Bulletin of the Seismological Society of America, Vol. 101, No. 5, pp. 2281–2295, October 2011, doi: 10.1785/0120090312

Repeatable Source, Site, and Path Effects on the Standard Deviation for Empirical Ground-Motion Prediction Models

by Po-Shen Lin, Brian Chiou, Norman Abrahamson, Melanie Walling, Chyi-Tyi Lee, and Chin-Tung Cheng

Abstract In this study, we quantify the reduction in the standard deviation for empirical ground-motion prediction models by removing ergodic assumption. We partition the modeling error (residual) into five components, three of which represent the repeatable source-location-specific, site-specific, and path-specific deviations from the population mean. A variance estimation procedure of these error components is developed for use with a set of recordings from earthquakes not heavily clustered in space. With most source locations and propagation paths sampled only once, we opt to exploit the spatial correlation of residuals to estimate the variances associated with the path-specific and the source-location-specific deviations. The estimation procedure is applied to ground-motion amplitudes from 64 shallow earthquakes in Taiwan recorded at 285 sites with at least 10 recordings per site. The estimated variance components are used to quantify the reduction in aleatory variability that can be used in hazard analysis for a single site and for a single path. For peak ground acceleration and spectral accelerations at periods of 0.1, 0.3, 0.5, 1.0, and 3.0 s, we find that the single-site standard deviations are 9%–14% smaller than the total standard deviation, whereas the single-path standard deviations are 39%–47% smaller.

Introduction

Ground-motion prediction models describe the mean and the aleatory variability of ground-motion amplitude in natural logarithmic units. For brevity, we will in this paper omit the reference to logarithm when we discuss the distribution of ground-motion amplitude. The aleatory variability is measured, in most empirical models, by the standard deviation of regression residual (observed amplitude minus predicted mean) from a data set that includes recordings at a broad range of site conditions and from earthquakes that occurred at different locations, sometimes in different tectonic regions. By using the observed standard deviation in a site-specific seismic hazard analysis, there is an implicit assumption that the variability seen in such a mixture is the same as the variability seen in the ground-motion amplitudes at a single site from multiple earthquakes occurring at a single source location or in a small source volume. This assumption has been referred to as the ergodic assumption of ground-motion amplitude by Anderson and Brune (1999), who have also suggested that it is not likely to hold in reality. The variability of ground-motion amplitude recorded at a single site or from a single path is expected to be smaller than the variability for a mixture of sites and source locations because the latter includes additional source location-to-location, site-to-site, and path-to-path variability.

It is well known that the value of the standard deviation of ground-motion amplitude has a large impact on the seismic hazard at long return periods (e.g., Restrepo-Velez and Bommer, 2003). A review of issues related to the estimation of standard deviation is given by Strasser *et al.* (2009). One promising approach for improving the estimate of standard deviation is to remove the ergodic assumption. Such improvement was not readily feasible in the past because data available then were insufficient to allow for a stable estimation of the variance components. For example, in the data set used by the Next Generation of Ground-Motion Attenuation Model project (Chiou *et al.*, 2008), only a few strong-motion sites recorded more than 10 earthquakes, and for these sites, the recorded earthquakes did not always occur close to each other. This limitation has been drastically improved in recent years, particularly in the small to moderate magnitude range ($3 \leq M \leq 5.5$), thanks to the installation and continuous operation of dense networks of broadband digital instruments, such as the networks in northern California (Boatwright *et al.*, 1999) and southern California (Wald *et al.*, 1999); K-NET (Kinoshita, 1998) and KiK-net (Aoi *et al.*, 2000) in Japan; and the Taiwan Strong-Motion Instrumentation Program (TSMIP) network in Taiwan (Liu *et al.*, 1999; Shin and Teng, 2001).

Using data from these networks, several recent studies (Chen and Tsai, 2002; Atkinson, 2006, Morikawa *et al.*,

2008) have estimated the standard deviations of site-to-site and path-to-path variability. These studies have found that the standard deviation of peak ground acceleration (PGA) is reduced by 7%–14% if site-to-site variability is removed and by 30%–60% if both site-to-site and path-to-path variability are removed. The results from previous studies also suggest that the path-to-path variability has a greater contribution to the total standard deviation than the site-to-site variability.

The main objective of this study is to quantify the reduction in standard deviation by removing ergodic assumption. We present a model of ground-motion amplitude that includes a decomposition of residuals into five components, three of which represent repeatable deviations whose variability adds to the total variability. We also propose a procedure by which the variance of each component can be estimated. We then apply this procedure to a large data set from Taiwan with at least 10 recordings per site. Finally, we present the results from the Taiwan data and compare these results with those from the previous studies in terms of the reduction in standard deviation due to the exclusion of certain variance components in an effort to remove the ergodic assumption.

Components of Variability

In this study, the nonlinear site amplification effects on the standard deviation of residual are ignored for the sake of simplicity. With this simplification, the mixed-effect model for the observed ground-motion amplitude (Abrahamson and Youngs, 1992) can be written as

$$y_{ik} = f(\mathbf{x}_{ik}, \boldsymbol{\theta}) + \eta_{E_i} + \xi_{ik}, \tag{1}$$

where y_{ik} is the natural logarithm of the observed 5%-damped pseudospectral acceleration at the kth site from the ith earthquake, $\mathbf{x}_{ik}$ is the vector of predictors (e.g., magnitude, distance, style-of-faulting factors, and site V_{S30}), $\boldsymbol{\theta}$ is the vector of fixed-effect coefficients to be estimated by the regression analysis, and $f(\mathbf{x}_{ik}, \boldsymbol{\theta})$ is the population mean. The modeling error of equation (1) consists of two terms: η_{E_i} (event term) is the mean deviation of the ith earthquake from the population mean, and ξ_{ik} is the modeling error (residual) after the event term is removed. Note that not all earthquakes are recorded at all sites. Though there are combinations of earthquakes and sites for which there are no data, we use k as a site index rather than as a record index (as was in Abrahamson and Youngs, 1992) for consistency with the notation used later.

The vector of residuals ξ_{ik} and the vector of event terms η_{E_i} are assumed to be multivariate zero-mean normal variables that are mutually independent and identically distributed, and ξ_{ik} is assumed to be independent of η_{E_i}. The total variance (σ_T) for equation (1) is

$$\sigma_T^2 = \tau_E^2 + \sigma^2, \tag{2}$$

where τ_E and σ denote the standard deviation of η_{E_i} and ξ_{ik}, respectively. This type of variance decomposition has been widely adopted in recent ground-motion prediction models. In general, τ_E and σ may be dependent on earthquake magnitude, distance, and/or ground-motion amplitude (e.g., Youngs *et al.*, 1995; Abrahamson *et al.*, 2008). Such features, however, are not the objective of our investigation, and they are not pursued in this study.

Adding Site Term

With multiple recordings at each site, residual ξ_{ik} can be partitioned to include the deviation of the site-specific amplification from the average soil amplification for the soil class (or V_{S30}) of a site. With the site-specific terms, equation (1) can be rewritten as

$$y_{ik} = f(\mathbf{x}_{ik}, \boldsymbol{\theta}) + \eta_{E_i} + \eta_{S_k} + \xi_{r_{ik}}. \tag{3}$$

The variable η_{S_k} (site term) represents the site-specific deviation of the kth site, and $\xi_{r_{ik}}$ is the residual after accounting for the event term η_{E_i} and the site term η_{S_k}. This model equation has been used by Chen and Tsai (2002) to investigate the event-to-event and site-to-site variabilities.

In equation (3), data are grouped by both the factors of event and site, and the site factor is not nested within the event factor. In other words, equation (3) is a two-way cross classification of spectral data, not a hierarchical classification. For this reason, we do not use the term intraevent or intrasite to describe the residual $\xi_{r_{ik}}$. We use the subscript r to distinguish it from the residual ξ_{ik} defined earlier and from the residual defined later.

The vectors of η_{S_k}, η_{E_i}, and $\xi_{r_{ik}}$ are assumed to be independent of each other, and the total variance based on equation (3) is the sum of three components:

$$\sigma_T^2 = \tau_E^2 + \tau_S^2 + \sigma_r^2, \tag{4}$$

where τ_S is the standard deviation of η_{S_k} (representing the site-to-site variability of the site-specific term), and σ_r is the standard deviation of $\xi_{r_{ik}}$.

Adding Path Term and Spatial Dependence

Residual $\xi_{r_{ik}}$ in equation (3) can be, in concept, further decomposed to include a path-specific term to represent the average deviation of a path from the population mean $f(\mathbf{x}_{ik}, \boldsymbol{\theta})$. With $\xi_{r_{ik}} = \eta_{P_{ik}} + \xi_{0_{ik}}$, equation (3) is rewritten as

$$y_{ik} = f(\mathbf{x}_{ik}, \boldsymbol{\theta}) + \eta_{E_i} + \eta_{S_k} + \eta_{P_{ik}} + \xi_{0_{ik}}, \tag{5}$$

where $\eta_{P_{ik}}$ (path term) is the mean deviation specific to the path between the ith earthquake location and the kth site, and $\xi_{0_{ik}}$ is the remaining residual after accounting for the event, site, and path terms.

Clearly, for the purpose of estimating $\eta_{P_{ik}}$ and its standard deviation (τ_P, representing the path-to-path variability of the path-specific term), equation (5) is overparameterized for a data set with the majority of paths sampled just once, and the standard deviation of $\xi_{0_{ik}}$ (denoted as σ_0) cannot be reliably separated from τ_P. Previous studies have defined source regions in such a way that one could assume similar path terms at a site for earthquakes in the same source region (Atkinson, 2006; Morikawa, 2008). Alternatively, Anderson and Uchiyama (2010) use weighted spatial average to provide an estimate of the path term for each earthquake-site combination. In this study, we use a regionless approach to estimate τ_P and σ_0 without actually estimating and removing the individual path term $\eta_{P_{ik}}$.

Our regionless approach is based on the postulation that path terms $\eta_{P_{ik}}$ are correlated and the degree of correlation increases with decreasing separation between paths, as adjacent paths share more similarity than far-apart paths. A metric that measures the path separation is defined later in Approach for Estimating the Variance Components. The vector of $\xi_{0_{ik}}$ is assumed to be mutually independent and independent of $\eta_{P_{ik}}$. Under these assumptions, the vector of residual $\xi_{r_{ik}} (= \eta_{P_{ik}} + \xi_{0_{ik}})$ is correlated; its covariance property, assuming stationarity, is described by the semivariogram (Cressie, 1993)

$$\gamma(\xi_{r_{ik}}, \xi_{r_{jk}}) = \frac{1}{2}\,\mathrm{Var}[\xi_{r_{ik}} - \xi_{r_{jk}}] = \frac{1}{2}E[(\xi_{r_{ik}} - \xi_{r_{jk}})^2]$$
$$= \tau_P^2[1 - \rho(\eta_{P_{ik}}, \eta_{P_{jk}})] + \sigma_0^2, \tag{6}$$

where $\rho(\eta_{P_{ik}}, \eta_{P_{jk}})$ is the correlation function. Note that the $1/2$ factor eliminates the need to carry a factor of 2 in the last equality of equation (6). If ground-motion amplitudes are limited to those from identical paths (zero separation), then $\gamma(\xi_{r_{ik}}, \xi_{r_{jk}})$ is equal to σ_0^2 because $\rho(\eta_{P_{ik}}, \eta_{P_{jk}}) = 1$ ($\eta_{P_{ik}}$ is identical to $\eta_{P_{jk}}$). If data are limited to those from far-apart (zero correlation) paths, then $\gamma(\xi_{r_{ik}}, \xi_{r_{jk}}) = \tau_P^2 + \sigma_0^2$.

Adding Source-Location Term

The event term η_{E_i} (for rock site condition) captures the effect of difference in the source properties for the different earthquakes in the data set. It is reasonable to assume that earthquakes occurring in a small source volume would have similar source properties and hence similar event terms. Incorporating this into the ground-motion prediction model, equation (5) becomes

$$y_{ik} = f(\mathbf{x}_{ik}, \boldsymbol{\theta}) + (\eta_{SR_i} + \eta_{E0_i}) + \eta_{S_k} + \eta_{P_{ik}} + \xi_{0_{ik}}, \tag{7}$$

where η_{SR_i} represents the mean event term for the cluster of earthquakes at the location of the ith earthquake, and η_{E0_i} is the event term of the ith earthquake after accounting for the repeatable source-location effect η_{SR_i}. The variance τ_E^2 is decomposed as

$$\tau_E^2 = \tau_{SR}^2 + \tau_{E0}^2, \tag{8}$$

where τ_{SR} and τ_{E0} are the standard deviation of η_{SR_i} and η_{E0_i}, respectively. The standard deviation τ_{SR} represents the location-to-location variability of the source-location-specific term. Estimation of τ_{SR} and τ_{E0} is difficult for a data set with an insufficient number of colocated or tightly clustered earthquakes. Using the same idea as previously discussed, we postulate that η_{SR_i} is correlated between two nearby source locations; we use a regionless approach based on the semivariogram of spatially correlated η_{E_i} for the estimation of τ_{SR} and τ_{E0}. More details on the estimation are discussed in Approach for Estimating the Variance Components.

Site terms η_{S_k} for two neighboring sites are likely to be correlated as well, and their correlation structure may also be included in our analysis. We have not pursued this additional step, because the focus of this study is on the site-specific hazard analysis.

Aleatory Variability for Single-Path and Single-Site Application

The modeling error of equation (7) is partitioned into three repeatable terms and two random error terms (η_{E0_i} and $\xi_{0_{ik}}$). The three repeatable terms represent the source-location-specific (η_{SR_i}), site-specific (η_{S_k}), and path-specific ($\eta_{P_{ik}}$) deviations from the population mean $f(\mathbf{x}_{ik}, \boldsymbol{\theta})$. The total aleatory variability of a ground-motion prediction model includes variability of these three repeatable terms if it is derived from a mixture of sites and earthquake locations. The total variance σ_T^2, with ergodic assumption, can be written as the sum

$$\sigma_T^2 = \tau_{SR}^2 + \tau_S^2 + \tau_P^2 + \tau_{E0}^2 + \sigma_0^2. \tag{9}$$

As explained in the Introduction, some applications do not warrant all components of variance. A discussion of two such applications follows.

Single-Site Application

Typically, in both deterministic and probabilistic analysis of site-specific earthquake hazard, the variance of predicted amplitude is assumed to be all aleatory variability. In a single-site application, the site-specific term η_{S_k} is a repeated term (a fixed effect), and it does not contribute to the aleatory variability for that site. The total variance for the single-site application, σ_{SS}^2, is thus given by

$$\sigma_{SS}^2 = \tau_{SR}^2 + \tau_P^2 + \tau_{E0}^2 + \sigma_0^2. \tag{10}$$

Note that the impact of removing the ergodic assumption goes beyond reducing the standard deviation. If we remove the ergodic assumption from the standard deviation, then we must also remove the ergodic assumption from the mean. Once we remove the site-to-site variability for a single-site application,

we can no longer use $f(\mathbf{x}_{ik}, \boldsymbol{\theta})$ as the mean amplitude for that site; the site-specific deviation must be included as part of the mean amplitude. The value of the site-specific term is estimated; it is not known exactly. As a result, the mean amplitude for the given site is subject to epistemic uncertainty. The accompanied epistemic uncertainty will offset the benefits of reduced aleatory variability in a hazard analysis if there are insufficient data or other information to constrain the site-specific mean amplitude.

Single-Path Application

In a single-path application (i.e., in a site-specific analysis involving only one source location or one small source volume) the terms η_{SR_i}, η_{S_k}, and $\eta_{P_{ik}}$ each have a fixed value, and they do not contribute to the aleatory variability for that path. Thus, for a single-path application the total aleatory variability comprises only the random variability of η_{E0_i} and $\xi_{0_{ik}}$,

$$\sigma_{SP}^2 = \tau_{E0}^2 + \sigma_0^2, \tag{11}$$

where σ_{SP} denotes the total standard deviation for application to a single path. Similar to the single-site application, once we remove the variability associated with the three repeatable terms in a single-path application, we can no longer use $f(\mathbf{x}_{ik}, \boldsymbol{\theta})$ as the mean amplitude; the deviations specific to the given path must be included as part of the mean amplitude. Because the values of these specific terms are estimated, the mean amplitude for the given path will be subject to epistemic uncertainty.

Data Set

The TSMIP (Liu *et al.*, 1999; Shin and Teng, 2001) operates a strong-motion network of approximately 700 strong-motion instruments that have recorded thousands of earthquakes in and near Taiwan since 1993, including the 1999 **M** 7.6 ChiChi earthquake. A large ground-motion database has been assembled by Lin (2009), who compiled and processed acceleration time series from shallow earthquakes that occurred from 1992 to 2003. Each ground-motion recording was baseline corrected, and the maximum usable period was determined following the procedures described in Chiou *et al.* (2008). The 5%-damped pseudospectral accelerations at periods ranging from 0.01 s to 10 s were calculated for both horizontal components of a recording. The orientation-independent measurement of average horizontal amplitude, GMRotI50 (Boore *et al.*, 2006), was also computed and included in the database. For each strong-motion station, an estimate of the V_{S30} value was either computed from the *PS* suspension logging data at or near the station or obtained from Lee *et al.* (2008), who mapped the distribution of V_{S30} in Taiwan using available geotechnical data and multivariate geostatistical technique.

In this study, we use Lin's database to estimate the variability components described previously. We apply the following criteria to select a data subset that fits our needs. First, we limit the data to distances of less than 200 km to be consistent with the distance range of applicability of the Chiou and Youngs (2008) ground-motion prediction model. Second, to ensure a stable estimation of the site terms, we restrict data to sites that had recorded at least 10 earthquakes. Finally, the 1999 Chi-Chi mainshock is excluded from the final data set because of complexities in defining the source-to-site path for extended ruptures. Spectral accelerations from the 16 March 2000, **M** 4.85 earthquake show little distance attenuation in distances between 10 km and 100 km. This event is also excluded. Further evaluation of this event is needed to understand why it is an anomaly.

Some of the predictor variables needed for our study are either missing or incomplete. Their values are filled in as followed. A strike-slip mechanism is assumed for the five earthquakes with unknown style-of-faulting factors. This only impacts their event terms. Hypocentral distance is used as the distance measure in this study, except for recordings from a few larger earthquakes with a finite source model. The hanging-wall effect impacts only the ground-motion amplitude a short distance from a large earthquake (magnitude larger than 6). Consequently, it is not an important effect for our data set, and we ignore it in this study. The relationship between soil depth and V_{S30} given in Chiou and Youngs (2008) is used to fill in the missing soil depth.

The final selected data set consists of 4756 recordings from 64 earthquakes (Table 1) and 285 strong-motion stations (Fig. 1). The magnitude-distance distribution of these recordings is shown in Figure 2. The selected earthquakes have moment magnitudes in the range of 3.9 to 6.3. A histogram of the number of recordings per station is shown in Figure 3. Most of the stations recorded between 10 and 30 earthquakes; only four stations recorded more than 30 earthquakes. A histogram of V_{S30} from the 285 stations is shown in Figure 4. The large peaks in the histogram at $V_{S30} =$ 220 m/s and $V_{S30} = 760$ m/s reflect the sites for which V_{S30} was estimated using correlation between surface geology and V_{S30}. We expect the site term to reflect the error in site V_{S30} and its standard deviation (τ_S) to include the contribution from V_{S30} measurement error as well as the site-to-site variation in soil amplification for a given V_{S30}.

The GMRotI50 data are only used for the period range that is determined to be reliable. The number of usable recordings for each of the six ground-motion parameters analyzed in this study is listed in Table 2. At a period of 3 s, the number of recordings is reduced approximately 10% from that at PGA.

Approach for Estimating the Variance Components

The error of the mixed-effects model of equation (1) is separated into η_{E_i} and ξ_{ik}, and their variances can be estimated by using either the two-stage weighted-least-squares method (Joyner and Boore, 1993) or the one-stage maximum-likelihood method (Abrahamson and Youngs, 1992) to account for the correlations of spectral acceleration data from

Table 1
List of Taiwan Earthquakes Selected for Analysis in This Study

ID	Date (yyyy/mm/dd)	Time (hr:mm:ss)	Latitude	Longitude	Hypo Depth (km)	**M**	Style of Faulting*	Z_{TOR} (km)	Number of Stations
1	1993/12/13	09:23:30	24.256	120.776	22.5	4.32	RO	22.5	18
2	1993/12/15	21:49:43	23.194	120.507	15.2	5.40	R	15.2	64
3	1993/12/20	03:32:04	23.227	120.506	19.2	4.27	SS	19.2	24
4	1993/12/21	03:14:28	23.216	120.509	14.4	4.35	R	14.4	21
5	1993/12/22	16:22:20	23.208	120.499	17.0	4.49	SS	17.0	30
6	1994/03/28	08:11:15	23.018	120.714	20.4	5.30	RO	20.4	43
7	1994/04/06	01:12:11	23.494	120.448	18.7	4.91	SS	18.7	52
8	1994/05/31	15:00:06	23.692	120.794	9.2	4.43	NO	9.2	22
9	1994/06/05	01:09:30	24.468	121.787	5.1	6.30	NO	5.1	73
10	1995/01/19	11:39:08	23.305	120.753	14.6	4.32	RO	14.6	33
11	1995/02/23	05:19:02	24.204	121.687	33.0	6.20	–	33.0	17
12	1995/02/26	08:08:18	23.088	121.384	22.3	4.74	RO	22.3	15
13	1995/03/22	03:30:21	23.831	121.435	7.4	4.71	R	7.4	28
14	1995/04/11	17:47:27	23.248	120.504	16.5	3.93	R	16.5	23
15	1995/04/23	02:47:40	23.233	120.459	11.9	4.10	RO	11.9	27
16	1995/04/23	02:57:52	23.233	120.437	9.8	4.17	SS	9.8	30
17	1995/04/23	03:01:46	23.234	120.441	12.1	4.16	SS	12.1	37
18	1995/05/01	14:50:45	24.052	121.569	13.0	4.76	RO	13.0	23
19	1995/05/27	18:11:11	23.058	121.342	19.8	5.70	R	19.8	27
20	1995/07/07	03:04:48	23.896	121.078	8.6	5.19	R	8.6	107
21	1995/07/14	16:52:46	24.368	121.743	9.8	5.71	SS	9.8	56
22	1995/07/14	17:40:48	24.358	121.719	5.2	4.74	–	5.2	21
23	1995/09/28	17:58:05	23.509	120.449	11.9	4.34	SS	11.9	42
24	1995/10/31	22:27:06	23.262	120.380	18.1	4.50	R	18.1	69
25	1995/11/14	07:26:26	24.044	121.456	10.3	4.06	SS	10.3	24
26	1996/04/07	16:55:36	23.475	120.670	4.4	4.57	R	4.4	32
27	1996/05/28	21:53:22	24.069	121.477	14.8	4.73	–	14.8	48
28	1996/10/19	19:16:05	23.183	120.532	13.6	4.06	R	13.6	37
29	1996/11/16	00:22:43	23.208	120.300	20.2	4.13	RO	20.2	29
30	1997/04/02	22:36:41	24.701	121.692	8.5	4.18	N	8.5	24
31	1997/10/29	23:18:37	23.618	120.628	13.3	4.17	RO	13.3	27
32	1997/11/14	04:29:50	24.209	121.662	10.2	5.10	–	10.2	33
33	1998/07/17	04:51:14	23.500	120.660	6.0	5.66	R	6.0	43
34	1999/09/20	17:57:15	23.940	121.010	8.0	5.90	R	6.5	107
35	1999/09/20	18:03:41	23.810	120.850	8.0	6.20	R	6.7	99
36	1999/09/20	21:46:36	23.600	120.820	18.0	6.20	SS	1.5	86
37	1999/09/22	00:14:40	23.810	121.080	10.0	6.20	R	7.7	95
38	1999/09/25	23:52:49	23.870	121.010	16.0	6.30	R	10.0	98
39	2000/02/15	21:33:18	23.316	120.740	14.7	5.14	RO	14.7	149
40	2000/03/09	05:08:44	23.222	121.493	27.3	4.86	R	27.3	50
41	2000/05/17	03:25:46	24.193	121.098	9.7	5.61	SS	9.7	80
42	2000/06/10	18:23:29	23.901	121.109	13.7	6.07	R	13.7	245
43	2000/06/19	21:56:24	23.920	121.092	25.7	4.91	RO	25.7	179
44	2000/07/28	20:28:07	23.411	120.933	7.4	5.65	SS	7.4	148
45	2000/08/23	00:49:16	23.636	121.635	24.8	5.11	–	24.8	133
46	2000/09/01	09:24:38	24.080	121.138	7.7	4.77	SS	7.7	95
47	2000/09/10	08:54:46	24.085	121.584	17.3	5.70	SS	17.3	128
48	2000/12/10	19:30:44	23.116	120.226	12.0	4.95	SS	12.0	120
49	2000/12/29	18:03:28	24.361	121.884	7.0	4.76	SS	7.0	59
50	2001/01/11	08:36:59	24.081	120.987	21.1	4.56	N	21.1	108
51	2001/02/18	20:25:10	23.585	120.719	15.7	4.37	SS	15.7	107
52	2001/03/01	16:37:50	23.838	120.997	10.9	5.00	R	10.9	161
53	2001/06/14	02:35:25	24.419	121.928	8.8	5.71	SS	8.8	116
54	2001/06/19	05:16:15	23.177	121.077	6.6	5.00	N	6.6	82
55	2001/06/30	04:07:37	24.055	121.543	22.2	4.51	R	22.2	93
56	2001/09/17	22:44:44	23.276	120.654	6.8	4.82	R	6.8	69
57	2001/11/04	08:45:35	23.936	121.043	7.5	4.44	R	7.5	79
58	2002/02/12	03:27:25	23.741	121.723	35.0	5.52	RO	35.0	224
59	2002/05/15	03:46:05	24.651	121.872	8.5	5.97	N	8.5	89
60	2002/09/06	11:02:01	23.890	120.729	27.8	4.65	R	27.8	116
61	2003/04/03	06:59:33	23.153	120.485	14.4	4.28	RO	14.4	60

(*continued*)

Table 1 (*Continued*)

ID	Date (yyyy/mm/dd)	Time (hr:mm:ss)	Latitude	Longitude	Hypo Depth (km)	**M**	Style of Faulting*	Z_{TOR} (km)	Number of Stations
62	2003/06/09	01:52:50	24.370	122.023	24.4	5.58	SS	24.4	139
63	2003/06/09	05:08:04	24.380	121.851	2.4	4.71	RO	2.4	42
64	2003/06/10	08:40:32	23.504	121.699	35.1	5.71	R	35.1	201

*R, Reverse; RO, Reverse/Oblique; SS, Strike-slip; N, Normal; NO, Normal/Oblique

a single earthquake and to account for uneven sampling of the different earthquakes. As we incorporate site-specific terms and path-specific terms into equation (1), the data covariance matrix becomes more complicated. Estimating the variance components of equation (8) requires a more flexible regression method that is capable of handling the correlation structure induced by the site term and the spatially correlated path terms. A rigorous regression method of this kind may be feasible but is not attempted in this study. Instead, we use a multistep approach that is more tractable and easy to implement. This procedure, which roughly parallels the stages of error decomposition described previously, is presented next.

Step 1: Developing a Ground-Motion Prediction Model

Initially, we chose to use the empirical ground-motion model of Chiou and Youngs (2008) (hereafter, CY08) to avoid the need to derive a new empirical model for Taiwan. However, after seeing some misfits we modestly revised CY08 by using the mixed-effects model of equation (1) to obtain better fits to the Taiwan data. The resulting η_{E_i} from the revised model are plotted as a function of magnitude in Figure 5, and the residuals ξ_{ik} are plotted as a function of distance in Figure 6a. The estimated value of τ_E is listed in Table 3.

Step 2: Estimating the Standard Deviations τ_S and σ_r

The residuals ξ_{ik} from the updated CY08 are modeled by using a linear mixed-effect model (corresponding to equation 3),

$$\xi_{ik} = C + \eta_{S_k} + \xi_{r_{ik}}. \tag{12}$$

Coefficient C accounts for the possible deviation from zero due to uneven sampling of the stations. The estimated deviation is less than 0.022 for all periods (hence, they are not listed). The estimated values for τ_S and σ_r are listed in Table 3. The residuals $\xi_{r_{ik}}$ are shown in Figure 6b. Accounting for the site terms results in a significant reduction in the scatter of the residuals.

Step 3: Estimating τ_P and σ_0

As explained previously, we use a regionless approach to estimate τ_P and σ_0. The key component of our approach is the metric CI that measures the separation between two paths. This metric is defined in the following. There are multiple earthquakes recorded at the kth site, as shown in

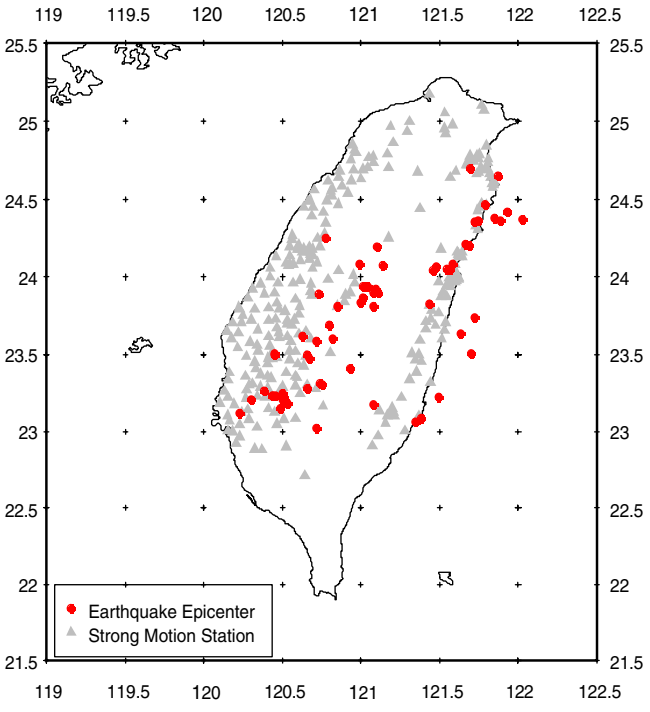

Figure 1. Distribution map of strong-motion stations and earthquakes used in this study. Solid circles indicate earthquake epicenters, and triangles indicate station locations. The color version of this figure is available only in the electronic edition.

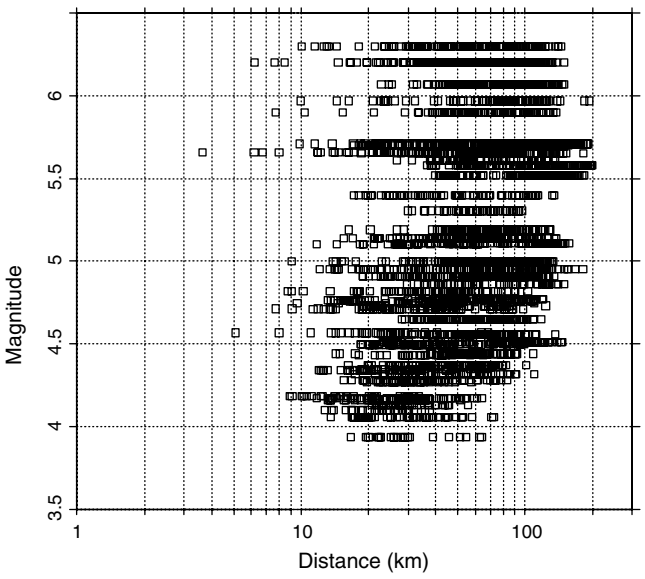

Figure 2. Magnitude-distance distribution of the selected Taiwan data.

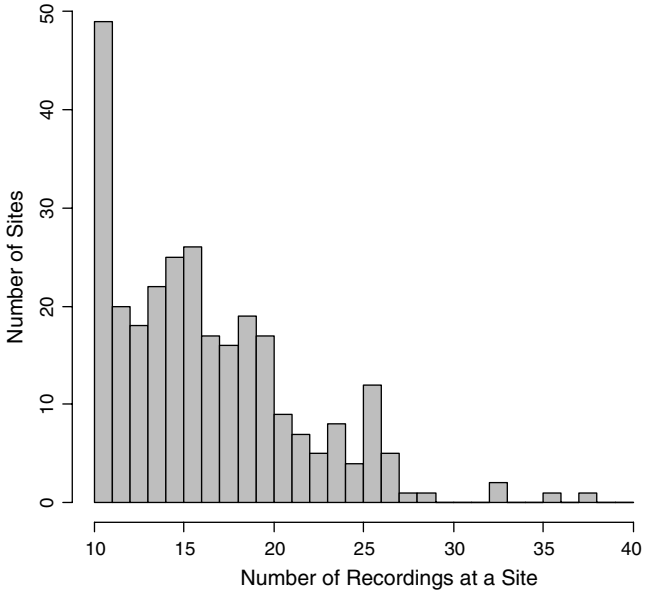

Figure 3. Histogram of the number of recordings per station for the Taiwan data set used in this study.

Figure 7. For each pair of earthquakes, the distance between the ith and jth hypocenters is denoted as ΔH_{ij}, and the hypocentral distance from the ith earthquake to the kth site is denoted as R_{ik}. To quantify the path separation between the $i-k$ and the $j-k$ paths, we define the closeness index as

$$\mathrm{CI}_{ijk} = \frac{\Delta H_{ij}}{(R_{ij} + R_{jk})/2}. \tag{13}$$

Using the closeness index avoids the need to define specific source regions. If two earthquakes are located close together when compared against the average hypocentral distance

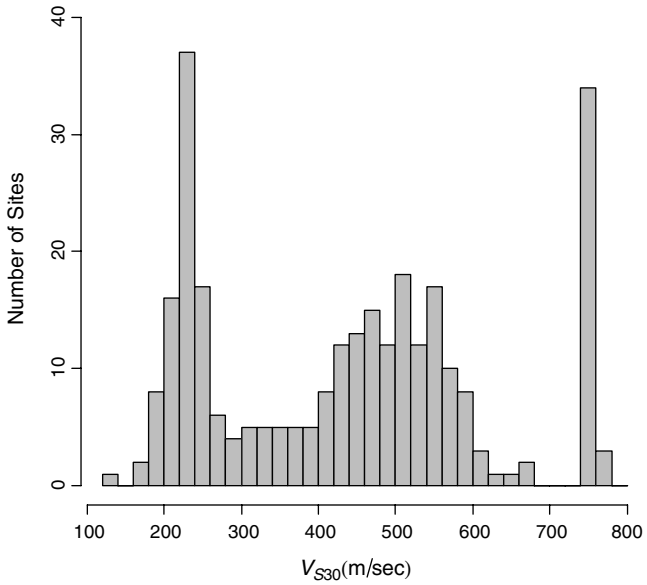

Figure 4. Histogram of the V_{S30} distribution for the 285 sites used in this study.

Table 2
Number of Usable Response Spectral Values for the Periods Analyzed

Spectral Period (s)	Number of Recordings	Number of Earthquakes
0.01 (*PGA*)	4756	64
0.1	4756	64
0.3	4756	64
0.5	4756	64
1.0	4753	64
3.0	4320	64

(small CI), then the two paths are regarded as similar. If the earthquakes are located far apart when compared with the average distance (large CI), then the two paths are regarded as dissimilar. The CI value ranges from 0 for colocated earthquakes to a value near 2 for earthquakes located in opposite (180° difference in azimuth) epicentral directions from the site. When two earthquakes are in opposite epicentral directions, ΔH_{ij} is close to $R_{ik} + R_{jk}$, hence, the CI value is near 2.

With the closeness index as our metric for path separation, we proceed to estimate τ_P and σ_0 by modeling the empirical semivarigram of $\xi_{r_{ik}}$. For convenience, we normalize the residual $\xi_{r_{ik}}$ by the σ_r estimated in Step 2. We then compute the difference between the normalized residual for all pairs of earthquakes recorded at the kth site as

$$\Delta\xi_{r_{ijk}} = \frac{\xi_{r_{ik}} - \xi_{r_{jk}}}{\sqrt{2}\cdot\sigma_r}, \qquad j > i. \tag{14}$$

The values of $\Delta\xi^2_{r_{ijk}}$ are grouped into bins equally spaced in the logarithm scale of CI (Fig. 8), except for the lowest bin

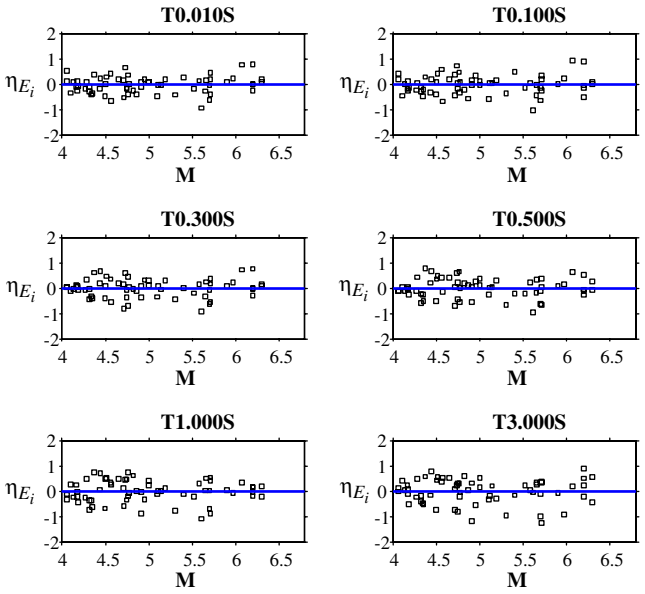

Figure 5. Event terms (η_{E_i}) from the updated Chiou and Youngs model for the Taiwan data. The color version of this figure is available only in the electronic edition.

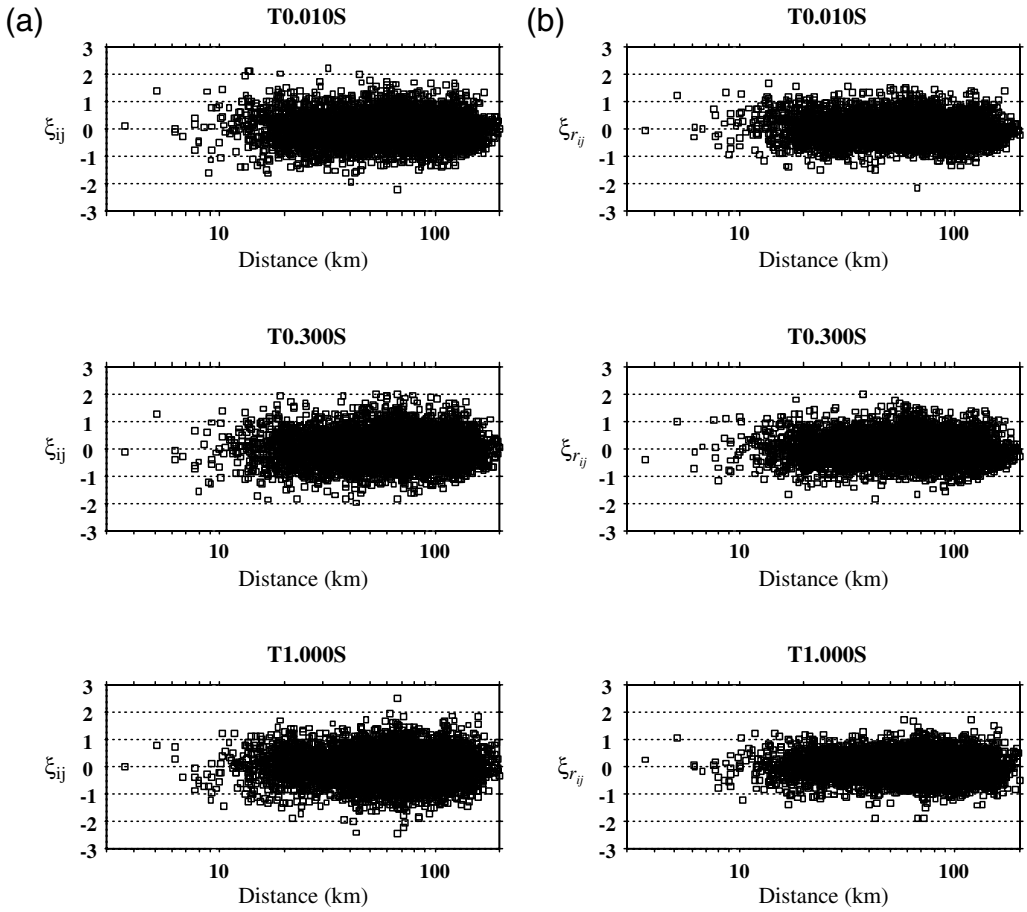

Figure 6. Residuals of PGA ($T = 0.01$ s) and response spectral accelerations at $T = 0.3$ and $T = 1$ s, (a) before and (b) after accounting for site terms.

for which the bin width is made wider to include a sufficient number of data points. The standard deviation of $\Delta\xi_{r_{ijk}}$ ($\sigma_{\Delta\xi_r}$), calculated as the square root of the mean $\Delta\xi^2_{r_{ijk}}$ (c.f., equation 6), is shown in Figure 9 for each bin. The binned $\sigma_{\Delta\xi_r}$ is clearly dependent on CI: $\sigma_{\Delta\xi_r}$ is smaller in low-CI bins and larger for high-CI bins. The trend revealed in Figure 9 suggests a parameterization of

$$\sigma_{\Delta\xi_r} = b_1 + \frac{(b_2 - b_1)\cdot \mathrm{CI}^n}{b_3 + \mathrm{CI}^n}, \qquad (15)$$

Table 3
Standard Deviations from the Mixed-Effects Regressions

	Step 1		Step 2	
Spectral Period (s)	τ_E	σ	τ_S	σ_r
0.01 (*PGA*)	0.344	0.518	0.259	0.449
0.1	0.381	0.580	0.353	0.460
0.3	0.367	0.553	0.280	0.476
0.5	0.386	0.557	0.302	0.468
1.0	0.437	0.573	0.364	0.442
3.0	0.497	0.570	0.389	0.416

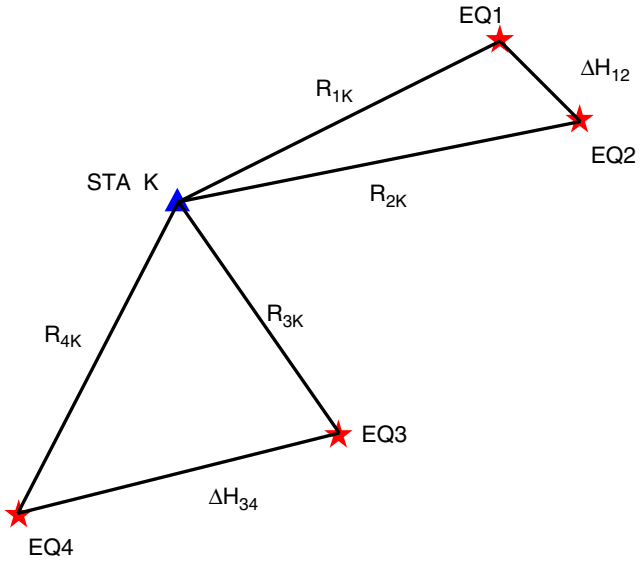

Figure 7. Example of the components used to compute the closeness index. The stars are the earthquakes, and the triangle is the station. The color version of this figure is available only in the electronic edition.

where b_1 is the $\sigma_{\Delta\xi_r}$ at CI = 0, and b_2 is the asymptotic value of $\sigma_{\Delta\xi_r}$ at CI $\gg b_3$. In large-CI bins, $\sigma_{\Delta\xi_r}$ is similar for all spectral periods. We thus impose the constraint that b_2 is identical for all periods. We also constrain n and b_3 to be period independent. With these as constraints, b_2, b_3, n, and six period-dependent b_1 are estimated simultaneously using ordinary least-squares fitting of the binned $\sigma_{\Delta\xi_r}$ values. The resulting coefficients are listed in Table 4, and the fitted model for each spectral period is shown in Figure 9 as the solid line.

The resulting standard deviation for the different spectral periods are compared in Figure 10. Overall, there is a trend for a smaller standard deviation at small CI values at the short spectral periods than at the long spectral periods. This indicates that there is greater repeatability of path effects at short periods than at long periods. One possible explanation for the larger variability at long period is that it reflects the radiation pattern effects on the ground-motion amplitude. If the earthquakes located in a small source volume have different focal mechanisms, then the variability for small CI values would include the effects of variability in the focal mechanisms. Because the focal mechanism effects are stronger at long periods than at short periods, we would expect greater variability in the long-period amplitude for a specific source-site combination.

As explained earlier, σ_0 is equal to b_1 (the $\sigma_{\Delta\xi_r}$ at CI = 0), multiplied by σ_r to remove the normalization. Therefore, for a given spectral period T, we have

$$\sigma_0 = b_1 \cdot \sigma_r. \tag{16}$$

$\sqrt{\tau_P^2 + \sigma_0^2}$ is equal to b_2 (the $\sigma_{\Delta\xi_r}$ at large CI) multiplied by σ_r to remove the normalization. We have

$$\tau_P = \sqrt{b_2^2\sigma_r^2 - \sigma_0^2} = \sqrt{b_2^2 - b_1^2} \cdot \sigma_r. \tag{17}$$

The values of τ_P and σ_0, which were estimated from the TSMIP data, are listed in Table 5.

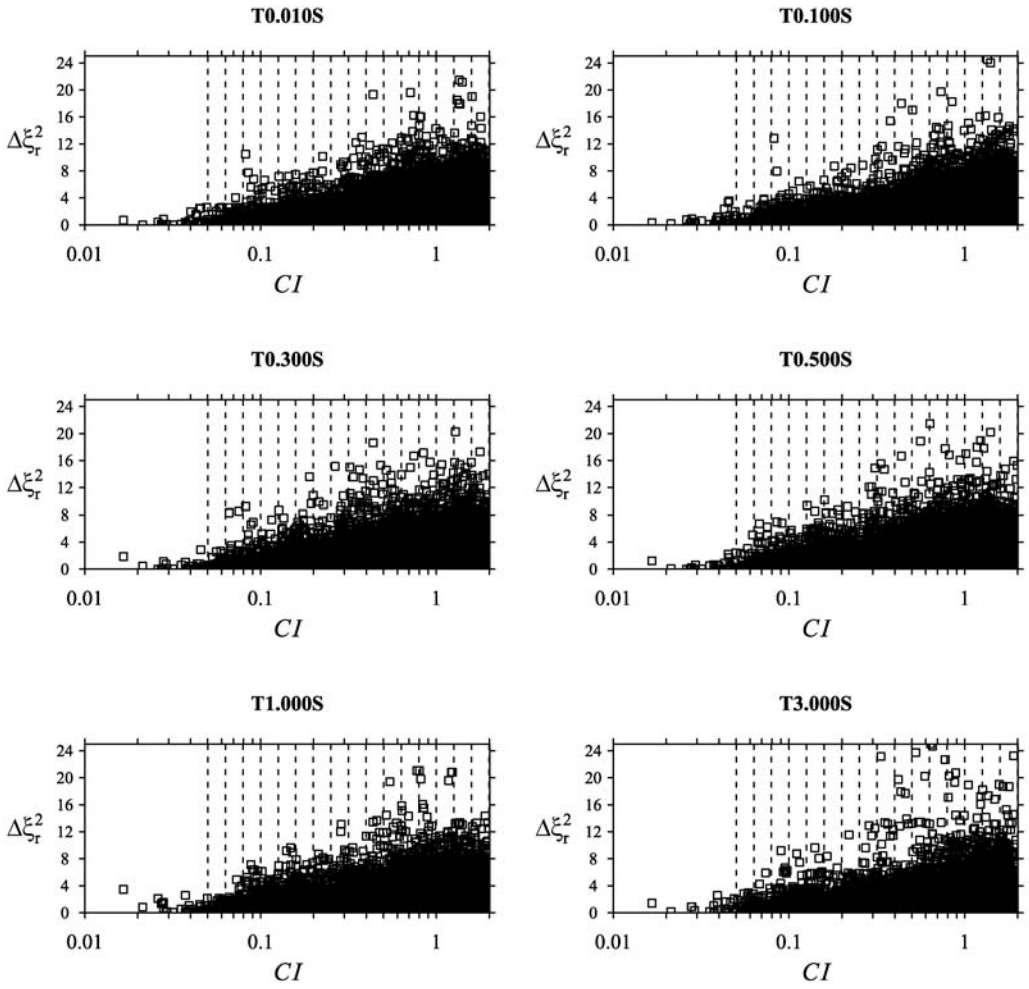

Figure 8. Values of $\Delta\xi_r^2$ shown as a function of the closeness index (CI). The vertical lines show the boundary of bins used for computing the binned standard deviations of$\Delta\xi_r$.

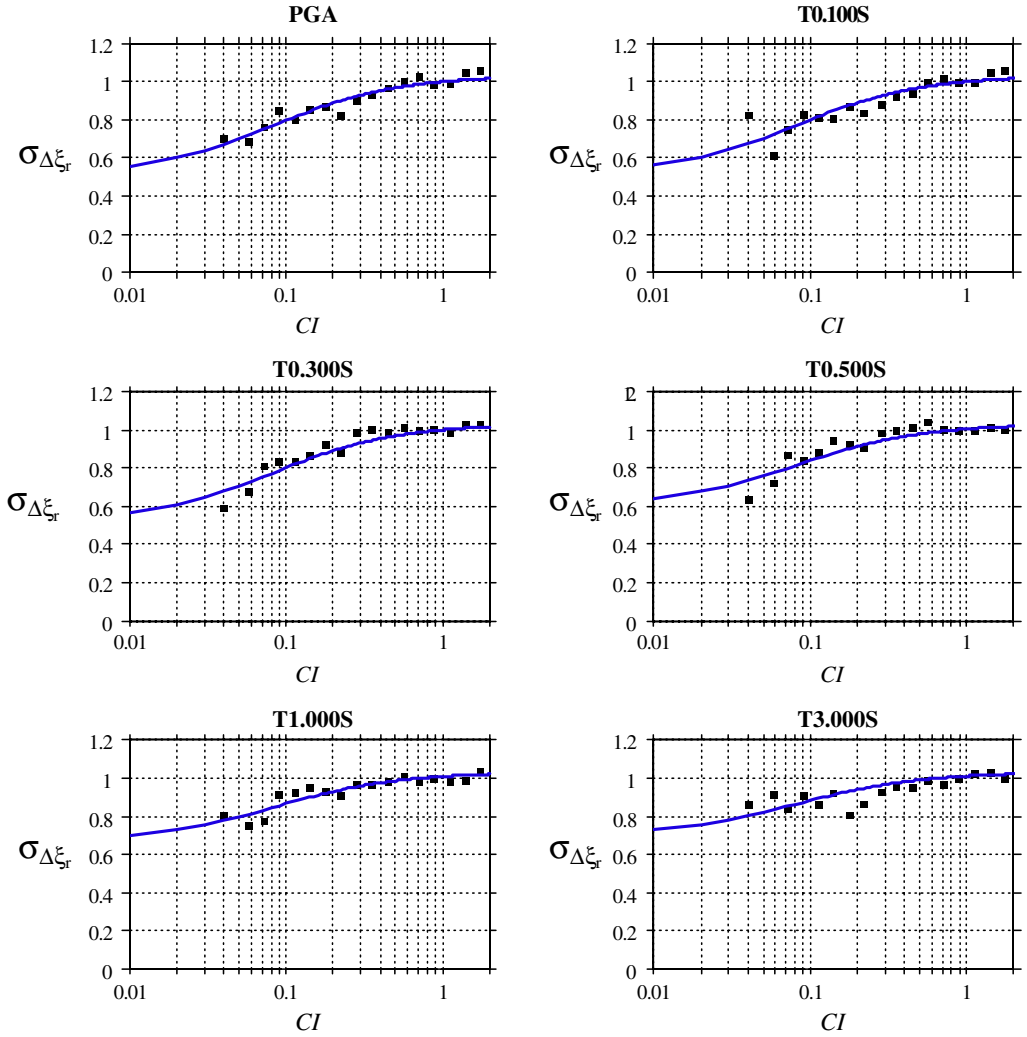

Figure 9. The binned $\sigma_{\Delta\xi_r}$ for PGA and spectral accelerations at periods of $T = 0.1$, 0.3, 0.5, 1.0, and 3.0 s. The solid line is the model fitted to binned $\sigma_{\Delta\xi_r}$. The color version of this figure is available only in the electronic edition.

Step 4: Estimating τ_{SR} and τ_{E0}

A procedure similar to that described previously is used for the estimation of τ_{SR} and τ_{E0}. The normalized difference between the event terms of the ith and the jth earthquakes is computed as

$$\Delta\eta_{ij} = \frac{\eta_{E_i} - \eta_{E_j}}{\sqrt{2} \cdot \tau_E}, \tag{18}$$

where τ_E is the event-term standard deviation estimated in Step 1 (Table 3). The values of $\Delta\eta_{ij}^2$ are shown as a function of ΔH in Figure 11. These $\Delta\eta_{ij}^2$ values are put into ΔH bins (defined by the vertical lines in Figure 11). The standard deviation of $\Delta\eta_{ij}$ ($\sigma_{\Delta\eta}$), computed as the square root of the mean $\Delta\eta_{ij}^2$, is shown in Figure 12 for each ΔH bin. The dependence of $\sigma_{\Delta\eta}$ on ΔH is similar for all six ground-motion parameters; there is no systematic period dependence. The standard deviation generally increases as ΔH increases from 3 to 100 km as expected.

The decrease in $\sigma_{\Delta\eta}$ as ΔH increases from 100 to 250 km is counterintuitive; we expect the source-location terms to be less repeatable (less correlated) as the separation between source-location increases. Our explanation for this trend is the shape of the island of Taiwan. Because the island is only about 100 km wide and 300 km long, the hypocenter separations of 100–250 km represent earthquake pairs in the

Table 4
Coefficients for the CI Dependence of the Standard Deviation $\sigma_{\Delta\xi_r}$

Spectral Period (s)	b_1	b_2	b_3	n
0.01 (*PGA*)	0.513	1.031	0.065	1.100
0.1	0.518	1.031	0.065	1.100
0.3	0.522	1.031	0.065	1.100
0.5	0.603	1.031	0.065	1.100
1.0	0.665	1.031	0.065	1.100
3.0	0.701	1.031	0.065	1.100

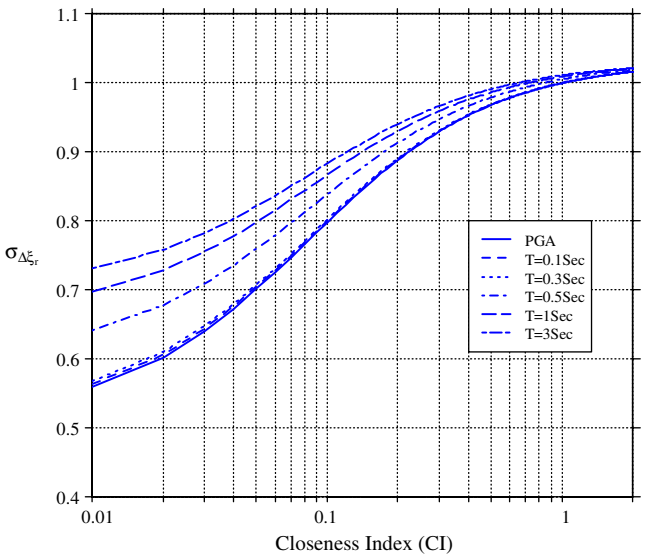

Figure 10. Comparison of the modeled dependence of $\sigma_{\Delta\xi_r}$ on the CI for the different spectral periods. The color version of this figure is available only in the electronic edition.

northeast (near 122° W, 24.5° N) and the southwest (near 120.5° W, 23.25° N). That is, for this set of earthquakes, we are only sampling two small regions where $\Delta H > 150$ km. The decrease in the standard deviation with increasing ΔH indicates that the average stress drops for earthquakes in these two small regions are more similar than expected for this range of separation. If there was a better azimuthal coverage of recorded earthquakes at large separation distances (for a larger area), then we expect that the observed trend would not be observed. While this trend is important for a model applicable to Taiwan, it is not considered to be representative of typical regions.

To develop a parametric model that is widely applicable, we limit the data to those with ΔH_{ij} less than 100 km; the dependence of $\sigma_{\Delta\eta}$ on ΔH is modeled using a form that is monotonically increasing with ΔH:

$$\sigma_{\Delta\eta}(\Delta H) = b_4 + b_5 \tanh(b_6 \Delta H). \tag{19}$$

Coefficient b_4 is the value of $\sigma_{\Delta\eta}$ at $\Delta H = 0$, and $(b_4 + b_5)$ is the asymptotic value at large ΔH. The estimated coefficients are listed in Table 6, and the resulting model, applicable to all periods, is shown as the solid line in Figure 12.

The standard deviation of the unexplained random event-term variation (τ_{E0}) is equal to the estimated b_4 multiplied by τ_E to remove the normalization:

$$\tau_{E0} = b_4 \cdot \tau_E. \tag{20}$$

The standard deviation $\sqrt{\tau_{SR}^2 + \tau_0^2}$ is equal to $b_4 + b_5$ the $\sigma_{\Delta\eta}$ at large ΔH multiplied by τ_E to remove the normalization, so

$$\tau_{\mathrm{SR}} = \sqrt{(b_4 + b_5)^2 \tau_E^2 - \tau_{E0}^2} = \sqrt{(b_4 + b_5)^2 - b_4^2} \cdot \tau_E. \tag{21}$$

The values of τ_{SR} and τ_{E0}, which were estimated from the TSMIP data, are listed in Table 5.

Results from Previous Studies

Estimates of single-site and single-path standard deviations have been developed by Chen and Tsai (2002), Atkinson (2006), and Morikawa *et al.* (2008). There has not been a consistent notation used in these previous studies. To facilitate comparisons between the results from this study and the previous studies, Table 7 shows how the notation used in this study is related to the notation used in previous studies, with the last column showing the notation proposed by Al Atik *et al.* (2010). The estimated variance components from the previous studies are summarized for PGA in Table 8. The results of the three previous studies are described next.

Single-Site Standard Deviations

Chen and Tsai (2002) fitted a different PGA data set from Taiwan to a ground-motion prediction model with both event and site terms. They found that σ_{SS} was 14% lower than σ_T (Table 9). The data set used by Chen and Tsai included data from sites located on rock and firm soil site conditions, but their regression model did not include a term

Table 5
Variance Components Estimated from the TSMIP Data

	Components of Variability					Mixture of Sites and Paths	Single Site	Single Path
Spectral Period (s)	τ_{SR}	τ_{E0}	τ_S	τ_P	σ_0	σ_T	σ_{SS}	σ_{SP}
0.01 (PGA)	0.254	0.247	0.259	0.401	0.230	0.637	0.583	0.337
0.1	0.282	0.273	0.353	0.410	0.238	0.710	0.616	0.363
0.3	0.272	0.264	0.280	0.423	0.249	0.680	0.620	0.362
0.5	0.286	0.277	0.302	0.392	0.282	0.695	0.626	0.396
1.0	0.323	0.314	0.364	0.348	0.294	0.737	0.640	0.430
3.0	0.367	0.357	0.389	0.315	0.292	0.773	0.668	0.461

$\sigma_T = \sqrt{\tau_{SR}^2 + \tau_{E0}^2 + \tau_S^2 + \tau_P^2 + \sigma_0^2}$;$\sigma_{SS} = \sqrt{\tau_{SR}^2 + \tau_{E0}^2 + \tau_P^2 + \sigma_0^2}$;$\sigma_{SP} = \sqrt{\tau_{E0}^2 + \sigma_0^2}$

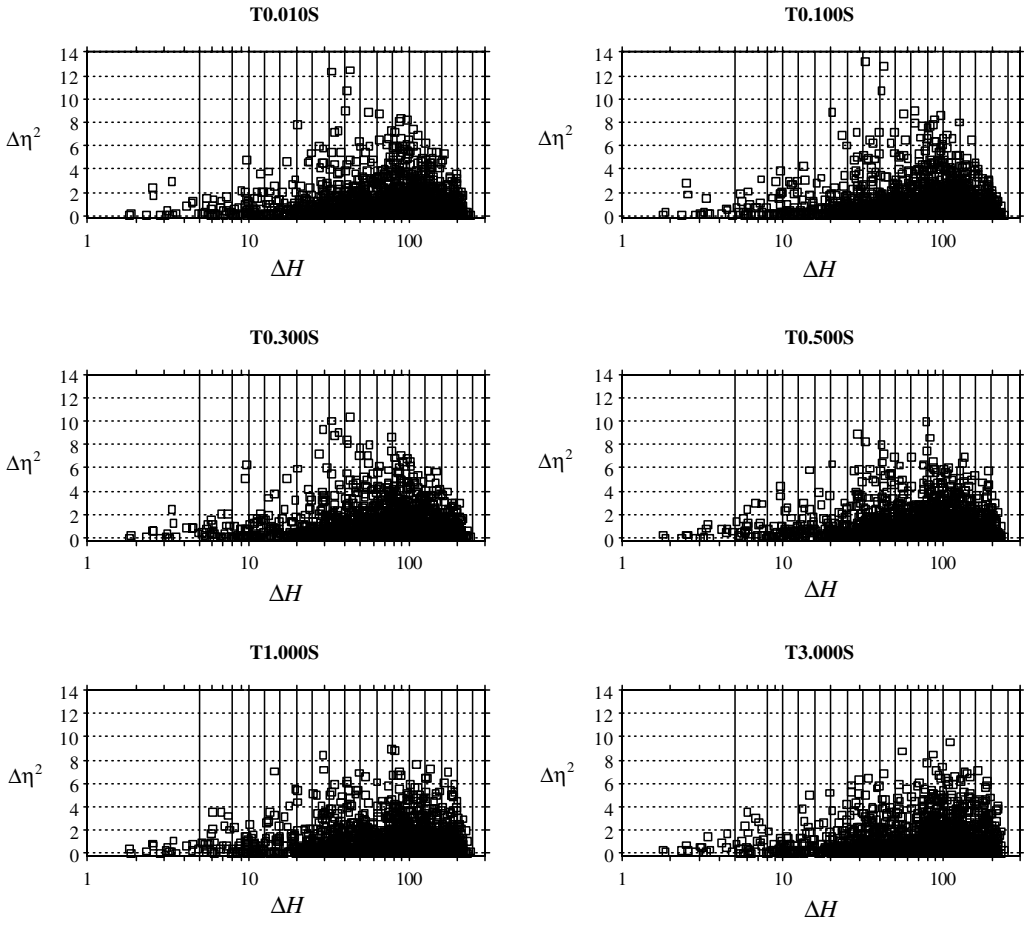

Figure 11. Values of $(\Delta\eta)^2$ as a function of ΔH. The vertical lines show the boundary of bins used for computing the binned standard deviations of $\Delta\eta$.

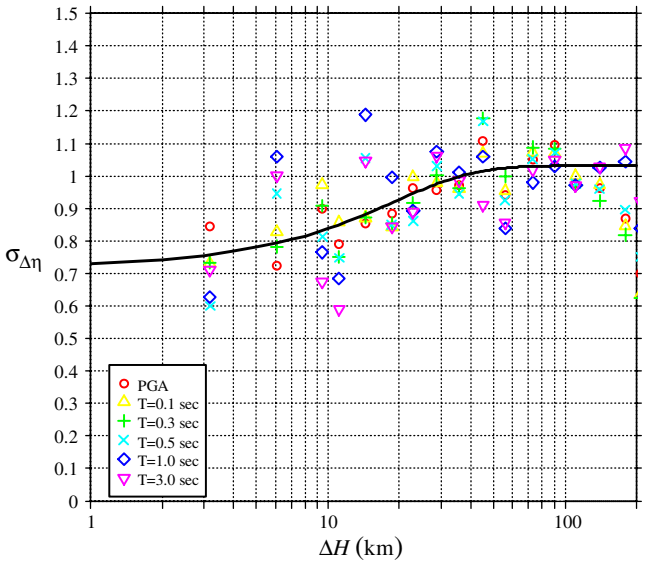

Figure 12. Dependence of $\sigma_{\Delta\eta}$ on hypocenter separation, ΔH, for PGA and spectral accelerations at periods of $T = 0.1$, 0.3, 0.5, 1.0, and 3.0 s. The color version of this figure is available only in the electronic edition.

for the site class or for the V_{S30}. As a result, their estimate of σ_T includes the additional variability due to differences in the average site response for different site class or for different V_{S30}. If a term had been included in their regression to model this average site response, then there would be slightly less of a reduction in the standard deviation for the single-site condition.

Atkinson (2006) used strong-motion data from stations in the Los Angeles basin that had recorded multiple earthquakes to estimate σ_{SS}. Unlike Chen and Tsai (2002), Atkinson corrected the data to a reference V_{S30} condition before conducting the regression analysis. She found that σ_{SS} was 13% lower for PGA and 8% lower for $T = 1$ s spectral acceleration when compared with the σ_T of the whole data set (Table 9).

Table 6
Coefficients for the ΔH Dependence of the Standard Deviation $\sigma_{\Delta\eta}$

Coefficient	Estimate
b_4	0.718
b_5	0.313
b_6	0.040

Table 7
Correspondence of Notation for Components of Variability from Previous Studies and This Study

Standard Deviation	This Study	Chen & Tsai (2002)	Atkinson (2006)	Morikawa *et al.* (2008)	Al Atik *et al.* (2010)
Event term	τ_E	σ_E	–	τ (no correction)	τ
Source-location term	τ_{SR}	–	–	–	τ_{L2L}
Event term (accounted for source-location term)	τ_{E0}	–	–	τ (applied correction)	τ_0
Residual (accounted for event terms)	σ	–	–	σ (no correction)	ϕ
Site term	τ_S	σ_S	–	–	ϕ_{S2S}
Residual (accounted for event and site terms)	σ_r	σ_r	–	–	ϕ_{SS}
Path term	τ_P	–	–	–	ϕ_{P2P}
Residual (accounted for event, site, and path terms)	σ_0	–	–	σ (applied correction)	–
Total (for mixture of sites and paths)	σ_T	–	σ_{reg}	δ	σ_{tot}
Total (for single site)	σ_{SS}	–	σ_I (weighted average)	–	σ_{SS}
Total (for single path)	σ_{SP}	–	σ_{ie}	$\varepsilon_{\eta_i} = \frac{\eta_i}{\tau}$ (applied correction)	–

Table 8
Examples of the Decomposition of the Standard Deviation of PGA into Different Components from Previous Studies*

	This Study	Chen & Tsai (2002)	Atkinson (2006)†	Morikawa *et al.* (2008)	Chiou and Youngs (2008)	Chiou and Youngs (2008)
Region	Taiwan	Taiwan	Southern CA	Japan	Worldwide (**M** 4)	Worldwide (**M** 7)
Magnitude range	3.9–6.3	(Not Specified)	3.0–7.1	5.5–7.5	4	7
τ_E	0.344	0.490	–	–	0.343	0.255
τ_{SR}	0.254	–	–	–	–	–
τ_{E0}	0.247	–	–	0.25	–	–
σ	0.518	–	–	–	0.598	0.455
τ_S	0.259	0.371	0.354	–	–	–
σ_r	0.449	0.397	–	–	–	–
τ_P	0.401	–	0.457	–	–	–
σ_0	0.230	–	–	0.28	–	–
σ_T	0.637	0.731	0.711	0.78	0.689	0.522
σ_{SS}	0.565	0.631	0.617	–	–	–
σ_{SP}	0.337	–	0.414	0.36	–	–

*For reference, the standard deviations from this study and from the Chiou and Youngs (2008) global model for two magnitudes are also shown.
†Atkinson (2006) presents several alternative sets of results. The results listed here are the Boore–Joyner–Fumal (Boore *et al.*, 1997) corrected set of results.

Single-Path Standard Deviations

Atkinson (2006) also found that if she restricted the data to a single-source region so that the paths are similar, then the standard deviation for PGA was 33% lower than σ_T compared with a 13% reduction for the single-site condition (Table 10).

Morikawa *et al.* (2008) used strong-motion data from Japan to estimate σ_{SP}. The data were grouped into six small regions, and region-specific terms were included in the regression model. They found, for PGA, σ_{SP} was 54% smaller than the estimated σ_T (Table 10).

Comparison of Current Results with Results from Previous Studies

The reduction in total standard deviation from the current study is compared with the results from the previous studies in Table 9 for the single-site condition. Our results are

Table 9
Comparison of Single-Site Standard Deviations (σ_{SS}) as a Fraction of σ_T

Spectral Period (s)	This Study	Chen & Tsai (2002)	Atkinson (2006)
PGA	$0.91\sigma_T$	$0.86\sigma_T$	$0.87\sigma_T$
0.1	$0.87\sigma_T$	–	–
0.3	$0.91\sigma_T$	–	$0.91\sigma_T$
0.5	$0.90\sigma_T$	–	–
1.0	$0.87\sigma_T$	–	$0.92\sigma_T$
3.0	$0.86\sigma_T$	–	$0.93\sigma_T$

Table 10
Comparison of Single-Path Standard Deviations (σ_{SP}) as a Fraction of σ_T

Spectral Period (s)	This Study	Atkinson (2006)	Morikawa *et al.* (2008)
PGA	$0.53\sigma_T$	$0.67\sigma_T$	$0.46\sigma_T$
0.1	$0.51\sigma_T$	–	$0.38\sigma_T$
0.3	$0.53\sigma_T$	$0.68\sigma_T$	$0.44\sigma_T$
0.5	$0.57\sigma_T$	–	$0.45\sigma_T$
1.0	$0.58\sigma_T$	$0.67\sigma_T$	$0.47\sigma_T$
3.0	$0.60\sigma_T$	–	$0.47\sigma_T$

consistent with the two previous studies, showing a 9% to 14% reduction in the total standard deviation.

The results for the single-path condition (Table 10) are not as consistent as for the single-path condition: σ_{SP} is about 45% smaller than σ_T, compared with 33% found by Atkinson (2006) and 55% reduction found by Morikawa *et al.* (2008). The differences in the reductions are likely due to data set differences. For example, the Japanese data used by Morikawa *et al.* (2008) tend to have larger σ_T than the other regions (Table 8). Comparing the σ_{SP} values for PGA, we found similar results for the three studies: 0.36 for Morikawa *et al.* (2008), 0.414 for Atkinson (2006), and 0.337 for this study. The large differences in the reduction factors (Table 10) reflect the large differences in σ_T, not differences in σ_{SP}.

Conclusions and Discussions

Recent analysis of strong-motion data sets has shown that the aleatory variability of a ground-motion prediction model can be reduced significantly if the ergodic assumption is removed. Using a large data set from Taiwan, this study estimates the variance components of ground-motion amplitude and uses the results to quantify the reduction in the total standard deviation due to the removal of ergodic assumption. For peak ground acceleration and spectral accelerations at the periods of 0.1, 0.3, 0.5, 1.0, and 3.0 s, we find that the standard deviations at a single site are 9%–14% smaller than the total standard deviation with ergodic assumption, whereas the standard deviations for a single path (i.e., for a unique combination of site and source location) are 39%–47% smaller.

Such reduction in the aleatory variability, however, comes at the price of increased epistemic uncertainty in the predicted mean amplitude. If the single-path standard deviation is used for the aleatory variability, both deterministic and probabilistic seismic hazard analyses will need to consider the increased epistemic uncertainty in the mean amplitude for each path. For regions with large ground-motion data sets, such as Taiwan, empirical estimates of source-site-specific mean amplitude can be made with only a small increase in the epistemic uncertainty. For regions with little or no data to constrain the path-specific mean amplitude, the additional epistemic uncertainty will be large.

There is an additional complexity for application to probabilistic seismic hazard analysis (PSHA) without ergodic assumption. The spatial correlation in the epistemic uncertainty will need to be addressed if estimates of the epistemic uncertainty in the hazard (e.g., hazard fractiles) are desired. The focus in this study is on the standard deviation of error components; we have not parameterized our results in terms of correlation coefficients. Models for the spatial correlation of the path terms and the source-location terms can be developed in the future and used in PSHAs to allow for the modeling of the spatial correlations of the epistemic uncertainties in the mean amplitude. A method for conducting nonergodic PSHA that includes spatially correlated logic trees is given by Walling (2009). For the mean hazard, the spatial correlation is not an issue.

Data and Resources

The strong-motion data used in this study were provided by the Central Weather Bureau, Taiwan (R.O.C) via Geophysical Database Management System found at http://gdms.cwb.gov.tw (last accessed April 2011).

Some of the analysis and graphics were produced using R found at http://www.r-project.org/ (last accessed April 2011).

Acknowledgments

This research was sponsored by the Sinotech Engineering Consultants, Inc., Pacific Gas & Electric Company, and the California Department of Transportation. We would like to thank Julian Bommer and two anonymous reviewers for their useful review comments that improved the paper. This research was supported by the Taiwan Earthquake Research Center (TEC) funded through the National Science Council (NSC) with grant number NSC96-2119-M008-004. The TEC contribution number for this article is 0075. This work was partially funded by the PG&E/DOE cooperative agreement "Development and Verification of an Improved Model for Extreme Ground Motions Produced by Earthquakes" (DOE Award Number DE-FC28-05RW 12358).

References

Abrahamson, N. A., and R. R. Youngs (1992). A stable algorithm for regression analyses using the random effects model, *Bull. Seismol. Soc. Am.* **82,** 505–510.

Abrahamson, N. A., G. Atkinson, D. Boore, Y. Bozorgnia, K. Campbell, B. Chiou, I. M. Idriss, W. Silva, and R. Youngs (2008). Comparisons of the NGA ground-motion relations, *Earthquake Spectra* **24,** no. 1, 45–66.

Al Atik, L., N. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Let.* **81,** 794–801.

Anderson, J. G., and J. N. Brune (1999). Probabilistic seismic hazard assessment without the ergodic assumption, *Seismol. Res. Let.* **70,** 19–28.

Anderson, J. G., and Y. Uchiyama (2010). A methodology to improve ground motion prediction equations by including path corrections, *Bull. Seismol. Soc. Am.*, in review.

Aoi, S., K. Obara, S. Hori, K. Kasahara, and Y. Okada (2000). New strong-motion observation network: KiK-net, abstract no. S71A-05, *Eos Trans. AGU*, Fall Meeting Supplement, **81,** 329

Atkinson, G. M. (2006). Single-station sigma, *Bull. Seismol. Soc. Am.* **96,** 446–455.

Boatwright, J., H. Bundock, J. H. Luetgert, D. H. Oppenheimer, L. M. Baker, J. B. Fletcher, J. R. Evans, C. D. Stephens, K. A. Fogleman, L. Gee, D. S. Dreger, D. L. Carver, V. M. Grazier, C. W. Scrivner, and M. K. McClaren (1999). Implementing ShakeMap in the Bay area, *Eos Trans. AGU* **80,** F700.

Boore, D., W. Joyner, and T. Fumal (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work, *Seismol. Res. Lett.* **68,** 128–153.

Boore, D. M., J. Watson-Lamprey, and N. A. Abrahamson (2006). GMRotD and GMRotI: Orientation-independent measures of ground motion, *Bull. Seismol. Soc. Am.* **96,** 1202–1511.

Chen, Y-H., and C-C. P. Tsai (2002). A new method for estimation of the attenuation relationship with variance components, *Bull. Seismol. Soc. Am.* **92,** 1984–1991.

Chiou, B., and R. R. Youngs (2008). An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthquake Spectra* **24,** no. 1, 173–215.

Chiou, B., R. Darragh, N. Gregor, and W. Silva (2008). NGA project strong-motion database, *Earthquake Spectra* **24,** no. 1, 23–44.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*, revised edition, John Wiley and Sons, New York.

Joyner, W. B., and D. M. Boore (1993). Methods for regression analysis of strong-motion data, *Bull. Seismol. Soc. Am.* **83,** 469–487.

Kinoshita, S. (1998). Kyoshin Net (K-NET), *Seismol. Res. Lett.* **69,** 309–332.

Lee, C.-T., and B.-R. Tsai (2008) Mapping V_{S30} in Taiwan, *Terr. Atmos. Ocean. Sci.* **19,** no. 6, 671–682.

Lin, P.-S. (2009). Ground-motion attenuation relationship and path-effect study using Taiwan data set, *Ph.D. Thesis*, Nation Central University, Taiwan (in Chinese).

Liu, K.-S., T.-C. Shin, and Y. B. Tsai (1999). A free-field strong-motion network in Taiwain: TSMIP, *Terr. Atmos. Ocean. Sci.* **10,** no. 2, 377–396.

Morikawa, N., T. Kanno, A. Narita, H. Fujiwara, T. Okumura, Y. Fukushima, and A. Guerpinar (2008). Strong motion uncertainty determined from observed records by dense network in Japan, *J. Seismol.* **12,** no. 4, 529–546.

Restrepo-Velez, L. F., and J. J. Bommer (2003). An exploration of the nature of the scatter in ground motion prediction equations and the implications for seismic hazard assessment, *J. Earthquake Eng.* **7,** no. S1, 171–199.

Shin, T. C., and T. L. Teng (2001). An overview of the 1999 Chi-Chi, Taiwan, Earthquake, *Bull. Seismol. Soc. Am.* **91,** 895–913.

Strasser, F. O., N. A. Abrahamson, and J. J. Bommer (2009). Sigma: Issues, insights, and challenges, *Seismol. Res. Let.* **80,** no. 1, 40–56.

Walling, M. A. (2009). Non-Ergodic Probabilistic Seismic Hazard Analysis and Simulation Spatial of Variation in Ground Motion, *Ph.D. Thesis*, University of California, Berkeley, California.

Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B. Worden (1999). TriNet "ShakeMaps": Rapid generation of peak ground motion and intensity maps for earthquakes in southern California, *Earthquake Spectra* **15,** 537–555.

Youngs, R. R., N. A. Abrahamson, F. Makdisi, and K. Sadigh (1995). Magnitude dependent variance of peak ground acceleration, *Bull. Seismol. Soc. Am* **85,** 1161–1176.

Sinotech Engineering Consultants, Inc.
Taipei, Taiwan
(P.-S.L., C.-T.C.)

California Department of Transportation
Sacramento, California
(B.C.)

Pacific Gas and Electric Company
San Francisco, California
(N.A.)

United States Geological Survey
Golden, Colorado
(M.W.)

Institute of Applied Geology
National Central University
Jhongli, Taiwan
(C.-T.L.)

Manuscript received 25 September 2009