Quantifying the Spatial Distribution of Site Response by Use of the Yokohama High-Density Strong-Motion Network

by Kenichi Tsuda, Ralph J. Archuleta, and Kazuki Koketsu

Abstract We have estimated the spatial variation of site response using threecomponent accelerograms from the Yokohama High-Density Strong Motion network: 150 surface and 9 borehole accelerometers located in a 20 × 30 km² area. The site response is the average from 28 earthquakes; 16 were recorded at all borehole stations. We inverted data recorded at the boreholes to determine (1) source parameters (seismic moment M_0 and corner frequency f_c) for each event, (2) Q(f) for the path, and (3) frequency-dependent site factors for the boreholes. The inversion scheme is independent of a reference station, that is, there are no *a priori* constraints about site response. For events with depths greater than 30 km, we find Q(f) almost independent of frequency, $Q(f) = 285f^{0.06}$; for the four shallow events that were analyzed, $Q(f) = 70 f^{0.28}$.

Having determined M_0, f_c , and Q(f) using the borehole data, we computed the site response for each of the 150 surface stations by dividing the observed surface spectrum by the path-modified source spectrum. We then averaged this ratio for all 24 events. To quantify the variability of the site response between sites we introduced a site-response ratio that depends on station separation and frequency. From this we computed cumulative probability functions for different frequencies for two cases: (1) the average Vs30 ratio is most similar for two sites with the same separation distance and (2) the average Vs30 ratio is least similar. When two sites have the most similar Vs30 ratio, these functions indicate that 95% of the time a frequencydependent site response ($f \le 1$ Hz) can be predicted within a factor of two for a site up to 5 km from a site with a known site response, and 90% of the time the unknown response is within a factor of three for frequencies 1-10 Hz. When two sites have the least similar average Vs30 ratio, the cumulative probability functions have similar shapes for all frequencies above 0.5 Hz; for frequencies up to 10 Hz the site amplification can be predicted within a factor of three for 80% of the station pairs separated by as much as 5 km.

Introduction

The effects of local site conditions on earthquake ground motions were documented as far back as the 1906 earthquake (Reid, 1910). To quantify site effects Borcherdt (1970) introduced the spectral ratio approach by taking the ratio of the Fourier amplitude spectrum of a soil site to a rock site. Borcherdt and Gibbs (1976) extended this approach for a wide range of site conditions and concluded that a site-response map "provides a crude form of seismic zonation for the region." Because of the inhomogeneous nature of the near-surface geology, every earthquake produces a spatial variation of ground motion. Because of the deployment of accelerometers, earthquakes such as the 1994 Northridge and 1995 Hyogo-ken Nanbu (Kobe) have clearly reinforced the role that site effects play in damaging ground motion (e.g., Kawase, 1996; Bonilla *et al.*, 1997; Hartzell *et al.*, 1997). One of the primary engineering questions is whether *a priori* knowledge of the site response can reduce the uncertainty in predicting ground motion from future events. Stewart and Baturay (2000, 2001) show that ground-response analysis can improve the accuracy of ground-motion predictions relative to attenuation relations for soil sites. Field and the Southern California Earthquake Center (SCEC) Phase III Working Group (2000) found that the average shear-wave velocity in the upper 30 m delineated different site-amplification factors. Likewise, they found that the depth to basin was highly correlated to amplification but cautioned that the depth of the basin might be substituting for some other site attribute. One approach to understanding

the spatial distribution of site response is to consider the coherency of seismic waves over a given distance (e.g., Harichandran, 1988; Abrahamson *et al.*, 1991; Kiureghian, 1996; Zerva and Zhang, 1997; Zerva, 2000; Zerva and Beck, 2003). These studies have demonstrated how phase differences in the time history can create differences in the site response as a function of distance between sites. A complementary and basic question is how the amplitude of the site response itself varies from site to site. The additional complication exists that the ground response might vary from earthquake to earthquake. In practice, we need to separate effects that are event-to-event from those that are site-to-site. Quantifying the spatial variation of site response from eventto-event and site-to-site requires multiple events recorded on a dense set of stations.

Attempts have been made to evaluate spatial variation of seismic ground motions by using seismic arrays, but none have had the quality of data or the extensive auxiliary data of the Yokohama array (Tsuboi et al., 2001). After the 1994 Northridge earthquake SCEC deployed a "quasi-dense" 20element array (Archuleta, 1994) in San Fernando Valley. Hough and Field (1996) and Field and Hough (1997) analyzed these data looking for the spatial variability of the site amplification. Their data consisted of 20 local events (hypocentral distances between 2 and 22 km) with magnitudes between 2.4 and 4.0. Only 11 events were recorded on 10 or more stations. To examine the variation in the site response they used the method of Abrahamson and Sykora (1993) in which the 5% damped acceleration-response spectrum at a given site is divided by the average response spectrum of the entire array. Field and Hough (1997) found that 95% of the response spectra are within a factor of 2.3 of the array average at 0.5 km and increasing to 4.2 at 5.0 km. They could not resolve any dependence on frequency.

Fletcher et al. (2003) examined the site response across an array in the Santa Clara Valley ($\sim 25 \times 25 \text{ km}^2$) with station spacing on the order of 2-3 km. Using data from 14 events, they simultaneously inverted 876 records to determine the source parameters, a κ to correct for near-surface attenuation and an attenuation parameter Q_0 for the whole path. Having determined the spectral parameters of the source model (as modified by the path effects), the ratio between the observed spectrum and the spectrum predicted is the site response. This approach is similar to the approach we use for the Yokohama array. However, because of the rather large station separation there was no attempt to look at site-to-site variation. Fletcher et al. (2003) looked at the event-to-event variation by computing the variance of the site response from the 14 events. About one-half of the stations have high variability (differences greater than 1.8) for the 1- to 6-Hz range with no obvious correlation to basin depth.

The Joint Working Group on Effects of Surface Geology recently operated a dense strong-motion array (23 sites within a radius of 5 km) in Kushiro City for $1\frac{1}{2}$ years (Sasatani *et al.*, 1998). They computed the site response based on spectral ratios relative to a rock site; the rock reference site was 40 km away from the city. Sasatani *et al.* (1998) displayed the spatial variation of the site response in 3D plots of amplitude versus frequency and station separation. This provided a qualitative estimate of the spatial variation of site response, but the variation did not correlate with any physical feature other than highlands versus lowland.

Harichandran (1999) listed conditions necessary for arrays measuring the spatial variation of ground motion: (1) be useful to both engineers and seismologists, (2) record all three components of ground motions, and (3) be located in areas with high seismicity and a wide variety of site conditions. The one array that satisfies these conditions best is the Yokohama High-Density Strong Motion Network (Tsuboi *et al.*, 2001).

In this study of the Yokohama array, we first separate source, path, and site effects by inverting the S waves recorded in the boreholes for source and path parameters. Once we have determined stable source and path parameters, we compute the spectrum for all surface sites for each event. The site response is the averaged difference between the observed and predicted spectrum. Having the site response for all 150 stations, we have looked for correlations between the site response and other physical parameters such as shear-wave velocities for upper 30 m and surface geology. To quantify the spatial variation of the amplitude of the site response, we introduce a site-response correlation and a siteresponse coefficient. Both are functions of frequency and station separation. In our approach, we can determine the spatial variation of the site response as a function of both frequency and spatial separation.

Seismic Data Set

The Yokohama High-Density Strong Motion Seismograph Network began operation in 1997 for seismic-hazard mitigation (Yokohama City Office, 1998; Tsuboi *et al.*, 2001). It consists of 150 surface and 9 borehole stations; each station has three-component accelerometers with the signal digitized at 200 Hz. Each site has been logged for *P*wave and *S*-wave velocities; the logging depth varies from 12 to 100 m.

For this study, we selected 28 events that were recorded on almost all of the borehole accelerographs and most surface stations (Table 1). The energy magnitudes for the 28 events range from 4.1 to 6.1. The distribution of epicenters and hypocenters is shown in Figure 1. Many of the events are located below Yokohama and provide nearly planar wavefronts impinging on the array. Hypocentral distances to the array are approximately 50 to 130 km. The events are divided into two groups based on their focal depth—deep events are those with hypocentral depth greater than 30 km. There are 24 deep events and four shallow events: events 1, 9, 14, and 20.

For our analysis, we used the two horizontal components at each site. We selected a 10-sec time window begin-

 Table 1

 Location and Magnitudes of Earthquakes Recorded on Yokohama Array

Event	Date mmddyy (time)	Lat.* (° N)	Long.* (° E)	Depth* (km)	$M_{ m w}$	Moment* (N m)	No. Surface Stations	No. Boreholes	PGA (cm/sec ²)	Distance [†] to HD01 (km)
1	05/03/98	34.95	130 18	5	5.5	$2.3E \pm 17$	149	9	55.6	62.4
2	05/16/98	34.95	139.10	71	47	1.5E + 16	150	9‡	73.0	94.5
3	08/29/98	35.60	140.05	67	53	9.8E + 16	149	Q‡	160.9	80.0
4	11/08/98	35.60	140.05	78	47	1.4E + 16	150	9‡	109.1	83.6
5	11/28/98	35.63	140.10	67	43	3.3E + 15	144	9 [‡]	58.0	83.2
6	12/03/98	35.61	140.04	67	4.4	3.9E + 15 3.9E + 15	142	9 [‡]	40.3	79.6
7	04/25/99 (18.00)	35.52	140.30	92	44	45E + 15	142	7	26.2	109.3
8	04/25/99(21.00)	35.46	140.63	58	5.2	6.2E + 16	104	5	8.9	152.2
9	05/22/99	35.45	139.19	23	4.1	1.6E + 15	148	9	74.5	49.8
10	07/15/99	35.92	140.46	56	5.1	5.3E + 16	126	6	21.4	108.8
11	08/09/99	35.83	139.96	116	4.6	8.2E + 15	106	9‡	15.5	127.4
12	08/11/99	35.40	139.83	62	4	9.9E + 14	133	6	39.7	66.2
13	09/13/99	35.57	140.20	77	53	1.1E + 17	150	9‡	165.7	93.2
14	02/11/00	35 50	139.05	18	4.2	2.3E + 15	149	9	61.4	53.1
15	04/10/00	35.19	140.07	55	47	1.3E + 16	142	9 [‡]	30.9	75.4
16	06/03/00	35.68	140.72	48	6.1	1.3E + 18 1.7E + 18	147	9	45.7	115.7
17	08/27/00	35.00	140.14	77	4.6	9.3E + 15	126	8	24.7	96.7
18	09/29/00	35.52	139.73	90	4.8	1.8E + 16	148	Q‡	64.4	90.6
19	09/18/01	35.32	139.81	51	4 5	5.6E + 15	144	9 [‡]	101.6	53.6
20	12/18/01	35.54	139.15	30	4.5	6.1E + 15	146	9	82.3	48.3
21	03/13/03	36.07	139.87	50	4.9	2.3E + 16	133	8	20.2	89.3
22	04/08/03	35.68	140.72	48	4.8	2.3E + 16 2 1E + 16	118	5	21.5	88.2
23	05/10/03	35.86	140.07	65	47	1.4E + 16	145	8	45.1	87.9
23	05/12/03	35.00	140.14	52	5.2	7.1E + 16	149	8	45.7	77.5
25	05/17/03	35.70	140.70	53	53	1.1E + 16	146	8	27.3	116.0
26	08/18/03	35.80	140.10	71	4.8	1.12 + 10 1.9F + 16	140	0 [‡]	54.0	92.4
23	09/20/03	35.00	140.30	59	5.7	3.6E + 17	146	9	92.5	93.6
28	10/15/03	35.60	140.10	68	5.1	5.2E + 16	146	9 [‡]	130.7	83.3

*These parameters were determined by National Institute for Earth Science and Disaster Prevention (NIED), Tsukuba, Japan.

[†]This is hypocentral distance to HD01 (Fig. 2) a borehole station near the center of the array. HD01 coordinates are 35.4564° N, 139.5990° E.

^{*}Indicates the reference deep events used for the inversion.

ning 1.0 sec before the first *S*-wave arrival for all events. A cosine window taper of 0.5 sec is applied to both ends of the 10-sec record. We calculated the Fourier amplitude spectrum of each component. Our amplitude spectrum is the vector summation of the two horizontal amplitude spectra in the frequency range 0.5–30 Hz.

For events occurring in the Kanto area, Kinoshita (1992) showed that f_{max} (Hanks, 1982) is usually greater than 25 Hz; thus, the effects from f_{max} are likely to be small for our analysis. We examined the spectra of the P-wave coda (noise for our analysis) preceding the S wave at the stiffest sites where one can expect the P-wave coda to have the greatest influence. Using the records in the boreholes, the P-wave coda spectra are always less than the S-wave spectra in the frequency band we analyzed. Our determination of the site response assumes linearity. Although we cannot assert that the response was linear, the largest peak ground acceleration (PGA) at the surface for any event was 165 cm/sec² with a corresponding peak ground velocity (PGV) of 4.6 cm/sec Moreover, of 3944 recordings only 27 PGAs were greater than 80 cm/sec² and 103 PGV were greater than 3.0 cm/sec. Following Beresnev (2002), we can be reasonably sure that nonlinearity is not a factor. Our basic results, based on averaging 28 events for 150 stations, will be unaffected even if there were one or two recordings with a minor nonlinear effect.

Method

The observed ground motion (for linear response) can be expressed as a convolution of the source, path, and site. In the frequency domain, we can write this convolution as a multiplication:

$$|A(f)| = |S(f)| \operatorname{Site}(f) | R^{-1} e^{-\pi f R/Q(f)\beta}, \qquad (1)$$

where *f* is frequency, |A(f)| is the acceleration-amplitude spectrum of the recorded ground motion, |S(f)| is the source spectrum, Q(f) is the quality factor, |Site(f)| is the siteresponse amplitude, *R* is the hypocentral distance, and β is the average shear-wave velocity between source and site. To estimate the spatial variation of site response, it is necessary to separate these three factors. In general, separating each element requires constraints to avoid trade-offs between the three elements. A standard constraint is to define |Site(f)| at a rock station (e.g., Kinoshita and Ohike, 2002); or another



Figure 1. Geography of the region near Yokohama. Epicenters are plotted in map view and cross section. Filled circles, with a corresponding number, show events listed in Table 1. Open circles denote the locations of events recorded at the surface stations of Yokohama array but with no borehole data.

approach is to make a reference source spectrum |S(f)| for a specific event (Moya *et al.*, 2000; Moya and Irikura, 2003). For the Yokohama array, there are no nearby rock sites nor do we have enough information to create |S(f)| for a reference event. Because of the combination of boreholes and surface sites, we have devised an alternative method to separate the source, path, and site effects.

We assume Boatwright's (1978) representation of a ω^2 source spectrum (Brune, 1970) (equation 2). The amplitude of the source spectrum has a nonlinear dependence on the corner frequency:

$$|S(f)| = CM_0 (2\pi f)^2 f_C^2 / (f^4 + f_C^4)^{0.5}, \qquad (2)$$

where C is a constant that depends on the hypocentral dis-

tance from source to site, radiation parameter of the source, material parameters, and free surface effects; M_0 is the seismic moment; and f_c is the corner frequency (Brune, 1970, 1971). To assess the quality of our approximation of using only the hypocentral distance, rather than a spreading factor, we compared spectral amplitudes at 1.0 Hz assuming 1/Rgeometrical spreading and those computed with a frequencywavenumber algorithm using the location of the hypocenters and the velocity structure of the region (Yamazaki *et al.*, 1992). The spectral amplitude, using geometrical spreading, is within 2.5% and 6% of the *f-k* spectral amplitude for the deep and shallow events, respectively. These are small differences especially because the site response for the surface stations is derived from the 24 deep events. The path effect also has a nonlinear dependence on frequency if we assume that the attenuation parameter has a power law dependence on frequency $Q(f) = Q_0 f^{\gamma}$.

Because of the nonlinear relationship between the data and some of the parameters, we use a Heat Bath algorithm (Sen and Stofa, 1995) to invert for the four parameters: M_0 , f_c , Q_0 , and γ as well as the frequency-dependent site response. We start using the data from 12 deep events that were recorded on all nine borehole accelerometers. Initially, we assume that the borehole response is independent of frequency. Under this assumption, we invert the borehole data for 12 deep events $(M_w 4.0-5.3)$ to determine M_0, f_c, Q_0 , and y. The difference between the observed and predicted spectrum is taken as our first estimate of the borehole site response. Using this estimate of the (borehole) site response, our next step is to invert the spectrum $f \leq 1.0$ Hz for M_0 . With this estimate of the seismic moment, we then invert all of the data to find f_c , Q_0 , and γ . This produces the second estimate of the site response as the difference between the predicted spectrum based on the current values of M_0, f_c, Q_0, f_c

and γ and the observed spectrum. We iterate this procedure until the residuals between observed and predicted spectra are negligible at all nine borehole sites (Fig. 2b). Thus, we derive a frequency-dependent site response for each of the borehole stations as well as solve for the path parameters Q_0 and γ . Once we have stable values of the borehole siteresponse and path parameters, we invert the borehole data from each of the 24 deep events to find, M_0 (M_w based on M_0), and f_c (Table 2). We applied the same procedure for the four shallow events.

Having the source and path parameters for all of the events, we use equation (1) to predict the spectrum at all 150 surface stations for all 24 deep events. We use only the deep events because the events are more uniformly illuminated by the Yokohama array from below. The difference between the predicted spectrum and the observed spectrum is the site response |Site(f)|. We average the results from the 24 events to determine the final estimate of the site response at each surface station.



Figure 2. (a) The Yokohama High-Density Strong Motion Network with 150 surface stations (open circles) and nine borehole stations (filled circles) are mapped on the surface geology. Green denotes reclaimed soil; orange, alluvium; turquoise, Holocene soil; blue, Tertiary rock. The description of the surface geology is provided by Pasco Corporation. (b) The borehole response is shown for each iteration of the algorithm during the inversion of the 12 reference earthquakes. The final residual between the computed path-modified source spectrum and the observed borehole spectrum becomes the borehole response.

Table 2Source Parameters Derived by Inversion

2							
Event	Date (mm/dd/yy)	$M_{\rm w}$ (NIED)	M _w (This Study)	$f_{\rm C}$	Stress Drop (MPa)		
1*	05/03/98	5.5	5.1	2.4	92.0		
2	05/16/98	4.7	4.9	2.3	25.3		
3	08/29/98	5.3	5.3	1.2	17.8		
4	11/08/98	4.7	4.9	1.8	14.0		
5	11/28/98	4.3	4.4	3	11.2		
6	12/03/98	4.4	4.5	2.5	9.2		
7	04/25/99_1	4.4	4.5	3.5	27.5		
8	04/25/99_2	5.2	5.0	1.5	12.7		
9*	05/22/99	4.1	4.3	2.5	7.0		
10	07/15/99	5.1	5.0	1.1	4.0		
11	08/09/99	4.6	4.6	4.5	80.1		
12	08/11/99	4.0	4.2	2.7	4.6		
13	09/13/99	5.3	5.2	1.4	18.9		
14*	02/11/00	4.2	4.5	2.3	9.9		
15	04/10/00	4.7	4.7	1.5	3.1		
16	06/03/00	6.1	5.8	0.5	4.9		
17	08/27/00	4.6	4.7	1.8	7.8		
18	09/29/00	4.8	5.0	1.4	7.6		
19	09/18/01	4.5	4.6	1.2	1.6		
20*	12/08/01	4.5	4.6	1.8	7.5		
21	03/13/03	4.9	5.0	1	3.4		
22	04/08/03	4.8	4.6	1.7	4.9		
23	05/10/03	4.7	4.8	1.7	7.3		
24	05/12/03	5.2	5.3	0.8	3.9		
25	05/17/03	5.3	5.1	1.1	5.1		
26	08/18/03	4.8	5.0	1.4	9.0		
27	09/20/03	5.7	5.5	0.8	8.6		
28	10/15/03	5.1	5.3	1.3	24.3		

*Indicates a shallow event (focal depth \leq 30 km).

Results

Source and Path Parameters

In Figure 3a, we compare seismic moments for the deep events found by our inversion (filled circles denote the deep events and open circles denote the shallow events) with those obtained by the National Institute for Earth Science and Disaster Prevention (NIED). The two estimates of seismic-moment values are generally within a factor of two. For the 24 deep events, the average difference of $\log M_0$ between NIED and our inversion results is -0.057 with a standard deviation of 0.25. NIED estimates of seismic moment have a standard deviation of 0.15 in log moment (Kubo et al., 2002). The agreement between NIED and our estimate provides an independent check on the inversion method. Although we could have used NIED estimates of seismic moment as a constraint, our method is entirely self-consistent in deriving all of the parameters necessary to determine the site response. In Figure 3b, we plot seismic moment versus corner frequency for 24 deep events. We also draw three lines corresponding Brune stress drops of 5 MPa, 10 MPa, and 20 MPa. Overall, the seismic moment appears to scale with the inverse corner frequency cubed with a constant stress drop about 10 MPa. The stress drop varies by approximately a factor of two, a small scatter compared with other results (e.g., Hanks, 1978; Archuleta *et al.*, 1982), where the stress-drop variation is more likely a factor of 10. Plotting the stress drop as a function of focal depth (Fig. 3c) indicates that the stress drop may have some depth dependence.

By inverting borehole data from 12 deep earthquakes that were recorded on all of the borehole accelerometers, we determined $Q_0 = 285$ and $\gamma = 0.06$; for the four shallow events, we found $Q_0 = 70$ and $\gamma = 0.27$. The frequency dependence of Q(f) is strongly influenced by the spectral shape for frequencies greater than 10 Hz. In Figure 4a, we illustrate the spectral fit by plotting the observed spectra from a deep $M_{\rm w}$ 4.7 earthquake (event 18) recorded at nine borehole sites as well as the predicted spectrum. Although the seismic moment and corner frequency are specific to this earthquake, the path parameter Q(f) is the average of all 12 events. The shallow events have much stronger frequency dependence $\gamma = 0.27$ than found for the deep events. This may be caused by a greater percentage of the ray path being contained in the crust than for the deep events. A caveat is that we have only four shallow events by which we determine the path parameters. To illustrate the spectral fit for a shallow event we plot the observed spectrum for a $M_{\rm w}$ 4.2 earthquake (event 14) along with the predicted spectrum at each of the boreholes (Fig. 4b).

To check the quality of the fit between the observed and synthetic amplitude spectrum, we use the log-amplitude ratio between observed and synthetic spectrum for nine borehole stations. The averaged ratio for 24 deep and four shallow events is shown in Figure 5a and b, respectively. The ratios are close to zero demonstrating that derived parameters for the path and source are consistent with the data for the entire frequency range.

Validation of Site Response

In this study, frequency-dependent surface site response is derived from the difference between the observed spectrum and synthetic spectrum that is based on the source and path parameters found by inverting the data. The final surface site response is the average of this difference for 24 deep events. This surface site response can be evaluated by two independent methods. At the nine borehole stations, we can determine the empirical site response by dividing the surface spectrum by the borehole spectrum. In addition, we have measured the shear-wave velocities for the shallow layers at each station including the nine borehole stations (Yokohama City Office, 1998). To check our derived site response, we computed a synthetic site response at each borehole station by using the shear-wave velocity profiles at each site. We computed the theoretical transfer function (site response) at the nine borehole stations assuming constant quality factors (Q = 20) and densities for all layers (Tsuboi et al., 2001). We also computed the empirical spectral ratio between the surface spectrum and borehole spectrum for the S wave. We compare the results from all three methods for



Figure 3. Source parameters obtained by inversion are plotted. (a) Seismic moments from this study are compared with those from NIED, which uses three stations in a regional broadband network. Solid circles correspond to the deep events, and open circles to shallow events. Dashed lines indicate a factor of two. (b) Seismic moment is plotted versus corner frequency for the 24 deep events. Lines of constant stress drop (Brune, 1970, 1971) are plotted. Within a factor of two the stress drops for the 24 events are ~ 10 MPa. (c) Stress drop is plotted versus depth for the 24 deep events.



Figure 4. The spectral fit between the observed spectrum (black line) and the predicted spectrum (gray line) is shown for a deep event (no. 18) $M_{\rm w}$ 4.7 (a) and for a shallow event (no. 14) $M_{\rm w}$ 4.2 (b).

frequencies 0.5–20 Hz (Fig. 6). The surface recordings have a stronger *P*-wave coda that interferes with the *S*-wave spectrum, thereby reducing the maximum useable frequency from 30 Hz to 20 Hz. Tsuboi *et al.* (2001) showed that the observed spectral ratios of the surface to borehole nearly coincide with the theoretical transfer functions; however, they also suggested that the assumption—constant Q and constant density for all the layers—is not appropriate for all stations. Although there is general agreement among all of the estimates, there is especially good agreement between the empirical spectral ratio and our derived site response.



Figure 5. To examine the misfit between the observed and computed spectra we plot the averaged log-amplitude ratio between observed spectrum and synthetic spectrum for 24 deep events recorded at the nine borehole sites (a) and for four shallow events (b). The solid line denotes the average value and dashed lines indicate the average $\pm \sigma$. The average misfit is near 0 for all frequencies.

Spatial Variation of Site Response

With confidence that the averaged site response is accurately estimated up to at least 10 Hz, we contoured the site response at the 150 surface stations averaged for three frequency ranges (0.5–1.0 Hz, 1.0–3.0 Hz, and 3.0–10.0 Hz) (Fig. 7). We also show a contour map of the average shearwave velocity for the upper 30 m (AVS30) based on the velocity profile at each station (Yokohama City Office, 1998). Even the low-frequency range (0.5–1.0 Hz) site response within the Yokohama array shows significant variation. The general features of the contour are similar to a previous study by Tsuda (2001).

To quantify the site-response variation as a function of both frequency and distance, we introduce a function; $SRCO_{ij}(f)$ (site-response correlation function). We define this site response correlation as

$$\operatorname{SRCO}_{ij}(f) = \frac{2 |\operatorname{Site}_i(f)| |\operatorname{Site}_j(f)|}{(|\operatorname{Site}_i(f)|)^2 + (|\operatorname{Site}_j(f)|)^2}, \quad (3)$$

where $|\text{Site}_i(f)|$ is the site response for the *i*th station and $|\text{Site}_i(f)|$ is the site response for the *j*th station. $\text{SRCO}_{ij}(f)$ is normalized so that if the two site responses are identical, $\text{SRCO}_{ij}(f)$ will take on a maximum value of 1.0; it has a minimum value of zero. This function measures the similarity of the amplitude of the site response at two stations as a function of frequency.

We use this function to quantify the spatial distribution

of site response. We limit our analysis to all station pairs within 5 km of each other; initially there were 2262 station pairs. However, many of the station pairs have a separation distance of less than 5 m. Rather than trying to average SRCO_{ij}(f) for different station pairs to get a single value for a given separation distance, we consider two criteria based on AVS30. If two or more pairs of stations have the same separation within 5 m of one another, we choose the pair that has the smallest difference in the AVS30, that is, we take the station pair that has the most similar site condition. This reduces the number of stations that has the maximum difference in AVS30 for the same separation distance. This represents the site response for stations with the least similar site condition based on AVS30.

For all station pairs we compute the ratio AVS30; for any pair we take the ratio such that it is always less than one. For all 2262 pairs we find that the mean AVS30 ratio is 0.73; for the 466 pairs with most similar AVS30 ratio we have a mean = 0.88; for the 466 pairs with least similar AVS30 ratio the mean = 0.55 (Fig. 8). One can see that the distribution of the 466 pairs with the most similar AVS30 is not at all like those for the least similar AVS30. The overall distribution of the 2262 pairs is weighted toward more similar values of AVS30.

Based on the SRCO function, we have tried to predict the site response everywhere. To do this, we assume that the unknown site response at a given location can be predicted based on a known site response:



Figure 6. Spectral ratios of surface response to borehole response. The black line is the site response derived after removing the source and path parameters found by inversion. The gray line is the empirical site response based on spectral ratios of surface to borehole recordings for 24 deep events (dashed gray lines are $\pm 1\sigma$); the dotted line corresponds to the computed spectral ratio based on the shear-wave velocity profile (Yokohama City Office, 1998) using the same assumptions as Tsuboi *et al.* (2001) for the attenuation and density of each layer the dotted line is the theoretical site response based on locally determined *S*-wave velocities. Very good agreement exists between the empirical site response and ours obtained from inverting for the source and path parameters.



Figure 7. Map views for the area within the city of Yokohama. In (a) we contour the average shear wave velocity for the upper 30 m. The average site response at the surface 150 stations is computed from 24 deep earthquakes. For each frequency band we computed an average response. This average response is contoured in the panels b, c, and d. Note that each panel has a different scale. The frequency band 3.0–10.0 Hz shows a wider range of amplification than the lower-frequency bands. These figures can be compared with the surface geology shown in Figure 2.

$$\operatorname{Site}_{\operatorname{unknown}}(f) = \operatorname{SRC}(f) \cdot \operatorname{Site}_{\operatorname{known}}(f),$$
 (4)

where SRC(f) (site-response coefficient) is a frequencydependent coefficient. Substitution of equation (4) into equation (3) leads to a quadratic equation with two roots; we choose the root that is greater than one. This is simply an assumption that "unknown" site response is always larger than "known" site response. In general, when a site response is unknown, one has to be concerned about amplification with respect to a known site. The coefficient $SRC(f)^+$ is

$$\operatorname{SRC}(f)^{+} = \left(1 + \sqrt{1 - (\operatorname{SRCO}_{ij}(f))^{2}}\right) / \operatorname{SRCO}_{ij}(f).$$
(5)

The positive root of the quadratic provides the form for estimating amplification. The negative root has a value that is the inverse of the positive root, that is, $\text{SRC}(f)^- = 1/$ SRC $(f)^+$ where the minus and plus superscripts refer to the negative and positive roots of the quadratic. This selection of $\text{SRC}(f)^+$ is not the same as dividing one spectrum by another (Fig. 9). As defined $\text{SRC}(f)^+$ is always greater than 1.0 for any frequency unlike a spectral ratio.

To see what the predictive capabilities might be, we contour $SRC(f)^+$ as a function of frequency and separation distance for 466 the station pairs with the most similar site condition (Fig. 10). We have done the same for the 466 pairs with the least similar site condition; the plot is not shown although we have analyzed these data as described in the following. As seen in the histogram for the number of stations versus station separation (Fig. 10), there is limited data for station pairs with a separation distance less than 600 m. Consequently we have examined only data for separation distances of 300 m and greater. There is a gap in the data between 500 and 600 m; contours in this range are an artifact of the contouring program. The number of station pairs with a separation distance greater than 600 m is rather uniform. The value of $SRC(f)^+$ determined by pairs of stations tends to increase with increasing frequency; for example, for the frequencies less than 1.0 Hz, this value is mostly less than 1.5.

To quantify this effect, we considered two possibilities: (1) what is the average site-response coefficient taken over the entire 5-km separation?, and (2) for a given frequency and for any two stations within 5 km of each other, what is the probability that the site response at one station will exceed another by a specified factor? The results for (1) are shown in Figure 11 where we consider both the case of similar AVS30 and dissimilar AVS30. In Figure 11a, we plot the average value of SRC(f)⁺ at each frequency where the average is over all separation distances up to 5 km. In Figure 11b and c, we plot the average value of SRC(f)⁺ (averaged over frequencies up to 10 Hz) as a function of distance for both the most similar and least similar site condition, respectively. The averaged (over distance) values of the site-response coefficients increase with frequency (Fig. 11a).

To understand better how the site-response coefficient varies with distance we consider six frequencies. For each frequency we plot the cumulative probability versus the site-response coefficient (Fig. 12). We do this for both the most similar and least similar site conditions. For each frequency, the upper panel shows $SRC(f)^+$ as a function of distance; the corresponding lower panel shows the cumulative probability as a function of $SRC(f)^+$. We mark the cumulative probability where the $SRC(f)^+$ value is a factor of two (Fig. 12).

We have tried to model each cumulative probability function as a smooth parametric function assuming the following form:

$$\operatorname{Cum}(\operatorname{SRC}(f)^+, f) = 1 - \left[\exp\{-(\operatorname{SRC}(f)^+ - 1)/a\}\right].$$
(6)



Figure 8. The distribution of the ratio of AVS30 for all 2262 pairs of stations with a separation less than 5 km is plotted in the upper panel. For each pair of stations the ratio of AVS30 for the two stations is always computed with the station having the largest AVS30 being the denominator, that is, the ratio is less than one. The distribution for the 466 pairs with the most similar AVS30 ratio is shown in the middle panel and for the least similar AVS30 ratio is shown in the bottom panel.



Figure 9. Comparison of a standard spectral ratio and the site correlation coefficient $SRC(f)^+$. The spectral ratio depends on which of the two sites is in the denominator, as such the spectral ratio can be greater or less than 1.0 as a function of frequency. $SRC(f)^+$ (solid black line) is always greater than 1.0 for the entire frequency range. We also show the spectral ratio for a pair of stations Ns06 and Na10 depending on whether NS06 is in the numerator (dashed line) or in the denominator (gray line).

For each of the frequencies plotted in Figure 12, we solved for the parameter *a* by nonlinear regression. The values of *a* and residuals ($=\Sigma$ (data-Cum)²/(number of data)) are given in Table 3 for both the most similar and least similar AVS30 ratio.

First consider the situation in which the two sites have a similar AVS30. At frequencies 0.5, 0.75, and 1.0 Hz, there is a 95% probability that two sites, separated by as much as 5 km, will have site responses that differ by less than a factor of 2. Even as the frequency increases, there is still a high probability (~90%) that the site response at any two sites will differ by less than a factor of 3. However, as shown in the plots of site-response coefficient versus distance, there are isolated differences as large as 10 for individual frequencies, and there are also many smaller differences. The cumulative probability shows that the large outliers have a probability of 1% or less of being observed.

For the 466 station pairs with the least similar AVS30, one sees, as expected, more variation in the site-response coefficient as a function of distance for all six frequencies. This translates into cumulative probabilities that are almost the same for the six frequencies. Although there is more variation, the cumulative probabilities show that there is more than an 80% probability that two sites will have am-

 Table 3

 Regression Parameters for Equation (6)

Frequency	Most Si	milar AVS30	Least Similar AVS30			
(Hz)	а	Residuals	а	Residuals		
0.5	0.17	7.5E-04	0.37	3.7E-04		
0.75	0.26	5.4E-04	0.67	2.3E-03		
1	0.28	9.1E-04	0.84	3.1E-03		
2	0.61	6.9E-04	0.86	7.1E-04		
5	0.65	3.5E-04	0.73	4.9E-04		
10	0.79	1.0E-03	0.99	4.4E-04		

plification that differs by less than a factor of 3 for all frequencies.

Discussion

Because of the density of stations and the detailed information about each site and the high seismicity, the Yokohama High-Density Strong-Motion Network provides a unique opportunity to study both source and site characteristics. Using the borehole records at nine stations, we inverted the characteristics for the path and source parameters for 28 earthquakes. Although range of magnitudes is limited $(M_{\rm w} 4.0-6.1)$, we found that the Brune stress drop is nearly constant ~ 10 MPa with a scatter of about a factor of two. The stress drop appears to increase with depth (e.g., Fig. 3c), a trend consistent with Moya et al. (2000). For the deep events, the attenuation parameter is nearly independent of frequency. The attenuation parameter has a stronger dependence for the shallow earthquakes, but we have examined only four events. Kinoshita and Ohike (2002) and Yamanaka et al. (1998) have found a stronger dependence on frequency; however, both had limited their analysis to frequencies less than 10 Hz. For comparison with the previous results on attenuation we plot all three synthetic spectra in Figure 13. The spectral shape for frequencies less than 10 Hz is similar for the three models, but if we extend all models to higher frequencies, the Q(f) model determined in our analysis provides a better fit to the data.

In our approach, we iterated on the inversions until the site response of the borehole record stabilized. This approach is independent of a reference site. Fortunately, we had two independent methods for verifying the derived site response. Each of the boreholes had been logged for elastic wave velocities and density. In addition, each borehole has an accelerometer at the surface and in the borehole. The former allowed us to compute a theoretical response; the latter provided an empirical response. There is very good agreement between the derived site responses based on our inversion and empirical spectral ratio (Fig. 6). Thus, we are confident that the inversion scheme successfully separated source, path, and site. Having determined the source and path parameters, we computed spectra for all of the 150 surface stations for 24 deep events. We averaged the differences



Figure 10. A contour plot of the site-response coefficients $SRC(f)^+$ (see text) from 466 distinct station pairs with a separation distance between 119 m and 5 km. We smoothed over a distance of 5 m and over a frequency interval of 0.25 Hz. Although there is spatial variability, the site-response coefficient is generally less than 2.0 (separation distances up to 5 km). As expected, there is more variation in the site-response coefficient for the higher frequencies. Histogram at the top indicates the number of station pairs within 100-m intervals. Although a few station pairs have distances less than 300 m, there are several gaps in the station-separation distance. To avoid contouring artifacts resulting from the gaps, such as that between 0.5 and 0.6 km, we consider station separations greater than 300 m.

between the computed and observed spectra to find an average site response that we then analyzed.

Because of the density of sites, we developed two functions—a site-response correlation and a site-response coefficient—that allowed us to quantify different site responses and to predict site responses in both the frequency and spatial domains. We use the site-response coefficient to predict an unknown site response for distances up to 5 km from a site with a known site response. If the two sites have similar AVS30 values, we found that for frequencies up to 1.0 Hz, the unknown site response has a 95% probability of differing by less than a factor of 2 from a known site response. By constructing cumulative probability curves (Fig. 12), we show that there is a 90% probability that site-response coefficients will differ by less than a factor of 3 for frequencies up to 10.0 Hz. In the case where the two sites have the least



Figure 11. The averaged site-response coefficients for the 466 station pairs with the most similar and least similar AVS30 are plotted as a function of frequency (a) and distance (b), (c). In (a) each site-response coefficient is averaged over all station-separation distances from 0 to 5 km; the dashed line is the mean $\pm \sigma$. The most similar and least similar AVS30 are the black and gray lines, respectively. In b (most similar) and c (least similar), each site-response ratio is averaged over frequencies of 0.5–10.0 Hz.

similar AVS30, we found that for all six frequencies, 80% of the predictions for an unknown site response would be within a factor of 3 of a known site response for separation distances up to 5 km. Thus for all 2262 station pairs we would infer that the site response can be predicted within a factor of 3 with 80% probability for all frequencies up to 10 Hz.

The variation of site response comes from event-toevent variation and site-to-site variation. The site-to-site variation has been conventionally estimated by using coherence between two sites (e.g., Harichandran, 1988; Abrahamson *et al.*, 1991; Zerva and Zhang, 1997; Zerva, 2000); such studies are based on the difference in phase of two signals at two sites. These studies show that the coherency, for frequencies of 1.0 Hz and greater, is less than 0.5 if two sites are separated by more than 1.0 km. In our study we have concentrated on the amplitude of the site response and how it varies from site to site. If one averages over all distances, the average site response coefficients (Fig. 11a) vary by less than a factor of two between 0.5 and 10 Hz with a standard deviation of about 1.5 and 2.0 for sites with similar AVS30 and dissimilar AVS30, respectively.

The density of the Yokohama accelerometer network coupled with the completeness of the site characterization has allowed us to quantify site response as a function of station separation and frequency. This analysis provides, for the first time, a detailed quantitative look at how the amplitude of the site response varies with regard to distance, frequency, and local shear-wave velocity. It may provide a better understanding of the predictability of ground motion from future earthquakes.

Acknowledgments

We are indebted to the city of Yokohama for allowing us access to this unique data set. We particularly appreciate the help of Pengcheng Liu in developing the nonlinear inversion. Thoughtful critiques from two anonymous reviewers and the associate editor have greatly improved this manuscript. We dedicate this article to the late Prof. Masayuki Kikuchi of the Earthquake Research Institute, University of Tokyo, who not only constructed this array but also continuously encouraged us (K.T., K.K.) to do this study. This work was supported by the U.S. Geological Survey, 03HQGR0053, and by the Southern California Earthquake Center, funded by the National Science Foundation and the U.S. Geological Survey. This is ICS Contribution Number 644 and SCEC Contribution Number 938.

References

- Abrahamson, N. A., and D. Sykora (1993). Variations of ground motions across individual sites, in *Proc. Fourth DOE Natural Phenomena Hazards Mitigation Conf. I*, 9192–9198.
- Abrahamson, N. A., J. F. Schneider, and J. C. Stepp (1991). Empirical spatial coherency functions for application to soil-structure interaction analyses, *Earthquake Spectra* 7, 1–27.
- Archuleta, R. J. (1994). SCEC portable instrumentation and analysis of the source and site effects of the Northridge earthquake, *EOS Trans. AGU* 75, no. 44, Suppl. 175.
- Archuleta, R. J., E. Cranswick, C. Mueller, and P. Spudich (1982). Source



Figure 12. For six central frequencies we show the site-response coefficients as a function of station separation (upper panel) and cumulative probability as function of site-response coefficients (lower panel). Red lines correspond to pairs of stations with the most similar AVS30; black lines correspond to pairs of stations with the least similar AVS30. The value listed on each of the lower panels corresponds to the cumulative probability where the site-response coefficient is equal to 2. The plots show that for frequencies up to 1.0 Hz, there is a 95% probability that an unknown site response will be within a factor of 2 of a known site response if the two sites, with similar AVS30, are within 5 km of each other. The site response will be within a factor of 3 for 90% of the station pairs with similar AVS30. For station pairs with the least similar AVS30, there is an 80% chance at all six frequencies that the site response will be within a factor of 3.



Figure 13. We compare Q(f) from our analysis with other studies. The black line corresponds to the observed spectrum recorded on 29 September 2000 (event 18, Table 1). The red line is the synthetic spectrum based on our source parameters, Q(f) model, and borehole-response correction. The blue line is the spectrum from Kinoshita and Ohike (2002); the green line is the spectrum from Yamanaka *et al.* (1998) using the same source parameters and borehole correction used to compute the red spectrum but limited to 1–10 Hz—the frequency band for their model of Q(f).

- Beresnev, I. A. (2002). Nonlinearity at California generic soil sites from modeling recent strong-motion data, *Bull. Seism. Soc. Am.* 92, 863– 870.
- Boatwright, J. (1978). Detailed spectral analysis of two small New York State earthquakes, *Bull. Seism. Soc. Am.* 68, 1117–1131.
- Bonilla, L. F., J. H. Steidl, G. T. Lindley, A. G. Tumarkin, and R. J. Archuleta (1997). Site amplification in the San Fernando valley, California: variability of site-effect estimation using the S-wave, coda, and H/V methods, *Bull. Seism. Soc. Am.* 87, 710–730.
- Borcherdt, R. D. (1970). Effects of local geology on ground motion near San Francisco Bay, *Bull. Seism. Soc. Am.* 60, 29–61.
- Borcherdt, R. D., and J. F. Gibbs (1976). Effects of local geological conditions in the San Francisco Bay region on ground motions and the intensities of the 1906 earthquake, *Bull. Seism. Soc. Am.* 66, 467–500.
- Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res. 75, 4997–5009.
- Brune, J. N. (1971). Erratum to Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res. 76, 5002.
- Field, E. H., and S. E. Hough (1996). The variability of PSV response spectra across a dense array deployed during the Northridge aftershock sequences, *Earthquake Spectra* 13, 243–257.
- Field, E. H., and the Southern California Earthquake Center (SCEC) Phase III Working Group (2000). Accounting for site effects in probabilistic seismic hazard analyses of Southern California: overview of the SCEC Phase III Report, *Bull. Seism. Soc. Am.* **90**, S1–S32.
- Fletcher, J. B., J. Boatwright, and A. G. Lindh (2003). Wave propagation and site response in the Santa Clara Valley, *Bull. Seism. Soc. Am.* 93, 480–500.
- Hanks, T. C. (1978). Earthquake stress drops, ambient tectonic stresses and stresses that drive plate motions, *Pure Appl. Geophys.* **115**, 441–458.
- Hanks, T. C. (1982). f_{max}, Bull. Seism. Soc. Am. 71, 1867–1879.
- Harichandran, R. S. (1988). Local spatial variation of earthquake ground motion, in *Earthquake Engineering and Soil Dynamics II, Recent Ad*vances in Ground-Motion Evaluation, J. L. Von Thun (Editor), ASCE Geotechnical Special Publication No. 20, New York, New York.
- Harichandran, R. S. (1999). Spatial Variation of Earthquake Ground Motion, Michigan State University, http://www.msu.edu/~harichan.
- Hartzell, S., E. Cranswick, A. Frankel, D. Carver, and M. Meremonte (1997). Variability of site response in the Los Angels urban area, *Bull. Seism. Soc. Am.* 8, 1377–1400.
- Hough, S. E., and E. H. Field (1996). On the coherence of ground motion in San Fernando Valley, *Bull. Seism. Soc. Am.* 86, 1724–1732.
- Kawase, H. (1996). The cause of the damage belt in Kobe: the basin edge effects, constructive interference of the direct S-wave with the basininduced diffracted/Rayleigh waves, *Seism. Res. Lett.* 67, 25–34.
- Kinoshita, S. (1992). Local characteristics of the f_{max} of bedrock motion in the Tokyo metropolitan area, Japan, J. Phys. Earth. **40**, 487–515.
- Kinoshita, S., and M. Ohike (2002). Scaling relations of earthquakes that occurred in the upper part of the Philippine Sea plate beneath the Kanto region, Japan, estimated by means of borehole recordings, *Bull. Seism. Soc. Am.* 92, 611–624.
- Kiureghian, D. A. (1996). A coherency model for spatially varying ground motions, *Earthquake Eng. Struct. Dyn.* 25, 99–111.
- Kubo, A., E. Fukuyama, H. Kawai, and K. Nonomura (2002). NIED seismic moment tensor catalogue for regional earthquakes around Japan: quality test and application, *Tectonophysics* 356, 23–48.
- Moya, A., and K. Irikura (2003). Estimation of site effects and *Q* factors using a reference event, *Bull. Seism. Soc. Am.* **93**, 1730–1745.

- Moya, A., J. Aguirre, and K. Irikura (2000). Inversion of source parameters and site effects from strong ground motion records using genetic algorithms, *Bull. Seism. Soc. Am.* **90**, 977–992.
- Reid, H. F. (1910). The California Earthquake of April 18, 1906, Publication 87, V21, Carnegie Institute of Washington, Washington, D.C.
- Sasatani, T., S. Higashi, and H. Nagumo (1998). Strong motion observations in Kushiro, Hokkaido, Japan, in *Effects of Surface Geology on Seismic Motion*, Vol. 1, K. Irikura, K. Kudo, H. Okada, and T. Sasatani (Editors), AA Rotterdam: Balkema, 279–284.
- Sen, K. M., and P. L. Stoffa (1995). Global Optimization Methods in Geophysical Inversion, Elsevier, New York.
- Stewart, J. P., and M. B. Baturay (2000). Evaluation of uncertainties in ground motion estimates for soil sites, PEER Utilities Program Report No. 2000/04, Pacific Earthquake Engineering Research Center, Richmond, California.
- Stewart, J. P., and M. B. Baturay (2001). Uncertainties and residual in ground motion estimates at soil sites, in 4th International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, P. Shamsher (Editor), University of Missouri–Rolla, Rolla, Missouri.
- Tsuboi, S., M. Saito, and Y. Ishihara (2001). Verification of horizontal-tovertical spectral ratio technique for estimation of the site response using borehole seismographs, *Bull. Seism. Soc. Am.* 91, 499–510.
- Tsuda, K. (2001) Inversion of ground motion spectra for site response at 150 stations of a dense accelerograph array, *Master's Thesis*, University of Tokyo.
- Yamanaka, H., A. Nakamaru, K. Kurita, and K. Seo (1998). Evaluation of site effects by an inversion of S-wave spectra with a constraint condition considering effects of shallow weathered layers, *Zisin* 51, 193– 202 (in Japanese with English abstract).
- Yamazaki, K., M. Minamishima, and K. Kudo (1992). Propagation characteristics of intermediate-period (1–10 seconds) surface waves in the Kanto Plain, Japan, J. Phys. Earth. 40, 117–136.
- Yokohama City Office (1998). Report for the investigation of soil, geology conditions around Yokohama area (in Japanese).
- Zerva, A. (2000). Spatial variability of seismic motions recorded over extended ground surface areas, in *Wave Motion in Earthquake Engineering*, E. Kausel and G. Manolis (Editors), WIT Press, Southhampton, UK, 97–141.
- Zerva, A., and J. L. Beck (2003). Identification of parametric ground motion random fields from spatially recorded seismic data, *Earthquake Eng. Struct. Dyn.* 32, 77–791.
- Zerva, A., and O. Zhang (1997). Correlation patterns in characteristic of spatially variable seismic ground motions, *Earthquake Eng. Struct. Dyn.* 26, 19–39.

Department of Earth Science and Institute for Crustal Studies University of California

Santa Barbara, California 93106 (K.T., R.J.A.)

Earthquake Research Institute University of Tokyo 1-1-1 Yayoi, Bunkyo Tokyo 113-0032, Japan (K.K.)

Manuscript received 15 November 2004.