# Performance of On-Site Earthquake Early Warning System Using Strong-Motion Records from Recent Earthquakes

Duruo Huang[1]; Gang Wang, M.ASCE[2]; and Feng Jin[3]

**Abstract:** The method of on-site earthquake early warning was proposed in the past decade to enhance seismic preparedness and safety measures for areas close to the earthquake epicenter. The method leverages critical information contained in the initial 3 s of a P wave that arrives early to predict possibly intense ground shaking at a site. In this study, two types of empirical relationships are developed: (1) $Pd_3$-PGV relationship for predicting the horizontal peak ground velocity (PGV) based on the peak displacement in the first 3 s of the vertical motion ($Pd_3$); and (2) $\tau_c - M_w$ relationship for estimating the moment magnitude ($Mw$) of an earthquake based on a ground-motion period parameter ($\tau_c$). Performance of these empirical relationships in on-site earthquake early warning framework is examined using global earthquake records from the most updated Next Generation Attenuation (NGA-West2) strong-motion database. Region-specific empirical predictive equations for California and Japan are proposed based on the state of the art mixed-effect regression to separate ground-motion variability between different earthquakes and that within an earthquake. Applicability and inapplicability of the empirical models in the near-fault condition are investigated using ground motions with strong velocity pulses. The overarching goal of study is to leverage the comprehensive database and state of the art regression techniques to facilitate understanding and engineering practice of on-site earthquake early warning in a variety of regions and in the near-fault condition. **DOI: 10.1061/(ASCE)NH.1527-6996.0000318.** 

**Author keywords:** On-site early warning; $Pd_3$–PGV and $\tau_c - M_w$ relationships; NGA-West2 database; Near-fault ground motion.

## Introduction

Conceptually, an earthquake early warning system can be categorized as a regional network or an on-site system by seismic station setup (Allen et al. 2009). A regional early warning system issues earthquake warnings before destructive seismic waves arrive at remote sites using seismic station network that covers a large area. It has been regarded as an accurate approach given a longer time window, but it can be highly demanding for the rapid reporting network to process and transmit data faster than seismic wave propagation in order to ensure adequate front time for the emergency response. Unavoidably, a blind zone always exists around the ruptured fault where no warning can be provided because several seconds are needed to observe and respond to a strong ground shaking before a warning is issued (Allen et al. 2009; Kamigaichi et al. 2009; Wu and Kanamori 2005a). Also, the performance of a regional earthquake early warning system degrades if a region is incapable of installing a densely populated seismic station network.

The idea of on-site earthquake early warning was initially proposed by Wu and Kanamori (2005a, b) to enhance seismic preparedness and safety measures of the blind zone close to the earthquake epicenter. The method leverages critical initial portion of a P wave that arrives early at a site to provide warning for the following intense ground shaking. The on-site early warning system has been regarded as economically efficient in a way that it is applicable to regions where only one or a few seismic observation stations are installed. Intuitively, an efficient and effective predictor in the empirical model has significant impact on the performance of such kind of early warning system. Previous researchers recommended using the peak displacement measured in the initial 3 s of a P wave ($Pd_3$) as a predictor of the peak ground velocity (PGV) of the incoming S wave (Wu and Kanamori 2005a, b; Wu et al. 2007). Several sets of empirical $Pd_3$-PGV relations have been proposed based on least-squares linear regression analyses of a couple of earthquakes in Taiwan, Japan, and California (Chen et al. 2012; Wu et al. 2007). Note that the P wave–based earthquake early warning system has been implemented at more than 580 seismic recording stations in Taiwan (Wu and Kanamori 2005a, b, Wu et al. 2016). Each early waring device is equipped with a three-component accelerometer at a sampling rate of 100 points per second. With the use of software algorithm embedded in the device, the signal recorded at each station can be processed for detecting the arrival of P wave and computing the precursor $Pd_3$. Once the captured signal exceeds a predefined threshold, the early warning device will send out alert. Most recently, the capability of on-site early warning approach has been demonstrated during the 2016 $M_w$ 6.4 Meinong earthquake in Taiwan. The on-site device provides several-second lead time before the arrival of peak shaking (Wu et al. 2016). Additionally, it is important to determine the magnitude of an earthquake. For this purpose, a parameter $\tau_c$ is defined to represent the average period of the initial portion of the P wave, which is empirically correlated to the earthquake magnitude (Huang et al. 2015; Wu and Kanamori 2008).

[1]Assistant Professor, Dept. of Hydraulic Engineering, Tsinghua Univ., Beijing 100084, China.

[2]Associate Professor, Dept. of Civil and Environmental Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Hong Kong, China (corresponding author). Email: gwang@ust.hk

[3]Professor, Dept. of Hydraulic Engineering, Tsinghua Univ., Beijing 100084, China.

Note. This manuscript was submitted on October 3, 2017; approved on July 25, 2018; published online on December 8, 2018. Discussion period open until May 8, 2019; separate discussions must be submitted for individual papers. This paper is part of the ***Natural Hazards Review***, © ASCE, ISSN 1527-6988.

Note that few attempts have been made to comprehensively examine empirical relations in an on-site early warning system using the renowned global database and the state of the art mixed-effect regression technique. In this study, we evaluate performance of the empirical $Pd_3$-PGV and $\tau_c - M_w$ relations by using a total of 4,004 selected ground-motion records from the most updated Next-Generation Attenuation database (NGA-West2) strong-motion databases. Note that these two empirical models serve different purposes and provide complementary information to the on-site early warning system. Specifically, the $Pd_3$—PGV model predicts the peak ground velocity throughout entire horizontal ground-motion time histories that strongly correlates to structural damage, while the $\tau_c - M_w$ relation predicts the earthquake magnitude ($M_w$) to provide general information for earthquake source characteristics.

The NGA-West2 database adopted in the current study is an update of the original the Pacific Earthquake Engineering Research Center NGA (PEER-NGA) database with addition of major earthquake recordings in recent years in seismic active regions including California, Japan, Taiwan, and other regions (Italy, Mexico, Iran, Greece, Turkey, Alaska, etc.) (Ancheta et al. 2014). A region-specific predictive model is proposed in this study to account for the influence of ground motions with regional characteristics. Ground motions close to a ruptured fault resulting from forward-directivity are adopted to examine the viability of empirical $Pd_3$-PGV relation in a near-source environment. The feasibility of using multivariant precursors in an on-site early warning framework will be also investigated. The overarching goal of the study is to leverage the comprehensive database to facilitate understanding the features, applicability, and inapplicability of empirical on-site prediction in different regions and in a near-fault condition.

## Strong-Motion Database

In this study, the most updated NGA-West2 database (Ancheta et al. 2014) from the Pacific Earthquake Engineering Research Center is used for developing empirical prediction equations in an on-site earthquake early warning system. The NGA-West2 database is an expansion and update of the existing PEER-NGA database (Chiou et al. 2008), and both of them have been extensively used in the past decade in developing ground-motion prediction equations (GMPEs) for a significant number of intensity measures such as peak ground acceleration (PGA), PGV, peak ground displacement (PGD), Arias intensity, cumulative absolute velocity, spectral acceleration (Wang 2011; Du and Wang 2013a; Abrahamson et al. 2014; Boore et al. 2014; Campbell and Bozorgnia 2014; Chiou and Youngs 2014; Idriss 2014), and their correlations (Wang and Du 2012, 2013; Du and Wang 2013b).

Table 1 summarizes strong-motion data used in this study. The compiled database contains 4,004 three-component recordings from 64 earthquakes, including recent California earthquakes after 2004, namely, the 2005 magnitude 5.2 Anza earthquake, the 2007 magnitude 5.4 Alum Rock earthquake, and the 2008 magnitude 5.4 Chino Hills earthquake. Five Japan earthquakes include the 1995 $M_w$ 6.9 Kobe earthquake, the 2007 $M_w$ 6.8 Chuetsu-Oki earthquake, the 2000 $M_w$ 6.6 Tottori earthquake, the 2004 $M_w$ 6.6 Niigata earthquake and the 2008 $M_w$ 6.9 Iwate earthquake. Fig. 1 illustrates the moment magnitude and hypocentral distance distribution of all ground-motion records adopted in this study. Note that all the records have been postprocessed by applying acausual Butterworth filter following the PEER processing methodology to remove long-period drift and high-frequency noises (Ancheta et al. 2014). Seismic information for these events and geological conditions for recording stations are provided by

**Table 1.** Earthquake events used in this study

| Event | Date (mm/dd/yyyy) | Moment magnitude | Region | Number of record sets[a] |
|---|---|---|---|---|
| California earthquakes[b] | Before 2004 | 4.27–7.28 | California | 963 |
| Anza | 06/12/2005 | 5.2 | California | 110 |
| Alum Rock | 10/30/2007 | 5.4 | California | 159 |
| Chino Hills | 07/29/2008 | 5.4 | California | 336 |
| Kobe | 01/16/1995 | 6.9 | Japan | 12 |
| Chuetsu-Oki | 07/16/2007 | 6.8 | Japan | 555 |
| Iwate | 06/13/2008 | 6.9 | Japan | 317 |
| Niigata | 10/23/2004 | 6.6 | Japan | 528 |
| Tottori | 10/06/2000 | 6.6 | Japan | 410 |
| Chi-Chi (main shock) | 09/21/1999 | 7.6 | Taiwan | 421 |
| Others[c] | — | — | Global | 193 |

[a]Each record set contains three components (two horizontal and one vertical) of acceleration time histories.
[b]California earthquakes and number of records in parentheses are as follows: San Fernando (29), Imperial Valley (33), Coalinga (44), Morgan Hill (23), N. Palm Springs (27), Whittier Narrows (105), Landers (67), Northridge (148), Hector Mine (82), Anza-02 (71), Gilroy (33), Big Bear City (33), and others in the PEER-NGA database.
[c]Other regions include Italy, Mexico, Iran, Greece, Turkey, Alaska, etc. as shown in the PEER-NGA database.

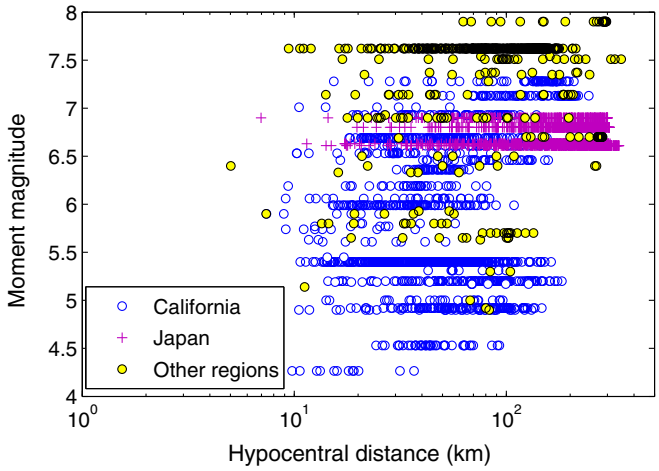


**Fig. 1.** Magnitude-distance distribution of strong-motion records used in this study.

Kaklamanos and Baise (2011). In particular, the effectiveness and efficiency of earthquake early warning systems highly rely on accurate detection of the P wave arrival at a site. In this study, the arrival time of the vertical seismic trace is determined using an automated algorithm by Allen (1978), then each record is inspected visually and adjusted manually to accurately differentiate the real P wave arrival from the recorded ambient noise. Fig. 2 shows an example of $Pd_3$ and $\tau_c$, peak ground acceleration, peak ground velocity, and peak ground displacement in a ground-motion time history recorded at the Parachute Test Site station during the 1987 Superstition Hills earthquake. Note that $Pd_3$ denotes the maximum absolute displacement in the first 3 s of the P wave; it is computed by double integration of the vertical acceleration time history over 3 s after the detected P wave arrival. The PGV is the maximum value of PGVs obtained from two horizontal time histories, which represents the most damaging effect of the horizontal shear wave. The notion $\tau_c$ will be explained in greater detail in a later section.

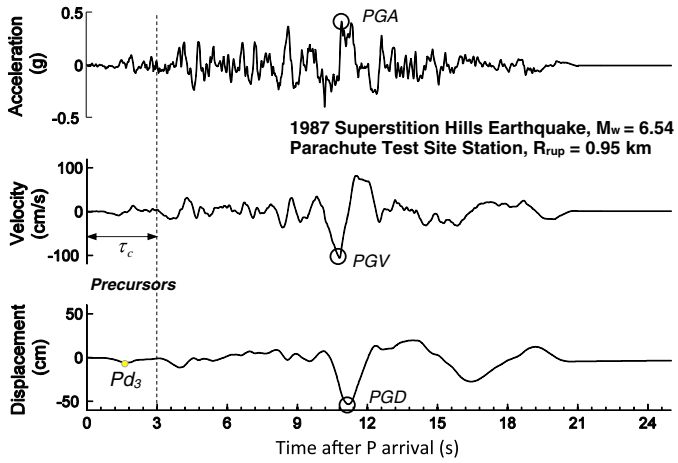


**Fig. 2.** Example of precursors (e.g., $\tau_c$, $Pd_3$), peak ground acceleration, peak ground velocity, and peak ground displacement in a ground-motion time history recorded at the Parachute Test Site station during the 1987 Superstition Hills earthquake.

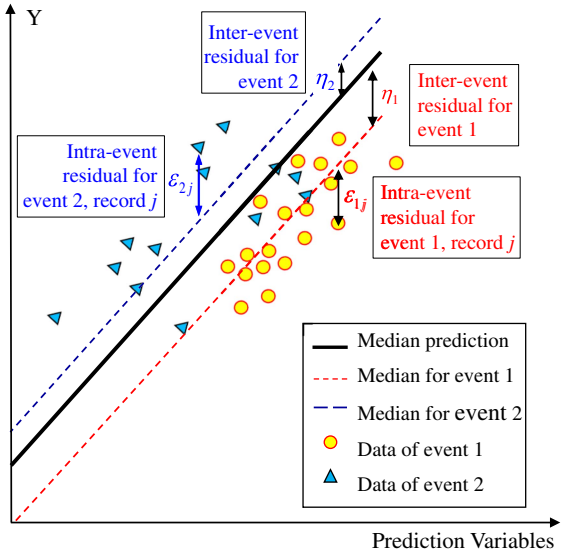


**Fig. 3.** Illustration of ground-motion variability: interevent and intraevent residuals.

## Empirical $Pd_3$-PGV Correlation Model

### Mixed-Effect Regression Analysis Based on the NGA-West2 Database

With the use of mixed-effect regression technique, the $Pd_3$-PGV prediction equation can be written as follows:

$$\log_{10}\mathrm{PGV}_{ij} = \overline{\log_{10}(\mathrm{PGV})_{ij}} + \eta_i + \varepsilon_{ij} \tag{1}$$

where $\log_{10}\mathrm{PGV}_{ij}$ and $\overline{\log_{10}\mathrm{PGV}_{ij}}$ denote logarithm of the measured and the median prediction of the maximum value of PGV obtained from two as-recorded *horizontal* time histories; $i$ and $j$ denote the $j$th recording in the $i$th event, respectively; and $\eta_i$ and $\varepsilon_{ij}$ represent the interevent residual and intraevent residual, which are assumed to be normally distributed with a mean of zero and standard deviations of $\tau$ and $\sigma$, respectively (Abrahamson and Youngs 1992). The predicted PGV will be empirically determined using $Pd_3$, the maximum displacement amplitudes of the initial 3 s in a P wave, via Eq. (2)

$$\overline{\log_{10}(\mathrm{PGV})_{ij}} = a + b\log_{10}(Pd_3)_{ij} \tag{2}$$

where $a$ and $b$ are the regression coefficients. Fig. 3 illustrates the mixed-effect regression model. The important feature of mixed-effect regression is the total residuals associated with the prediction model are separated into two mutually independent components: the interevent residuals and intraevent residuals. The interevent residuals measure variability between different earthquake events, and the intraevent residuals measure variability within an earthquake. Therefore, the mixed-effect regression will not be dominated by earthquakes with a large amount of records. This regression approach has been actively used in the past two decades to develop a significant number of GMPEs in contemporary earthquake engineering practice (e.g., Abrahamson et al. 2014; Boore et al. 2014; Campbell and Bozorgnia 2014; Chiou and Youngs 2014; Du and Wang 2013a, 2016; Idriss 2014). Therefore, the standard deviation of the total residuals (i.e., $\eta_i + \varepsilon_{ij}$) is determined using $\sigma_{\text{total}} = \sqrt{\tau^2 + \sigma^2}$, where $\tau$ and $\sigma$ are the interevent and intraevent standard deviations, respectively. Note that the two-stage mixed-effect regression requires maximum likelihood estimate of the model parameters via Eq. (3) (Abrahamson and Youngs 1992)

$$\begin{aligned}\ln L = &-\frac{1}{2}N\ln(2\pi) - \frac{1}{2}(N-M)\ln(\sigma^2) - \frac{1}{2}\sum_{i=1}^{M}\ln(\sigma^2 + n_i\cdot\tau^2)\\ &-\frac{1}{2\sigma^2}\sum_{i=1}^{M}\sum_{j=1}^{n_i}\left(y_{ij} - \frac{1}{n_i}\sum_{j=1}^{n_i}y_{ij}\right)^2\\ &-\frac{1}{2}\sum_{i=1}^{M}\frac{n_i\left(\frac{1}{n_i}\sum_{j=1}^{n_i}y_{ij} - \frac{1}{n_i}\sum_{j=1}^{n_i}\mu_{ij}\right)^2}{\sigma^2 + n_i\cdot\tau^2}\end{aligned} \tag{3}$$

where $M$ = number of events; $N$ = number of data points; $n_i$ = number of recordings for the $i$th event; $y_{ij}$ = observed $j$th recording in the $i$th event; and $\mu_{ij}$ represents the corresponding predicted value. In this study, these notations represent $\mathrm{PGV}_{ij}$ and $\overline{\mathrm{PGV}_{ij}}$, respectively. Using the mixed-effect regression analyses, the model coefficients ($a$ and $b$) can be obtained as follows:

$$\overline{\log_{10}(\mathrm{PGV})} = 1.189 + 0.561\times\log_{10}(Pd_3) \tag{4}$$

where $\overline{\mathrm{PGV}}$ is in cm/s and $Pd_3$ is in cm. The standard deviations of interevent and intraevent residuals, $\tau$ and $\sigma$, are determined as 0.16 and 0.30 from regression, respectively. The total standard deviation, $\sigma_{\text{total}}$, is computed as 0.34. Fig. 4(a) shows the $Pd_3$ and PGV for the global dataset, along with the regression line bounded by $\pm 1$ standard deviation. The interevent/intraevent residuals against earthquake magnitude, and intraevent residuals against hypocentral distance are presented in Fig. 4(b). It is shown that the mixed-effect regression is unbiased in terms of both interevent and intraevent residuals.

The log-log space is useful to handle data with a large range, e.g., the PGV ranges from 0.1 to 100 cm/s, covering four orders of magnitude, while $Pd_3$ covers four to five orders of magnitude. Also, PGV distribution is found to be lognormally distributed. Therefore, log-log is used for $Pd_3$-PGV relationship. During the course of study, many other functional forms in log-log space were attempted, such as hyperbolic, power law, quadratic, and trilinear forms. Note that these regressions are quite similar within the major portion of data. The major difference appears only when $Pd_3$ is very large or very small, where data is too sparse to strongly favor one functional form over another. Therefore, we proposed to use the simplest functional form Eq. (2) that is linear in the log-log space.

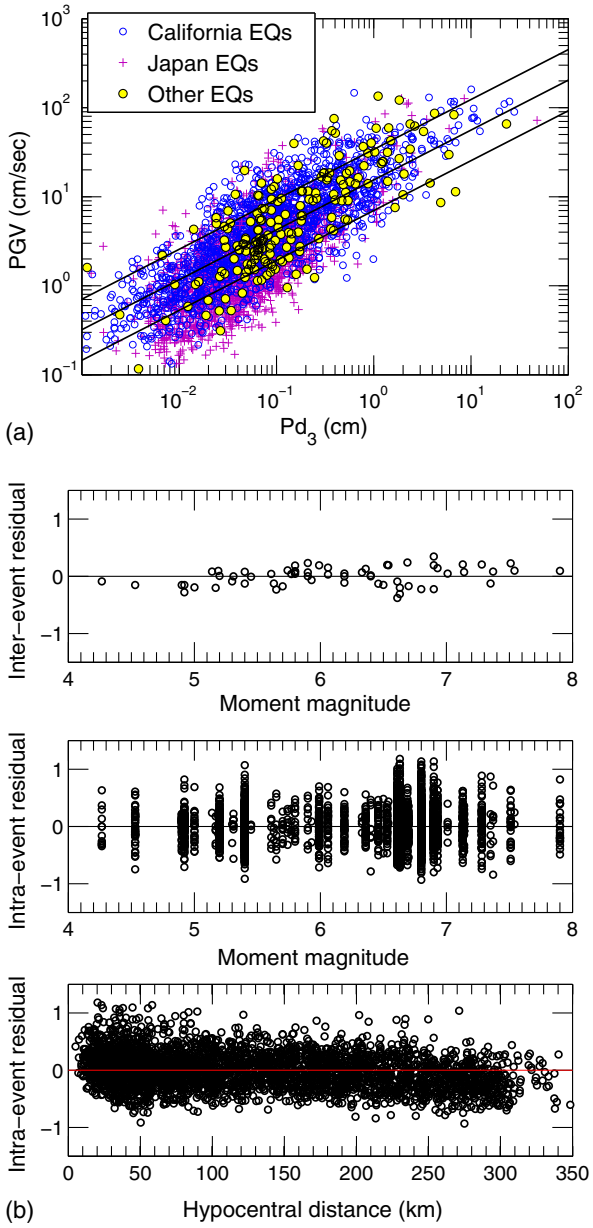


**Fig. 4.** (a) $Pd_3$-PGV relation for the earthquake recordings from the NGA-West2 database (black lines are median and $\pm 1$ standard deviation curves); and (b) distribution of interevent and intraevent residuals over moment magnitude and hypocenter distance.

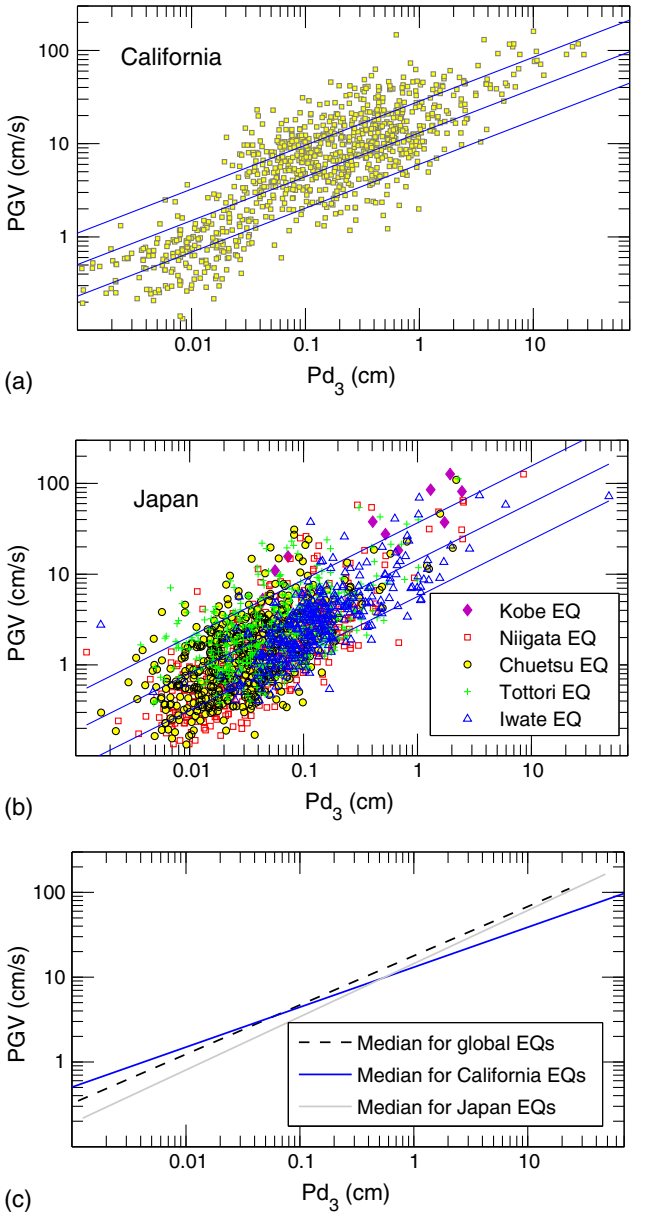


**Fig. 5.** Region-specific $Pd_3$-PGV predictive relations obtained using (a) California; (b) Japan; and (c) global strong-motion data.

### *Region-Specific $Pd_3$-PGV Correlation Model*

Previous studies indicate significant regional difference in the energy content, time–frequency characteristics, and nonstationarity of ground-motion time histories based on geostatistical analyses of earthquake data in California, Mexico, Japan, and Taiwan (Huang and Wang 2015a, b, 2017). Such observation calls for development of region-specific $Pd_3$-PGV predictive models to account for the influence of regional site conditions and regional characteristics of strong motions. One of the major challenges in developing such a model is the availability of densely populated ground-motion observation data in different regions.

In this study, a total of 4,004 ground-motion records have been compiled from recent earthquakes occurred in California, Japan, Taiwan and other regions, with updated geological information in each region. The $Pd_3$-PGV relations for different regions have been developed by leveraging sufficient ground-motion data recorded in difference regions. The functional form in Eq. (1) is adopted.

Fig. 5 demonstrates median prediction models obtained using California, Japan, and global strong-motion data, while fitting parameters $a$ and $b$, and standard deviations of interevent and intraevent residuals are summarized in Table 2. In general, the slope for the California model is more gentle than that of the Japan model,

**Table 2.** Regression coefficients in $Pd_3$-PGV relation Eq. (2) in the region-specific earthquake early warning models

| Model | $a$ | $b$ | $\tau$ | $\sigma$ | $\sigma_{\text{total}}$ |
|---|---|---|---|---|---|
| Global | 1.189 | 0.561 | 0.16 | 0.30 | 0.34 |
| California | 1.117 | 0.471 | 0.20 | 0.27 | 0.34 |
| Japan | 1.160 | 0.627 | 0.24 | 0.33 | 0.41 |
| Other regions | 1.252 | 0.580 | 0.13 | 0.30 | 0.33 |

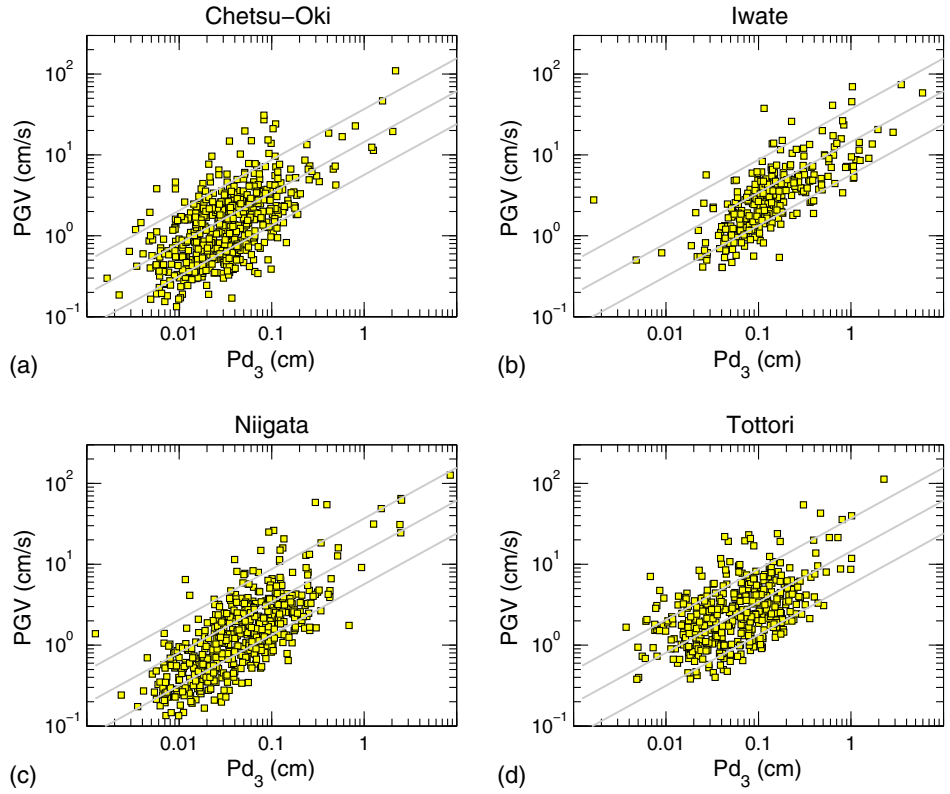


**Fig. 6.** Region-specific $Pd_3$–PGV predictive relation for the strong-motion records in (a) the 2007 $M$ 6.8 Chuetsu-Oki earthquake, (b) the 2008 $M$ 6.9 Iwate earthquake, (c) 2004 $M$ 6.6 Niigata earthquake, and (d) the 2000 $M$ 6.6 Tottori earthquake in Japan; gray lines are median and $\pm 1$ standard deviation curves.

indicating that a relatively small PGV value will be predicted using the California model when $Pd_3$ is larger than 0.2 cm. For example, given a $Pd_3$ of 5 cm, a PGV of 40 cm/s is estimated using the Japan empirical relation, around 50% larger than that estimated using the California one. The regional difference may be explained by the fact that the shallowest part of the crust in the western United States typically has shear wave velocities in the range of 760–1,500 m/s, corresponding to rock of the National Earthquake Hazards Reduction Program (NEHRP) site Class B (BSSC 2003). The softer rocks tend to exhibit higher damping, which may reduce peak amplitude of ground motions. Fig. 6 shows the region-specific $Pd_3$-$PGV$ relation for Japan and strong-motion records in the four recent earthquakes, namely the 2007 magnitude 6.8 Chuetsu-Oki earthquake, the 2008 magnitude 6.9 Iwate earthquake, the 2004 magnitude 6.6 Niigata earthquake, and the 2000 magnitude 6.6 Tottori earthquake. The fitting parameters $a$ and $b$ in Eq. (2) are provided in Table 2. It can be observed that the four recent earthquakes for a total of 1,810 sets of records are generally contained within $\pm 1$ standard deviation of median prediction for the entire Japan dataset. If emphasis is placed on intense ground shaking with PGVs greater than 5 cm/s, the predictability is more satisfactory, as shown in Fig. 6. In addition, Fig. 7 shows the $Pd_3$-PGV median prediction and $\pm 1$ standard deviation for earthquakes in California.

### *Observation from the 1999 Chi-Chi Earthquake in Taiwan*

Note that the empirical $Pd_3$-PGV relation is incapable of accurately predicting intense ground shaking for some catastrophic events, such as the 1999 magnitude 7.6 Chi-Chi earthquake in Taiwan. Fig. 8(a) presents the $Pd_3$-PGV data obtained from the Chi-Chi earthquake (marked as circles) against the global data. It can be observed that Chi-Chi earthquake data are outliers with PGV values significantly larger than those obtained from other earthquakes given a same displacement precursor ($Pd_3$). This phenomenon might be explained by the long rise time in the fault rupture process during the Chi-Chi earthquake (Zeng and Chen 2001). The rise time refers to the time required for completion of a slip at a point on the fault plane. Previous local researchers reported that the Chi-Chi earthquake had a long rise time associated with low dynamic stress drop (Zeng and Chen 2001). Fig. 8(b) shows the rise time over $Pd_3$ relation obtained using global earthquakes in the NGA database. The documented overall rise time of the Chi-Chi earthquake is 3.52 s, which is substantially longer than those of other earthquakes. It has also been reported that the rise time over the primary slip area in the Chi-Chi earthquake can reach about 8 s (Chi et al. 2001). Therefore, the first 3 s of the P wave may not be sufficient and adequate to characterize a long ruputure process. Moreover, there are some unique features associated with the fault rupture process in the Chi-Chi earthquake. It has been identified that the Chelungpu fault rupture contains a significant amount of lubrication materials, which reduce friction and causes large slip velocity (Ma et al. 2003). This result may have important implications for improving earthquake early warning if the fault rupture mechanism is similar to the Chelungpu fault.

### *Observation from Near-Fault Pulse-Like Records*

During an earthquake event, fault rupture propagation and earthquake source radiation pattern cause spatial variation of horizontal

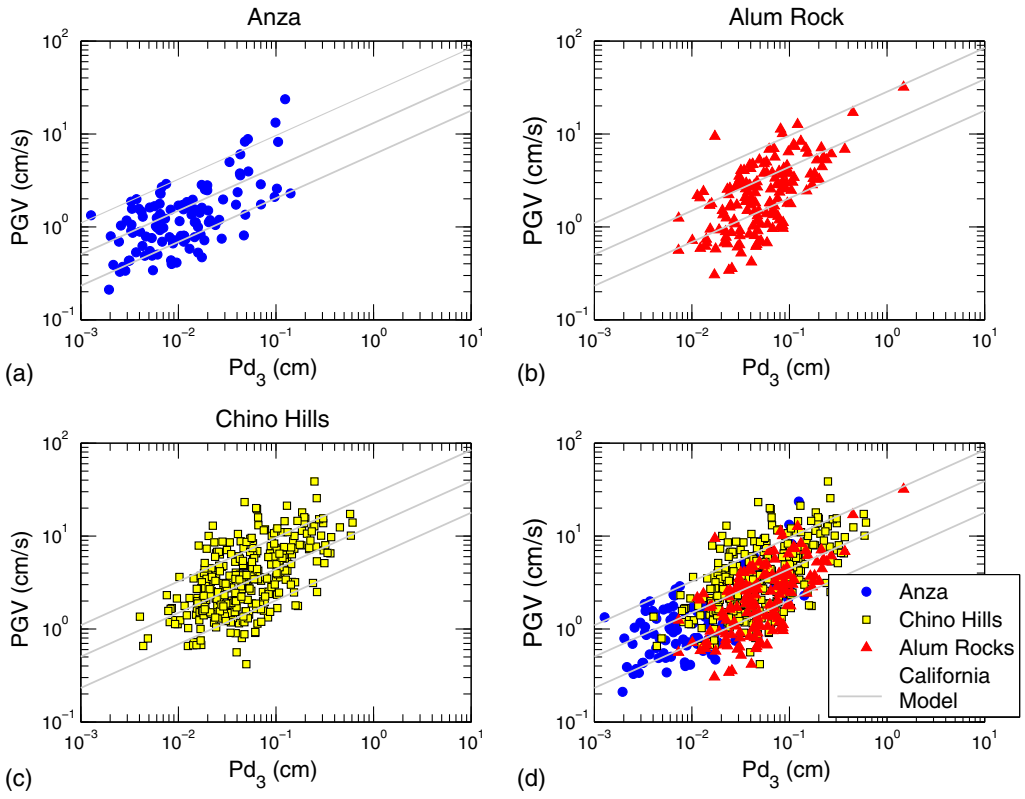


**Fig. 7.** Region-specific $Pd_3$–PGV predictive relation for the strong-motion records in the (a) 2005 $M$ 5.2 Anza earthquake, (b) the 2007 $M$ 5.4 Alum Rock earthquake, (c) the 2008 $M$ 5.4 Chino Hills earthquake, (d) and all three events; gray lines are median and $\pm 1$ standard deviation curves.

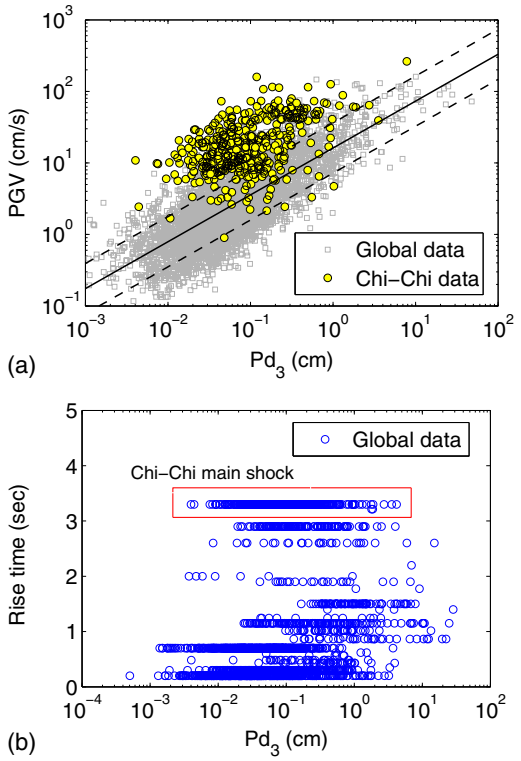


**Fig. 8.** (a) $Pd_3$-PGV relationship for global earthquake recordings from the NGA West2 database and from the 1999 Chi-Chi earthquake; median and $\pm 1$, $\pm 2$ standard deviation curves are plotted; (b) rise time versus $Pd_3$ for global records and for the Chi-Chi records.

ground motions in amplitude and duration around the fault (Huang and Wang 2015b; Du and Wang 2017). There also exists a systematic difference (or polarization) of horizontal motions in the direction perpendicular to the fault strike, termed as the fault-normal (FN) direction, and in the direction parallel to the fault strike, termed as the fault-parallel (FP) direction (Wang et al. 2015). Ground motions within a distance of 30 km from the ruptured fault may be substantially different from other ground motions due to directivity effects (Somerville et al. 1997). These near-fault motions have been identified to have strong velocity pulses (Baker 2007), which can impose severe demands on structures. In the PEER-NGA database, certain ground motion records have been identified as having strong velocity pulses that may be associated with fault rupture directivity effects. Detailed results and documentation can be found in Baker (2007). Besides the pulse-like records identified by Baker (2007), several records with strong FN pulses included in the Design Ground Motion Library (Wang et al. 2015) have also been adopted in this study. In total, the pulse-like records include 63 having pulses in the FN component only, 23 having pulses in the FP component only, and 30 with pulses in both FN and FP components.

In this study, performance of the empirical $Pd_3$-PGV relation is examined using the abovementioned collection of near-fault pulse-like records. Fig. 9 presents the $Pd_3$-PGV relation using a total of 116 pulse-like records and the global data from the NGA-West2 database. It is shown that PGVs of the pulse-like records are generally from 30 to 200 cm/s, which are substantially larger than those of other strong-motion records. Meanwhile, the corresponding precursor $Pd_3$ falls into a wide range from 0.02 to 60 cm, indicating that $Pd_3$ fails to predict potential large ground shaking in a near-fault scenario. The empirical $Pd_3$-PGV relation systematically underestimates PGVs for pulse-like records and seems unconservative

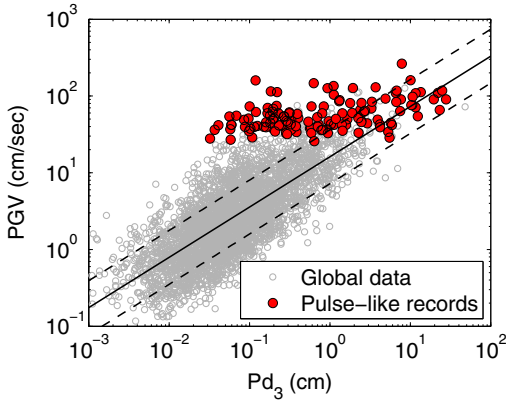


**Fig. 9.** $Pd_3$-PGV relation for near-fault pulse-like records and the global records in the NGA-West2 database; median (solid) and ±1 standard deviation (dashed) curves are plotted.

for the purpose of earthquake early warning. In such a case, the empirical $Pd_3$-PGV relation is more suitable for estimating PGVs less than 40 cm/s. It is also suggested that vector-valued precursors may be used to improve the model capacity by incorporating supplemental information of the faulting mechanism and physical rupture process.

## Empirical $\tau_c$-Magnitude Correlation Model

Apart from the aforementioned displacement precursor (i.e., $Pd_3$), a ground-motion period parameter $\tau_c$ has been extensively used by researchers for quick determination of the earthquake magnitude in an earthquake early warning system (Huang et al. 2015; Wu and Kanamori 2005a, b). It has been identified that earthquake magnitude is closely related to the average period of the initial portion of P wave, as large events yield long-period initial motions (Wu and Kanamori 2005a). A ratio $r$ is first defined by integrating displacement time history $u(t)$ and velocity time history $\dot{u}(t)$ of initial part of a record

$$r = \frac{\int_0^{t_0} \dot{u}^2(t)dt}{\int_0^{t_0} u^2(t)dt} \tag{5}$$

where the integration is over the time interval $(0, t_0)$ of a record. With the use of Parseval's theorem, the ratio $r$ can be defined as

$$r = 4\pi^2 \frac{\int_0^{\infty} f^2|\hat{u}(f)|^2 df}{\int_0^{\infty} |\hat{u}(f)|^2 df} = 4\pi^2 \cdot \overline{f^2} \tag{6}$$

where $\overline{f^2}$ = squared average frequency; and the average-period parameter $\tau_c$ in the initial portion of a record can be defined using the following equation:

$$\tau_c = \frac{1}{\sqrt{\overline{f^2}}} = \frac{2\pi}{\sqrt{r}} = \frac{2\pi}{\sqrt{\left[\int_0^{t_0} \dot{u}^2(t)dt\right] / \left[\int_0^{t_0} u^2(t)dt\right]}} \tag{7}$$

where $t_0$ represents the time lapse and is selected as 3 s in this study. Regression analysis was conducted using the mixed-effect algorithm, and fitting parameters are provided in the following equation:

$$\overline{\log_{10}(\tau_c)} = 0.301 \times M_w - 1.666 \tag{8}$$

where $M_w$ = moment magnitude; $\tau_c$ = average initial period in second; standard deviations of interevent and intraevent residuals

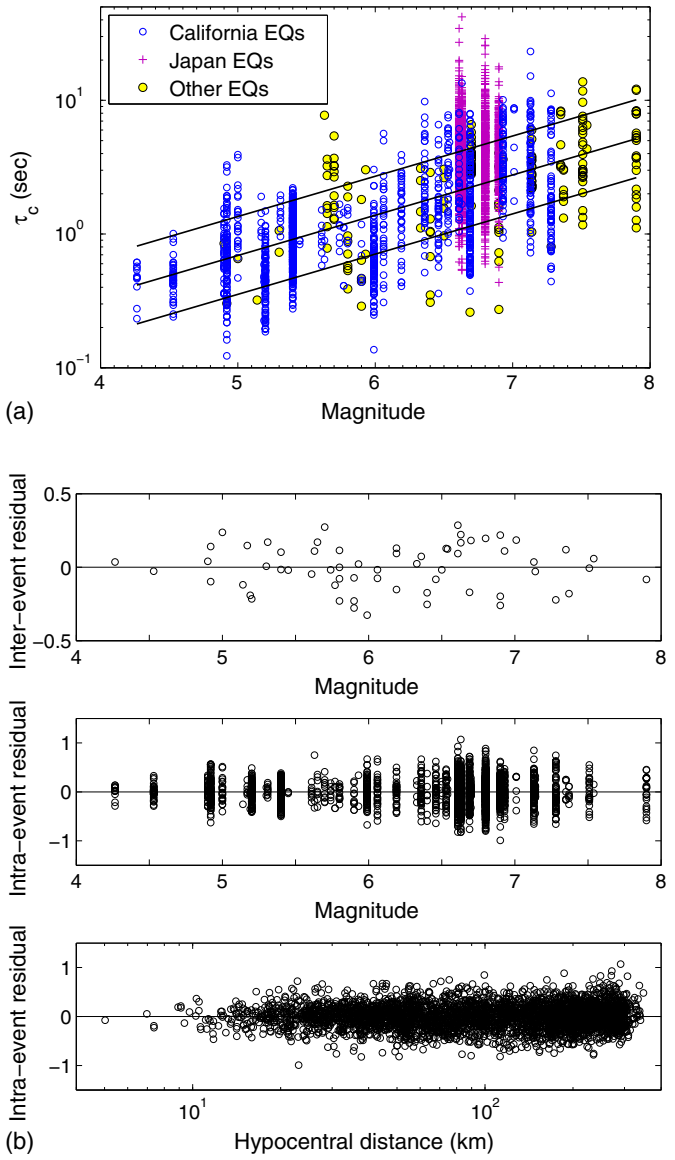


**Fig. 10.** (a) Magnitude-$\tau_c$ relation for the earthquake recordings from the global database where the solid black line represents the median predicted $\tau_c$ using mixed effect regression; median and ±1 standard deviation curves are plotted in black lines; and (b) distributions of interevent and intraevent residuals with respect to earthquake moment magnitudes and hypocenter distances.

(i.e., $\tau$ and $\sigma$) are computed as 0.15 and 0.25, respectively; and the total standard deviation $\sigma_{\text{total}}$ is 0.29. Note that $M_w$ in this study covers the most useful range for engineering application (i.e., $M_w$ = 4–8). But numerically, the variation of magnitude is limited within a narrow range, therefore it would not make much difference whether $\log(M_w)$ or $M_w$ is used in the regression. In this study, we adopted the log-linear relationship for $\tau_c - M_w$.

Fig. 10(a) shows the median and ±1 standard deviation lines for the $\tau_c - M_w$ relation for the global dataset. We also compared $\tau_c - M_w$ relations using only California data, or using California and Japan data. We found that the regressed models are not particularly different from that of the global dataset, which indicates the $\tau_c - M_w$ relationship may not be region specific. Therefore, the global model can be used for all regions. The distributions of interevent and intraevent residuals over earthquake magnitudes and

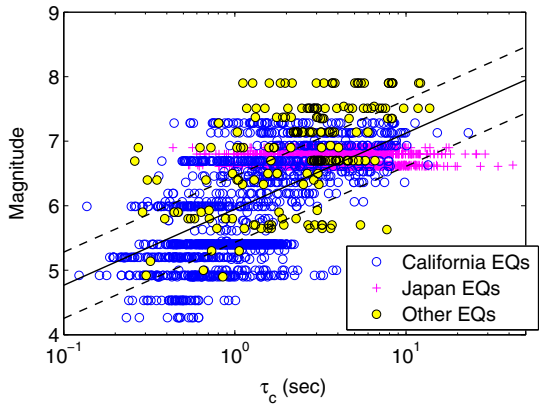


**Fig. 11.** $\tau_c$-magnitude relation for earthquake recordings from the global dataset, showing median and $\pm 1$ standard deviation curves.

intraevent residual over hypocentral distances are presented in Fig. 10(b). The empirical prediction is found unbiased in terms of interevent and intraevent residuals over a broad range of hypocentral distance from 0 to 300 km and earthquake moment magnitudes from 4 to 8.

Note that Eq. (8) is to estimate $\tau_c$ from earthquake magnitudes. In reality, it is more useful to estimate the earthquake magnitude from $\tau_c$ directly. Therefore, the $\tau_c$–magnitude relationship is rewritten as follows:

$$\overline{M_w} = 5.946 + 1.179 \times \log_{10}(\tau_c) \tag{9}$$

with a standard deviation $\sigma_M = 0.51$; and the relationship is plotted in Fig. 11. Note that Eq. (9) is obtained by direct regression analysis, instead of being an inversion of Eq. (8).

## Discussions

### *Can Multivariant Precursors Improve Predictability and Performance of the Empirical Earthquake Early Warning Model?*

In the current study, performance of the empirical $Pd_3$-PGV model has been examined. It has been reported that empirical models perform reasonably well for most cases, with an exception for the Chi-Chi earthquake or pulse-like near-fault records. There remains a question whether the model capacity can be improved by using an alternative $Pd_n$ predictor, or incorporating multiple precursors for more information about the initial portion of a ground-motion time history. To this end, the peak displacement measures in the initial 2, 4, 5, 6, 7, and 8 s of the vertical motion, namely $Pd_2$, $Pd_4$, $Pd_5$, $Pd_6$, $Pd_7$, and $Pd_8$, are obtained to investigate goodness of the linear model using these predictors, as shown in Fig. 12(a). It is shown that standard deviations of all prediction models are within a narrow range (0.34–0.36), indicating that the predictability would not change much by using different $Pd_n$ precursor. However, the advance warning time would be reduced by using a longer precursor. More discussion on the PGV arrival time will be provided in the following section.

Additionally, total residuals estimated using the mixed-effect regression based on a single precursor, $Pd_i$, are correlated with other precursors ($Pd_j$) in Fig. 12(b). It is shown that the total residual associated with precursor $Pd_i$ has virtually no correlation with the precursor itself, as expected. More importantly, the total residual using precursor $Pd_i$ has very weak correlation with another precursor $Pd_j$. For example, residuals associated $Pd_3$ have correlation less than 0.2 with $Pd_8$, indicating that they cannot be explained by this additional precursor. The weak correlation is observed for all pairs of $Pd$ precursors. Therefore, using two or more $Pd$ precursors cannot improve the accuracy of the earthquake early warning system.

### *What Is the Difference between the Conventional Least-Squares Approach and the Mixed-Effect Approach for Strong-Motion Data Analysis?*

Throughout the study the mixed-effect regression technique is used, which is based on a maximum likelihood procedure and accounts for interevent and intraevent variabilities. Conversely, the method of conventional least-squares regression is a procedure to determine the best fit line by minimizing the sum of squares of residuals associated with all data points. Fig. 13(a) compares median regression lines using the two previously mentioned techniques. Above all, the mixed-effect regression provides unbiased estimates of PGVs over a broad range of earthquake magnitudes and hypocentral distance, whereas the between-earthquake and within-earthquake variabilities are not distinguished by using the linear least-squares approach, and results are to some extent biased over magnitudes and distance. It can be visually inspected that the least-squares median line is slightly lower than the mixed-effect line if $Pd_3 < 0.1$ cm, owing to the fact that the least-squares regression result is dominated by a few events with a significant number of small $Pd_3$-PGV values.

### *When Does PGV Arrive after Detection of a P Wave?*

As a few seconds of advance warning can allow people and systems to take precautions before arrival of the most intensive shaking; knowing the arrival time of PGV is yet another important consideration for the efficiency of an early warning system. Figs. 14(a and b) show the distribution of PGV arrival time after detection of the P wave over earthquake moment magnitude and over hypocentral distance. About 20 records whose PGVs arrive within the initial 3 s in a time sequence are highlighted, most of which are ground-motion records from California. Obviously, the $Pd$-based method is not suitable for such events due to insufficient front time for an emergency response. The arrival time of PGV seems to linearly increase with hypocentral distance in a log-log space as shown in Fig. 14(b). On average, the PGV arrival time is 5 s for a site located at the hypocenter distance of 20 km, leaving 2 s for emergency response; while for a site located at 100 km away, the PGV arrival can be 10–40 s.

## Conclusions

In this study, performance of the empirical $Pd_3$-PGV and $\tau_c - M_w$ relationships in earthquake early warning have been evaluated using a significant number of strong-motion records from recent earthquakes in the NGA-West2 database. These two models serve different purposes for the on-site early warning. The $Pd_3$-PGV model aims at predicting the maximum velocity of the ground motion that is strongly correlated to structural damage. Therefore, the model is most useful for precautious operation, e.g., emergency shut down upon an earthquake strike. Conversely, $\tau_c - M_w$ predicts the earthquake magnitude, which would provide a general information of the earthquake source characteristics. These two pieces of information are complementary to each other.

The study leverages the comprehensive database and provides a global model and region-specific empirical models for Japan and California based on state of the art mixed-effect regression

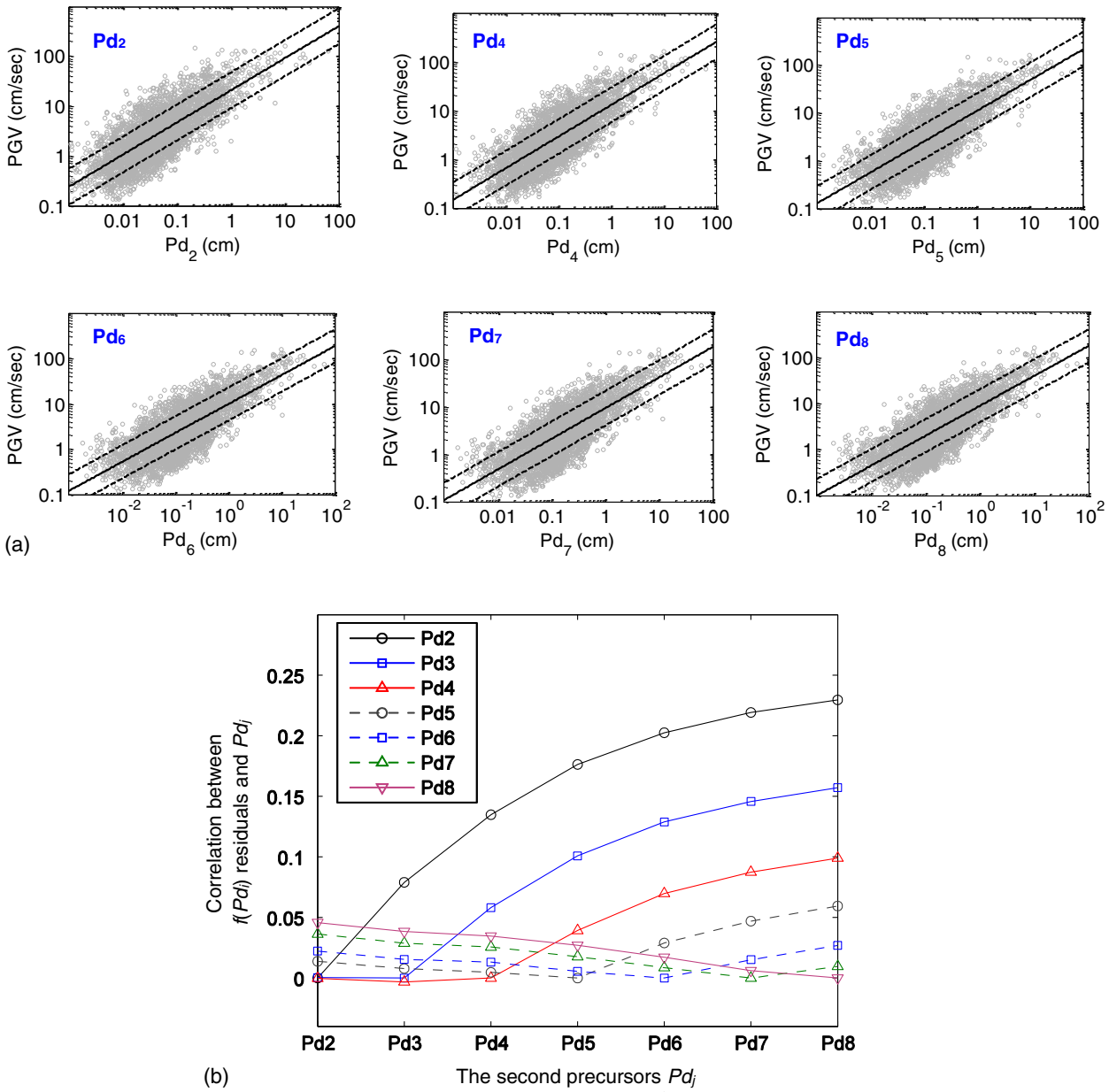


**Fig. 12.** (a) $Pd_n$-PGV relations for precursors $Pd_2$, $Pd_4$, $Pd_5$, $Pd_6$, $Pd_7$, and $Pd_8$; and (b) correlations between prediction residuals of $Pd_n$ and $Pd_n$.

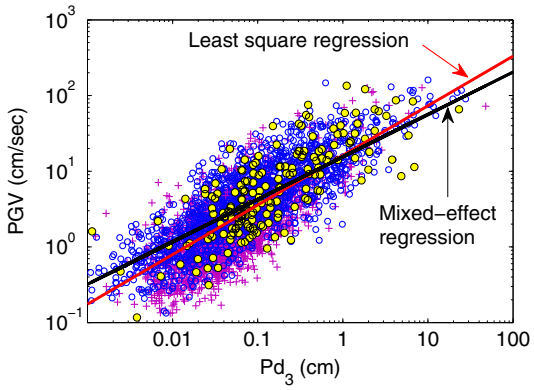


**Fig. 13.** Comparisons of linear least-squares regression with mixed-effect regression using global data.

analyses. These predictive models are found unbiased in terms of interevent and intraevent residuals over a wide range of hypocentral distance and earthquake moment magnitude. Conversely, performance of empirical $Pd_3$-based early warning model for characteristic earthquakes and near-source pulse-like motions are found to be inadequate. The amount of advance warning time can be measured using the arrival time of PGV after the detection of P wave arrival, which is found to increase with hypocentral distance.

The study further demonstrates that incorporating multivariant precursors will not substantially improve predictability of an early warning model. Emphases also have been placed on comparison between regression results using the linear least-squares and the mix-effect regression techniques. The least-squares approach yields a slight underestimate of ground shaking (i.e., PGV) for small $Pd_3$ because the result is governed by densely populated small-amplitude records in the database.

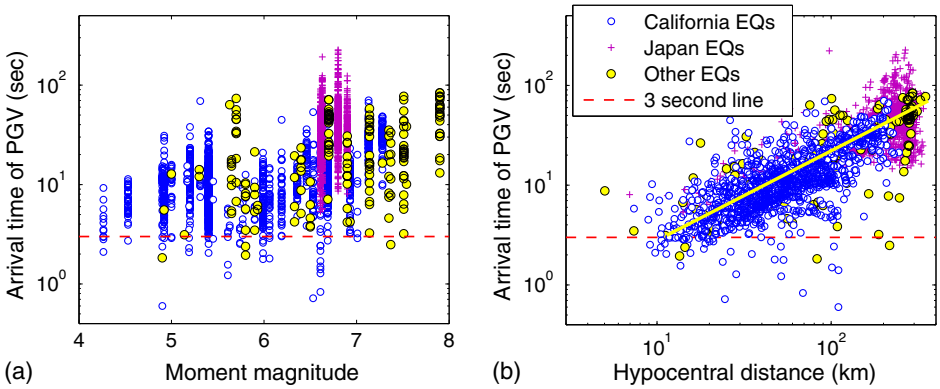


**Fig. 14.** Arrival time of PGVs over (a) moment magnitude; and (b) hypocenter distance.


## Acknowledgments

This research was supported by the Hong Kong Research Grants Council (RGC) under the General Research Fund Grant Nos. 16213615 and 16214118.


## References


Abrahamson, N. A., W. J. Silva, and R. Kamai. 2014. "Summary of the ASK14 ground motion relation for active crustal regions." *Earthquake Spectra* 30 (3): 1025–1055. https://doi.org/10.1193/070913EQS198M.

Abrahamson, N. A., and R. R. Youngs. 1992. "A stable algorithm for regression analyses using the random effects model." *Bull. Seismol. Soc. Am.* 82 (1): 505–510.

Allen, R. M., P. Gasparini, O. Kamigaichi, and M. Bose. 2009. "The status of earthquake early warning around the world: An introductory overview." *Seismol. Res. Lett.* 80 (5): 682–693. https://doi.org/10.1785/gssrl.80.5.682.

Allen, R. V. 1978. "Automatic earthquake recognition and timing from single traces." *Bull. Seismol. Soc. Am* 68( 5): 1521–1532.

Ancheta, T. D., et al. 2014. "NGA-West2 database." *Earthquake Spectra* 30 (3): 989–1005. https://doi.org/10.1193/070913EQS197M.

Baker, J. W. 2007. "Quantitative classification of near-fault ground motions using wavelet analysis." *Bull. Seismol. Soc. Am.* 97 (5): 1486–1501. https://doi.org/10.1785/0120060255.

Boore, D. M., J. P. Stewart, E. Seyhan, and G. M. Atkinson. 2014. "NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes." *Earthquake Spectra* 30 (3): 1057–1085. https://doi.org/10.1193/070113EQS184M.

BSSC (Building Seismic Safety Council). 2003. *NEHRP recommended provisions for seismic regulations for new buildings and other structures*. Technical Rep. No. FEMA 450. Washington, DC: Building Seismic Safety Council, FEMA.

Campbell, K. W., and Y. Bozorgnia. 2014. "NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra." *Earthquake Spectra* 30 (3): 1087–1115. https://doi.org/10.1193/062913EQS175M.

Chen, D. Y., T. L. Lin, Y. M. Wu, and N. C. Hsiao. 2012. "Testing a P-wave earthquake early warning system by simulating the 1999 Chi-Chi, Taiwan, $M_w$ 7.6 earthquake." *Seismol. Res. Lett.* 83 (1): 103–108. https://doi.org/10.1785/gssrl.83.1.103.

Chi, W. C., D. Dreger, and A. Kaverina. 2001. "Finite-source modeling of the 1999 Taiwan (Chi-Chi) earthquake derived from a dense strong-motion network." *Bull. Seismol. Soc. Am.* 91 (5): 1144–1157. https://doi.org/10.1785/0120000732.

Chiou, B. S. J., R. Darragh, N. Gregor, and W. Silva. 2008. "NGA project strong motion database." *Earthquake Spectra* 24 (1): 23–44. https://doi.org/10.1193/1.2894831.

Chiou, B. S. J., and R. R. Youngs. 2014. "Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra." *Earthquake Spectra* 30 (3): 1117–1153. https://doi.org/10.1193/072813EQS219M.

Du, W., and G. Wang. 2013a. "A simple ground-motion prediction model for cumulative absolute velocity and model validation." *Earthquake Eng. Struct. Dyn.* 42 (8): 1189–1202. https://doi.org/10.1002/eqe.2266.

Du, W., and G. Wang. 2013b. "Intra-event spatial correlations for cumulative absolute velocity, Arias intensity, and spectral accelerations based on regional site conditions." *Bull. Seismol. Soc. Am.* 103 (2A): 1117–1129. https://doi.org/10.1785/0120120185.

Du, W., and G. Wang. 2016. "A one-step Newmark displacement model for probabilistic seismic slope displacement hazard analysis." *Eng. Geol.* 205: 12–23. https://doi.org/10.1016/j.enggeo.2016.02.011.

Du, W., and G. Wang. 2017. "Prediction equations for ground motion significant durations using the NGA-West2 database." *Bull. Seismol. Soc. Am.* 107 (1): 319–333. https://doi.org/10.1785/0120150352.

Huang, D., and G. Wang. 2015a. "Region-specific spatial cross-correlation model for stochastic simulation of regionalized ground-motion time histories." *Bull. Seismol. Soc. Am.* 105 (1): 272–284. https://doi.org/10.1785/0120140198.

Huang, D., and G. Wang. 2015b. "Stochastic simulation of regionalized ground motions using wavelet packets and cokriging analysis." *Earthquake Eng. Struct. Dyn.* 44 (5): 775–794. https://doi.org/10.1002/eqe.2487.

Huang, D., and G. Wang. 2017. "Energy-compatible and spectrum-compatible (ECSC) ground motion simulation using wavelet packets." *Earthquake Eng. Struct. Dyn.* 46 (11): 1855–1873. https://doi.org/10.1002/eqe.2887.

Huang, P. L., T. L. Lin, and Y. M. Wu. 2015. "Application of $\tau_c * Pd$ in earthquake early warning." *Geophys. Res. Lett.* 42 (5): 1403–1410. https://doi.org/10.1002/2014GL063020.

Idriss, I. M. 2014. "An NGA-West2 empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes." *Earthquake Spectra* 30 (3): 1155–1177. https://doi.org/10.1193/070613EQS195M.

Kaklamanos, J., and L. G. Baise. 2011. "Model validations and comparisons of the next generation attenuation of ground motions (NGA-West) project." *Bull. Seismol. Soc. Am.* 101 (1): 160–175. https://doi.org/10.1785/0120100038.

Kamigaichi, O., M. Saito, K. Doi, T. Matsumori, S. Tsukada, K. Takeda, T. Shimoyama, K. Nakamura, M. Kiyomoto, and Y. Watanabe. 2009. "Earthquake early warning in Japan: Warning the general public and future prospects." *Seismol. Res. Lett.* 80 (5): 717–726. https://doi.org/10.1785/gssrl.80.5.717.

Ma, K. F., E. E. Brodsky, J. Mori, J. Chen, T. A. Song, and H. Kanamori. 2003. "Evidence for fault lubrication during the 1999 Chi-Chi, Taiwan, earthquake ($M_w$ 7.6)." *Geophys. Res. Lett.* 30 (5): 1–4. https://doi.org/10.1029/2002GL015380.

Somerville, P. G., N. F. Smith, R. W. Graves, and N. A. Abrahamson. 1997. "Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity." *Seismol. Res. Lett.* 68 (1): 199–222. https://doi.org/10.1785/gssrl.68.1.199.

Wang, G. 2011. "A ground motion selection and modification method capturing response spectrum characteristics and variability of scenario earthquakes." *Soil Dyn. Earthquake Eng.* 31 (4): 611–625. https://doi.org/10.1016/j.soildyn.2010.11.007.

Wang, G., and W. Du. 2012. "Empirical correlations between cumulative absolute velocity and spectral accelerations from NGA ground motion database." *Soil Dyn. Earthquake Eng.* 43: 229–236. https://doi.org/10.1016/j.soildyn.2012.07.029.

Wang, G., and W. Du. 2013. "Spatial cross-correlation models for vector intensity measures (PGA, Ia, PGV, and SAs) considering regional site conditions." *Bull. Seismol. Soc. Am.* 103 (6): 3189–3204. https://doi.org/10.1785/0120130061.

Wang, G., R. Youngs, M. Power, and Z. Li. 2015. "Design ground motion library: An interactive tool for selecting earthquake ground motions." *Earthquake Spectra* 31 (2): 617–635. https://doi.org/10.1193/090612EQS283M.

Wu, Y. M., and H. Kanamori. 2005a. "Experiment on an onsite early warning method for the Taiwan early warning system." *Bull. Seismol. Soc. Am.* 95 (1): 347–353. https://doi.org/10.1785/0120040097.

Wu, Y. M., and H. Kanamori. 2005b. "Rapid assessment of damage potential of earthquakes in Taiwan from the beginning of P waves." *Bull. Seismol. Soc. Am.* 95 (3): 1181–1185. https://doi.org/10.1785/0120040193.

Wu, Y. M., and H. Kanamori. 2008. "Development of an earthquake early warning system using real-time strong motion signals." *Sensors* 8 (1): 1–9. https://doi.org/10.3390/s8010001.

Wu, Y. M., H. Kanamori, R. M. Allen, and E. Hauksson. 2007. "Determination of earthquake early warning parameters, $\tau_c$ and Pd, for southern California." *Geophys. J. Int.* 170 (2): 711–717. https://doi.org/10.1111/j.1365-246X.2007.03430.x.

Wu, Y. M., W. T. Liang, H. Mittal, W. A. Chao, C. H. Lin, B. S. Huang, and C. M. Lin. 2016. "Performance of a low-cost earthquake early warning system (P-alert) during the 2016 ML 6.4 Meinong (Taiwan) earthquake." *Seismo. Res. Lett.* 87 (5): 1050–1059. https://doi.org/10.1785/0220160058.

Zeng, Y., and C. H. Chen. 2001. "Fault rupture process of the 20 September 1999 Chi-Chi, Taiwan, earthquake." *Bull. Seim. Soc. Am.* 91 (5): 1088–1098. https://doi.org/10.1785/0120000743.