J Seismol (2019) 23:39–57
https://doi.org/10.1007/s10950-018-9792-3

ORIGINAL ARTICLE

Continuous integration of data into ground-motion models using Bayesian updating

Peter J. Stafford

Received: 30 April 2018 / Accepted: 10 September 2018 / Published online: 1 October 2018

Abstract The development of empirically constrained ground-motion models has historically followed a cyclic process in which every few years, existing models are updated to reflect new data and knowledge that has become available. Ground-motion developers make use of their prior knowledge to identify appropriate functional forms for the models, but the actual regression analysis and model calibration is effectively performed from a fresh start with each update. With the anticipated increase in data availability coming in the future, this traditional approach will become increasingly cumbersome. The present article presents a framework in which Bayesian updating is used to continuously update existing ground-motion models as new data becomes available. This new approach is shown to provide similar results to the more traditional approach, but is far less data-intensive and will scale well into the future. The approach is demonstrated through an example in which an initial regression analysis is conducted on a portion of the NGA-West2 dataset representative of the information available in 1995. Model parameters, variance components and crossed random effects are then updated with data from every other event in the NGA-West2 dataset and the results from Bayesian updating and traditional regression analysis are compared. The two methods are shown to provide similar results, but the advantages of the Bayesian approach are subsequently highlighted. For the first time, the article also demonstrates how prior distributions of model parameters can be obtained for existing ground-motion models that have been derived using both classical, as well as more elaborate multi-stage, procedures with and without constrained parameters.

Keywords Ground-motion models · Bayesian updating · Ground-motion prediction equation · Continuous integration

P. J. Stafford (✉)
Department of Civil, Environmental Engineering, Imperial College London, South Kensington campus, London SW7 2AZ, UK
e-mail: p.stafford@imperial.ac.uk

1 Introduction

Ground-motion models typically used in seismic hazard and risk applications are empirically calibrated models with functional forms that are designed to replicate the physics of strong ground-motion scaling. There is a very long history of development of these models (Douglas 2018), dating back to 1964. The initial models developed aimed to capture basic scaling of peak ground motion parameters with magnitude and distance, and only focussed upon the predictions of median motions for these scenarios. However, in time, the models have evolved to enable them to also capture generic scaling with respect to the style-of-faulting, geometric effects of finite faults, near-surface site response effects, and the influence of

the deeper velocity structure, among other features. At the same time, it is now recognised very strongly that these models must specify the expected distribution of ground-motion intensities for each scenario they consider, rather than just the median motion (Stafford 2015b).

A simplistic view of ground-motion model development is to suggest that a large database of accelerograms, and their associated meta-data, is compiled and uniformly processed before a regression analysis is performed. In some cases, this is actually a true reflection of how models have been developed in the past. However, it is increasingly rare for ground-motion model development to follow such a simplistic model-fitting exercise. The Next Generation of Attenuation (NGA) project (Power et al. 2008) made a significant impact upon the way in which ground-motion models were developed as it became abundantly clear that this simple regression-based approach was no longer sufficient. Indeed, the majority of the NGA models appear over-parameterised from a simple regression perspective, but this is due to the fact that many of the model parameters are constrained from numerical simulations or theoretical arguments. What the NGA project

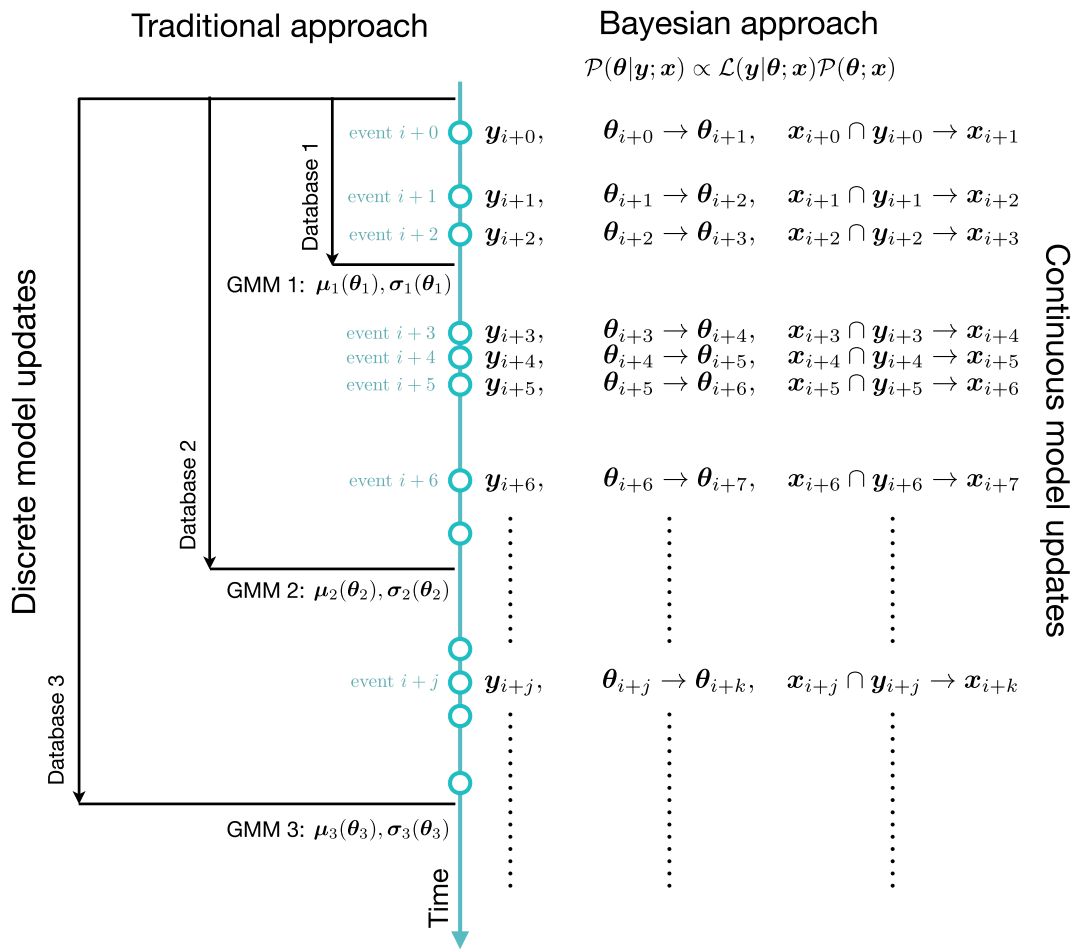

Fig. 1 Conceptual illustration of the differences between the traditional empirical approach to ground-motion model development and the continuous Bayesian updating approach presented herein. The passage of time flows down the figure, with the circular markers representing earthquake events that are recorded on accelerograph networks and provide observations $\boldsymbol{y}_i$. On the left, the traditional approach of periodically compiling a ground-motion data base and developing a new model is shown. Here, $\boldsymbol{\mu}_i$ and $\boldsymbol{\sigma}_i$ represent the revised functional forms at each update, and $\boldsymbol{\theta}_i$ are the revised model parameters. On the right, we have the continuous refinement through Bayesian updating. With each new event, the entire probability distribution of the parameters, $\mathcal{P}(\boldsymbol{\theta}|\boldsymbol{x})$, is updated, and new observations are absorbed into the prior for the next update. The functional form on the right-hand-side remains constant

did was to emphasise the fact that in order for models to behave well, when extrapolated to scenarios for which limited empirical constraint exists, functional forms (and potentially model parameters) need to be designed with extrapolation in mind from the outset.

The limitations of existing strong-motion databases have therefore driven a lot of effort into trying to understand how ground-motions should scale for a very broad range of rupture scenarios and site properties. The net result is that we are now relatively well-placed to understand which functional forms are likely to be successful for predicting median ground motions for scenarios of relevance to hazard and risk applications. At the same time, the data limitations previously referred to will inevitably relax with time. As more earthquakes are recorded on increasingly dense accelerograph networks, the database of available records will continue to increase. For some rupture scenarios, the challenges will shift from being associated with insufficient numbers of records, to having too many records to manage.

The purpose of the present article is to outline a new paradigm in ground-motion model development in which existing models are continuously updated as new data becomes available. This is in contrast to the current approach whereby every few years databases are consolidated and revisions to existing models are released. This difference in approaches is schematically illustrated in Fig. 1, with the details of the process elaborated upon in the following section.

The key tool used in the present study is Bayesian updating in which a model is incrementally updated as new data becomes available, as shown in Fig. 1. Bayesian methods have previously been used within Engineering Seismology for a number of applications. For example, Wang and Takada (2009) demonstrated how Bayesian updating can be used to adapt an existing ergodic ground-motion model to target a particular site (or region). In their study, the existing ergodic ground-motion model is used to define a prior distribution for the model parameters, while local site-specific recordings are then used to update the prior to find the site-specific posterior distribution of the model parameters. Arroyo and Ordaz (2010a, b) showed how a Bayesian framework can be employed to derive regression parameters for a vector of intensity measures simultaneously, rather than working on a period-by-period basis, as is commonly done. Kuehn et al. (2011) used Bayesian networks to explore the strength of dependencies between ground-motion intensities and variables within the NGA database, and a number of other applications can be found (e.g. Moss and Der Kiureghian 2006; Moss 2011; Hermkes et al. 2013).

Most recently, Stafford (2014), Landwehr et al. (2016) and Kuehn and Abrahamson (2017) have used Bayesian methods to fit ground-motion models with complex variance structures. Specifically, partially non-ergodic models with regional (Stafford 2014), or path-specific (Landwehr et al. 2016), parameters have been developed, and uncertainties in independent variables have been accounted for in Stafford (2014) and Kuehn and Abrahamson (2017).

The present study builds upon these previous contributions and can be regarded as an extension of the work of Wang and Takada (2009) to consider multiple sites simultaneously whilst also accounting for the more complex random effects structure discussed in Stafford (2014).

In addition to enabling continuous integration of data into ground-motion models, the use of Bayesian regression approaches also enables the probability distribution of the model parameters to be represented. For the purposes of prescribing prior distributions, the present study makes use of a second-moment representation of the full multivariate distribution. However, the updating process provides the full distribution of the parameters at each stage of the updating process. This enables analysts to understand which parameters of the model are less well-constrained than others, and which parameter combinations shown non-trivial correlations. Information regarding parameter uncertainties and parameter correlations is typically provided by traditional regression software. However, the recent developments in ground-motion modelling in which multi-stage regression analyses are conducted has prevented these metrics from being computed. Moving to a Bayesian approach makes these features available once more and improves our understanding of the robustness of the derived models. Having access to the parameter uncertainties also enables prediction intervals to be computed that are useful for defining model-specific epistemic uncertainties within seismic hazard analysis (Arroyo and Ordaz 2011).

In the following section, the Bayesian updating framework employed here is described along with an example application. This description makes use of information that is not currently available for

many ground-motion models and so Section 3 is then devoted to explaining how the required information may be obtained. The implications for the framework presented herein is then discussed in Section 4.

2 Bayesian updating of ground motion models

A general ground motion model can be represented as in Eq. 1:

$$\boldsymbol{y} = \boldsymbol{\mu}(\boldsymbol{X};\boldsymbol{\beta}) + \boldsymbol{Zb} + \boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{y}$ is an $n \times 1$ vector of observed ground motions (usually the logarithm of some intensity measure), $\boldsymbol{\mu}(\boldsymbol{X};\boldsymbol{\beta})$ is an $n \times 1$ vector of mean predictions for a particular set of n_v independent variables $\boldsymbol{X}$ that relate to the rupture scenario and site conditions. The matrix $\boldsymbol{X}$ is often regarded as an $n \times n_v$ dimension matrix, although for a nonlinear hierarchical model one can also think of $\boldsymbol{X}$ as representing a list of independent variables that are required for defining the model. For instance, in the $n \times n_v$ matrix form of $\boldsymbol{X}$, one column would typically represent magnitude, $\boldsymbol{M}$, while another might represent logarithmic average shear-wave velocity, $\ln V_{S,30}$. However, if we only have n_e earthquakes and n_s recording stations in the database, then there is no need to replicate values of these variables simply to populate the $n \times n_v$ matrix.

The p model parameters are defined in the $p \times 1$ vector $\boldsymbol{\beta}$, and these represent the 'fixed effects' or the parameters that reflect the entire 'population' of ground-motions for the region in question. The term $\boldsymbol{Zb}$ corresponds to the q random effects of the model and represents systematic deviations of certain observations within $\boldsymbol{y}$ away from the population mean. For example, a particular event may have larger than average source motions and this will be reflected by a positive element within $\boldsymbol{b}$. The $q \times 1$ vector $\boldsymbol{b}$ is the actual vector of random effects, while the $n \times q$ matrix $\boldsymbol{Z}$ will typically contain partial derivatives of $\boldsymbol{\mu}$ with respect to the random effects. See Stafford (2015a) for an explicit example of this formulation in the context of nonlinear site response.

The general framework for performing mixed effects regression analysis is shown in Eq. 2 (Bates et al. 2015). In this framework, two random variables are considered, $\boldsymbol{\mathcal{Y}}$ which represents an n-dimensional random vector of observed (logarithmic) intensity measure values, and $\boldsymbol{\mathcal{B}}$ representing a q-dimensional vector of random effects. The use of the capital calligraphic font indicates that the model regards these terms as being random variables. The actual observed ground-motions are denoted as $\boldsymbol{y}$ and the *unobserved* random effects are $\boldsymbol{b}$.

$$(\boldsymbol{\mathcal{Y}}|\boldsymbol{\mathcal{B}} = \boldsymbol{b}) \sim \mathcal{N}\left(\boldsymbol{\mu}(\boldsymbol{X};\boldsymbol{\beta}) + \boldsymbol{Zb}, \sigma^2\boldsymbol{I}\right) \tag{2}$$

Equation 2 states that the conditional distribution of $\boldsymbol{\mathcal{Y}}$ given a vector of random effects $\boldsymbol{b}$ is normal with a mean of $\boldsymbol{\mu}(\boldsymbol{X};\boldsymbol{\beta}) + \boldsymbol{Zb}$ and a variance of $\sigma^2\boldsymbol{I}$. If we have non-constant within-event variance (heteroskedasticity), then the variance term changes from $\sigma^2\boldsymbol{I}$ to $\sigma^2\boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is a matrix representing the within-event covariance structure (Stafford 2015a).

For mixed effects models in general, we do not aim to directly estimate the unobserved vector of random effects, $\mathbf{b}$. Rather, the focus is upon estimating the population parameters in $\boldsymbol{\beta}$ along with the variance components contained within the symmetric $q \times q$ covariance matrix, $\boldsymbol{\Sigma}_{\boldsymbol{\vartheta}}$, of the random effects. In general models, the distribution of $\boldsymbol{\mathcal{B}}$ is defined as in Eq. 3. This equation states that the random effects are represented by a multivariate normal distribution with zero-means and a covariance matrix of $\boldsymbol{\Sigma}_{\boldsymbol{\vartheta}}$.

$$\mathcal{B} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\vartheta}}) \tag{3}$$

The subscript $\boldsymbol{\vartheta}$ is used to denote a vector of parameters than can potentially be used to define the covariance matrix. For example, in the case that a non-constant (heteroskedastic) variance structure for the random effects is considered, parameters defining this structure are represented by $\boldsymbol{\vartheta}$.

In early adoptions of mixed effects approaches (e.g. Abrahamson and Youngs 1992), this $q \times q$ 'matrix' was simply a scalar value that has usually been referred to as the between-event variance and is commonly denoted by τ^2. In the simplest crossed formulation in which we have independent random effects for both event-to-event, τ^2, and station-to-station variation, ϕ^2_{S2S}, (Stafford 2014), the covariance matrix for the random effects becomes:

$$\boldsymbol{\Sigma}_{\boldsymbol{\vartheta}} = \begin{bmatrix} \tau^2 & 0 \\ 0 & \phi^2_{S2S} \end{bmatrix} \tag{4}$$

This is the random effects structure that is used in the example application that follows.

In the context of the present study, the parameters of the model consist of the fixed effects parameters

$\boldsymbol{\beta}$, the random effects parameters $\boldsymbol{b}$, and the variance components associated with the random effects (here, the between-event standard deviation, τ, and the station-to-station standard deviation, ϕ_{S2S}) as well as the residual standard deviation ϕ. Generically defining this entire set of parameters as $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \tau, \phi_{S2S}, \phi, \boldsymbol{b}\}$, we can define the joint probability distribution of the model parameters as $\mathcal{P}(\boldsymbol{\theta}; \boldsymbol{x})$, where $\boldsymbol{x}$ represents some initial dataset (or information) that is used to calibrate this distribution. When a new set of observations of ground-motions, $\boldsymbol{y}$, becomes available, it is possible to define a likelihood function, $\mathcal{L}(\boldsymbol{y}|\boldsymbol{\theta}; \boldsymbol{x})$, that describes how likely that set of observations is given the current information in $\boldsymbol{x}$ and the current estimates of the model parameters $\boldsymbol{\theta}$. Bayes' theorem can then be used to update the prior distribution of the model parameters using Eq. 5 and to obtain a posterior (updated) distribution of the model parameters, $\mathcal{P}(\boldsymbol{\theta}|\boldsymbol{y}; \boldsymbol{x})$, given both our initial information and the new data.

$$\mathcal{P}\left(\boldsymbol{\theta}|\boldsymbol{y}; \boldsymbol{x}\right) \propto \mathcal{L}\left(\boldsymbol{y}|\boldsymbol{\theta}; \boldsymbol{x}\right) \mathcal{P}\left(\boldsymbol{\theta}; \boldsymbol{x}\right) \tag{5}$$

In Eq. 5, the expression is shown as a proportionality only. The strict definition of Bayes' rule also includes a normalizing term as a denominator of this expression. However, in the present study, the updating from prior to posterior estimates of the parameter distribution is performed using Markov-Chain Monte Carlo (MCMC) simulation which avoids the need to compute this normalizing term (Stan Development Team 2018).

The basis of the method presented herein is that it is assumed that some prior distribution, $\mathcal{P}(\boldsymbol{\theta}; \boldsymbol{x})$, of the parameters associated with an existing model is available. Commonly used software packages, such as `lme4` (Bates et al. 2015) in `R` (R Core Team 2017), that are used for ground-motion model develop are able to provide this information. In other cases, when this information is not already available, other options exist to derive this in a retrospective manner—as discussed in Section 3. The likelihood function that appears within Eq. 5 is the traditional likelihood function used for mixed effects models.

The logarithm of this likelihood function for a dataset containing n records can therefore be written as in Eq. 6:

$$\begin{aligned}\ln \mathcal{L}(\boldsymbol{y}|\boldsymbol{\theta}; \boldsymbol{x}) \; = \; & -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln \left|\boldsymbol{C}_{\boldsymbol{x};\boldsymbol{\theta}}\right| \\ & -\frac{1}{2} \left[\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{x}; \boldsymbol{\theta})\right]^T \boldsymbol{C}_{\boldsymbol{x};\boldsymbol{\theta}}{}^{-1} \left[\boldsymbol{y} - \boldsymbol{\mu}(\boldsymbol{x}; \boldsymbol{\theta})\right] \end{aligned} \tag{6}$$

Here, $\boldsymbol{C} \equiv \boldsymbol{C}_{\boldsymbol{x};\boldsymbol{\theta}}$ is the global covariance matrix for the entire dataset, and the subscripts $\boldsymbol{x}; \boldsymbol{\theta}$ are used to emphasise the dependence upon components of the parameter vector and the initial information. In simple cases where only event-specific random effects are considered (i.e. non-crossed cases), the covariance matrix has a block diagonal structure and so efficient methods for computing the determinant and inverse are available (Abrahamson and Youngs 1992). However, when the random effects $\boldsymbol{b}$ have the more general covariance matrix $\boldsymbol{\Sigma}_{\vartheta}$, and the residual errors have a covariance defined by $\sigma^2 \boldsymbol{\Lambda}$, then the covariance of the logarithmic motions is defined using Eq. 7.

$$\text{var}\left(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\beta}\right) = \boldsymbol{Z}^T \boldsymbol{\Sigma}_{\vartheta} \boldsymbol{Z} + \sigma^2 \boldsymbol{\Lambda} = \boldsymbol{C} \tag{7}$$

In the present study the evaluation of Eq. 5 is performed using MCMC with the `Stan` language via the `rstan` package (Stan Development Team 2018). Note that the result of this updating is a full distribution of the model parameters and that in some cases empirical correlations will arise among parameters that should, in theory, be independent. For example, over the entire parameter vector $\boldsymbol{\theta}$, herein it is assumed that the fixed effects are independent of the random effects (apart from a perfect correlation that exists for one particular term, to be discussed in the following section). The random effects are also assumed to be independent consistent with the specification in Eq. 4. Furthermore, while the full posterior distribution can be obtained from the results of MCMC, it is not provided in any parametric form. For the purposes of the present study, it is therefore assumed that the fixed effects are jointly distributed according to a multivariate normal distribution such that only a second-moment representation of the full distribution is used. It is assumed that the variance components and random effects are normally distributed, and independent of one another, and so the mean and standard error for each parameter is extracted from the outputs of the MCMC.

When a new event takes place and new recordings become available, the prior distribution of all parameters is defined as the posterior distribution from either the initial analysis (or the previous update). Equation 5 is then evaluated using MCMC only for the records from the new event. Therefore, while databases used for recent models included thousands of records—implying the construction and manipulation of a very large covariance matrix in a traditional analysis—the

proposed Bayesian updating approach works with far fewer records at any given time. This updating process has implications for the computation of random effects, as discussed in the next section.

2.1 Treatment of random effects

Events and stations that appeared in the initial database will have had random effects estimated for them, and these estimates will have some degree of uncertainty that will depend upon how many records each event provided or how many times a recording was made at a given station. Each time a new set of recordings associated with a new event is used to update the model, a new random effect for this event must be computed, but the median prediction for the population model will also change. In order for the previously computed random effects for older events to make sense, it is necessary to adjust these previous estimates to account for the change in the median model predictions. For stations, the newly added event may include recordings at stations for which random effects have previously been computed. In this case, the previous estimates of these station random effects must be updated to account for the change in model median predictions as well as the new data for each station. Therefore, while the fixed and random effects are assumed to be decoupled, there is an intrinsic coupling between the model 'intercept' (essentially a parameter that centres the model with respect to the residuals) and the values of the random effects for both event and station.

The implication of new data changing previous estimates of random effects for events means that random effects for events will change even though the new data that is added is only associated with some new different event. These changes take place due to the adjustment to the intercept just discussed, but also due to the fact that the total variance is partitioned among between-event, station-to-station, and within-event components. If the newly added data changes the relative magnitudes of the variance components then this can also have an impact upon the random effects.

However, a problem associated with the random effects for events is that the variance component associated with these effects represents both the variation in the actual random effects as well as their error estimates. As new data is now introduced for the previous events, the error estimates in the event random effects do not tend to reduce but variations in the model medians are still mapped into corrections to the event random effects. For this reason, the estimates of the between-event variance can be slightly larger than what would be obtained from a traditional analysis. That said, the main situation in which this issue arises is when relatively poorly recorded events are used for which little constraint upon the event random effect is available. In the future as recording networks become increasingly dense, and as databases of ground-motions grow, it becomes possible to only make use of events for which a sufficiently high number of recordings have been made and this will mitigate against this variance inflation. Note that this issue does not arise to the same extent for station random effects and their associated variance as newly added events can contain recordings on stations for which random effects have been computed and so their error estimates can also evolve.

2.2 Example application

To demonstrate how the approach outlined above can work, an example application is presented here using both the traditional and Bayesian approaches. In order to do so, we consider a hypothetical situation in which we travel back in time and start applying the new approach from the mid-1990s.

2.2.1 Dataset used

The empirical database used is that of the NGA-West2 project (Ancheta et al. 2014), and various filters are applied to obtain a selection of records for this example. Records from Class 1 (C1) mainshocks (Wooddell and Abrahamson 2014), having distances within 200km are considered. These records should have a $PGA \geq 1 \times 10^{-4}$ g, and their significant durations should not be more than 2.75 standard deviations above average levels—defined using the prediction equation of Afshari and Stewart (2016). The application of this filter results in 8548 records from 384 events recorded at 3097 stations. Therefore, the average number of records per event is far greater than the average number of records per station.

In the first instance, the total database defined above is restricted to only include records that were obtained in 1995 or earlier. This arbitrary date was selected in order to represent the level of data that was

available at the time that models such as Abrahamson and Silva (1997) were developed. The initial database contained 924 records of 103 events at 629 different recording stations. This database is far smaller than what is commonly used now, but is sufficient to constrain the initial regressions and to provide a starting point for the Bayesian updating approach to be demonstrated.

The initial and total datasets used herein are shown in Fig. 2. Markers in the figure are colour coded according to whether or not they were in the initial dataset. Events that contribute at least 100 records are also annotated by name. While there are clearly many smaller events that are added to the initial dataset, there are also a number of well-recorded events spread over the full magnitude range considered.

Two sets of results are obtained and presented in this section. The first set corresponds to the application of a traditional regression approach in which the entire database currently available is used to constrain the model parameters. Therefore, an initial regression analysis is performed upon the dataset including pre-1996 records and then for each of the 281 earthquakes that occurred since then a new dataset is compiled by adding the new events to the previous dataset and the regression analysis is repeated. The set of results for the traditional approach therefore consists of 282 nonlinear mixed effects regression analyses obtained using the `lme4` package (Bates et al. 2015) in `R` (R Core Team 2017). A crossed mixed effects formulation is adopted in which event-specific and station-specific effects are considered (Stafford 2014).

The second set of results corresponds to the application of Bayesian updating. An initial regression analysis is performed on the initial dataset using Markov-Chain Monte Carlo via the `rstan` package (Stan Development Team 2018) in `R` (R Core Team 2017). Note that this means that the starting point for the traditional and Bayesian approaches is slightly different due to the different regression techniques employed—despite using the same database. From this initial regression, the prior distribution $\mathcal{P}(\boldsymbol{\theta}; \boldsymbol{x})$ was constructed. Thereafter, for each of the 281 new earthquakes, only the records associated with each of these earthquakes is used within the framework of Eq. 5 to update the prior and obtain the posterior $\mathcal{P}(\boldsymbol{\theta}|\boldsymbol{y}; \boldsymbol{x})$. This posterior then becomes the new

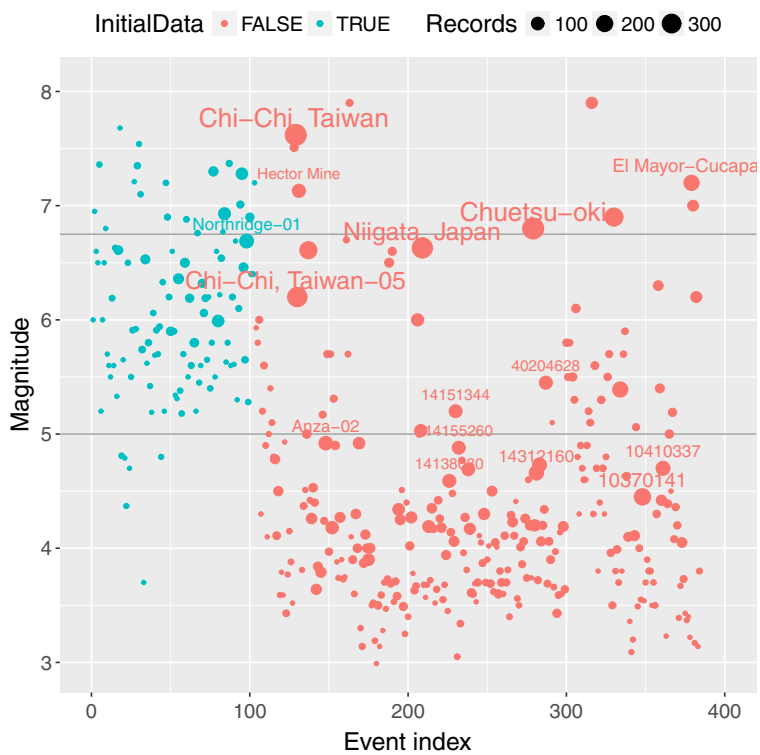

Fig. 2 Distribution of dataset used in the example application. Points are coloured according to whether or not they were used in the initial regression analysis, and are sized by the number of records they contribute. Annotated events contribute at least 100 records. Grey horizontal lines show threshold magnitudes for the functional form considered

prior distribution for the next event and the process is repeated. For the MCMC sampling, four parallel chains were simulated in each case with a total of 4000 samples per chain (of which, only the last 2000 were retained for the model updating).

Ordinarily, for the same dataset, the `lme4` package far outperforms `rstan` from a computational perspective, as the sampling in MCMC can be time consuming. However, in the present application, the overall computational time is actually shorter for the MCMC approach because the dataset used for each case is considerably smaller than the total dataset being used by `lme4`.

Note that the Stan files used for both the initial regression analysis as well as the Bayesian updating are available from https://github.com/pstafford/BayesianStanRegression. The prior distributions used for the initial Bayesian regression are included in these files.

2.2.2 Functional form adopted

The functional form used for the present example is chosen to be a simplification of the recent model of Abrahamson et al. (2014). The reason for selecting this model as the basis of the present example is that the core functional form of the model has not changed significantly since the Abrahamson and Silva (1997) model was published over 20 years ago. The Abrahamson et al. (2014) model therefore provides a good example of a model that has periodically received updates on the basis of discrete changes to ground-motion databases. In evolving from the Abrahamson and Silva (1997) model (where nonlinear site effects were first incorporated into a ground-motion model) to the first NGA model of Abrahamson and Silva (2008), a significant increase in the number of available records took place. Importantly, a number of supporting numerical studies were also performed that enabled certain functional expressions, such as the nonlinear site response (Walling et al. 2008), to be constrained outside of a traditional regression analysis (some of these constraints are retained herein). The latest model of Abrahamson et al. (2014) builds upon the Abrahamson and Silva (2008) version, with the major changes being related to improving the performance for smaller magnitude and longer distance scenarios. Additional refinements to various other components, such as hanging wall effects, have also been made, but are not relevant for the present analysis.

The objective herein is to demonstrate the method and to that end, it was deemed sufficient to omit some of the secondary functional terms related to hanging wall effects, sediment depth and depth of rupture, simply to reduce the number of free regression coefficients. At the same time, it was desirable to not simplify the model to the extent that it would be unrealistic by modern standards. Therefore, the main complexity associated with nonlinear site response is retained in addition to the base magnitude and distance scaling and style-of-faulting effects. Although sediment depth effects are not included, the representation of site effects is actually made more complex through the use of the crossed mixed effects formulation that includes random effects for both events and recording stations (Stafford 2014, 2015a).

The overall functional form for spectral acceleration (although only PGA is considered herein) is a function of moment magnitude, $\boldsymbol{M}$, rupture distance, R_{rup}, binary style-of-faulting flags (F_N for normal and normal-oblique events, and F_R for reverse and reverse-oblique events), and the average shear-wave velocity over the uppermost 30 m, $V_{S,30}$, and can be described by Eq. 8.

$$\ln Sa = f_1(\boldsymbol{M}, R_{rup}) + F_N f_7(\boldsymbol{M}) + F_R f_8(\boldsymbol{M}) + f_5(\widehat{Sa}_r, V_{S,30}) \quad (8)$$

The base magnitude and distance scaling represented by $f_1(\boldsymbol{M}, R_{rup})$ has coefficients and terms that change depending upon the level of magnitude. For the largest events, $\boldsymbol{M} > M_1$ where $M_1 = 6.75$, the function is defined as:

$$f_1^{(\boldsymbol{M}>M_1)} = a_1 + a_5(\boldsymbol{M} - M_1) + a_8(8.5 - \boldsymbol{M})^2 + [a_2 + a_3(\boldsymbol{M} - M_1)]\ln(R) + a_{17}R_{rup} \quad (9)$$

Over the intermediate magnitude range, where $M_2 \leq \boldsymbol{M} \leq M_2$ and $M_1 = 5.0$, the function is defined as:

$$f_1^{(M_2 \leq \boldsymbol{M} \leq M_1)} = a_1 + a_4(\boldsymbol{M} - M_1) + a_8(8.5 - \boldsymbol{M})^2 + [a_2 + a_3(\boldsymbol{M} - M_1)]\ln(R) + a_{17}R_{rup} \quad (10)$$

Finally, for small magnitudes, an additional linear term is added as an extension of the scaling for the intermediate range such that the functional form becomes:

$$f_1^{(\boldsymbol{M}<M_1)} = a_1 + a_4(M_2 - M_1) + a_8(8.5 - M_2)^2 + a_6(\boldsymbol{M} - M_2) + [a_2 + a_3(M_2 - M_1)]\ln(R) + a_{17}R_{rup} \quad (11)$$

In each of the above cases, the distance metric R is derived from the rupture distance R_{rup} to account for magnitude-dependent near-source saturation effects. The expression is defined as:

$$R = \sqrt{R_{rup}^2 + c_{4M}(\boldsymbol{M})^2} \tag{12}$$

with

$$c_{4M}(\boldsymbol{M}) = \begin{cases} c_4 & \text{for } \boldsymbol{M} > 5 \\ c_4 + (c_4 - 1)(\boldsymbol{M} - 5) & \text{for } 4 < \boldsymbol{M} \leq 5 \\ 1 & \text{for } \boldsymbol{M} \leq 4 \end{cases} \tag{13}$$

The parameter $c_4 = 4.5$ is regarded as a constant herein.

The style-of-faulting terms $f_7(\boldsymbol{M})$ and $f_8(\boldsymbol{M})$ have similar functional forms and can be expressed as in Eqs. 14 and 15, respectively.

$$f_7(\boldsymbol{M}) = \begin{cases} a_{11} & \text{for } \boldsymbol{M} > 5 \\ a_{11}(\boldsymbol{M} - 4) & \text{for } 4 < \boldsymbol{M} \leq 5 \\ 0 & \text{for } \boldsymbol{M} \leq 4 \end{cases} \tag{14}$$

$$f_8(\boldsymbol{M}) = \begin{cases} a_{12} & \text{for } \boldsymbol{M} > 5 \\ a_{12}(\boldsymbol{M} - 4) & \text{for } 4 < \boldsymbol{M} \leq 5 \\ 0 & \text{for } \boldsymbol{M} \leq 4 \end{cases} \tag{15}$$

The site response term, $f_5(\widehat{Sa}_r, V_{S,30})$, is a nonlinear function of the reference expected spectral acceleration $\widehat{Sa}_r$ which is defined for a reference velocity value of $V_{S,30} = 1180$ m/s. In the present study, this expected acceleration level is taken as the event- and station-corrected reference acceleration such that the random effect for each event and the random effect for each station (that reflects systematic linear site effects) is propagated through the site response term. Linear site response is assumed for sites having $V_{S,30} > V_{lin}$ where $V_{lin} = 660$ m/s, and values of the shear-wave velocity are capped such that $V^*_{S,30} = V_{S,30}$ if $V_{S,30} < V_1$ and $V^*_{S,30} = V_1$ otherwise. The value of V_1 is taken as 1500 m/s. These values of V_1 and V_{lin} are taken from Abrahamson et al. (2014), as are two other terms within their site response model, $c = 2.4$ and $n = 1.5$, as these parameters arise from external numerical constraints. The site response model varies with the level of $V_{S,30}$. For values of $V_{S,30} \geq V_{lin}$ the response is linear and is defined by:

$$f_5^{V_{S,30} \geq V_{lin}} = (a_{10} + bn) \ln\left(\frac{V^*_{S,30}}{V_{lin}}\right) \tag{16}$$

while for softer sites the response is nonlinear and is defined by:

$$f_5^{V_{S,30} < V_{lin}} = a_{10} \ln\left(\frac{V^*_{S,30}}{V_{lin}}\right) - b \ln\left(\widehat{Sa}_r + c\right) + b \ln\left[\widehat{Sa}_r + c\left(\frac{V^*_{S,30}}{V_{lin}}\right)^n\right] \tag{17}$$

Note that in the model development process followed by Abrahamson et al. (2014), the calibration of parameters was made in a number of steps. Part of the reason for this was to ensure reliable performance across multiple response periods (while only PGA is considered as an example here), but another reason is that their model included additional functional terms. Such a multi-stage approach is not adopted here and 12 fixed effects regression coefficients are solved for simultaneously. The free coefficients are $\boldsymbol{a} = \{a_i, b\}$ where $i \in \{1, \ldots, 6, 8, 10, \ldots, 12, 17\}$.

2.2.3 Regression results

Following the process outlined above, a large set of regression results are obtained using both approaches. Figure 3 shows the variation of the fixed effects estimates, and their associated errors, as a function of the index of the event that is added. As events are added in chronological order, the horizontal axis can also be interpreted as time (although spacing between events is non-constant). Clearly, the results obtained from the two approaches are not identical, and nor should we expect them to be. The leftmost point in all panels of Fig. 3 shows the results from the initial regression using the same dataset for both approaches. The only difference in this particular case is type of model fitting that is performed in the `lme4` and `rstan` packages. These differences should be kept in mind when interpreting the results for other points in the figure.

It can be appreciated that the coefficients primarily associated with distance scaling (a_2, a_3, and a_{17}) tend to show the best agreement. Coefficients a_{10} and b that relate to the site response effects are also actually fairly consistent when one appreciates that there should be a negative correlation between these coefficients if the considered dataset contains records with nonlinear site effects. The coefficients a_{11} and a_{12} that relate to style-of-faulting effects indicate fairly

Fig. 3 Fixed effects estimates for all model parameters using the traditional and Bayesian updating approaches. Vertical lines represent the standard errors in the coefficient estimates

consistent trends. The coefficients that show the greatest deviations are a_4, a_5, and a_8 that relate to the magnitude scaling for moderate to large magnitudes. Part of these differences arise from relatively large differences in the starting estimates from the initial regression analysis and most likely reflects issues associated with over-parameterisation for the initial dataset. The other reason why these parameters show greater volatility is that they have less constraint as a result of lower effective sample sizes. The constraint on the magnitude scaling coefficients comes from the numbers of earthquake events while the constraint upon 'within-event' effects, like distance scaling and site response, comes from the numbers of records. The large difference in numbers of events versus numbers of records explains why the size of the errors for the magnitude scaling terms are relatively large (note the different ordinate scales used in Fig. 3). However, it is noteworthy that despite these reasons, the differences in the parameter estimates from the two approaches are actually very small as one moves to the right-hand-side of the panels. Therefore, as more data is considered in both the traditional and Bayesian approaches, the results tend to converge.

Figure 4 shows the individual variance components obtained as well as the total standard deviation. Note that while the panels shown represent τ (the between-event variability), ϕ_{S2S} (the station-to-station variability), and ϕ (the event- and site-corrected within-event variability), the ϕ_{S2S} values are not directly comparable with other values sometimes reported in the literature (Rodriguez-Marek et al. 2013). As shown in Eqs. 1 and 3, the random effects are operated upon by $\boldsymbol{Z}$. Studies such as Rodriguez-Marek et al. (2013) work directly upon the $\boldsymbol{Zb}$ terms and so can include the effects of $\boldsymbol{Z}$. Furthermore, the inclusion of non-linear site effects in the functional form means that $\boldsymbol{Z}$ will contain elements that decrease with increasing $\widehat{Sa}_r$ (Stafford 2015a), and the values in Fig. 4 do not account for these effects. The total standard deviation computed in this case is therefore *not* actually the standard deviation that might be encountered for many realistic scenarios, but rather reflects the maximum values that can arise under linear site conditions. That said, the estimates in both cases are obtained in a consistent manner and so can be directly compared.

In Fig. 4, the numbers of records that are added to the dataset with each added event are shown as

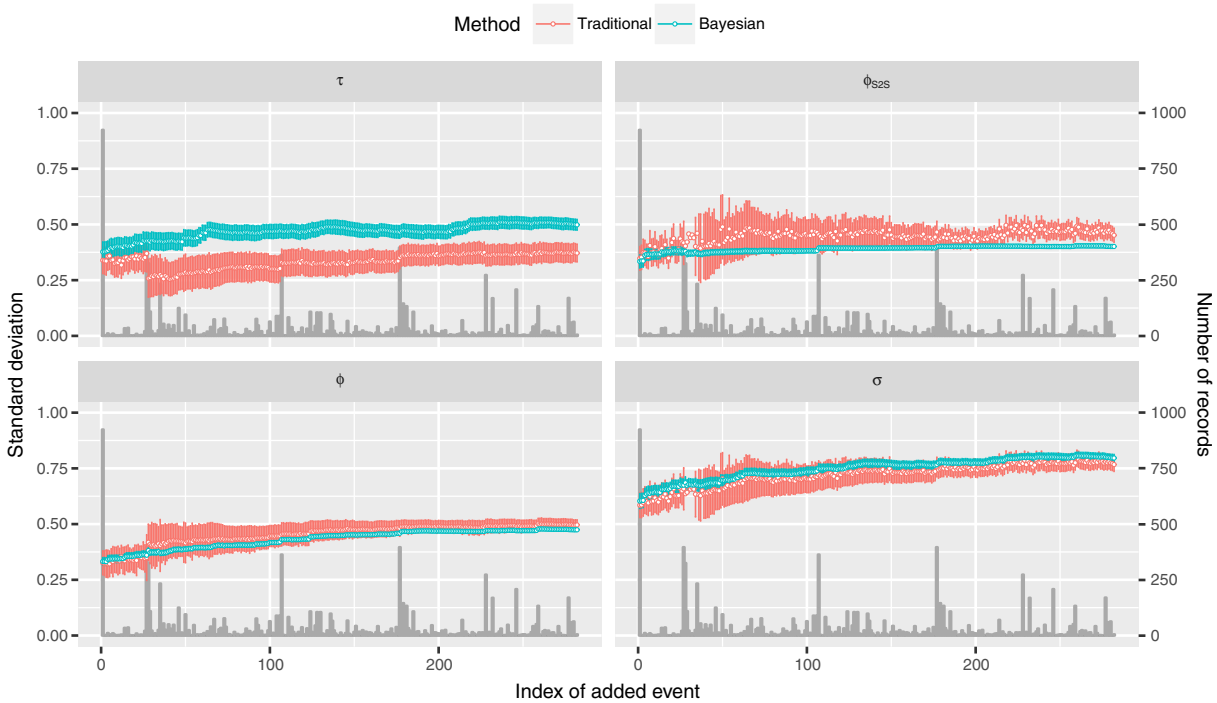

Fig. 4 Variance components, and standard errors, estimated using the traditional and Bayesian approaches. Dark grey bars on the secondary axis show the numbers of records added for each new event

vertical bars with their amplitudes defined by the secondary axis. What is very clear in the panel related to τ is that the introduction of particular well-recorded events can have a significant impact upon the τ estimate. For this dataset, the first major change occurs under the traditional regression approach when the Chi-Chi earthquake records are added. Their introduction causes a significant imbalance in the dataset, with the vast majority of prior events having far fewer records (see Fig. 2). This imbalance, and the particular characteristics of the Chi-Chi records, changes the partitioning of total variance among the various components with a particular increase in the ϕ_{S2S} values. Note that for these variance components, the error estimates are obtained using the two-way crossed classification (without interaction) of Searle (1971) (section 11.6) and are only approximate for this particular nonlinear model. The data imbalance and nonlinearity within the Chi-Chi data causes numerical issues with these approximate error estimates in Fig. 4. This only applies to the traditional regression results, not the Bayesian analyses that directly provide more appropriate estimates.

As discussed in Section 2.1, the values of the random effects for event and station evolve in time as more data is integrated into the model. Figure 5 provides examples of this evolution for the three stations within the database with the greatest number of recordings. In Fig. 5, it can be appreciated that there is a significant degree of consistency between the random effects estimated using both approaches, and that the addition of new data (as shown by the vertical

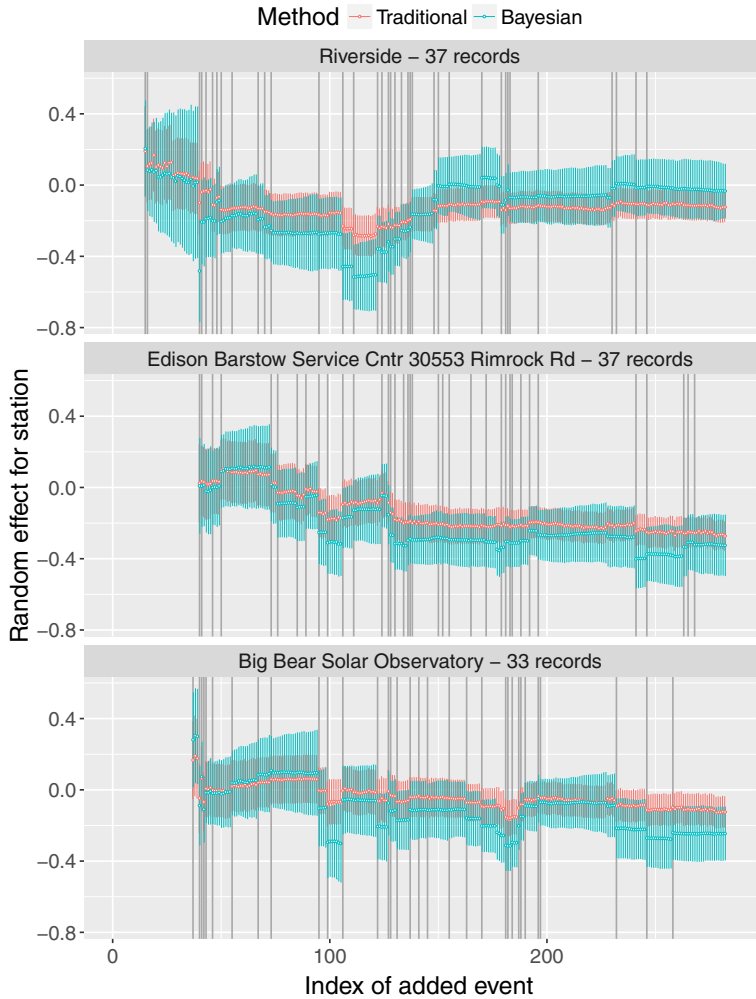

Fig. 5 Evolution of station random effects for the three stations with the most recordings. Vertical grey lines are shown when the added event includes a recording at the corresponding station. Error bars show the standard errors in the random effects

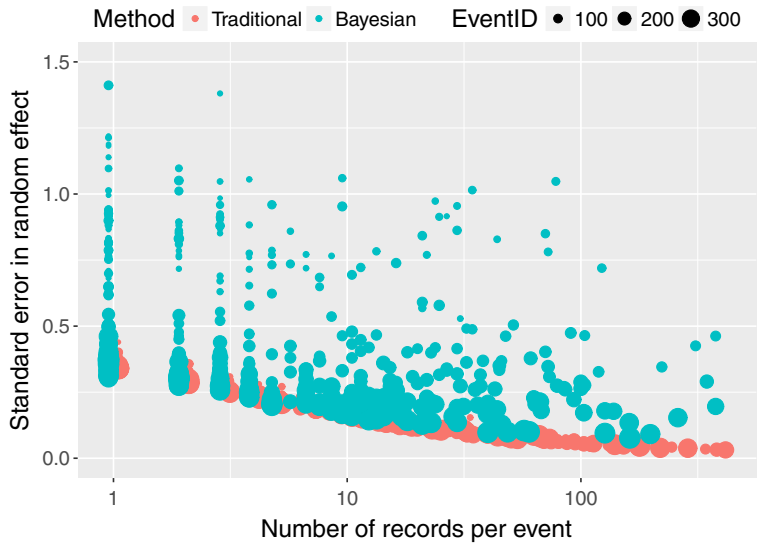

Fig. 6 Standard errors in the estimates of event random effects as a function of the numbers of records per event using both the traditional and Bayesian approaches. The marker size scales with the order in which the event was added to the analysis database. Note that markers have been offset slightly in the horizontal direction to avoid overlap

dark grey lines) has a similar impact regardless of the method. It is therefore apparent that site-specific adjustments can be obtained for individual locations within the dataset by using the Bayesian approach.

However, the three example sites shown in Fig. 5 are also cases for which relatively strong empirical constraint exists. In general, with current databases the numbers of available records for each event, or for each station, are far lower. Figures 6 and 7 show how the error estimates for the random effects vary as a function of the numbers of records per event and station, respectively. Figure 6 also shows plots each point according to its event identifier (which also defines the temporal order with which events were recorded). This plot shows that the events within the initial dataset, having event identifiers of 103 or less, typically have the largest errors in the random effects, but as more and more events are added the error estimates reduce and tend towards those associated with the traditional approach.

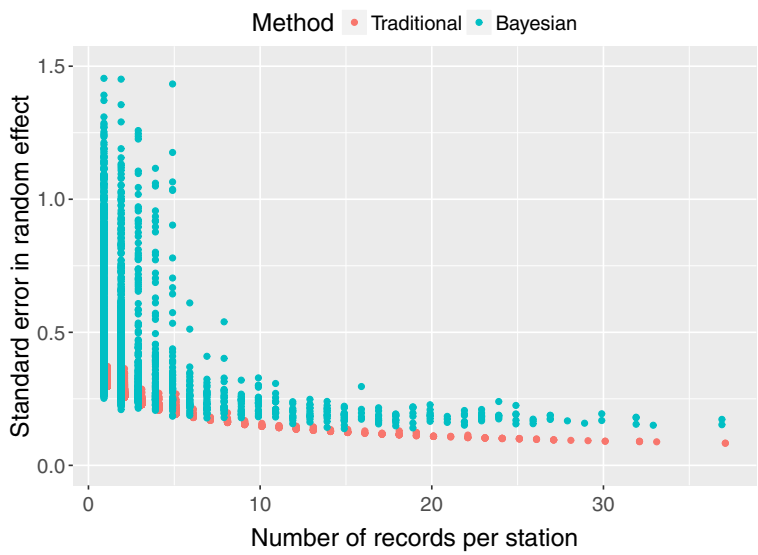

Fig. 7 Standard errors in the estimates of station random effects as a function of the numbers of records per event using both the traditional and Bayesian approaches. Note that markers have been offset slightly in the horizontal direction to avoid overlap

In Fig. 7, it can be seen that stations with very few recordings have very little constraint upon their random effects and that the error estimates in the Bayesian case are far greater than those from the traditional regression analysis.

3 Retrospective estimation of parameter covariance

In the analyses presented thus far it has been assumed that some initial prior $\mathcal{P}(\boldsymbol{\theta}; \boldsymbol{x})$ exists. However, many ground-motion models already exist and so it is useful to outline how this prior might be constructed given some existing model. Empirical ground-motion models have traditionally been derived through a regression analysis using large numbers of recordings from many earthquakes. The process that is commonly adopted is to postulate a functional form for the ground-motion model and to then undertake a regression analysis to infer the values of the model parameters. Finding the final form of the model is often a process involving a balance between including enough functional terms (and model parameters) to capture the main trends in the empirical data, and not over-fitting the model. The traditional way to assess whether a model is being over-fit is to test for the statistical significance of the model parameters that are found from the regression analysis. These tests essentially compare the size of the standard errors in the estimates of the coefficients (factored by a student-t statistic that reflects the database size and the degree of confidence that one wishes to use—normally 95% confidence) with the magnitude of the fitted coefficients.

Once one defines the final model then the outputs of the regression analysis include the final fitted model parameters (often called the 'regression coefficients'), the variance components of the model (such as the between-event and within-event variability), but also the covariance matrix of the model parameters. This covariance matrix includes (on its diagonal) the squared standard errors of the model coefficients, and the covariances (correlations) among coefficients on the off-diagonals. For a well-fitted and robust model, one aims to have small standard errors, and small correlations among the coefficients. The former point means that the individual parameters are well-defined, while the second points about the low correlations means that the model parameters will be less sensitive to a change in the dataset and that the functional terms tend to reflect distinct scaling effects. Furthermore, when one wishes to make a future prediction this covariance matrix of the model parameters can be used to define a prediction interval and this prediction interval is a way of estimating the model-specific epistemic uncertainty, and can be used within an approach to infer logic-tree weights, among other things (Arroyo and Ordaz 2011). This covariance matrix of the fixed effects can also be used as the basis of the $\mathcal{P}(\boldsymbol{\theta}; \boldsymbol{x})$.

However, in recent years, it has become increasingly common to develop semi-empirical models, in which some components of the model are constrained through numerical simulation, or theoretical considerations. This is especially the case for nonlinear site response effects, as discussed earlier. In addition, the functional forms are much more complex and in order to ensure a 'smooth' behaviour in terms of spectral shapes, among other things, it has become common to determine the model parameters through a series of regression steps (Abrahamson et al. 2014). In this approach, one first focusses upon the scaling with respect to a particular parameter, say magnitude, and then fixes the scaling with respect to this parameter before then inferring the remaining parameters in future steps. In these subsequent steps, the previously defined model parameters are assumed to be known constants, as are the theoretically defined, or numerically defined constraints.

A problem with these developments, however, is that it is no longer possible to compute standard regression metrics such as the covariance matrix, or standard errors, for the model coefficients. This also means that one cannot define model-specific epistemic uncertainty through the use of prediction intervals. It therefore becomes hard to understand whether these complex functional forms are really supported by the empirical data they were derived from, and which components of the model are better constrained than others. However, for the present article, the key issue is that it becomes harder to identify an appropriate $\mathcal{P}(\boldsymbol{\theta}; \boldsymbol{x})$ as the starting point of future updates.

To overcome the limitations of not being able to compute these metrics during the regression or model-development stage, the present section highlights an option for estimating the parameter covariances (or at least variances) in a retrospective manner. That is, following the development of the model one can then

look back and assess how stable the various components of the model are and use this information to construct the prior $\mathcal{P}(\boldsymbol{\theta};\boldsymbol{x})$. In addition, this approach need not necessarily be applied to the dataset used for the model development. It could be the case that one is interested in assessing how stable a particular model is for use in another region for which there is some empirical data available. For such cases, the following approach can also hold.

3.1 Fisher Information matrix

The Fisher Information matrix is defined in terms of the log-likelihood function of a particular dataset given some model parameters (or regression coefficients).

Consider that an existing ground-motion model is defined by its (logarithmic) mean, and standard deviation, both of which may be functions of a set of n_p model parameters $\boldsymbol{\theta}$ (as previously, this vector can contain the fixed effects and variance components). For a given dataset, we can consider the logarithmic spectral ordinates to be defined as $\boldsymbol{y} \equiv \ln \boldsymbol{S_a}(T)$. We then define the log-likelihood function for this data, given the parameters $\boldsymbol{\theta}$ as $\ln \mathcal{L}(\boldsymbol{y}|\boldsymbol{\theta};\boldsymbol{x})$ as in Eq. 6. In principle, if the dataset used for the model development is also used for the evaluation of this log-likelihood, then we should expect the computed likelihood to be a global maximum.

The Fisher Information matrix is then a $n_p \times n_p$ matrix that is defined in terms of this log-likelihood function as:

$$\boldsymbol{I}_{ij}(\boldsymbol{\theta}) = -\mathbb{E}\left[\frac{\partial^2 \ln \mathcal{L}(\boldsymbol{y}|\boldsymbol{\theta};\boldsymbol{x})}{\partial\theta_i \partial\theta_j}\right] \tag{18}$$

Note that the expectation operator in practice really just means that these mixed partial derivatives are evaluated at the expected values of the model parameters. That is, we use the fitted model parameters to determine the derivatives that feed into the Fisher Information matrix. Therefore, this Fisher Information matrix is equivalent to the Hessian matrix of the log-likelihood function with respect to the n_p model parameters. For some elaborate ground-motion models, the analytical evaluation of these second derivatives can be a cumbersome process. However, the derivatives can be relatively easily obtained through the use of algorithmic or automatic differentiation (Molkenthin et al. 2014, 2015).

3.2 Cramér-Rao bounds

With the Fisher Information matrix defined, one can then apply a well-known approach to infer bounds upon the covariances of the model parameters. The Cramér-Rao bounds place constraints upon the variance (and covariance) of the model parameters using the following relation.

$$\mathrm{cov}(\boldsymbol{\theta}) \geq \boldsymbol{I}(\boldsymbol{\theta})^{-1} \tag{19}$$

That is, we can state that the covariances of the model parameters must be equal to or greater than the values implied by the inverse of the Fisher Information matrix. This result makes intuitive sense when we appreciate that the Fisher Information matrix is describing the curvature of the log-likelihood surface around the global maximum. The greater the curvature for variations of a given parameter, the stronger the constraint upon the optimal value of that parameter.

Importantly, we can also make use of the lower bound of this covariance matrix as a reasonable prior. That is, we can define the prior distribution of an existing model according to Eq. 20, as a multi-variate (of dimension n_p) normal distribution with the published parameter values, $\hat{\boldsymbol{\theta}}$, as the mean and with the covariance matrix coming from the Fisher Information matrix (evaluated at these published parameter values).

$$\mathcal{P}(\boldsymbol{\theta};\boldsymbol{x}) \approx \mathcal{N}_{n_p}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{I}(\hat{\boldsymbol{\theta}})^{-1}\right) \tag{20}$$

Note that this equation strictly applies when the mean vector used to compute the information matrix is cast as $\boldsymbol{\mu}(\boldsymbol{\theta}) \equiv \boldsymbol{\mu}(\boldsymbol{x};\boldsymbol{\theta}) + \boldsymbol{Zb}$. That is, the curvature should be evaluated around the conditional estimates of the mean after correcting for the random effects. However, in most cases, while published articles present figures showing the random effects, the actual numerical values are not made available. In these cases, an approach that can be adopted is to compute the total residuals for the model and to then run a random effects regression analysis on these residuals in order to obtain the estimates of the random effects.

It is also important to note that this approach can be used on all of the model parameters (even those that have been constrained during the model development). In cases where there is little or no empirical constraint upon the parameters then this will be reflected by very small curvatures that are then translated into large variances (and hence uninformative priors). The

use of such weak priors for the subsequent regression may cause numerical instabilities however and so care should be taken when relaxing previously constrained parameters. A more robust way to handle such cases is to place an informative prior centred at the constrained value of the parameter, but to allow for sampling away from this value. The Cauchy distribution is useful in such cases as the vast majority of samples will still be draw near the previously constrained value, whilst still allowing for the consideration of values some distance from this past estimate.

3.3 Multivariate normal distribution

For the special case in which the observations may be assumed jointly normally distributed (which is generally assumed within engineering seismology) the expressions for the Fisher Information matrix can be cast in terms of first-order partial derivatives rather than through the use of the Hessian matrix.

Regard the n observations of logarithmic spectral ordinates as being distributed according to an n dimensional multivariate normal distribution:

$$\boldsymbol{y} \sim \mathcal{N}_n \left[\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{C}(\boldsymbol{\theta})\right] \tag{21}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta}) \equiv \boldsymbol{\mu}(\boldsymbol{x}; \boldsymbol{\theta})$ is the vector of mean logarithmic predictions and $\boldsymbol{C}(\boldsymbol{\theta}) \equiv \boldsymbol{C}_{\boldsymbol{x};\boldsymbol{\theta}}$ is the covariance matrix for the data.

In this particular case, the Fisher Information matrix has elements that are defined by:

$$\boldsymbol{I}_{ij}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})^T}{\partial \theta_i} \boldsymbol{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j} + \frac{1}{2} \mathrm{tr}\left[\boldsymbol{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{C}(\boldsymbol{\theta})}{\partial \theta_i} \boldsymbol{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{C}(\boldsymbol{\theta})}{\partial \theta_j}\right] \tag{22}$$

where here the tr signifies the trace operator. So, under this special case, one only requires the computation of the Jacobian. In many cases, the analytical forms of most of these partial derivatives will have already been computed as they are often needed within nonlinear regression analyses. For example, in the example application of Section 2.2, the relevant partial derivatives with respect to the mean, $\partial \boldsymbol{\mu}(\boldsymbol{\theta})/\partial \theta_i$, are already required by the `lme4` package. The derivatives of the covariance matrix can be more challenging to compute, but this depends upon the model formulation. Either way, if automatic differentiation is adopted then these terms are readily obtainable.

4 Discussion

The Bayesian updating approach advocated in the previous sections has been shown to provide results that are broadly consistent with those from more traditional regression analyses, whilst also providing additional information such as error estimates in the variance components. The method is predicated upon the assumption that we now possess a sufficient understanding of how ground-motions scale to enable a robust functional form to be proposed. Given this functional form, the model parameters can then be continuously updated as new events provide recordings.

The example application presented in Section 2.2, while making use of real data was hypothetical and started with an initial regression based upon a dataset representative of the state-of-the-art from more than two decades ago. Many of the differences between the Bayesian updating and traditional results presented in that example can be explained by the lack of constraint provided by the relatively small initial data set. The results should improve significantly when starting with existing models that already have far stronger empirical constraint.

The example application presented in Section 2.2 provides a proof of concept demonstration of how Bayesian updating can be used to constrain parameters of ground-motion models as well as to obtain site-specific predictions for multiple sites simultaneously. This is currently also possible using traditional regression techniques, but these require the entire database to be reanalysed for each new event that is integrated into the database. This feature of the Bayesian updating approach has implications for partially non-ergodic hazard analysis, as discussed in the following section.

4.1 Implications for partially non-ergodic hazard analysis

Figure 7 shows that as the numbers of recordings at particular stations increase, the Bayesian approach is able to appropriately refine the estimates of the random effects. This means that the Bayesian updating approach can be used to help determine site-specific adjustments that should be made to ground-motion models. Wang and Takada (2009) demonstrated how multiple model coefficients could be adjusted to target

a specific site within a Bayesian updating framework, and their method could be considered as being equivalent to that presented here. However, while the approach of Wang and Takada (2009) focusses upon making adjustments to ground-motion models in order to target a single specific site, their approach focussed upon updating the fixed effects coefficients of the model and therefore creates an entirely new model that is appropriate for a given location. In contrast, the approach advocated here focusses upon station-specific random effects and this means that when the new site-specific data is added to the model, both the fixed effects and random effects benefit from this data. In the example application provided here, the random effect for each station was essentially applied to the linear site amplification term only. However, as demonstrated by Stafford (2014), site-specific random effects could be included into other parameters of the model as well. For example, Fig. 3 suggests that the parameters a_{10} and b had a negative correlation. If the extent of this correlation is not the same for all sites (and we should not expect that it would be), then this sort of relationship may be represented by including coupled station-specific random effects on both a_{10} and b. This naturally adds a significant amount of complexity to the variance structure for the model, but the Bayesian approach, particularly as implemented using the `Stan` language (Stan Development Team 2018), is readily able to model such complexity.

The key difference between the present approach and that of Wang and Takada (2009) is therefore that the approach herein simultaneously refines estimates of station effects for all stations for which recordings have been made. Importantly, the approach is also able to provide estimates of the uncertainties associated with these station effects which is important for constraining epistemic uncertainty within hazard applications. In cases where a future station of interest is located away from recording stations, an estimate of the random effect and associated uncertainty can be obtained via kriging, or similar spatial techniques such as Gaussian process regression (Landwehr et al. 2016).

Station effects have been the focus of the present discussion as the example application targeted these effects. The Bayesian approach currently provides estimates of the station random effects and their errors at locations for which recordings have been made. In practical cases where a partially non-ergodic hazard analysis is conducted (e.g. Rodriguez-Marek et al. 2014), a full site response analysis will be performed rather than simply using a site correction factor. However, there are still applications where the station random effects (and their associated uncertainties) can be useful. For spatial applications, such as the generation of seismic hazard maps, or simulations of ground-motion fields for portfolio loss estimation, it is not usually practically realistic to conduct detailed site response analyses across the entire spatial region - although at least one example of this exists (Bommer et al. 2017). In these cases, an important improvement to the ground-motion modelling approach, that relaxes the ergodic assumption to some extent, is to obtain a spatial map of station random effects and their uncertainties (using methods like kriging, as previously mentioned). The hazard maps, or ground-motion fields, can then take these spatial variations into account without having to include the full station-to-station variability within the aleatory variability of the ground-motion model.

The error estimates for the station random effects can also be used to constrain the level of epistemic uncertainty that is modelled when more detailed site-specific site response analyses are performed. This is particularly useful as defining the appropriate levels of epistemic uncertainty for these site response calculations is a non-trivial exercise (Rodriguez-Marek et al. 2014).

In addition to event and station effects that have been considered in the present article, the Bayesian updating approach can also be used to progressively make path-specific adjustments to models. This could be done using a regional classification of travel paths, such as considered by Stafford (2014), but is more elegantly handled through a spatially-continuous representation of model parameters as shown by Landwehr et al. (2016).

4.2 Expansion for Big Data and continuous updating

It is inevitable that ground-motion databases will increase in size significantly over the coming years and decades. This increase in numbers of records, coupled with increasingly complex approaches to modelling variance structures (e.g. Stafford 2014, 2015a) dictate that the computational demands associated with model development will become non-trivial. The Bayesian updating approach outlined herein eliminates a lot of this demand and scales very well with

both increases in numbers of records and complexity of variance structures.

The method also affords one the opportunity to have new ground-motion data integrated into models as soon as the meta-data for the event is available. Note that this data could also be integrated in near real time if one also accounts for the uncertainties associated with initial magnitude and finite-fault attributes, as the Bayesian regression framework is well-suited to accommodate uncertainties in the predictor variables (Moss and Der Kiureghian 2006; Moss 2011; Stafford 2014; Kuehn and Abrahamson 2017).

4.3 Relaxation of constrained parameters

In Section 2.2, a simplification of the Abrahamson et al. (2014) ground-motion model was employed, and it was noted that various parameters of this model have been externally constrained by numerical simulations. The framework presented thus far can also be adapted to allow these constant parameters to gradually be calibrated by the empirical data. This can be done either by adding a new parameter to the vector $\boldsymbol{\theta}$, and adding corresponding entries into $\mathcal{P}(\boldsymbol{\theta}; \boldsymbol{x})$, or by initially representing these 'constants' as random variables, but imposing very strong prior distributions on these. The framework presented here therefore provides a mechanism for smoothly transferring constraint from theoretical or numerical bases to empirical bases as appropriate.

5 Conclusions

This article presents a new approach to the development of ground-motion models that can scale to meet the 'Big Data' needs of the future. Rather than re-calibrating an entire ground-motion model as new data becomes available, the existing models are represented in their traditional forms in addition to the covariance matrix of the model parameters, and error distributions of any random effects parameters. As new data becomes available this can be continuously integrated by using Bayesian updating such that only the new records from the most recent earthquake are considered at a given time. The complete history of all previous events is reflected through the joint probability distribution of all model parameters, variance components, and random effects. For the example considered in this study, this approach has proven to have significant computational advantages even with the size of existing datasets.

While Bayesian updating is a well-established method in other fields, it has been difficult to advocate its usage for empirical ground-motion modelling until now because the ground-motion databases have not been sufficiently rich and functional forms used for these models were still evolving. There is likely to be more such evolution in the future, but it is also clear that functional forms are stabilising and that the key attributes of these functional expressions enable successful predictions of motions for a very broad range of scenarios. The challenges that will arise with time related to the treatment of large amounts of data can be well met through the adoption of Bayesian approaches.

As functional forms have become more complex, the challenges in calibrating the parameters of these models have also increased. This has occurred to the extent that multi-stage regression procedures must be adopted and this has meant that information about the parametric uncertainty has been lost. The adoption of Bayesian approaches enables the full distribution of the model parameters to be represented and ensures that analysts can continue to evaluate the overall robustness of their models, as well as the relative robustness of individual functional terms.

Acknowledgements This article has benefited from the comments of two anonymous reviewers. I am very grateful for the time and effort that these reviewers put into helping me improve this article. All of the analyses in this article have been made possible by the existence of the `R` (particularly the `lme4`, `rstan`, `Rcpp` and `ggplot2` packages) and `Stan` software, and the ground-motion data from the NGA-West2 project. I am sincerely grateful to the developers and maintainers of these packages, and to the hosts and curators of the NGA-West2 database.

References

Abrahamson NA, Silva WJ (1997) Empirical response spectral attenuation relations for shallow crustal earthquakes. Seismol Res Lett 68(1):94–127

Abrahamson N, Silva W (2008) Summary of the Abrahamson & Silva NGA ground-motion relations. Earthq Spectra 24(1):67–97

Abrahamson NA, Youngs RR (1992) A stable algorithm for regression analyses using the random effects model. Bull Seismol Soc Am 82(1):505–510

Abrahamson NA, Silva WJ, Kamai R (2014) Summary of the ASK14 ground motion relation for active crustal regions, vol 30. Earthquake Engineering Research Institute

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

Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BSJ, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T, Donahue JL (2014) NGA-West2 database. Earthq Spectra 30(3):989–1005

Arroyo D, Ordaz M (2010a) Multivariate Bayesian regression analysis applied to ground-motion prediction equations, part 1: theory and synthetic example. Bull Seismol Soc Am 100(4):1551–1567

Arroyo D, Ordaz M (2010b) Multivariate Bayesian regression analysis applied to ground-motion prediction equations, part 2: numerical example with actual data. Bull Seismol Soc Am 100(4):1568–1577

Arroyo D, Ordaz M (2011) On the forecasting of ground-motion parameters for probabilistic seismic hazard analysis. Earthq Spectra 27(1):1–21

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48

Bommer JJ, Stafford PJ, Edwards B, Dost B, van Dedem E, Rodriguez-Marek A, Kruiver P, van Elk J, Doornhof D, Ntinalexis M (2017) Framework for a ground-motion model for induced seismic hazard and risk analysis in the Groningen Gas Field, The Netherlands. Earthq Spectra 33(2):481–498

Douglas J (2018) Ground motion prediction equations 1964-2018. Tech. rep. University of Strathclyde, Glasgow

Hermkes M, Kuehn NM, Riggelsen C (2013) Simultaneous quantification of epistemic and aleatory uncertainty in GMPEs using Gaussian process regression. Bull Earthq Eng 12(1):449–466

Kuehn NM, Abrahamson NA (2017) The effect of uncertainty in predictor variables on the estimation of ground-motion prediction equations. Bull Seismol Soc Am 108(1):358–370

Kuehn NM, Riggelsen C, Scherbaum F (2011) Modeling the joint probability of earthquake, site, and ground-motion parameters using Bayesian networks. Bull Seismol Soc Am 101(1):235–249

Landwehr N, Kuehn NM, Scheffer T, Abrahamson N (2016) A nonergodic ground-motion model for California with spatially varying coefficients. Bull Seismol Soc Am 106(6):2574–2583

Molkenthin C, Scherbaum F, Griewank A, Kuehn N, Stafford PJ (2014) a study of the sensitivity of response spectral amplitudes on seismological parameters using algorithmic differentiation. Bull Seismol Soc Am 104(5):2240–2252

Molkenthin C, Scherbaum F, Griewank A, Kuehn N, Stafford PJ, Leovey H (2015) Sensitivity of probabilistic seismic hazard obtained by algorithmic differentiation: a feasibility study. Bull Seismol Soc Am 105(3):1810–1822

Moss RES (2011) Reduced sigma of ground-motion prediction equations through uncertainty propagation. Bull Seismol Soc Am 101(1):250–257

Moss RES, Der Kiureghian A (2006) Incorporating parameter uncertainty into attenuation relationships. In: U.S. National conference on earthquake engineering, San Francisco, pp 1–10

Power M, Chiou B, Abrahamson N, Bozorgnia Y, Shantz T, Roblee C (2008) An overview of the NGA project. Earthq Spectra 24(1):3–21

R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rodriguez-Marek A, Cotton F, Abrahamson NA, Akkar S, Al Atik L, Edwards B, Montalva GA, Dawood HM (2013) A model for single-station standard deviation using data from various tectonic regions. Bull Seismol Soc Am 103(6):3149–3163

Rodriguez-Marek A, Rathje EM, Bommer JJ, Scherbaum F, Stafford PJ (2014) Application of single-station sigma and site-response characterization in a probabilistic seismic-hazard analysis for a new nuclear site. Bull Seismol Soc Am 104(4):1601–1619

Searle SR (1971) Linear models. Wiley, New York

Stafford PJ (2014) Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models. Bull Seismol Soc Am 104(2):702–719

Stafford PJ (2015a) Extension of the random-effects regression algorithm to account for the effects of nonlinear site response-short note. Bull Seismol Soc Am 105(6):3196–3202

Stafford PJ (2015b) Variability and uncertainty in empirical ground-motion prediction for probabilistic hazard and risk analyses. In: Advances in performance-based earthquake engineering. Springer International Publishing, Cham, pp 97–128

Stan Development Team (2018) RStan: the R interface to Stan

Walling M, Silva W, Abrahamson N (2008) Nonlinear site amplification factors for constraining the NGA models. Earthq Spectra 24(1):243–255

Wang M, Takada T (2009) A Bayesian framework for prediction of seismic ground motion. Bull Seismol Soc Am 99(4):2348–2364

Wooddell KE, Abrahamson NA (2014) Classification of main shocks and aftershocks in the NGA-West2 database. Earthq Spectra 30(3):1257–1267