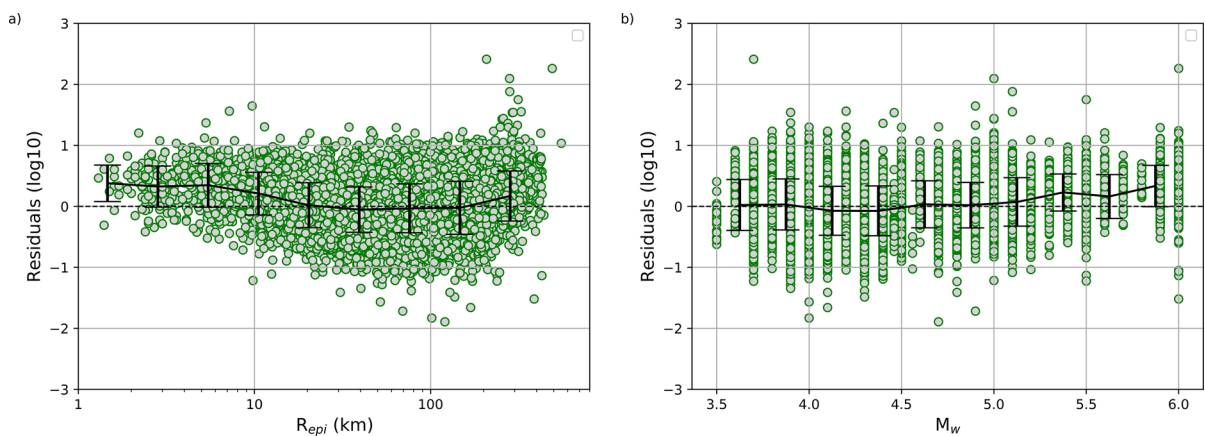
Fig. 8 The area metric plot (a) before and (b) after calibration of the SMSIM parameters with the down-sampled data

re-calibration. We were therefore able to obtain a good fit for the complete ESM dataset for Italy by optimising the SMSIM parameters using only 4%-5% of the original dataset.

Figure 10 shows the differences in the AM values before (0.447) and after (0.0674) updating the parameters using both the down-sampled dataset (Fig. 10a) and the complete dataset (0.0516 from 0.226) (Fig. 10b). When the changes in the input parameters are compared before and after re-calibration, the maximum change is seen in the parameters related to the geometrical spreading, i.e. the slopes of each distance segment

(*gamma* and *gamma*1) compared to the other parameters considered in this study.

As we have seen, the changes in the various seismological parameters are variable (both in amplitude and direction), and hence, it is important to understand quantitatively how each of the model parameters varies in different confidence levels of the data. By creating confidence bands and determining the parameters providing simulations that are consistent with each band, we can obtain an idea about the behaviour of individual parameters, as the calibration migrates towards unbiased simulations. The range of parameters at each

Fig. 9 The distribution of the PGA residuals (complete data) after optimising the parameters. (Left) Epicentral distance vs residuals; (right) M_w vs residuals
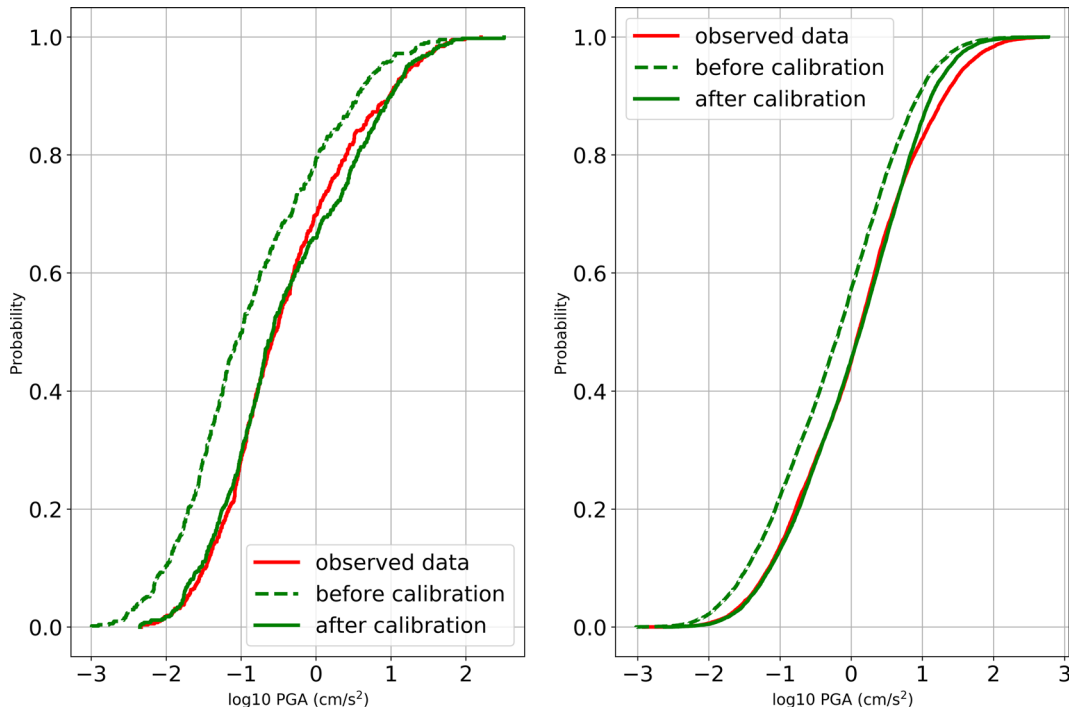
Fig. 10 The area metric plot of (a) down-sampled and (b) complete dataset before and after updating the parameters. The dashed green line represents the empirical distribution of data before updating and the solid line represents the simulation distribution with the updated parameters

confidence level then indicates the degree of epistemic uncertainty that can affect the simulations.

As discussed in Sect. 2.2, confidence levels for the empirical ground motion data are created using the DKW inequality. 99.9% and 95% confidence bands on the empirical data are analysed here and, subsequently, we identify the distribution of simulation parameters that are consistent with each confidence level. The simulations are considered to agree with the defined data confidence band once 90% of the simulated recordings fall within the band. We allowed a small fraction of tolerance (10%) since the number of models with 100% simulations inside the band is very few and may be limited by simplifications of the model form, as defined in SMSIM. 87 parameter combinations lie within the 99.9% confidence band of the data, i.e. around 10% of all trial parameter combinations. The number of alternative parameter combinations is reduced when decreasing the confidence level required (and subsequently decreasing the uncertainty around the ECDF). Specifically, the number of candidate parameter combi-

nations drops from 87 when fitting a 99.9% confidence band around the ECDF, to 41 at 95%.

After obtaining candidate parameter combinations within each confidence band, we analysed their correlation plots using the Pearson correlation coefficient technique. The parameter correlation was low for high confidence bands (e.g. 99.9%) but increased as the confidence level decreased (e.g. 95%). When the correlation is plotted with simulations consistent with the 95% confidence band, we observe a correlation between different parameters used in our study (Fig. 11). The maximum correlation we observe is between $gamma$ and $gamma_2$ and also between $alpha$ and $gamma$. Both these values exhibit a negative correlation, which shows that $gamma$ and $gamma_2$, and $alpha$ and $gamma$, are inversely dependent to each other. All of these parameter combinations provide us with improved simulations compared to the original simulations, since the AM values of the 41 models (for the 95% data confidence) are lower than the AM value obtained using the initial combination of parameters. We observe a wide range

a)

b)
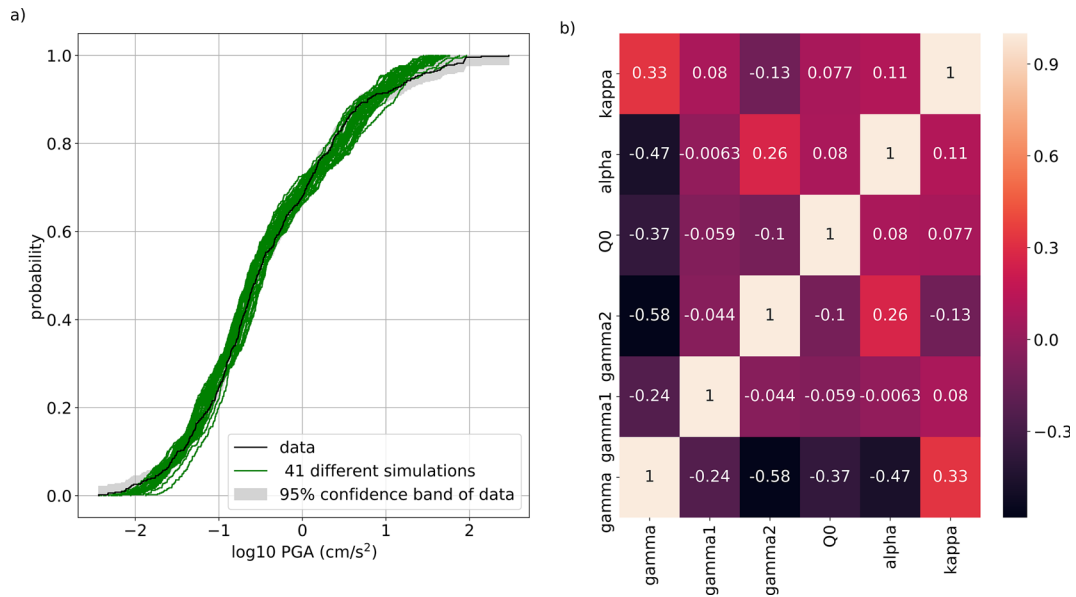
Fig. 11 (a) Suit of simulations that fit into the 95% confidence band of the data. Green distributions are the selected models that fit in the confidence band (allowing up to 10% of the simulated

intensities to be outside the bands). (b) correlation plot of the input parameters by using simulations that fit into the 90% confidence band

of possible values for individual parameters, which is to be expected given the numerous possible parameter combinations and strong covariance in such models (Cotton et al. 2013).

Extending the validity of the calibrated model to different periods, we test all parameter combinations falling within the 95% confidence band of the PGA ECDF on PSA data at 0.1 s. We repeated this process iteratively for periods 0.1, 0.5, 0.8 and 1 s and ended up with only two of the initial 41 models that fall within the 95% confidence band of all periods. As a result of the iterative process, our model now provides updated parameters for PGA, and PSA at 0.1s, 0.5s, 0.8s, and 1 s. The corresponding calibrated model parameter values (with minimum average AM over all periods) are $Q_0 = 241.36$, $alpha = 0.26$, $gamma = -0.58$, $gamma1 = -0.48$, $gamma2 = -1.61$, $kappa = 0.03$. The simulations with the updated model provide a better fit for all these periods compared with the initial simulations.

3.1 Epistemic uncertainty

Given the suite of candidate simulation parameter combinations determined for any given period, we can

resample numerous alternative parameter configurations by using their correlation (or covariance) matrix. Specifically, we use a resampling technique that, based on a random sample of a base parameter, uses the Cholesky decomposition of the covariance matrix to successively build correlated parameter combinations (Rietbrock et al. 2013). In following this approach, the resampled parameter combinations will preserve the statistical properties of the original set of simulation parameters, while the number of parameter combinations falling inside the confidence band will increase. Note, however, that despite ensuring statistical consistency of the simulation parameters, this does not ensure that all the resultant simulations fall within the confidence bands of the data (ECDF), as shown in Figure S1 in the supplementary material. The red scatters are the parameters that fall inside the confidence band. Nevertheless, these new parameter combinations can be used in a manner similar to the use of the optimum parameter combination, where the resultant CDF can subsequently checked against the ECDF's confidence band.

The Cholesky decomposition (L) of the covariance matrix (A) of our retained simulation parameters (i.e. those producing simulations falling within the defined confidence limit) is given by $A = LL^T$, where L is

a real lower triangular matrix with positive diagonal entries. We start by generating n random (uncorrelated) variables (here we simulate $n = 1000$) for each simulation parameter, similar to the initial stage of the calibration algorithm described earlier. To create correlated simulation parameters, these independent random variables are multiplied by the lower triangular matrix L. As a result, the use of Cholesky decomposition has generated sample sets of variables that have consistent covariance with our previous optimisation. This should lead to far more simulation parameter sets that subsequently fall within the confidence limits at each given step in the calibration.

Using this approach, it is possible to not only generate additional models that fit within the confidence band of the empirical cumulative distribution function (ECDF), but also to reduce the width, or standard deviation, of the parameters used, as compared to the initial trials. For instance, by utilising correlated samples, the number of parameter combinations that fit within the ECDF's confidence band over all periods, from PGA to 1 s increased from 2 to 14. Figure 12 shows a scatter pair plot for 1 s, which highlights the interdependence of each simulation parameter within the suite of parameter combinations examined in this study. The parameter combinations that fit within the ECDF's confidence band increased seven-fold, compared to the initial set of samples that were determined without correlation (2 uncorrelated samples). There are variations in the values while exploring the aspect of epistemic uncertainty. This variability may arise due to the extensive scope of analysis, especially in a larger region such as Italy, leading to significant fluctuations in parameters during the exploration of epistemic uncertainty. The range of kappa from 0.012 to 0.031 is significant, but certainly less than the range of site-specific values that would be determined regionally. Consider, also that the method inherently explores the trade-off between parameters — while there is significant variation of gamma2 (which only affects decay at distances above 119 km), this trades off with Q — such that the combined attenuation models will be similar. While it may be considered that in Italy there are abundant data to constrain values of attenuation, this is in terms of the bulk effect and not the individual components. This is well-demonstrated by Shible et al. (2022), who used several approaches to define seismological (Fourier-based) model, including the stress drop, Q and gamma. They showed significant ranges of parameters, depend-

ing on the method chosen and assumptions used. The complete set of final models, valid over the range PGA to 1 s can be found in Table S1 of the supplementary material, and a scatter plot displaying the correlated initial samples and the correlated samples that fit within the 95% confidence band of PGA is also provided in Figure S1.

4 Conclusion

This paper discussed the importance of regionally adjusted ground-motion models for robust and precise ground-motion predictions for engineering applications. The sparse data problem and the need for regionally adjusted ground-motion models are considered, and an algorithm has been developed based on the concept of Area Metric for the validation and calibration of ground-motion models. The calibration is based on a stochastic simulation approach, which was performed using the SMSIM programs. The approach was tested using the ground motion recordings of Italy from the ESM dataset. We have analysed ground motion simulation results by taking advantage of the available information and the properties of recorded signals from ESM data, along with the duration parameters and the spectral decomposition results for strong motion data in Europe, which form our prior. Initial simulation results, using the prior simulation parameters of Bindi and Kotha (2020), have been analysed and showed the predictions appear broadly consistent, but under-predicting on average, when using the site effects model from Boore et al. (2014). In the first instance, this study, therefore, validates the use of the spectral decomposition results of Bindi and Kotha (2020), paired with the site amplification model of Boore et al. (2014), for simulation of PGA. Importantly, the model of Bindi and Kotha (2020) uses a magnitude-dependent stress drop, without which, PGA would have been underestimated for the larger events. Of particular note when comparing CDFs of observed and simulated data, our simulation CDF takes into account aleatory uncertainty, inherent in empirical data, by adding a sigma component, as commonly used in the empirical ground motion models.

The initial fit of the simulation model based on the spectral decomposition by Bindi and Kotha (2020), was further refined using a calibration technique, that builds on the Area Metric (Sunny et al. 2021). The calibra-

Fig. 12 Pair-wise comparison showing the dependencies between calibrating parameters. The parameters being examined here are Q_0, $gamma$, $gamma_1$, $gamma_2$, $alpha$, and $kappa_0$. Within the plot, green circles indicate the number of corre-lated parameters that fall within the confidence band of the 1 s data, while grey squares represent the initial correlated parameters (correlation obtained from the uncorrelated samples that fit within the 95% confidence band of the PGA data)

tion is an iterative approach that, based on an initial prior (with wide parameter variability), refines the correlated simulation parameters on a period-by-period basis. This results in a decreasing width (and epistemic uncertainty) of the individual simulation parameter distributions. Residuals and AM plots, analysed after all these considerations, show that the calibrated simulations improve the fit to the empirical data. We have also used correlation plots to show the behaviour of various input parameters used in this study. In order to under-

stand the interdependence among various parameters in more detail, further research in this area is necessary, and this will be the main focus of subsequent studies.

Our approach uses a uniformly down-sampled dataset from the full Italian ESM dataset, considering the computationally expensive stochastic process, for faster execution. This was found to be a successful way of improving the speed and efficiency of simulation-based calibration approaches.

Importantly, the presented optimisation approach does not provide a unique solution but rather a suite of simulation parameters, each of which performs better than the initial prior. This allows us to account for the inherent parameter covariance matrix and the associated epistemic uncertainty. We have analysed this suit of parameter combinations by considering a confidence band around the data using DKW inequality and then selecting the simulations and the corresponding parameters that fall in the particular confidence level of the data. The calibrated simulations using this optimisation are designed to give a minimum area metric value for the given region and this framework for the calibration and updating of parameters can help achieve robust and transparent regionally adjusted stochastic models.

Data and resources Simulation model parameters determined for the Italian region as part of this study are provided in the supplementary material. The response spectra used in this study are available on the website https://esm-db.eu/esmws/flatfile/1. All other data used in this study are from the sources listed in the references. In order to allow for the results presented in this paper to be reproduced, the authors have made available the software used in relation to the work. The code for the efficient computation of the stochastic area metric is available at the doi: 10.5281/zenodo.4419645. The codes used for the calibration algorithm are accessible at https://github.com/Jaleena/Calibration.

Declarations

Conflict of interest The authors declare no competing interests.

References

Afshari Kioumars, Stewart Jonathan P (2016) Physically parameterized prediction equations for significant duration in active crustal regions. Earthq Spectra 32(4):2057–2081

Akkar Sinan, Sandıkkaya M Abdullah, Şenyurt M, Azari Sisi A, Ay Bekir Özer, Traversa Paola, Douglas John, Cotton Fabrice, Luzi Lucia, Hernandez Bruno et al (2014) Reference database for seismic ground-motion in Europe (RESORCE). Bull Earthq Eng 12:311–339

Anderson John G (2015) The composite source model for broadband simulations of strong ground motions. Seismol Res Lett 86(1):68–74

Atkinson Gail M (2008) Ground-motion prediction equations for eastern North America from a referenced empirical approach: implications for epistemic uncertainty. Bull Seism Soc Am 98(3):1304–1318

Atkinson Gail M, Assatourians Karen, Boore David M, Campbell Ken, Motazedian Dariush (2009) A guide to differences between stochastic point-source and stochastic finite-fault simulations. Bull Seism Soc Am 99(6):3192–3201

Atkinson Gail M, Boore David M (1995) Ground-motion relations for eastern North America. Bull Seism Soc Am 85(1):17–30

Atkinson Gail M, Boore David M (2006) Earthquake ground-motion prediction equations for eastern North America. Bull Seism Soc Am 96(6):2181–2205

Atkinson Gail M, Silva Walter (2000) Stochastic modeling of California ground motions. Bull Seism Soc Am 90(2):255–274

Beresenev I, Atkinson GM (1997) Modeling finite fault radiation from WN spectrum. Bull Seismol Soc Am 87:67–84

Bindi Dino, Kotha SR (2020) Spectral decomposition of the engineering strong motion (ESM) flat file: regional attenuation, source scaling and Arias stress drop. Bull Earthquake Eng 18(6):2581–2606

Boatwright John (1982) A dynamic model for far-field acceleration. Bull Seism Soc Am 72(4):1049–1068

Boore David M (1983) Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. Bull Seism Soc Am 73(6A):1865–1894

Boore David M (2003) Simulation of ground motion using the stochastic method. Pure Appl Geophys 160(3):635–676

Boore David M, Thompson Eric M (2014) Path durations for use in the stochastic-method simulation of ground motions. Bull Seism Soc Am 104(5):2541–2552

Boore David M, Stewart Jonathan P, Seyhan Emel, Atkinson Gail M (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. Earthq Spectra 30(3):1057–1085

Campbell Kenneth W (2003) Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in eastern North America. Bull Seism Soc Am 93(3):1012–1033

Cotton Fabrice, Archuleta Ralph, Causse Mathieu (2013) What is sigma of the stress drop? Seismol Res Lett 84(1):42–48

de Angelis, Marco, Gray, Ander (2021) Why the 1-Wasserstein distance is the area between the two marginal CDFs

Edwards B, Zurek B, Van Dedem E, Stafford PJ, Oates S, Van Elk J, DeMartin B, Bommer JJ (2019) Simulations for the development of a ground motion model for induced seismicity in the Groningen gas field, The Netherlands. Bull Earthquake Eng 17(8):4441–4456

Graves Robert W, Pitarka Arben (2010) Broadband ground-motion simulation using a hybrid approach. Bull Seism Soc Am 100(5A):2095–2123

Hanks Thomas C, McGuire Robin K (1981) The character of high-frequency strong ground motion. Bull Seism Soc Am 71(6):2071–2095

Kolli Mohan Krishna, Bora Sanjay Singh (2021) On the use of duration in random vibration theory (RVT) based ground motion prediction: a comparative study. Bull Earthquake Eng 19(4):1687–1707

Lanzano, G, Luzi, L, Russo, E, Felicetta, C, D'Amico, MC, Sgobba, S , Pacor, F (2018) Engineering strong motion database (ESM) flatfile [data set]. istituto nazionale di geofisica e vulcanologia (INGV)

Luzi Lucia, Puglia Rodolfo, Russo Emiliano, D'Amico Maria, Felicetta Chiara, Pacor Francesca, Lanzano Giovanni, Çeken Ulubey, Clinton John, Costa Giovanni et al (2016) The engineering strong-motion database: a platform to access Pan-European accelerometric data. Seismol Res Lett 87(4):987–997

Mai P Martin, Imperatori Walter, Olsen Kim B (2010) Hybrid broadband ground-motion simulations: combining long-period deterministic synthetics with high-frequency multiple S-to-S backscattering. Bull Seism Soc Am 100(5A):2124–2142

McGuire Robin K, Hanks Thomas C (1980) RMS accelerations and spectral amplitudes of strong ground motion during the San Fernando, California earthquake. Bull Seism Soc Am 70(5):1907–1919

Olsen Kim, Takedatsu Rumi (2015) The SDSU broadband ground-motion generation module BBtoolbox version 1.5. Seismol Res Lett 86(1):81–88

Shahram Pezeshk, Arash Zandieh, Behrooz Tavakoli (2011) Hybrid empirical ground-motion prediction equations for eastern North America using NGA models and updated seismological parameters. Bull Seism Soc Am 101(4):1859–1870

Razafindrakoto Hoby NT, Cotton Fabrice, Bindi Dino, Pilz Marco, Graves Robert W, Bora Sanjay (2021) Regional calibration of hybrid ground-motion simulations in moderate seismicity areas: application to the upper Rhine graben. Bull Seism Soc Am 111(3):1422–1444

Rietbrock Andreas, Strasser Fleur, Edwards Benjamin (2013) A stochastic earthquake ground-motion prediction model for the United Kingdom. Bull Seismol Soc Am 103(1):57–77

Shible Hussein, Hollender Fabrice, Bindi Dino, Traversa Paola, Oth Adrien, Edwards Benjamin, Klin Peter, Kawase Hiroshi, Grendas Ioannis, Castro Raul R et al (2022) GITEC: a generalized inversion technique benchmark. Bull Seismol Soc Am 112(2):850–877

Somerville Paul, Sen Mrinal, Cohee Brian (1991) Simulation of strong ground motions recorded during the 1985 Michoacan, Mexico and Valparaiso, Chile earthquakes. Bull Seism Soc Am 81(1):1–27

Strasser Fleur O, Abrahamson Norman A, Bommer Julian J (2009) Sigma: issues, insights, and challenges. Seismol Res Lett 80(1):40–56

Sunny Jaleena, de Angelis Marco, Edwards Benjamin (2021) Ranking and selection of earthquake ground-motion models using the stochastic area metric. Seismol Res Lett

Vanmarcke Erik H, Lai Shih-Sheng P (1980) Strong-motion duration and RMS amplitude of earthquake records. Bull Seism Soc Am 70(4):1293–1307

Wald David J, Allen Trevor I (2007) Topographic slope as a proxy for seismic site conditions and amplification. Bull Seism Soc Am 97(5):1379–1395