Variation of teleseismic short-period waveforms at Matsushiro Seismic Array System

Mamoru Kato$^1$, Ichiro Nakanishi$^2$, and Hiroyuki Takayama$^3$

$^1$Earth Science Group, Graduate School of Human and Environmental Studies, Kyoto University, Kyoto 606-8501, Japan
$^2$Department of Geophysics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan
$^3$Meteorological Research Institute, 1-1 Nagamine, Tsukuba 305-0052, Japan

(Received March 15, 2004; Revised March 25, 2005; Accepted April 25, 2005)

We investigate the origin of variation of teleseismic P phase waveforms recorded at Matsushiro Seismic Array System. Short-period, vertical component waveforms from three source regions were analyzed separately in the frequency domain. Observed variation is mainly attributed to the site effects, which are frequency dependent and strongly depend on the incoming azimuth. Such variation is not seen for the pre-event noise spectra. The secondary wavefield that is excited by the incident P phase is the likely cause of the observed waveform variation.

Key words: Waveform analysis, seismic coda, seismic noise, seismograph network.

1. Introduction

Microearthquake networks that have been originally developed to monitor the local seismicity can also be utilized as large aperture arrays to study the velocity structure in the Earth’s mantle (e.g., Walck, 1984; Nakanishi, 1988; Vidale and Benz, 1992; Sugiyama and Nakanishi, 2001). One major advantage to use the seismic array recordings in global structural seismology is that we can achieve an efficient signal enhancement (Rost and Thomas, 2002). Such array techniques also provide estimates of the slowness vector of these phases, which are crucial in determining their origins.

When the array techniques are applied to a massive quantity of short-period recordings, it is often implicitly assumed that the observed recordings are composed only of the signal and the random noise; by stacking the array waveforms we can suppress the incoherent noise, and the signal-to-noise ratio for the target phase would improve as the number of traces becomes large (Kanesewich, 1975). One way to represent teleseismic records is to assume that an observed seismogram, $O$, is an output of a linear system, $O = S * P * R + N$, where $S$, $P$, $R$ and $N$ represent the source, propagation, and receiver characteristics and the noise, respectively, and $*$ is the convolution operator. The receiver characteristics is defined as the product of sensor response and site response. In global structural seismology, we often stack waveforms from more than one events together to enhance weak phases from deep Earth, or in other words, to enhance the propagation characteristic $P$. To ensure an effective enhancement, it is required to deconvolve source characteristics from each waveform prior to stacking, and we usually estimate the source wavelet empirically from the observed direct P waveforms. Such empirical estimation of the source wavelet, however, should be effective only when variations of receiver characteristics, $R$, are small among the array stations. This is not always guaranteed, as variations of the amplitude of short-period recordings are usually large. Among two terms in receiver characteristics, we usually know sensor response parameters and difference of the sensor responses among array stations can be corrected if necessary. Effect of the remaining site effect appears both on the phase as well as on the amplitude of the spectrum, and waveform processing in time domain, which appears to be a common practice, might not be sufficient to cancel out their variations. It would be desirable to recognize how site effects distort the waveforms before we apply array techniques to the observed wavefield, since our goal in processing teleseismic waveforms is to evaluate properties of incident wavefield.

We investigate effects of the receiver characteristics on teleseismic recordings of P phases at Matsushiro Seismic Array System (MSAS) in Japan. Despite the small diameter of MSAS, correlations among the observed waveforms are not high, which clearly shows the influence of the receiver characteristics at each station (Kato and Nakanishi, 2000). Our approach is to identify the relative receiver characteristics among the array stations in frequency domain. Site effects at MSAS stations appear to be the major contributor on waveform variation, and are stably estimated from the observed spectra, which are frequency dependent and depend also on the back azimuth strongly. On the other hand, such site effects are not seen in the pre-event noise spectra. These observations imply the local secondary wavefield that is excited by the incoming (primary) wavefield is the primary constituent of the site characteristics.

2. Method

The observed power spectrum of the $i$-th event recorded at the $j$-th station at an array, $O_{ij}(f)$, is expressed as the product of two terms, $O_{ij}(f) = S_i(f)P_j(f)$, where $S_i(f)$ and $P_j(f)$ are the spectrum of the incident wave-
field for the \(i\)-th event, and of the receiver terms at the \(j\)-th station, respectively. In the notation originally used in strong motion seismology (Andrews, 1982), \(S_i(f)\) and \(P_j(f)\) are the source and propagation terms, respectively, the latter of which include anelastic dissipation (Andrews, 1986; Iwata and Irikura, 1986, 1988). In using the array recording of teleseismic events, we cannot isolate the source characteristics from the propagation characteristics so that the incident wavefield term would include anything that is not station dependent. By taking logarithm of both sides, we can obtain a set of linear equations, \(\log O_{ij}(f) = \log S_i(f) + \log P_j(f)\), for each frequency. When \(I\) events are recorded at all of \(J\) stations, the linearized system is expressed as \(A \delta m = d + n\), where the model vector \(m\) consists of \(I + J\) unknowns, the data vector \(d\) consists of \(I \times J\) observations (knowns), and \(n\) is the noise vector. At least one singular value of \(\mathbf{A}\) is zero because of its symmetry, and the inversion for \(m\) is an underconstrained problem (e.g., Andrews, 1982; Iwata and Irikura, 1986, 1988). This implies that we cannot estimate the absolute receiver characteristics solely from the observations.

Two approaches are often used to solve this system, both of which are mathematically equivalent. One approach is to take spectral ratios. A reference station is designated in the array and the observed spectra for each event are normalized by the spectrum at the reference station to cancel out the incident wavefield terms. By solving the normalized set of the equations, \((\log O_{ij} - \log O_{iR}) = (\log P_j - \log P_R)\), where \(R\) indicates the reference station, we can obtain the relative receiver characteristics. In strong motion seismology, for example, a hard rock site is taken as the reference station to isolate the effect of the soft surface layers at other stations.

Another approach to solve this system of equations which we take in this study is the array average method. We estimate for the \(i\)-th source spectrum by taking the log mean of the \(J\) observed power spectra

\[
(\log S_i)_o = \frac{1}{J} \sum_{j=1}^{J} \log O_{ij}
\]

(i.e., the average of observed spectra over the array). Similarly, we estimate the receiver characteristics at the \(j\)-th station as the average of the residuals

\[
(\log P_j)_o = \frac{1}{I} \sum_{i=1}^{I} \log O_{ij} - (\log S_i)_o.
\]

These estimates are elements of the model vector \(\mathbf{m}\); the residual model vector \(\delta \mathbf{m} = \mathbf{m} - \mathbf{m}\) satisfies the system \(\mathbf{A} \delta \mathbf{m} = \delta \mathbf{d} + \mathbf{n}\), where \(\delta \mathbf{d} = \mathbf{d} - \mathbf{A} \mathbf{m}\). We can pursue the model perturbation vector \(\delta \mathbf{m}\) via inversion of the system if necessary. For the present study, the initial estimates \(\mathbf{m}\) explained more than 90% of the observations, and considering the uncertainty of the spectral estimates (see below), we did not pursue the model perturbation vector \(\delta \mathbf{m}\). The array average method is shown successful when the number of stations is large (Wilson and Pavlis, 2000).

We utilized the multitaper method in estimating the power spectra. This method is suited to estimate a spectrum of a short time series. A set of spectral estimates is obtained using tapers that are orthogonal to each other, whose weighted average is the final spectral estimate (Thomson, 1982; Park et al., 1987; Percival and Walden, 1993; Lees and Park, 1995). The first 7 tapers of 4-\(\pi\) prolate taper series were used in this study. An outline with the pertinent equations is presented in Park et al. (1987). We also applied a jackknife method to estimate errors in the spec-

![Fig. 1. Location of Matsushiro Seismic Array System (MSAS), and of array stations (inset). All but MAT are borehole stations. See also Table 1.](image1)

![Fig. 2. Location of the events used in this study. Diamonds, squares, and circles denote event for northeast Japan (NEJ), western Philippine Sea (PHS), and eastern Philippine Sea (MAR) paths, respectively. Radius of the outer circle is 90 degrees, and shaded area indicates region of epicentral distances between 40 and 55 degrees from MSAS (the triangle).](image2)
Table 1. Stations of MSAS.

| Code | Lat. (°N) | Lon. (°E) | Dep (m) | Elev (m) | Vp (km/s) | T (sec.) |
|------|-----------|-----------|---------|----------|-----------|----------|
| TKM  | 36.560    | 138.242   | 46.7    | 598      | 6.1       | 1.00     |
| JZO  | 36.522    | 138.250   | 43.3    | 616      | 3.6       | 1.02     |
| SGD  | 36.539    | 138.319   | 75.4    | 1231     | 5.3       | 0.98     |
| DIR  | 36.482    | 138.298   | 53.0    | 838      | 4.8       | 1.02     |
| IRK  | 36.468    | 138.254   | 43.3    | 761      | 4.1       | 0.97     |
| MAT  | 36.543    | 138.207   | 0.0     | 406      | 6.0       | 1.01     |
| WDR  | 36.487    | 138.214   | 65.0    | 961      | 4.3       | 1.01     |

Dep: Depth of sensor measured from surface, Elev: Elevation of sensor, Vp: Velocity of boring sample, T: Characteristic period of vertical component sensor.

Fig. 3. Examples of the waveforms of an Alaskan event (NEJ path) recorded at MSAS. Unfiltered traces are aligned at hand-picked arrival of P phase, and normalized with a common amplitude scale. Numbers noted below station code for each waveform is the epicentral distances. Shaded areas portions of waveforms used for spectral analysis of the P phases and the pre-event noise.

3. Observations

MSAS, operated since 1983 by Japan Meteorological Agency (JMA) (Osada et al., 1984), is located in central Japan (Fig. 1), and is the main contributor to the global CTBT monitoring network in Japan. MSAS consists of seven stations and the diameter of the array is approximately 10 kilometers (Fig. 1). The array is located in a highland area, and the elevations of the sensors vary within the array (Table 1). At station MAT the sensor is located in the horizontal vault that houses the broadband sensors of GSN station MAJO, and at the other six stations they are installed in boreholes at depths of 40 to 70 meters (Table 1). Each station is equipped with three component velocity-type sensors of the same type, which are of the standard short-period type, with the damping factor of 0.7 and the characteristic frequency of 1 Hz (Table 1). Calibration signals are monitored in the routine operation and no severe change of these sensor parameters has been detected during our target period. Sampling frequency is 80 Hz.

Reflecting the tectonic history in this area, the surface geology at stations is heterogeneous. The compressional velocity of the drilling samples varies from 3.6 km/sec to 6.1 km/sec (Table 1), and significant lateral velocity changes, in particular irregular subsurface layers, is detected by the refraction surveys (Asano et al., 1969). Locations of teleseismic events determined solely by MSAS data with array seismology methodology often disagree with those determined by the national network of JMA or US Geological Survey (USGS) even after the observed travel times are empirically corrected for the shallow structure (Nagai and Kashiwabara, 1985; Wakui and Nishiwaki, 1986; Maki et al., 1987; Kobayashi et al., 1993).

Receiver characteristics is the product of receiver and site terms. In a narrow sense the site effect is effects of surface soft layer, as is used in strong motion seismology. Our definition of the site in this report is that everything that is station dependent and causes variation of waveforms at array stations. MSAS stations are installed at different
Fig. 4. Estimated spectra for waveforms in Fig. 3. Thick and dash lines denote estimates for P phase and pre-event noise, respectively. Shaded areas indicate ±1σ error bound for the P phase spectra, estimated by jackknife method. The signal-to-noise ratios are sufficiently high, and the variances are adequately small below 5 Hz.

buried depths and amplification at the surface for the vault station and superposition of the incoming and surface reflected waves for boreholes should contribute to differences of the waveforms. These factors should be peculiar to each station and would affect the results of the array processing, and thus are treated as elements of the site characteristics in the following analysis.

We analyzed the vertical component P wave signals from the shallow events whose epicentral distances from MSAS are larger than 40 degrees (Fig. 2); waveforms from these events are less susceptible to triplicated wave propagation in the upper mantle transition zone. Events in three separate regions occurring between 1984 and 1995 were selected, Alaskan events for northeastern Japan (NEJ) path, Indonesian events for western Philippine Sea (PHS) path, and New Guinea events for eastern Philippine Sea (MAR) path (Fig. 2). Numbers of events were 20, 17, and 20, respectively. These events were carefully chosen so that the ray paths of the P phase were similar and the propagation effects were equalized within each group. The records show adequately high signal-to-noise ratio in the frequency bands that are often used in high-frequency studies of the global structural seismology (~1 Hz), as is shown below. We set the length of the analysis time window to 6 seconds, which should be long enough to cover the source process time for the events of these sizes (mb 5.0–6.1). Small change of the window length do not change our conclusions qualitatively.

Figure 3 shows examples of observed P waveforms for an Alaskan event, along with the calculated spectra (Fig. 4). A high correlation among seven waveforms is seen only at the first half cycle at the arrival of the P phase. Variation of the waveforms becomes large in the following coda part, and
the correlation is extremely low. For this event the maximum amplitude in the time window for the P phase at station MAT is almost twice as large as that at IRK. Such large contrast of the amplitude between MAT and IRK is a common feature in the NEJ records. A large amplification at MAT is seen over the entire frequency range in the power spectrum (Fig. 4). The spectra of the pre-event noise waveforms vary little in these frequencies and the signal-to-noise ratio is adequately high below 5 Hz for the present dataset.

The receiver characteristics apparently have an azimuthal dependence. For an Indonesian event for the PHS path, the amplitude at MAT is close to the array average (Figs. 5 and 6), which is different from the records for the NEJ path. Similar amplitude variation among the array stations is repeated seen in the same group, and these variations are unlikely to be due to the complex source process or the wavefield influenced by the structure in the source region. These observations also implies that differences in sensor response parameters do not solely cause the variation of observed waveforms. Data from three regions therefore were separately treated in the subsequent analysis.

We did not apply any prior constraints on the shape of the spectra of incoming waveforms, and the data for each frequency band was treated independently. The frequency range for the analysis was set to 0.2 to 5 Hz. The upper limit is determined by the signal-to-noise ratio. The lower limit is set below the frequency band that the sensor response is flat (2 Hz and up), but this frequency range is often used in recent studies of global structural seismology, and thus of our interest. Results were plotted as the normalized values relative to station JZO, which is located at the center of the array, for the purpose of comparison. We obtain similar results for cases that sensor responses are corrected beforehand using the published parameters and are cases they are not corrected, as difference of sensor responses at stations are negligible (these tabulated sensor response parameters explain at most one third of the observed variation). In the subsequent figures, we display results for cases that sensor responses are not corrected.

4. Results and Discussion

Estimated receiver characteristics are both functions of frequency and azimuth (Fig. 7). For NEJ path, the amplification at MAT is higher than those at the other stations, and the amplification at IRK the lowest, both of which agree with the observation in the time domain. Although the azimuths from MSAS to the source regions are only slightly different for events of the PHS and the MAR paths, results for these paths do not agree well, the difference being largest at MAT and SGD. The relatively high amplification at high frequencies at all stations would imply that the amplification at the JZO, the reference station, is lower than the array average in this frequency band, but no other common pattern is seen in the three independent results. These results indicate that the incoming azimuth is one of the controlling parameters of the receiver effects. This supports our findings with waveforms, that is, sensors are not solely responsible for variation of the observed waveforms.

We adopted the jackknife approach to estimate uncertainties of these estimates. \( I \) subsets of data, in each of which the spectra for the \( i \)-th event are discarded \((i = 1 \sim I)\), were made from the original dataset, and each of \( I \) subsets was analyzed for the receiver responses in the same manner. Statistics of the estimates for these subsets would provide an appropriate estimate of the error of the results with the entire dataset. Such an empirical approach is suited to test, for example, whether the signal-to-noise ratio is adequate to infer the result or whether results were influenced by a particular observation. Results for NEJ path are shown in Fig. 8 as the example, and similar results are obtained for the other paths as well. Variances of the site characteristics...
for the subsets are very small, and the major features of the results appear to be robust. The stability of the results suggests that the observed variations of the spectra are unlikely to be caused by chance due to the large noise or an unknown random wavefield.

The site characteristics are most likely due to the secondary wavefield that is excited by the incident direct P phase. When the power spectra of the pre-event noise waveforms of the same length, 6 seconds, were analyzed in the same manner, we find that variation of the site characteristics are similar among three independent datasets (Fig. 9). In the first 6 second from the onset of direct P phase, energy of P phase and immediate coda is contained. Stability of the results imply that a deterministic wavefield in the P coda is included in the time window, which is likely to be excited at the onset of the P phase.

We should note that results for pre-event noise have similar shapes for three datasets at all stations; relative responses are flat below approximately 1.2 Hz, and gradually increase until 2 or 3 Hz. For conventional short-period sensors, small changes in sensor parameters produce large changes in amplitude and phase responses at the vicinity of their characteristic frequency, which is for our case 1 Hz. Though no such change has been reported, there is a possibility that sensor response parameters might have small changes from the published values, and are partially responsible for the observed spectral variations (Fig. 9). Knowledge on properties of noises at these stations would be required to further verify this idea. This approach could be used to investigate properties of array stations when sensor responses are inaccurately known.

An ideal situation for a seismic array is that the site re-
Fig. 7. Receiver characteristics derived for the NEJ path (solid line), the PHS path (dashed line), and the MAR path (dotted line). Results are scaled relative to JZO. We obtain identical results with the spectral ratio method (not shown here).

Fig. 8. Receiver response spectra that are inverted from 20 subsets of events for NEJ corridor, with a jackknife approach. These estimates agree well with each others as well as with our estimates from full dataset (Fig. 7), suggesting small variances of the estimates.

Fig. 9. Receiver response spectra that are inverted from pre-event noise spectra. Solid, dashed, and dotted lines denote estimates for NEJ, PHS, and MAR paths, respectively.
sponses are similar at all stations; in such case scaling the amplitude in the time domain would effectively equalize the site effects so that a linear stacking will favorably enhance the incoming wavefield. This is probably not the case for MSAS, as the site responses for array stations are strongly site dependent and frequency dependent. Caution must be taken in detecting weak later phases when the array seismograms are contaminated by the site effects. Onsets of phases should be clear when the high-frequency energy is abundant, and stations with relative high amplifications at high frequencies might strongly influence appearance of weak phases on the stacked trace. Variation of site responses appears to be stronger in high frequencies, and it is desirable to have an empirical estimate on how severe local effects are before analyzing array data (e.g., Der et al., 1987; Der et al., 1990). Robustness of results from array processing should be tested with statistical means.

On the other hand, we demonstrate that the site effects are likely to represent the local secondary wavefield that is deterministically excited by the primary P wavefield. Stability of our estimates of the site characteristics in the frequency domain implies that the removal of the site effects in the time domain could be realistically possible. Such attempts have been made previously to detect converted phase abundance in the time domain would effectively equalize the incoming wavefield. This is probably not the case for MSAS, as the site responses for array stations are strongly site dependent and frequency dependent. Caution must be taken in detecting weak later phases when the array seismograms are contaminated by the site effects. Onsets of phases should be clear when the high-frequency energy is abundant, and stations with relative high amplifications at high frequencies might strongly influence appearance of weak phases on the stacked trace. Variation of site responses appears to be stronger in high frequencies, and it is desirable to have an empirical estimate on how severe local effects are before analyzing array data (e.g., Der et al., 1987; Der et al., 1990). Robustness of results from array processing should be tested with statistical means.

On the other hand, we demonstrate that the site effects are likely to represent the local secondary wavefield that is deterministically excited by the primary P wavefield. Stability of our estimates of the site characteristics in the frequency domain implies that the removal of the site effects in the time domain could be realistically possible. Such attempts have been made previously to detect converted phase abundance in the time domain would effectively equalize the incoming wavefield. This is probably not the case for MSAS, as the site responses for array stations are strongly site dependent and frequency dependent. Caution must be taken in detecting weak later phases when the array seismograms are contaminated by the site effects. Onsets of phases should be clear when the high-frequency energy is abundant, and stations with relative high amplifications at high frequencies might strongly influence appearance of weak phases on the stacked trace. Variation of site responses appears to be stronger in high frequencies, and it is desirable to have an empirical estimate on how severe local effects are before analyzing array data (e.g., Der et al., 1987; Der et al., 1990). Robustness of results from array processing should be tested with statistical means.

5. Conclusions

We investigated receiver characteristics at MSAS in central Japan. Multiple taper was used for the spectral estimate, results of which were used to estimate the site characteristics using the array average method. Results of three subsets of data indicate that site effect is strong, frequency dependent, and azimuthal dependent. Errors that were estimated with an empirical approach are small, indicating the site characteristics represent the secondary wavefield that is deterministically excited by the primary wavefield. Lack of such features in the pre-event noises suggests that these site effects are likely to reflect the local wavefield that is excited by the direct P phase in the vicinity of the stations.

Acknowledgments. We deeply appreciate efforts by staff at Matsushiro Observatory to keep MSAS in operation for years. Akio Kobayashi, Akira Yamada, Takehi Isse, and Noriaki Serizawa assisted data retrieval from the archive. We would like to thank Akio Kobayashi, and Tomotaka Iwata for fruitful discussions. Constructive comments by Gary Pavlis on the earlier draft improved the manuscript, and we are deeply appreciated. Comments by Andreas Rietbrock, anonymous reviewer, and the Editor Ichichi Fukuyama are gratefully acknowledged. Figures are drawn with GMT software package (Wessel and Smith, 1991), MEXST and JSPS, and Kurata Memorial Hitachi Science and Technology Foundation provided financial support to MK and IN, respectively.

References

Andrews, D. J., Separation of source and propagation spectra of seven Mammoth Lakes aftershocks, USGS Open File Rep., 82-591, 437, 1982.

Andrews, D. J., Objective determination of source parameters and similarity of earthquakes of different size, in Earthquake Source Mechanics (Maurice Ewing Series 6), edited by S. Das, J. Boatwright, and C. H. Scholz, Am. Geophys. Union, Washington DC, pp. 259–267, 1986.

Asano, S., S. Kubota, H. Okada, M. Nogoshi, H. Suzuki, K. Ichikawa, and H. Watanabe, Explosion seismic studies of the underground structure in the Matsushiro Earthquake swarm area, J. Phys. Earth, 17, 77–90, 1969.

Bear, L. K. and G. L. Pavlis, Estimation of slowness vectors and their uncertainties using multi-wavelet seismic array processing Bull. Seis. Soc. Am., 87, 755–769, 1997.

Der, Z. A., R. H. Shumway, and A. C. Lees, Multi-channel deconvolution of P waves at seismic arrays, Bull. Seis. Soc. Am., 77, 195–211, 1987.

Der, Z. A., M. R. Hirano, and R. W. Shumway, Coherent processing of regional signals at small seismic arrays, Bull. Seis. Soc. Am., 80, 2161–2176, 1990.

Iwata, T. and K. Irikura, Separation of source, propagation and site effects from observed S-waves, Zisin, Ser. II, 39, 579–593, 1986 (in Japanese with English abstract).

Iwata, T. and K. Irikura, Source parameters of the 1983 Japan Sea Earthquake sequence, J. Phys. Earth, 36, 155–184, 1988.

Kanepschewich, E. R., Time series analysis in geophysics, The University of Alberta Press, Edmonton, 1975.

Kato, M. and I. Nakanishi, Upper mantle velocity structure in the western Pacific rim estimated from short-period recordings at Matsushiro Seismic Array System, Earth Planet Space, 52, 459–466, 2000.

Kobayashi, A., N. Mikami, and Y. Ishikawa, Epicentral correction for the hypocenter of Matsushiro Seismic Array System, Tech. Rep. Seis. Obs., Japan Meteorol. Agency, 12, 1–9, 1993 (in Japanese).

Kobayashi, Akira Yamada, Takehi Isse, and Noriaki Serizawa assisted data retrieval from the archive. We would like to thank Akio Kobayashi, and Tomotaka Iwata for fruitful discussions. Constructive comments by Gary Pavlis on the earlier draft improved the manuscript, and we are deeply appreciated. Comments by Andreas Rietbrock, anonymous reviewer, and the Editor Ichichi Fukuyama are gratefully acknowledged. Figures are drawn with GMT software package (Wessel and Smith, 1991), MEXST and JSPS, and Kurata Memorial Hitachi Science and Technology Foundation provided financial support to MK and IN, respectively.

References

Andrews, D. J., Separation of source and propagation spectra of seven Mammoth Lakes aftershocks, USGS Open File Rep., 82-591, 437, 1982.

Andrews, D. J., Objective determination of source parameters and similarity of earthquakes of different size, in Earthquake Source Mechanics (Maurice Ewing Series 6), edited by S. Das, J. Boatwright, and C. H. Scholz, Am. Geophys. Union, Washington DC, pp. 259–267, 1986.