Combining Dynamic Rupture Simulations with Ground-Motion Data to Characterize Seismic Hazard from $M_w$ 3 to 5.8 Earthquakes in Oklahoma and Kansas

by Samuel A. Bydlon, Kyle B. Withers, and Eric M. Dunham

Abstract  Many seismically active areas suffer from a lack of near-source ground-motion recordings, making ground-motion prediction difficult at distances within $\sim 40$ km from an earthquake. We aim to aid the development of near-source ground-motion prediction equations (GMPEs) by generating synthetic ground-motion data via simulation. Building on previous work using point-source moment tensor sources to simulate small ($M_w$ 3–4) earthquakes for a target region encompassing north central Oklahoma and south central Kansas, we perform dynamic rupture simulations of earthquakes up to $M_w$ 5.8. We introduce complexity into the rupture process by adding stochastically generated heterogeneity to initial stress conditions. Our simulated ground-motion data are added to a catalog of recorded ground-motion data to construct a composite recorded-simulated ground-motion catalog that we use to develop a GMPE for the target region. This procedure can be generalized and used for GMPE development in other regions with near-source ground-motion data scarcity, which could directly benefit critical applications such as the National Seismic Hazard Maps produced by the U.S. Geological Survey (USGS).

Electronic Supplement: Figures of moment rate functions and seismograms at five distances (3, 6, 9, 12, and 15 km) for fault perpendicular velocity, fault perpendicular acceleration, fault parallel velocity, and fault parallel acceleration for nine simulated earthquakes, and tables containing the time of first rupture and cumulative slip at each point on the fault for each simulated earthquake.

Introduction

Understanding potential ground-motion amplitudes in a seismically active area is a critical step toward proper estimation of the region’s seismic hazard. Expected ground-motion intensities are often conveyed through ground-motion prediction equations (GMPEs), in which ground-motion intensity measures (e.g., peak ground velocity [PGV], peak spectral accelerations [PSAs]) are expressed as functions of distance from an earthquake source and parameters such as moment magnitude and style of faulting. GMPEs often suffer from a lack of data, in particular at distances less than 40 km from an earthquake source that can be used to constrain ground-motion behavior at near-source distances. Such distances are of particular importance when assessing potential seismic hazard to engineered structures because ground-motion amplitudes are largest close to the fault.

There are several different methods of constructing GMPEs, including regression of recorded ground-motion data, regression of synthetic ground-motion data produced via ground-motion simulation (Atkinson and Boore, 1995, 2006; Toro et al., 1997; Atkinson and Silva, 2000; Silva et al., 2002; Frankel, 2009), and methods that involve adjusting recorded ground-motion data from one area for use in physically distinct areas of interest, such as the hybrid empirical approach (Campbell, 2003; Pezeshk et al., 2011) and referenced empirical approach (Atkinson, 2008, 2010; Atkinson and Motazedian, 2013; Hassani and Atkinson, 2015). There have also been GMPEs constructed using elements of all of the above-mentioned approaches, such as Yenier and Atkinson (2015). Bydlon et al. (2017) present a methodology for incorporating synthetic ground-motion data produced via deterministic simulations into a catalog of recorded ground-motion data. The method requires gathering data on the material structure, earthquake sources, and ground motions from those sources in a target area designated for GMPE construction (Oklahoma/Kansas). These data are used when simulating a set of earthquakes from that area in a manner that achieves consistency with available recorded ground-motion data. The simulations provide synthetic ground-motion data
with intensity measures similar to those of the recorded data from the target area, but with much denser spatial and azimuthal coverage. After verifying that the synthetic ground-motion data have similar amplitudes and decay rates as the recorded data, the synthetic and recorded data are combined in a composite catalog to construct a GMPE for the target area.

The methodology in Bydlon et al. (2017) focuses on small earthquakes (\(M_w \leq 3-4\)) that are idealized as point-source moment tensors (characterized in terms of the final moment and source duration). For larger magnitude earthquakes, effects rising from the finite spatial extent and detailed rupture history, such as directivity, can significantly affect the character of ground motions in both the near and far fields. Such finite-source effects are incorporated into earthquake simulations through either kinematic or dynamic rupture models. Kinematic models require slip history to be prescribed at each point on a fault, but there is no guarantee that resulting models respect the physical processes governing earthquake rupture propagation. In contrast, dynamic rupture models incorporate a physics-based relationship between the evolution of stress and slip on a fault via a friction law and determine the rupture history along with seismic radiation. In this study, we generate synthetic ground motions by performing simulations of dynamic ruptures and seismic-wave propagation using the 3D finite-difference code WaveQLab3D (Duru and Dunham, 2016).

Many studies employed kinematic models to study ground motions for purposes of ground-motion prediction because they tend to be easier to implement and less computationally intensive than dynamic models. Kinematic models can also be linked to constraints from slip inversions. Synthetic ground-motion generation via kinematic models often involves the convolution of sources represented as a set of subfaults or subevents with assumed source time functions and Green’s functions (Hartzell, 1978; Kanamori, 1979; Hadley and Helmerger, 1980; Irikura, 1983; Imagawa et al., 1984; Joyner and Boore, 1986; Heaton and Hartzell, 1989; Somerville et al., 1991; Hutchings, 1994; Zeng et al., 1994). Tumarkin and Archuleta (1994) extended this work by designing a site-specific ground-motion prediction method that simulated earthquakes as a series of small events with an assumed (Aki–Brune) source time function convolved with all available Green’s functions at a site, thus accounting for many different source-site paths. Building on these ideas, the Southern California Earthquake Center developed a methodology within the Cybershake project that incorporates probabilistic earthquake scenarios and rupture variation into ground-motion modeling based on kinematic sources. This approach was used to simulate ground motions and hazard estimates at over 200 sites in the Los Angeles region (Graves et al., 2011).

More recent studies used kinematic source models obtained via finite-fault source inversions to study the dynamics of rupture processes and the connection of source properties to ground motions (Causse et al., 2010, 2013; Menashe and Mai, 2011; Gabriel et al., 2012; Mena et al., 2012; Goda and Atkinson, 2014; Mai and Thingbaijam, 2014; Taborda and Bielak, 2014; Ji et al., 2015; Graves and Pitarka, 2016; Vyas et al., 2016; Crempien and Archuleta, 2017). A downside to using kinematic rupture models is that they are not necessarily consistent with contemporary understandings of earthquake rupture physics, namely the interplay between friction and elasticity that can lead to complexity in the rupture process at many scales. The introduction of heterogeneity into kinematic source models can introduce such complexity but requires a statistical characterization of the kinematic source parameters, which can be complex, partially correlated, and generally poorly constrained, particularly at the small scales associated with radiation of high-frequency waves. Song et al. (2009) analyzed the spatial coherence of source parameters in kinematic models of the 1999 Iżmit, Turkey, and 1992 Landers, California, earthquakes in an effort to understand how to generate kinematic models for ground-motion prediction that capture the physics of the rupture process. Song and Somerville (2010) expanded on this idea by using 2D spatial data analysis tools common in geostatistics to generate physically realistic kinematic source models. Lee and Song (2017) used nonparametric coregionalization, an approach used in geostatistics, to quantify variability in dynamically simulated earthquake rupture scenarios to improve estimates of variability in earthquake simulations without having to perform numerous dynamic simulations. Studies such as Mena et al. (2012), Schmedes et al. (2012), and Trugman and Dunham (2014) also statistically analyzed the results of dynamic rupture simulations with the goal of generating kinematic source models with statistically similar properties.

Dynamic rupture models can also be used directly for ground-motion prediction. Dynamic rupture models differ from their kinematic cousins in that dynamic models do not take as inputs full source time histories of slip at each point on a fault but instead rely on a set of initial conditions (e.g., stress and friction conditions) that are used to compute slip histories according to specified relationships governing friction and material response. Although dynamic rupture models do away with the reliance on ad hoc kinematic source parameter correlations to achieve realistic rupture complexity, complexity is achieved in a dynamic rupture setting by the introduction of heterogeneity into initial stress and friction parameters. Methods of characterizing initial stress and friction parameter heterogeneity in dynamic rupture models have been extensively studied and continue to be an evolving area of research. Oglesby and Day (2002) computed ground motions from dynamic sources with stochastic prestress patterns to explore the sensitivity to fault strength on rupture dynamics and ground motions. Guatteri et al. (2003) generated ground motions from a dynamic source up to 2 Hz for an \(M_w 7.0\) scenario earthquake by defining fracture energy and stress drop as spatial random fields. Other studies (e.g., Ampuero et al., 2006; Ripperger et al., 2007; Schmedes et al., 2010; Dunham et al., 2011a,b; Gabriel et al., 2013; Shi and Day, 2013; Song and Dalguer, 2013; Bydlon and Dunham, 2015; Imperatori and Mai, 2015; Lozos et al., 2015) focused on understanding the effects of the inputs to dynamic rupture...
models (such as initial shear stress) on the rupture process and ground motions.

Dynamic rupture models have been employed to generate synthetic ground motions that are validated against recorded ground-motion data and GMPEs and have been used to predict ground motions for scenario earthquakes (Olsen et al., 2008, 1997; Duan, 2012; Shi and Day, 2013; Baumann and Dalguer, 2014; Heinecke et al., 2014; Douilly et al., 2015; Andrews and Ma, 2016; Roten et al., 2017; Tanrén et al., 2017; Tsuda et al., 2017). Ripperger et al. (2008) used dynamic ruptures with heterogeneous initial shear stresses to examine near-field ground motions from large, vertical strike-slip earthquakes.

Andrews and Boatwright (1998), building on ideas in Andrews (1980), examined dynamic ruptures with self-similar stress variations. They argued the long-wavelength component of stress that controls the rupture length and magnitude of rupture should come from the same spectrum as the shorter wavelength fluctuations used to introduce complexity in the rupture process. This idea was extended by Andrews and Barall (2011) to a procedure for generating heterogeneous initial stresses for planar fault dynamic rupture simulations. This procedure is designed to intake a desired rupture length and magnitude and output a heterogeneous stress distribution that, when used in a dynamic simulation, results in an earthquake of approximately such size that terminates naturally instead of stopping due to the inclusion of a hard barrier on the fault. Using a corrected version of this procedure that accounts for the alteration of the Fourier spectrum due to a depth conditioning function that determines the depth extent of a desired rupture, Andrews and Ma (2016) simulated multiple realizations of $\sim M_w$ 7.0 on vertical, strike-slip faults and found good agreement between the resulting synthetic ground motions and the GMPEs presented in Boore et al. (2014).

In this study, we develop a GMPE for a target region of Oklahoma/Kansas via the methodology of combining synthetic and recorded ground motions presented in Bydlon et al. (2017), but now using dynamic rupture simulations on a finite fault instead of point moment tensor sources. Rupture complexity is introduced using the stochastic stress generation procedure of Andrews and Barall (2011) and Andrews and Ma (2016). We modify the depth conditioning function in the stress generation procedure because earthquakes in the target region are smaller than the events simulated in the presenting studies and because there is no evidence of earthquakes from the target region breaking the surface. Additionally, we modify the procedure to account for a 1D velocity structure we use in our simulations of Oklahoma/Kansas earthquakes. We simulate earthquakes up to $\sim M_w$ 5.8 because the largest earthquake in this area that has occurred since the beginning of recent injection-related seismicity was the 3 September 2016 Pawnee earthquake (Grandin et al., 2017). By complementing recorded ground-motion data from the target region with realistic, but synthetic, ground-motion data produced via earthquake simulations, usable GMPEs can be derived and constrained for areas that are smaller and thus more appropriate than, for example, the entirety of the central and eastern United States. Although this study demonstrates this methodology for the Oklahoma/Kansas target area, we emphasize that this framework can be ported to other areas of interest. This framework could thus be used to generate GMPEs for use in applications such as the National Seismic Hazard Maps produced by the U.S. Geological Survey (USGS; Petersen et al., 2015). We emphasize, however, that the simulations presented in this article neglect certain aspects of the true physical problem, such as site effects and scattering by small-scale heterogeneities that could affect ground-motion characteristics such as between-event and within-event variability. Instead, our objectives are (1) to demonstrate an approach to earthquake ground-motion simulation that integrates physics-based dynamic rupture modeling with available information on a target area, such as velocity structure, to simulate hypothetical earthquakes having ground motions consistent with available aggregate ground-motion data, and (2) to demonstrate how the synthetic data can be integrated into a ground-motion catalog. After accounting for additional aspects such as those mentioned earlier (site effects and scattering), and doing further validation, GMPE constructed with this methodology could be utilized in real-world applications.

Target Region Ground Motion and Material Structure

This study focuses on an area encompassing central Oklahoma, north central Oklahoma, and south central Kansas, the same as described in Bydlon et al. (2017). We use the ground-motion database for the central and eastern United States collected and processed as per Gupta et al. (2017). These data were collected via Incorporated Research Institutions for Seismology (IRIS) data services (see Data and Resources) using the Standing Order for Data (SOD) interface. Earthquake magnitudes, locations, and depths were updated to those contained in the USGS Advanced National Seismic System (ANSS) composite catalog (see Data and Resources). The instruments that collected these measurements had a range of sampling rates, the lowest being 40 Hz. Therefore, the data were filtered using a fourth-order causal Butterworth filter with low- and high-pass frequencies of 0.3 and 20 Hz, respectively. This filtering could lead to an underestimation of peak ground acceleration, but for the frequencies of interest in this study (up to $\sim$5 Hz) spectral values of the filtered and unfiltered data do not differ significantly (Gupta et al., 2017). We extract the subset of ground-motion records associated with earthquakes with epicenters between 35° and 38° N latitude, 96° and 99.5° W longitude, and $M_w$ 4 and greater. The majority of records are for earthquakes occurring during 2009 and later. This catalog includes earthquakes up through the end of calendar year 2016. The catalog consists of ground-motion intensity measures from 2163 ground-motion records associated with 77 unique earthquakes. Figure 1 shows a map of the earthquake epicenters in the target region catalog. Figure 2 shows PGVs and PSAs at 0.2 s (PSA
\[ \sigma_{22}(x_2) = g \int_0^{x_2} (\rho_r(z) - \rho_w) dz, \]

in which \( g \) is acceleration due to gravity, \( \rho_r \) is the density of the overburden rock (for density profile, see Table 1), and \( \rho_w \) is the density of water (assumed to be constant at 1000 \( \text{kg/m}^3 \)). We assume \( \sigma_{11} = \sigma_{22} = \sigma_{33} \) and that the only nonzero component of initial shear stress is \( \sigma_{13} \); therefore, the earthquakes we simulate have \( \sigma_{12} = 0 \) if the fault is vertical.

The procedure to generate the heterogeneous stress field starts by assigning a complex value at every point on a square array of \( lx \) by \( lx \). Both the real and imaginary parts of each point are random with a Gaussian distribution with variance 1/2. The array is Fourier transformed, and the Fourier amplitudes of the constant mode \((k_1 = 0, k_2 = 0)\) is assigned zero amplitude. \( k_1 \) are wavenumbers in the along-strike direction and \( k_2 \) are wavenumbers in the down-dip direction. The second lowest modes are then modified to confine the rupture in the along-strike and down-dip direction. As noted in Andrews and Ma (2016), there was an error in the
Ellsworth (2017) indicates that virtually all earthquakes in Arkansas have ruptured to the surface and recent work by Schoenball and Ellsworth (2017) indicates that virtually all earthquakes in Oklahoma/Kansas have been considered. In our Oklahoma/Kansas simulations, we account for this by modifying the depth conditioning function such that the aspect ratio of the rupture area is 1:1. Our reason for this modification is twofold. First, we need to account for the Fourier amplitude normalization error such that the aspect ratio of the rupture area is 1:1. Our reason for this modification is twofold. First, we need to account for the down-dip and along-strike directions. Our depth function is of the form

\[ D_{\text{mod}}(z) = \begin{cases} 
\frac{1}{1 + (z/z_0)^4} - \cos(\pi z/z_T)^2, & \text{if } z \leq z_T \\
\frac{1}{1 + (z/z_0)^4}, & \text{if } z > z_T 
\end{cases} \]

in which \( z \) is the depth, \( z_0 \) is a depth term that helps determine the down-dip extent of rupture propagation, and \( z_T \) is a depth term that helps determine the upper extent of rupture. As noted in Andrews and Barall (2011), \( z_0 \) and \( z_T \) do not precisely determine the rupture extent, but instead shape the function that dictates the depths below which the initial shear stress is significantly less than dynamic friction stress. The rupture extent more closely resembles the values of the half-length of the lowest mode of the distribution. In this model, such depth extents are less (i.e., shallower) than \( z_0 \) and \( z_T \). We choose \( z_T \) to be 2.7 km. \( z_0 \) is chosen to achieve approximate desired rupture areas and magnitudes; however, exact rupture areas and magnitudes are not determined a priori. Although the parameters described in this section can help provide a rupture with the desired characteristics, some trial-and-error is required to obtain the specific rupture area and magnitude.

Figure 3 shows an example of a heterogeneous initial shear stress and stress-drop distribution that is used to simulate an earthquake in this study. We use a grid spacing of 25 m, with 512 points in the along-strike and down-dip directions; thus the half-length of the lowest mode is 6.4 km.
We choose the following parameters such that the above-mentioned procedure generates the distribution shown in Figure 3:

\[ \mu_d = 0.55, \quad \alpha = 0.05, \quad \mu_{sref} = 0.665, \quad z_0 = 12.5 \text{ km}, \quad (3) \]

in which \( \mu_d \) is the dynamic coefficient of friction, \( \alpha \) is a factor that scales the ratio of initial shear stress to normal stress, and \( \mu_{sref} \) is a reference static friction value that serves as the minimum static coefficient of friction in the output distribution. The parameters \( \alpha, \mu_d, \) and \( z_0 \) define the initial shear-to-normal stress ratio (i.e., \( \tau/\sigma_n \)) via the relationship

\[ \tau/\sigma_n = [\mu_d + \alpha w(x_2, x_3)]D(z), \quad (4) \]

in which \( w(x_2, x_3) \) is a random function constructed per the method described earlier in this section. To simulate earthquakes of different magnitudes, we vary \( \alpha \) and \( z_0 \); however, \( \alpha \) is ideally kept constant and is only adjusted to precisely tune simulated earthquake magnitude. Figure 4 shows profiles of yield stress, dynamic friction stress, and initial shear stress as a function of (a) depth at along-strike distance of 0 km and (b) along-strike distance at depth of 5.4 km. Both profiles are through the hypocenter of a simulated \( M_w 5.4 \) earthquake. The color version of this figure is available only in the electronic edition.
We simulate earthquakes on dipping faults that attempt to mimic estimated dips of earthquakes in the target region. Moment tensors from the USGS archives for $M_w > 4$ earthquakes in the target region indicate that dips for most events are between 75° and 90°. Initial stress conditions are resolved onto the fault, but because the only nonzero component of initial shear stress is $\sigma_{13}$, increasing the dip angle increases the ratio of shear to normal stress on the fault, which leads to larger earthquakes assuming all other aspects of the simulation remain unchanged.

Nucleation and Friction Law Specifications

Simulated spontaneous ruptures follow a slip-weakening friction law but are initiated by forcing weakening using a time-dependent friction coefficient in the hypocentral region. The shear stress $\tau$ at any point on the fault during sliding is defined as:

$$\tau = \mu \sigma_{\text{eff}},$$

in which $\mu$ is the coefficient of friction and $\sigma_{\text{eff}}$ is the effective normal stress acting on the fault. The coefficient of friction at time $t$ (in which $t = 0$ at the beginning of a simulation) is

$$\mu = \mu_s + (\mu_d - \mu_s) \max(f_1, f_2),$$

in which $D$ is the path integrated distance that a node has slipped, $d_0$ is the slip-weakening distance (we choose $d_0 = 5$ cm), and $T$ is the time of forced rupture, as explained below. If $f_2 > f_1$, the rupture is being forced via the time-weakening law. If $f_2 < f_1$, the rupture is propagating spontaneously according to the slip-weakening friction law. Our ruptures transition to spontaneous rupture soon after being nucleated.

We initiate earthquakes by forcing the fault to rupture in a circular area around the hypocenter. The process is designed such that the rupture transitions smoothly from forced to spontaneous rupture to reduce oscillations in rupture velocity. This initiation procedure has been demonstrated in several of the benchmark exercises designed and performed by the Southern California Earthquake Center’s dynamic rupture code validation project, for instance in the benchmarks TPV26 and TPV27 (Harris et al., 2018). The rupture is forced by artificially reducing the friction coefficient beginning at time $T$. The coefficient of friction is reduced from its static to its dynamic value over a specified time, $t_0 = 0.1$ s. The time of the forced rupture at any point on the fault ($T$ in seconds) is defined as:

$$T = \begin{cases} \frac{r}{0.95V_S} + \frac{0.081r_{\text{crit}}}{\mu} \left(\frac{1}{1-t/T_{\text{crit}}}-1\right), & \text{if } r < r_{\text{crit}} \\ \frac{10^3}{\mu}, & \text{if } r \geq r_{\text{crit}} \end{cases},$$

in which $r$ is the distance from the hypocenter. We choose $r_{\text{crit}}$ in our simulations to be 900 m, but our simulated ruptures transitions from forced to spontaneous quickly after initiation, well within the 900 m radius. Near the hypocenter, the rupture is forced to propagate at a speed of $0.95V_S$, in which $V_S$ is the local shear-wavespeed.

Simulated Earthquakes Included in Composite Catalog

Once a heterogeneous initial stress distribution has been generated via the process described in the Generating Initial Stresses section, we identify regions of positive stress drop and choose a hypocentral location where such a region is within the estimated depth range of earthquakes in the target region, namely 2–6 km (Schoenball and Ellsworth, 2017). Hypocenter locations are selected by hand as regions of high-initial shear stress in the range of hypocentral depths of real earthquakes in the catalog. Dips are chosen between 80° and 90°, consistent with the USGS moment tensor solutions in the area. For each different earthquake simulated, we choose a different random seed such that each realization of the initial stress conditions is unique.

As an example, we simulate an earthquake on a fault plane dipping 80°, the same earthquake that was produced via the initial conditions shown in Figures 3 and 4. The parameters used to generate the initial stress distribution are as follows:

$$\mu_d = 0.55, \quad \alpha = 0.05,$$

$$\mu_{\text{ref}} = 0.665, \quad z_0 = 12.5 \text{ km}.$$ (10)

The hypocenter is located at 5.4 km depth. Figure 5 shows rupture contours spaced every 0.1 s and final slip. The moment of the earthquake is $1.55 \times 10^{17}$ N·m. The total area with nonzero slip is 24.1 km² and the average slip is 31.5 cm.

The descriptions of the full suite of $M_w > 4.0$ simulated events can be found in Table 2. Figures 5–14 illustrate the resulting ruptures from the dynamic earthquake simulations included in our composite ground-motion catalog. In all simulations, $\mu_d = 0.55$ and $\mu_{\text{ref}} = 0.665$.

We use the 3D finite-difference code WaveQLab3D (Duru and Dunham, 2016) to simultaneously determine the rupture history and synthetic seismograms. WaveQLab3D is a high-order accurate finite-difference code that utilizes coordinate transformed structured meshes to handle dipping faults, heterogeneous frictional properties and velocity.
Figure 5. (a) Cumulative slip and (b) rupture contours spaced every 0.1 s for an \( M_w \) 5.40 earthquake simulation with initial stress conditions defined by the parameters found in Table 2. The hypocenter is located at 5.4 km depth and 0 km along-strike distance. The color version of this figure is available only in the electronic edition.

Figure 6. (a) Cumulative slip and (b) rupture contours spaced every 0.1 s for an \( M_w \) 4.23 earthquake simulation with initial stress conditions defined by the parameters found in Table 2. The hypocenter is located at 4.1 km depth and 0 km along-strike distance. The color version of this figure is available only in the electronic edition.

Table 2

<table>
<thead>
<tr>
<th>( M_w )</th>
<th>Dip (°)</th>
<th>( \alpha )</th>
<th>( z_0 ) (km)</th>
<th>( D_{hypo} ) (km)</th>
<th>Moment (N \cdot m)</th>
<th>Slip Area (km²)</th>
<th>Average Slip (cm)</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.23</td>
<td>87</td>
<td>0.045</td>
<td>11</td>
<td>4.1</td>
<td>( 2.77 \times 10^{15} )</td>
<td>2.5</td>
<td>3.5</td>
<td>6</td>
</tr>
<tr>
<td>4.38</td>
<td>87</td>
<td>0.045</td>
<td>11</td>
<td>4.1</td>
<td>( 4.68 \times 10^{15} )</td>
<td>2.9</td>
<td>5.1</td>
<td>7</td>
</tr>
<tr>
<td>4.64</td>
<td>88</td>
<td>0.05</td>
<td>12</td>
<td>5.4</td>
<td>( 1.16 \times 10^{16} )</td>
<td>3.4</td>
<td>10.9</td>
<td>8</td>
</tr>
<tr>
<td>4.73</td>
<td>86</td>
<td>0.05</td>
<td>12</td>
<td>5.4</td>
<td>( 1.59 \times 10^{16} )</td>
<td>4.4</td>
<td>11.5</td>
<td>9</td>
</tr>
<tr>
<td>5.10</td>
<td>84</td>
<td>0.045</td>
<td>12.25</td>
<td>4.8</td>
<td>( 5.64 \times 10^{16} )</td>
<td>11.1</td>
<td>16.0</td>
<td>10</td>
</tr>
<tr>
<td>5.33</td>
<td>84</td>
<td>0.045</td>
<td>12.25</td>
<td>4.8</td>
<td>( 1.25 \times 10^{17} )</td>
<td>21.0</td>
<td>18.8</td>
<td>11</td>
</tr>
<tr>
<td>5.36</td>
<td>87</td>
<td>0.05</td>
<td>12.5</td>
<td>5.4</td>
<td>( 1.36 \times 10^{17} )</td>
<td>22.3</td>
<td>19.2</td>
<td>12</td>
</tr>
<tr>
<td>5.40</td>
<td>85</td>
<td>0.05</td>
<td>12.5</td>
<td>5.4</td>
<td>( 1.55 \times 10^{17} )</td>
<td>24.1</td>
<td>20.3</td>
<td>5</td>
</tr>
<tr>
<td>5.76</td>
<td>84</td>
<td>0.045</td>
<td>20.5</td>
<td>5.6</td>
<td>( 5.44 \times 10^{17} )</td>
<td>54.5</td>
<td>31.5</td>
<td>13</td>
</tr>
<tr>
<td>5.79</td>
<td>82</td>
<td>0.045</td>
<td>20.5</td>
<td>5.6</td>
<td>( 6.00 \times 10^{17} )</td>
<td>61.3</td>
<td>30.9</td>
<td>14</td>
</tr>
</tbody>
</table>

Parameters for the simulated events \( M_w > 4 \), including magnitude, dip angle \( \alpha \), down-dip depth cutoff \( (z_0) \), hypocentral depth \( (D_{hypo}) \), total moment, total area with nonzero slip, average slip, and the corresponding figure number that shows the final slip and rupture contours of the simulated event.
structures, and anelastic attenuation. In these simulations, stations are distributed in a grid pattern along the surface spaced every 1 km. We discard all stations with a hypocentral distance greater than 20 km such that we evenly sample the radiation pattern. The maximum domain size is constrained by computational limitations. Because the lowest $S$ wave-speed on our simulations is $\sim 1.25$ km/s, we resolve frequencies up to a maximum of $\sim 10$ Hz, assuming five points per wavelength are required for accurate resolution. Figures 15 and 16 show synthetic seismograms at various distances for azimuths 45° and 90° from strike for the $M_w$ 5.4 earthquake described above. Figure 17 shows the moment rate function for the $M_w$ 5.4 simulated earthquake. Similar figures for all simulated earthquakes can be found in the electronic supplement to this article.

Composite GMPE for Oklahoma/Kansas

We combine all synthetic ground-motion data from the simulations described in the Simulated Earthquakes Included in Composite Catalog section, simulated ground-motion data described in Bydlon et al. (2017) for earthquakes $M_w$ 3–4, and recorded ground-motion data for earthquakes $M_w \geq 3$ in the target region into a single composite ground-motion catalog. We convert synthetic ground-motion data to RotD50 ground-motion intensity metrics using the two horizontal components, an azimuthal rotation increment of 10°, and 5% damping for spectral accelerations. Using this composite ground-motion catalog, we develop a GMPE for the Oklahoma/Kansas region for earthquakes greater than $M_w$ 3 and up to the largest recorded event in the catalog ($M_w$ 5.8). We fit the data to produce a GMPE for the target area using a modified version of the functional form presented in Shahjouei and Pezeshk (2016). The modified functional form is similar to that used in Bydlon et al. (2017). With the addition of ground-motion data for earthquakes $M_w > 4$, we are no longer able to discern a clear change in the rate of attenuation at distances greater than 100 km; so we remove the terms that correspond to a change in the rate of
attenuation at such distances and keep only the terms that correspond to a change of rate of attenuation at 40 km. We use nonlinear regression to fit a GMPE to our composite catalog to predict median ground-motion intensities using the functional form

\[
\log(\bar{Y}) = c_1 + c_2 M + c_3 M^2 + (c_4 + c_5 M) \\
\times \min\{\log(R), \log(40)\} + (c_6 + c_7 M) \\
\times \max\{\log(R/40), 0\} + c_8 R, \tag{11}
\]

with

\[
R = \sqrt{R_{hyp}^2 + c_9^2}, \tag{12}
\]

in which \(\bar{Y}\) is the median value of the specified ground-motion intensity measure (in units of cm/s\(^2\) for PSA, cm/s for PGV), \(M\) is the moment magnitude, \(R_{hyp}\) is the hypocentral distance in kilometers, and \(c_1-c_9\) are the coefficients obtained when fitting the ground-motion data using equation (11). PSA is computed with a 5% damping parameter. The coefficients computed for PGV and PSA at \(T = 1, 0.5, \) and \(0.2\) s (1, 2, and 5 Hz, respectively) are reported in Table 3. Figure 18 plots composite catalog data for earthquakes \(M_w > 4\) and shows GMPEs \((M_w 4.5\) and \(5.5)\) for PGV and PSA \((T = 0.2 \) s) constructed by fitting the composite catalog with equation (11). Figure 19 plots composite catalog data for earthquakes \(M_w > 3\) and shows GMPEs \((M_w 3.5, 4.5,\) and \(5.5)\) for PGV and PSA \((T = 0.2 \) s).

We chose to use hypocentral distance as the distance metric used in this study over alternative options such as Joyner–Boore distance \((R_{JB})\). Although metrics such as \(R_{JB}\) are commonly used in the context of finite-fault simulations, we use hypocentral distance to be consistent with Atkinson (2015). The GMPE presented in Atkinson (2015) is currently the most heavily weighted GMPE used during the generation procedure of the USGS’s one-year seismic hazard forecast for the central and eastern United States from induced and

**Figure 9.** (a) Cumulative slip and (b) rupture contours spaced every 0.1 s for an \(M_w 4.73\) earthquake simulation with initial stress conditions defined by the parameters found in Table 2. The hypocenter is located at 5.4 km depth and 0 km along-strike distance. The color version of this figure is available only in the electronic edition.

**Figure 10.** (a) Cumulative slip and (b) rupture contours spaced every 0.1 s for an \(M_w 5.10\) earthquake simulation with initial stress conditions defined by the parameters found in Table 2. The hypocenter is located at 4.8 km depth and 0 km along-strike distance. The color version of this figure is available only in the electronic edition.
natural earthquakes in zones experiencing induced seismicity (Petersen et al., 2017), a characteristic of our target region.

Figure 20 shows the residuals, defined as the difference (in log units) between the observed and predicted ground-motion intensities, for PGV and PSA ($T = 0.136$ s) with means and standard deviations of all ground-motion data ($M_w > 3$) binned every 10 km. Figure 21 shows residuals for ground-motion data $M_w > 4$. We compute means and standard deviations for the combined dataset, the simulated data only, and the recorded data only. The means for all of those permutations are near zero or well within 1 standard deviation and there are no trends in residuals over the magnitude range $M_w 3–5.8$, indicating that our simulated data have amplitude and decay characteristics similar to the recorded data. Therefore, we conclude that our simulations are producing realistic ground motions and that our GMPE performs well within the above-mentioned magnitude range.

There are notable differences between the GMPE developed in this study and other GMPEs used to estimate expected ground motions in areas of the central United States experiencing induced seismicity, such as the GMPEs presented in Atkinson (2015) and Gupta et al. (2017). The GMPE presented in Shahjouei and Pezeshk (2016) was developed for earthquakes $M_w 5–8$, limiting comparisons to the GMPE presented here. Figures 18 and 19 qualitatively demonstrate differences between the GMPE presented here and that of Atkinson (2015). Low-frequency metrics, such as PGV, show good agreement between two GMPEs at distances <40 km. At farther distances, the Atkinson (2015) GMPE predicts lower ground motions than is indicated by the composite catalog presented in this study, which at distances <40 km is composed of exclusively recorded ground-motion data from the target area. This behavior was also demonstrated by Gupta et al. (2017) by comparing recorded ground-motion data with the Atkinson (2015) GMPE. At these distances, similar behavior is observed when comparing higher frequency metrics such as PSA ($T = 0.2$ s). At distances < 40 km, however, there are
significant differences between the GMPE presented in this study and the Atkinson (2015) GMPE and such higher frequency metrics. We find that the GMPE derived from our composite ground-motion catalog predicts higher ground motions than the Atkinson (2015) GMPE, a difference that could have a significant effect on hazard calculations.

Comparing Finite Fault and Point-Source Simulations

In an effort to understand the limitations of using point moment tensors as sources in earthquake simulations, we perform an experiment comparing the resulting ground motions from a finite-fault simulation to a simulation of an equivalent point moment tensor source with a similar moment rate function. We focus on the M$_w$ 4.38 finite-fault simulation described in the Simulated Earthquakes Included in Composite Catalog section and depicted in Figure 7. We perform a simulation using a Brune source time function, a commonly used function and the functional form of source time functions used to perform simulations of small magnitude earthquakes (M$_w$ 3–4) in Bydlon et al. (2017), which are included in our composite ground-motion catalog presented in this study. In these simulations, stations are distributed in a grid pattern along the surface spaced every 1 km.

The Brune (1970) moment rate function is

$$\dot{M}_0(t) = M_0\omega_b^2(t - t_0)e^{-\omega_b(t - t_0)}H(t - t_0),$$

in which $t_0$ is an offset time and $\omega_b = 2\pi f_c$ for corner frequency $f_c$.

Figure 22 shows the moment rate functions and their spectra for the finite-fault simulation and Brune moment rate functions chosen to approximate the form of the finite-fault simulation. For the Brune point-source simulation, $\omega_b = 9.8$ s$^{-1}$ and $t_0 = 0.215$ s. Moment tensors for all simulated earthquakes can be found in the electronic supplement.

The moment tensor simulations are intended to mimic an earthquake on a fault dipping 87°. The moment tensor components are as follows for both point-source simulations:

- **Figure 13.** (a) Cumulative slip and (b) rupture contours spaced every 0.1 s for an M$_w$ 5.76 earthquake simulation with initial stress conditions defined by the parameters found in Table 2. The hypocenter is located at 5.6 km depth and 0 km along-strike distance. The color version of this figure is available only in the electronic edition.

- **Figure 14.** (a) Cumulative slip and (b) rupture contours spaced every 0.1 s for an M$_w$ 5.79 earthquake simulation with initial stress conditions defined by the parameters found in Table 2. The hypocenter is located at 5.6 km depth and 0 km along-strike distance. The color version of this figure is available only in the electronic edition.
in which $x_1$ and $x_3$ are the horizontal directions and $x_2$ is the vertical direction.

Ground-motion intensity measures from the two simulations are shown in Figure 23. The Brune source time function produces data somewhat similar to the finite-fault simulation, albeit with greater amplitudes and variability, for both PGV and PSA ($T = 0.2$ s). We speculate that the point-source simulation features greater variability due to a more pronounced radiation pattern close to the source than for the finite-fault simulation. Future work could examine these differences in more detail and quantify finite-source effects as a function of magnitude or source dimension. Additionally, we speculate that as the magnitude of simulated events increases, the differences between point-source and finite-fault models will become more pronounced as rupture propagation effects lead to amplitude saturation at near distances. Future work could focus on performing these comparative experiments with increasing magnitude to determine if such divergence is observed and, if so, the approximate magnitude where such divergence occurs.

Conclusions

In this study, we built on the framework presented in Bydlon et al. (2017) for integrating synthetic ground motions produced via ground-motion simulations into recorded ground-motion catalogs for use in regionally specific GMPE development. To simulate earthquakes larger than $M_w 4$, we use finite-fault simulations with stochastically generated initial stresses. By modifying the stress generation procedure of Andrews and Barall (2011) to mimic earthquakes in the Oklahoma/Kansas, we produce ground motions that exhibit similar trends in amplitudes and decay rates with distance as the recorded ground-motion data from the target region. The approach we use to structure these simulations could be used in the future to simulate ground-motion data in other areas where recorded data are limited, but information about the behavior of ground motions in the area is desired. In this study, we demonstrate how this framework can be used to develop a GMPE for a specific target region, Oklahoma/Kansas, but we emphasize that the procedure...
can be ported to other locations to aid regionally specific GMPE development using physics-based simulations of earthquakes. Because GMPEs are a fundamental element of seismic hazard assessment, it is vital to have accurate and well-constrained information on ground-motion behavior. The simulations presented in this article are a step forward toward producing realistic ground motions usable in real-world applications, but there are still specific elements of the physical problem, such as site effects and scattering from small-scale heterogeneities, that could be included to produce even more accurate simulations. Integrating physics-based simulations into GMPEs is a step that will directly benefit seismic hazard assessment moving forward, and in this work we outlined a repeatable approach to producing and integrating synthetic data into the GMPE development procedure.

Figure 16. Synthetic seismograms showing (a) fault perpendicular and along-strike velocities (cm/s) and (b) accelerations (cm/s²) for a simulated $M_w$ 5.4 earthquake at distances between 3 and 15 km. Seismograms are aligned with the events epicenter at an azimuthal angle of 90°. The color version of this figure is available only in the electronic edition.

Figure 17. Moment rate function for a simulated $M_w$ 5.4 earthquake. The color version of this figure is available only in the electronic edition.
**Figure 18.** (a) PGVs and (b) PSAs ($T = 0.2$ s) as functions of hypocentral distance for ground-motion data from the Oklahoma/Kansas target area composite ground-motion dataset including recorded data $M_w 4–5$, $M_w 5–5.5$, $M_w 5.5–5.8$, and simulated $M_w 4–5.8$ ground motions. GMPEs for $M_w 4.5$ and 5.5 constructed by fitting our composite catalog using equation (11) (coefficients shown in Table 3) are shown. Atkinson (2015) $M_w 4.5$ and 5.5 GMPEs included for reference. The color version of this figure is available only in the electronic edition.

**Table 3**

<table>
<thead>
<tr>
<th></th>
<th>Peak ground velocity (cm/s)</th>
<th>Peak spectral acceleration ($T = 0.2$ s) (cm/s$^2$)</th>
<th>Peak spectral acceleration ($T = 0.2$ s) (cm/s$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M_w 4.5$</td>
<td>$M_w 5.5$</td>
<td>$M_w 5.5$</td>
</tr>
<tr>
<td>$c_1$</td>
<td>$-1.87$</td>
<td>$-3.92$</td>
<td>$-2.42$</td>
</tr>
<tr>
<td>$c_2$</td>
<td>$1.16$</td>
<td>$2.38$</td>
<td>$2.04$</td>
</tr>
<tr>
<td>$c_3$</td>
<td>$0.06$</td>
<td>$0.18$</td>
<td>$0.16$</td>
</tr>
<tr>
<td>$c_4$</td>
<td>$7.63 \times 10^4$</td>
<td>$-3.08 \times 10^5$</td>
<td>$-1.16 \times 10^4$</td>
</tr>
<tr>
<td>$c_5$</td>
<td>$0.17$</td>
<td>$0.15$</td>
<td>$0.25$</td>
</tr>
<tr>
<td>$c_6$</td>
<td>$7.63 \times 10^4$</td>
<td>$-3.08 \times 10^5$</td>
<td>$-1.16 \times 10^4$</td>
</tr>
<tr>
<td>$c_7$</td>
<td>$0.35$</td>
<td>$0.09$</td>
<td>$0.04$</td>
</tr>
<tr>
<td>$c_8$</td>
<td>$-7.63 \times 10^4$</td>
<td>$3.08 \times 10^5$</td>
<td>$1.16 \times 10^4$</td>
</tr>
<tr>
<td>$c_9$</td>
<td>$1.53$</td>
<td>$4.71$</td>
<td>$4.04$</td>
</tr>
</tbody>
</table>

Coefficients $c_1$–$c_9$ obtained by fitting the composite recorded/simulated ground-motion catalog for the Oklahoma/Kansas target region using equation (11) for RotD50 horizontal component ground-motion intensity measures PGV (in cm/s) and 1, 2, and 5 Hz peak spectral acceleration (PSA) (5% damping, in cm/s$^2$). GMPE, ground-motion prediction equation; PGV, peak ground velocity.

**Figure 19.** (a) PGVs and (b) PSAs ($T = 0.2$ s) as functions of hypocentral distance for ground-motion data from the Oklahoma/Kansas target area composite ground-motion dataset including recorded data $M_w 3–4$, $M_w 4–5$, $M_w 5–5.5$, $M_w 5.5–5.8$, and simulated $M_w 3–5.8$ ground motions. GMPEs for $M_w 3.5, 4.5$, and 5.5 constructed by fitting our composite catalog using equation (11) (coefficients shown in Table 3) are shown. Atkinson (2015) $M_w 3.5, 4.5$, and 5.5 GMPEs included for reference. The color version of this figure is available only in the electronic edition.

**Data and Resources**

Ground-motion data used in this study were collected from Incorporated Research Institutions for Seismology (IRIS) Data Services (http://ds.iris.edu/ds/, last accessed February 2017). Processing methods are detailed in Gupta et al. (2017). Earthquake magnitudes, locations, and depths were updated to those contained in the U.S. Geological Survey (USGS) Advanced National Seismic System (ANSS) composite catalog (http://earthquake.usgs.gov/data/comcat/, last accessed February 2017). The well log data were obtained from a member of the Stanford Center for Induced and Triggered Seismicity.

**Acknowledgments**

Funding was provided by the Stanford Center for Induced and Triggered Seismicity (SCITS). The authors thank the directors and...
Figure 20. Plots of residuals (difference between observed and predicted in log units, binned every 10 km) for (a) PGV and (b) PSA ($T = 0.2$ s) for all ground-motion data ($M_w > 3$) in the composite target region ground-motion catalog. Squares indicate the mean values of each bin with errors bars indicating ±1 standard deviation. The top panel in each column shows the residuals between the composite GMPE and the recorded data, the middle panel shows the residuals between the composite GMPE and the simulated data, and the bottom panel shows the residuals between the composite GMPE and the combined dataset. The means are near zero or well within 1 standard deviation for all cases, indicating that our simulations are producing ground-motion data that have similar amplitude and decay characteristics as recorded data. The color version of this figure is available only in the electronic edition.
Figure 21. Plots of residuals (difference between observed and predicted in log units, binned every 10 km) for (a) PGV and (b) PSA ($T = 0.2$ s) for ground-motion data from events $M_w > 4$ in the composite target region ground-motion catalog. Squares indicate the mean values of each bin with errors bars indicating ±1 standard deviation. The top panel in each column shows the residuals between the composite GMPE and the recorded data, the middle panel shows the residuals between the composite GMPE and the simulated data, and the bottom panel shows the residuals between the composite GMPE and the combined dataset. The means are near zero or well within 1 standard deviation for all cases, indicating that our simulations are producing ground-motion data that have similar amplitude and decay characteristics as recorded data. The color version of this figure is available only in the electronic edition.
participants of SCITS, especially Abhineet Gupta, Bill Ellsworth, Greg Beroza, and Jack Baker, for their feedback on this project. The authors thank Editor-in-Chief Thomas Pratt and Editors Fabrice Cotton and Ralph Archuleta for their thorough reviews that substantially improved this article.

References


