Bulletin of the Seismological Society of America, Vol. 100, No. 5A, pp. 2185–2195, October 2010, doi: 10.1785/0120090209

Investigation of Geometrical Spreading and Quality Factor Functions in the New Madrid Seismic Zone

by Arash Zandieh and Shahram Pezeshk

Abstract The accuracy and applicability of geometrical spreading and quality factor functions are investigated for the New Madrid seismic zone (NMSZ) using recorded small and moderate earthquakes. These functions represent the path effect in frequency domain.

The database used in this study consists of 500 broadband seismograms from 63 events of magnitude M_w 2.5 to 5.2, recorded by the Center for Earthquake Research and Information (CERI) at the University of Memphis. The hypocentral distances range from 10 to 400 km. All the broadband stations are located within the Mississippi embayment with different site conditions. The vertical components of the records are processed and used to define the path effect term in frequency range of 0.2 to 30 Hz. A hinged-trilinear geometrical spreading and frequency-dependent quality factor functions are used to describe the path term. The regression analysis using a genetic algorithm (GA) indicates that at distances less than 70 km the spectral amplitudes decay as R^{-1}; between 70 and 140 km spectral amplitudes increase with distance and the geometric spreading is defined as $R^{+0.25}$; beyond 140 km, the attenuation is described by $R^{-0.5}$. The quality factor function is expressed as $Q = 614f^{0.32}$ for frequencies greater than 1 Hz after the regression analysis. For the broader range of frequency used in this study (0.2 to 30 Hz), the Q function is described by a third-degree polynomial described as $\log Q(f) = 2.898 - 0.464 \log f + 1.238(\log f)^2 - 0.540(\log f)^3$. The results of this study are compared with those of Atkinson (2004) and Samiezade-Yazd *et al.* (1997). The path term obtained in this study can be used in the stochastic method to predict ground motions in the NMSZ and eastern North America (ENA).

Introduction

In seismic areas of the world such as eastern North America (ENA), where strong ground-motion data are sparse, developing models to predict ground motion are of great significance. These models estimate the ground motion corresponding to seismological characteristics of the region such as magnitude, distance, and local site conditions. One widely used model to estimate ground motion is the stochastic model. In this method earthquake source, propagation path, and near-surface elements are modeled in the frequency domain (Boore, 2003). Atkinson and Boore (1995, 1998, 2006), Frankel *et al.* (1996), and Toro *et al.* (1997) are examples of using the stochastic method to develop ground-motion prediction equations in ENA. The accuracy of the stochastic procedure relies on models used to describe source, path, and near-surface site condition. In this study, we only focus on the propagation path effects. Path effect is modeled in the frequency domain by multiplying a geometrical attenuation function by an anelastic attenuation function (Boore, 2003)

$$P(R, f) = Z(R) \exp[-\pi f R / Q(f) \beta_s], \qquad (1)$$

where $Z(R)$ is the geometrical spreading function, $Q(f)$ is the quality factor function that accounts for anelastic crustal path attenuation, β_s is the shear-wave velocity of the crust, f is the frequency, and R is the hypocentral distance. In practice, there is a trade-off between the anelastic attenuation and the geometrical spreading; that is, the empirical data does not have the ability to distinguish between the two, and only the overall path effect attenuation can be described by the analysis of the ground-motion amplitudes (Atkinson and Mereu, 1992). According to equation (1), at near-source distances, the path effect is mostly dominated by the geometrical spreading, and the anelastic attenuation affects the path term considerably at larger regional distances. Atkinson and Mereu (1992) used a regression analysis on a dataset of 1200 vertical component seismograms from 100 small and moderate earthquakes to study the attenuation of the Fourier amplitude of acceleration in southeastern Canada. They found that the geometrical spreading is described by $R^{-1.1}$ at distances less than 70 km, which is slightly faster than the theoretical attenuation of the direct waves in an ideal

whole-space (R^{-1}). Between 70 and 130 km, the spectral amplitudes are approximately constant (R^0) because the direct waves are joined by postcritical reflections from Moho discontinuity. Beyond 130 km, amplitudes decay as $R^{-0.5}$ consistent with theoretical attenuation of the Lg surface waves in a half-space (Atkinson and Mereu, 1992). The corresponding quality factor function is given by $Q = 670f^{0.33}$. Atkinson and Boore (1995) modified the geometrical spreading at distances less than 70 km to R^{-1} and refitted the Q model to $Q = 680f^{0.36}$. Atkinson (2004) studied 1700 seismograms from 186 earthquakes of magnitude m_N 2.5–5.6 recorded on hard rock sites of southeastern Canada and northeastern United States. Atkinson (2004) determined the amplitude decay as $R^{-1.3}$ within 70 km of the source, which is faster than previous studies. From 70 km to 140 km, Moho bounced waves cause a decrease in attenuation to$R^{+0.2}$. In the transition zone, spectral amplitudes increase with distance for low frequencies (Atkinson, 2004). Beyond 140 km, the attenuation is described by $R^{-0.5}$. For frequency above 1 Hz, Atkinson (2004) represented the Q model as $Q = 893f^{0.32}$, and over all frequency ranges she described the Q function by a third-degree polynomial. Sonley and Atkinson (2006) studied the path attenuation at distances less than 100 km from the source in the Charlevoix seismic zone of eastern Canada. They analyzed three-component seismograms from 350 earthquakes and found the geometrical attenuation of $R^{-1.36}$ for distances less than 70 km. Allen *et al.* (2007) studied 1200 vertical component records from 84 earthquakes in southeastern Australia, which is also considered a stable continental region much like ENA. They found that the low-frequency amplitudes decay can be described as $R^{-1.3}$ within 90 km of the source. From 90 to 160 km, because of the postcritical reflections the geometrical spreading is approximated by $R^{+0.1}$; beyond 160 km, low-frequency amplitudes decay as $R^{-1.6}$. Allen and Atkinson (2007) compared the source spectra and attenuation in ENA and southeastern Australia and concluded that the properties of these two regions are very similar at short distances ($R < 70$ km). At larger distances, observed attenuation differences may be due to differences in crustal structure. Samiezade-Yazd *et al.* (1997) examined nearly 2200 vertical component waveforms from 237 earthquakes in the New Madrid seismic zone (NMSZ) and defined the distance scaling for earthquakes in the NMSZ including a multihinged geometrical spreading function, a multihinged duration function, and an exponential quality factor function. They determined the geometrical spreading function as $R^{-1.0}$ for distances less than 50 km, $R^{-0.25}$ for distances from 50 to 120 km, $R^{0.0}$ for distances from 120 to 200 km, $R^{-0.5}$ for distances from 200 to 220 km, and $R^{-1.0}$ for distances from 220 to 500 km. The quality factor function coherent with this geometrical spreading function is described as $Q = 900f^{0.3}$ (Samiezade-Yazd *et al.*, 1997). This latter study may suggest there is a regional variability in path attenuation term in between the central United States, specifically the NMSZ, and the other regions of ENA, such as southeastern Canada.

To develop proper regional ground-motion prediction equations that play a key role in the seismic hazard analysis, an accurate regional path term is essential. An appropriate path term is also necessary to derive a proper source spectrum of the recorded earthquakes. In this study, we estimate the path term, including the geometrical spreading and quality factor functions for the NMSZ, by examining recorded seismograms in the region. The purpose of this study is not to invalidate the use of the models already developed for ENA in the NMSZ; rather, it is to provide an alternative model for the path effect along with the other existing models. This will help in providing a better definition of epistemic uncertainties.

The procedure used in this study is similar to the method used by Atkinson and Mereu (1992) to model the functional form of attenuation curves in southeastern Canada. A Monte Carlo simulation is used to test the statistical significance of the evaluated functional forms for the geometrical spreading and the quality factor functions. The resulting path effect model is compared with the Atkinson (2004) and the Samiezade-Yazd *et al.* (1997) models.

Database

In this project the seismograms recorded by the Center for Earthquake Research and Information (CERI) broadband stations are used. CERI's broadband stations consist of 11 stations. These stations and their identifications are listed in Table 1. The locations of broadband stations are shown in Figure 1. All the broadband stations are located in the Mississippi embayment with different site conditions. The database consists of 500 broadband seismograms from 63 local events of magnitude M_w 2.5 to 5.2 that occurred from 2000 to 2009. The hypocentral distances range from 10 to 400 km. The distribution of earthquakes in magnitude and distance is shown in Figure 2. The most notable event in the database is the M_w 5.2 earthquake that struck near Mount Carmel in southern Illinois on Friday morning, 18 April 2008 at 09:36:59 UTC (04:37 CDT). A map of CERI's broadband stations as well as the location of earthquakes used in this study are shown in Figure 1.

The sensor type of all broadband stations is the Guralp CMG40T. This instrument has a velocity response that is flat between 30 s and 50 Hz. We used the high-gain channel with a sampling rate of 100 samples per second.

It should be noted that only the vertical component ground motions are used in this study. One of the techniques used to estimate the site response is the observed horizontal-to-vertical component (H/V) ratio (Nakamura, 1989; Lermo and Chávez-García, 1993; Atkinson and Cassidy, 2000). In this approach the site amplification is assumed to be negligible for the vertical component; hence, the vertical component may mostly reflect the effects of the source and path attenuation. Therefore, the spectral ratio of the horizontal-to-vertical component can be used to estimate the site amplification of the horizontal component. Lermo and

Table 1
CERI Broadband Station List

Station Name	Latitude	Longitude	Elevation (m)	Sensor	Location
GLAT	36.26937	−89.28756	120.0	CMG40T	Glass, TN
GNAR	35.9652	−90.0178	71.0	CMG40T	Gosnell, AR
HALT	35.91060	−89.33953	85.0	CMG40T	Halls, TN
HBAR	35.5550	−90.6572	74.0	CMG40T	Harrisburg, AR
HICK	36.5409	−89.2288	141.0	CMG40T	Hickman, KY
HENM	36.7160	−89.4717	88.0	CMG40T	Henderson Mound, MO
LNXT	36.10138	−89.49127	144.0	CMG40T	Lenox, TN
LPAR	35.6019	−90.3002	66.5	CMG40T	Lepanto, AR
PARM	36.6635	−89.7522	85.0	CMG40T	Stahl Farm, MO
PEBM	36.11312	−89.86229	76.0	CMG40T	Pemiscott Bayou, MO
PENM	36.4502	−89.6280	85.0	CMG40T	Penman, Portageville, MO

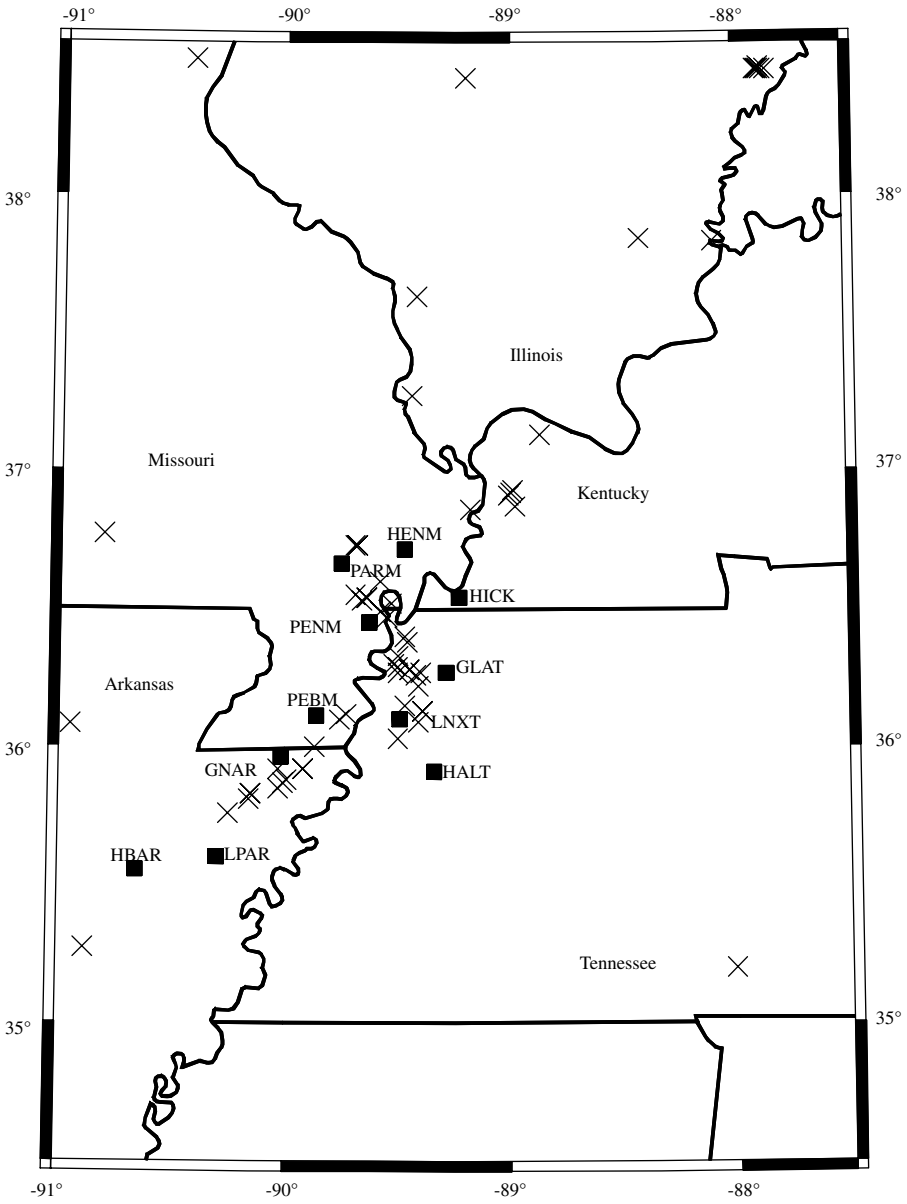

Figure 1. Map of CERI's broadband stations (squares) and earthquakes used in this study (crosses).

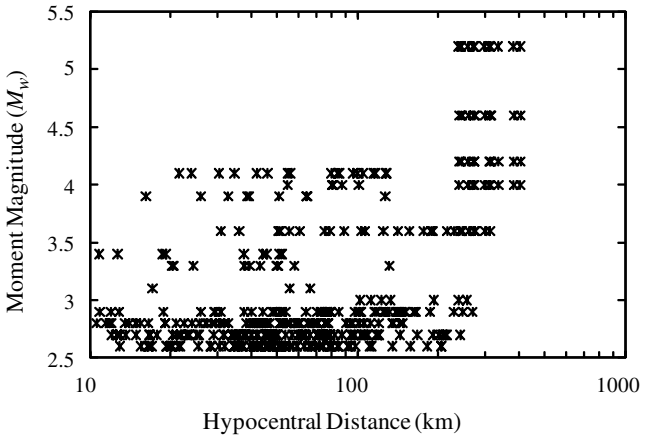

Figure 2. Magnitude and distance distribution of earthquakes used in this study.

Chávez-García (1993) showed that the H/V ratio technique works well in estimating the frequency and amplitude of the first resonant response mode in sedimentary deposits. Atkinson and Cassidy (2000) examined the applicability of the H/V ratio to determine the soil amplification of the Fraser River Delta, British Columbia, and concluded that the H/V ratio is a reasonable first approximation to amplification and is useful for sites with unknown soil profiles. Atkinson (2004) developed the attenuation model for the vertical component, due to the paucity of the horizontal data, and applied it to the horizontal component database. After examining the residuals, no trends were found with distance; she concluded that an independent model for the horizontal component is not necessary. The H/V ratio, as an approximation of the site amplification, is applied to the vertical component model to predict the horizontal component (Atkinson, 2004). In this study we only used vertical component ground motions to estimate the path term. We assumed that the site effects are negligible for the vertical component. The study of the H/V ratio and its ability to estimate the site effects in the NMSZ is the subject of our future studies. In fact, by checking the variability of H/V ratio with distance, a conclusion can be made if the path effect derived in this study is applicable for the horizontal component ground motions as well.

Data Processing

Each waveform contains direct, reflected, and refracted phases of both S and P waves. Atkinson and Mereu (1992) and Atkinson (2004) used the shear-window, a part of the signal containing shear-wave phases, to model the attenuation caused by the path. The shear-window starts when the first direct shear-wave arrives and ends when all significant shear phases arrive. In this study, we examined the difference between using shear-window and the whole waveform length. This is done for several waveforms with different hypocentral distances while the waveforms are dominated by different phases at different distances. As distance increases, the direct waves will be joined by reflected and refracted waves. It should be noted that the earthquakes used in this study are all local earthquakes with hypocentral distances less than 400 km with maximum focal depth of 30 km. Therefore, the observed phases in waveforms are limited. We concluded that the Fourier amplitudes are very close when using shear-window or the whole record. Figure 3 shows this comparison for three records with different hypocentral distances. As it can be observed from Figure 3, the acceleration power spectral density spectrums (power spectra) of the records are very close for shear-window and whole length of the signal. The details on calculating the power spectra are discussed later in this section.

To obtain Fourier spectrum of the signal while excluding the preevent noise, we used Welch's method of power spectral density (PSD) estimation (Welch, 1967). Welch's method consists of dividing the time series data into overlapping segments, computing periodogram (squared amplitude of discrete-time Fourier transform) of each windowed segment, and then averaging the PSD estimates. In this study, we used 256-sample segments with 50% overlap. Each segment is windowed using the Hamming window to minimize the spectral leakage between frequencies. For the sampling rate of 100 samples per second, the duration of each segment

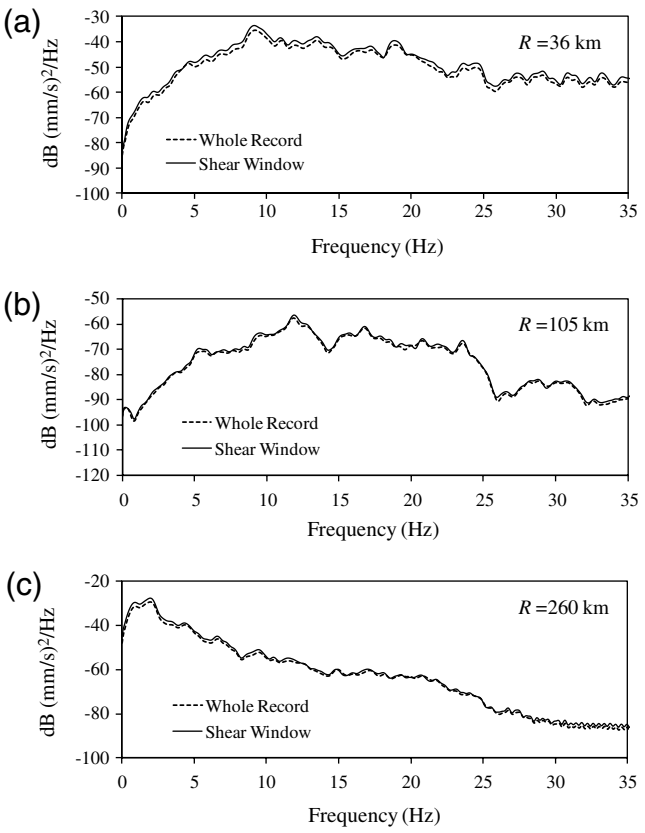

Figure 3. Comparison of the shear-window and the whole length record power spectra for three records with different hypocentral distances.

is 2.5 s, and 1024 Fourier transform points are used while evaluating the fast Fourier transform (FFT) of the windowed segment; therefore, the frequency bins from FFT are 0.0977 Hz wide.

The recorded waveforms are first corrected for the instrument response using the instrument transfer function. Then the velocity is converted to acceleration and the PSD of acceleration, P_a, is calculated using Welch's method. The obtained power spectra should be corrected for preevent noise. The noise window for each waveform is the segment of the record ahead of the P arrival. The noise window is processed in the same way as the signal and the noise PSD, P_N, is calculated. The PSD of the noise-corrected acceleration signal, P_s, is given by

$$P_s = P_a - P_N. \tag{2}$$

The Fourier amplitude of the noise-corrected acceleration is the square root of power spectra calculated in equation (2). At this stage the signal-to-noise ratio was evaluated for each record for different frequencies. Data were selected for further analysis only at frequencies for which the signal-to-noise ratio exceeded 2.

The Fourier spectrum of ground acceleration for each waveform in the database is obtained using the discussed method and tabulated for frequencies of 0.2 to 30 Hz. It should be noted that data are most abundant in the frequency range of 1 to 10 Hz.

Analysis of Path Effect

In this study a procedure similar to that introduced by Atkinson and Mereu (1992) is used to analyze the path effect. The geometrical spreading function and quality factor function are estimated using regression of data to an equation that models the observed spectral amplitudes. This equation is given by (Atkinson and Mereu, 1992)

$$\log A_{ij}(f) = \log A_i^s(f) - B(f)\log R_{ij} - C(f)R_{ij} + \log S_j(f), \tag{3}$$

where $A_{ij}(f)$ is the observed spectral amplitude of earthquake i at station j at frequency f, $A_i^s(f)$ is the source spectral amplitude of earthquake i, R is the hypocentral distance, $B(f)$ is the geometrical spreading coefficient, $C(f)$ is the coefficient of anelastic attenuation, and $S_j(f)$ is the site response term for station j (Atkinson and Mereu, 1992). The relation between coefficient $C(f)$ and quality factor is defined by Atkinson and Mereu (1992) as follows:

$$Q(f) = [\log(e)\pi f]/[C(f)\beta_s], \tag{4}$$

where β_s is the shear-wave velocity. Some studies of attenuation in ENA used a frequency-dependent quality function of the form $Q(f) = Q_0 f^n$ $(n \geq 0)$ where $Q(f)$ increases with frequency (Atkinson and Mereu, 1992; Atkinson and Boore, 1995; Samiezade-Yazd *et al.*, 1997).

Atkinson and Mereu (1992) described the geometrical spreading function, by a hinged-trilinear functional form, in which the slope is different in three distance segments of less than 70 km, 70 to 130 km, and beyond 130 km. Atkinson (2004) used the same functional form while studying attenuation of ground motion in southeastern Canada and the northeastern United States. The hinged-trilinear form of the geometrical spreading is given by

$$B(f)\log R_{ij} = \begin{cases} b_1(f)\log R_{ij} & R_{ij} \leq R_1 \\ b_1(f)\log R_1 + b_2(f)\log(R_{ij}/R_1) & R_1 < R_{ij} \leq R_2 \\ b_1(f)\log R_1 + b_2(f)\log(R_2/R_1) + b_3(f)\log(R_{ij}/R_2) & R_{ij} > R_2 \end{cases}. \tag{5}$$

At close distances the direct waves are dominant and based on theoretical studies of the geometrical spreading in a homogeneous elastic whole-space $b_1 = 1.0$ (Atkinson and Mereu, 1992). Atkinson (2004) found the attenuation steeper than R^{-1} at distances less than 70 km. At distances beyond 70 km, the postcritical reflections from Moho discontinuity join the direct waves and cause an energy enhancement and decrease the attenuation rate (Burger *et al.*, 1987; Atkinson and Mereu, 1992). At large and regional distances, multiple refractions and reflections of the shear waves trapped in crustal guides, Lg waves, dominate the records (Atkinson and Mereu, 1992). Atkinson (2004) and Atkinson and Mereu (1992) assumed the value of 0.5 for the geometric spreading coefficient b_3 for Lg waves, based on several studies. In this study, we also fixed the value of b_3 to 0.5.

Several studies used the hinged-trilinear geometrical spreading form of equation (5) in their ground-motion modeling studies (Atkinson and Mereu, 1992; Atkinson and Boore, 1995; Atkinson, 2004). Atkinson and Mereu (1992) found that there is no significant dependence of $b_1(f)$, $b_2(f)$, and $b_3(f)$ on frequency; their resulting coefficients were: $b_1 = 1.1 \pm 0.1$, $b_2 = 0.0 \pm 0.2$, $b_3 = 0.5$, $R_1 = 70 \pm 5$ km, and $R_2 = 130 \pm 10$ km. They also described the quality function as $Q(f) = 670 f^{0.33}$. Atkinson (2004) confirmed

the independence of coefficient to frequency and found the best result when $b_1 = 1.3$, $b_2 = -0.2$, $b_3 = 0.5$, $R_1 = 70$ km, and $R_2 = 140$ km. Atkinson (2004) expressed the quality factor for frequencies greater than 1 Hz as $Q(f) = 893f^{0.32}$. For a wider range of frequencies (0.2 to 20 Hz), Q is modeled by a third-degree polynomial of $\log Q(f) = 3.052 - 0.393 \log f + 0.945(\log f)^2 - 0.327(\log f)^3$. Allen *et al.* (2007) also described the Q for southeastern Australia as a third-degree polynomial given by $\log Q(f) = 3.66 - 1.44 \log f + 0.768(\log f)^2 + 0.058(\log f)^3$. Samiezade-Yazd *et al.* (1997) determined a multihinge geometrical spreading function similar to equation (5), but with four hinge points instead of two, to descried the distance scaling in the NMSZ up to 500 km. In their model, the b coefficients and hinge points are: $b_1 = 1.0$, $b_2 = 0.25$, $b_3 = 0.0$, $b_4 = 0.5$, $b_5 = 1.0$, $R_1 = 50$ km, $R_2 = 120$ km, $R_3 = 200$ km, and $R_4 = 220$ km. The quality factor function in their study is described by $Q = 900f^{0.3}$.

For modeling the attenuation caused by path, the method proposed by Atkinson and Mereu (1992) is used. In equation (3) the source term is the only term that depends on the event exclusively and not on the station location. We take advantage of this dependency and normalize all the observed amplitudes to the source level. In the algorithm used in this study to find the path effect, and explained later, the source term is evaluated for each event at each station, then the path term is found by minimizing the differences between source terms, $\log A_{ij}^s(f)$, evaluated for all j stations recorded event i. By rearranging equation (3) and replacing $\log A_i^s(f)$ with $\log A_{ij}^s(f)$,

$$\log A_{ij}^s(f) = \log A_{ij}(f) + B(f) \log R_{ij} + C(f)R_{ij} - \log S_j(f), \tag{6}$$

where $\log A_{ij}^s(f)$ is the source parameter evaluated for event i at station j. It should be noted that the site term is assumed to be constant at each station and is event invariant, while the ground-motion intensities of the events in the database are low enough to prevent any nonlinear soil behavior.

Our goal is to find the best model for path attenuation, consisting of the geometrical spreading and the quality factor, using the available NMSZ database. Therefore, the objective is to find the variables $b_1(f)$, $b_2(f)$, $b_3(f)$, R_1, R_2, $C(f)$, and $S_j(f)$ using equations (3) through (6). Applying different combinations of the mentioned parameters in equation (6) results in an estimation of the source amplitude corresponding to station j, $A_{ij}^s(f)$. The best combination is the one that minimizes the mean absolute value of residuals defined as

$$\text{res}_i = \frac{1}{N} \sum_{j=1}^{N} \left| \log A_{ij}^s(f) - \overline{\log A_i^s(f)} \right|, \tag{7}$$

where $\overline{\log A_i^s(f)}$ is the average of $\log A_{ij}^s(f)$ over all N stations subscribed by j,

$$\overline{\log A_i^s(f)} = \frac{1}{N} \sum_{j=1}^{N} \log A_{ij}^s(f). \tag{8}$$

In other words, the best solution is the one that results in almost the same source terms for stations that recorded the same events.

In this study a genetic algorithm (GA) is used to find the best combination of variables $b_1(f)$, $b_2(f)$, $b_3(f)$, R_1, R_2, $C(f)$, and $S_j(f)$ (Holland, 1975; Goldberg, 1989). The GA focuses on a population of variables, which are created randomly in the range defined by the constraints. Variables are grouped in sets, each of which is called a string and composed of a series of characters that defines a possible solution for the problem. The performance of the variables, as described by the objective function and the constraints, is represented by the fitness of each string. A mathematical expression, called a fitness function, calculates a value for a solution of the objective function. The fitter solution gets the higher value, and the ones that violate the objective function and constraints are penalized. Therefore, like what happens in nature, the fittest and best solutions will survive and get the chance to be a parent of the next generation. In a crossover procedure, two selected parents reproduce the next generation. The procedure first divides the selected parent strings into segments, then some of the segments of a parent string are exchanged with corresponding segments of another parent string. The mutation operation guarantees diversity in the generated populations. This is done by flipping (0 to 1 or vice versa) a randomly selected bit in the selected binary string to create a muted string. Mutation prevents a fixed model of solutions from being transferred to the next generation (Holland, 1975; Goldberg, 1989).

The genetic algorithm tries different combinations of the variables and finds the best solution that minimizes the mean value of the errors defined in equation (7) over all events. Therefore, the objective function that has to be minimized by the GA is defined as

$$\text{Objective Function} = \frac{1}{M} \sum_{i=1}^{M} \text{res}_i, \tag{9}$$

where M is the number of events.

We used the hinged-trilinear functional form of equation (5) for geometrical spreading function. In the GA, the constraints imposed on the coefficient b of the hinged-trilinear geometrical spreading function of equation (5) are: $1.0 \le b_1 \le 2.0$, $-0.5 \le b_2 \le 0.5$, and $b_3 = 0.5$. Parameters R_1 and R_2 in equation (5) were not treated as the regression variables in the GA. Instead, the GA process was repeated for all combinations of $R_1 = 50, 60, 70, \ldots 120$ and $R_2 = 70, 80, 90, \ldots 160$. The combination that yields the minimum residuals is selected. The overlapped values of the parameters R_1 and R_2 were included to examine the possibility of a bilinear shape of equation (5). As mentioned before, we assumed that the soil amplification for the vertical component is negligible. To fulfill this assumption in the regression

analysis, the average site term $\log S_j(f)$ over all stations should be set to zero. To impose this constraint on the regression analysis, we added the term $\frac{1}{N}\sum_{j=1}^{N}|\log S_j(f)|$ to the right side of equation (7) to minimize it along with the $\frac{1}{N}\sum_{j=1}^{N}|\log A_{ij}^{s}(f) - \overline{\log A_i^s(f)}|$ term during the GA analysis.

Results

The proposed GA procedure is applied to the database considering a frequency range of 0.2 to 30 Hz using 0.1 log increments. Analyses show that there is no significant dependence of coefficients b_1, b_2, and b_3 on frequency. The resulting parameters for the geometrical spreading function are: $b_1 = 1.0$, $b_2 = -0.25$, $b_3 = 0.5$, $R_1 = 70$ km, and $R_2 = 140$ km. The result for the $Q(f)$ function is shown in Figure 4. The Q function has larger amplitudes at low frequencies ($f < 1$ Hz). We fitted a third-degree polynomial to the estimated Q values as

$$\log Q(f) = 2.898 - 0.464\log f + 1.238(\log f)^2 - 0.540(\log f)^3. \quad (10)$$

For the frequencies higher than 1 Hz, the Q function can be described by the exponential form as $Q(f) = 614f^{0.32}$.

Because the site term was constrained in the regression analysis (average site term was set to zero over all stations), the evaluated site terms are very small. At all frequencies and for all stations, the site terms are less than 0.03 log units.

For each event in the database, the averaged source spectrum over all the recording stations, defined by equation (8), is estimated after the GA regression analysis. Any azimuthal bias engendered in source spectrum determined by an individual station would be diminished by averaging the estimates from multiple stations (Atkinson and Cassidy, 2000).

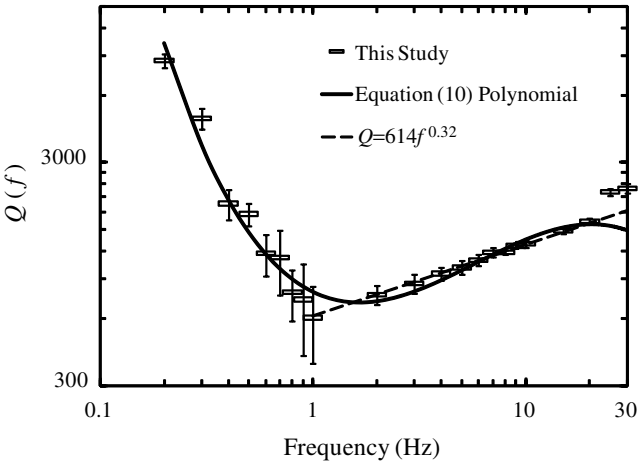

Figure 4. Q values obtained by regression analysis in this study (rectangles) with their standard errors bars; third-degree polynomial fitted to Q values (solid line); and linear Q model for $f > 1$ Hz (dashed line).

The estimated source spectra of the events used in this study for magnitudes M_w 3.0–4.1 are shown in Figure 5, as well as the Brune (1970, 1971) source model. The overall shape of the evaluated source spectra is consistent with the Brune spectrum. In Figure 5, the Brune spectra are plotted for a stress drop of 50 bars, which matches the evaluated high-frequency spectral levels for magnitudes M_w 3.0–4.1. In fact, the stress drop of 50 bars is not recommended for use with the stochastic point or finite-source ground-motion simulations for moderate and large events in the NMSZ. Atkinson (2004) also found stress drops of significantly less than 150 bars for events with $M_w < 4.0$ in southeastern Canada and the northeastern United States. She determined that the stress drop increases with moment magnitude for events with $M_w < 4.3$, and it reaches an approximately constant level of 100 to 200 bars for larger events. In this study, we did not evaluate the stress drop for events with $M_w > 4.1$ due to paucity of data in this magnitude range. The comparison shown in Figure 5 is only made to illustrate the consistency between overall shapes of the evaluated source spectrum in this study with the Brune spectrum.

To have a visual comparison of the observed data with the model obtained for the path effect, the source and site terms are removed from recorded amplitudes. The remaining spectral amplitudes represent the path effect. The site and source corrected amplitudes are referred to as normalized amplitudes (Atkinson and Mereu, 1992) and are given by

$$\log A_{ij}^{\text{norm}}(f) = \log A_{ij}^{s}(f) - \log A_{ij}(f) + \log S_j(f), \quad (11)$$

where $\log A_{ij}^{\text{norm}}(f)$ is the normalized amplitude. The averaged source, equation (8), is used as the source term, $\log A_{ij}^{s}(f)$, for each event. The normalized amplitudes for several frequencies are shown in Figure 6 along the path term of $B(f)\log R + C(f)R$, computed using the coefficients derived in this study. The scatter observed in Figure 6 shows

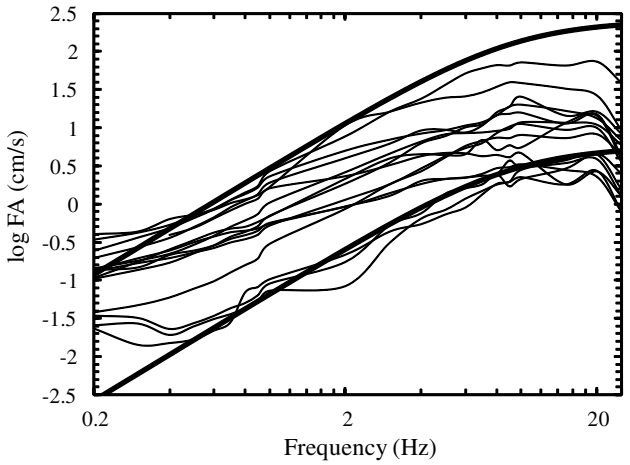

Figure 5. Estimated source spectra for events of $3.0 \le M_w \le 4.1$ (light lines) and the Brune source spectra (heavy lines). The Brune source model is plotted for stress drop of 50 bars.

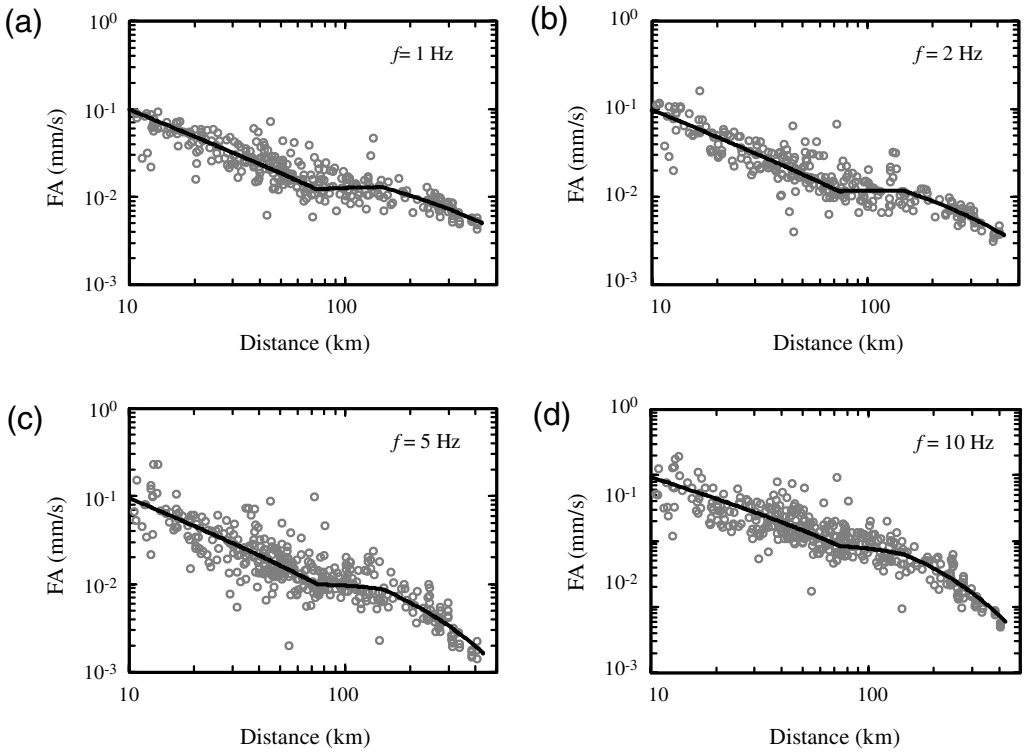

Figure 6. Observed normalized amplitudes (circles) and the predicted path effect (solid lines) for frequencies (a) 1, (b) 2, (c) 5, and (d) 10 Hz.

the uncertainty involved with the data. To study the statistical significance of this scatter, we subtract the estimated path term from normalized amplitudes and define them as residuals given by

$$\text{res}^{\text{norm}} = \log A_{ij}^{\text{norm}}(f) - B(f)\log R_{ij} + C(f)R_{ij}. \quad (12)$$

By substituting equation (11) into equation (12) and comparing it with equation (3), the residuals defined in equation (12) are also the residuals of the regression defined as the logarithm of the observed Fourier amplitudes minus the logarithm of the predicted Fourier amplitudes defined by equation (3). Figure 7 shows the residuals defined by equation (12) versus distance for two frequencies of 1 and 5 Hz. As it can be observed, there is no apparent trend in residuals versus the distance. Statistical analysis of the residuals in equation (12) shows that these residuals are normally distributed with a zero mean. It should be noted that the residuals defined in equation (12) have the logarithmic units; therefore, the residuals in nonlogarithmic scale have the lognormal distribution. The residuals defined in equation (12) are used to evaluate the statistical properties, including standard errors and confidence interval, of the estimated path effect parameters.

The averaged residuals and their standard deviation at each station are plotted in Figure 8 for frequencies 1 and 5 Hz. While the site response term is neglected (the evaluated

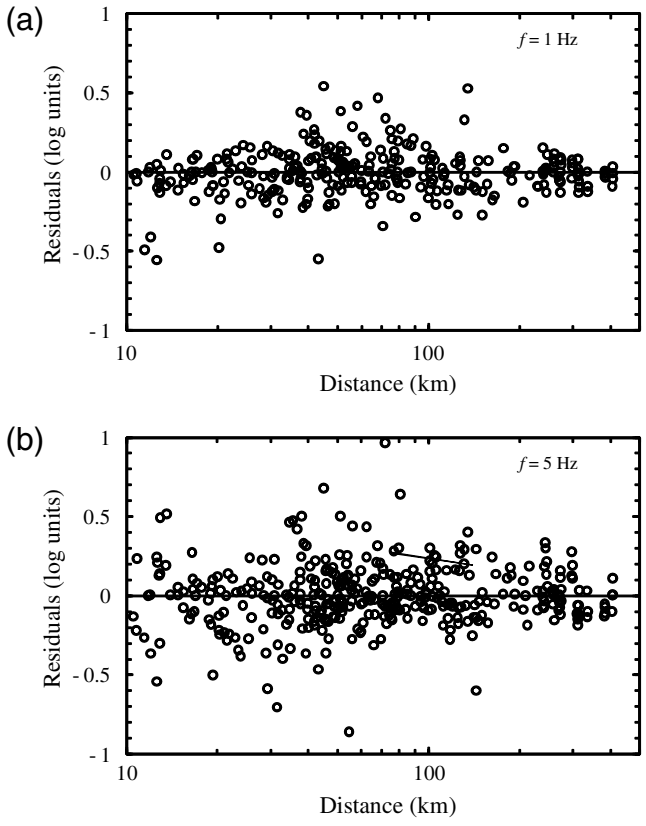

Figure 7. Residuals (log units) of the regression versus distance (equation 12), for frequencies of (a) 1 and (b) 5 Hz.

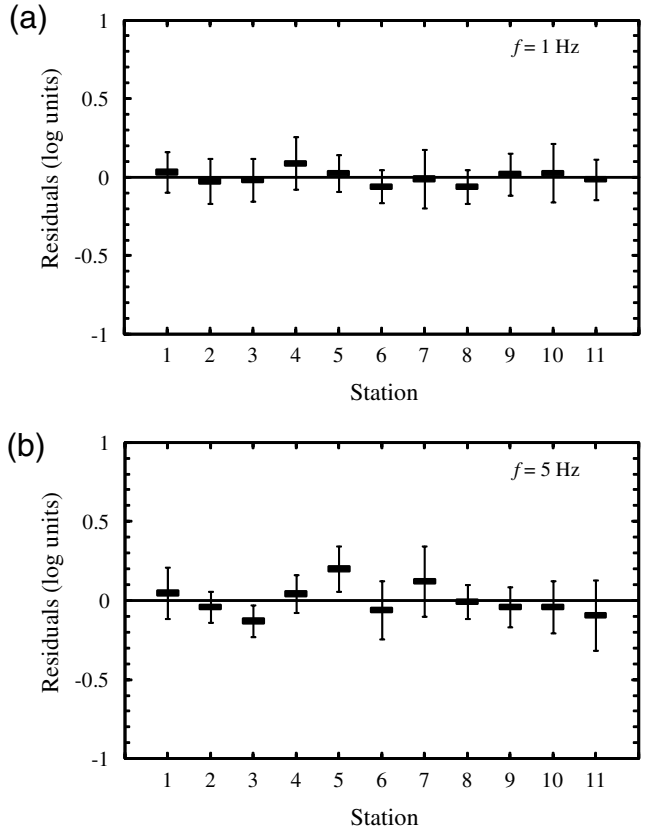

Figure 8. Averaged residuals (log units) and their standard deviation at each station, for frequencies of (a) 1 Hz and (b) 5 Hz. Station numbers are based on the station names provided in Table 1.

site terms, $\log S_j(f)$, are very small for all stations), the averaged residuals by station would be a measure of the significance of the site response at the stations. It should be noted that the variability observed in Figure 8 shows uncertainty exists in the Fourier amplitudes of ground motions at each station. The uncertainty existing in the site term and the way it is treated in the study consists of only one part of the total uncertainty. By averaging the residuals by stations, we were investigating if there was an unusual deviation at a station. While the residuals are averaged by stations, an observed deviation might be caused by an unpredicted site term at that specific site station. In other words, if a significant deviation in one station with respect to the others is observed, one can conclude relatively that the site effect has a larger impact on the observed amplitudes at that site. In Figure 8 we can see deviations for all the sites. This is more sensible when looking at the standard deviations of the residuals as well as the mean. However, the variability seen by standard deviations in Figure 8 suggests there may be some site responses, especially at high frequencies, for few stations.

To test the precision of the estimated parameters, their standard errors and confidence intervals are evaluated using the Monte Carlo simulation technique (Press *et al.*, 2007). For this purpose, 1000 datasets of the normalized amplitudes, equation (11), were simulated for the given estimated parameters $b_1(f)$, $b_2(f)$, $b_3(f)$, and $C(f)$. The datasets are simulated by bootstrapping the residuals (Efron and Tibshirani, 1993). The simulations were performed for frequency ranges of 1 to 10 Hz in 0.1 log frequency increments, where the data are most plentiful. Thus, we used the form of $Q(f) = 614f^{0.32}$ to define $C(f)$. At each frequency the normalized amplitudes, equation (11), are generated by adding randomly generated residuals to the given $B(f)\log R + C(f)R$ term, which is evaluated for the same distances, R, as the original database. The random residuals are drawn from a normal distribution with zero mean and standard deviation equal to the standard deviation of the original database residuals defined by equation (12). For each simulated normalized amplitude dataset, a regression analysis is used to find the path attenuation coefficients $b_1(f)$, $b_2(f)$, $b_3(f)$, and $C(f)$. The standard errors of the estimated parameters (the bootstrap estimated of the standard error) are the standard deviation of the estimated parameters of the simulated datasets (Efron and Tibshirani, 1993). In regression analysis of the simulated datasets, the Q function is constrained to $Q(f) = Q_0 f^{0.32}$ to evaluate the standard error of the Q_0. Also, the parameter b_3 is fixed to value of 0.5 in the regression. It should be noted that the mean of the estimated parameters attained from the simulated datasets are the same as the original database ($b_1 = 1.0$, $b_2 = -0.25$, and $Q_0 = 614$). The calculated standard errors ($\hat{\text{se}}$) are: $\hat{\text{se}}(b_1) = 0.09$, $\hat{\text{se}}(b_2) = 0.25$, and $\hat{\text{se}}(Q_0) = 25$.

There are several methods to evaluate the bootstrap confidence interval. The simplest one is the standard bootstrap confidence interval that works when the estimator is normally distributed. The $100(1-\alpha)\%$ standard bootstrap confidence interval on the estimator parameter is defined by Efron and Tibshirani (1993) as

$$\text{Estimate} = \pm z_{\alpha/2}.\hat{\text{se}}, \tag{13}$$

where $z_{\alpha/2}$ is the upper $100_{\alpha/2}$ percentage point of the standard normal distribution and $\hat{\text{se}}$ is the bootstrap estimate of the standard error. In this study, the distribution of the regression estimator parameters of the simulations b_1, b_2, and Q_0, are close to normal distribution. Therefore, equation (13) is used to estimate the confidence intervals. For example, the 90% confidence interval on the path attenuation parameters are $b_1 = 1.0 \pm 0.15$, $b_2 = -0.25 \pm 0.41$, $b_3 = 0.5$, and $Q = (614 \pm 41.12)f^{0.32}$. It should be noted that the hypocentral distance ranges in the database used in this study are from 10 to 400 km. Therefore, the result of this study is applicable for distances greater than 10 km; any use of this model for distances less than 10 km should be done with caution.

The path term, $P(R, f)$, derived in this study is compared with those obtained in Atkinson (2004) and Samiezade-Yazd *et al.* (1997) (see Fig. 9) for frequencies of 1 and 5 Hz. To make a consistent comparison of two path attenuation models according to the available data, the attenuation models of Atkinson (2004) and Samiezade-Yazd *et al.* (1997) are scaled to a level that provides the minimum sum

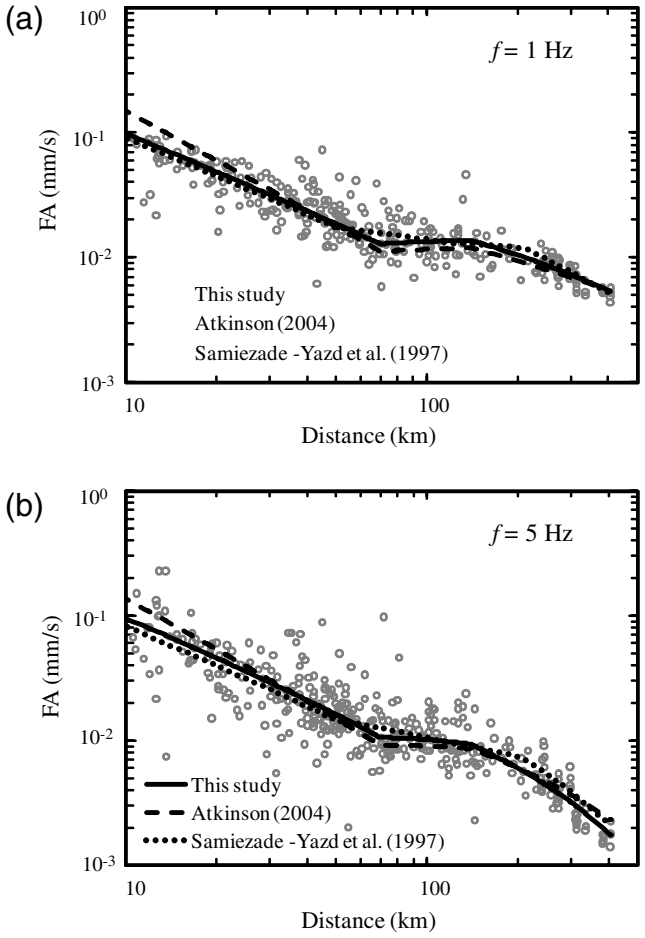

Figure 9. Comparison of the path term obtained in this study (solid line) with those of Atkinson (2004) (dashed line) and Samiezade-Yazd *et al.* (1997) (dotted line). Attenuation models of Atkinson (2004) and Samiezade-Yazd *et al.* (1997) are scaled to a level that provides the minimum sum of the squared residuals for the data on the plot (circles).

of the squared residuals for the data on the plot. There are some differences between the Atkinson (2004) model and the result of this study, especially at close distances, caused by a lower value for coefficient $b_1 = 1.0$ compared with that of Atkinson (2004) ($b_1 = 1.3$). The attenuation predicted by the Atkinson (2004) study is on average 20% higher than the result of this study for distances less than 40 km. However, based on the error analysis on the coefficient b_1, the 90% confidence interval of this coefficient is $b_1 = 1.0 \pm 0.15$. Therefore, the value of $b_1 = 1.3$ of the Atkinson (2004) model lies close to the upper limit of the 90% confidence interval on coefficient b_1 in this study. In general, the difference between the two models is not very significant for distances greater than 40 km. The Samiezade-Yazd *et al.* (1997) path model is very close to the developed model in this study for distances less than 50 km because the b_1 coefficients are the same for the two models. The greatest differences between the two models are for distances from 60 to 100 km and also greater than 180 km. The attenuation predicted by the Samiezade-Yazd *et al.* (1997) model is about 12% and 22% higher than that predicted in our model, respectively, for the distance ranges of 60–100 km and greater than 180 km.

It should be noted that source, site, and path term are correlated, and there is always a trade-off between these terms. Therefore, the reader is cautioned against using the path model derived in this study along with the source or site parameters deduced from other studies assuming different types of the path models.

Conclusions

In this study, 500 seismograms recorded by CERI's broadband stations for events of magnitude M_w 2.5 to 5.2 are used to determine the path term in the New Madrid seismic zone in the frequency domain. The vertical components of the records are corrected for the preevent noise, and their Fourier amplitudes of acceleration are computed over a broad frequency range of 0.2–30 Hz. A regression procedure using the GA is used to determine the best values for parameters defining the shape of attenuation caused by path. This attenuation is described by geometrical spreading and quality factor functions. A hinged-trilinear functional form is assumed for geometrical spreading, and the regression indicated that the spectral amplitudes decay as $1/R$ at distances less than 70 km; between 70 and 140 km, spectral amplitudes increase with distance and the geometrical spreading is defined as $R^{+0.25}$; beyond 140 km, the geometrical spreading is described as $R^{-0.5}$. The increased attenuation in the transition zone is caused by postcritical reflections from Moho discontinuity (Burger *et al.*, 1987; Atkinson and Mereu, 1992). The quality factor function is expressed as $Q = 614f^{0.32}$ for frequencies greater than 1 Hz after the regression analysis. For the broader range of frequency used in this study (0.2 to 30 Hz), the Q function is described by a third-degree polynomial described as $\log Q(f) = 2.898 - 0.464 \log f + 1.238(\log f)^2 - 0.540(\log f)^3$.

The confidence intervals and standard errors of the estimated path attenuation parameters are evaluated using the Monte Carlo simulation technique. The result for path attenuation effect is compared with the results of Atkinson (2004) and the Samiezade-Yazd *et al.* (1997) model. This comparison is shown in Figure 9 and discussed throughout the study.

The result of this study can be used within the stochastic method to predict ground motions in the NMSZ and ENA. It also can be used, along with other path models that already exist in ENA regions, to help provide a better representation of epistemic uncertainties caused by the regional variability of path effects in ENA.

Data and Resources

Seismograms used in this study were collected as part of the ANSS Central and Eastern United States. Data can be

obtained from the Advanced National Seismic System from http://earthquake.usgs.gov/monitoring/anss/regions/mid/ (last accessed April 2010).

Acknowledgments

We wish to thank many people who contributed data, information, or criticisms. In particular, we wish to thank Mitch Withers for providing broadband seismograms and Heather DeShon and Christy Chiu for helping in data processing. We thank Gail Atkinson for her insightful and constructive guidance. We also thank Gail Atkinson and an anonymous reviewer for their thorough comments and suggestions, which greatly helped improve the article.

References

Allen, T. I., and G. M. Atkinson (2007). Comparison of earthquake source spectra and attenuation in eastern North America and southeastern Australia, *Bull. Seismol. Soc. Am.* **97,** 1350–1354.

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

Atkinson, G., and D. Boore (2006). Ground motion prediction equations for earthquakes in eastern North America, *Bull. Seismol. Soc. Am.* **96,** 2181–2205.

Atkinson, G., and J. Cassidy (2000). Integrated use of seismograph and strong motion data to determine soil amplification in the Fraser Delta: Results from the Duvall and Georgia Strait earthquakes, *Bull. Seismol. Soc. Am.* **90,** 1028–1040.

Atkinson, G., and R. Mereu (1992). The shape of ground motion attenuation curves in southeastern Canada, *Bull. Seismol. Soc. Am.* **82,** 2014–2031.

Atkinson, G. M. (2004). Empirical attenuation of ground motion spectral amplitudes in southeastern Canada and the northeastern United States, *Bull. Seismol. Soc. Am.* **94,** 1079–1095.

Atkinson, G. M., and D. M. Boore (1995). New ground motion relations for eastern North America, *Bull. Seismol. Soc. Am.* **85,** 17–30.

Atkinson, G. M., and D. M. Boore (1998). Evaluation of models for earthquake source spectra in eastern North America, *Bull. Seismol. Soc. Am.* **88,** 917–934.

Boore, D. M. (2003). Prediction of ground motion using the stochastic method, *Pure Appl. Geophys.* **160,** 635–676.

Brune, J. (1970). Tectonic stress and the spectra of seismic shear waves, *J. Geophys. Res.* **75,** 4997–5009.

Brune, J. (1971). Correction: Tectonic stress and the spectra of seismic shear waves, *J. Geophys. Res.* **76,** 5002.

Burger, R., P. Somerville, J. Barker, R. Herrmann, and D. Helmberger (1987). The effect of crustal structure on strong ground motion attenuation relations in eastern North America, *Bull. Seismol. Soc. Am.* **77,** 420–439.

Efron, B., and R. Tibshirani (1993). *An Introduction to the Bootstrap*, Chapman & Hall, New York.

Frankel, A., C. Mueller, T. Barnhard, D. Perkins, E. Leyendecker, N. Dickman, S. Hanson, and M. Hopper (1996). National seismic hazard maps: Documentation June 1996, *U.S. Geol. Surv. Open–File Rept.* 96–532.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Massachusetts.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, Michigan.

Lermo, J., and F. Chávez-García (1993). Site effect evaluation using spectral ratios with only one station, *Bull. Seismol. Soc. Am.* **83,** 1574–1594.

Nakamura, Y. (1989). A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface, *Railway Technical Research Institute, Quarterly Report* **30,** 25–33.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2007). *Numerical Recipes*, Cambridge University Press, Cambridge.

Samiezade-Yazd, M., R. B. Herrmann, L. Malagnini, and W. Liu (1997). A regional comparison of vertical ground motion in North America, 1434-94-G-2403, Saint Louis University, St. Louis, Missouri, www.eas.slu.edu/People/RBHerrmann/GroundMotion.

Sonley, E., and G. M. Atkinson (2006). Path-specific attenuation in eastern Canada, *Bull. Seismol. Soc. Am.* **96,** 1375–1382.

Toro, G. R., N. A. Abrahamson, and J. F. Schneider (1997). Model of strong ground motions from earthquakes in central and eastern North America: Best estimated and uncertainties, *Seis. Res. Lett.* **68,** 41–57.

Welch, P. D. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoust.* **AU-15,** 70–73.

Department of Civil Engineering
The University of Memphis
Memphis, Tennessee 38152

Manuscript received 3 August 2009