Bull Earthquake Eng
DOI 10.1007/s10518-012-9368-5

ORIGINAL RESEARCH PAPER

On the influence of site conditions and earthquake magnitude on ground-motion within-earthquake correlation: analysis of PGA data from TSMIP (Taiwan) network

Vladimir Sokolov · Friedemann Wenzel · Kuo-Liang Wen · Wen-Yu Jean

Received: 17 February 2012 / Accepted: 20 July 2012

Abstract We analysed the within-earthquake correlation of ground motion using the strong-motion records accumulated by the TSMIP (Taiwan Strong Motion Instrumentation Program) network in Taiwan during 1993–2009. Two ground-motion prediction equations, which were recently developed for peak ground acceleration (PGA) in the region and based on moment and local magnitude and hypocentral distance, were used for the calculation and analysis of ground-motion residuals. We also used the database containing shear-wave velocity data averaged for the top 30 m of the soil column (Vs30) for the TSMIP stations. We showed that the within-earthquake correlation may vary significantly depending on site classes, gross geological features of the area, and magnitude of earthquakes, records of which dominate the analysed dataset. On the one hand, there is a prominent correspondence between the within-earthquake correlation of PGA residuals and spatial correlation of Vs30 values, which was estimated for particular geological structures (e.g., sedimentary filled basins and large plain areas). On the other hand, the high level of ground-motion correlation (or significant non-random component of residuals) may be caused by the joint influence of soft surface soil and thick sediments and by the path or azimuthal effects. The point-source approximation of extended fault and neglected hanging- and foot-wall effects may also result in non-random residuals. The application of empirical correction factors, which consider the magnitude of earthquakes, source-to-site distance and Vs30 value for given stations, allows for the effective reduction in the level of within-earthquake correlation, as well as the within-earthquake standard deviation. The results of the analysis may be used in practical estimates of seismic

V. Sokolov (✉) · F. Wenzel
Geophysical Institute, Karlsruhe Institute of Technology (KIT), Hertzstrasse 16, 76187 Karlsruhe, Germany
e-mail: Vladimir.Sokolov@kit.edu

K.-L. Wen · W.-Y. Jean
National Center for Research on Earthquake Engineering, Taipei, Taiwan, ROC

K.-L. Wen
Institute of Geophysics, National Central University, Jhongli, Taiwan, ROC

Published online: 09 August 2012

hazard, damage and loss for spatially distributed structures (portfolios, lifelines) in Taiwan, as well in other regions with similar geological characteristics.

Keywords Ground-motion correlation · Seismic hazard and risk assessment · TSMIP network

1 Introduction

Throughout the last decade, the interest in studies of ground-motion correlation has grown significantly. This correlation is related to the ground-motion variability, or residuals between observations and the results of modelling, for different earthquakes (between-earthquake correlation) and different locations (within-earthquake spatial correlation). The correlation of ground-motion residuals reflects a non-random component in the residuals, which is caused by factors not accounted for by the ground-motion model and which therefore constitutes epistemic uncertainty (attributable to incomplete knowledge) that may be reduced, in principle. The parameters of the probability distribution function for the loss to a portfolio (e.g., fractiles and standard deviation) during an earthquake, which are very important for decision-making and mitigation activities, are affected by ground-motion uncertainty and correlation. The proper treatment of ground-motion correlation is essential for the estimation of seismic hazard, damage and loss for widely located building assets (portfolios) and spatially distributed structures (lifelines). Modelling of the correlation is also useful for the development of rapid post-earthquake ground-motion estimation tools (e.g., ShakeMap).

The influence of ground-motion uncertainty and correlation on characteristics of loss distribution for portfolios and lifelines has been analysed in several studies (see, for example, reviews in Sokolov and Wenzel 2011a,b). Different models of within-earthquake correlation were applied in loss estimations by Park et al. (2007), Goda and Hong (2008b), Goda and Atkinson (2009), Molas et al. (2006), Crowley et al. (2008a) and Sokolov and Wenzel (2011a,b). The general findings may be summarised as follows. For every particular earthquake scenario, the between-earthquake variability changes the level of ground motion at all locations in a similar way, i.e., the level will be lower or higher than the median estimates made by a predictive equation. The larger level of ground motion at all locations will cause greater damage everywhere and higher total loss values and vice versa. The within-earthquake variability leads to fluctuations in ground motion from one location to another, stronger than the median estimates at one site and weaker at another site. Even if the large level of ground motion causes great damage at some locations, the damage at other locations may be small. Thus, a higher between-earthquake correlation and a within-earthquake spatial correlation results in a larger variation in losses to a portfolio and a higher probability of extreme loss values; this would lead to a high variability in total loss. Sokolov and Wenzel (2011a) performed sensitivity analysis of the impact of variation of ground-motion correlation components on probabilistic estimates of seismic loss and damage for extended objects and showed that the influence of correlation depends, on one hand, on level of hazard and probability level of interest (return period).

The modern ground-motion prediction equations (GMPEs) allow for the recognition of the between-earthquake correlation because the equations specify the between-earthquake and within-earthquake components of variability (e.g., Boore et al. 1997; Douglas 2003, 2006; Tsai et al. 2006; Abrahamson et al. 2008). The within-earthquake correlation should be evaluated for a given area empirically, and several correlation models, in which the correlation

depends uniquely on intersite separation distance, were obtained for regions characterised by a dense observation of records from numerous earthquakes, e.g., California, Japan, and Taiwan (see Sokolov et al. 2010; Goda 2011; Esposito and Iervolino 2011, for collection of recent works).

The results reported by these studies reveal a different rate of decay of the within-earthquake correlation with separation distance. Among the reasons for this difference, it is worth noting different ground-motion prediction models as well as different ground-motion parameters that were used for the analysis. It has been shown that the difference is related to the frequency content of ground motion (Goda and Hong 2008a; Baker and Jayaram 2008). In contrast, the difference may be caused by regional peculiarities (Goda and Atkinson 2009; Jayaram and Baker 2009). Analysis of the within-earthquake correlation in Taiwan (Sokolov et al. 2010) indicates that the correlation structure is highly dependent on local geology and on peculiarities of propagation path (azimuth-dependent attenuation). Thus, a single generalised model of the correlation may not be adequate in some cases.

To reduce random variability in modelled ground motion and therefore reduce the uncertainty in estimated losses, it is necessary to find appropriate and acceptable approaches to reduce the standard deviation (sigma) associated with GMPEs (e.g., Alatik et al. 2010; Strasser et al. 2009). Atkinson (2006) and Morikawa et al. (2008) showed that the total standard deviation of ground-motion prediction may be reduced by the application of a single site-specific model or a region-specific correction factor, which is determined by grouping ground-motion data at specific stations of a dense strong-motion array. Tsai et al. (2006), Anderson and Uchiyama (2011), and Lin et al. (2011) have found that a reduction in the total standard deviation could be achieved if the path effect could be specified.

The characteristics of ground-motion variability (i.e., standard deviation, between-earthquake and within-earthquake correlation) are used together with corresponding ground-motion prediction equations in strong ground-motion modelling. Thus, the proper selection of the correlation models is a matter of great importance. In most studies (see Sokolov et al. 2010; Goda 2011; Esposito and Iervolino 2011, for collection of recent works), the within-earthquake correlation models were developed by considering all ground-motion data without classifying the data according to specific soil conditions or narrow geographical areas.

The goal of this study was to investigate to what extent the within-earthquake correlation of ground-motion residuals (peak amplitudes of ground acceleration, PGA) in Taiwan depends on array- or site-specific grouping and corresponding corrections, which allow for a reduction in the standard deviation associated with existing GMPEs. Compared with our previous work (Sokolov et al. 2010), in this study, we used an extended ground-motion database, and we considered the dependence of ground-motion correlation on the shear-wave velocity averaged for the top 30 m of the soil column (Vs30) and earthquake magnitude. The database includes strong-motion records accumulated by the TSMIP (Taiwan Strong Motion Instrumentation Program) network from 54 shallow earthquakes ($M_L > 5.0$, focal depth <30 km, more than 7,000 records) that occurred between 1993 and 2009. The correlation structure was investigated for residuals obtained using two ground-motion models recently developed for Taiwan. Recommendations for practical estimates of seismic hazard, damage and loss for widely located building assets and spatially distributed structures in Taiwan, as well in other regions with similar geological characteristics, are provided.

2 Ground-motion correlation

2.1 Basic definitions

The ground-motion parameter Y at n locations during m earthquakes can be expressed as

$$\log Y_{i,j} = f(e_i, p_{i,j}, s_{i,j}) + \eta_i + \varepsilon_{i,j} \quad i = 1, 2, ..., m; \quad j = 1, 2, ..., n; \tag{1}$$

where e_i are properties of the earthquake source; $p_{i,j}$ are the properties of propagation path; $s_{i,j}$ are the properties of site location j during earthquake i; and f is a suitable function that describes the dependence of the mean value of the ground-motion parameter $\log \overline{Y}_{i,j}$ on magnitude, distance, local site conditions, etc, i.e., $\log \overline{Y}_{i,j} = f(e_i, p_{i,j}, s_{i,j})$. Usually, an appropriate ground-motion prediction equation (GMPE) is used as the function f. The random variables η_i and $\varepsilon_{i,j}$ represent the between-earthquake and within-earthquake components of variability (independent and normally distributed with variances σ_η^2 and σ_ε^2), respectively. The value of η_i is common to all sites during a particular earthquake i, and the value of $\varepsilon_{i.j}$ depends on the site j. Assuming the independence of the two random terms, the total variance σ_T^2 is given by $\sigma_T^2 = \sigma_\eta^2 + \sigma_\varepsilon^2$ (e.g., Brillinger and Preisler 1984, 1985; Abrahamson and Youngs 1992; Joyner and Boore 1993; Boore et al. 1997; see also Strasser et al. 2009).

Let us consider two recording sites separated by distance Δ. The relationship between ground-motion residuals for a randomly selected earthquake i at the sites j and k (i.e., between $\eta_i + \varepsilon_{i,j}$ and $\eta_i + \varepsilon_{i,k}$ values) can be expressed as the total correlation coefficient $\rho_T(\Delta)$ (e.g., Park et al. 2007; Goda and Hong 2008a,b)

$$\rho_T(\Delta) = \frac{\sigma_\eta^2 + \rho_\varepsilon(\Delta)\sigma_\varepsilon^2}{\sigma_T^2} = \rho_\eta + \rho_\varepsilon(\Delta)\frac{\sigma_\varepsilon^2}{\sigma_T^2} = \rho_\eta + \rho_\varepsilon(\Delta)(1 - \rho_\eta) \tag{2}$$

where $\rho_\eta = \frac{\sigma_\eta^2}{\sigma_T^2}$ is the between-earthquake correlation (Wesson and Perkins 2001) and $\rho_\varepsilon(\Delta)$ represents the correlation coefficient between $\varepsilon_{i,j}$ and $\varepsilon_{i,k}$ (within-earthquake spatial correlation). Alternatively, the correlation coefficient $\rho_T(\Delta)$ is expressed as

$$\rho_T(\Delta) = 1 - \frac{\sigma_d^2}{2\sigma_T^2} \tag{3}$$

where σ_d^2 denotes the variance of $(\eta_i + \varepsilon_{i,j}) - (\eta_i + \varepsilon_{i,k})$ (Boore et al. 2003). The correlation coefficient $\rho_\varepsilon(\Delta)$, which by definition is given by

$$\rho_\varepsilon(\Delta) = \frac{COV[\varepsilon_{ij}, \varepsilon_{ik}]}{\sigma_\varepsilon^2} \tag{4}$$

where $COV(A, B)$ denotes the covariance of A and B, expressed as

$$\rho_\varepsilon(\Delta) = 1 - \frac{\sigma_d^2}{2\sigma_\varepsilon^2} \tag{5}$$

When estimating the correlated ground motion at k-sites, in addition to the median value of ground motion $\overline{Y}_{ij}$, we need to generate the standard normal variates (errors) of η_i and $\varepsilon_{i,j}$ (see Eq. 1). Descriptions of the procedure for the generation of the k-site random field of ground-motion error values that are spatially correlated may be found throughout the literature (e.g., Johnson 1987; Park et al. 2007).

2.2 Evaluation of within-earthquake correlation

Several methods that can be adopted to estimate the within-earthquake correlation of ground-motion residuals have been described in the literature (e.g., Boore et al. 2003; Wang and Takada 2005; Goda and Hong 2008a; Goda and Atkinson 2009, 2010; Baker and Jayaram 2008; Jayaram and Baker 2009). Let $Z(x)$ denote the total residuals between the log of the observed ground-motion parameter at a site x and that of the value predicted from the attenuation relationship, i.e., $Z(x_j) = \log Y_{ij} - \log \overline{Y}_{ij}$. Let us consider a pair of Z-values obtained for sites x_u separated by distance Δ, i.e., $z_u, \quad z_{u+\Delta}$. On one hand, the sample correlation coefficient may be evaluated for given separation distance bin (e.g.,Wang and Takada 2005; Goda and Hong 2008a). The correlation function is obtained by normalising the covariance function with the variance σ_z^2

$$\rho(z_u, z_{u+\Delta}) = \rho(\Delta) = \frac{COV(z_u\ , z_{u+\Delta})}{\sigma_z^2} \tag{6}$$

For a large number of recordings from an earthquake, the between-earthquake residual η becomes the mean residual for that event, i.e., $\hat{\eta} = \mu_Z$, where $\hat{\eta}$ is the estimate of the between-earthquake residual. Thus, the sample correlation coefficient is estimated as

$$\hat{\rho}(\Delta) = \frac{COV(\varepsilon_u, \varepsilon_{u+\Delta})}{\hat{\sigma}_\varepsilon^2} \tag{7}$$

where $\hat{\sigma}_\varepsilon$is the estimate of the within-earthquake standard deviation and $\varepsilon_u$and $\varepsilon_{u+\Delta}$ are the within-earthquake residuals at two sites separated by distance Δ.

The within-earthquake correlation can also be estimated via the sample semivariogram $\hat{\gamma}(\Delta)$ (Goovaerts 1997) as

$$\hat{\gamma}(\Delta) = \frac{1}{2}\sigma_d^2(\Delta) \tag{8}$$

where $\sigma_d^2(\Delta)$ is the variance of $z_u - z_{u+\Delta}$ that falls within a separation-distance bin represented by Δ. The relationship between the sample coefficient of the within-earthquake correlation $\hat{\rho}_\varepsilon(\Delta)$and semivariogram $\hat{\gamma}(\Delta)$ can be expressed as

$$\hat{\rho}_\varepsilon(\Delta) = 1 - \frac{\hat{\gamma}(\Delta)}{\hat{\sigma}_\varepsilon^2} \tag{9}$$

Boore et al. (2003), Baker and Jayaram (2008), Jayaram and Baker (2009), Goda and Hong (2008a), Goda and Atkinson (2009, 2010) and Esposito and Iervolino (2011) considered an approach based on sample semivariograms.

The estimates of $\hat{\rho}_\varepsilon(\Delta)$ or $\hat{\gamma}(\Delta)$ provide a set of experimental values for a finite number of separation distances Δ. A continuous function can be fitted based on these experimental values. The function $\rho_\varepsilon(\Delta)$ may be represented by an exponential decay function (e.g., Boore et al. 2003; Wang and Takada 2005; Goda and Hong 2008a)

$$\rho_\varepsilon(\Delta) = \exp(a\Delta^b) \tag{10}$$

where a and b are the model parameters. Wang and Takada (2005) suggested characterising the function $\rho_\varepsilon(\Delta)$by a single parameter: the so-called "correlation distance" R_C. The correlation distance is the site-to-site distance for which the correlation coefficient $\rho_\varepsilon(\Delta)$ decreases to $1/e = 0.368$.

3 The data

The strong-motion database, which was used in this study, includes records obtained at TSMIP stations from 54 shallow earthquakes ($M_L > 5.0$, focal depth <30 km, more than 7,000 records) that occurred between 1993 and 2009 (see Fig. 1a). The data were collected during the implementation of the Taiwan Strong Motion Instrumentation Program (TSMIP), which was conducted by the Seismological Observation Center of the Central Weather Bureau (CWB), Taiwan, R.O.C. (Liu et al. 1999). More than 650 digital free-field strong motion instruments were installed within seven arrays (Fig. 1b).

We considered the peak amplitudes of ground acceleration (PGA) in this study, and the following record selection criteria were used. First, the record should have clear P- and S-wave onsets. Second, only records for which the signal-to-noise ratio exceeded 2 were further processed. To check the ratio, the Fourier amplitude spectra of the strongest part of shaking (S-wave) and of the pre-event noise were calculated and compared. Third, the number of records that were obtained during a seismic event after applying the criteria mentioned above was at least 50. Thus, we focused on well-recorded events, which provided enough data to estimate event-dependent between-earthquake residuals (see next section).

The within-earthquake correlation, in principle, depends on the chosen ground-motion prediction model because the correlation describes the behaviour of residuals between observations. Several empirical ground-motion models were developed recently for Taiwan (see, for example, review in Cheng et al. 2007 and Sokolov et al. 2010). When selecting a ground-motion model, we considered the following. First, the database used for the development of the model should be as large as possible, covering a broad range of magnitudes and distances. Second, the models, which used distances to rupture (e.g., the closest distance to the rupture plane or horizontal distance to the surface projection of the rupture), were not considered because it is difficult to apply such models for numerous records and earthquakes.

In our study, we used two ground-motion models that were recently proposed for estimating the peak amplitudes of ground acceleration in Taiwan; the models use hypocentral distance as a parameter and include a description of the between-earthquake and the within-earthquake components of residuals. The first model, which is based on moment magnitude M_W and is hereinafter referred to as the MW2010 model, was developed by Sokolov et al. (2010) using the database from earthquakes that occurred between 1993 and 2004

$$\ln PGA = -3.07 + 0.83M_W - 1.33\ln[R + 0.15\exp(0.54M_W)] + 0.0023R \pm \sigma_T,$$
$$\sigma_T = 0.67; \quad \sigma_\eta = 0.39; \quad \sigma_\varepsilon = 0.55 \tag{11}$$

where PGA is measured in units of g and R is the hypocentral distance in kilometers. The second model, described by Tsai et al. (2006), is based on local magnitude M_L and records of earthquakes that occurred before 2002

$$\log_{10} PGA = 0.4063 + 0.7936\ M_L - 0.02146\ M_L^2$$
$$-1.7056\log_{10}(R + 5.7814 \times 10^{-0.05656M_L}) + 0.0004183\ R \pm \sigma_T$$
$$\sigma_T = 0.316; \quad \sigma_\eta = 0.176; \quad \sigma_\varepsilon = 0.263, \tag{12}$$

or in units of natural logarithm $\sigma_T = 0.727$; $\sigma_\eta = 0.405$; $\sigma_\varepsilon = 0.605$, The model is hereinafter is referred to as the ML2002 model. Both models describe the geometric mean of PGA of horizontal components, and the residuals for the models were estimated as $\ln(PGA_{OBS}/PGA_{MOD})$, where indexes *OBS* and *MOD* denote the observed data and the modelled values, respectively.

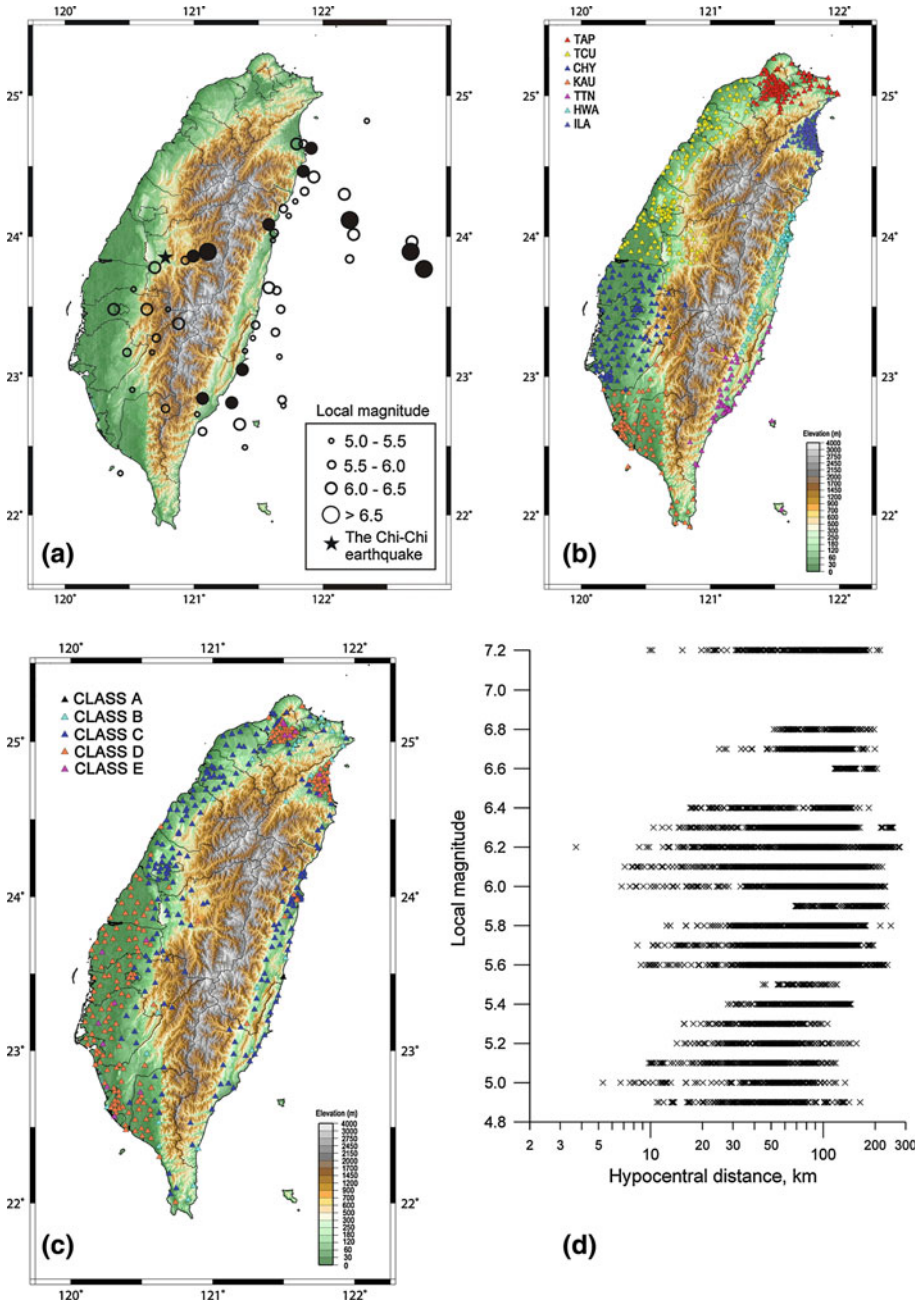

Fig. 1 The database used in this work. **a** Distribution of earthquake epicentres, records of which were analysed. *Filled circles* denote earthquakes, which compose the group of intermediate-to-large events. **b–c** TSMIP free-field strong-motion stations, arrays **b** and assigned site class **c** division. **d** Distribution of records versus local magnitude and epicentral distance

Both local magnitude and moment magnitude were used in our analysis. The earthquake catalogue collected by the CWB shows local magnitudes. Therefore, we considered several sources to determine the moment magnitude M_W values: the Harvard seismic catalogue http://www.seismology.harvard.edu/; the Institute of Earth Sciences, Academia Sinica http://tecws.earth.sinica.edu.tw (earthquake catalogue of Broadband Array in Taiwan for Seismology, BATS); and the regional relationships between seismic moment and local magnitude (Lin and Lee 2008) and between moment and local magnitudes (Campbell et al. 2002).

As shown in Fig. 1a, we considered shallow earthquakes that occurred beneath Taiwan and to east of the island under the ocean. Recent studies of the attenuation of earthquake ground motion for the Taiwan region (e.g., Wang 1998; Chen 1998; Sokolov et al. 2006; Chung et al. 2009) show that the attenuation of S waves is relatively stronger in Central Taiwan than in the northern part of island and offshore. Sokolov et al. (2006) developed models of Fourier amplitude spectra for Taiwan by considering the events beneath the island and offshore events separately. However, the generalised ground-motion models selected for this study (Eqs. 11–12) were developed without consideration of path or azimuthal effects.

The ground-motion models also do not consider different site conditions. The free-field strong-motion station sites in the Taiwan region were initially classified (Lee et al. 2001) using a scheme compatible with the 1994 and 1997 NEHRP provisions (BSSC 1997), which are based on the properties of the top 30 m of the soil column, disregarding the characteristics of deeper geology. Recently, the classification was revised (Kuo et al. 2011), and a database containing Vs30 data for most TSMIP stations was created (Kuo et al. 2012; see also http://egdt.ncree.org.tw/). Lee and Tsai (2008) constructed a Vs30 map for the entire island. We used these sources to estimate the Vs30 value and to identify a site class for each strong-motion station. The site classification is based on Vs30 values as follows (Lee and Tsai 2008): Site Class B (rock), Vs30 760–1,500 m/s; Site Class C (very dense soil and soft rock), Vs30 360–760 m/s; Site Class D (stiff soil), Vs30 180–360 m/s; and Site Class E (soft soil), Vs30 < 18 m/s. The distribution of TSMIP stations with respect to the assigned site classes is shown in Fig. 1c. As observed, the Class D and E stations are mainly distributed on plains and basins with unconsolidated sediments; class C stations are located on the northwestern part of Taiwan and around the Central Mountain Range; and the class B stations are mostly located in mountain areas and on the north and south of Taiwan (see also Lee and Tsai 2008).

In addition to the site classification based on the Vs30 values, we also considered gross geological conditions (see, for example, Lee and Tsai 2008) and analysed the data of four arrays separately: CHY, ILA, TAP and TCU (see Fig. 1b). The CHY array is located on the so-called Chianan Plain, which is covered with diverse Quaternary alluvial sediments (total thickness may almost reach 1,500 m, e.g., Lin et al. 2009). Most stations of the ILA array are installed on the Quaternary alluvial Ilan basin, in which the thickness of deposits may reach beyond 1,200 m in the central part. The TAP array—most of the ground-motion stations located within the Taipei Basin—is a triangular asymmetric alluvium-filled basin; the thickness of sediments in the deepest part of the basin reaches up to 700 m. The geological structure inside the basin consists of Quaternary layers above Tertiary base rock (e.g., Wang 2008). The TAP array also includes several stations located to the east and south of the basin in mountainous and hilly areas. The stations of the TCU array are located on relatively stiff soils in extended hilly areas. The relative distributions of strong-motion records obtained at different site classes and arrays are shown in Fig. 2.

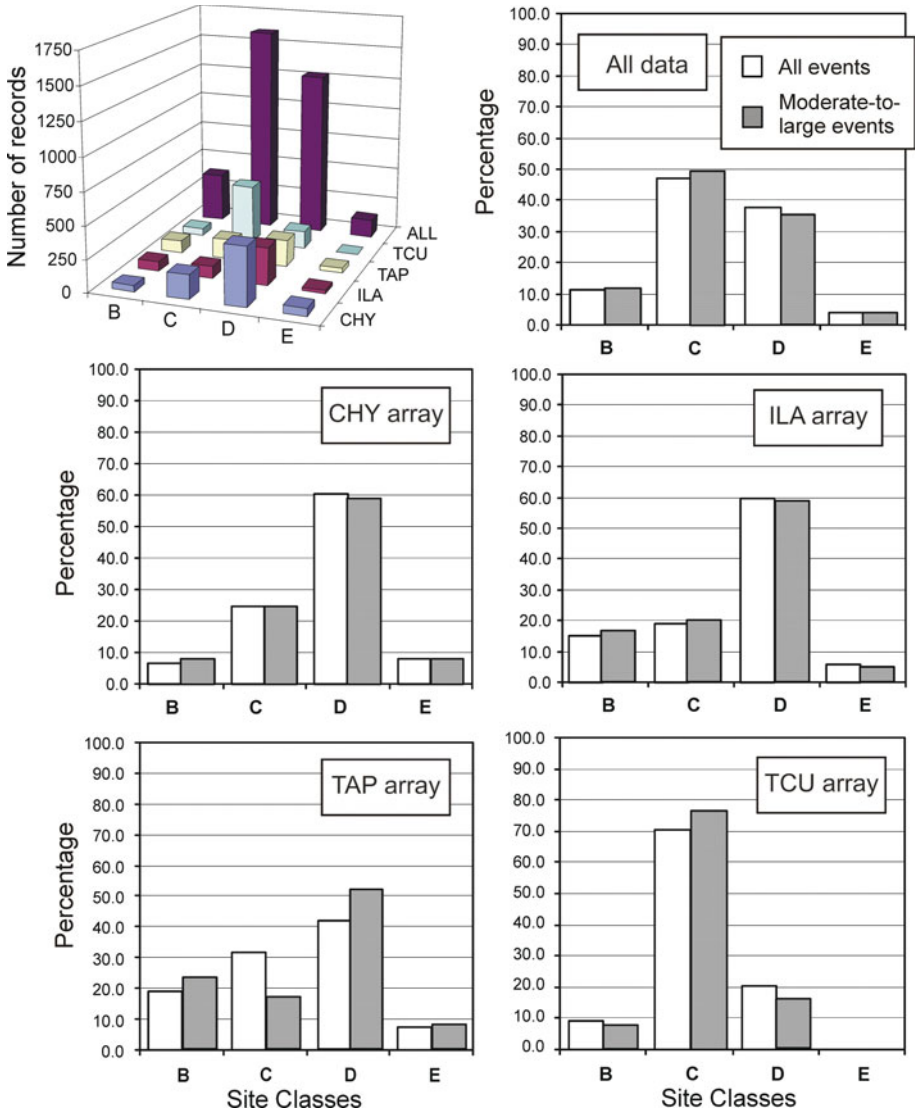

Fig. 2 Distribution of ground-motion records (absolute numbers and relative percentage) versus site classes and arrays

4 Technique and results

Two approaches for estimating within-earthquake correlation were mentioned above: one is based on the estimation of sample semivariograms, while the other is the direct evaluation of the linear correlation coefficient. Goda and Atkinson (2009) showed that the direct evaluation of the correlation coefficient tends to result in a slightly more rapid decay of the function $\rho_\varepsilon(\Delta)$. At the same time, they noted that the possible differences in $\rho_\varepsilon(\Delta)$ based on the two methods can be largely attributed to the effects of between-earthquake standard deviation

and suggested creating an overall $\rho_\varepsilon(\Delta)$ model by averaging the results of the methods. In our study, we applied both approaches and considered the averaged values of the spatial correlation coefficients, which were obtained as follows: the coefficients were calculated for every bin using both methods; then, the average value for a given bin was calculated as the arithmetic average. These averaged values were used to estimate the empirical correlation function $\rho_\varepsilon(\Delta)$.

The evaluation of the within-earthquake standard deviation $\hat{\sigma}_\varepsilon$ (i.e., the sample within-earthquake standard deviation) is very important because it can significantly affect the estimates of within-earthquake correlation (e.g., Goda and Hong 2008a; Goda and Atkinson 2009, 2010). On the one hand, the sample standard deviation provided by the GMPE can be used, as was performed, for example, by Sokolov et al. (2010) and Esposito and Iervolino (2011). On the other hand, $\hat{\sigma}_\varepsilon$ for a given seismic event may be assessed by evaluating the sample semivariogram for the data pairs with sufficiently large separation distances Δ (Goda and Atkinson 2009, 2010). It is assumed that within-earthquake residuals are uncorrelated at sufficiently long separation distances, and the semivariogram is expected to attain a constant plateau level. Alternatively, $\hat{\sigma}_\varepsilon$ may be calculated directly as the event-based standard deviation of regression residuals. However, the GMPE-based and the directly calculated event-based values may be underestimated if the residuals are strongly correlated (Kawakami and Mogi 2003; Hong et al. 2009).

At the same time, the within-earthquake standard deviation inferred from the large-separation-distance plateau of the semivariogram is rather a subjective estimate because of uncertainty in the used range of the plateau. In this work, we used all of these approaches to estimate the sample within-earthquake standard deviation, and we considered the alternative values, i.e., the generalised GMPE-based estimates $\hat{\sigma}_{\varepsilon,GMPE}$; the direct estimates from a single event $\hat{\sigma}_{\varepsilon,DSE}$ and from the used datasets $\hat{\sigma}_{\varepsilon,DDS}$; and the semivariogram-based estimates from a single event $\hat{\sigma}_{\varepsilon,SVSE}$ and from the used datasets $\hat{\sigma}_{\varepsilon,SVDS}$.

Let us analyse residuals obtained using the MW2010 model. Figure 3a shows sample semivariograms $\hat{\gamma}(\Delta)$ that were used to estimate the event-based within-earthquake standard deviation $\hat{\sigma}_{\varepsilon,SVSE}$. In this case, the spatial variability increases gradually with separation distance Δ and reaches a plateau at $\Delta > 50$ km. Thus, the value of $\hat{\sigma}_{\varepsilon,SVSE}$ for a given event was calculated as the weighted average of the $\sqrt{\hat{\gamma}(\Delta)}$ values between 50 and 100 km. The weights are given by the number of strong-motion data pairs within the bins.

To estimate the direct within-earthquake standard deviation of regression residuals $\hat{\sigma}_{\varepsilon,DSE}$ or $\hat{\sigma}_{\varepsilon,DDS}$, the between-earthquake residual $\hat{\eta}$ for a particular event was calculated as the mean residual for that event, and the value was subtracted from the regression residuals. Figure 3b compares the event-based direct estimates of $\hat{\sigma}_{\varepsilon,DSE}$ values with the estimates $\hat{\sigma}_{\varepsilon,SVSE}$ obtained from the semivariograms. In most cases, the direct estimates are slightly lower than those obtained from the semivariograms; however, in general, the difference is less than 10 %.

Compared with the whole dataset, the event-based estimates of the within-earthquake standard deviation are affected by a limited number of stations, which register a particular earthquake. Figure 3c shows the distribution of the direct event-based estimates $\hat{\sigma}_{\varepsilon,DSE}$ for both ground-motion models used versus the number of records collected for the events and versus the magnitude of the events. As expected, when the number of records used is relatively small, the event-based estimates reveal smaller values of the sample within-earthquake standard deviation than the generalised GMPE's standard deviation. In these cases, the records were obtained at a limited number of stations located within a narrow geographical area.

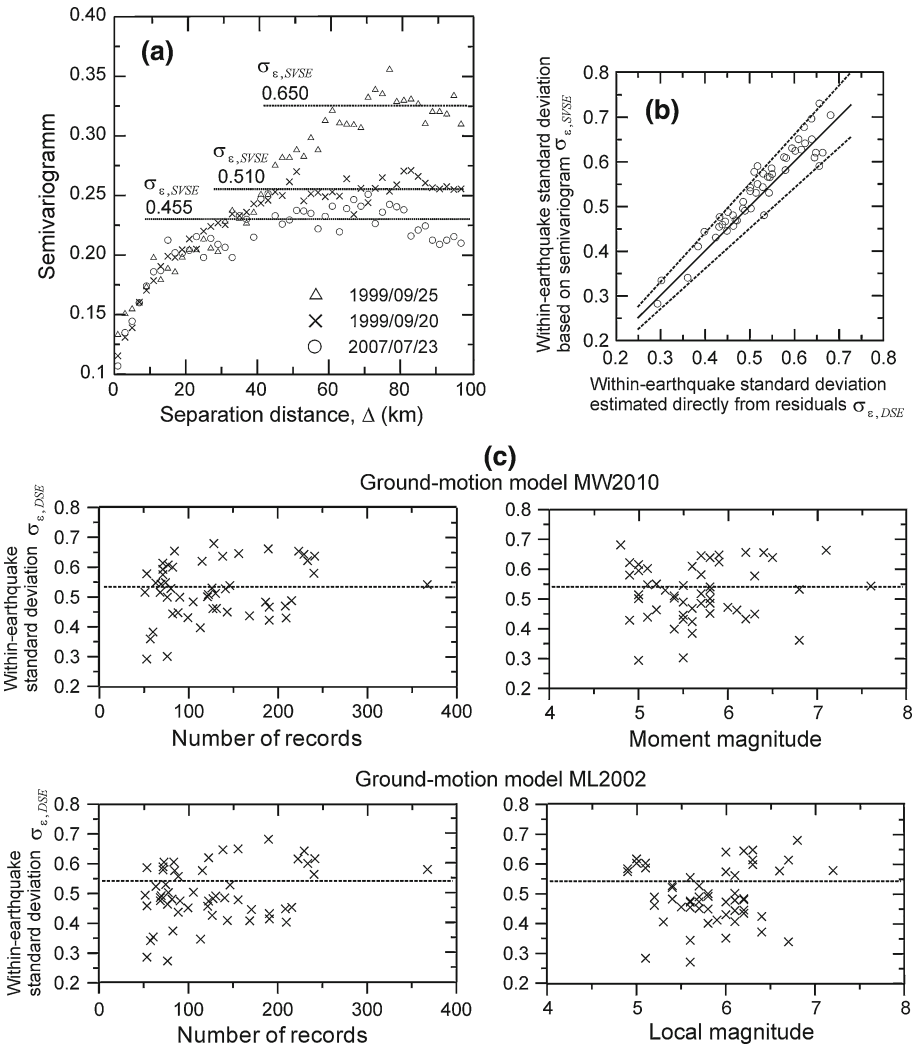

Fig. 3 Statistics of within-earthquake residuals. **a** Experimental semivariogram $\hat{\gamma}(\Delta)$ (examples of three earthquakes) used to estimate within-earthquake standard deviation. **b** Comparison of event-dependent estimates of within-earthquake standard deviation $\hat{\sigma}_\varepsilon$, which were obtained using different techniques (see text). The *solid line* denotes direct correspondence; *dashed lines* mark ±10% interval. **c** Event-dependent values of within-earthquake standard deviation, which were estimated directly from residuals: distribution versus number of records and magnitude

Such a limited amount of data is mostly typical for small-magnitude earthquakes or for distant moderate-to-large earthquakes.

The small-magnitude earthquakes ($M_L < 6.0$, see Fig. 3c) are characterised, at least in the case of shallow well-recorded events in Taiwan, by a relatively large scatter in the event-based standard deviation. Therefore, in addition to conducting an analysis of the possible influence of different techniques used to estimate the within-earthquake standard deviation, we also verified the effect of magnitude selection when compiling the database for the study.

In this work, we used an exponential decay function $\rho_\varepsilon(\Delta) = \exp(a\Delta^b)$ to describe the within-earthquake correlation; the model parameters a and b were evaluated using the

weighted least squares technique from experimental correlation values calculated for a finite number of separation distances Δ. The weights are given by the number of strong-motion data pairs within the bins. Figure 4a compares the correlation functions $\rho_\varepsilon(\Delta)$ for ground-motion residuals obtained with respect to both ground-motion models considered (MW2010 and ML2002). The correlation was estimated using (a) four approaches for the calculation of the sample within-earthquake standard deviation, which were described above (GMPE-based, direct estimates from a single event and semivariogram-based estimates from a single event and from dataset) and (b) three datasets consisting of different amounts of records. The first dataset contains all selected records; the second dataset (12 earthquakes, see Fig. 1a) was compiled using records from moderate-to-large earthquakes ($M_W \geq 6.0$ for the model MW2010 and $M_W \geq 6.3$ for the model ML2002); and the third dataset contains records from small earthquakes ($M_W < 6.0$ for the model MW2010 and $M_W < 6.3$ for the model ML2002). The magnitude range $M_W \geq 6.0$ (and, correspondingly, $M_W \geq 6.3$) was selected because the direct event-based within-earthquake standard deviations estimated for these events show relatively stable scatter around the generalised GMPE's standard deviation (Fig. 3c).

The difference between estimates of within-earthquake correlation, which were obtained using different approaches for calculating the within-earthquake standard deviation, seems to be negligible. Thus, in subsequent analyses, we used the averaged estimates calculated from the results of these approaches. However, there is a prominent difference between the correlation functions obtained from the three datasets considered: the level of within-earthquake correlation increases with the increase in earthquake magnitude, records of which dominate the dataset. The possible reasons for this phenomenon will be discussed below. Because the datasets, which contain all selected records and records from small earthquakes, show characteristics similar to those of the within-earthquake correlation models, we further consider only the first dataset (all records).

Let us consider subsets of the data, which contain records obtained under specific site conditions, namely, different site classes and records collected by particular arrays (CHY, ILA, TAP, and TCU, see Fig. 1b). The following groups of data were classified with respect to different site conditions: class BC; class C; class CD; and class DE. Each group contains records collected by stations located at corresponding site classes. The number of records collected for site classes B and E is not sufficient to analyse the data separately. For the considered datasets, the generalised GMPE-based within-earthquake standard deviation ($\hat{\sigma}_{\varepsilon,GMPE}$) could not be used; thus, we applied the direct estimates $\hat{\sigma}_{\varepsilon,DDS}$ from the datasets. Note that for these datasets, the values of the event-based within-earthquake standard deviation (both direct $\hat{\sigma}_{\varepsilon,DSE}$ and based on semivariogram $\hat{\sigma}_{\varepsilon,SVSE}$ variates) were calculated using all data from a given earthquake.

Figure 4b and c shows correlation functions $\rho_\varepsilon(\Delta)$ obtained for different site classes and arrays, and Table 1 summarises the characteristics of the datasets and the corresponding within-earthquake correlation. The correlation functions $\rho_\varepsilon(\Delta)$, which were estimated for two datasets (all data and moderate-to-large earthquakes), are compared in Fig. 5. Residuals based on both ground-motion models reveal similar phenomena. First, the level of within-earthquake correlation depends on site class, i.e., the estimated correlation distances are smallest (8–10 km for all data and 15–20 km for moderate-to-large earthquakes) for the class BC group of data and largest (22–30 km and more than 50 km, respectively) for the class DE group of data. Second, the level of within-earthquake correlation varies depending on the strong motion array. The level of correlation is lowest (correlation distances 6–10 km for all data and 7–15 km for moderate-to-large earthquakes) for the ILA and the TAP arrays, which are the alluvium-filled basins with large spatial variations in the thickness of deposits. The CHY array (Chianan Plain, thick Quaternary strata) is characterised by the largest

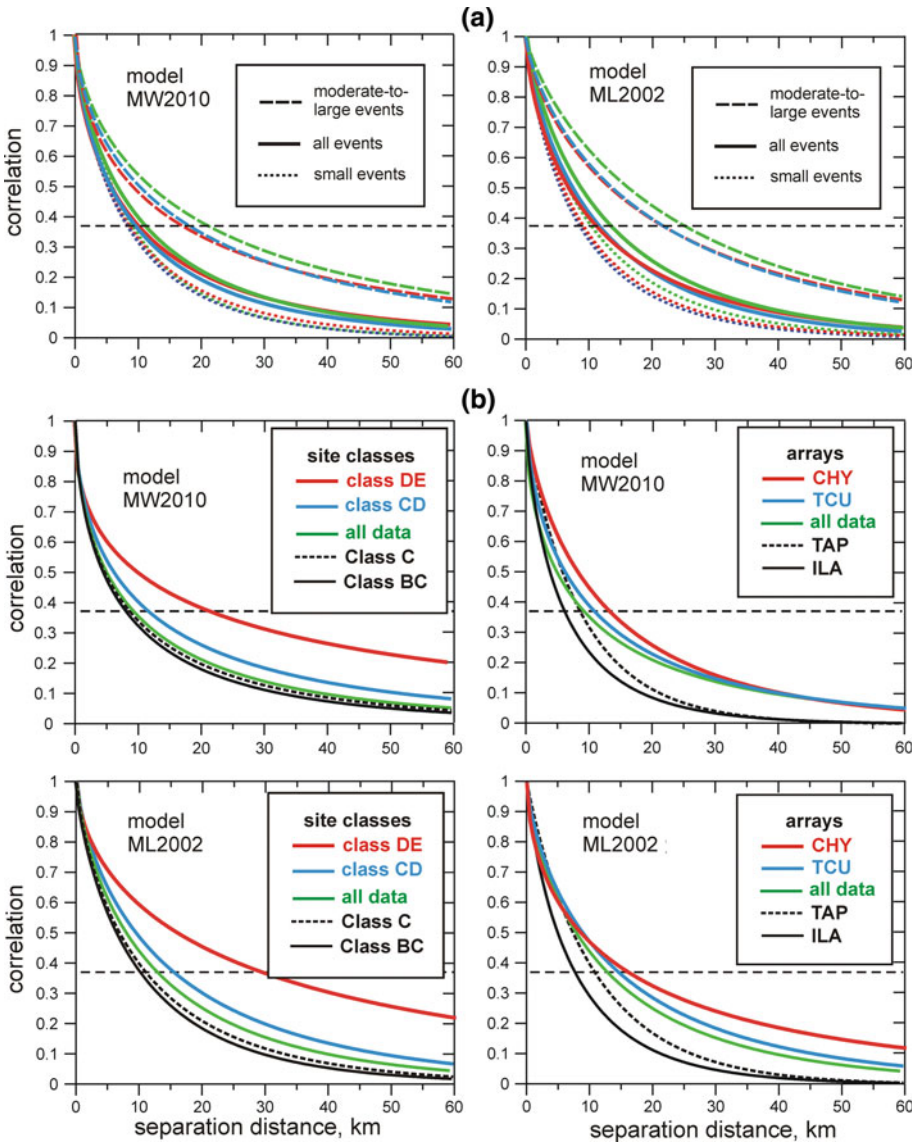

Fig. 4 Site-to-site (spatial) correlation functions $\rho_\varepsilon(\Delta)$ estimated for considered ground-motion models and various groups of data. **a** Magnitude-dependent grouping of the data; results based on three variants of sample within-earthquake standard deviation (GMPE-based, direct estimates from a single event, and semivariogram-based estimates from a single event; see text) are shown in different colours; **b** site-class and array-dependent grouping of data. *Thin dashed lines* mark the level $1/e = 0.368$

correlation distances (13–17 km for all data and approximately 40 km for moderate-to-large earthquakes). As observed, the level of within-earthquake correlation, which was estimated for the considered datasets, varies with earthquake magnitude, records of which dominate the dataset. The possible reasons for the peculiarities of the within-earthquake correlation in Taiwan will be discussed in the next section. Note that the level of correlation resulting from the ML2002 model is consistently higher than that resulting from the MW2010 model.

Table 1 Statistical characteristics of within-earthquake residuals and within-earthquake correlation, raw data

Data set, site classes, arrays	Number of records	Mean residual from the data	Within-earthquake standard deviation		within-earthquake correlation function		Correlation distance, km[a]
			$\sigma_{\varepsilon,DDS}$	$\sigma_{\varepsilon,SVDS}$	a	b	
Model MW2010, $\sigma_{\varepsilon,GMPE} = 0.55$, $M_W > 4.8$							
All data	6,787	0.0	0.541	0.573	−0.268	0.583	9.6
BC	3,958	−0.091	0.572	0.583	−0.283	0.601	8.2
C	3,200	−0.054	0.569	0.578	−0.278	0.586	8.9
CD	5,749	0.026	0.532	0.569	−0.246	0.566	11.9
DE	2,830	0.129	0.463	0.508	−0.234	0.469	22.1
CHY	1,486	0.059	0.481	0.495	−0.142	0.750	13.4
ILA	970	0.078	0.504	0.540	−0.247	0.763	6.2
TAP	968	0.036	0.509	0.522	−0.137	0.920	8.7
TCU	1,270	0.0	0.514	0.530	−0.217	0.639	10.9
Model ML2002, $\sigma_{\varepsilon,GMPE} = 0.605$, $M_L > 5.0$							
All data	7,070	0.0	0.527	0.570	−0.149	0.743	12.9
BC	4,078	−0.112	0.553	0.565	−0.157	0.792	10.3
C	3,298	−0.076	0.548	0.560	−0.160	0.758	11.2
CD	5,998	0.025	0.517	0.563	−0.134	0.728	15.8
DE	2,996	0.154	0.445	0.499	−0.133	0.593	30.0
CHY	1,712	0.105	0.466	0.456	−0.199	0.575	16.5
ILA	968	0.014	0.479	0.524	−0.188	0.817	7.7
TAP	958	0.101	0.495	0.535	−0.106	0.938	10.9
TCU	1325	0.0	0.521	0.563	−0.139	0.731	14.9
Model MW2010, $\sigma_{\varepsilon,GMPE} = 0.55$, $M_W > 6.0$							
All data	1,876	0.0	0.564	0.590	−0.132	0.669	20.6 (2.1)
BC	1,144	−0.10	0.592	0.640	−0.140	0.715	15.6 (1.9)
C	926	−0.05	0.586	0.595	−0.096	0.784	19.9 (2.2)
CD	1,585	0.03	0.553	0.589	−0.133	0.644	22.9 (1.9)
DE	732	0.157	0.477	0.482	−0.113	0.526	63.1 (2.9)
CHY	400	0.16	0.525	0.490	−0.122	0.577	38.3 (2.9)
ILA	266	0.057	0.512	0.525	−0.297	0.630	6.9 (1.1)
TAP	254	0.036	0.488	0.510	−0.201	0.707	9.7 (1.1)
TCU	402	0.022	0.539	0.595	−0.159	0.678	15.1 (1.4)
Model ML2002,$\sigma_{\varepsilon,GMPE} = 0.605$, $M_L > 6.3$							
All data	2,081	−0.03	0.574	0.630	−0.068	0.809	27.7 (2.1)
BC	1,240	−0.176	0.581	0.605	−0.102	0.759	20.2 (2.0)
C	994	−0.144	0.584	0.657	−0.070	0.842	23.5 (2.1)
CD	1,748	0.0	0.570	0.618	−0.055	0.838	31.9 (2.0)
DE	841	0.179	0.492	0.533	−0.043	0.761	62.5 (2.1)
CHY	427	0.136	0.524	0.460	−0.053	0.796	40.0 (2.4)
ILA	302	−0.041	0.475	0.466	−0.174	0.781	9.4 (1.2)

Table 1 continued

Data set, site classes, arrays	Number of records	Mean residual from the data	Within-earthquake standard deviation		within-earthquake correlation function		Correlation distance, km[a]
			$\sigma_{\varepsilon,DDS}$	$\sigma_{\varepsilon,SVDS}$	a	b	
TAP	374	0.172	0.493	0.516	−0.078	0.933	15.5 (1.4)
TCU	469	−0.072	0.554	0.657	−0.111	0.734	19.9 (1.3)

[a] Values in parentheses indicate difference (ratio) between corresponding estimates for intermediate-to-large earthquakes and for all earthquakes

5 Analysis and discussion

The within-earthquake correlation of ground-motion residuals reflects a non-random component in the residuals, which is caused by factors not accounted for by the ground-motion model. The apparent dependence of spatial correlation level on site classes (see Fig. 4b, c) may be qualitatively explained as follows. Peak amplitudes of ground acceleration are sensitive to the relatively high-frequency region of soil amplification. Sokolov et al. (2004, 2007) noted the prominent influence of geological and geomorphological factors on high-frequency amplitudes of site amplification for rock (class B) and soft rock or very dense soil (class C) sites in Taiwan. The influence results in large variations in amplification amplitudes and predominant frequencies between particular stations with the same assigned site class. Table 1 shows that the scatter of residuals (within-earthquake standard deviation) is the highest for the class BC subset of the data.

On the other hand, even the relatively high variance of the amplitudes of site amplification at intermediate and low frequencies, which may be typical for site classes D and E (stiff and soft soil), may not result in considerable modifications in the high-frequency PGA amplitudes. Therefore, the scatter of residuals is relatively low for these soil types (Table 1), which together with the prominent non-random component of the residuals (positive non-zero mean residuals) is responsible for the relatively high level of within-earthquake correlation in comparison with the other site classes (Fig. 4). At the same time, as it will be shown below, other factors (i.e., influence of deep sediments, etc.) may contribute to the non-randomness of residuals for the class DE subset of the data.

5.1 Spatial distribution of geological characteristics (Vs30 values and thickness of sediments)

The subsets of data collected for some arrays (CHY, ILA and TAP) are characterised by approximately a similar distribution of ground-motion records for different site classes (Fig. 2). Thus, it is reasonable to assume that the difference in the level of within-earthquake correlation for these arrays, in addition to other factors, may be caused by a corresponding difference in the spatial variations of the average shear-wave velocity Vs30. For example, Lee and Tsai (2008) showed a prominent spatial variation in Vs30 values along the Taipei Basin. On the other hand, Baker and Jayaram (2008) and Jayaram and Baker (2009) analysed the Vs30 values for stations that recorded the Chi–Chi earthquake and showed that the values have significant spatial correlation.

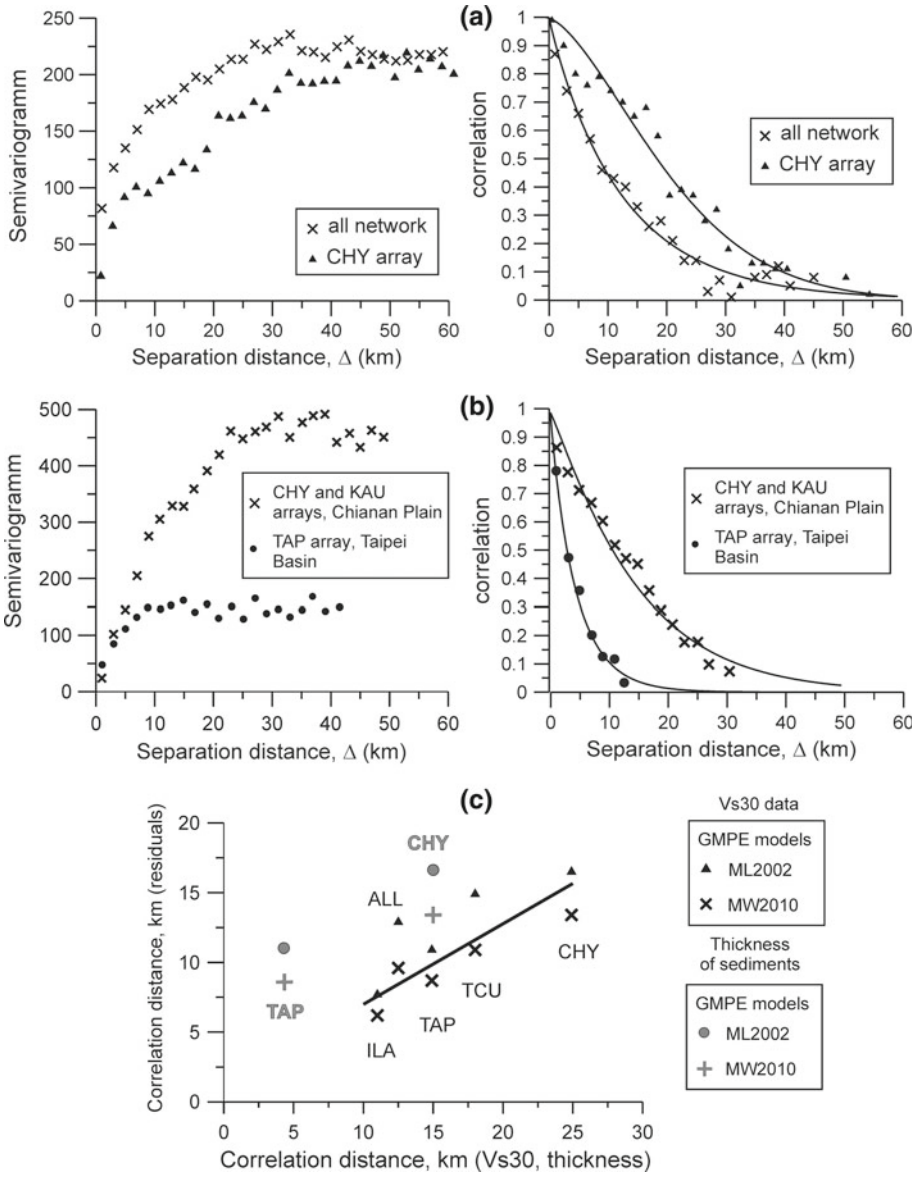

Fig. 5 Statistics of spatial distribution of geological data; analysis of Vs30 values (**a**) and thickness of sediments (**b**). **c** Relationship between correlation distances inferred from correlation functions: within-earthquake residuals versus Vs30 (*black symbols*) and within-earthquake residuals versus thickness of sediments (*grey symbols*), different ground-motion arrays and considered ground-motion models. The *solid line* shows a linear relationship (Eq. 13c, see text)

In this work, we used the recent estimates of the Vs30 values (Kuo et al. 2011, 2012; Lee and Tsai 2008) and analysed the spatial correlation of Vs30 for the entire TSMIP network and for the CHY, ILA, TAP and TCU arrays separately. The following scheme was applied. First, a sample semivariogram $\hat{\gamma}_{Vs30}(\Delta)$ was calculated using a particular dataset;

then, the variance of spatially distributed Vs30 values σ^2_{Vs30} was estimated from the large-separation distance plateau of the semivariogram. The variance was used to calculate the sample correlation coefficients and the corresponding correlation functions $\rho_{Vs30}(\Delta)$. Thus, it may be possible to compare the characteristics of spatial correlation between the PGA within-earthquake residuals and the average shear-wave velocities.

Figure 5a shows sample semivariograms $\hat{\gamma}_{Vs30}(\Delta)$ and spatial correlation functions $\rho_{Vs30}(\Delta)$ calculated for the entire TSMIP network and the CHY array; Fig. 5c compares the correlation distances R_C inferred from the corresponding correlation functions. There is a prominent correspondence between the within-earthquake correlation of the PGA residuals $\rho_\varepsilon(\Delta)$ (both ground-motion models) and the Vs30 values $\rho_{Vs30}(\Delta)$ estimated for particular arrays; a larger correlation between the Vs30 values corresponds to a larger correlation between the PGA residuals, and the relationship is apparently linear. However, the relation $R_{C,PGA}/R_{C,Vs30}$, which was estimated using the data from the entire TSMIP network, reveals a relatively lower level of Vs30 correlation for a given level of residual correlation compared with similar relations for certain arrays. It seems that the spatial variation of the Vs30 values estimated over a large area, which covers various gross geological structures, is not an adequate descriptor of ground-motion correlation. The following simple relationships between the Vs30 correlation distances $R_{C,Vs30}$ and the PGA residuals correlation distances $R_{C,PGA}$ were obtained (without consideration of the values for the entire TSMIP network):

$$R_{C,\ PGA(MW2010)} = 0.931 + 0.516R_{C,Vs30}\ \ [\pm 0.883],\ \ r = 0.985, \text{ the MW2010 model} \tag{13a}$$

$$R_{C,\ PGA(ML2002)} = 1.522 + 0.638R_{C,Vs30}\ \ [\pm 1.602],\ \ r = 0.944, \text{ the ML2002 model} \tag{13b}$$

$$R_{C,\ PGA} = 1.227 + 0.577R_{C,Vs30}\ \ [\pm 1.883],\ \ r = 0.874, \text{ for both models} \tag{13c}$$

where the error term is shown in brackets and r is the linear correlation coefficient. The linear equation provides the best representation of the relationship between $R_{C,PGA}$ and $R_{C,Vs30}$ for the MW2010 model.

However, in addition to the softness of the surface rocks, the thickness of sediments is another important local geological factor that affects the level of shaking experienced in earthquakes. As shown in Figs. 1c and 2 (see also Lee and Tsai 2008), the class D and class E stations are mainly distributed on plains and basins with unconsolidated sediments. Thus, the joint influence of soft surface soil and thick sediments, which cover large areas, may be responsible for homogeneous site effects, which are not considered in the ground-motion models. The resulting non-random residuals produce a relatively high level of within-earthquake correlation. At the same time, uneven basement topography and complex local structure may produce considerable variability in local site amplification and therefore in the peak amplitudes of ground motion. As recently reported, large spatial variations in ground-motion characteristics are apparent in the Taipei (see, for example, a short review in Sokolov et al. 2009) and Ilan (e.g., Furumura et al. 2001; Liu and Tsai 2005; Hsu et al. 2008) basins, for which the level of spatial correlation is found to be the smallest.

In this work, we performed a preliminary analysis of the spatial variation in the thickness of sediments for the Taipei Basin and Chianan Plain. The detailed structure of the Taipei Basin is known from shallow reflection seismic experiments and from borehole drilling data (Wang et al. 2004, see also Miksat et al. 2010). The triangular basin is floored by a Tertiary basement with a maximum depth of approximately 700 m in the northwest. The depth of the Miocene basement of the Chianan Plain was estimated by Lin et al. (2009). The analysis was performed using a scheme similar to that used for the Vs30 values. Figure 5b shows the

semivariograms $\hat{\gamma}(\Delta)$ and spatial correlation function $\rho(\Delta)$ (thickness of sediments) calculated for the Taipei Basin (TAP array) and Chianan Plain (CHY and KAU arrays). Figure 5c compares the correlation distances R_C inferred from the corresponding correlation functions. A direct relationship between the correlation of the thickness of sediments and the correlation of the PGA residuals is evident.

5.2 Moderate-to-large earthquakes

When analysing the dependence of spatial correlation on earthquake magnitude (see Fig. 6 and Table 1, for comparison), it is necessary to note the following. First, all considered groups of site classes (i.e., BC, CD, and DE) show an increase in correlation level with the increase in earthquake magnitude, records of which dominate the dataset. Thus, the data for all site classes are characterised by additional non-random residuals for moderate-to-large earthquakes compared with the data from the smaller earthquakes. The relative increase in the correlation level is approximately similar for all site classes for the model ML2002, while the MW2010 model reveals the highest relative increase for the class DE group of data.

Second, the data from the ILA array are characterised by almost the same level of within-earthquake ground-motion correlation for both considered datasets (all earthquakes and moderate-to-large earthquakes), and the relation $R_{C,PGA}/R_{C,Vs30}$ does not depend on magnitude. The data from the other arrays reveal an apparent dependence of the correlation on magnitude; this dependence seems to be smallest for the TAP array and largest for the CHY array, for which the level of within-earthquake correlation (correlation distance) for moderate-to-large earthquakes is almost 3 times higher than that estimated from all of the data.

Thus, it is reasonable that, in addition to the spatial variation of the Vs30 values, some additional factors that affect the non-random component of residuals and therefore ground-motion within-earthquake correlation should be considered. Bearing in mind the distribution of epicentres of earthquakes with $M_L \geq 6.3 (M_W \geq 6.0)$, the records of which were analysed separately (Fig. 1a), it is possible to suggest that the path or azimuthal effects play an important role in ground-motion variability and correlation in Taiwan (see also Tsai et al. 2006; Sokolov et al. 2010). At the same time, because the used ground-motion prediction equations consider hypocentral distance, the point-source approximation of extended fault (e.g., Chi–Chi earthquake) as well as neglected hanging- and foot-wall effects may result in non-random residuals in the near-fault region.

5.3 Site class- and array-dependent correction

The knowledge of the loss distribution about the mean (e.g., the variance or standard deviation) is very important for decision making and mitigation activities. For example, primary insurers are concerned with the central part of the distribution (mean and median values), while re-insurers deal mostly with the tail of the distribution. It has been already mentioned in Introduction that a higher between-earthquake and within-earthquake correlation results in a higher probability of extreme loss values and a larger variation in losses to a portfolio.

To reduce random variability in the modelled ground motion and thus reduce the uncertainty in estimated losses, it is necessary to find appropriate and acceptable approaches to reduce the standard deviation (sigma) associated with GMPEs and to reduce the within-earthquake correlation by removing the non-random component of residuals. In their previous study on the ground-motion correlation in Taiwan, Sokolov et al. (2010) showed that the application of a site- or array-dependent empirical correction may affect the within-earth-

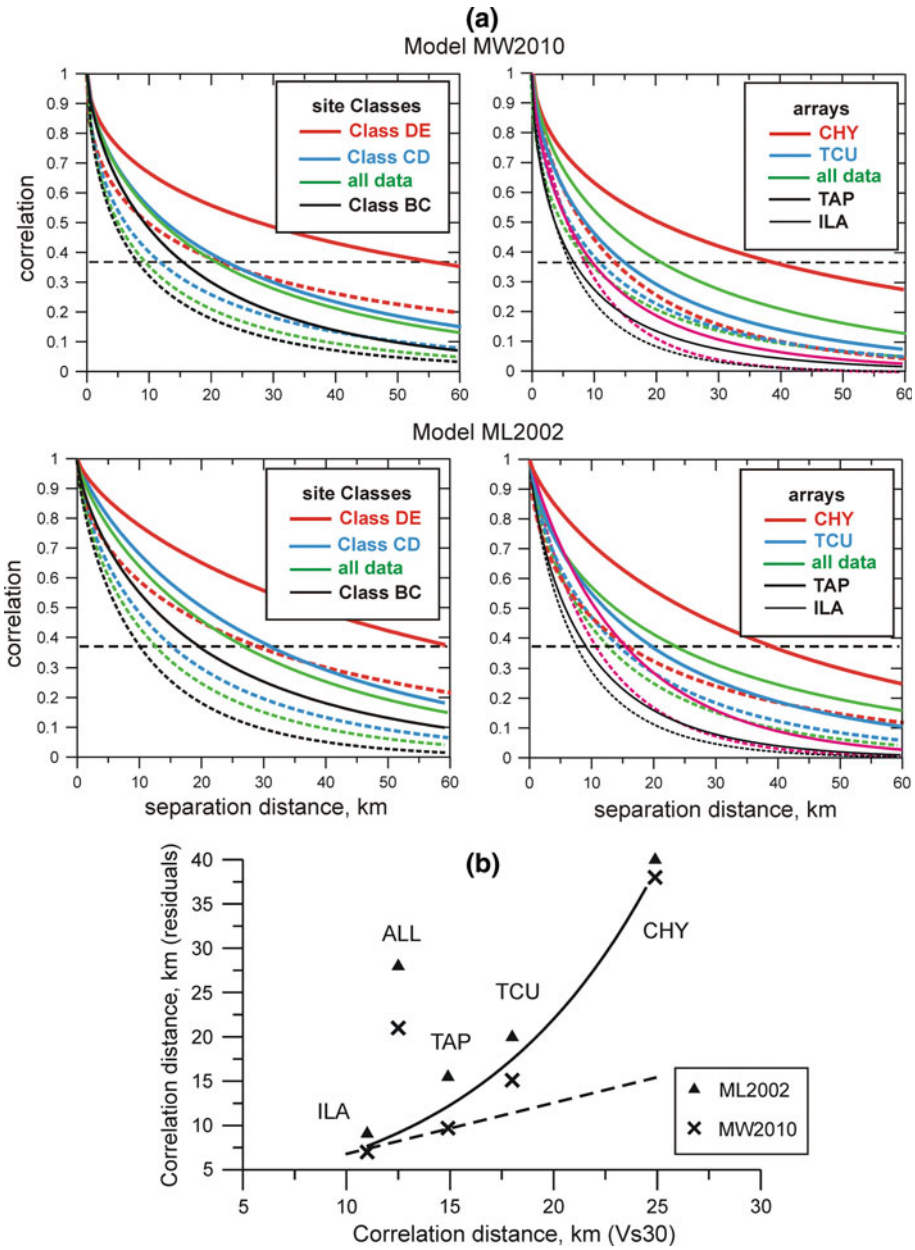

Fig. 6 Site-to-site correlation, analysis of intermediate-to-large earthquakes. **a** Comparison between the correlation functions $\rho_\varepsilon(\Delta)$ estimated for considered ground-motion models, site classes and arrays using two datasets: all data (*dashed lines*) and moderate-to-large earthquakes (*solid lines*). *Thin dashed lines* mark the level $1/e = 0.368$. **b** Relationship between correlation distances inferred from correlation functions (Vs30—$R_{C,Vs30}$ and within-earthquake residuals—R_C), different ground-motion arrays and considered ground-motion models. The *dashed line* shows the linear relationship estimated for all earthquakes (Eq. 13c, see also Fig. 5c); the *solid line* shows the exponential relationship (both models) estimated for intermediate-to-large earthquakes as follows: $R_C = \exp(0.824 + 0.114 R_{C,Vs30})$ [±2.35]

quake correlation. In this work, we verified, in greater detail, the possibility of reducing the level of within-earthquake correlation using a larger amount of data.

On the one hand, bearing in mind the direct dependence of within-earthquake correlation on spatial variations of average shear-wave velocity Vs30, it is reasonable to consider Vs30 values in the correction. Additionally, the response of deep sediments (the Chianan Plain, CHY array) to relatively long-period ground-motions during the moderate-to-large earthquakes should be taken into account. On the other hand, the range of predominant frequencies of earthquake radiation shifts toward the low-frequency part of the spectrum with the increase in magnitude. Thus, the low-frequency site amplification at stiff and soft soil sites (site classes D and E) may play an increasingly important role in the generation of ground-motion acceleration for intermediate-to-large earthquakes compared with smaller-magnitude events, producing non-random residuals. Thus, considering the dependence of residuals on earthquake magnitude appears to be important.

The following characteristics of earthquake records were considered when estimating the empirical correction factors CF for given event i and site j: magnitude M, source-to-site (hypocentral distance) R, and average shear-wave velocity Vs30 assigned for each strong-motion station. We tested various combinations of the parameters as follows:

$$CF_{i,j} = a + bM_i + c \ln R_{i,j}, \tag{14a}$$

$$CF_j = a + d \ln Vs30_j, \tag{14b}$$

$$CF_{i,j} = a + bM_i + c \ln R_{i,j} + d \ln Vs30_j \tag{14c}$$

where a, b, c, and d are the empirical coefficients. The functions CF were estimated by a least-squares regression analysis of within-earthquake residuals $\varepsilon_{i,j}$ separately for various datasets (all records and records from moderate-to-large earthquakes), site classes (BC, C, D, and DE) and arrays (all TSMIP network, CHY, ILA, TAP, and TCU arrays). It has been found that the joint consideration of magnitude, distance and average shear-wave velocity (Eq. 14c) provides the smallest values for the residual sum of squares, i.e., $\sum (\varepsilon_{i,j} - CF_{i,j})^2$. The results of the application of Akaike's information criterion (AIC, Akaike 1974) and the Bayesian information criterion (BIC, Schwarz 1978) for these models are presented in Table 2 (examples for the MW2010 ground-motion model). Smaller AIC and the BIC values result in the model becoming more robust. As shown in Table 2, consideration of the Vs30 values in the models is necessary; however, the models that include magnitude and distance as parameters are among the best used.

Let us consider two variants of the site-dependent correction factor $CF_{i,j}(M_i, R_{i,j}, Vs30_j)$, which were calculated using Eq. 14c. The first variant (CF_{CLASS}) contains four equations that were estimated separately for every considered site class or group of site classes, i.e., BC, C, D, and DE. The second variant (CF_{ARRAY}) contains five equations that were estimated for the entire dataset and separately for the arrays CHY, ILA, TAP and TCU. These estimates were calculated for both considered datasets (all earthquakes and moderate-to-large earthquakes) separately. Figures 7 and 8 show the distribution of uncorrected ground-motion residuals versus independent variables R and $Vs30$ for some datasets and a comparison between correlation functions $\rho_\varepsilon(\Delta)$, which were estimated for different site classes and arrays using uncorrected residuals and residuals after the correction.

The analysis of the data from all earthquakes shows that both variants of the correction factor would prominently reduce the site-to-correlation only for the class DE group (Fig. 7c). When averaging the effects of source-to-site propagation path and reducing the influence of an extended source, i.e., when using the data from all earthquakes, the non-random component of the residuals appears to be the most sensitive to relatively small Vs30 values (site classes

Table 2 Results of applying Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) for ranking models with respect to empirical correction factors *CF*

Parameters in the model (Eqs. 14a, 14b, 14c)	Number of records	SD	AIC	BIC
All data				
M, R	6,787	0.540	−1.233	4.463
M, R, Vs30	6,787	0.521	−1.302	2.436
Vs30	6,787	0.525	−1.289	9.962
DE site classes				
M, R	2,830	0.462	−1.540	3.281
M, R, Vs30	2,830	0.459	−1.552	1.654
Vs30	2,830	0.460	−1.550	8.080
CHY array				
M, R	1,486	0.479	−1.466	3.021
M, R, Vs30	1,486	0.453	−1.577	1.353
Vs30	1,486	0.459	−1.556	7.282

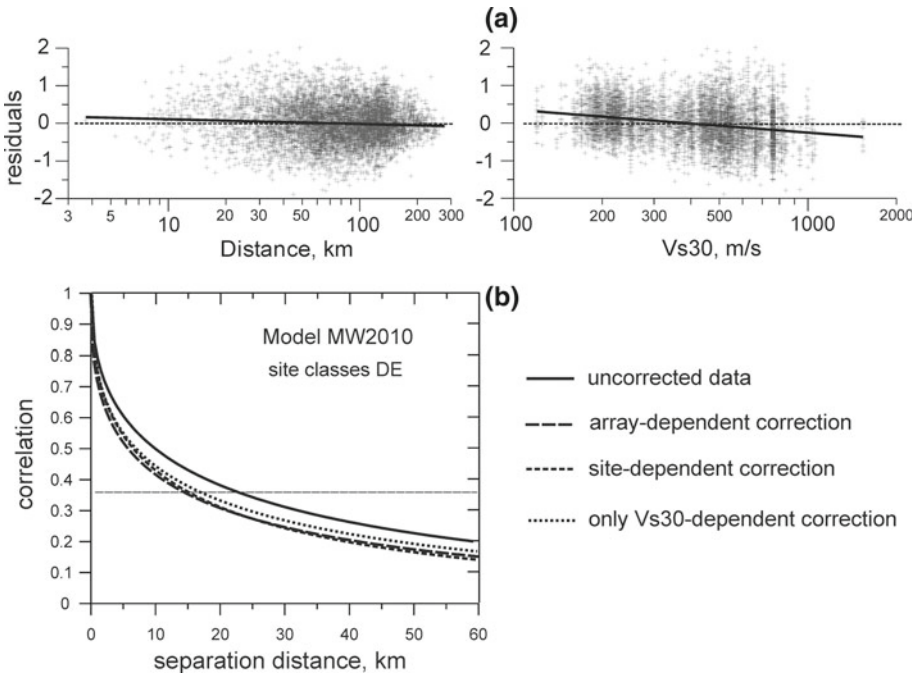

Fig. 7 Empirical correction for residuals, all data. **a** Distribution of residuals versus hypocentral distance and average shear-wave velocity Vs30. *Solid lines* show linear regression of residuals. **b** Influence of the correction (various models, see text) on estimates of correlation functions $\rho_\varepsilon(\Delta)$

D and E, Vs30 < 360 m/s). When the dataset contains residuals related to various site classes (i.e., all data or particular arrays), the influence of the correction factors on within-earthquake correlation is negligible.

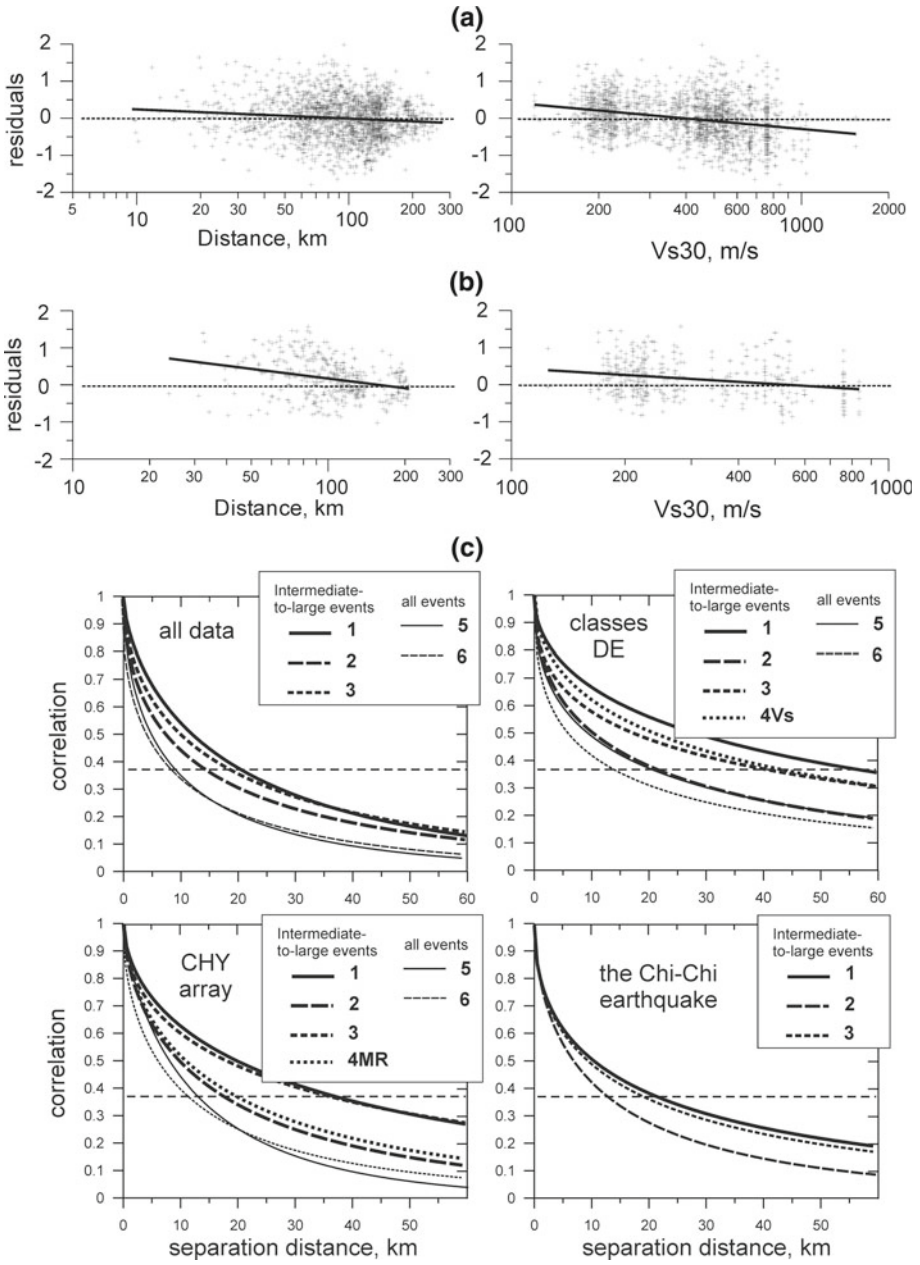

Fig. 8 Empirical correction for residuals, intermediate-to-large earthquakes. **a**, **b** Distribution of residuals versus hypocentral distance and average shear-wave velocity Vs30 for all data (**a**) and CHY array (**b**). *Solid lines* show linear regression of residuals. **c** Influence of the correction (various models, see text) on estimates of site-to-site correlation functions $\rho_\varepsilon(\Delta)$ for different datasets. Intermediate-to-large earthquakes: *1*—uncorrected data; *2*—array-dependent correction (CF_{ARRAY}), Eq. 14c; *3*—site-class-dependent correction (CF_{CLASS}), Eq. 14c; *4Vs*—correction based only on Vs30 values, Eq. 14b; *4MR*—correction based only on magnitude and distance, Eq. 14a. All earthquakes: *5*—uncorrected data; *6*—array-dependent correction (CF_{ARRAY}), Eq. 14c

Table 3 Characteristics of array-dependent correction factor (CF_{ARRAY}), which may be applied to reduce the within-earthquake correlation

Data set	Coefficients of Eq. 14c			
	a	*b*	*c*	*d*
Model MW2010, $M_W > 4.8$				
All data	2.003 ± 0.073	0.029 ± 0.007	-0.104 ± 0.008	-0.290 ± 0.009
CHY	2.006 ± 0.141	0.085 ± 0.013	-0.099 ± 0.014	-0.354 ± 0.019
ILA	1.438 ± 0.168	0.053 ± 0.018	-0.091 ± 0.017	-0.222 ± 0.021
TAP	1.340 ± 0.188	0.145 ± 0.020	-0.104 ± 0.027	-0.289 ± 0.019
TCU	1.828 ± 0.215	0.041 ± 0.015	0.008 ± 0.021	-0.346 ± 0.027
Model MW2010, $M_W > 6.0$				
All data	2.794 ± 0.179	-0.009 ± 0.016	-0.159 ± 0.017	-0.338 ± 0.018
CHY	9.635 ± 0.470	-0.510 ± 0.035	-0.827 ± 0.046	-0.396 ± 0.033
ILA	0.877 ± 0.350	0.306 ± 0.044	-0.154 ± 0.031	-0.376 ± 0.038
TAP	0.864 ± 0.293	0.438 ± 0.036	-0.359 ± 0.052	-0.345 ± 0.029
TCU	0.301 ± 0.457	0.154 ± 0.037	0.038 ± 0.034	-0.241 ± 0.053
Model ML2002, $M_L > 5.0$				
All data	2.269 ± 0.073	-0.038 ± 0.008	-0.030 ± 0.007	-0.321 ± 0.009
CHY	2.570 ± 0.134	0.023 ± 0.014	-0.082 ± 0.013	-0.396 ± 0.017
ILA	1.491 ± 0.175	-0.014 ± 0.021	-0.025 ± 0.016	-0.223 ± 0.020
TAP	1.027 ± 0.197	0.342 ± 0.025	-0.245 ± 0.026	-0.322 ± 0.018
TCU	1.421 ± 0.212	-0.070 ± 0.018	0.168 ± 0.019	-0.291 ± 0.026
Model ML2002, $M_L > 6.3$				
All data	2.239 ± 0.198	-0.011 ± 0.0234	-0.028 ± 0.016	-0.391 ± 0.017
CHY	8.019 ± 0.549	-0.485 ± 0.052	-0.469 ± 0.047	-0.449 ± 0.035
ILA	-2.141 ± 0.365	0.753 ± 0.058	-0.123 ± 0.026	-0.396 ± 0.029
TAP	0.956 ± 0.339	0.568 ± 0.057	-0.497 ± 0.050	-0.364 ± 0.029
TCU	-2.129 ± 0.541	0.279 ± 0.056	0.273 ± 0.035	-0.183 ± 0.049

The influence of correction factors developed using the dataset containing data from moderate-to-large earthquakes is shown in Fig. 8c. The array-dependent correction (CF_{ARRAY}) reduces the level of within-earthquake correlation to an even greater degree than the site-dependent correction (CF_{CLASS}), both for particular datasets collected for different site classes and arrays and for a single earthquake recorded by many stations (e.g., the Chi–Chi earthquake). The array-dependent correction reflects, to a certain degree, the above-mentioned peculiarities of ground-motion residuals, which result in non-random residuals, i.e., joint influence of soft surface soil and thick sediments, the path or azimuthal effects; the consequences of the point-source approximation of extended faults; and neglected hanging- and foot-wall effects.

However, the array-dependent correction, even if it functions much better than the class-dependent one, could not reduce the within-earthquake correlation for intermediate-to-large earthquakes to the same degree, as can be achieved by the corresponding correction in the case of all-event data (Fig. 8c). Thus, the additional factors, which can result in non-random

Table 4 Statistical characteristic of within-earthquake residuals and within-earthquake correlation after application of the array-dependent correction (see Table 3)

Data set	Number of records	Mean residual from the data	Within-earthquake standard deviation		within-earthquake correlation function		Correlation distance, km[a]
			$\sigma_{\varepsilon,DDS}$	$\sigma_{\varepsilon,SVDS}$	a	b	
Model MW2010, $\sigma_{\varepsilon,GMPE} = 0.55$, $M_W > 4.8$							
All data	6,787	−0.018	0.518	0.538	−0.315	0.519	9.3 (1.03)
BC	3,958	−0.012	0.556	0.565	−0.300	0.570	8.3 (1.0)
C	3,200	−0.002	0.556	0.560	−0.266	0.583	9.7 (0.92)
CD	5,749	−0.008	0.515	0.544	−0.259	0.559	11.2 (1.06)
DE	2,830	0.026	0.458	0.491	−0.322	0.428	14.1 (1.57)
CHY	1,486	0.0	0.453	0.480	−0.238	0.583	11.7 (1.15)
ILA	970	−0.004	0.489	0.520	−0.258	0.714	6.7 (0.92)
TAP	968	0.0	0.479	0.486	−0.207	0.714	9.1 (0.95)
TCU	1,270	0.0	0.497	0.520	−0.211	0.640	11.4 (0.97)
Model ML2002, $\sigma_{\varepsilon,GMPE} = 0.605$, $M_L > 5.0$							
All data	7,070	−0.022	0.493	0.533	−0.271	0.534	11.5 (1.12)
BC	4,078	−0.022	0.534	0.544	−0.256	0.613	9.2 (1.12)
C	3,298	−0.009	0.533	0.543	−0.236	0.602	11.0 (1.00)
CD	5,998	−0.001	0.491	0.529	−0.225	0.585	12.8 (1.23)
DE	2,996	0.022	0.431	0.491	−0.264	0.470	17.0 (1.76)
CHY	1,712	0.0	0.432	0.448	−0.269	0.524	12.3 (1.34)
ILA	968	0.0	0.469	0.533	−0.214	0.747	7.8 (0.88)
TAP	958	0.0	0.448	0.460	−0.224	0.651	9.9 (0.98)
TCU	1,325	0.0	0.497	0.515	−0.248	0.547	12.8 (1.18)
Model MW2010, $\sigma_{\varepsilon,GMPE} = 0.55$, $M_W > 6.0$							
All data	1,876	−0.037	0.506	0.538	−0.197	0.598	15.1 (1.36)
BC	1,144	−0.035	0.551	0.554	−0.216	0.610	12.3 (1.27)
C	926	−0.017	0.549	0.551	−0.146	0.695	15.9 (1.25)
CD	1,585	−0.005	0.504	0.536	−0.206	0.575	15.6 (1.47)
DE	732	0.004	0.427	0.502	−0.214	0.500	21.8 (2.89)
CHY	400	0.0	0.425	0.460	−0.167	0.619	18.0 (2.13)
ILA	266	0.0	0.452	0.462	−0.437	0.485	5.5 (1.25)
TAP	254	0.0	0.422	0.433	−0.289	0.639	6.9 (1.40)
TCU	402	0.0	0.528	0.564	−0.285	0.501	12.2 (1.23)
Model ML2002, $\sigma_{\varepsilon,GMPE} = 0.605$, $M_L > 6.3$							
All data	2,081	−0.041	0.506	0.556	−0.188	0.575	18.3 (1.51)
BC	1,240	−0.059	0.546	0.567	−0.160	0.644	17.2 (1.17)
C	994	−0.051	0.555	0.570	−0.106	0.757	19.4 (1.21)
CD	1,748	−0.032	0.511	0.58	−0.139	0.647	21.1 (1.51)
DE	841	0.014	0.440	0.530	−0.178	0.519	27.8 (2.25)
CHY	427	0.0	0.461	0.517	−0.193	0.520	23.6 (1.69)
ILA	302	0.0	0.378	0.418	−0.441	0.488	5.4 (1.75)

Table 4 continued

Data set	Number of records	Mean residual from the data	Within-earthquake standard deviation		within-earthquake correlation function		Correlation distance, km[a]
			$\sigma_{\varepsilon,DDS}$	$\sigma_{\varepsilon,SVDS}$	a	b	
TAP	374	0.0	0.433	0.446	−0.297	0.532	9.8 (1.60)
TCU	469	−0.015	0.557	0.640	−0.206	0.563	16.5 (1.20)

The characteristics for the raw data are listed in Table 1

[a] Values in parentheses indicate difference (ratio) between corresponding estimates for intermediate-to-large earthquakes and for all earthquakes

residuals, should be considered and incorporated into Eq. 14 for intermediate-to-large earthquakes.

The characteristics of array-dependent correction factors, which may be applied to reduce the within-earthquake correlation, are presented in Table 4. The corrected PGA_{CMOD} values are obtained as

$$\ln PGA_{CMOD}(M_i, R_{i,j}, Vs30_j) = \ln PGA_{MOD}(M_i, R_{i,j}) + CF(M_i, R_{i,j}, Vs30_j); \tag{15}$$

where PGA_{MOD} are the values calculated using the corresponding ground-motion model. Table 4 summarises the parameters of the within-earthquake correlation after applying the correction factors and compares the resulting correlation distances with those estimated from the uncorrected data (see Table 1). The correction allows for the effective reduction in the level of within-earthquake correlation. For example, the correlation distances vary by more than a factor of 2 for the class DE and CHY array residuals and intermediate-to-large earthquakes. Note that the correction also reduces the within-earthquake standard deviation.

6 Conclusions

We analysed peculiarities of the within-earthquake correlation of ground motion in Taiwan using the strong-motion records accumulated by the TSMIP network from 54 shallow earthquakes ($M_L > 5.0$, focal depth <30 km, more than 7,000 records) that occurred between 1993 and 2009. Two ground-motion prediction equations, which were recently developed for the region and which are based on moment and local magnitude and hypocentral distances, were used to calculate the peak ground acceleration and to analyse ground-motion residuals. We also used the database containing average shear-wave velocity Vs30 data for most TSMIP stations (Kuo et al. 2012; see also http://egdt.ncree.org.tw/) as well as the Vs30 map constructed for the whole island (Lee and Tsai 2008) to estimate the Vs30 values and identify a site class for each strong-motion station. The PGA residuals were combined into several groups depending on earthquake magnitude, site classes and particular arrays of the TSMIP network, which are characterised by different geological conditions (e.g., alluvium-filled basins; thick Quaternary strata; relatively stiff soils in extended hilly areas).

We have shown that there is a prominent correspondence between the within-earthquake correlation of PGA residuals and spatial correlation of the Vs30 values, which was estimated for an area characterised by a specific geological structure. The larger correlation between the Vs30 values corresponds to the larger correlation of the PGA residuals, and the relationship

between the correlation distances is apparently linear. On the other hand, the level of within-earthquake correlation may vary significantly depending on site classes, general geological conditions and earthquake magnitude, records of which dominate the analysed dataset. The level of correlation increases with the decrease in average shear-wave velocity and with the increase in earthquake magnitude. At the same time, the correlation is smallest for areas with large spatial variations in deposit thickness (Taipei and Ilan alluvium basins) and largest for thick sediments, which cover large areas (CHY array, Chianan Plain). A direct relationship between the correlation of the thickness of sediments and the within-earthquake correlation of the PGA residuals has been observed in the study. The correlation distances (i.e., the site-to-site distance for which the within-earthquake correlation coefficient $\rho_\varepsilon(\Delta)$ decreases to $1/e = 0.368$) may vary from 4 to 60 km depending on a particular area, earthquake, or group of earthquakes.

The high level of ground-motion correlation reflects the presence of a non-random component in residuals, which may be caused by the joint influence of soft surface soil and thick sediments and by path or azimuthal effects. The point-source approximation of an extended fault (e.g., Chi–Chi earthquake) as well as the neglected hanging- and foot-wall effects may also result in non-random residuals. Thus, a single generalised model of spatial correlation across geologically heterogeneous regions may not be adequate in some cases. At the same time, it has been found that the application of empirical correction factors, which consider earthquake magnitude, source-to-site distance and average shear-wave velocity for a given station and which can be developed for particular arrays, allows for the effective reduction in the level of within-earthquake correlation (more than 2 times) and within-earthquake standard deviation.

The results of the analysis, which are summarised in Tables 1, 3 and 4, may be used to make practical estimates of seismic hazard, damage and loss for widely located building assets and spatially distributed structures in Taiwan as well in other regions with similar geological characteristics. However, the sensitivity of the results to variations in input data and particular models should be considered, and several correlation models should be used in the Logic Tree scheme. Another important application of the developed models of within-earthquake correlation is the characterisation of uncertainty in Shake Maps (e.g., Wald et al. 2008) using conditional simulations of ground-motion fields, given the observations from recording stations (Park et al. 2007; Crowley et al. 2008b).

Future tasks include evaluating the correlation structure for other parameters of ground motion, which are used for seismic loss assessments, namely, peak velocities and spectral acceleration at various oscillation frequencies. It seems rational to select data based on geological characterisation (e.g., rocky sites covered by very thin superficial layers, sedimentary-filled basins with gradual or steep increase of thickness of sediments, large plain area) rather than based on whether or not they belong to a particular array. Establishing quantitative relationships between parameters of spatial correlation of ground motion and characteristics of spatial variability of geology (e.g., Eqs. 13) would be useful for regions for which strong motion data of real engineering significance are completely unavailable or very scarce but which are characterised by sufficient geological and geotechnical databases.

Acknowledgments We would like to thank Chun-Te Chen, Chun-Hsiang Kuo, Chyi-Tyi Lee and Che-Min Lin for providing the necessary data used in this study. The comments made by the anonymous reviewers greatly improved this article. TSMIP ground-motion data were obtained from the Central Weather Bureau of the Republic of China. This work was sponsored by Deutsche Forschungsgemeinschaft (DFG), Germany, project WE 1394/18-1.

References

Abrahamson NA, Youngs RR (1992) A stable algorithm for regression analysis using the random effects model. Bull Seismol Soc Am 82:505–510

Abrahamson N, Atkinson G, Boore D, Bozorgnia Y, Campbell K, Chiou B, Idriss IM, Silva W, Youngs R (2008) Comparison of the NGA ground-motion relations. Earthq Spectra 24(1):45–66. doi:10.1193/1.2924363

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723. doi:10.1109/TAC.1974.1100705

Alatik L, Abrahamson N, Bommer JJ, Scherbaum F, Cotton F, Kuehn N (2010) The variability of ground-motion prediction models and its components. Seismol Res Lett 81(5):794–801. doi:10.1785/gssrl.81.5.794

Anderson JG, Uchiyama Y (2011) A methodology to improve ground-motion prediction equations by including path corrections. Bull Seismol Soc Am 101:1822–1846. doi:10.1785/0120090359

Atkinson GM (2006) Single-station sigma. Bull Seismol Soc Am 96:446–455. doi:10.1785/0120050137

Baker JW, Jayaram N (2008) Effects of spatial correlation of ground motion parameters for multi-site seismic risk assessment: collaborative research with Stanford University and AIR. Final Technical Report. Report for U.S. Geological Survey National Earthquake Hazards Reduction Program (NEHRP) External Research Program Award 07HQGR0031

Boore DM, Gibbs JF, Joyner WB, Tinsley JC, Ponti DJ (2003) Estimated ground motion from the 1994 Northridge, California, Earthquake at the site of the Interstate 10 and La Cienega Boulevard Bridge collapse. West Los Angeles, California. Bull Seismol Soc Am 93:2737–2751. doi:10.1785/0120020197

Boore DM, Joyner WB, Fumal TE (1997) Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work. Bull Seismol Soc Am 68:128–153

Brillinger DR, Preisler HK (1984) An exploratory analysis of the Joyner-Boore attenuation data. Bull Seismol Soc Am 74:1441–1450

Brillinger DR, Preisler HK (1985) Further analysis of the Joyner-Boore attenuation data. Bull Seismol Soc Am 75:611–614

Campbell KW, Thenhaus PC, Barnard TP, Hampson DB (2002) Seismic hazard model for loss estimation and risk management in Taiwan. Soil Dyn Earthq Eng 22:743–754

Chen KJ (1998) S-wave attenuation structure in the Taiwan area and its correlation with seismicity. Terrest Atmos Ocean Sci (TAO) 9(1):97–118

Cheng CT, Chiou SJ, Lee CT, Tsai YB (2007) Study of probabilistic seismic hazard maps of Taiwan after Chi-Chi earthquake. J GeoEng 2(1):19–28 http://www.sinotech.org.tw/gerc-ctr/2007.files/papers_pdf/cheng/2007-8.pdf

Chung JK, Chen YL, Shin TC (2009) Spatial distribution of coda Q estimated from local earthquakes in Taiwan area. Earth Planets Space 61:1077–1088

Crowley H, Bommer JJ, Stafford PJ (2008) Recent developments in the treatment of ground- motion variability in earthquake loss model. J Earthq Eng 12(S):71–80. doi:10.1080/13632460802013529

Crowley H, Stafford PJ, Bommer JJ (2008) Can earthquake loss models be validated using field observations? J Earthq Eng 12:1078–1104. doi:10.1080/13632460802212923

Douglas J (2003) Earthquake ground motion estimation using strong-motion records: a review of equations for estimation of peak ground acceleration and spectral ordinates. Earth Sci Rev 61:43–104

Douglas J (2006) Errata and additions to "Ground motion estimation equations 1966–2003", BRGM/RP-54603-FR

Esposito S, Iervolino I (2011) PGA and PGV spatial correlation models based on European multievent datasets. Bull Seismol Soc Am 101:2532–2541. doi:10.1785/0120110117

Furumura M, Furumura T, Wen KL (2001) Numerical simulation of Love wave generation in the Ilan Basin, Taiwan, during the 1999 Chi-Chi earthquake. Geoph Res Lett 28(17):3385–3388

Goda K (2011) Interevent variability of spatial correlation of peak ground motions and response spectra. Bull Seismol Soc Am 101:2522–2531. doi:10.1785/0120110092

Goda K., Atkinson GM (2009) Probabilistic characterization of spatially correlated response spectra for earthquakes in Japan. Bull Seismol Soc Am 99:3003–3020. doi:10.1785/0120090007

Goda K, Atkinson GM (2010) Intraevent spatial correlation of ground-motion parameters using SK-net data. Bull Seismol Soc Am 100:3055–3067. doi:10.1785/0120100031

Goda K, Hong HP (2008a) Spatial correlation of peak ground motions and response spectra. Bull Seismol Soc Am 98:354–365. doi:10.1785/0120070078

Goda K, Hong HP (2008b) Estimation of seismic loss for spatially distributed buildings. Earthq Spectra 24:889–910. doi:10.1193/1.2983654

Goovaerts P (1997) Geostatistics for Natural Resources Evaluation. Oxford Univ. Press, New-York
Hong HP, Zhang Y, Goda K (2009) Effect of spatial correlation on estimated ground motion-prediction equations. Bull Seismol Soc Am 99:928–934. doi:10.1785/0120080172
Hsu HM, Chao SJ, Hwang H (2008) Assessment of ground shaking in Ilan County, Taiwan. J Mar Sci Technol 16(4):308–313
Jayaram N, Baker JW (2009) Correlation model for spatially-distributed ground-motion intensities. Earthq Eng Struct Dyn 38:1687–1708. doi:10.1002/eqe.922
Johnson ME (1987) Multivariate statistical simulation. Wiley Series in Probability and Mathematical Statistics, Los Alamos National Laboratory, Los Alamos
Joyner WB, Boore DM (1993) Methods for regression analysis of strong-motion data. Bull Seismol Soc Am 83:469–487
Kawakami H, Mogi H (2003) Analysing spatial intraevent variability of peak ground acceleration as a function of separation distance. Bull Seismol Soc Am 93:1079–1090. doi:10.1785/0120020026
Kuo CH, Wen KL, Hsieh HH, Chang TM, Lin CM, Chen CT (2011) Evaluating empirical regression equations for Vs and estimating Vs30 in northeastern Taiwan. Soil Dyn Earthq Eng 31(3):431–439. doi:10.1016/j.soildyn.2010.09.012
Kuo CH, Wen KL, Hsieh HH, Lin CM, Chang TM, Kuo KW (2012) Site classification and Vs30 estimation of free-field TSMIP stations using the logging data of EGDT. Eng Geol (in press). doi:10.1016/j.enggeo.2012.01.013
Lee CT, Tsai BR (2008) Mapping Vs30 in Taiwan. Terrest Atmos Ocean Sci (TAO) 19:671–682. doi:10.3319/TAO.2008.19.6.671(PT)
Lee CT, Cheng CT, Liao CW, Tsai YB (2001) Site classification of Taiwan free-field strong-motion stations. Bull Seismol Soc Am 91:1283–1297
Lin CM, Chang TM, Huang YC, Chiang HJ, Kuo CH, Wen KL (2009) Shallow S-wave velocity structures in the western coastal plain of Taiwan. Terrest Atmos Ocean Sci (TAO) 20:299–308. doi:10.3319/TAO.2007.12.10.01(T)
Lin PS, Chiou B, Abrahamson N, Walling M, Lee CT, Cheng CT (2011) Repeatable source, site and path effects on the standard deviation for empirical ground-motion prediction models. Bull Seismol Soc Am 101:2281–2295. doi:10.1785/0120090312
Lin PS, Lee CT (2008) Ground motion attenuation relationships for subduction-zone earthquakes in northeastern Taiwan. Bull Seismol Soc Am 98:220–240. doi:10.1785/0120060002
Liu KS, Shin TC, Tsai YB (1999) A free-field strong motion network in Taiwan: TSMIP. Terrest Atmos Ocean Sci (TAO) 10:377–396
Liu KS, Tsai YB (2005) Attenuation relation of peak ground acceleration and velocity for crustal earthquakes in Taiwan. Bull Seismol Soc Am 95:1045–1058. doi:10.1785/0120040162
Miksat J, Wen KL, Sokolov V, Chen CT, Wenzel F (2010) Simulating the Taipei Basin response by numerical modeling of wave propagation. Bull Earthq Eng 8(4):847–858. doi:10.1007/s10518-009-9171-0
Molas GL, Anderson R, Seneviratna P, Winkler T. (2006) Uncertainty of portfolio loss estimates for large earthquakes. In: Proceedings of first European conference on earthquake engineering and seismology, Geneva, Switzerland, 3–8 September 2006: Paper 1117
Morikawa N, Kanno T, Narita A, Fujiwara H, Okumura T, Fukushima Y, Guerpinar A (2008) Strong motion uncertainty determined from observed records by dense network in Japan. J Seismol 12:529–546. doi:10.1007/s10950-008-9106-2
Park J, Bazzurro P, Baker JW (2007) Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimations. In: Kanda J, Takada T, Furuta H (eds) Applications of statistics and probability in civil engineering. Taylor & Francis Group, London pp 1–8
Schwarz GE (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464. doi:10.1214/aos/1176344136
Sokolov V, Loh CH, Wen KL (2004) Evaluation of generalized site response functions for typical soil classes (B, C and D) in the Taiwan region. Earthq Spectra 20(4):1279–1316
Sokolov V, Loh CH, Wen KL (2006) Strong ground motion source scaling and attenuation models for earthquakes located in different source zones in Taiwan. In: Proceedings of the 4th intenational conference on earthquake engineering, Taipei, Taiwan, October 12–13 2006, paper 003
Sokolov V, Loh CH, Jean WY (2007) Application of Horizontal-to-Vertical (H/V) Fourier spectral ratio for analysis of site effect on rock (NEHRP-class B) sites in Taiwan, Soil. Dyn Earthq Eng 27(4):314–323. doi:10.1016/j.soildyn.2006.09.001
Sokolov V, Wen KL, Miksat J, Wenzel F, Chen CT (2009) Analysis of Taipei basin response for earthquakes of various depths and locations using empirical data. Terrest Atmos Ocean Sci (TAO) 20(5):687–702. doi:10.3319/TAO.2008.10.15.01(T)

Sokolov V, Wenzel F (2011) Influence of spatial correlation of strong ground-motion on uncertainty in earthquake loss estimation. Earthq Eng Struct Dyn 40:993–1009. doi:10.1002/eqe.1074
Sokolov V, Wenzel F (2011) Influence of ground-motion correlation on probabilistic assessments of seismic hazard and loss: sensitivity analysis. Bull Earthq Eng 9:1339–1360. doi:10.1007/s10518-011-9264-4
Sokolov V, Wenzel F, Jean WY, Wen KL (2010) Uncertainty and spatial correlation of earthquake ground motion in Taiwan. Terrest Atmos Ocean Sci (TAO) 21:905–921. doi:10.3319/TAO.2010.05.03.01(T)
Strasser FO, Abrahanson NA, Bommer JJ (2009) Sigma: issues, insights, and challenges. Seismol Res Let 80(1):40–56. doi:10.1785/gssrl.80.1.40
Tsai CCP, Chen YH, Liu CH (2006) The path effect in ground-motion variability: an application of the variance-component technique. Bull Seismol Soc Am 96:1170–1176. doi:10.1785/0120050155
Wald DJ, Lin KW, Quitoriano V (2008) Quantifying and qualifying USGS ShakeMap uncertainty. USGS Open-File Report 2008-1238
Wang CY, Lee YH, Ger ML, Chen YL (2004) Investigating subsurface structures and P- and S-wave velocities in the Taipei Basin. Terr Atmos Ocean Sci 15(4):609–627
Wang JH (1998) Q values of Taiwan: a review. J Geol Soc China 36:15–24
Wang JH (2008) Urban seismology in the Taipei metropolitan area: review and prospective. Terrest Atmos Ocean Sci (TAO) 19:213–234. doi:10.3319/TAO.2008.19.3.213(T)
Wang M, Takada T (2005) Macrospatial correlation model of seismic ground motions. Earthq Spectra 21(4):1137–1156. doi:10.1193/1.2083887
Wesson RL, Perkins DM (2001) Spatial correlation of probabilistic earthquake ground motion and loss. Bull Seismol Soc Am 91:1498–1515. doi:10.1785/0120000284