

## Prediction equations for the effective number of cycles of ground motions for shallow crustal earthquakes



Wenqi Du<sup>a,\*</sup>, Xiaohui Yu<sup>b</sup>, Gang Wang<sup>c</sup>

<sup>a</sup> State Key Laboratory of Water Resources and Hydropower Engineering Science, Institute of Engineering Risk and Disaster Prevention, Wuhan University, 299 Bayi Road, Wuhan, China

<sup>b</sup> School of Civil Engineering, Harbin Institute of Technology, Nangang District, Harbin, China

<sup>c</sup> Dept. of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

### ARTICLE INFO

#### Keywords:

Effective number of cycles  
Empirical equations  
NGA-West2 database  
Mixed-effects model  
Rainflow range-counting

### ABSTRACT

The number of ground motion cycles is one of important characteristics of seismic loadings. This paper presents new prediction equations for the effective numbers of cycles using the mixed-effects model and 7447 ground-motion recordings selected from the NGA-West2 database. Four measures of the effective numbers of ground motion cycles, including two absolute and two relative measures, were computed based on the rainflow range-counting approach. The proposed functional forms employ four predictor variables consisting of moment magnitude  $M$ , rupture distance  $R_{rup}$ , site condition parameter  $V_{s30}$ , depth-to-top-of-rupture parameter  $Z_{tor}$ , and a rupture directivity term  $I_{dir}$ . An additional sediment depth parameter  $Z_1$  is incorporated in the predictive model for the absolute measures. The proposed models are applicable in predicting the effective numbers of cycles subjected to shallow crustal earthquakes with  $M$  ranging from 4 to 7.9, and rupture distance up to 300 km. It is also found that the standard deviations of the relative measures are much smaller than the absolute ones, indicating a higher level of predictability for the relative measures of ground motion cycles.

### 1. Introduction

Seismic loadings are complex and transient excitations, so a complete characterization of earthquake ground motions requires multiple intensity measures (IMs), including measures of a time-series about its peak amplitude, duration, cumulative energy, and frequency content, etc. Among these ground motion characteristics, the number of ground motion cycles, or alternatively, ground motion duration, has been widely regarded as an important parameter in geotechnical earthquake engineering [1]. Numerous studies have stated that the number of cyclic loadings has a great influence on the buildup of pore water pressure in liquefiable soils (e.g. Refs. [2–4]). The effect of the number of uniform amplitude cycles, represented by the so-called magnitude scaling factor, has been commonly incorporated in empirical liquefaction potential assessment [5–7].

In laboratory tests, liquefaction potential assessment is usually conducted in such a manner that soil specimens are subjected to uniform cycles of loadings. Yet, earthquake ground motion is a transient excitation consisting of seismic cycles with irregular amplitude; it is therefore necessary to convert all irregular amplitude cycles to an

equivalent number of uniform cycles (e.g. Refs. [8,9]). Hancock and Bommer [10] reviewed and compared the existing cycle-counting definitions in the literature, and they concluded that the rainflow range-counting approach is the most desirable method in earthquake engineering because it properly quantifies both high-frequency and low-frequency cyclic waves.

As an important parameter in earthquake engineering, the existing prediction equations for the effective number of ground motion cycles are, however, relatively few. The pioneering work was conducted by Seed and Idriss [11], in which the relationship between the number of ground motion cycles and earthquake magnitude was studied. After that, some researchers [12,13] have proposed predictive models based on various ground motion databases, cycle-counting definitions, and functional forms. The detailed information of both models is summarized in Table 1.

The recent compilation of the NGA-West2 database including thousands of global earthquake recordings enables researchers to develop the new generation of predictive models for ground motion IMs. This paper then aims at developing a set of prediction equations for the effective numbers of cycles based on the expanded NGA-West2 ground

\* Corresponding author. State Key Laboratory of Water Resources and Hydropower Engineering Science, Institute of Engineering Risk and Disaster Prevention, Wuhan University, 299 Bayi Road, Wuhan 430072, China.

E-mail address: [wqdu@ntu.edu.sg](mailto:wqdu@ntu.edu.sg) (W. Du).

<https://doi.org/10.1016/j.soildyn.2019.105759>

Received 5 February 2018; Received in revised form 23 September 2018; Accepted 2 July 2019

Available online 12 July 2019

0267-7261/ © 2019 Elsevier Ltd. All rights reserved.

**Table 1**  
Summary of the recently proposed models for the effective number of ground motion cycles.

Cycle parameter <sup>a</sup>	Cycle-counting definitions	No. of earthquakes used	No. of records used	Reference
A relative measure	Peak	150	1664	Liu et al. [12]
N <sub>A</sub> (2.0)	Rainflow range	115	2406	Stafford and Bommer [13]
N <sub>R</sub> (2.0)	Rainflow range	169	7447	This study
N <sub>A</sub> (3.0)				
N <sub>R</sub> (2.0)				
N <sub>R</sub> (3.0)				

<sup>a</sup> N<sub>A</sub> (2.0), N<sub>A</sub> (3.0): absolute measures of ground motion cycles for exponents as 2 and 3, respectively. N<sub>R</sub> (2.0), N<sub>R</sub> (3.0): relative (normalized) measures of ground motion cycles for exponents as 2 and 3, respectively.

motion database. Four measures of the numbers of effective cycles, including two absolute and two relative ones, are first calculated using the rainflow range-counting approach. Appropriate functional forms for these number-of-cycle measures are then derived based on the statistical programming software R [14]. Finally, the performance of the proposed models is examined and compared with existing models, and some recommendations and discussions about their practical use are presented.

## 2. Ground motion database

A subset of the Pacific Earthquake Engineering Research Center's PEER NGA-West2 database [15] is selected in this study. The NGA-West2 ground motion recordings can be downloaded from the PEER website (<https://ngawest2.berkeley.edu/>). The whole NGA-West2 database includes 21,533 three-component uniformly processed recordings from 599 worldwide shallow crustal earthquakes in active tectonic regions. The exclusion criteria employed by Campbell and Bozorgnia [16] are first adopted to exclude some low-quality, unreliable, incomplete or poorly-recorded data, aftershocks, and non-free-field recordings. Besides, recent studies have stated that some low-amplitude ground motions in the NGA-West2 database are possibly polluted by signal noise [17,18]. Inclusion of such motions would inevitably affect the calculated numbers of cycles. Thus, to avoid any bias caused by low-amplitude ground motions, the following selection criteria are further applied: (1) moment magnitude  $M \geq 4$ ; (2) rupture distances  $R_{rup} \leq 300$  km; and (3) peak ground acceleration  $PGA \geq 0.001$  g. Recordings that fail to satisfy these criteria are then excluded from the database.

The resulting final database consists of 7447 recordings from 169 earthquakes with  $M$  ranging from 4 to 7.9, and  $R_{rup}$  ranging from 0.1 km to 299.54 km. The  $M$ - $R_{rup}$  distribution, and the histograms of the time-averaged shear wave velocity in top 30 m  $V_{s30}$ , depth to the 1.0 km/s shear wave isosurface  $Z_1$  for ground motion recordings, and the depth to the top of the fault rupture  $Z_{tor}$  for earthquake events are summarized in Fig. 1. Note that for each recording, the geometrical mean of the calculated numbers of cycles for the two horizontal components is used for developing the empirical models. Although the physical interpretation of the geometric mean of the number of cycles seems ambiguous, it represents the average value from the two horizontal components of a ground motion in logarithmic scale. Besides, a NGA-West2 flatfile including the detailed information of these ground motion recordings was downloaded from the PEER website (<http://peer.berkeley.edu/ngawest2/databases/>).

Fault rupture directivity can greatly influence the ground motion characteristics, generally resulting in pulse-like ground motion recordings in some near-fault locations. Two conditions are required for producing forward directivity effects: the fault ruptures towards a site,

and the slip direction is aligned with this site [19]. Based on a recent study [20] that proposed a new algorithm to identify directivity-like ground motions from the NGA-West2 database, 131 directivity-like ground motions are identified to examine the effect of forward directivity on the effective numbers of ground motion cycles. Fig. 2 displays the  $M$ - $R_{rup}$  distribution of the directivity-like recordings contained in the present database. It can be seen that most of the directivity-like recordings have a rupture distance less than 40 km.

## 3. Development of predictive models

### 3.1. Calculation of the effective numbers of cycles

As summarized by Hancock and Bommer [10], there are dozens of cycle counting definitions in the literature, which can be broadly classified into several categories: peak counting, level crossing counting, range counting, and indirect counting approaches. In this study, the rainflow range-counting method [21] is adopted, because it properly quantifies the contribution of both high-frequency and low-frequency cyclic waves contained in a broad-band signal. The rainflow method is widely used to assess the fatigue life of a structure subject to complex loadings. This algorithm counts a history of peaks and troughs in sequence for a given time-series, and they are regarded as starting and ending points for defining each cycle. Thus, a series of half-cycles and full-cycles with various amplitudes are obtained, and the number of cycles can be calculated by summation. One can refer to Refs. [10,22] for the detailed algorithm of the rainflow range-counting approach.

The absolute definition of the effective number of cycles can be expressed as:

$$N_A(j) = \sum_{i=1}^{2T_n} u_i^j \quad (1)$$

where  $N_A$  denotes the absolute measures of the effective numbers of cycles;  $u_i$  is the amplitude of the  $i$ -th half cycle obtained by the aforementioned rainflow range-counting approach;  $T_n$  is the total number of cycles; and  $j$  is the exponent coefficient, affecting the relative importance of different amplitude cycles. Note that a higher value of  $j$  represents a greater contribution of the large amplitude cycles to the effective number of cycles calculated. Many studies have investigated the relationship between cyclic resistance ratio and the number of cycles for soil liquefaction potential (e.g. Ref. [23]). For instance, Liu et al. [12] reported that the exponent values are in the range of 2–3 based on laboratory tests and field observations; Idriss and Boulanger [6] recommended an exponent coefficient of 3 for clean sands based on high-quality samples. Thus, two exponent coefficients, namely 2 and 3, are employed to compute the numbers of cycles for both absolute and relative measures in this study.

A typical relative definition of the effective numbers of cycles, in which each amplitude  $u_i$  is normalized by the maximum amplitude of all half-cycles  $u_{max}$ , is expressed as:

$$N_R(j) = \frac{1}{2} \sum_{i=1}^{2T_n} \left( \frac{u_i}{u_{max}} \right)^j \quad (2)$$

where  $N_R$  denotes the relative measures of the effective cycles. As mentioned above, the values of 2 and 3 are used for the exponent coefficient  $j$ . Therefore, four definitions of the effective numbers of cycles, namely  $N_A$  (2.0),  $N_A$  (3.0),  $N_R$  (2.0), and  $N_R$  (3.0), are considered in this study. The empirical values for these measures will be used for subsequent regression analyses.

### 3.2. Magnitude and distance scalings for empirical data

Unlike other ground motion IMs (e.g., the source duration term in predictive models for significant duration [17]) that can be theoretically estimated based on seismological models, there are no theoretical

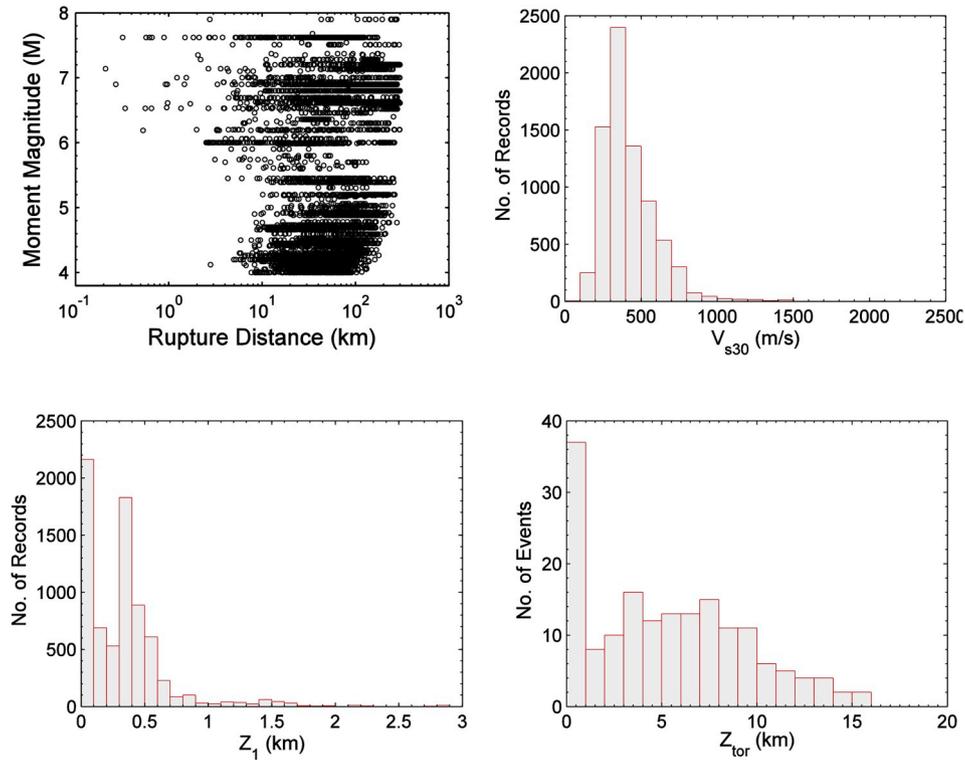


Fig. 1. Moment magnitude-rupture distance distribution, the histograms of  $V_{s30}$  and  $Z_1$  for ground motion recordings, and the histogram of  $Z_{tor}$  for earthquake events used in this study.

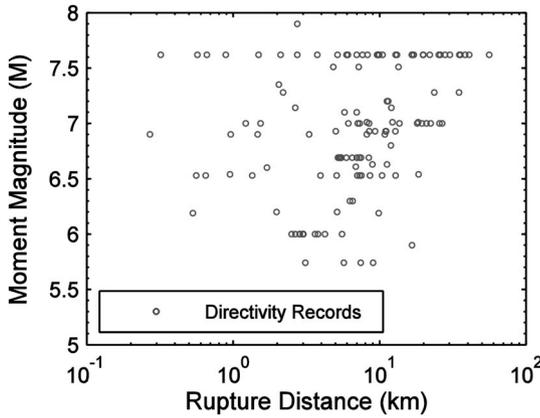


Fig. 2. Moment magnitude and rupture distance distribution of the forward directivity recordings used in this study.

source functions associated with the prediction of the number of ground motion cycles. Thus, we focused on the derivation of functional forms mainly based on the empirical data. The empirical data are first split into two rupture distance (magnitude) bins, and then the empirical data trends with respect to moment magnitude (rupture distance) for  $N_A(2.0)$  and  $N_R(2.0)$  are shown in Figs. 3 and 4, respectively. At each plot, the whole M (or  $R_{rup}$ ) range is evenly partitioned into several bins, and the trend line consisting of the mean value of each binned empirical data is displayed. For the absolute measures, it can be seen that  $N_A(2.0)$  generally increases with increasing M, except that the  $N_A(2.0)$ - $M_w$  trend keeps almost constant for large-magnitude ( $M > 6.5$ ) and short-source distance ( $R_{rup} < 80$  km) earthquake scenarios. Besides, Fig. 3b illustrates that  $N_A(2.0)$  generally decreases with increasing  $R_{rup}$ , but the decreasing trend varies at different magnitude and distance intervals. At small magnitudes ( $M < 5.5$ ),  $N_A(2.0)$  decreases fastest in the short-distance range ( $R_{rup} < 60$  km), while the slope of the decreasing

trend becomes much flatter over the other distance range. At large magnitudes ( $M > 5.5$ ),  $N_A(2.0)$  decreases almost linearly versus  $R_{rup}$ , as shown in the right plot.

As is shown in Fig. 4a, the mean curve for binned  $N_R(2.0)$  keeps approximately constant at relatively small magnitudes ( $M < 5.5$ ), while it increases as M increases in the  $M > 5.5$  magnitude range. In particular, the increasing trend in the short distance range ( $R_{rup} < 80$  km) is more pronounced than that of the far distance range ( $80 < R_{rup} < 300$  km). In addition, as shown in Fig. 4b, the  $N_R(2.0)$ - $R_{rup}$  trend increases linearly within the short distance range ( $R_{rup} < 50$  km), whereas it keeps almost constant at moderate-to-far distances for both small and large magnitude ranges. The plots in Figs. 3 and 4 demonstrate the M- and  $R_{rup}$ -dependence of  $N_A(2.0)$  and  $N_R(2.0)$ , which could provide some insights when choosing appropriate functional forms during the regression process.

### 3.3. Functional forms and the implications

The mixed-effects model widely used in earthquake engineering is adopted for developing new predictive models for  $N_A$  and  $N_R$ . The mixed-effects model can separate the total residuals into between-event and within-event residuals. Under the assumption that the distribution of the number of ground motion cycles is lognormal, the models proposed take the forms as:

$$\ln(N_x)_{ij} = \overline{\ln(N_x)_{ij}} + \eta_i + \varepsilon_{ij} \quad (3)$$

Where  $N_x$  denotes the effective numbers of cycles considered in this study (i.e.,  $N_A(2.0)$ ,  $N_A(3.0)$ ,  $N_R(2.0)$ , and  $N_R(3.0)$ );  $\overline{\ln(N_x)_{ij}}$  is the logarithm of the predicted  $N_x$  for the  $j$ -th recording from the  $i$ -th event;  $\eta_i$  and  $\varepsilon_{ij}$  denote the between-event and within-event residuals, which are normally distributed with zero means and standard deviations  $\tau$  and  $\phi$ , respectively. The total standard deviation  $\sigma$  is calculated as  $\sigma = \sqrt{\tau^2 + \phi^2}$ .

The regression analysis was conducted for  $N_A$  and  $N_R$  separately, using the “nlme” package implemented in the statistical programming

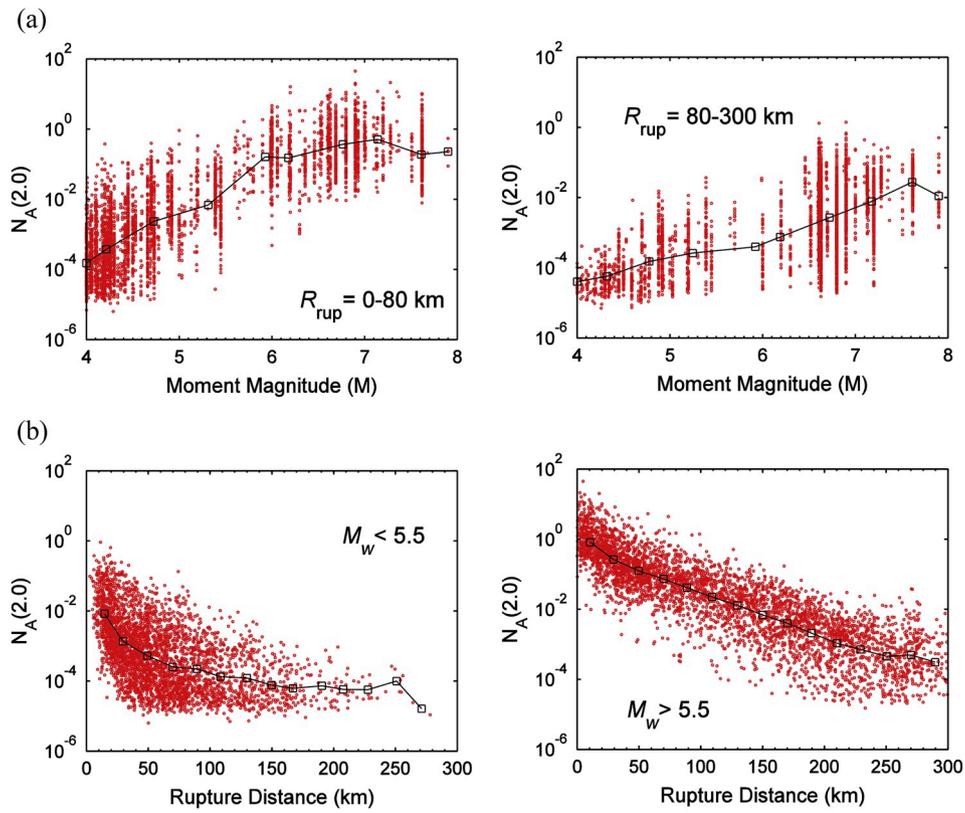


Fig. 3. Distributions of the empirical  $N_A(2.0)$  with respect to (a) moment magnitude, and (b) rupture distance, respectively.

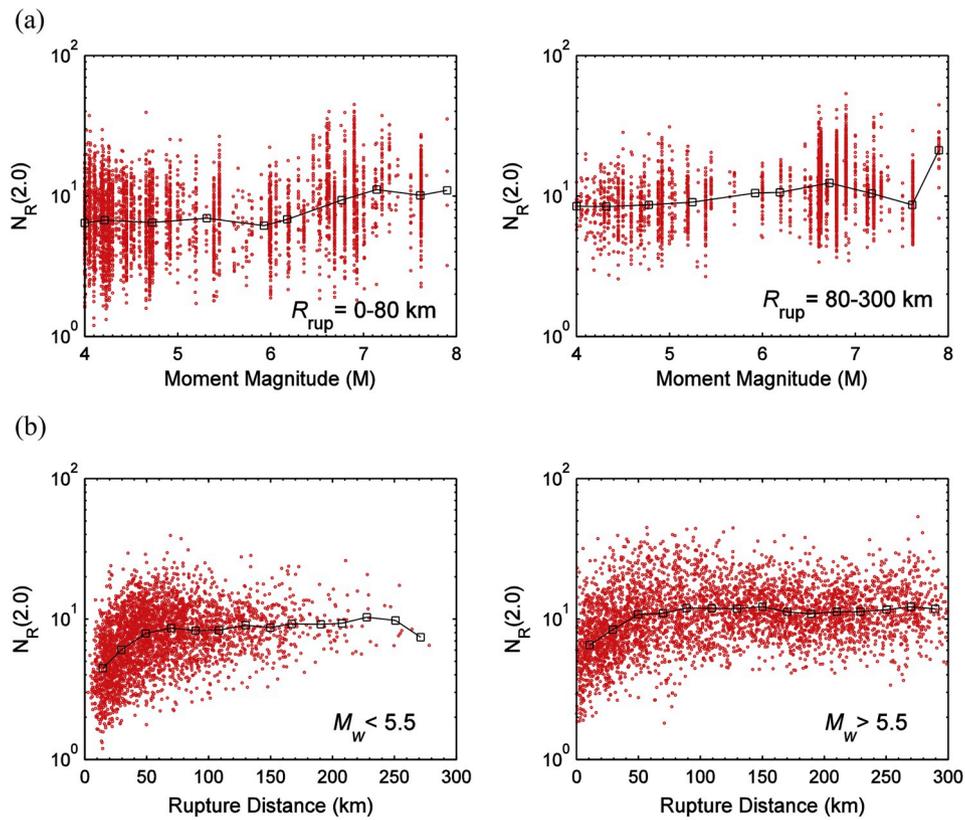


Fig. 4. Distributions of the empirical  $N_R(2.0)$  with respect to (a) moment magnitude, and (b) rupture distance, respectively.

software **R** [14]. Based on a trial and error approach, many candidate functional forms were tried and tested for fitting the empirical data. The final functional forms are determined according to a combined consideration of simplicity, predictive efficiency, unbiased distribution of residuals, and physical basis.

### 3.3.1. Functional form for $N_A$

The functional form chosen for the absolute measures  $N_A$  is presented as:

$$\ln(N_A) = f_{mag} + f_{dis,mag} + f_{V_{s30},Z_1} + f_{Z_{tor}} + f_{dir} \quad (4)$$

$$f_{mag} = c_1 + c_2 \cdot M_1 + c_3 \cdot M_1^2 \quad \text{where } M_1 = \min(M, 7.2) \quad (5)$$

$$f_{dis,mag} = \begin{cases} (c_4 + c_5 \cdot M_1) \cdot (R_{rup} - 30) + c_6 \cdot R_{rup}; & R_{rup} \leq 30 \text{ km} \\ c_6 \cdot R_{rup}; & 30 \text{ km} < R_{rup} \leq 60 \text{ km} \\ c_6 \cdot R_{rup} + c_7 \cdot (R_{rup} - 60); & 60 \text{ km} < R_{rup} \leq 200 \text{ km} \\ c_6 \cdot R_{rup} + c_7 \cdot (R_{rup} - 60) + c_8 \cdot (R_{rup} - 200); & R_{rup} > 200 \text{ km} \end{cases} \quad (6)$$

$$f_{V_{s30},Z_1} = c_9 \cdot \ln(V_{s30}) + f_{\delta Z_1} \quad (7)$$

$$f_{\delta Z_1} = \begin{cases} 0 & \delta Z_1 \leq 0.3 \text{ km} \\ c_{10} \cdot (\delta Z_1 - 0.3) & \delta Z_1 > 0.3 \text{ km} \end{cases} \quad (8)$$

$$f_{Z_{tor}} = c_{11} \cdot Z_{tor} \quad (9)$$

$$f_{dir} = c_{12} \cdot I_{dir} \quad (10)$$

where  $M$  is moment magnitude;  $R_{rup}$  is rupture distance (km);  $V_{s30}$  is the time-averaged shear wave velocity of the upper 30 m (m/s);  $Z_1$  is the depth to the 1.0 km/s shear wave isosurface;  $Z_{tor}$  is the depth to the top of the fault rupture (km); and  $I_{dir}$  is an indicator variable representing directivity conditions ( $I_{dir} = 1$  for forward directivity effect, and  $I_{dir} = 0.0$  otherwise). We observed an influence of the sediment depth effect on the  $N_A$  magnitude, so such effect is quantified by sediment depth differential  $\delta Z_1$ . The term  $\delta Z_1$ , which is also employed by recently developed ground motion prediction equations (e.g. Refs. [24–26]), measures the difference between the empirical  $Z_1$  and the predicted median depth  $\mu_{z_1}$  as:

$$\delta Z_1 = Z_1 - \mu_{z_1}(V_{s30}) \quad (11)$$

where  $\mu_{z_1}(V_{s30})$  can be calculated based on empirical relationships for specific regions. As summarized in Boore et al. [24], the empirical  $Z_1$ - $V_{s30}$  equations for California and Japan are expressed as:

$$\text{California: } \ln(\mu_{z_1}) = \frac{-7.15}{4} \ln\left(\frac{V_{s30}^4 + 570.94^4}{1360^4 + 570.94^4}\right) - \ln(1000) \quad (12a)$$

$$\text{Japan: } \ln(\mu_{z_1}) = \frac{-5.23}{2} \ln\left(\frac{V_{s30}^2 + 412.39^2}{1360^2 + 412.39^2}\right) - \ln(1000) \quad (12b)$$

where  $\mu_{z_1}$  and  $V_{s30}$  are in the units of km and m/s, respectively. For applications where the measured or predicted  $Z_1$  is not available, a default value of  $\delta Z_1$  as 0 is recommended. Note that some other seismological parameters, such as fault category indicators, were also tried and tested during the regression process, but it was observed that they are statistically insignificant. The calculated coefficients for  $N_A$  (2.0) and  $N_A$  (3.0), associated with the corresponding standard errors and  $p$ -values, are summarized in Tables 2 and 3, respectively. It can be seen that all coefficients yield very small  $p$ -values, concluding that they are statistically significant.

The proposed model includes the scaling of  $N_A$  with respect to  $M$ ,  $R_{rup}$ , site condition,  $Z_{tor}$ , and  $I_{dir}$ . It is seen that a quadratic  $M$ -scaling is employed in this model; such formulation can properly fit the empirical

**Table 2**

Regression parameters of the proposed  $N_A$  (2.0) model.

Parameter	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
Value	-21.06	5.220	-0.218	-0.221	0.025	-0.038	0.0074
Std. Error	1.950	0.733	0.067	0.013	0.002	0.002	0.002
$p$ -value	0.000	0.000	0.001	0.000	0.000	0.000	0.000
Parameter	$c_8$	$c_9$	$c_{10}$	$c_{11}$	$c_{12}$	$\tau$	$\phi^a$
Value	0.010	-0.764	0.309	0.082	-0.411	0.549	1.033
Std. Error	0.001	0.035	0.073	0.018	0.118	-	-
$p$ -value	0.000	0.000	0.000	0.000	0.000	-	-

<sup>a</sup> Note: the total standard deviation  $\sigma$  can be computed as  $\sigma = \sqrt{\phi^2 + \tau^2}$ .

**Table 3**

Regression parameters of the proposed  $N_A$  (3.0) model.

Parameter	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
Value	-34.62	8.426	-0.380	-0.377	0.044	-0.062	0.016
Std. Error	2.929	1.100	0.100	0.020	0.003	0.003	0.003
$p$ -value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Parameter	$c_8$	$c_9$	$c_{10}$	$c_{11}$	$c_{12}$	$\tau$	$\phi$
Value	0.015	-1.232	0.422	0.132	-0.479	0.795	1.597
Std. Error	0.002	0.054	0.112	0.026	0.182	-	-
$p$ -value	0.000	0.000	0.000	0.000	0.009	-	-

magnitude scaling over the broad  $M$  range. A piece-wise additive  $R_{rup}$ -scaling term is incorporated, which allows more flexibility in capturing the various decreasing trends at different distance intervals (Fig. 3b). Besides, the sediment thickness  $Z_1$  and  $V_{s30}$  are used together to model the effect of site conditions. The proposed model includes a negative  $N_A$ - $V_{s30}$  correlation and a positive  $N_A$ - $Z_1$  correlation (see the coefficients  $c_9$  and  $c_{10}$  listed in Tables 2 and 3), indicating that the ground motions recorded at softer soil sites or at sites with deeper sediment thickness will, on average, have larger  $N_A$  values. This can be explained by the fact that  $N_A$ , as an absolute measure, is highly correlated with peak-amplitude IMs such as PGA [27]. The effects of site conditions on  $N_A$  and PGA are thus similar. Yet, unlike PGA, the soil nonlinearity effect on  $N_A$  is found to be insignificant. This is perhaps because compared to PGA,  $N_A$  quantifies not only the peak-amplitude cycle, but also a number of secondary ground motion cycles. Thus, a linear  $V_{s30}$  term only is incorporated in the  $N_A$  functional form.

We also observed a positive  $N_A$ - $Z_{tor}$  correlation, which is presently difficult to explain, yet such observation is consistent with a previous study [13]. In addition, a constant term is incorporated in the functional form to quantify the effect of rupture directivity. The coefficient  $c_{12}$  is negative, indicating that the rupture directivity effect causes a reduction in the number of ground motion cycles, which is expected due to the more low-frequency components of directivity-like recordings.

### 3.3.2. Functional form for $N_R$

The final functional form for the relative measures  $N_R$  is presented as:

$$\ln(N_R) = f_{mag} + f_{dis,mag} + f_{V_{s30}} + f_{Z_{tor}} + f_{dir} \quad (13)$$

$$f_{mag} = \begin{cases} c_1 & M < 5.5 \\ c_1 + c_2 \cdot (M - 5.5) & M \geq 5.5 \end{cases} \quad (14)$$

$$f_{dis,mag} = \begin{cases} c_3 \cdot R_{rup} \cdot \log\left(\frac{M_2}{8}\right); & R_{rup} \leq 50 \text{ km} \\ c_3 \cdot 50 \cdot \log\left(\frac{M_2}{8}\right); & R_{rup} > 50 \text{ km} \end{cases} \quad (15)$$

$$f_{V_{s30}} = c_4 \cdot \ln(V_{s30}) \quad (16)$$

$$f_{Z_{tor}} = c_5 \cdot Z_{tor} \quad (17)$$

**Table 4**  
Regression parameters of the proposed  $N_R$  models.

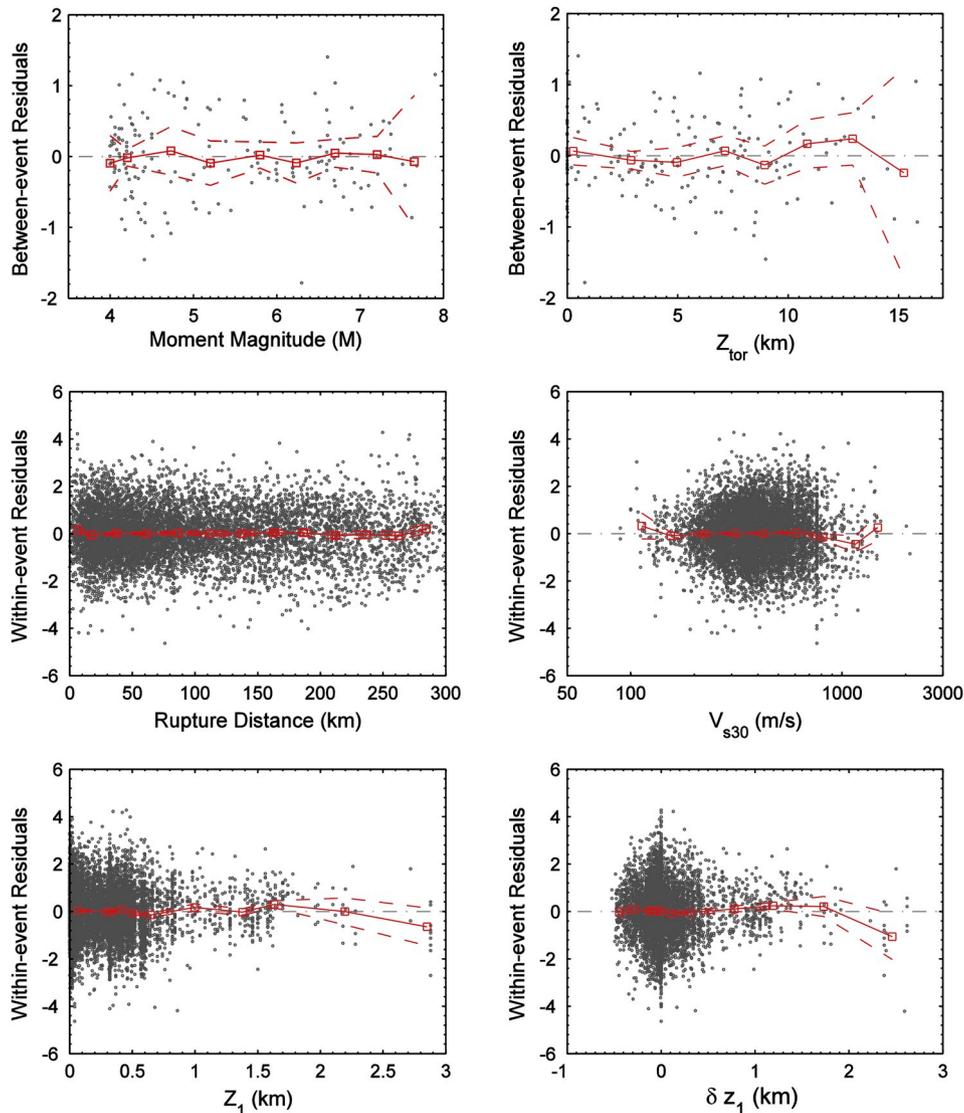
$N_R$ (2.0)	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$\tau$	$\phi$
Value	0.846	0.414	-0.048	0.095	-0.015	-0.266	0.157	0.392
Std. Error	0.092	0.036	0.001	0.013	0.005	0.039	-	-
$p$ -value	0.000	0.000	0.000	0.000	0.003	0.000	-	-
$N_R$ (3.0)	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$\tau$	$\phi$
Value	0.503	0.378	-0.040	0.060	-0.017	-0.261	0.128	0.378
Std. Error	0.086	0.031	0.001	0.012	0.005	0.038	-	-
$p$ -value	0.000	0.000	0.000	0.000	0.001	0.000	-	-

$$f_{dir} = c_6 \cdot I_{dir} \quad (18)$$

where the definitions of  $M$ ,  $R_{rup}$ ,  $V_{s30}$ ,  $Z_{tor}$ , and  $I_{dir}$  are explained above, and  $M_2$  is a truncated magnitude parameter defined as  $M_2 = \min[\max(M, 5.5), 7.2]$ . Note that the sediment thickness parameter  $Z_1$ , which was employed for predicting  $N_A$  in Eq. (4), is not included in the functional form for  $N_R$  due to its statistical insignificance. The regressed coefficients, along with the standard errors and  $p$ -values, are listed in Table 4. Again, the small  $p$ -values denote that these coefficients are statistically significant.

The proposed functional form includes a bilinear  $M$  scaling term:  $N_R$  is independent of  $M$  at  $M < 5.5$ , and it increases with increasing  $M$  in the  $M \geq 5.5$  range. This bilinear  $M$ -term is supported by the distribution of the empirical data shown in Fig. 4a. The increasing trend of  $N_R$  at moderate-to-large magnitudes is physically anticipated. Based on theoretical seismological models (e.g. Ref. [28]), large-magnitude earthquakes can generate ground motions with relatively broadband seismic waves and long source duration, which tends to yield an increase in the relative number of ground motion cycles.

The increasing  $R_{rup}$ - $N_R$  trend is restricted at short distances only ( $R_{rup} < 50$  km), and a  $R_{rup}$ -independent term for  $N_R$  is considered at moderate-to-far distances. The notable increase of  $N_R$  at short distances ( $R_{rup} < 50$  km) is expected, which is mainly caused by the increasing refractions and reflections of seismic waves over the travelling path, resulting in more effective cycles of ground motions. Yet, readers may be surprised about the choice of the  $R_{rup}$ -independent scaling at  $R_{rup} > 50$  km. As illustrated in Fig. 4b,  $N_R$  keeps roughly constant as  $R_{rup}$  increases for bins with  $R_{rup}$  larger than 50 km. This observation can be explained as follows. As travelling distance increases, although the increasing refractions and reflections of seismic waves yield more ground motion cycles, the high-frequency component of seismic waves



**Fig. 5.** Distributions of the between-event residuals with respect to  $M$ ,  $Z_{tor}$ , and within-event residuals with respect to  $R_{rup}$ ,  $V_{s30}$ ,  $Z_1$ , and  $\delta Z_1$  for the absolute measure  $N_A$  (2.0). Note that in these, and subsequent plots, the red square symbol denotes the local mean value of each binned residuals, and the dashed curves denote their 95% confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

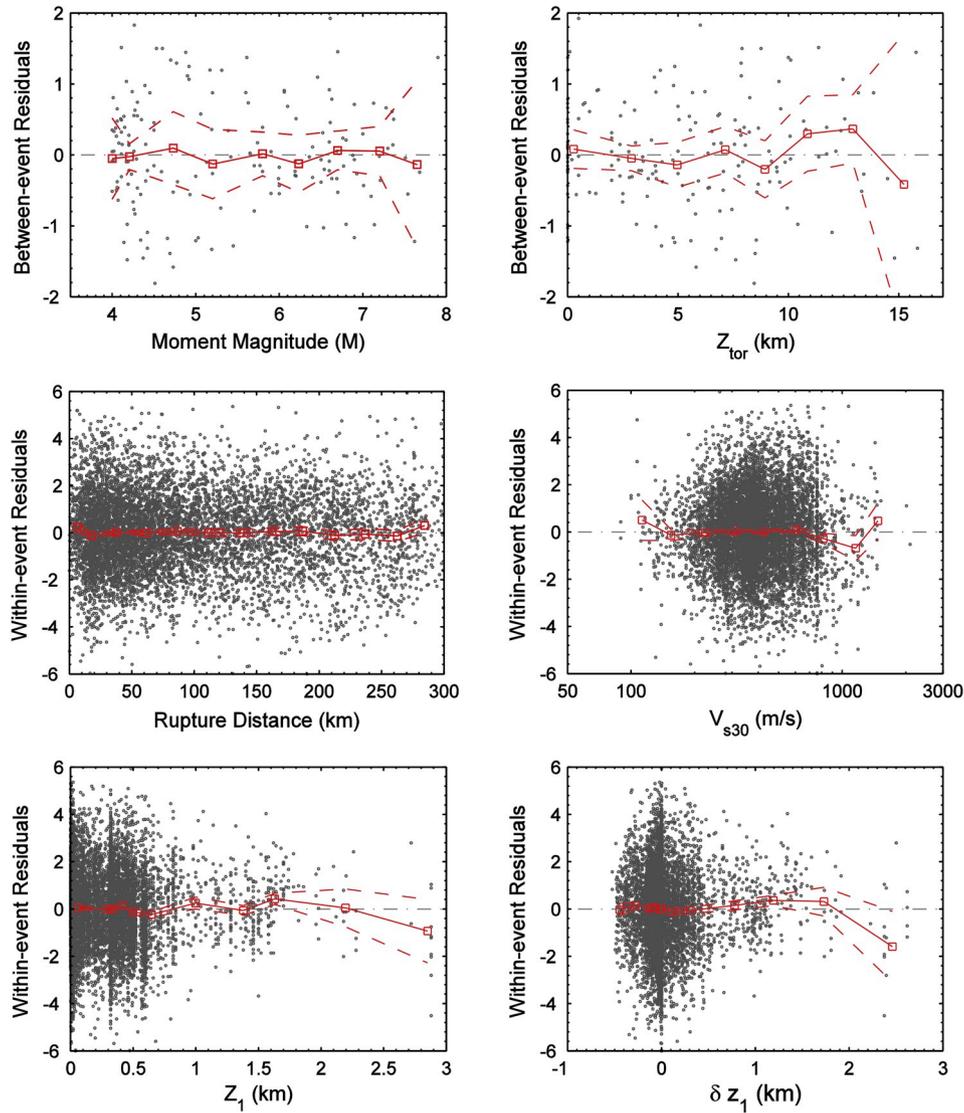


Fig. 6. Distributions of the between-event residuals with respect to  $M$ ,  $Z_{tor}$ , and within-event residuals with respect to  $R_{rup}$ ,  $V_{s30}$ ,  $Z_1$ , and  $\delta z_1$  for the absolute measure  $N_A$  (3.0).

decays much faster than the low-frequency component, which inevitably reduces the contribution of the high-frequency component to the effective number of cycles. We speculate that at moderate-to-far distances ( $R_{rup} > 50$  km), both effects on the magnitude of  $N_R$  tend to counteract each other, so  $N_R$  is relatively insensitive to the change of  $R_{rup}$  within such a distance range. Besides, a  $M_2$ -dependent term is incorporated in  $f_{dis,mag}$  to better capture  $N_R$ - $R_{rup}$  increasing trend ( $R_{rup} < 50$  km) at various magnitude intervals.

It is well recognized that soft site conditions can generally amplify the amplitudes of ground motion cycles. The functional form indicates that there is a positive correlation between  $V_{s30}$  and  $N_R$ , which might be explained as a result of different amplifications of ground motion cycles. At soft sites (with small  $V_{s30}$  values), the amplification of the peak cycle amplitude ( $u_{max}$ ) might be, by average, greater than those of the other individual cycles, resulting in a slight decrease of  $N_R$ .

The coefficient of the  $Z_{tor}$  scaling for  $N_R$  is slightly negative, which is perhaps because the depth of buried ruptures influences the scattering of seismic wave propagation. As explained by Bommer and Stafford [13], the waves of a deeper event are possibly less scattered and dispersed, leading to a decrease in the magnitude of  $N_R$ . Besides, similar to  $N_A$ , the rupture directivity effects also cause a reduction in the magnitude of  $N_R$ .

### 3.4. Residual analysis

The models developed need to be validated by checking the distributions of residuals. Figs. 5 and 6 show the between-event residuals with respect to  $M$ ,  $Z_{tor}$ , and the within-event residuals with respect to  $R_{rup}$ ,  $V_{s30}$ ,  $Z_1$ , and  $\delta z_1$  for  $N_A$  (2.0) and  $N_A$  (3.0), respectively. To better inspect the residual trend, the residuals are partitioned into several equally spaced bins; the local means as well as the 95% confidence intervals of the binned residuals are shown in each plot. It can be seen that both the between- and within-event residuals are generally unbiased versus the predictor variables. There are a few slight biases (e.g., within-event residuals at  $\delta z_1 > 2$  km), which are inevitable due to a paucity of data.

Figs. 7 and 8 illustrate the between-event and within-event residuals with respect to  $M$ ,  $Z_{tor}$ ,  $R_{rup}$ ,  $V_{s30}$ , and  $Z_1$  for  $N_R$  (2.0) and  $N_R$  (3.0), respectively. The local binned means and the 95% confidence intervals are also shown in each plot. It is interesting to see that, although very simple  $M$  and  $R_{rup}$  terms are incorporated in the functional form, no clear biases or trends with respect to these predictor variables can be observed, validating the appropriateness of the functional form adopted. Besides, the generally unbiased distribution of the residuals versus  $Z_1$  confirms that the  $N_R$  values are independent of  $Z_1$ . The plots in Figs. 5–8 indicate that the proposed models fit the empirical data

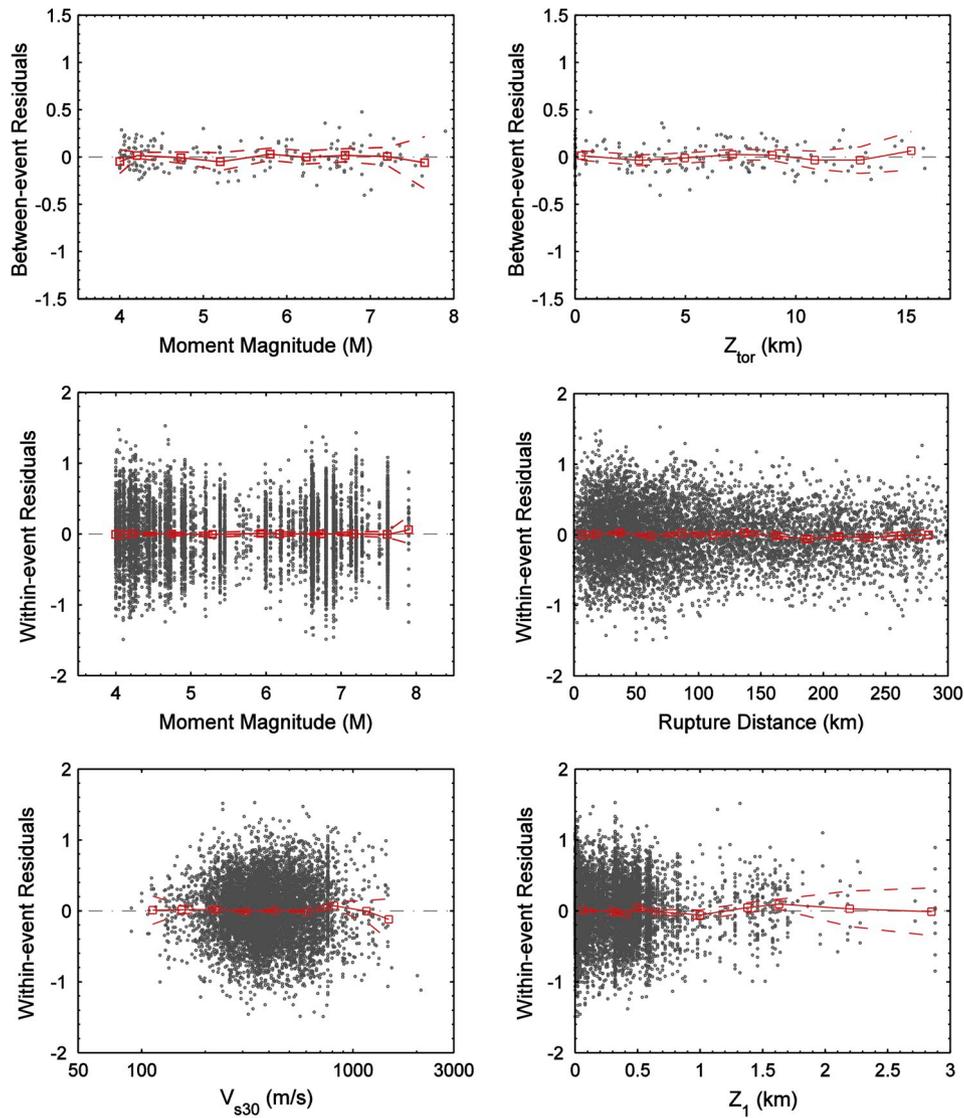


Fig. 7. Distributions of the between-event residuals with respect to  $M$ ,  $Z_{tor}$ , and within-event residuals with respect to  $M$ ,  $R_{rup}$ ,  $V_{s30}$ , and  $Z_1$  for the relative measure  $N_R$  (2.0).

reasonably well.

Fig. 9a and b further display the calculated within-event residuals of the forward directivity recordings against rupture distance for  $N_A$  (2.0) and  $N_R$  (2.0), respectively. The computed means of the binned residuals are also shown in each plot. In general, the residuals are not significantly deviated from the zero line, and the few biased local means (e.g.,  $30 < R_{rup} < 40$  km in Fig. 9a) are mainly caused by the limited data. Therefore, the proposed models reasonably address the rupture directivity effect on the effective number of ground motion cycles.

The normality assumption also needs to be examined for the  $N_A$  and  $N_R$  models developed. Q-Q plot (quantile-quantile plot) is used herein for visually checking the univariate normality of the between-event and within-event residuals. The Q-Q plot is a scatterplot that compares the empirical data quantiles and the corresponding theoretical quantiles [29]. It will lie on the “ $y = x$ ” line if it is a perfect case. Fig. 10a and b shows the normalized Q-Q plots of the between-event and within-event residuals for  $N_A$  (2.0) and  $N_R$  (2.0), respectively. The theoretical quantiles are calculated from the standard normal distribution. It can be seen that the empirical data trends match closely with the “ $y = x$ ” line, especially in the range of  $-3$  to  $3$ . Besides, although not shown here, the residuals of  $N_A$  (3.0) and  $N_R$  (3.0) exhibit similar trends of Q-Q plots. Thus, it is concluded that the normality assumption of the

regression models for logarithmic  $N_A$  and  $N_R$  is reasonable.

### 3.5. Aleatory variability terms

As mentioned above, the total standard deviation consists of two components, namely the between-event standard deviation  $\tau$  and the within-event standard deviation  $\phi$ . In this subsection, the between- and within-event residuals are partitioned into several  $M$  bins, and in each bin the standard deviation of the residuals is calculated. Fig. 11a and b shows the computed binned standard deviations of  $\tau$  and  $\phi$  with  $M$  for  $N_A$  (2.0) and  $N_R$  (2.0), respectively. It can be seen that for both measures  $N_A$  and  $N_R$ , the standard deviations  $\tau$  and  $\phi$  do not show a clear dependence on magnitude. Consequently, it is recommended to use the average value of these binned standard deviations (the solid curve in each plot), and such constant  $\tau$  and  $\phi$  values are appropriate for all earthquake scenarios. The standard deviation values of  $\tau$  and  $\phi$  are listed in Tables 2 and 3 for  $N_A$  (2.0) and  $N_A$  (3.0), and in Table 4 for  $N_R$  (2.0) and  $N_R$  (3.0), respectively.

## 4. Model performance

The performance of the new models is examined and compared with

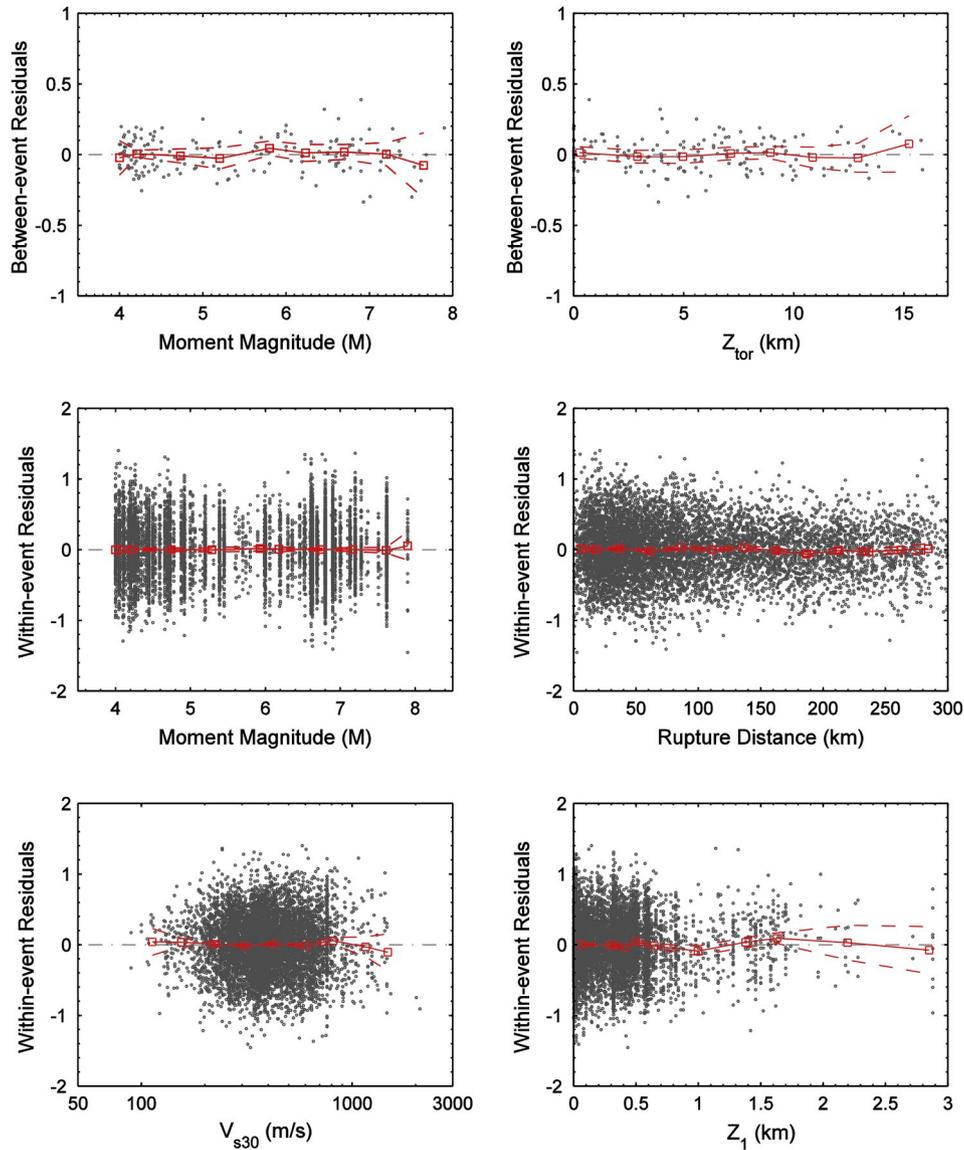


Fig. 8. Distributions of the between-event residuals with respect to  $M$ ,  $Z_{tor}$ , and within-event residuals with respect to  $M$ ,  $R_{rup}$ ,  $V_{s30}$ , and  $Z_1$  for the relative measure  $N_R$  (3.0).

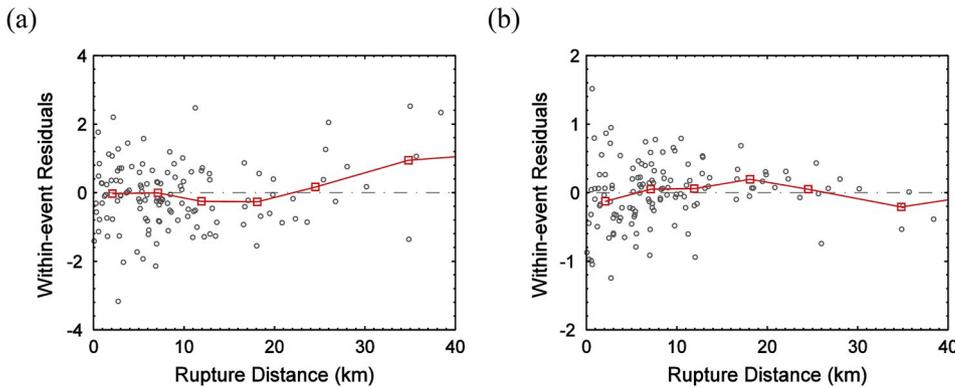


Fig. 9. Distribution of the within-event residuals of forward directivity recordings with respect to rupture distance  $R_{rup}$  for (a)  $N_A$  (2.0) and (b)  $N_R$  (2.0), respectively. The red square symbol denotes the local mean value of each binned residuals. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

existing models in this section. Figs. 12 and 13 display the predicted median and median  $\pm 1$  standard deviation curves with respect to distance scaling under three earthquake scenarios for  $N_A$  and  $N_R$ , respectively. The empirical data selected from the present database are also shown in each plot. Several observations can be made from these

plots. First, the predicted trends of  $N_A$  and  $N_R$  are generally in agreement with the empirical data, indicating that the functional forms adopted for scalings with magnitude and distance are appropriate. Second, as shown in Fig. 12c, the median predictions of  $N_A$  for the M 7.5 event seem to be slightly larger than the observed data. This is

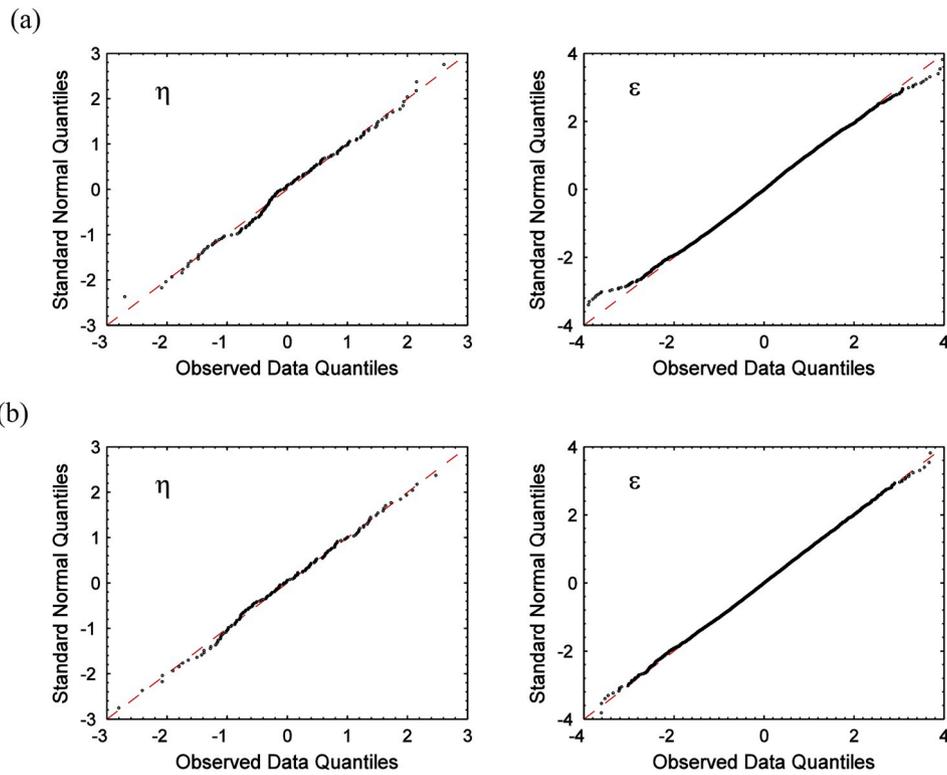


Fig. 10. Q-Q plots of the normalized between-event and within-event residuals for (a)  $N_A(2.0)$ , and (b)  $N_R(2.0)$ , respectively. Note:  $\eta$  and  $\epsilon$  represent the between-event residuals and within-event residuals, respectively.

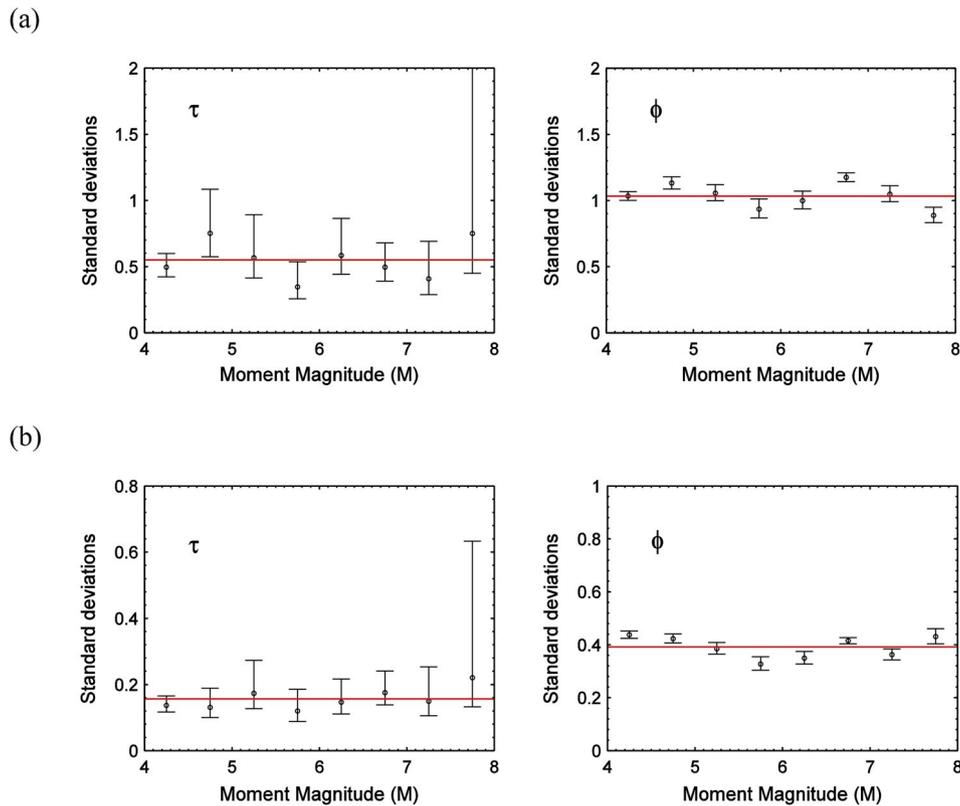
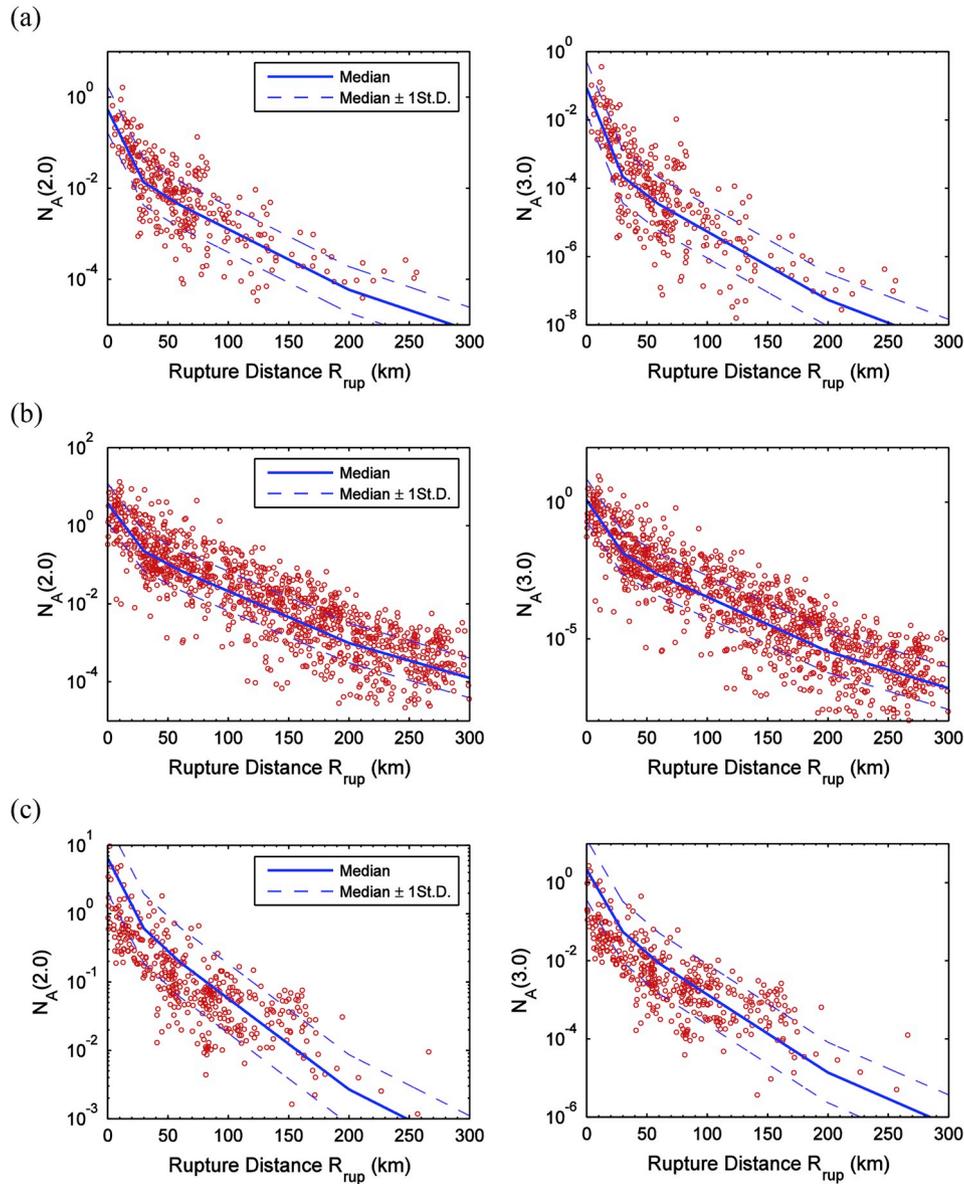


Fig. 11. Between-event and within-event binned standard deviations versus magnitude for (a)  $N_A(2.0)$  and (b)  $N_R(2.0)$ , respectively. At each plot, the points denote the calculated standard deviations for the binned residuals, and the vertical confidence intervals are estimated based on an assumed chi-squared distribution. Note:  $\tau$  and  $\phi$  denote the between-event and within-event standard deviations, respectively.



**Fig. 12.** Comparisons of the predicted  $N_A(2.0)$  and  $N_A(3.0)$  with the empirical data for three earthquake scenarios considered: (a)  $M = 5.5$ ,  $Z_{tor} = 6$  km; (b)  $M = 6.5$ ,  $Z_{tor} = 4$  km; (c)  $M = 7.5$ ,  $Z_{tor} = 0$  km. The  $V_{s30}$  value is set as 400 m/s, and  $\delta z_i$  is assigned as 0 for these scenarios considered. In these and subsequent plots, the empirical data are selected following a combination of  $[M-0.25, M+0.25]$  and  $[V_{s30}-200, V_{s30}+200]$ . For example, Fig. 12a displays the empirical data of  $N_A(2.0)$  and  $N_A(3.0)$  for  $5.25 < M < 5.75$  and  $200 < V_{s30} < 600$  m/s.

because the data within this magnitude range ( $7.25 < M < 7.75$ ) are dominated by recordings from the 1999 Chi-Chi earthquake, which is known to have produced smaller than expected ground motions compared to a similar earthquake event (e.g. Refs. [30,31]). Besides, as shown in Fig. 13b, the proposed  $N_R$  relationships are, to some extent, lower than the empirical data in the moderate-to-far distance range. This is mainly attributed to the  $R_{rup}$ -independent term (for  $R_{rup} > 50$  km) employed in the functional form (Eq. (15)).

The performance of the models is further compared with a prior study in the literature. The  $N_A(2.0)$  and  $N_R(2.0)$  models developed by Stafford and Bommer [13] (termed as SB09 hereafter) are enlisted. They are appropriate for such comparison because of the same cycle-counting definition considered. Note that only the SB09 model without the consideration of directivity effect (i.e., Eqs. (4) and (11) in Ref. [13]) is employed herein. The median predictions of the new models and the SB09 model under three earthquake scenarios are compared in Fig. 14. For  $N_A(2.0)$ , the predictive curves between the present study and the SB09 model are generally consistent. Yet, as is shown in Fig. 14b, the predictions for  $N_R$

(2.0) exhibit somewhat discrepancies, especially within the short distance range. Specifically, the SB09 model has a steep parabolic trend at short distances for  $M = 7.5$ . The predictive differences between the two models might be due to the different distance scaling terms employed; we choose a bilinear  $R_{rup}$  term, whereas a  $\cos(M)$ -dependent  $R_{rup}$  term is incorporated in the SB09 model.

### 5. Discussions

Quantifying the effective number of ground motion cycles is important in assessing liquefaction potential of soil deposits or damage degree of structures [32]. This paper therefore proposed new predictive models for four measures of the numbers of effective cycles, including two absolute ( $N_A(2.0)$ ,  $N_A(3.0)$ ) and two relative measures ( $N_R(2.0)$ ,  $N_R(3.0)$ ). They could be used for a variety of applications. In particular, the relative measures have some favorable features. First, the predictive models for  $N_R(2.0)$  and  $N_R(3.0)$  have significantly smaller standard deviations than those of the  $N_A$  models. A smaller standard deviation

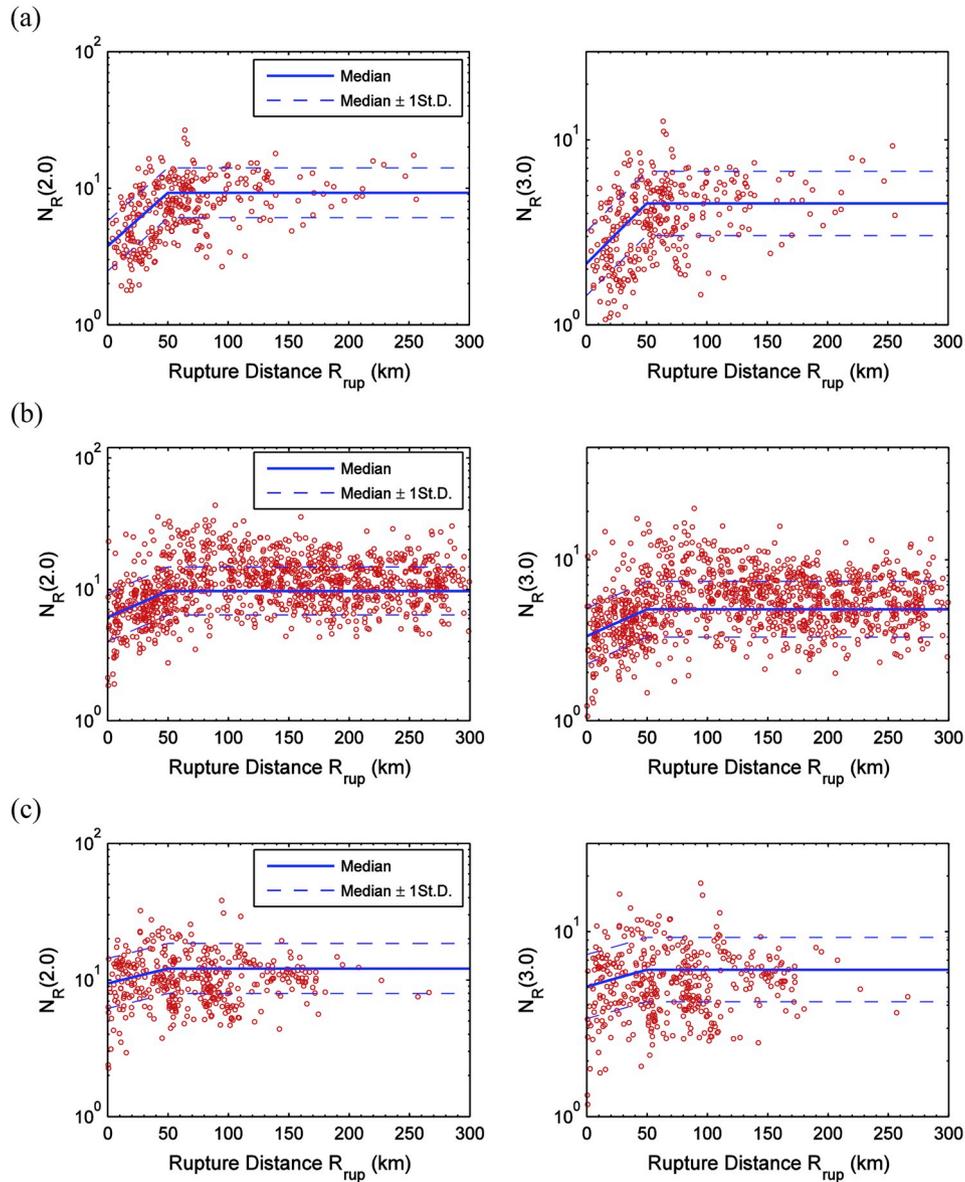


Fig. 13. Comparisons of the predicted  $N_R(2.0)$  and  $N_R(3.0)$  with the empirical data for three earthquake scenarios: (a)  $M = 5.5$ ,  $Z_{tor} = 6$  km; (b)  $M = 6.5$ ,  $Z_{tor} = 4$  km; (c)  $M = 7.5$ ,  $Z_{tor} = 0$  km. The  $V_{s30}$  value is set as 400 m/s for these scenarios considered.

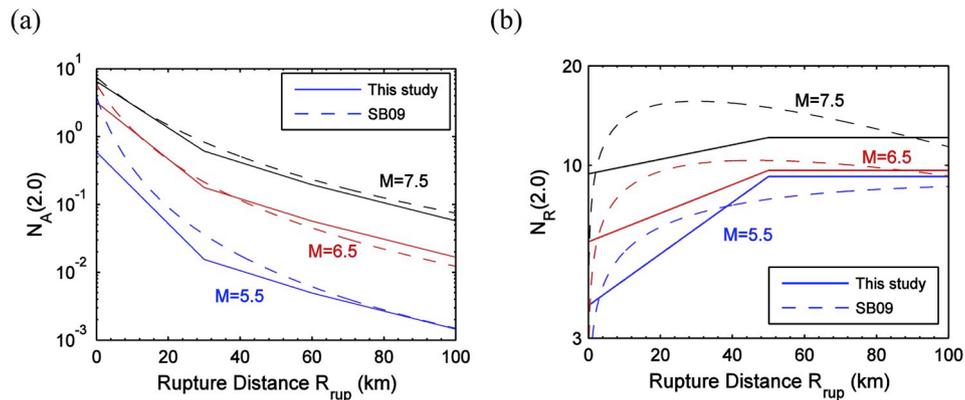


Fig. 14. Median predictions of the new models developed compared with the SB09 model for (a)  $N_A(2.0)$ , and (b)  $N_R(2.0)$ , respectively. In these plots, the number of cycles are predicted with input parameters assigned:  $V_{s30} = 400$  m/s,  $\delta z_1 = 0$ , and  $Z_{tor}$  as 6, 4, and 0 km for  $M = 5.5$ , 6.5, and 7.5, respectively. The near-fault directivity effect is not considered for both models.

indicates a higher level of predictability, which is preferable in engineering applications. Second, a recent study stated that  $N_R$  provides some supplementary information regarding the ground motion characteristics compared with commonly used IMs such as PGA and spectral accelerations [27]. Therefore,  $N_R$ , together with peak response parameters (e.g., PGA), can be regarded as a vector-IM with more complete earthquake information provided. Many studies have advocated the advantage of the use of a vector-IM in seismic hazard analysis (e.g. Refs. [33,34]). Thus, the vector-IM consisting of  $N_R$  (3.0) and PGA could be appropriate for evaluating the liquefaction potential of clean sands.

This paper focused on studying the effective numbers of ground motion cycles directly based on earthquake recordings. Yet, the “equivalent number of cycles” concept originally denotes the number of equivalent stress cycles (termed as  $N_{cyc}$ ) in soil masses subjected to a random shaking [35]. Recently some scholars have studied empirical relationships for  $N_{cyc}$  by performing site response analysis (e.g. Refs. [9,36]). For specific soil profiles, the magnitude of  $N_{cyc}$  is closely related to the occurrence of soil liquefaction. The exploration of predictive models for  $N_{cyc}$  will be a subject of future study.

## 6. Conclusions

This study presented new predictive models for the effective numbers of cycles of ground motions using 7447 earthquake recordings selected from the NGA-West2 database. Four measures of the effective cyclic numbers, including two absolute ( $N_A$  (2.0) and  $N_A$  (3.0)), with exponent coefficients of 2 and 3, respectively) and two relative measures ( $N_R$  (2.0) and  $N_R$  (3.0)), were considered and calculated using the rainflow range-counting approach. The models were developed based on the mixed-effects regression analysis. The functional forms predict the effective numbers of cycles as a function of moment magnitude ( $M$ ), rupture distance ( $R_{rup}$ ), site parameters ( $V_{s30}$  and sediment depth  $Z_1$ ), depth to the top of the fault rupture ( $Z_{tor}$ ), and forward directivity indicator ( $I_{dir}$ ). A quadratic  $M$ -scaling and a piece-wise bilinear  $M$  and  $R_{rup}$  terms are employed in the proposed model for  $N_R$ . Visual inspection of the between-event and within-event residuals indicates that they are generally unbiased versus the predictor variables employed. The prediction equations developed in this paper are applicable to shallow crustal earthquakes with  $M$  ranging from 4 to 7.9,  $V_{s30}$  in the range of 100 to 2100 m/s, and rupture distance up to 300 km.

The performance of the proposed models is validated and compared with an existing study. The new models could be regarded as an improvement over existing models owing to the use of the expanded database. The relative measures demonstrate an advantage over the absolute measures due to the much smaller standard deviations (i.e., higher predictability) reported. Therefore, the relative measures  $N_R$ , along with peak amplitude IMs (e.g., PGA), can constitute a meaningful vector-IM that could be applied in the assessment of liquefaction potential in engineering applications.

## Acknowledgments

The authors acknowledge support from General Research Fund (Grant No. 16214118, 16213615) from Hong Kong Research Grants Council and Joint Research Fund for Overseas Chinese Scholars and Scholars in Hong Kong and Macao (Grant No. 51828902) from National Natural Science Foundation of China. The authors greatly thank two anonymous reviewers for their helpful comments to improve this manuscript.

## References

[1] Kramer SL. Geotechnical earthquake engineering. New Jersey: Practice Hall; 1996. p. 653.

[2] Seed HB, Lee KL. Liquefaction of saturated sands during cyclic loading. *J Soil Mech Found Div* 1966;92(6):105–34.

[3] Martin GR, Finn WL, Seed HB. Fundamentals of liquefaction under cyclic loading. *J Geotech Geoenviron Eng* 1975;101. (ASCE# 11231 Proceeding).

[4] Polito CP, Green RA, Lee J. Pore pressure generation models for sands and silty soils subjected to cyclic loading. *J Geotech Geoenviron Eng* 2008;134(10):1490–500.

[5] Cetin KO, Seed RB, Der Kiureghian A, Tokimatsu K, Harder Jr. LF, Kayen RE, Moss RE. Standard penetration test-based probabilistic and deterministic assessment of seismic soil liquefaction potential. *J Geotech Geoenviron Eng* 2004;130(12):1314–40.

[6] Idriss IM, Boulanger RW. Semi-empirical procedures for evaluating liquefaction potential during earthquakes. *Soil Dynam Earthq Eng* 2006;26:115–30.

[7] Onder Cetin K, Tolga Bilge H. Performance-based assessment of magnitude (duration) scaling factors. *J Geotech Geoenviron Eng* 2012;138(3):324–34.

[8] Annaki M, Lee KL. Equivalent uniform cycle concept for soil dynamics. *J Geotech Eng Div* 1977;103(6):549–64.

[9] Green RA, Terri GA. Number of equivalent cycles concept for liquefaction evaluations—Revisited. *J Geotech Geoenviron Eng* 2005;131(4):477–88.

[10] Hancock J, Bommer JJ. The effective number of cycles of earthquake ground motion. *Earthq Eng Struct Dyn* 2005;34(6):637–64.

[11] Seed HB, Idriss IM. Simplified procedure for evaluating soil liquefaction potential. *J Soil Mech Found Div* 1971;97(SM9):1249–73.

[12] Liu AH, Stewart JP, Abrahamson NA, Moriawaki Y. Equivalent number of uniform stress cycles for soil liquefaction analysis. *J Geotech Geoenviron Eng* 2001;127(12):1017–26.

[13] Stafford PJ, Bommer JJ. Empirical equations for the prediction of the equivalent number of cycles of earthquake ground motion. *Soil Dynam Earthq Eng* 2009;29(11):1425–36.

[14] Pinheiro J, Bates D, DebRoy S, Sarkar D. R Core Team. NLME: linear and nonlinear mixed effects models. R package version 2014;3:1–117.

[15] Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BSJ, et al. NGA-West2 database. *Earthq Spectra* 2014;30(3):989–1005.

[16] Campbell KW, Bozorgnia Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthq Spectra* 2014;30(3):1087–115.

[17] Afshari K, Stewart JP. Physically parameterized prediction equations for significant duration in active crustal regions. *Earthq Spectra* 2016;32(4):2057–81.

[18] Du W, Wang G. Prediction equations for ground-motion significant durations using the NGA-West2 database. *Bull Seismol Soc Am* 2017;107(1):319–33.

[19] 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. *Seismol Res Lett* 1997;68(1):199–222.

[20] Shahi SK. A probabilistic framework to include the effects of near-fault directivity in seismic hazard assessment PhD. dissertation USA: Stanford University; 2013. p. 226.

[21] ASTM. Cycle counting in fatigue analysis. Annu Book ASTM (Am Soc Test Mater) Stand 1985;03.01. Designation E1049-85.

[22] Amzallag C, Gerey JP, Robert JL, Bahauud J. Standardization of the rainflow counting method for fatigue analysis. *Int J Fatigue* 1994;16(4):287–93.

[23] Idriss IM, Boulanger RW. Soil liquefaction during earthquakes. *Earthquake Engineering Research Institute*; 2008. MNO-12.

[24] Boore DM, Stewart JP, Seyhan E, Atkinson GM. NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthq Spectra* 2014;30(3):1057–85.

[25] Chiou BSJ, Youngs RR. Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. *Earthq Spectra* 2014;30(3):1117–53.

[26] Du W. An empirical model for the mean period ( $T_m$ ) of ground motions using the NGA-West2 database. *Bull Earthq Eng* 2017;15(7):2673–93.

[27] Du W, Wang G. Empirical correlations between the effective number of cycles and other intensity measures of ground motions. *Soil Dynam Earthq Eng* 2017;102:65–74.

[28] Boore DM. Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. *Bull Seismol Soc Am* 1983;73:1865–94.

[29] Johnson RA, Wichern DW. Applied multivariate statistical analysis. Upper Saddle River, New Jersey: Prentice Hall; 2007.

[30] Boore DM. Comparisons of ground motions from the 1999 Chi-Chi earthquake with empirical predictions largely based on data from California. *Bull Seismol Soc Am* 2001;91(5):1212–7.

[31] Travarasou T, Bray JD, Abrahamson NA. Empirical attenuation relationship for Arias intensity. *Earthq Eng Struct Dyn* 2003;32(7):1133–55.

[32] Ning CL, Du W, Li B. Biaxial hysteretic model for the characterization of quasi-static testing of shear-critical reinforced concrete columns. *Adv Struct Eng* 2018. <https://doi.org/10.1177/1369433218788369>. published online.

[33] Bazzurro P, Cornell CA. Vector-valued probabilistic seismic hazard analysis (VPSHA). Proceedings of the 7th US national conference on earthquake engineering, Boston, Massachusetts. 2002. p. 21–5.

[34] Du W, Wang G. Fully probabilistic seismic displacement analysis of spatially distributed slopes using spatially correlated vector intensity measures. *Earthq Eng Struct Dyn* 2014;43(5):661–79.

[35] Seed HB, Idriss IM, Makdisi F, Banerjee. Representation of irregular stress time histories by equivalent uniform stress series in liquefaction analyses. EERC 75-29, earthquake engineering Research center. Berkeley, CA: Univ. of California; 1975.

[36] Kishida T, Tsai CC. Seismic demand of the liquefaction potential with equivalent number of cycles for probabilistic seismic hazard analysis. *J Geotech Geoenviron Eng* 2014;140(3):04013023.