On the Estimation of Ground-Motion Duration Models with an Application to the M9 Simulations

Nicolas M. Kuehn*[0] and Melanie Walling†[1]

[1] *GeoEngineers Inc.*

Abstract

Different parameterizations, related to mixed-effects structure, of ground-motion models for the duration of strong shaking are revisited. A new parameterization is proposed which can better account for the separation of source and path duration in additive ground-motion duration models. In addition, different distributions for the likelihood of ground-motion duration given the predictor variables are investigated. Traditionally, duration has been modeled as lognormally distributed, which makes sense for a multiplicative model, but less so for an additive model. Models using a lognormal and Gamma likelihood are compared, using the M9 simulations (Frankel et al., 2018) as an underlying data set. In general, differences between models in terms of predictions are small, but the models using a Gamma distribution perform slightly better than the ones employing a lognormal distribution. Furthermore, the new parameterization outperforms traditional models. All models are estimated via Bayesian inference, accounting for epistemic uncertainty through the posterior distribution of the parameters.

1 Introduction

Typically, seismic hazard analysis is calculated form ground-motion parameters that characterize the amplitude of the shaking. A complete description of ground motion includes a measure of duration. Ground-motion duration can influence structural damage (e.g. Chandramohan et al., 2016), liquefaction (e.g. Green et al., 2020), and slope-stability (e.g. Wang et al., 2021). Several different duration measures have been proposed, see Bommer and Martínez-Pereira (1999) for a comprehensive overview. Different duration measures can be more useful for different applications.

Similar to ground-motion models (GMMs) for peak parameters or response spectral ordinates (pseudo-spectral acceleration, PSA), one can derive empirical GMMs for ground-motion duration measures (e.g. Bommer et al., 2009; Kempton and Stewart, 2006; Afshari and Stewart, 2016; Jaimes and García-Soto, 2021; Bahrampouri et al., 2021; Anbazhagan et al., 2017; Yaghmaei-Sabegh et al., 2022; Du and Wang, 2017). PSA GMMs model the target variable as a multiplication of different terms corresponding of source, path, and site effects. By contrast, duration GMMs come in two distinct fashions: multiplicative, where total duration is modeled as source duration *times* path duration (e.g. Bommer et al., 2009; Jaimes and García-Soto, 2021; Bora et al., 2014; Yaghmaei-Sabegh and Hassani, 2020), and additive, where total duration equals source duration *plus* path duration (e.g. Abrahamson and Silva, 1996; Bahrampouri

*kuehn@ucla.edu

†mwalling@geoengineers.com

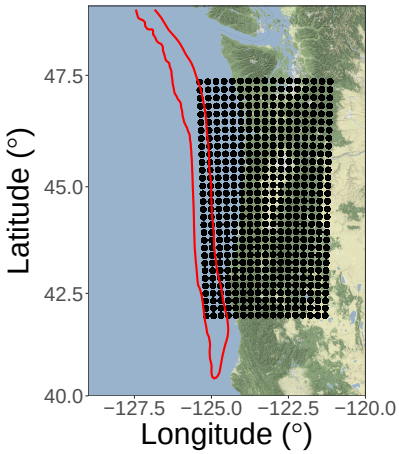

Figure 1: Location of M9 stations used in the analysis. Red line is a contour line of R_{RUP} = 10km.

et al., 2021; Kempton and Stewart, 2006; Afshari and Stewart, 2016). An additive model assumes that the dominant contributor to path duration is the time difference of arrival times of compression and shear waves (Bahrampouri et al., 2021), in which case the path duration should be magnitude-independent.

Multiplicative duration models have the same structure as PSA GMMs, and can be fit in the same manner, to account for between-event and site-to-site residuals via random effects models (Stafford, 2014). Additive models, however, require more care to properly account for event terms associated with event-specific source durations (Walling et al., 2018). Here, we discuss the implications of different parameterizations of additive duration models, and propose a model structure that better accounts for the separation of source and path durations than in previous models. In addition, we investigate the Gamma distribution as a way to model the likelihood of ground-motion durations. Duration models typically assume a lognormal distribution, which is very reasonable for multiplicative models, but its application to an additive model of positive terms is less natural. We fit models with both a lognormal and Gamma likelihood, and compare the estimated models in terms of their predictions.

We estimate the different models based on data from the M9 simulations (Frankel et al., 2018). This provides a great data set to illustrate the different models. Compared to real data sets with observed data, the data distribution is more homogeneous. In particular, there is abundant data at short distances, which is important for duration models, since the source duration represents the model prediction at zero distance.

The paper is organized as follows: we very briefly describe the data, then lay the foundation of the duration models in Section Intuition of Duration Model. Details about the M9 models are found in Section Duration Models for M9 Simulations, followed by results and a discussion.

2 Data

We illustrate the different duration models on simulated ground motions from the M9 project (Frankel et al., 2018). The M9 simulations comprise simulations of 30 different magnitude 9 events on the Cascadia subduction zone. The nice thing about simulated data is that the distance range is "complete", i.e. we do not have to worry about missing data in particular at short distances. We use a subset of the M9 simulations (mainly to reduce the number of data). Figure 1 shows a map of the simulated stations that we use. Selection is also guided by the desire to have only stations in one direction of the rupture (i.e. only towards the coast).

In total, we use 15,810 observations from 30 simulations, i.e. 527 observations per simulation.

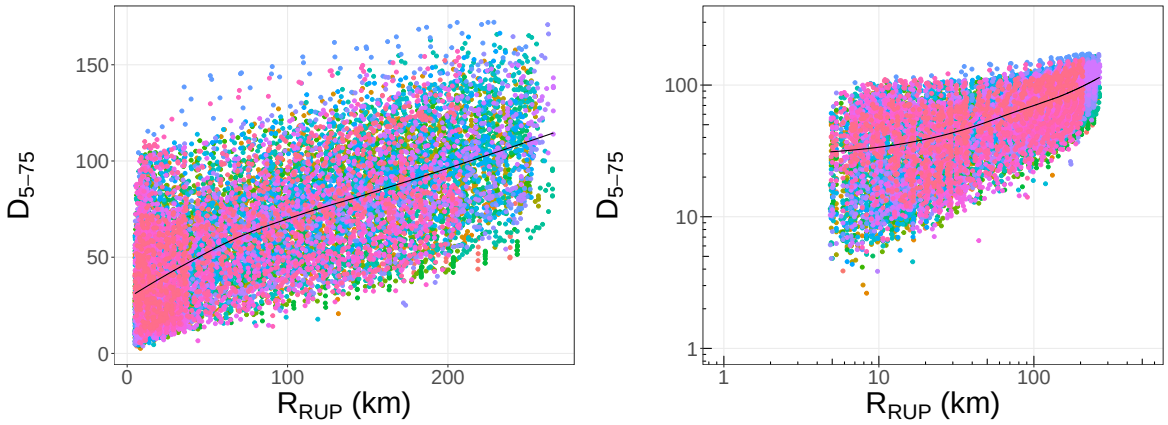

Figure 2: Used D_{5075} duration values of the M9 simulations against rupture distance R_{RUP}, color coded by simulation. The black line is a loess fit (Cleveland, 1979) to the data.

Each simulation is treated as an event, with an event term. The predictor variable is the closest distance to the rupture R_{RUP}, also called rupture distance. As the duration measure, we use the significant duration (Trifunac and Brady, 1975), which which is based on the normalized arias intensity $I(t)$. Here, we use D_{5-75}, which is commonly used (e.g. Bahrampouri et al., 2021; Kempton and Stewart, 2006; Afshari and Stewart, 2016; Jaimes and García-Soto, 2021).

3 Intuition of Duration Model

Generally, there are two approaches to model duration. The first is to model duration as a multiplicative model, similar to PSA GMMs (e.g. Jaimes and García-Soto, 2021; Anbazhagan et al., 2017; Meimandi-Parizi et al., 2020; Bora et al., 2015, 2019). Such an approach has the advantage that the model can be fit in the same way as a response spectrum GMM (using mixed-effects regression to partition into between-event and within-event residuals, corresponding to source and path duration). A multiplicative duration model has the basic form

$$\ln D = f(\vec{r}; \mathbf{M}, R, \ldots) + \delta B + \delta W \tag{1}$$

where $f(\vec{r}; \mathbf{M}, R)$ is a function with coefficients $\vec{r}$, dependent on magnitude, distance, and possibly other parameters such as time-averaged shear wave velocity in the upper thirty meters V_{S30}. δB and δW are the event term and within-event residual, respectively. These are typically (universally) assumed to be normally distributed with standard deviations σ_S and σ_P. Thus, the overall duration is assumed to be lognormally distributed in a multiplicative model.

A different approach to model duration is an additive model, where the total duration is modeled as the sum of a source and path duration, $D_{total} = D_{source} + D_{path}$ (e.g. Afshari and Stewart, 2016; Kempton and Stewart, 2006; Boore and Thompson, 2014; Abrahamson and Silva, 1996; Bahrampouri et al., 2021; Boore, 2003; Atkinson and Boore, 1995). Empirical duration models based on regression analysis using such a model sometimes try to account for between-event terms (ostensibly corresponding to a source term) with the following modeling approach (Afshari and Stewart, 2016; Kempton and Stewart, 2006; Bahrampouri et al., 2021; Abrahamson and Silva, 1996)

$$\ln D = \ln \left[f_{source}(\vec{c}; \mathbf{M}) + f_{path}(\vec{r}; R) \right] + \delta B + \delta W \tag{2}$$

where $f_{source}(\vec{c}; \mathbf{M})$ and $f_{path}(\vec{r}; R)$ are functions for the source and path duration, respectively, and δB and δW are the between-event and within-even residual. The model of Equation 2

has the advantage that it can be fit using a traditional mixed-effects regression (albeit with a nonlinear functional form). However, Equation (2) implies

$$D = f_{source}(\vec{c}; \mathbf{M})\ e^{\delta B}\ e^{\delta W} + f_{path}(\vec{r}; R)\ e^{\delta B}\ e^{\delta W} \tag{3}$$

so the "full" source and path duration are both affected by the event term δB and within-event residual δW. In particular, the event term δB, which is typically interpreted as a term that determines the variability of the source duration, has the effect of changing the path duration (and thus the distance scaling). This means that δB can be understood as a random effect that models an individual distance scaling for each event. Since in a model like Equation (2) the residuals δW and δB affect both source and path durations, we call this model a "mixed model".

One way to separate the source and path duration in an additive model would be to write it as

$$D = f_{source}(\vec{c}; \mathbf{M})\ e^{\delta B} + f_{path}(\vec{r}; R)\ e^{\delta W} \tag{4}$$

in which the two terms are completely decoupled. Assuming δB and δW are normally distributed, this means that both source and path duration are lognormally distributed. The sum of two logarithmic distributions is not a logarithmic distribution, and as a consequence, the total duration D is not lognormally distributed. Such a model was used in Walling et al. (2018). A (rather severe) constraint in Equation (4) is that all terms are positive. It then follows that the source duration, $f_{source}(\vec{c}; \mathbf{M})$, must be smaller than the smallest observed duration, D. This is a severe constraint on the source duration, and ignores possible sources of random variation in observed duration, in particular at shorter distances (see for example Figure 4(b) of Boore and Thompson (2014), which shows negative path durations after correcting total durations using a reasonable, physics based source duration model).

To alleviate the constraint that the source duration must be smaller than the observed duration in an additive model, we propose to partially decouple the two terms, and write the model as

$$D = \left[f_{source}(\vec{c}; \mathbf{M})\ e^{\delta B} + f_{path}(\vec{r}; R)\right]\ e^{\delta W} \tag{5}$$

In this model, the path duration is unaffected by the between-event residual δB (which is thus a proper measure for the variability of the source duration), while the structure of the model allows for the observed duration D to be smaller than the event-specific source duration, which is given by $D_{source} = f_{source}(\vec{c}; \mathbf{M})\ e^{\delta B}$, in the case that δW is negative. In this model, the total duration D is lognormally distributed (assuming δW is normally distributed).

We can write the model of Equation (5) as a hierarchical model

$$D_{source} \sim LN(\ln f_{source}(\vec{c}; \mathbf{M}), \sigma_S) \tag{6}$$

$$\text{med}_{path} = D_{source} + f_{path}(\vec{c}; R) \tag{7}$$

$$D \sim LN(\ln \text{med}_{path}, \sigma_P) \tag{8}$$

One can understand these Equations as follows: the source duration is lognormally distributed, with a median which is a function of magnitude, and standard deviation σ_S. Given a value of the source duration for a specific event, the median path duration is the sum of the source duration and a function that depends on distance. The actual observed duration is then lognormally distributed with the calculated median and standard deviation σ_P.

One should note some differences of the model of Equation (5) to traditional PSA GMMs as well as the mixed model of Equation (2). The model cannot be fit using traditional mixed effects, and the total variance is not the sum of the individual variances (i.e. $\sigma^2_{total} \neq \sigma^2_S + \sigma^2_P$).

3.1 Why Lognormal?

In general, duration models assume that D is lognormally distributed (we are not aware of any model that does not make this assumption). One reason for this assumption is ease of computation, and the advantage that one can use methods that have been well established for PSA GMMs. However, for response spectral GMMs the lognormal distribution makes physical sense, as the spectrum is a multiplication of several terms (Boore, 2003). For duration models, the situation is not as clear, in particular if one wants to use an additive model (for a multiplicative model as in Equation (1) the lognormal distribution makes sense). It is easy to reformulate Equations (6) to (8) using a different distribution, as long as it is defined for a positive random variable. Instead of the lognormal distribution, we also model duration with a Gamma distribution. In the following paragraphs, we briefly describe the Gamma distribution and how it can be used in a duration model context.

If the random variate X is distributed according to a Gamma distribution,

$$X \sim \text{Gamma}(\alpha, \beta) \tag{9}$$

then the corresponding probability density function is

$$pdf(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta x} \tag{10}$$

where $\Gamma(\cdot)$ is the Gamma function, α is called the shape parameter, and β is the rate parameter. The mean of the Gamma distribution is calculated as

$$\mathbb{E}[X] = \mu = \frac{\alpha}{\beta} \tag{11}$$

If we assume that duration is distributed according to a Gamma distribution, then we model the mean as a function of magnitude (for source duration) and distance (for path duration). Hence, we can write

$$\mu_{source} = f_{source}(\vec{c}; \mathbf{M}) \tag{12}$$

We can then rewrite the rate parameter β of the Gamma distribution as $\beta = \frac{\alpha}{\mu}$ (cf. Equation (11)), and thus formulate the hierarchical model with Gamma distributions as

$$D_{source} \sim G(\alpha_S, \frac{\alpha_S}{\mu_{source}}) \tag{13}$$

$$\mu_{path} = D_{source} + f_{path}(\vec{r}; R) \tag{14}$$

$$D \sim G(\alpha_P, \frac{\alpha_P}{\mu_{path}}) \tag{15}$$

analogous to the lognormal model in Equations (6) to (8). It would also be possible to model the source and path duration with a different distribution (e.g. a lognormal distribution for the source and a Gamma distribution for the path duration).

The Gamma model has the same number of parameters to fit as the lognormal model: coefficients for the source and path duration, event-specific source durations, and a distributional parameter (σ for lognormal, α for Gamma) for both the source and path duration. One difference is that we model the mean as a function of magnitude and distance for the Gamma model, while we model the median in a lognormal model.

We do not think that there is an a-priori a preference for either the lognormal or Gamma model for an additive duration GMM. In the literature, duration has been universally modeled with a lognormal distribution (e.g. Kempton and Stewart, 2006; Afshari and Stewart, 2016; Bahrampouri et al., 2021; Jaimes and García-Soto, 2021), but one could argue that this is due to familiarity (based on PSA GMMs), as well as ease of computation. In later sections, we compare duration models based on the Gamma model and the lognormal distribution.

3.2 Source Duration Model

The model of the source duration is often based on the inverse of the corner frequency of the source spectrum (e.g. Abrahamson and Silva, 1996; Kempton and Stewart, 2006), following Brune (1970, 1971)

$$D_{source} = \frac{1}{f_c} \tag{16}$$

Following Abrahamson and Silva (1996), the corner frequency can be written as

$$f_c = 4.9\ 10^6 \beta \left(\frac{\Delta\sigma}{M_0}\right)^{1/3} \tag{17}$$

where M_0 is the seismic moment ($M_0 = 10^{1.5\mathbf{M}+16.05}$), β is the shear-wave velocity at the source (from now on we assume $\beta = 3.2$km/s), and $\Delta\sigma$ is the stress parameter. Thus, this give a model for the source duration as

$$D_{source} = \left(\frac{\Delta\sigma(\mathbf{M})}{10^{1.5\mathbf{M}+16.05}}\right)^{-1/3} / \left(4.9\ 10^6 \beta\right) \tag{18}$$

Magnitude scaling in the source duration is introduced by the moment M_0, as well as a potentially magnitude dependent stress parameter. Assuming an exponential dependence of the stress parameter on magnitude ($\Delta\sigma = \exp(a_1 + a_2\mathbf{M})$), the scaling can be combined (Bahrampouri et al., 2021; Bommer et al., 2009) into a model of the form

$$D_{source} = \exp(c_1 + c_2\mathbf{M}) \tag{19}$$

A constant (magnitude-independent) value of the stress parameter $\Delta\sigma$ implies a value of the slope of $\ln D_{source}$ with magnitude of $c_2 = 1.151$ (Bommer et al., 2009), while a slope of $a_2 = 3.45$ for the magnitude dependence of $\Delta\sigma$ leads to a constant (magnitude-independent) source duration (Afshari and Stewart, 2016).

We want to stress that Equation (18) is merely a crutch to connect source duration to magnitude, and to inform the form of the magnitude scaling for the source duration. While Equation (18) is developed(?) for circular ruptures, its use in empirical duration models does not mean that the events used to estimate the empirical model are assumed to have a circular rupture. Similarly, the stress parameter $\Delta\sigma$ is an empirical parameter that makes Equation (18) work.

It should also be noted that the use of a double-corner frequency model, as used in Atkinson and Silva (2000) to model a finite fault comprised of Brune subfaults, leads to the same model structure. In this case, the source duration is modeled as (Boore and Thompson, 2014)

$$D_{source} = 0.5/f_a \tag{20}$$

where f_a is the lower of two corner frequencies in the double-corner frequency model of (Atkinson and Silva, 2000), which is used to model a finite fault comprised of Brune subfaults. An empirical model for f_a based on Californian events is

$$\ln f_a = 2.181 - 0.496\mathbf{M} \tag{21}$$

which leads to the same magnitude dependence as Equation (19) but with a different slope. In an empirical duration model based on events of different magnitudes the slope would be a parameter to be estimated.

As stated before, in an empirical duration model one wants to account for between-even source duration variability, i.e. the fact that an event can have a longer or shorter source

duration than what the model implies. This can be modeled by including a random effect for each event. The distribution of the random effects allows one to include source duration variability in a forward prediction. There are different possibilities how to include an event-specific random effect in the source duration model based on the corner frequency. One can model an event-specific stress parameter $\Delta\sigma_e$, or an event-specific source duration $D_{source,e}$, where subscript e indicates that the parameter corresponds to event e. If the stress parameter is modeled as a lognormal distribution, then the source duration is also a lognormal distribution, i.e. from

$$\Delta\sigma \sim LN(\mu_{\ln \Delta\sigma}, \sigma_{\ln \Delta\sigma}) \tag{22}$$

it follows that

$$D_{source} \sim LN(\mu_{\ln D_S}, \sigma_{\ln D_S}) \tag{23}$$

where $\mu_{\ln D_S}$ is calculated according to Equation (18) (using $\exp \mu_{\ln \Delta\sigma}$ as the stress parameter), and $\sigma_{\ln D_S} = \frac{\sigma_{\ln \Delta\sigma}}{3}$. This is not the case for other distributions, i.e. a Gamma distributed stress parameter does not imply a source duration that is Gamma distributed or vice versa.

In general, we think that the fact that the two parameterizations of the model (using stress parameter or source duration directly) are equivalent for the lognormal distribution makes this model preferable. It is difficult to test this part of the model, as the source durations are not directly observable.

The M9 simulations (obviously) have a constant magnitude $\mathbf{M} = 9$, so we cannot infer a magnitude scaling. However, we can estimate a model using the stress-parameter parameterization, which induces a scaling of the source duration with magnitude. For ease and stability of computation, we use the following model for the source durations of the M9 simulations

$$\ln D_{source} \sim N(\mu_{\ln D_S}, \sigma_{\ln D_S}) \tag{24}$$

i.e. with a constant median logrithmic source duration $\mu_{\ln D_S}$. One can calculate a median stress parameter from $\mu_{\ln D_S}$ using Equation (18).

3.3 Path Duration Model

The model for the path duration needs to incorporate some sort of distance scaling that ensures that duration increases with distance. This is often modeled as a combination of linear segments (Kempton and Stewart, 2006; Atkinson and Boore, 1995; Boore and Thompson, 2014; Bahrampouri et al., 2021; Abrahamson and Silva, 1996; Afshari and Stewart, 2016) in additive models. Based on Figure 2, we use a bilinear model for the distance scaling. The model is

$$f_{path}(\vec{r}; R_{RUP}) = r_1 R_{RUP} + (r_2 - r_1)\delta \ln\left[1 + \exp\left(\frac{R_{RUP} - R_b}{\delta}\right)\right] \tag{25}$$

which is a bilinear function with a smooth transition at $R = R_b$, where r_1 and r_2 are the slope below and above the breakpoint R_b, respectively, and δ controls the smoothness of the transition. We fix the value of $\delta = 1$, the other parameters are estimated during model fitting. The model of Equation (25) is differentiable with respect to all parameters, which helps during model fitting.

One note on computation: for large distances, the term $\exp\left(\frac{R_{RUP}-R_b}{\delta}\right)$ can be numerically unstable and lead to overflow. Hence, one should use a logSumexp function, as implemented e.g. in the package `matrixStats` (Bengtsson, 2022) for the computer environment `R` (R Core Team, 2021).

3.4 Model Predictions

Making prediction with the partially decoupled model is more complicated than for the mixed or a multiplicative model. In the mixed model, where the event term "sits outside the logarithm" in Equation (2), the median prediction can just be calculated as $med = \ln\left[f_{source}(\vec{c}; \mathbf{M}) + f_{path}(\vec{r}; R)\right]$, and the standard deviation of the corresponding normal distribution (in log-space) is calculated as $\sigma = \sqrt{\sigma_S^2 + \sigma_P^2}$.

By contrast, in the partially decoupled model the (event-specific) source duration is part of the prediction for the full duration, and thus there are different possibilities for calculating predictions of a new event: one could use the mean of the source duration distribution, the median, the mode, or some other point estimate. The mean of the total duration can be calculated as

$$\mu(D) = \int_0^\infty \int_0^\infty x \; p(D_S) p(x|D_p) dx dD_S \tag{26}$$

This is the same value as when using the mean of the source duration in the prediction

$$\mu(D) = \mu(D_S) + f_{path}(\vec{r}; R_{RUP}) \tag{27}$$

where $\mu(D_S) = \exp\left[\mu_{\ln D_S} + 0.5\sigma^2_{\ln D_S}\right]$ for the M9 simulation model (as the source duration model is a lognormal distribution). This is not true for the median, i.e. $\text{med}(D) \neq \text{med}(D_S) + f_{path}(\vec{r}; R_{RUP})$.

If the full distribution is desired, one can use Monte Carlo simulation, i.e. one can sample from the source duration distribution, calculate desired quantities (such as mean or fractiles) of the total duration distribution given the sampled source durations, and then average over the samples. However, this approach is more computationally involved, and results may change from run to run, or depend on the number of samples.

4 Duration Models for M9 Simulations

We estimate different duration models on the durations from the M9 simulations. We then compare the models in terms of their scaling, their predictive distributions, and their predictive performance. The estimated models are

1. A mixed model (cf. Equation (2)).
2. A partially decoupled model, with the source and path duration modeled as a lognormal distribution.
3. A partially decoupled model, with the source duration modeled as a lognormal distribution and the path duration modeled as a Gamma distribution.

The mixed model has the form

$$\ln D = \ln\left[c_1 + f_{path}(\vec{r}; R_{RUP})\right] + \delta B + \delta W \tag{28}$$

where δB and δW are normally distributed with mean zero and standard deviations σ_S and σ_P, respectively, and c_1 is the source duration function, which is modeled as a constant since all vents have the same magnitude.

As seen in the Figure 2, there is more variability in the logarithmic duration values at short distances. Such a behavior cannot be modeled with a constant standard deviation σ_P or shape

parameter α_P. Hence, we model these parameters as distance dependent, with the following functional forms

$$\sigma_P(R_{RUP}) = s_1 + s_2 \frac{1}{1 + \exp(-s_3(R_{RUP} - R_s))} \tag{29}$$

$$\alpha_P(R_{RUP}) = a_1 + a_2 R_{RUP} \tag{30}$$

The parameters s_1, s_2, s_3 and R_s, or a_1 and a_2, respectively, are determined during model fitting, and are constrained such that the overall standard deviaton or shape parameter is positive.

4.1 Model Implementation

The models are fitted via Bayesian inference (e.g. Spiegelhalter and Rice, 2009; Gelman et al., 2013) using the program Stan (Carpenter et al., 2017; Stan Development Team, 2022). Stan implements Markov Chan Mote Carlo (MCMC) sampling (Neal, 1993) in the form of Hamiltonian Monte Carlo (Neal, 2011; Betancourt, 2017b,a). Thus, the results of the model fitting process are samples from the posterior distribution, which can be used to assess the epistemic uncertainty associated with the models.

We run the regressions through the `R`-interface `cmdstanr` (Gabry and Češnovar, 2021). We run four chains for each model, with 200 sampling iterations. Convergence of the chains is evaluated via the $\hat{R}$ (R-hat) statistic (Vehtari et al., 2020). Comparisons between the models is done via the PSIS-LOO information criterion (Vehtari et al., 2017, 2021), which is an approximation to leave-one-out cross-validation. We also assess the posterior predictive distribution of the models, which is a measure of the overall uncertainty/variability of the model, and comprises both epistemic uncertainty of the parameters and aleatry variability due to the lognormal or Gamma distribution. The posterior predictive distribution for an observation $\tilde{y}$ is defined as (Gelman et al., 2013)

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)d\theta \tag{31}$$

so it is calculated by integrating out the uncertainty of the estimated parameters.

The parameters to be estimated are the coefficients for the path duration model (r_1, r_2, R_b), the log median of the source duration ($\mu_{\ln D_S}$) for the partially decoupled model or c_1 for the mixed model, and the shape/scale parameters σ_S, σ_P and α_P. In addition, the individual source durations D_S (for the decoupled models) or the event terms δB for the mixed model are estimated as well. The prior distributions for the source durations D_S is given by Equation (24), while it is a normal distribution with mean zero and standard deviation σ_S for δB. For the

other parameters, the prior distributions are as follows:

$$\begin{aligned} r_1, r_2 &\sim N(0, 0.2) \\ R_B &\sim N(60, 40) \\ c_1 &\sim N(25, 5) \\ \sigma_S &\sim N(0, 1) \\ \mu_{\ln D_S} &\sim N(3.2, 2) \\ \sigma_{\ln D_S} &\sim N(0, 2) \\ a_1 &\sim N(0, 5) \\ a_2 &\sim N(0, 1) \\ s_1, s_2 &\sim N(0, 1) \\ s_3 &\sim N(1, 2) \\ R_s &\sim N(60, 40) \end{aligned}$$

5 Results

Figure 3 shows the posterior density, assessed from the posterior samples, of the parameters associated with the path duration as well as the average source duration. In general, the path duration coefficients are very similar between the three models, with slightly smaller values of r_1 for the Gamma model, and slightly wider uncertainty of r_2 for the mixed model.

The average source duration is captured by parameter c_1 (for the mixed model) and $\exp(\mu_{\ln D_S})$ for the partially decoupled models (Gamma and lognormal). It models the median of the source duration distribution, and is the model prediction for $R_{RUP} = 0$. However, we need to remember that for the Gamma model, the prediction represents the mean of the total duration, while for the lognormal and mixed model it represents the median. Hence, we also calculate the mean of the total duration for the lognormal and mixed model via

$$\mu(D(R_{RUP} = 0)) = \exp\left[\mu_{\ln D_S} + 0.5\sigma_P(R_{RUP} = 0)^2\right] \tag{32}$$

We calculate the mean for each posterior sample, which gives a posterior distribution for this parameter. In Figure 3, the posterior distribution of the mean prediction for $R_{RUP} = 0$ are quite similar between the models.

Figure 4 shows the estimated standard deviations/shape parameters. The mixed model shows a smaller value for σ_S than the other models. The reason is that the event term δB in the mixed model is not just a source term, but models a different distance scaling for each event, and thus σ_S is more a measure of the variability of the distance scaling. This can be seen in Figure 5, which shows predictions for each of the 30 events (including the event-specific D_S or δB). The Gamma and logormal (partially decoupled) models show a similar range of predictions, which stays constant with increasing distance. By contrast, the mixed model predictions show a different distance scaling for each event, and increasing range with distance.

Within-event residuals are shown in Figure 6. The residuals for the Gamma and lognormal models are calculated via

$$resid = \ln D_{obs} - \ln\left[D_S + f_{path}(\vec{r}; R_{RUP})\right] \tag{33}$$

where D_S is the estimated event-specific source duration. For the mixed model the residual is the term δW in Equation (28). Hence, the residuals are calculated in log-scale with respect to the logarithm of the mean prediction for the Gamma model, and with respect to the logarithm of the median prediction of the other two models. There are no large differences between the

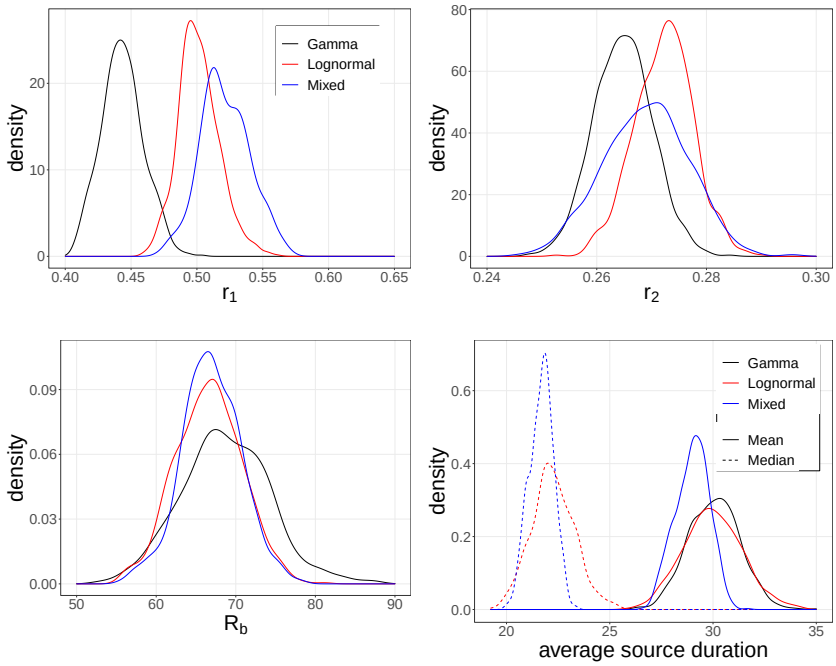

Figure 3: Posterior density of distance slopes r_1, r_2, distance break point R_b, and average source duration, for the three different models. For the averge source duration, solid lines are the mean of the total duration at $R_{RUP} = 0$, while dashed lines are the median.

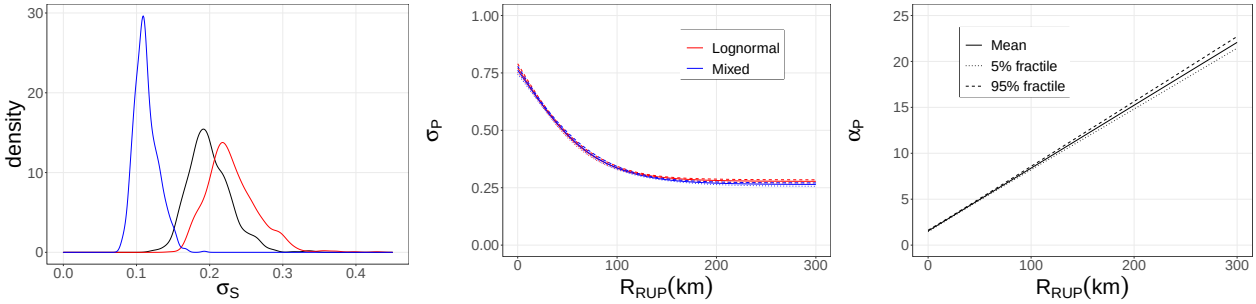

Figure 4: Posterior distribution of the standard Deviation σ_S of the source duration distribution, as well as shape parameter α_P (for the Gamma model) and scale parameter σ_P (for the lognormal and mixed model), dependent on distance. For the shape/scale parameters, the mean and 5%/95% fractiles are shown, calculated from the posterior distributions of the parameters associated with the shape/scale parameters.

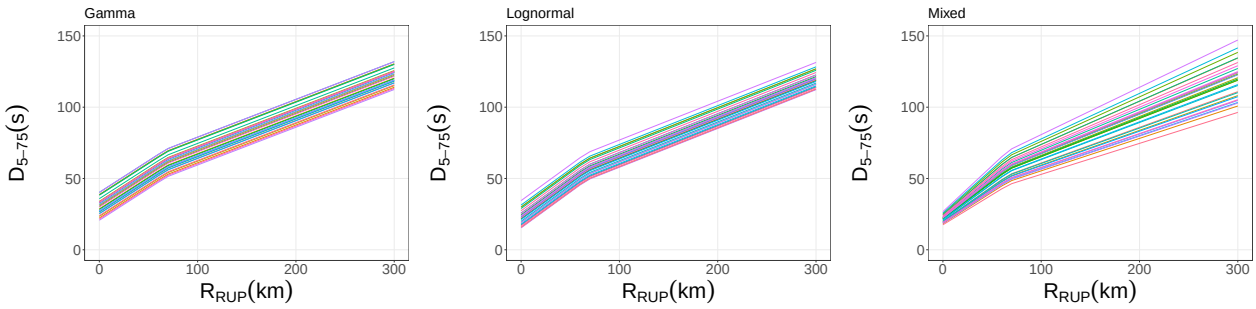

Figure 5: Individual mean (Gamma) and median (lognormal and mixed) prediction of the 30 different simulations for the three different models. Predictions are calculated with the mean of the posterior samples of the parameters.

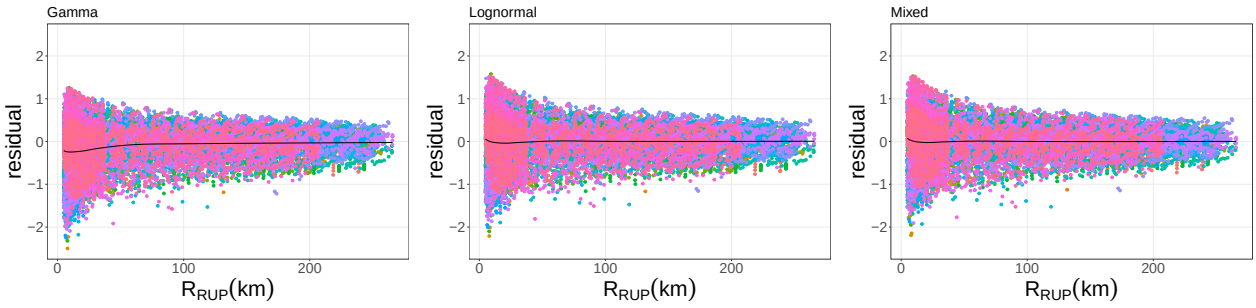

Figure 6: Residuals of the different models. The black line is a loess fit (Cleveland, 1979) to the residuals.

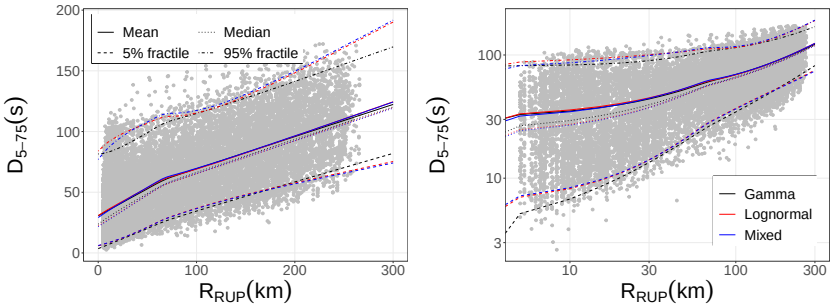

Figure 7: Predictive density and data for the different models.

different models in terms of the residuals. The residuals of all models are well well centered, with larger variability at shorter distances. Such variability is also seen in peak and response spectral values in Frankel et al. (2018), and due to different distances to subevents.

Figure 7 shows the distance scaling of the predictive distribution, together with the data. The predictive distribution comprises both the epistemic uncertainty and the aleatory variability of the model. For each distance, we calculate the mean or median prediction for each posterior sample. Then, we randomly sample 100 source durations or event terms, and sample again 100 overall durations for each source duration/event term. From the resulting samples, we calculate the mean, median, and 5% and 95% fractiles. The mean and median predictions of all models are quite similar, but there are some differences in the fractiles, in particular at very small and very large distances.

Figure 8 shows graphical posterior predictive checks (Gabry et al., 2019). For these plots, we draw samples from the posterior predictive distribution (cf. Equation (31)), and compare to the data. For each sample of the posterior distribution, we randomly draw a new data set (i.e. we calculate the mean/median and shape/standard deviation for each data point and sample from the corresponding distribution). Thus, we have 800 synthetic data sets of 15,810 simulated durations. In Figure 8, we compare the median, 5% and 95% fractiles across those 800 simulated data sets with the median and fractiles of the observed data. Overall, for a good model we would expect that the fractiles of the simulated data sets are similar to the ones data. In general, the differences between the models are small, but the Gamma model seems to be able to capture the 5% and 95% fractiles of the data slightly better. The plots shown in Figure 8 show median and fractiles of the whole data set(s). The models can perform differently for different ranges of the data; we provide plots similar to Figure 8 for different distance bins in the electronic supplement[1].

To quantify model comparison, we can look at the expected log pointwise predictive density

[1]available at https://github.com/nikuehn/Duration_M9

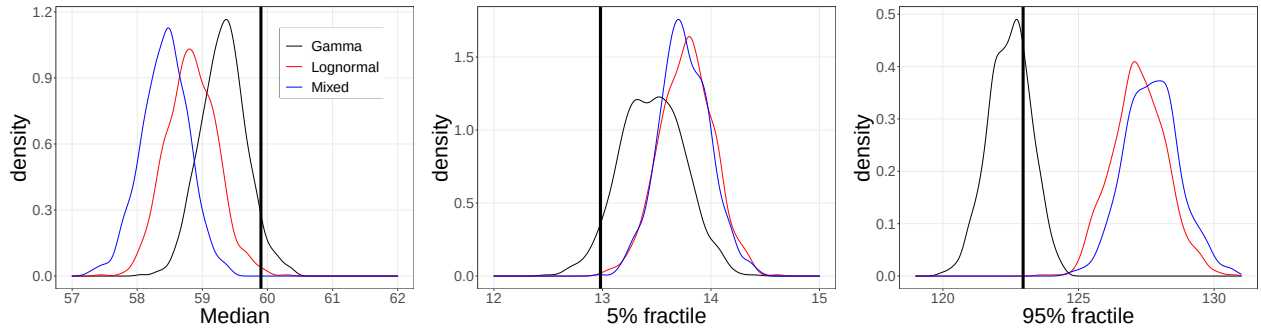

Figure 8: Posterior predictive checks. Shown are densities of the median and 5%/95% fractiles of 800 simulated data sets. The black vertical line displays the median or fractile of the M9 data.

Table 1: Differences in expected log pointwise predictive density $\Delta\widehat{elpd}_{PSIS-LOO}$, based on PSIS-LOO, standard error of the difference, and difference in log-likelihood ΔLL on test data when leaving out one event at a time.

Model	$\Delta\widehat{elpd}_{PSIS-LOO}$	SE($\Delta\widehat{elpd}_{PSIS-LOO}$)	ΔLL	SE(ΔLL)
Gamma	0.0	0.0	0.0	0.0
Mixed	-23.7	24.8	-52.0	9.9
Lognormal	-242.9	20.3	-53.8	9.5

(*elpd*, Vehtari et al., 2017) for a new data set. Table 1 lists the differences in *elpd* for the different models, relative to the best model in terms of *elpd*. The *elpd* values are an approximation to leave-one-out cross-validation. Thus, they are conditioned on the estimated source duration values/event terms. For details, see Vehtari et al. (2017), Vehtari et al. (2021), and the documentation of the `loo` package (Vehtari et al., 2022)[2]. Table 1 shows that the Gamma model has a better predictive performance than the lognormal model, which is in line with the posterior predictive checks shown in Figure 8. The mixed model has similar *elpd* compared to the Gamma model (within the standard error of the *elpd* difference), and performs much better than the lognormal model. Here one needs to remember that *elpd* is a measure of the predictive distribution conditional of the source durations or event terms, which in the case of the mixed model contain differences in the distance scaling for the events (Figure 5). The similar *elpd* values between the Gamma and mixed model, and the much larger value of the mixed model compared to the lognormal model indicates that modeling differences in distance scaling may be advantageous.

To assess model performance on new events (i.e. with unknown source durations/event terms), we also perform a leave-one event-out test. Here, we estimate the models by leaving out all records from one event as test values, and then calculate the log-likelihood of the test data. The differences i the log-likelihood to the Gamma model are shown in Table 1, for five runs (i.e. five test events), calculated as $\Delta LL = LL_{mod} - LL_{Gamma}$. For details on the calculation of the log-likelihood values of the different models, see the electronic supplement[3]. The Gamma model attains the highest log-likelihood value, but now the lognormal and mixed model are comparable, which is consistent with Figure 7.

[2]https://mc-stan.org/loo/index.html
[3]available at https://github.com/nikuehn/Duration_M9

5.1 NGA Sub Data

The models estimated in this paper are estimated for 30 simulations with magnitude $\mathbf{M} = 9$. Assuming a constant (magnitude independent) stress parameter $\Delta\sigma$, we can still make predictions for other magnitude values, via Equation (18). In this section, we calculate residuals to data from the NGA-Subduction (NGA-Sub) project (Bozorgnia et al., 2022; Contreras et al., 2022; Ahdi et al., 2022; Mazzoni et al., 2022, 2021). We use the interface data from the KBCG NGA-Sub model (Kuehn et al., 2020, 2022), and calculate residuals for the Gamma model (residuals to the other models are provided in the electronic supplement[4]). Predictions can be calculated as follows: the median stress parameter can be calculated from the estimated median source duration via

$$\mu_{\ln \Delta\sigma} = -3\left(\mu_{\ln D_S} + \ln 4.9 + 6\ln 10 + \ln \beta\right) + (1.5 * 9 + 16.05)\ln 10 \tag{34}$$

which is just rearranging of Equation (18). For each record, we can then calculate the mean prediction and residuals as

$$D_{source} = \left(\frac{\mu_{\ln \Delta\sigma}}{10^{1.5\mathbf{M}+16.05}}\right)^{-1/3} / \left(4.9\ 10^6 \beta\right) \exp \frac{\sigma^2}{2} \tag{35}$$

$$\mu = D_{source} + f_{path}(\vec{r}; R) \tag{36}$$

$$resid = \ln Y - \ln \mu \tag{37}$$

The residuals are shown in Figure 9 with respect to distance, magnitude and V_{S30}. The residuals are clearly biased (the mean of the residuals is -0.887), so the model based on the M9 simulations is overpredicting the data. The distance scaling of the data at larger distances is similar to the model, but is quite different for short distances. Part of this could be due to the geometry of the subuction zone (the M9 simulations are done for Cascadia/Pacific Northwest, while the NGA-Sub data is mostly comprised of Japanese records). Another possibility is that the distance breakpoint R_B is magnitude dependent, which would lead to a different distance scaling at short distances. Otherwise, the residuals show similar behavior as for the M9 simulations, with larger variability at short distances.

The residuals do not show a great trend with magnitude, which might indicate that a constant stress drop is not an unreasonable model. The increase of the loess fit at large magnitudes is mainly due to the residuals of the Tohoku ($\mathbf{M} = 9.12$) event, which suggests a larger than average source duration for Tohoku (compared to the NGA Sub data). Durations have been found to be negatively correlated with PGA (Bradley, 2011) based on shallow active crustal ground motions. The event term of Tohoku is negative in Kuehn et al. (2020), which would imply a positive event term for duration, so the observation of Figure 9 is in line with expectation.

The residuals show a positive trend at low values of V_{S30}, indicating longer durations at softer sites. This trend is in agreement with previous studies (Bahrampouri et al., 2021; Kempton and Stewart, 2006; Afshari and Stewart, 2016).

6 Discussion

We have presented and compared different parameterizations of additive duration models, based on the M9 simulations. The models differ in the placement of the random effect of the event term (affecting both source and path durations, as done previously in the literature, or only affecting the source duration), as well as in the observation likelihood (lognormal and Gamma).

[4]available at https://github.com/nikuehn/Duration_M9

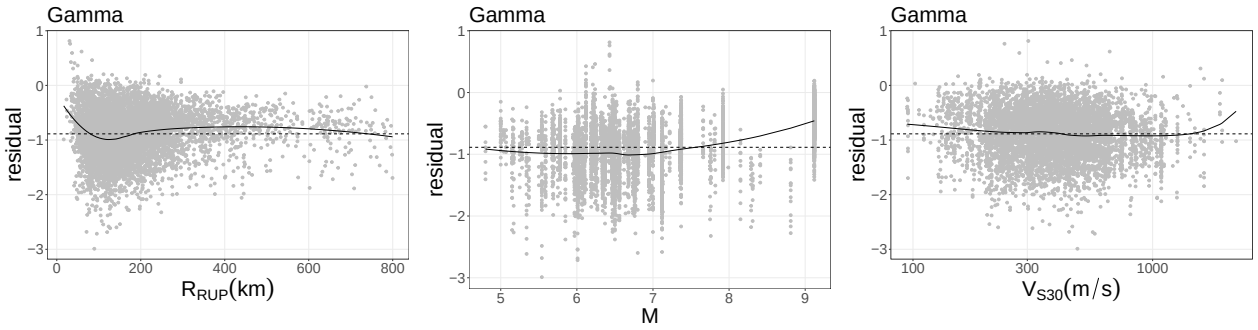

Figure 9: Residuals to Interface NGA-Sub data from Kuehn et al. (2020) with the Gamma model. The solid black lines show a loess fit (Cleveland, 1979) to the residuals, the dashed black lines are the mean of the residuals.

We think that the partially decoupled model is more physical (or its interpretation is more in line with the physical intention). The reason is that the partially decoupled model does not introduce a different path scaling for each event, and thus the variability of the estimated source-specific effects represents source duration variability, and is not affected by variability in distance scaling. We find that the partially decoupled Gamma model performs best of the three models, both in terms of graphical posterior predictive checks (Figure 8) as well as when assessed by PSIS-LOO and leave-one event-out training training/test splits (Table 1). However, the differences between the model predictions are not large (Figure 7), which raises the question whether it matters which model is used. The mean and median predictions of all models are very similar, and differences in the fractiles of the predictive distribution only become pronounced a larger distances. Hence, while the partially decoupled Gamma model is preferable in terms of its physical interpretation and a slightly better predictive performance, this is only of concern of the full distribution of the model prediction is needed. Related to this point is that the predictive distribution is harder to compute for the partially decoupled models; in the mixed model, the total variance can be calculated as $\sigma_T^2 = \sigma_S^2 + \sigma_P^2$, which is not valid for the decoupled models, and one has to resort to numerical or Monte Carlo integration.

We also find that the Gamma model has better predictive performance compared to the lognormal model. Published duration models (both multiplicative and additive) generally model duration with a lognormal distribution. As stated before, we do not think that there is an a-priori reason to prefer one over the other. Both lognormally and Gamma distributed variables are positive, continuous, and right skewed, which makes these distributions reasonable choices for duration data. The two models are both similarly easy (or complex) to fit, with similar results, so we believe more research is necessary to assess which model is better. In particular, the Gamma model should be applied to a real data set of observed duration models, to check whether the conclusions drawn from the M9 simulations also apply there.

There are advantages to using a lognormal model over the Gamma model. A lognormal model allows to easily model correlations of (log) duration residuals with residuals from PSA GMMs (Bradley, 2011). Such correlations can help to constrain duration model regressions (Stafford, 2008). In addition, the development of spatial correlation models for within-event residuals are easier with normally distributed residuals, though such models are not common (one exmple is Huang et al. (2020)).

By contrast to the lognormal and mixed model, which model the median of the duration distribution, the Gamma model is parameterized by a function for the mean of the distribution. This requires some care when comparing models, but it is easy to convert from median to mean for the lognormal model. This raises a more general question: when a point estimate of ground-motion duration is used, which one should be used? Possible choices are the median, the mean, or the mode, which are all different fro the lognormal and Gamma distribution. Bora et al.

(2014) used the median prediction of their duration model to calculate response spectra via random vibration theory (RVT) in conjunction with an empirical Fourier amplitude spectrum (FAS) model, but switched to the mean in Bora et al. (2015) and Bora et al. (2019). Lavrentiadis and Abrahamson (2021) used the median duration (D_{5-85} of (Abrahamson and Silva, 1996) in their RVT calculations. Duration model predictions may be used in stochastic simulations (Boore, 1983, 2003), as for example in Sokolov and Chernov (2001) or (Bommer et al., 2017). Often it is not clear which duration (mean or median is used), but given that most models use a lognormal distribution, it is implicitly the median. Green et al. (2020) sample from the duration distribution to calculate liquefaction hazard. It is probably application dependent whether it makes sense to neglect duration variability (it might not matter, it might be computationally too demanding), but if the variability is taken into account, the different models can potentially lead to different results.

As shown in Table 1, the predictive performance of the mixed model is much better compared to the lognormal model when conditioning on estimated event terms, but similar when testing against a new event. Since the event terms of the mixed model also account for event-specific distance scaling, this suggests the simplistic path model (with a constant distance scaling) can be improved. The 30 events are all simulated at the same stations, so the path effect is less of regional effect (like different scaling due to different paths because events are located in different regions), but is due to local differences in paths, which can arise due to different location, orientation and depth of the subevents used for simulation. Adding an event-specific distance scaling term to the partially decoupled models (modeled as a random effect to coefficient r_2) leads to a large improvement with values of $\Delta\widehat{elpd}_{PSIS-LOO} = +265.8$ for the Gamma model and $\Delta\widehat{elpd}_{PSIS-LOO} = +75.1$ for the lognormal model, compared to the Gamma model without event-specific distance scaling (the best model according to Table 1).

When estimating a model on observed data such as the NGA-Sub data, we would expect differences in distance scaling between events/regions, similar to (partially) nonergodic effects in PSA models (Stafford, 2014). Several of the NGA-Sub models (e.g. Kuehn et al., 2022; Parker et al., 2021; Abrahamson and Gülerce, 2022) include regional adjustment terms to account for differences in anelastic attenuation and linear site scaling. It is easy to include regional similar random effects structures in both the mixed and partially decoupled models. Moving to more fully nonergodic models (Anderson and Brune, 1999), incorporating a cell-specific distance scaling model (Kuehn et al., 2019; Dawood and Rodriguez-Marek, 2013) into both models should pose no conceptual problem. Nonergodic models based on spatially varying coefficient models (Bussas et al., 2017; Landwehr et al., 2016) can also be included, but should be done in the parameterization of Kuehn (2021), where the nonergodic latent effects are not integrated out.

The models presented in this paper do not include site scaling, i.e. scaling whith V_{S30} and/or $Z_{1.0/2.5}$. In the literature of (mixed) additive duration models, site scaling has been included as an additive (Abrahamson and Silva, 1996; Kempton and Stewart, 2006) or multiplicative term (Bahrampouri et al., 2021; Afshari and Stewart, 2016). In general, the partially decoupled models can accommodate both possibilities, and one should assess the implications when developing new models.

We only briefly touched upon within-model epistemic uncertainty associated with the models. The epistemic uncertainty is quantified via the posterior distribution of the parameters, which is shown in Figures 3 and 4. The uncertainty of the coefficients translates into uncertainty of model predictions, which generally should not be ignored. Plots of the epistemic uncertainty associated with the model predictions, similar to Figure 7, are shown in the electronic supplement (Figure S12)[5]. For the presented models, the epistemic uncertainty is rather small,

[5]available at https://github.com/nikuehn/Duration_M9

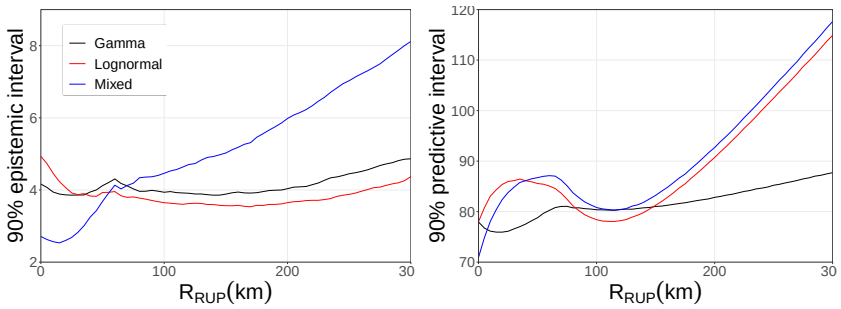

Figure 10: 90% credible intervals for the epistemic model predictions and the full predictive distribution.

which is due to the data distribution. There are 30 events, all of the same magnitude, and all records are simulated at the same stations, so the data coverage is quite extensive compared to an empirical model. The influence of the epistemic uncertainty on model predictions in the M9 duration models is small, but we would expect it to be larger for models based on observed data. Figure 10 shows the 90% credible intervals against distance, calculated as the 95% fractile minus the 5% fractile. Shown are the intervals for the epistemic uncertainty (Figure S12) and the full predictive distribution (Figure 7). At longer distances, the mixed and lognormal models show a larger difference in the predictive distribution compared to the Gamma model. The epistemic uncertainty of the mixed model increases stronger with distance than for the other models. This is due to the slightly wider posterior distribution of r_2 for the mixed model seen in Figure 3. As discussed before, the mixed model formulation leads to event-specific distance scaling, so we can interpret the r_2 coefficient as a mean distance scaling coefficient. It is informed by the distance scaling of the 30 events, and hence is based on less data, which leads to larger uncertainty. This is similar to the larger uncertainty of the global model in a partially nonergodic GMM compared to an ergodic global model (e.g. Kuehn et al., 2022).

When applied to observed data, the M9-models are strongly biased (Figure 9), though the scaling with magnitude as well as the long-distance scaling seem to be in agreement with the data scaling. Hence, we do not advise to use the models directly, but we believe that it can be used to provide informative prior distributions for some parameters of an empirical subduction duration model.

This paper is concerned with the estimation of additive duration models. We have illustrated the different models with the D_{5-75}-significant duration. Some experiments indicate that similar conclusions can be drawn for D_{5-95}, but we have not made a comprehensive analysis like for D_{5-75}. We do not expect major differences for other significant duration measures apart from the values of the parameters. Absolute and bracketed durations (i.e. duration measures based on exceedance of a certain threshold acceleration) decrease with distance (e.g. Bommer et al., 2009; Stafford, 2008), so models for them require more adjustment.

Code Availability

The Stan and `R` codes to run the analyses are available at `https://github.com/nikuehn/Duration_M9`.

Acknowledgements

Regressions are carried out using R-package `cmdstanr` (Gabry and Češnovar, 2021) and assessed with package `posterior` (Bürkner et al., 2022). ELPD values are calculated with package `loo` (Vehtari et al., 2022). Plots are made with `ggplot2` (Wickham, 2016) and `bayesplot` (Gabry and Mahr, 2022). Loess fits are calculated with `msir` (Scrucca, 2011), and the `tidyverse` (Wickham et al., 2019) is used for data wrangling.

We would like to thank Yousef Bozorgnia, Ken Campbell, Brian Chiou, Art Frankel, and the participants of the NGA-Subduction Project for helpful comments and discussions. We would also like to thank Art Frankel and Erin Wirth for providing the M9 durations.

We would like to thank the participants of the NGA-subduction program for numerous discussions that led to improvements of the methodology presented in this paper. The NGA-Subduction project was supported by FM Global, the U.S. Geological Survey, the California Department of Transportation, and the Pacific Gas and Electric Company. The supports are gratefully acknowledged. The opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the study sponsors.

References

Abrahamson, N. A. and Z. Gülerce (2022). Summary of the abrahamson and gulerce nga-sub ground-motion model for subduction earthquakes. *Earthquake Spectra*.

Abrahamson, N. A. A. and W. J. Silva (1996). Empirical ground motion models. Technical report, Report to Brookhaven National Lboratory.

Afshari, K. and J. P. Stewart (2016). Physically parameterized prediction equations for significant duration in active crustal regions. *Earthquake Spectra 32*(4), 2057–2081.

Ahdi, S., D. Y. Kwak, T. D. Ancheta, V. Contreras, T. Kishida, A. O. Kwok, F. Ruz, and S. J. P (2022). Nga-sub site database. *Earthquake Spectra*.

Anbazhagan, P., M. Neaz Sheikh, K. Bajaj, P. J. Mariya Dayana, H. Madhura, and G. R. Reddy (2017). Empirical models for the prediction of ground motion duration for intraplate earthquakes. *Journal of Seismology 21*(4), 1001–1021.

Anderson, J. G. and J. N. Brune (1999). Probabilistic Seismic Hazard Analysis without the Ergodic Assumption. *Seismological Research Letters 70*(1), 19–28.

Atkinson, G. M. and D. M. Boore (1995). Ground motion relations for eastern North America. *Bulletin of the Seismological Society of America 85*(1), 17–30.

Atkinson, G. M. and W. Silva (2000). Stochastic modeling of California ground motions. *Bulletin of the Seismological Society of America 90*(2), 255–274.

Bahrampouri, M., A. Rodriguez-Marek, and R. A. Green (2021, 5). Ground motion prediction equations for significant duration using the KiK-net database. *Earthquake Spectra 37*(2), 903–920.

Bengtsson, H. (2022). *matrixStats: Functions that Apply to Rows and Columns of Matrices (and to Vectors)*. R package version 0.62.0.

Betancourt, M. (2017a). A Conceptual Introduction to Hamiltonian Monte Carlo.

Betancourt, M. (2017b). The Convergence of Markov chain Monte Carlo Methods: From the Metropolis method to Hamiltonian Monte Carlo. pp. 1–12.

Bommer, J. J., B. Dost, B. Edwards, P. P. Kruiver, M. Ntinalexis, A. Rodriguez-Marek, P. J. Stafford, and J. Van Elk (2017). Developing a model for the prediction of ground motions due to earthquakes in the Groningen gas field. *Geologie en Mijnbouw/Netherlands Journal of Geosciences 96*(5), s203–s213.

Bommer, J. J. and A. Martínez-Pereira (1999, 4). The effective duration of earthquake strong motion. *Journal of Earthquake Engineering 3*(2), 127–172.

Bommer, J. J., P. J. Stafford, and J. E. Alarcon (2009, 12). Empirical Equations for the Prediction of the Significant, Bracketed, and Uniform Duration of Earthquake Ground Motion. *Bulletin of the Seismological Society of America 99*(6), 3217–3233.

Boore, D. M. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. *America 73*(6), 1865–1894.

Boore, D. M. (2003). Simulation of Ground Motion Using the Stochastic Method. *Pure and Applied Geophysics 160*(3), 635–676.

Boore, D. M. and E. M. Thompson (2014). Path durations for use in the stochastic-method simulation of ground motions. *Bulletin of the Seismological Society of America 104*(5), 2541–2552.

Bora, S. S., F. Cotton, and F. Scherbaum (2019, 2). NGA-West2 Empirical Fourier and Duration Models to Generate Adjustable Response Spectra. *Earthquake Spectra 35*(1), 61–93.

Bora, S. S., F. Scherbaum, N. Kuehn, and P. Stafford (2014). Fourier spectral- and duration models for the generation of response spectra adjustable to different source-, propagation-, and site conditions. *Bulletin of Earthquake Engineering 12*(1), 467–493.

Bora, S. S., F. Scherbaum, N. Kuehn, P. Stafford, and B. Edwards (2015). Development of a Response Spectral Ground-Motion Prediction Equation (GMPE) for Seismic-Hazard Analysis from Empirical Fourier Spectral and Duration Models. *Bulletin of the Seismological Society of America 105*(4), 2192–2218.

Bozorgnia, Y., N. A. Abrahamson, S. Ahdi, T. Ancheta, G. Atkinson, D. Boore, K. W. Campbell, B. Chiou, V. Contreras, R. Darragh, S. Derakhshan, N. Gregor, Z. Gulerce, I. M. Idriss, C. Ji, R. Kamai, T. Kishida, N. Kuehn, D. Y. Kwak, A. Kwok, P. S. Lin, H. Magistrale, S. Mazzoni, S. Midorikawa, S. Muin, G. Parker, S. Rezaeian, H. Si, W. Silva, J. Stewart, M. Walling, K. Wooddell, and R. Youngs (2022). Nga-subduction research program. *Earthquake Spectra*.

Bradley, B. A. (2011). Correlation of significant duration with amplitude and cumulative intensity measures and its use in ground motion selection. *Journal of Earthquake Engineering 15*(6), 809–832.

Brune, J. N. (1970, 9). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research 75*(26), 4997–5009.

Brune, J. N. (1971, 7). Correction [to "Tectonic stress and the spectra, of seismic shear waves from earthquakes"]. *Journal of Geophysical Research 76*(20), 5002–5002.

Bussas, M., C. Sawade, N. Kühn, T. Scheffer, and N. Landwehr (2017, 10). Varying-coefficient models for geospatial transfer learning. *Machine Learning 106*(9-10), 1419–1440.

Bürkner, P.-C., J. Gabry, M. Kay, and A. Vehtari (2022). posterior: Tools for working with posterior distributions. R package version 1.2.1.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software 76*(1), 1–32.

Chandramohan, R., J. W. Baker, and G. G. Deierlein (2016, 7). Impact of hazard-consistent ground motion duration in structural collapse risk assessment. *Earthquake Engineering & Structural Dynamics 45*(8), 1357–1379.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association 74*(368), 829–836.

Contreras, V., J. P. Stewart, T. Kishida, R. B. Darragh, B. S. J. Chiou, S. Mazzoni, R. R. Youngs, N. M. Kuehn, S. K. Ahdi, R. Boroschek, F. Rojas, and O. J (2022). Nga-sub source and path database. *Earthquake Spectra*.

Dawood, H. M. and A. Rodriguez-Marek (2013). A Method for Including Path Effects in Ground-Motion Prediction Equations: An Example Using the Mw 9.0 Tohoku Earthquake Aftershocks. *Bulletin of the Seismological Society of America 103*(2B), 1360–1372.

Du, W. and G. Wang (2017, 2). Prediction Equations for Ground-Motion Significant Durations Using the NGA-West2 Database. *Bulletin of the Seismological Society of America 107*(1), 319–333.

Frankel, A., E. Wirth, N. Marafi, J. Vidale, and W. Stephenson (2018, 10). Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust Based on 3D Simulations and Stochastic Synthetics, Part 1: Methodology and Overall Results. *Bulletin of the Seismological Society of America 108*(5A), 2347–2369.

Gabry, J. and T. Mahr (2022). bayesplot: Plotting for bayesian models. R package version 1.9.0.

Gabry, J., D. Simpson, A. Vehtari, M. Betancourt, and A. Gelman (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society. Series A: Statistics in Society 182*(2), 389–402.

Gabry, J. and R. Češnovar (2021). *cmdstanr: R Interface to 'CmdStan'*. https://mc-stan.org/cmdstanr, https://discourse.mc-stan.org.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013, 11). *Bayesian Data Analysis*. Number February. Chapman and Hall/CRC.

Green, R. A., J. J. Bommer, P. J. Stafford, B. W. Maurer, P. P. Kruiver, B. Edwards, A. Rodriguez-Marek, G. de Lange, S. J. Oates, T. Storck, P. Omidi, S. J. Bourne, and J. van Elk (2020). Liquefaction Hazard in the Groningen Region of the Netherlands due to Induced Seismicity. *Journal of Geotechnical and Geoenvironmental Engineering 146*(8), 1–15.

Huang, C., K. Tarbali, and C. Galasso (2020). Correlation properties of integral ground-motion intensity measures from Italian strong-motion records. *Earthquake Engineering and Structural Dynamics 49*(15), 1581–1598.

Jaimes, M. A. and A.-D. García-Soto (2021, 2). Ground-Motion Duration Prediction Model from Recorded Mexican Interplate and Intermediate-Depth Intraslab Earthquakes. *Bulletin of the Seismological Society of America 111*(1), 258–273.

Kempton, J. J. and J. P. Stewart (2006). Prediction Equations for Significant Duration of Earthquake Ground Motions Considering Site and Near-Source Effects. *Earthquake Spectra 22*(4), 985–1013.

Kuehn, N. (2021). Comparison of Bayesian Varying Coefficient Models for the Development of Nonergodic Ground-Motion Models. *Engrxiv*, 1–25.

Kuehn, N., Y. Bozorgnia, K. W. Campbell, and N. Gregor (2020). Partially Non-Ergodic Ground-Motion Model for Subduction Regions using the NGA-Subduction Database. Technical Report August, Pacific Earthquake Engineerig Research Center.

Kuehn, N. M., N. A. Abrahamson, and M. A. Walling (2019, 4). Incorporating Nonergodic Path Effects into the NGA-West2 Ground-Motion Prediction Equations. *Bulletin of the Seismological Society of America 109*(2), 575–585.

Kuehn, S. M., Y. Bozorgnia, K. W. Campbell, and N. Gregor (2022). A partially nonergodic ground-motion model for subduction earthquakes using nga-sub database. *submitted to Earthquake Spectra*.

Landwehr, N., N. M. Kuehn, T. Scheffer, and N. Abrahamson (2016, 12). A Nonergodic Ground-Motion Model for California with Spatially Varying Coefficients. *Bulletin of the Seismological Society of America 106*(6), 2574–2583.

Lavrentiadis, G. and N. A. Abrahamson (2021). A Non-ergodic Spectral Acceleration Ground Motion Model for California Developed with Random Vibration Theory. *arXiv*.

Mazzoni, S., T. Kishida, V. Contreras, S. Ahdi, D. Y. Kwak, Y. Bozorgnia, and S. J. P (2021). NGA-Sub Flatfile: R211022. Technical Report Dataset, The B. John Garrick Institute for the Risk Sciences.

Mazzoni, S., T. Kishida, J. P. Stewart, B. Darragh, T. Ancheta, B. S. J. Chiou, W. J. Silva, and B. Y (2022). Nga-sub relational database. *Earthquake Spectra*.

Meimandi-Parizi, A., M. Daryoushi, A. Mahdavian, and H. Saffari (2020, 2). Ground-Motion Models for the Prediction of Significant Duration Using Strong-Motion Data from Iran. *Bulletin of the Seismological Society of America 110*(1), 319–330.

Neal, R. (2011, 5). MCMC Using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo*, pp. 113–162.

Neal, R. M. (1993). Probabilistic Inference Using Markov Chain Monte Carlo Methods. Technical report, Department of Computer Science, Universiy of Toronto.

Parker, G. A., J. P. Stewart, D. M. Boore, G. M. Atkinson, and B. Hassani (2021). Nga-subduction global ground-motion models with regional adjustment factors. *Earthquake Spectra*.

R Core Team (2021). R: A Language and Environment for Statistical Computing.

Scrucca, L. (2011). Model-based sir for dimension reduction. *Computational Statistics & Data Analysis 5*(11), 3010–3026.

Sokolov, V. Y. and Y. K. Chernov (2001). Probabilistic microzonation of urban territories: A case of Tashkent City. *Pure and Applied Geophysics 158*(12), 2295–2311.

Spiegelhalter, D. and K. Rice (2009). Bayesian statistics. *Scholarpedia 4*(8), 5230.

Stafford, P. J. (2008). Conditional prediction of absolute durations. *Bulletin of the Seismological Society of America 98*(3), 1588–1594.

Stafford, P. J. (2014). Crossed and Nested Mixed-Effects Approaches for Enhanced Model Development and Removal of the Ergodic Assumption in Empirical Ground-Motion Models. *Bulletin of the Seismological Society of America 104*(2), 702–719.

Stan Development Team (2022). Stan modeling language users guide and reference manual, version 2.29.

Trifunac, M. D. and A. G. Brady (1975, 3). A study on the duration of strong earthquake ground motion. *Bulletin of the Seismological Society of America 65*(3), 581–626.

Vehtari, A., J. Gabry, M. Magnusson, Y. Yao, P.-C. Bürkner, T. Paananen, and A. Gelman (2022). loo: Efficient leave-one-out cross-validation and waic for bayesian models. R package version 2.5.1.

Vehtari, A., A. Gelman, and J. Gabry (2017, 9). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing 27*(5), 1413–1432.

Vehtari, A., A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner (2020, 7). Rank-Normalization, Folding, and Localization: An Improved $\widehat{R}$ for Assessing Convergence of MCMC. *Bayesian Analysis*, 1–23.

Vehtari, A., D. Simpson, A. Gelman, Y. Yao, and J. Gabry (2021, 7). Pareto Smoothed Importance Sampling.

Walling, M., N. Kuehn, N. Abrahamson, and S. Mazzoni (2018). Regional ground motion duration prediction model for subduction regions. *11th National Conference on Earthquake Engineering 2018, NCEE 2018: Integrating Science, Engineering, and Policy 2*, 1287–1297.

Wang, W., D.-Q. Li, Y. Liu, and W. Du (2021). Influence of ground motion duration on the seismic performance of earth slopes based on numerical analysis. *Soil Dynamics and Earthquake Engineering 143*, 106595.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani (2019). Welcome to the tidyverse. *Journal of Open Source Software 4*(43), 1686.

Yaghmaei-Sabegh, S. and B. Hassani (2020, 1). Investigation of the relation between Vs30 and site characteristics of Iran based on horizontal-to-vertical spectral ratios. *Soil Dynamics and Earthquake Engineering 128*(October 2019), 105899.

Yaghmaei-Sabegh, S., S. Karimzadeh, M. Ebrahimi, V. Ozsarac, and W. Du (2022, 5). A new region-specific empirical model for prediction of ground motion significant duration in Turkey. *Bulletin of Earthquake Engineering* (0123456789).