Soil Dynamics and Earthquake Engineering 43 (2012) 229–236

Contents lists available at SciVerse ScienceDirect

Soil Dynamics and Earthquake Engineering

journal homepage: www.elsevier.com/locate/soildyn

Empirical correlations between cumulative absolute velocity and spectral accelerations from NGA ground motion database

Gang Wang *, Wenqi Du

Hong Kong University of Science and Technology, Department of Civil and Environmental Engineering, Clear Water Bay, Kowloon, Hong Kong SAR, China

ARTICLE INFO

Article history:
Received 30 May 2011
Received in revised form
12 July 2012
Accepted 21 July 2012

ABSTRACT

Considering multiple ground motion intensity measures is important in seismic hazard analysis and ground motion selection process. Using the NGA strong motion database and recently developed ground-motion prediction models, empirical correlations are developed between cumulative absolute velocity (CAV) and spectral accelerations (Sa) at periods from 0.01 to 10 s. The CAV–Sa correlations at long periods are significantly influenced by rupture distance due to modification of the frequency content and duration of the acceleration time history through travel path. Similarly, the presence of strong velocity pulses in near-source ground motions also affects the correlations at moderate to long periods. On the other hand, the correlations are not particularly sensitive to the earthquake magnitude, orientation of the ground-motion recordings, selection of ground-motion prediction models and local site conditions. Piecewise linear fitting equations are provided to quantify the correlations for various cases. The application of the CAV–Sa correlations in ground motion selection process is also discussed.

1. Introduction

Ground motion intensity measures (IMs) are often used to represent different characteristics of earthquake ground motions. The commonly used ground motion intensity measures (IMs) can be summarized into different categories: (a) the peak values of ground motion time histories, including the peak ground acceleration (PGA), peak ground velocity (PGV), and peak ground displacement (PGD); (b) the response spectrum representing the peak response of a single-degree-of-freedom (SDOF) system; (c) the integral forms of the spectral acceleration values over period ranges of interest, including Spectrum Intensity [1], Acceleration Spectrum Intensity [2], and more recently, Displacement Spectrum Intensity [3]; (d) intensity measures related to the time integration of acceleration time histories, including Arias Intensity [4], Cumulative Absolute Velocity and its variants [5,6], as well as the time rate of Arias Intensity [7]; and (e) other intensity measures such as the duration of earthquake ground motions.

Correlating critical ground motion IMs with structural damages has been one of the most important topics in earthquake engineering. Particularly, the spectral acceleration (Sa) is one of the most popular ground motion IMs in earthquake engineering. Previous studies indicate that inelastic structural response is closely correlated with the spectral accelerations of input motions over an appropriate period range [8]. However, due to complexity of ground motion time histories, no single IM was found to be sufficient to characterize all important aspects of earthquake ground motions. For example, Cumulative Absolute Velocity (CAV) includes the cumulative effects of the ground shaking duration, which cannot be captured by the peak values of ground motion, such as PGA, and its response spectrum. CAV and its variants are found to have superior capacities than PGA in predicting the earthquake-induced soil liquefaction [9,10]. Therefore, it is necessary to incorporate multiple ground motion parameters in practice to better characterize earthquake ground motions. For example, the US Nuclear Regulatory Commission adopted both CAV and spectral acceleration checks in early warning systems to mandate the shutdown of a nuclear power plant when an earthquake occurs [11].

However, little effort has been devoted in the past to the statistical relationships between CAV and spectral accelerations. This study will develop empirical correlations between CAV and spectral accelerations based on the NGA strong motion database and recently developed ground-motion prediction models. Influencing factors such as earthquake magnitude, rupture distance, local site conditions, and presence of velocity pulses in the ground motion time histories will be studied. The proposed correlation model provides useful means to characterize the joint occurrence of these two important IMs, whose application in ground motion selection will be briefly discussed.

2. Empirical CAV–Sa correlation

Cumulative Absolute Velocity (CAV) is defined as the time integration of the absolute acceleration time history as follows:

$$\mathrm{CAV} = \int_0^{t_{total}} |a(t)|\,dt \tag{1}$$

* Corresponding author. Tel.: +852 2358 7161; fax: +852 2358 1534.
E-mail address: gwang@ust.hk (G. Wang).

0267-7261/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.soildyn.2012.07.029

where $|a(t)|$ is the absolute value of the acceleration time history, and t_{total} is the total duration of the ground motion. CAV assumes the unit of $g \cdot s$ throughout the paper, where g is gravitational acceleration. Since small-amplitude accelerations usually do not contribute to structural damages, several variations of CAV were proposed to exclude these small-amplitude accelerations into the time integration in Eq. (1) so as to better represent long period characteristics of the ground motion. EERI (1991) proposed a standardized CAV, which only integrates accelerations whose peak value exceeds a threshold value of 0.025g within a one-second time interval. Another useful CAV variant, CAV_5, is defined to exclude these acceleration values smaller than a threshold value of 5 cm/s^2 [9]. However, exclusion of low-amplitude accelerations from contributing to CAV variants also results in a relative small number of nonzero values, leading to a decreased stability in predicting these IMs [12]. Therefore, in this study we will limit our efforts to the development of correlations between CAV and spectral accelerations. Correlations for CAV variants will be a subject of a future study.

Among a few available ground-motion prediction models for CAV, the model recently developed by Campbell and Bozorgnia [12] is used in this study (henceforth referred to as the CB2010 model). The model is built on the efforts of the PEER-Next Generation Attenuation (NGA) relationships for shallow crustal earthquakes in active tectonic regions. One distinctive feature of the CB2010 model is that it utilizes the same functional form and database as these used in developing the NGA model for spectral accelerations (termed as the CB2008 model) by the same authors [13]. Detailed information about the strong motion database is provided in [14]. The ground motion database was carefully compiled to exclude aftershock records, non-free field site conditions, low quality data, non-shallow crustal events and data with missing horizontal components. This study employs a subset of the above database containing 1363 pairs of horizontal recordings from 59 events with magnitudes ranging from 5–7.9 and rupture distances ranging from 0 to 200 km. Due to limitations of strong motion instruments, each recorded time history has a maximum usable period that is related to filtering in signal processing to remove long-period noise. Filtering results in suppression of ground motion amplitudes and energy at periods longer than the maximum usable period, such that the motion is not representative of the real ground response over those period ranges. Therefore, only records with a maximum usable period greater than the specified spectral period will be used to compute the CAV–Sa correlation at that period. Hence the number of usable records used in computing CAV–Sa correlations decreases as the spectral period increases. For example, a total of 1363 pairs of records are used at period of 0.01 s, and the number decreases to 533 at period of 10 s.

The ground-motion prediction models typically assume the observed logarithmic spectral acceleration at a spectral period T, denoted as $\ln Sa(T)$, and the logarithmic cumulative absolute velocity, denoted as $\ln CAV$, of a ground motion record follow lognormal distribution, and can be written as follows:

$$\ln \mathrm{Sa}(T) = \mu_{\ln Sa}(M,R,\theta,T) + \varepsilon_{\ln Sa(T)}\sigma_{\ln Sa(T)} \tag{1}$$

$$\ln \mathrm{CAV} = \mu_{\ln CAV}(M,R,\theta) + \varepsilon_{\ln CAV}\sigma_{\ln CAV} \tag{2}$$

where $\mu_{\ln Sa}$ and μ_{lnCAV} are the median estimates of lnSa and lnCAV by ground-motion prediction models, e.g. [12,13], which are a function of the earthquake magnitude (M), certain measures of source-to-site distance (R), and other parameters (θ) that might be related to styles of faulting and local site conditions. The total residuals, denoted by terms $\varepsilon_{\ln Sa(T)} \cdot \sigma_{\ln Sa(T)}$ and $\varepsilon_{\ln CAV} \cdot \sigma_{\ln CAV}$, quantify the differences between the observed and the predicted IMs. The total residuals represent the aleatory uncertainty of the intensity measure and are often found to be normally distributed. The standard deviation of the total residuals is also often provided by the ground-motion prediction model, and is denoted as $\sigma_{\ln Sa(T)}$ and $\sigma_{\ln CAV}$ respectively. The CB2008 and CB2010 models showed that the aleatory uncertainty associated with CAV is significantly smaller than that associated with spectral accelerations at any period, e.g., $\sigma_{\ln CAV}$ is 20%, 33%, 40% and 49% smaller than $\sigma_{\ln Sa(T)}$ at $T=0.01$ s, 1 s, 5 s and 10 s, respectively. Therefore, CAV has less aleatory uncertainty and a higher degree of predictability than spectral accelerations.

The epsilon terms, $\varepsilon_{\ln Sa(T)}$ and $\varepsilon_{\ln CAV}$, in Eqs. (1) and (2) respectively represent the number of standard deviations by which the difference between the observed and the predicted logarithmic IMs is measured [15]. Epsilons are also called standardized total residuals, standard scores, or z-scores. They have often been used in probabilistic seismic hazard analysis and hazard degradation. To compute the epsilons, the predicted logarithmic IMs and their standard deviations of total residuals are obtained for each record and each spectral period using prediction models. We performed statistical analysis and verified that $\varepsilon_{\ln Sa(T)}$ or $\varepsilon_{\ln CAV}$ can be well approximated by a standard normal distribution with an expected zero mean and a unit standard deviation. Fig. 1 shows the scatterplots of the obtained $\varepsilon_{\ln Sa(T)}$ and $\varepsilon_{\ln CAV}$ at different periods. No obvious nonlinear pattern is observed between these two variables, indicating that the correlation coefficient could provide a reliable measure of the strength of the linear relationship between them. The correlation coefficients between epsilon terms are evaluated using the well-known Pearson product-moment correlation coefficient formulation [16]

$$\rho_{\ln CAV,\ln Sa(T)} = \frac{\sum_{i=1}^{n}(\varepsilon_{\ln CAV}^{(i)} - \overline{\varepsilon_{\ln CAV}})(\varepsilon_{\ln Sa(T)}^{(i)} - \overline{\varepsilon_{\ln Sa(T)}})}{\sqrt{\sum_{i=1}^{n}(\varepsilon_{\ln CAV}^{(i)} - \overline{\varepsilon_{\ln CAV}})^2 \sum_{j=1}^{n}(\varepsilon_{\ln Sa(T)}^{(j)} - \overline{\varepsilon_{\ln Sa(T)}})^2}} \tag{3}$$

where N is the total number of usable records for the correlation at period T, the superscript (i) denotes that the term is evaluated for the i-th record. The bar overhead designates the sample mean of each variable.

Using Eq. (3), the correlation coefficient is evaluated at 21 discrete spectral periods from 0.01 s to 10 s, as is illustrated in Fig. 2. It may be tempting to tabulate these values directly for each period, yet a simple fitting function is much more convenient for practical application. As is observed in Fig. 2, the correlation coefficient attains a peak value at $T=0.01$ s, and decreases as T increases to 0.1 s. The correlation coefficient increases again at moderate periods around 0.5 s, which approximately corresponds to the predominant period of earthquakes. The correlation coefficient generally decreases when T increases from 0.5 s to 10 s. Note that the correlations are between epsilon terms of two IMs, not their absolute values. The results can be interpreted as such that if CAV of a ground motion is greater than its expected value, its spectral accelerations at short-to-moderate periods, compared with these at long periods, are more likely to be greater than their expected values. Based on above observations, the following piecewise linear equation is proposed to approximate the CAV–Sa correlation:

$$\rho(T) \equiv \rho_{\ln CAV,\ln Sa(T)} = \rho_i + \frac{\log(T/T_i)}{\log(T_{i+1}/T_i)}(\rho_{i+1} - \rho_i) \ \text{ if } \ T_i \le T \le T_{i+1} \tag{4}$$

The above equation shows the correlation coefficient is linearly interpolated (in log T scale) between ρ_i and ρ_{i+1} values at T_i and T_{i+1} respectively. Usually, a linear interpolation through five or six discrete points can well approximate the empirical data. As is shown in Fig. 2, six periods T_1, T_2, T_3, T_4, T_5, and T_6 are chosen as 0.025 s, 0.12 s, 0.5 s 2 s, 4 s, and 10 s respectively to capture the overall shape of the correlation coefficients over the period range

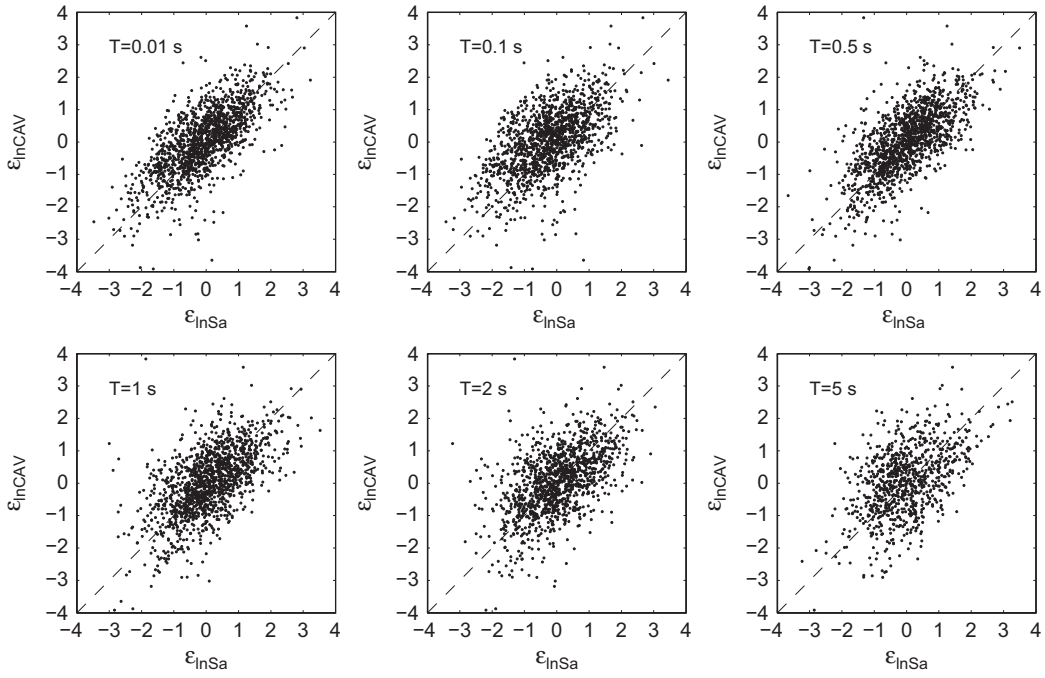

Fig. 1. Scatterplots of epsilons for lnCAV and lnSa(T) at different spectral periods.

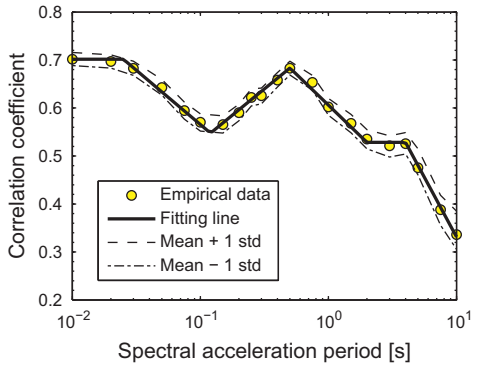

Fig. 2. CAV–Sa correlations using the CB2008 NGA model.

from 0.01 s to 10 s. The correlation coefficient assumes to be constant if $T < T_1$. Nonlinear regression analysis was performed to determine ρ_i values that best fit all empirical data, and the results are listed in Table 1. The misfit error is generally smaller than 0.01 at any period. The piecewise linear function will also be used in the later sections to quantify the correlation coefficients for various cases. We have to clarify that the fitting function is provided in this study solely as a convenient way to represent the empirical data, without intention for any theoretical implication.

As only a finite number of recordings are used to determine the correlation coefficient in Eq. (3), the result contains epistemic uncertainty, which can be estimated using Fisher transformation [16]. The Fisher transformation converts the correlation coefficient ρ into a transformed variable z

$$z = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right) = \tanh^{-1}\rho \tag{5}$$

where z approximately follows a normal distribution. One important property of the Fisher transformation is that it is an approximate variance-stabilizing method: while the variance of the correlation coefficient ρ usually becomes smaller when it approaches 1 or -1, the variance of the transformed variable z is approximately constant. The standard deviation of z can be estimated based on sample size via the following equation:

$$std(z) = \frac{1}{\sqrt{N-3}} \tag{6}$$

where N is the sample size. It is obvious that the degree of uncertainty in sample correlation decreases when sample size increases. Accordingly, the inverse Fisher transformation can be used to estimate the uncertainty in ρ. As is shown in Fig. 2, the estimated standard deviation of ρ is less than 0.05 at all periods, indicating the uncertainty is relatively low. Alternatively, one can estimate such finite sample uncertainty using other statistical methods such as the bootstrap method [17]. Although not shown in the paper, results from these two methods agree very well.

Selection of a particular ground-motion prediction equation and strong-motion database is another source of uncertainty in determining the empirical correlation. In the previous session, the total residuals associated with spectral accelerations are based on the CB2008 NGA model. Three other NGA models, namely, the Abrahamson–Silva model (AS2008), Boore–Atkinson model (BA2008) and Chiou–Youngs model (CY2008) [18] were also used to derive the CAV–Sa correlations. As is shown in Fig. 3, the uncertainty resulted from using different NGA models is practically insignificant. The fitting line to the averaged case (solid line) is plotted in Fig. 3 and its fitting parameters are also provided in Table 1. It is interesting to observe that using the CB2008 model results in the highest correlation coefficients at almost all periods, probably due to the advantage that both the CB2008 NGA model and CB2010 CAV model employ the same functional form and database in developing these models. Therefore, their residuals tend to have a higher degree of correlation than other models. To better estimate the uncertainty associated with using different NGA models and databases, analyses were performed using a combination of four NGA models and their associated databases (16 cases in total). The estimated mean correlation coefficients ± 1 standard deviation (dashed lines) are also plotted in Fig. 3. In general, the standard deviation of the estimated correlation coefficients associated with using different NGA models is around 0.04 at short periods, and 0.02 to 0.03 at long periods, as is reported in Table 1.

Table 1
CAV–SA correlations using different NGA models.

NGA models	T_1	T_2	T_3	T_4	T_5	T_6	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
CB2008	0.025	0.12	0.5	2	4	10	0.70	0.55	0.68	0.53	0.53	0.33
Averaged case	0.025	0.12	0.5	2	4	10	0.63	0.49	0.63	0.50	0.50	0.30
(std.)							(0.04)	(0.04)	(0.04)	(0.02)	(0.02)	(0.03)

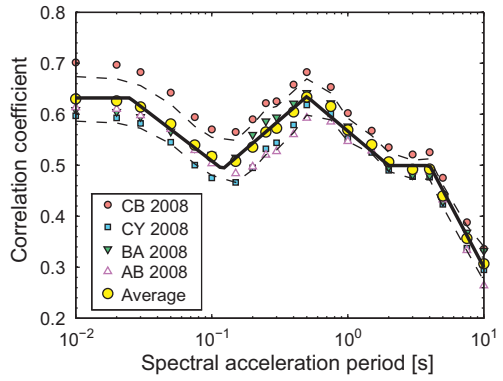

Fig. 3. Comparison of CAV–Sa correlations using different NGA models (solid line: the average fitting line; dashed lines: mean $\pm$ 1 standard deviation).

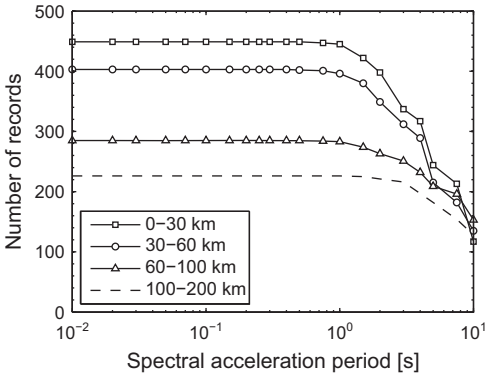

Fig. 4. Number of records over spectral periods in each distance bin.

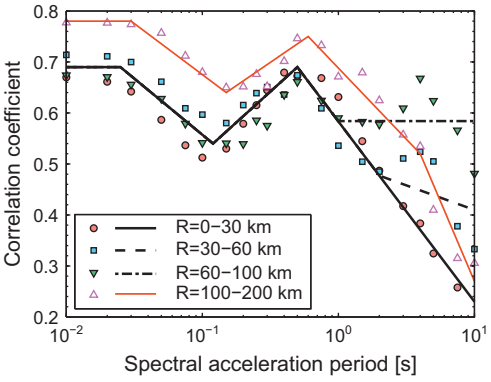

Fig. 5. CAV–Sa correlations for different rupture distance bins (empirical data is represented as symbols, fitting functions as lines).

As the recorded ground motions usually have two orthogonal horizontal components, we further examine whether the CAV–Sa correlations depend on the orientation of the records. It is noted that the correlations are based on "averaged" quantities of two horizontal components. The averaged spectral accelerations used in NGA models, termed as GMRotI50 [19], are independent of installed sensor orientations. On the other hand, the CAV predicted by the CB2010 model is from the geometric mean of two as-recorded horizontal components. Similar to previous studies regarding the spectral accelerations [20], CAVs of records orientated in the fault normal direction is found to be systematically higher than these orientated in the fault parallel direction. However, almost identical CAV–Sa correlations are obtained regardless of the component used. The influence of local site conditions on the CAV–Sa correlations is also studied and no systematic difference has been found. Therefore, we concluded that the correlations are not significantly dependent on the local site conditions.

3. Factors Influencing CAV–Sa correlation

3.1. Influence of rupture distances

Previous studies showed that the correlation between spectral accelerations is not sensitive to the ground motions' causal magnitudes and distances [15,21]. Similarly, the influence of earthquake's magnitude on the CAV–Sa correlation is found to be practically unimportant. However, the CAV–Sa correlation is significantly influenced by rupture distance (defined as the closest distance from the record station to the fault rupture plane), may be due to the fact that the travel path modifies the frequency content and the duration of the ground motion time history. It is worth mentioning that the effects of distance may be well incorporated into modern ground-motion prediction models, yet the correlation measures the degree of the linear relationship between the residuals of ground motion IMs relative to their predicted median values, which are not accounted for by the ground-motion prediction models.

The ground motion database is divided into four distance bins according to rupture distances R=0–30 km, 30–60 km, 60–100 km, 100–200 km. The size of the distance bin is carefully chosen to ensure a sufficient number of data points within the distance bin for statistically stable results. As is shown in Fig. 4, the number of usable records is 226–449 at T=0.01 s in each distance bin, while it reduces to 117 to 153 at T=10 s considering the maximum usable period. Fig. 5 compares the obtained correlation coefficients for each rupture distance bin. Several features are readily observed:

(1) The CAV–Sa correlations are very similar at the short-to-moderate periods (0.01–1 s) for distances within 100 km. For all cases, the largest correlation occurs between CAV and PGA (which can be regarded as Sa evaluated at T=0.01 s). The correlation attains a second peak value at spectral periods of 0.5–0.6 s, indicating that the variability in CAV residuals has the highest degree of association with that in residuals of spectral accelerations around that period range.
(2) For short-distance (R=0–30 km) ground motions, the correlation coefficients at 0.5–10 s decreases linearly with period (in log scale). The correlation coefficients for moderate-distance (R=30–60 km) ground motions are larger than the short-distance results for periods beyond 2 s. Even larger correlations for far-distance (R=60–100 km) records are observed at periods greater than 1 s. The difference probably can be attributed to different wave characteristics as long

travel path tends to filter out high frequency (low period) components of the wave motions. Therefore, far-distance ground motions consist primarily of surface waves excited by travel paths and site effects with stronger long-period frequency content and a longer duration [22–24]. The study shows that the far-distance ground motions have considerably higher CAV–Sa correlations at long periods than the near-source ground motions. The effect may have important implication as long-period ground motions have been increasingly important in seismic design of high-rise buildings, long bridges and base-isolated structures etc.

(3) For distances greater than 100 km, correlations at short-to-moderate periods are generally larger than that for distances within 100 km. It should be noted that relatively few data is available in this distance bin (cf. Fig. 4), therefore the results can be less reliable.

Piecewise fitting functions are also illustrated in Fig. 5 to approximate the empirical data points for each distance bin, and their parameters are summarized in Table 2. Again, these parameters are provided simply to facilitate the representation of the empirical data for the convenience of practical usage. They do not convey any theoretical or physical implication.

3.2. Influence of strong velocity pulse

Near-fault ground motions are of particular interest in earthquake engineering since they have been found to cause excessive damages to structures [25]. Fault rupture directivity effects might be the major reason for the presence of strong velocity pulses in near-fault ground motions. Previous researches have indicated that the strongest pulses are generally more closely aligned with fault-normal directions rather than fault-parallel directions [20]. Systematic differences in response spectrum due to directivity effects have been identified [26]. The quantitative classification scheme proposed by [27] is used in this study to identify and characterize records with pulses. The method used wavelet analysis to extract the largest velocity pulse from a given ground motion time history and classify it as "pulselike" or "non-pulse-like" using the following general criteria: (1) the pulse is large relative to the residual features of the ground motion after the pulse is extracted; (2) the pulse arrives early in the time history, as would be expected for pulses associated with rupture directivity effects; and (3) the absolute velocity amplitudes are large (PGV of the identified record must be equal to or greater than 30 cm/s). Although the causative mechanism of the identified pulselike motion by this method is still not quite certain unless other seismological information can be reviewed, it is believed that many of the identified pulse-like motions are associated with near-fault directivity effects.

A total of 91 fault-normal pulselike records identified in [27] are used in this study. A detailed list of these pulselike records can be found in the reference. Since ground-motion prediction models are usually based on statistical regression using both pulselike and non-pulselike records from a ground motion database, they tend to underestimate the spectral accelerations of pulselike records over a period range around the pulse period [28]. Similarly, we found that the ground-motion prediction model, such as CB2010 model, also tends to underestimate CAVs of pulselike ground motions. As are shown in Fig. 6(a), the total residuals of CAVs associated with pulselike records are positively biased, especially for records with moment magnitude smaller than seven. By comparison, the distribution of total residuals is not biased for near-fault records whose rupture distances are within 30 km, as shown in Fig. 6(b). The near-fault motions used for comparison are restricted to rupture distances within 30 km since the probability of occurrence of a pulse-like record is nearly zero for distances greater than 30 km [27].

Fig. 7 compares the CAV–Sa correlations for all pulselike records with these for near-fault records (R=0–30 km). Although the correlations for two cases generally have similar shapes, the correlations for pulselike records are constantly larger over all spectral periods, with the largest difference close to 0.2 at the spectral period of 2 s. Previous studies showed that the observed response spectra of pulselike records are systematically higher than the predicted values within a narrow-band period range centered around the periods of pulses [28,29]. Similarly, the presence of long-period strong velocity pulses also contributes to the increase in CAV–Sa correlation around the pulse periods.

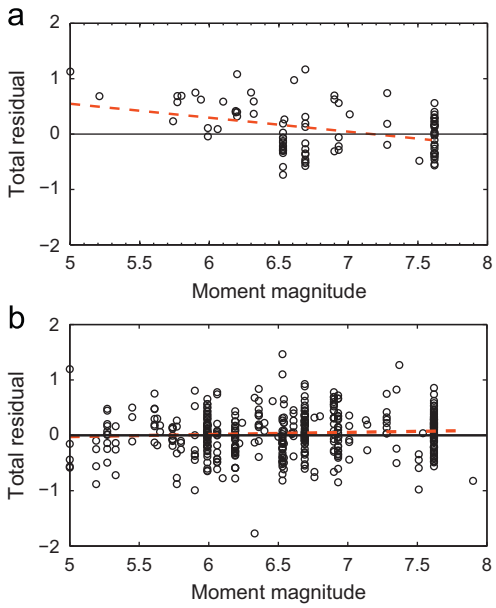

Fig. 6. Distribution of CAV total residuals for (a) pulselike records, and (b) records within rupture distance of 30 km.

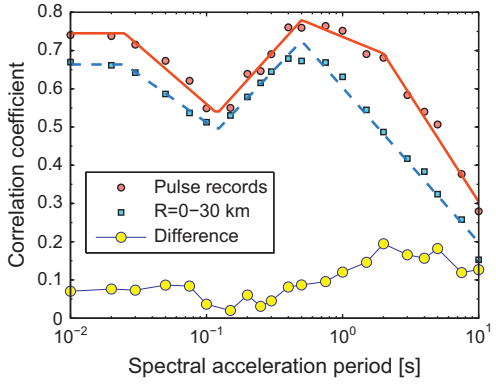

Fig. 7. CAV–Sa correlations for pulse records.

Table 2
CAV–Sa correlations for different rupture distance bins.

Distances (km)	T_1	T_2	T_3	T_4	T_5	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
0–30	0.025	0.12	0.5	–	10	0.69	0.54	0.69	–	0.23
30–60	0.025	0.12	0.5	2	10	0.69	0.54	0.69	0.48	0.41
60–100	0.025	0.12	0.5	1	10	0.69	0.54	0.69	0.58	0.58
100–200	0.03	0.15	0.6	4	10	0.78	0.64	0.75	0.52	0.27

Table 3 summarizes the fitting parameters of Eq. (4) for the CAV–Sa correlation of the pulselike ground motions.

4. Intra-event and inter-event Correlations

Following the concept of the random effect model [30], the total residual terms in Eqs. (1) and (2) can be partitioned into inter-event residual terms and intra-event residual terms as follows,

$$\varepsilon_{\ln CAV}\sigma_{\ln CAV} = \eta_{\ln CAV}\tau_{\ln CAV} + \varepsilon'_{\ln CAV}\sigma'_{\ln CAV} \quad (7)$$

$$\varepsilon_{\ln Sa(T)}\sigma_{\ln Sa(T)} = \eta_{\ln Sa(T)}\tau_{\ln Sa(T)} + \varepsilon'_{\ln Sa(T)}\sigma'_{\ln Sa(T)} \quad (8)$$

The inter-event residuals $\eta\tau$ and intra-event residuals $\varepsilon'\sigma'$ are usually assumed to be normally-distributed with means of zero and standard deviations of τ and σ' , respectively. Similar to the epsilon terms, the normalized inter- and intra-event residuals, η and ε' , follow a standard normal distribution.

The inter-event residual terms $\eta\cdot\tau$ represent the variation in residuals *between* different earthquake events, and it remains the same for all records from a given event. The intra-event residual term $\varepsilon'\sigma'$ represents the variation in residuals within a given event, and it is distinctive for each observed ground motion. Since the inter-event and intra-event residuals are independent, the standard deviation of the total residuals can be evaluated as follows:

$$\sigma_{\ln CAV} = \sqrt{(\tau_{\ln CAV})^2 + (\sigma'_{\ln CAV})^2} \quad (9)$$

$$\sigma_{\ln Sa} = \sqrt{(\tau_{\ln Sa})^2 + (\sigma'_{\ln Sa})^2} \quad (10)$$

First, the statistical distribution of CAV residual terms is examined using Kolmogorov–Smirnov (*K–S*) test. Fig. 8 plots the CDFs of the inter-event residuals and intra-event residuals associated with CAV. A theoretical normal distribution CDF, which has a mean of zero and standard deviation obtained from samples, is also plotted for comparison. The plots show the CDFs of empirical data lies within $\pm$ 10% *K–S* bounds. Therefore, the null hypothesis that the inter-event residuals and intra-event residuals associated with CAV follow normal distribution can not be rejected at 10% significant level. Moreover, statistical tests have been performed to test the normality of the inter- and intra-event residual terms associated with spectral accelerations [31]. Similar statistical tests were also performed and the normality between lnSa(*T*) and lnCAV were verified.

Table 3
CAV–Sa correlations for pulse records.

	T_1	T_2	T_3	T_4	T_5	ρ_1	ρ_2	ρ_3	ρ_i	ρ_5
Pulse records	0.025	0.12	0.5	2	10	0.75	0.54	0.78	0.69	0.30

The η and ε' terms associated with CAV and spectral accelerations were back-calculated for each ground-motion record in the database using the maximum likelihood approach proposed by [30]. Pearson product-moment correlation coefficient formulation Eq. (3) can be applied to obtain the inter-event correlation between the η terms of CAV and Sa, denoted as $\rho_{\eta_{\ln CAV},\eta_{\ln Sa}}$, and the intra-event correlation between the ε' terms of CAV and Sa, denoted as $\rho_{\varepsilon'_{\ln CAV},\varepsilon'_{\ln Sa}}$. Fig. 9 illustrates the obtained inter- and intra-event correlations, where the shaded areas show the range of the estimated mean $\pm$ one standard deviation of the correlations due to the finite sample uncertainty described in Session 2 using the bootstrap method. Relatively large uncertainty in the estimated inter-event correlations is observed due to the limited number of event terms (59 events are used in this study). On the other hand, the estimated intra-event correlation has much less uncertainty due to the large sample size. It is interesting to note that the inter-event correlations are rather uniformly distributed over a wide period range. The finding is in contrast with previous observation regarding the spectral accelerations, whose intra-event and inter-event correlations are similar [18].

Using Eqs. (7) and (8) and the property that the inter-event and intra-event residuals are independent variables, the total correlation for the normalized total residuals can be expressed as a combination of the inter-event and the intra-event correlations:

$$\rho_{\ln CAV,\ln Sa} = c_\eta \rho_{\eta_{\ln CAV},\eta_{\ln Sa}} + c_{\varepsilon'}\rho_{\varepsilon'_{\ln CAV},\varepsilon'_{\ln Sa}} \quad (11)$$

where

$$c_\eta = \frac{\tau_{\ln CAV}\tau_{\ln Sa}}{\sigma_{\ln CAV}\sigma_{\ln Sa}},\ c_{\varepsilon'} = \frac{\sigma'_{\ln CAV}\sigma'_{\ln Sa}}{\sigma_{\ln CAV}\sigma_{\ln Sa}} \quad (12)$$

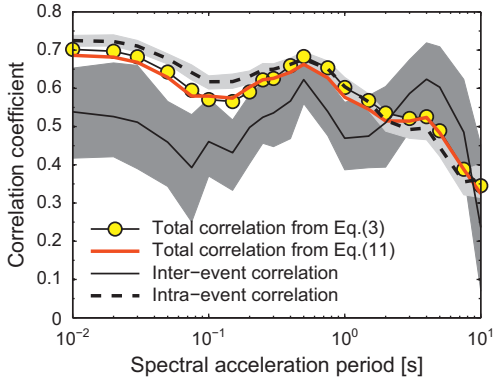

Fig. 9. Comparison of inter-/intra-event correlations and total correlations.

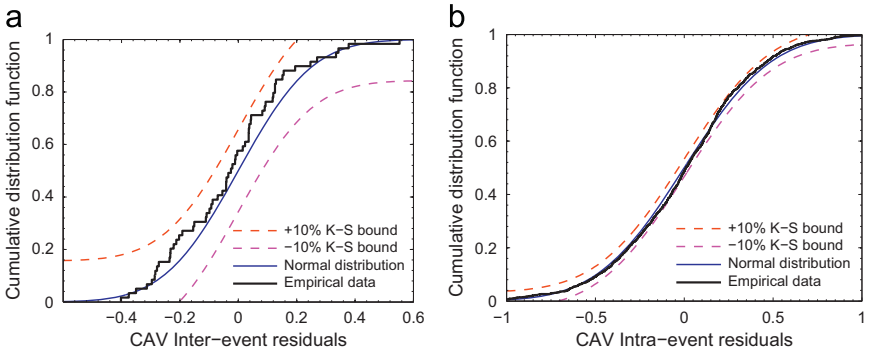

Fig. 8. Cumulative distribution functions of (a) inter-event residuals, and (b) intra-event residuals associated with CAV.

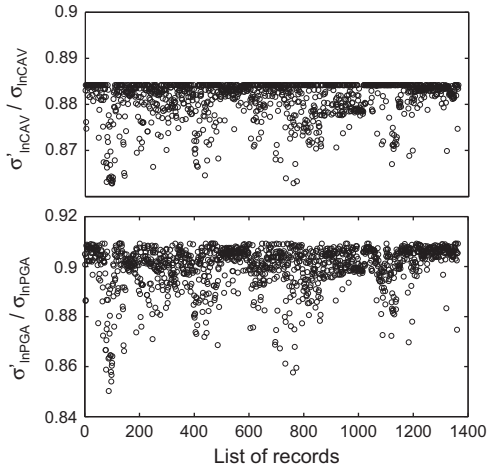

Fig. 10. Ratios of the standard deviations of intra-event residuals over the standard deviations of total residuals

Eq. (11) can be regarded as an alternative approach to obtain the CAV–Sa correlation, which is evaluated by Eq. (3) previously. Fig. 10 examines the variation of $\sigma'_{\ln CAV}/\sigma_{\ln CAV}$ and $\sigma'_{\ln Sa}/\sigma_{\ln Sa}$ terms for all records used in the CB2008 and CB2010 models. It is shown that the record-to-record variation of these terms is rather small, so their average values can be used to evaluate Eq. (11). Therefore, c_η and $c_{\varepsilon'}$ terms in Eq. (12) can be well approximated by constant numbers. Moreover, since the average values of $\sigma'_{\ln CAV}/\sigma_{\ln CAV}$ and $\sigma'_{\ln Sa}/\sigma_{\ln Sa}$ are close to each other (cf. Fig. 10, 0.90 vs. 0.88), we can approximate $c_\eta + c_{\varepsilon'} \approx (\tau_{\ln CAV}/\sigma_{\ln CAV})^2 + (\sigma'_{\ln CAV}/\sigma_{\ln CAV})^2 = 1$. The relationship can be used to further simplify Eq. (11). For example, the following approximation holds for $T=0.01$ s:

$$\rho_{\ln CAV,\ln Sa} \approx 0.2\rho_{\eta_{\ln CAV},\eta_{\ln Sa}} + 0.8\rho_{\varepsilon'_{\ln CAV},\varepsilon'_{\ln Sa}} \tag{13}$$

The above equation indicates that the total correlation is primarily (80%) contributed by the intra-event correlation, so they have similar estimated values and associated uncertainties. In fact, as is shown in Fig. 9, the total correlation obtained from combination of the inter- and intra-event correlations via Eq. (11) agrees well with the previous results obtained by using Eq. (3).

5. Discussions

In this study, empirical correlations for logarithmic CAV and logarithmic spectral accelerations are developed using the NGA strong motion database and recently developed ground-motion prediction models. It is observed that the CAV–Sa correlations at long periods are dependent on rupture distance, probably due to modification of the frequency content and duration of the ground motion time histories through the travel path. Furthermore, the presence of strong velocity pulses in the near-fault ground motions also leads to an increase in the correlation over moderate-to-long period ranges. On the other hand, the correlations are not particularly sensitive to the earthquake magnitude, orientation of the ground-motion recordings, selection of ground-motion prediction models and local site conditions.

A piecewise fitting function is proposed to quantify the empirical correlations for various cases. The obtained correlation enables consideration of the joint distribution of CAV and Sa in ground motion selection process following the proposal in [32]. Given the spectral acceleration value at a particular period T_0, the conditional mean of ln*CAV* can be evaluated from the following equation [33]:

$$\mu_{\ln CAV|\ln Sa(T_0)} = \mu_{\ln CAV} + \rho_{\ln CAV,\ln Sa(T_0)}\sigma_{\ln CAV}\varepsilon_{\ln Sa(T_0)} \tag{14}$$

where $\mu_{\ln CAV|\ln S_a(T_0)}$ is the conditional mean of lnCAV conditioned on a given $\ln S_a(T_0)$ value; $\mu_{\ln CAV}$ and $\sigma_{\ln CAV}$ are the median and the standard deviation of ln CAV respectively predicted by ground-motion models; $\varepsilon_{\ln S_a(T_0)}$ is the specified number of standard deviations (epsilon) of the spectral acceleration at the period T_0; $\rho_{\ln CAV,\ln S_a(T_0)}$ is the CAV–Sa correlation at period T_0. Eq. (14) shows that if an observed spectral acceleration at period T_0 is larger than its predicted value, i.e. $\varepsilon_{\ln S_a(T_0)} > 0$, the observed CAV of that record is also likely to be larger than its predicted value. The conditional standard deviation of lnCAV conditioned on a given $\ln S_a(T_0)$ value, denoted as $\sigma_{\ln CAV|\ln S_a(T_0)}$, can be written as:

$$\sigma_{\ln CAV|\ln S_a(T_0)} = \sigma_{\ln CAV}\sqrt{1-\rho^2_{\ln CAV,\ln S_a(T_0)}} \tag{15}$$

Compared with the unconditional distribution ($\sigma_{\ln CAV} = 0.42$ given by [12]), the variability of the conditional distribution of CAV reduces. For example, if PGA of a record is known and $\rho_{\ln CAV,\ln PGA} = 0.69$ based on this study, the standard deviation of the conditional CAV is reduced by almost 30% as $\sigma_{\ln CAV|PGA} = 0.30$. In fact, the above is an extension of the conditional spectrum concept proposed in [15]. Furthermore, the conditional CAV–Sa correlation can be determined using the unconditional correlation via the following equation:

$$\rho_{\ln CAV,\ln S_a(T)|\ln S_a(T_0)} = \frac{\rho_{\ln CAV,\ln S_a(T)} - \rho_{\ln CAV,\ln S_a(T_0)}\rho_{\ln S_a(T),\ln S_a(T_0)}}{\sqrt{1-\rho^2_{\ln CAV,\ln S_a(T_0)}}\sqrt{1-\rho^2_{\ln S_a(T),\ln S_a(T_0)}}} \tag{16}$$

Eqs. (14)–(16) fully describe the statistical distribution of CAV conditioned on a spectral acceleration ordinate at a specified period. The empirical correlation provides a viable means to select and modify ground motion time histories accordingly to capture the joint distribution of both CAV and response spectra over a period range of interest. Such a ground motion selection scheme as an extension to the authors' previous work [34] will be a subject of future studies. It is expected that capturing the joint distribution of CAV and response spectra can provide more realistic characterization of the ground motions in seismic hazard analysis.

Acknowledgments

The authors thank Dr. Brian Chiou for providing useful materials to facilitate this study. The authors also thank reviewers to provide constructive comments that improve the quality of the paper. Financial supports from the Hong Kong Research Grants Council Grant 620311 and the Li Foundation Heritage Prize Award are greatly appreciated.

References

[1] Housner GW. Spectrum intensities of strong-motion earthquakes. In: Proceedings of the Symposium on earthquakes and blast effects on structures, Los Angeles, California, 1952.
[2] Von Thun J, Roehm L, Scott G, Wilson J. Earthquake ground motions for design and analysis of dams. Earthquake Engineering and Soil Dynamics II—Recent Advances in Ground-Motion Evaluation Geotechnical Special Publication 1988;20:463–81.
[3] Bradley BA. Empirical equations for the prediction of displacement spectrum intensity and its correlation with other intensity measures. Soil Dynamics and Earthquake Engineering 2011;31(8):1182–91.
[4] Arias A. A measure of earthquake intensity. In: Hansen RJ, editor. Seismic design for nuclear power plants. Cambridge, MA: MIT Press; 1970. p. 438–83.
[5] Electrical Power Research Institute (EPRI). A criterion for determining exceedance of the operating basis earthquake, Report No. EPRI NP-5930, Palo Alto, California, 1988.
[6] Electrical Power Research Institute (EPRI). Standardization of the cumulative absolute velocity, Report No. EPRI TR-100082-T2, Palo Alto, California, 1991.

[7] Dashti S, Bray JD, Pestana JM, Riemer M, Wilson D. Centrifuge testing to evaluate and mitigate liquefaction-induced building settlement mechanisms. Journal of Geotechnical and Geoenvironmental Engineering 2010;136(7): 918–29.
[8] Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. Earthquake Spectra 1998;14(3):469–500.
[9] Kramer SL, Mitchell RA. Ground motion intensity measures for liquefaction hazard evaluation. Earthquake Spectra 2006;22(2):413–38.
[10] Liyanapathirana DS, Poulos HG. Assessment of soil liquefaction incorporating earthquake characteristics. International Journal of Soil Dynamics and Earthquake Engineering 2004;24:867–75.
[11] US Nuclear Regulatory Commission. Pre-earthquake planning and immediate nuclear power plant operator postearthquake actions, Regulatory Guide 1.166. Washington, DC, 1997.
[12] Campbell KW, Bozorgnia Y. A ground motion prediction equation for the horizontal component of cumulative absolute velocity (CAV) based on the PEER-NGA strong motion database. Earthquake Spectra 2010;26(3):635–50.
[13] Campbell KW, Bozorgnia Y. NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. Earthquake Spectra 2008;24:139–71.
[14] Campbell KW, Bozorgnia Y. Campbell-Bozorgnia NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters, PEER Report No. 2007/02, Pacific Earthquake Engineering Research Center. Berkeley: University of California; 2007.
[15] Baker JW, Cornell CA. A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. Earthquake Engineering and Structural Dynamics 2005;34:1193–217.
[16] Kutner MH, Nachtsheim CJ, Neter J. Applied linear regression models. 4th EditionMcGraw-Hill; 2004.
[17] Eferon B, Tibshirani RJ. An introduction to the bootstrap. Chapman & Hall; 1993.
[18] Abrahamson N, Atkinson G, Boore D, Campbell K, Chiou B, Idriss IM, et al. Comparisons of the NGA ground–motion relations. Earthquake Spectra 2008;24(1):45–66.
[19] Boore DM, Watson-Lamprey J, Abrahamson NA. Orientation-independent measures of ground motion. Bulletin of the Seismological Society of America 2006:1502–11.
[20] Somerville PG, Smith NF, Graves RW, Abrahamson NA. Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity. Seismological Research Letters 1997;68(1):199–227.
[21] Baker JW, Jayaram N. Correlation of spectral acceleration values from NGA ground motion models. Earthquake Spectra 2008;24(1):299–317.
[22] Koketsu K, Miyake H. A seismological overview of long-period ground motion. Journal of Seismology 2008;12:133–43.
[23] Rathje EM, Russell S, Bray JD. Empirical relationships for frequency content parameters of earthquake ground motions. Earthquake Spectra 2004;20(1): 119–44.
[24] Kempton JJ, Stewart JP. Prediction equations for significant duration of earthquake ground motions considering site and near-source effects. Earthquake Spectra 2006;22(4):985–1013.
[25] Luco N, Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthquake Spectra 2007;23(2):357–92.
[26] Spudich P, Chiou BSJ. Directivity in NGA earthquake ground motions: analysis using isochrone theory. Earthquake Spectra 2008;24(1):279–98.
[27] Baker JW. Quantitative classification of near-fault ground motions using wavelet analysis. Bulletin of the Seismological Society of America 2007;97(5): 1486–501.
[28] Baker JW. Identification of near-fault velocity pulses and prediction of resulting response spectra. In: Proceedings of the Geotechnical earthquake engineering and structural dynamics IV, Sacramento, May 18–22, 2008.
[29] Shahi SK, Baker JW. An empirically calibrated framework for including the effects of near-fault directivity in probabilistic seismic hazard analysis. Bulletin of the Seismological Society of America 2011;101(2):742–55.
[30] Abrahamson NA, Youngs RR. A stable algorithm for regression analysis using the random effects model. Bulletin of the Seismological Society of America 1992;82(1):505–10.
[31] Jayaram N, Baker JW. Statistical tests of the joint distribution of spectral acceleration values. Bulletin of the Seismological Society of America 2008;98(5): 2231–43.
[32] Bradley BA. A generalized conditional intensity measure approach and holistic ground motion selection. Earthquake Engineering and Structural Dynamics 2010;39(12):1321–42.
[33] Ang AHS, Tang WH. Probability concepts in engineering: emphasis on applications in civil and environmental engineering. John Wiley & Sons; 2007.
[34] Wang G. A ground motion selection and modification method capturing response spectrum characteristics and variability of scenario earthquakes. Soil Dynamics and Earthquake Engineering 2011;31(4):611–25.