BROADBAND NONPARAMETRIC GROUND-MOTION EVALUATION OF HORIZONTAL SHEAR-WAVE FOURIER SPECTRA

Kwan-Hee Yun¹), Dong-Hee Park¹), Jeong-Moon Seo²)
1) Korea Electric Power Research Institute, Korea 2) Korea Atomic Energy Research Institute, Korea

Abstract

The nonparametric method was employed to obtain detailed broadband attenuation characteristics of horizontal Fourier spectra based on the large spectral data set for small earthquakes in Korea. Before applying to the real data set, the nonparametric method was numerically tested by using artificial data set and the bias estimated. The inverted path terms show three linear regions roughly divided by 65km and 117km from the source. Also, the use of parametrically inverted Q functions was validated over the frequency band between 3-30Hz and distance range beyond 200km while complex behavior within 100km from the source was revealed which can not be properly fitted by combination of Q functional model and geometrical spreading models. Also, the near-source path terms show attenuation steeper than $R^{-1.3}$ for broad range of frequencies. Finally, mixed wave phases were found in the inverted path terms for low frequencies at distances beyond 200km.

Introduction

Until recently, in Korea, several studies on the attenuation of shear-wave has been conducted mostly by parametric methods. In the parametric methods, the attenuation characteristic is modeled by physically based functional forms and the parameters for the models were obtained by fitting the functional to the observed data. For example, in the stochastic ground-motion model ([1]), one of the parametric methods, combination of Q model ($Q_0 f^{\eta}$) and geometrical spreading function given by a piecewise continuous series of straight lines is assumed for the path effect. Previously, $348^{0.54}$ and bilinear geometrical spreading model ($R^{-1}, R^{-0.5}$) with 50km crossover distance have been inverted as the parameters of the path function of the stochastic ground-motion model ([2]). Generally speaking, parametric method is preferred where there are small number of available recorded data for the analysis and has the advantage of providing physical insights, but the model itself should be justified before applying to a local region even though the model is validated elsewhere.

Contrarily, nonparametric method is capable of separating the path effect from the source and site effects based only on the observed data without assuming any functional forms and show the undistorted nature of path effect. Furthermore, it can also provide additional information that cannot be identified by the parametric method and can be used as a reference to validate the results of the parametric methods. However, it needs relatively large amount of good quality data for the analysis in the regional distances.
In this paper, nonparametric evaluation of ground motion attenuation is conducted over the southern Korean Peninsula using a large amount of earthquake database accumulated after the modern earthquake observation stations operated mainly since 1999 ([3]). Also since we already performed parametric study of ground motion, systematic comparison between the results of two methods will be discussed.

Analysis method

The logarithm of the observed Fourier spectra \( Y = \log(Amp(f)) \) at a frequency \( f \) is the linear summation of the log of excitation term \( E(f) \), propagation term \( D(R,f) \), and site term \( Site(f) \) and the ground motion regression model can be written for a distance \( R \) as

\[
Y = \log(Amp(f)) = E(f) + Site(f) + D(R,f)
\]  
(1)

The nonparametric method is to separate the path term by removing the excitation and site terms from the observed Fourier spectra according to the simple formula (1). In equation (1), \( D(R,f) \) can be parameterized as a piecewise linear function with numerous nodes for a central frequency \( (fc) \), and a linear interpolation is taken between adjacent nodes.

\[
D(R,fc) = \sum_{k=0}^{n} L_k(R)D_k(fc)
\]  
(2)

In equation (2), \( L_k(R) \) is the linear interpolation function and \( D_k(fc) \) is the value of path term at the distance of node \( k \). In this study, instead of commonly imposed constraint that the path term should be zero at an arbitrary reference distance, a dummy parameter \( F(fc) \) for each frequency was introduced that is considered to be constant for all the earthquake events. As a result, \( E(fc) \) in equation (1) is expressed in terms of the summation of event-dependent \( E'(fc) \) and event-independent \( F(fc) \) and a regression model of equation (3) is formulated by combining equation (1) and (2). The equation (3) can be formed into a large matrix along with the data and excitation, site, and propagation terms can be inverted.

\[
Y(fc)_{ij} = \left[ E'(fc)_i + F(fc) \right] + Site(f)_j + \sum_{k=0}^{n} L_k(R)D_k(fc) (i, j = \text{index for excitation and site})
\]  
(3)

For the regression, a constraint that the sum of the site terms is forced to be zero is imposed and a further constraint of minimum roughness is forced on the path term \( D(R,f) \). When regressing the equation (3), it is important to take into account the possible bias due to the nonuniform distribution of the recorded data according to the stations, distance and earthquake magnitude. In this study, the possible bias is investigated by applying the nonparametric method to artificial data set generated by using known path, site, and source parameters calibrated to the real data set.

Finally, since no assumption was made on the reference distance, normalization of the regression results to a distance is needed so as to give physical meaning.

Seismic Data and Processing

For nonparametric evaluation of the ground-motion, Fourier acceleration spectral data set of
S-wave trains is compiled for the shallow regional earthquakes occurring around the Korean Peninsula. The earthquake data set consists of 253 earthquakes with 6,289 seismograms mostly since 1996 dated back to 1992 and is being updated for renewal of inversion each time earthquake occurs. Most of the data come from the earthquakes of magnitude less than 4.0. There are several seismic networks available for digital earthquake data within and outskirts of the Korean Peninsula. Major seismic networks are operated by Korea Meteorological Agency (KMA), Korea Institute of Geology and Mining (KIGAM), Korea Electric Power Research Institute (KEPRI), and Korea Institute of Nuclear Safety (KINS).

Seismic instrumentation is optimized for recording frequent micro-to-small earthquakes but also the rare strong ground motions without clipping with the system of 140-dB dynamic range. Many of the stations are equipped with one set of Kinematics ES-T accelerometer and seismometer. Seismometers are typed into Streckeisen STS-2 broadband sensor and short-period sensors such as Mark-Rand JC-V100 and Kinematics SS-1 with the natural frequency of 1Hz. Sensitivity of these sensors are calibrated to give full range of dynamic range of Quanterra 24bits recorders. In many cases, signals are stored at 100Hz after sharply downsampled by FIR filters, so the highest available frequency usually goes up to 40 Hz which is favorable for the study on the detailed attenuation structure of the region.

Most of the sensors are founded on outcrop of hard rock foundation or on excavated fresh bedrock if the soft overburden layer exists. The ray paths of the data set cover entire region of the southern part of the Korean Peninsula. Because the amplitude of Fourier spectrum is concerned the most, quality control of the seismic data was given foremost priority and took most of the time and effort in this study. Calibration status of all the seismic stations has been routinely checked each time an earthquake was reported by using various methods ([4]). Since the hypocentral distance is required to be fixed as inputs and critical for the nonparametric inversion, all of the local public reports of earthquake source parameters are carefully examined. Among these, epicenters and focal depths announced by KIGAM are given the most preference followed by KMA and ISC information resources in the following order. When the focal depths were unknown, 10km is assumed. Instrumental correction for short-period seismometer and STS-2 were perform to enlarge the available frequency band for frequency band below 1 Hz and above 30 Hz, respectively. Mainly acceleration seismograms are used for spectrum calculation, low frequency spectrum supplemented by velocity data.

The data segment for spectrum evaluation was selected by a time window of varying span depending on the epicenter with the onset time manually picked to include main wave energy packets. The length of the window is designed based upon the crustal velocity structure by S.K. Kim ([5]) to exclude Rayleigh wave components. For the distant records and when the recording time is available from the header, automatic time window according to the group velocities of 3.6km/sec and 2.6km/sec is used to include the S wave motions mainly composed of Lg waves. Fourier acceleration spectral amplitudes for two horizontal components were calculated with the frequency interval of 0.049Hz after cosine tapering with 5% at each end and vector summed. Spectrums are used of frequency range with more than three to four S/N ratio. Smoothing process was not applied to the original data set with concern that it would smooth out the detailed shape of the inverted path terms that could be revealed without.

Numerical Test

Before applying to the real data set, for the purpose of validation, the nonparametric method was applied to artificial data set generated by using the stochastic ground-motion model parameters calibrated to the observed data set. So the artificial data set is quasi-replica of the
The artificial data set consists of the synthetic spectra calculated by using the stochastic
ground-motion model parameters at the same frequencies of the observed data set. For nonparametric ground motion evaluation on the artificial data set, we choose six frequency bands named as fq1, fq3, fq5, fq10, fq20, and fq30 listed in Table 1 among 80 evenly spaced frequency bands from 0.5 to 40Hz in linear scale. The width of each band is approximately 0.5 Hz. As node points of the path terms, 17 values of distance (in km) - 10.0, 15.86, 21.54, 31.62, 46.42, 68.13, 100.0, 130.0, 160.0, 190.0, 220.0, 250.0, 280.0, 310.0, 340.0, 370.0, 400.0 - were selected. Throughout this paper, reference distance of 21.54km is used and the path terms are corrected for the 1/R geometrical spreading attenuation to investigate the detailed shape of the path terms.

A criterion was used to screen the data set according to the minimum number of recorded stations and spectral data in each frequency band. This screening was to preclude the possible bias of the nonparametric inversion results due to the nonuniform distribution of data according to the distance and magnitudes and seismic stations. Table 1 is the summary of the data eventually used for the analysis.

Table 1. Data information used for nonparametric inversion of path terms. The earthquakes were selected by the criteria that # of seismic stations is greater than 10 and # of spectral data greater than 100.

<table>
<thead>
<tr>
<th>ID</th>
<th>Frequency range (Hz)</th>
<th># of earthquake</th>
<th># of stations</th>
<th># of spectral data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fq1</td>
<td>0.99 - 1.49</td>
<td>109</td>
<td>131</td>
<td>34382</td>
</tr>
<tr>
<td>Fq3</td>
<td>2.97 - 3.46</td>
<td>149</td>
<td>131</td>
<td>54248</td>
</tr>
<tr>
<td>Fq5</td>
<td>4.94 - 5.44</td>
<td>148</td>
<td>130</td>
<td>54503</td>
</tr>
<tr>
<td>Fq10</td>
<td>9.88 - 10.38</td>
<td>123</td>
<td>128</td>
<td>38691</td>
</tr>
<tr>
<td>Fq20</td>
<td>19.76 - 20.25</td>
<td>99</td>
<td>126</td>
<td>27153</td>
</tr>
<tr>
<td>Fq30</td>
<td>29.63 - 30.13</td>
<td>61</td>
<td>115</td>
<td>15503</td>
</tr>
</tbody>
</table>

Figure 2. Normalized path bias terms (a) calculated by the difference between $D_{k}^{inv}(M, SD, \kappa_0)$ and $D_{k}^{inv}(M_0, SD_0, \kappa_0)$ terms (b) according to the equation (4)
To study the bias of the nonparametrically inverted path terms, a method of evaluating the bias effect was devised by use of the parametrically inverted parameters as in equation (4).

\[ B_k = D_{k}^{\text{inv}}(M, SD, \kappa_0) - D_{k}^{\text{inv}}(M_0, SD_0, \kappa_0) \quad (B_k = 0) \quad (4) \]

Where \( k \) is the distance node index, \( D_{k}^{\text{inv}}(M, SD, \kappa_0) \) is the nonparametric inversion results applied to the generated synthetic data set by using source \( (M, SD, \kappa_0) \) parameter sets inverted based on the observed data set and equal values of the site parameter \( \kappa_0 \) for all the seismic stations. \( D_{k}^{\text{inv}}(M_0, SD_0, \kappa_0) \) is the nonparametric inversion results applied to the generated synthetic data set by using equal values of the source parameters \( (M_0, SD_0) \) for all the earthquake events. In this study, \( M_0 = (3.5, \ldots, 3.5) \) and \( SD_0 = (50, \ldots, 50) \), and \( \kappa_0 = (0.016, \ldots, 0.016) \) were used. \( B_k \) is normalized to zero at a reference distance.

Figure 2(a) is the calculated bias terms for the six frequency bands and Figure 2(b) is the example of \( D_{k}^{\text{inv}}(M, SD, \kappa_0) \) and \( D_{k}^{\text{inv}}(M_0, SD_0, \kappa_0) \) for Fq1. The values of the bias terms are negligible with the value of less than 0.02 in log unit and within the error bound of around 0.04 estimated later in this study. However it could distort the path terms in near-source distance range where the path terms differ slightly each other according to the frequencies. The inverted path terms are corrected by using the calculated bias terms before the path terms are shifted to be zero at a reference distance. For example, the nonparametric path terms of Figure 2(b) was shifted by +0.22 after bias correction for normalization.

Figure 3 shows the result of nonparametrically inverted path terms applied to the artificial data set compared with the exact path terms. The exact path terms is considered to be \( D_{k}^{\text{inv}}(M_0, SD_0, \kappa_0) \) in equation (4) normalized to a zero at a reference distance. The two results show good agreement, verifying the use of nonparametric method of this study.
Results and Discussion

The verified nonparametric method through the numerical test was applied to the observed Fourier spectra data set by using the calculated bias terms from the numerical test. Figure 4 is the distribution of the raw data according to the earthquake magnitude and distance used for nonparametric evaluation of the Fourier spectra at around 3Hz.

Figure 5(a) is the results of the inverted path terms. From this figure, we can identify the three linear segments roughly divided by 65km and 117km as in the shape of the geometrical spreading function in Figure 1(b). To estimate the uncertainty of the nonparametric regression result, the bootstrap method was implemented for one of the frequency bands (Figure 5(b)). By 1000 bootstrap simulation, we obtained average of 0.024 log standard deviation for the inverted path terms of fq3.
Figure 4. An example of raw data of Fourier acceleration spectrum between 2.97-3.46Hz according to the moment magnitude ranges. The moment magnitudes were obtained by the stochastic ground-motion inversion.

Figure 5. Nonparametric regression results (a) obtained following the same process in the numerical test and its uncertainty (b) for fq3 band estimated by 1000 bootstrap simulation.
The simple thing that can be manipulated from the nonparametric result is the parameterization of the path terms in the far distance range approximately greater than 200km where $R^{-0.5}$ geometrical attenuation of the Fourier spectra is justified and the path terms in equation (2) can be parameterized as in equation (4).

$$D(R,fc) = -\frac{\pi \cdot fc \cdot R}{\beta \cdot Q(fc)} \cdot \log(e) + 0.5 \cdot \log \left(\frac{1}{R}\right) + C \quad (5)$$

Where $\beta$ is the shear wave velocity with the value of 3.5 km/sec and $C$ is an arbitrary constant. According to equation (5), $Q$ can be calculated based on the slope of the best linear fit of the path terms for each frequency and is independent of the geometrical spreading model at shorter distances. So to find out the detailed shape of the $Q$ in the frequency domain, nonparametric regression method was conducted for more frequencies. The results of the regression are shown in Figure 6. In Figure 6(b), we can identify the frequency range roughly between 3-30 Hz for which functional form of $Q_0 f^{\eta}$ (linear in log-log scale) model is well fitted. Especially, the parametrically determined $Q$ function in Figure 1 is very close to the nonparametric path terms over the same frequency range. This observation does not necessarily mean that parametric $Q$ function is only valid only for the frequency range between 3-30Hz. Actually, later in this paper, the cause of the misfit at the low frequency range will be discussed.

Figure 6. Nonparametric regression results for 25 central frequencies (a) and $Q$s calculated by the slope of the nonparametric results assuming $R^{-0.5}$ geometrical spreading (b).
Since now that both of the parametric and nonparametric methods were conducted, we can compare the path terms by the two methods. Based on the results of the parametric method, two kinds of path terms can be derived. One is “parametric path function”, calculated result of the parametric function regarding path effect shown in Figure 1 and the other is “parametric path residuals”, regression result on the path residuals produced by removing from the observed data the calculated values of the parametric functions regarding excitation and site. The regression on the path residual is conducted in the same way with the nonparametric method employed in this study but without considering excitation and site terms in equation (3).

The comparisons of nonparametric path terms with parametric path function and parametric path residuals are displayed in Figure 7(a) and 7(b), respectively. Upon seeing the results shown in Figure 7(a), it is clear that simple path function is not capable of representing the revealed complex pattern of nonparametric path terms over the short distance range less than 100km. However similarly to the fitting results as shown in Figure 6(b), the parametric path function fits the nonparametric regression results for broad frequency range at far distances greater than 200km except for the low frequency of 1Hz at which the parametric path function is only fitted to the far distant data set. This means that inverted Q-value at 1Hz is overestimated at short distances less than 200km and lower value should be used for simulating ground-motion.

**Figure 7.** Comparison of parametric path function (solid line) and nonparametric path terms (dotted line). The parametric path residuals in (b) is the regression result of the observed data with the source and site effect removed by using the relevant functions.
The parametric path residuals in Figure 7(b) fit the nonparametric results with the degree similar to the fitting result of parametric path function in Figure 7(a) for the high frequencies equal or above 3Hz (fq3-fq30). However for the frequency below 3Hz (fq1), the parametric path residuals show better fitting result over the wide distance range, which is natural since the parametric path residuals has more characteristics of the observed data than the parametric path function.

The observation mentioned above is clearer in Figure 8(a) where the three types of path terms are compared for 1Hz (fq1) in an enlarged Y-axis scale. Notable point in the Figure 8(a) is the similar pattern of abrupt increase of the nonparametric and parametric path residuals at distances beyond 200km. It indicates that other wave phases are emerging at this distance range within the time window for spectral calculation. This explains why the values of Q calculated from the slopes of the nonparametric path terms are greater than the Q functional value for lower frequencies below 3Hz in Figure 6(b). So the lower limit of the valid frequency range where the Q functional form is justified is estimated to be lower than 3Hz for the far distance range.

In Figure 8(a), it is also to be noted that rapid attenuation of the spectral amplitude of the nonparametric and parametric residual path terms is observed at distances less than 50km of the source compared to the functional path terms that have the R^{-1.1} geometrical spreading model. And they are almost comparable to R^{-1.5} for fq1.

**Figure 8.** Detailed view of the three kinds of path terms shown in Figure 7 and comparison with various geometrical spreading models according to different slopes at near-source distances for fq1 (a) and comparison of nonparametric path terms for wide range of frequencies with the various geometrical models (b).
In Figure 8(b), attenuation characteristics of the nonparametric terms for the higher frequencies is displayed in detail at distances less than 100km. Less rapid attenuating trend but still steeper than $R^{-1.3}$ is observed for the higher frequency path terms. It is also observed in Figure 8(b) that attenuation of path terms for higher frequencies is less rapid than for lower frequencies, which is the reverse of what the inverted path function in Figure 1 predicts. This indicates the possibility that Q might change in short distance range or geometrical spreading model is different according to the frequencies based on the assumption that only single wave phase is included during the windowing process to calculate the Fourier spectral amplitudes.

To quantitatively analyze the above phenomena, we calculated the exponent values of geometrical spreading ($n$ of $R^{-n}$) according to frequencies in distance range less than 100km assuming anelastic attenuation of $217^{0.7}$ and the result is summarized in Table 2. Q function of $217^{0.7}$ is estimated as the best fit of the Qs only for frequencies between 3-30Hz based on the slopes of the nonparametric path terms at distances beyond 200km in Figure 6.

Table 2. Calculated values of exponent geometrical spreading for nonparametrically inverted path terms at distances less than 100km according to frequencies assuming anelastic attenuation of $217^{0.7}$.

<table>
<thead>
<tr>
<th>$R_{hyp}$ (km)</th>
<th>fq1</th>
<th>fq3</th>
<th>fq5</th>
<th>fq10</th>
<th>fq20</th>
<th>fq30</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.54 - 31.62</td>
<td>1.608</td>
<td>1.555</td>
<td>1.278</td>
<td>1.420</td>
<td>1.144</td>
<td>1.172</td>
<td>1.363</td>
</tr>
<tr>
<td>31.62 - 46.42</td>
<td>1.272</td>
<td>1.249</td>
<td>1.147</td>
<td>0.999</td>
<td>0.840</td>
<td>0.867</td>
<td>1.062</td>
</tr>
<tr>
<td>46.42 - 68.13</td>
<td>0.607</td>
<td>0.626</td>
<td>0.659</td>
<td>0.338</td>
<td>0.499</td>
<td>0.632</td>
<td>0.560</td>
</tr>
<tr>
<td>68.13 - 100.00</td>
<td>-0.089</td>
<td>-0.279</td>
<td>-0.139</td>
<td>0.093</td>
<td>0.289</td>
<td>0.602</td>
<td>0.080</td>
</tr>
</tbody>
</table>
Conclusion

In this study, a nonparametric inversion method was applied to a large data set of horizontal Fourier spectral amplitudes for earthquakes occurring in and around the Korean Peninsula following the stochastic ground-motion model inversion. This is an effort to reveal the broadband attenuation characteristics of the Fourier spectra without assuming any functional forms regarding source, site and path effects and to verify the parametric path functions by comparison. Especially, the nonparametric inversion method was validated by applying to the artificial data set generated with known stochastic ground-motion model parameters calibrated to the observed data set. And based on the parametric results, a method was devised to evaluate the path bias terms attributed to the nonuniform distribution of data according to the distance and magnitudes and seismic stations. Numerically tested nonparametric method was applied to the real data set and the result shows following features of the Fourier spectra in broad frequencies.

• The inverted path terms with the bias correction show three distinct linear regions roughly divided by hypocentral distances of 65km and 117km.

• The use of parametrically inverted Q functions was validated over the frequency band between 3-30Hz and distance range beyond 200km.

• Complex behavior at the near distance range less than 100km was revealed which can not be properly fitted by combination of Q functional model and geometrical spreading models.

• Mixing of more than two wave phases was found in the low frequency band within the time window for spectral calculation beyond 200km in hypocentral distances.

• Parametrically estimated Q at 1Hz is overestimated because of fitting to the spectral data at the far distance range.

• Steep attenuation comparable to R−1.3 geometrical spreading is found at distances less than 50km.

• Unresolved phenomena is found that implies possible change of Q according to distances and different geometrical spreading according to the frequencies at short distances less than 100km.

The result of this study suggests that the parametric and nonparametric method have their own limitation and should be used in complementary way to successfully simulating the earthquake ground motion. And it could also contribute to reducing the uncertainty in simulating strong-ground.

References


