Empirical correlations between the effective number of cycles and other intensity measures of ground motions

Wenqi Du⁎, Gang Wang

⁎ Institute of Catastrophe Risk Management, Nanyang Technological University, 50 Nanyang Avenue Singapore
⁎⁎ Dept. of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon Hong Kong

A R T I C L E   I N F O

Keywords:
Effective number of cycles
Correlation coefficient
NGA database
Ground motion selection
Liquefaction potential

A B S T R A C T

The effective number of cycles is an important ground motion parameter for the assessment of liquefaction potential. In this paper, empirical correlations for two measures of the effective number of cycles with seven amplitude-, cumulative-, and duration-based intensity measures (IMs) are studied and compared, based on the NGA strong motion database and several ground motion prediction equations. The adopted definitions of the effective number of cycles include an absolute measure (NeA) and a relative measure (NeR). It is shown that NeA is highly correlated with high-frequency IMs, such as spectral acceleration (SA) at short periods, Arias intensity, and negatively correlated with signification durations (Ds). On the other hand, NeR shows generally negative correlations with both amplitude- and cumulative-based IMs. NeR also exhibits small-to-moderate positive correlations with Ds, which are commonly regarded as similar parameters to the effect number of cycles. Simple parametric functions are provided to describe the NeA-SA and NeR-SA correlations for various cases. The importance of considering multiple IMs rather than SA only in ground-motion selection is also briefly demonstrated.

1. Introduction

The number of cycles of ground motions has been widely recognized as one of the critical parameters in geotechnical earthquake engineering. Many studies (e.g., [1,2]) have concluded that the number of cycles of shaking is strongly correlated with the buildup of pore water pressure in liquefiable soils. As summarized by Hancock and Bommer [3], there are dozens of definitions to count the effective number of cycles, by converting all irregular amplitude cycles to an equivalent number of uniform cycles. The concept of equivalent number of cycles is commonly used for evaluating liquefaction potential [4–6].

Due to the complex features of ground motion time histories, single ground motion intensity measure (IM) cannot adequately characterize earthquake loadings. Therefore, a set of IMs (vector-IMs) is often required in some practical applications, such as the estimation of earthquake-induced slope displacement [7,8]. Since current ground motion prediction equations (GMPEs) only provide the means and standard deviations for specific IMs, empirical correlations among the residuals of these IMs are then the key requirement to contrast the joint distribution of vector-IMs. These empirical correlations are indispensable in vector-based probabilistic seismic hazard analysis [9] and scenario-based ground motion selection approaches, e.g., [10–12].

Recently a number of researchers have studied empirical correlations between the residuals of multiple IMs, such as spectral accelerations (SA) at multiple periods, Arias intensity (Ia), cumulative absolute velocity (CAV), and significant durations (Ds), e.g., [13–18]. However, to the knowledge of the authors, there are no existing correlation models involving the number of effective cycles. Bommer et al. [19] has studied the correlations between several duration parameters and effective numbers of ground motion cycles. Yet, they did not aim at evaluating the correlations between the residuals of these IMs, making it difficult to be used in some applications such as ground motion selection.

The objective of this paper is to examine the empirical correlations between the effective number of cycles and other commonly used IMs. The definitions of these employed IMs are firstly discussed, associated with the utilized GMPEs and ground motion database. Secondly, the estimated correlation coefficients between the residuals of these IMs are presented; simple parametric models are also proposed to readily predict the empirical correlations. The influence of rupture distance (Rrup) on the resulting correlation coefficients is then examined. Finally, based on the correlation results, some recommendations are provided regarding the use of different definitions of the effective number of cycles for practical applications.
2. Selected IMs and ground motion database

2.1. Effective number of cycles

As summarized in Hancock and Bommer [3], there are many cycle-counting definitions in the literature, which can be mainly classified into several categories: peak counting, level crossing counting, range counting, and indirect counting methods. These cycle-counting definitions were developed for low-cycle fatigue testing [20]. Among these definitions, the rainfall range-counting method is the most popular since it quantifies both the high-frequency and low-frequency cyclic waves in broadband signals. This method counts a history of peaks and troughs in sequence which can be regarded as starting and ending points for defining each cycle. The algorithm can be simplified as: (i), the signal is turned clockwise as 90°; (ii), an imagined source of water will flow down the “pagoda roofs” from their upper tops; (iii), the water will drip down when it reaches the edge. It will stop when it comes to a point that is already wet (quantified by previous flow), or it reaches opposite beyond the vertical of the starting point; (iv), the steps (ii)-(iii) can be repeated to get a series of half-cycles. The detailed algorithm of this approach can be found in References [3,21]. Fig. 1 shows a simple example about the application of the rainfall-counting technique. Total five half-cycles and one full-cycle are identified for this wave. Besides, prediction equations for the effective number of cycles based on the rainfall-counting approach have been proposed [22], which can be directly used to account for the statistical distributions of these IMs.

Similar to the cyclic damage parameter for low-cycle fatigue failure used by Malhotra [23], the absolute definition of the effective number of cycles can be expressed as:

\[ N_e = \sum_{i=1}^{2} u_i^2 \]  

(1)

where \( u_i \) is the amplitude of the \( i \)-th half cycle obtained by the rainfall range-counting method; \( T_n \) is the total number of cycles; and \( N_e \) is the absolute measure of the effective number of cycles. It is noted that the exponent coefficient is set as 2 herein, which reflects the relative importance of different amplitude cycles. A higher value of the exponent coefficient represents a larger contribution caused by large-amplitude cycles.

Relative definitions of the effective number of cycles are commonly used in earthquake engineering. A typical relative definition of the number of cycles, in which each amplitude \( u_i \) is normalized by the maximum amplitude of all half-cycles, \( u_{max} \), is expressed as:

\[ N_R = \frac{1}{2} \sum_{i=1}^{2} \left( \frac{u_i}{u_{max}} \right)^2 \]  

(2)

where \( N_R \) is the relative measure of the effective cycles. A value of 2 is also adopted for the exponent coefficient.

It is worth noting that only the effective number of cycles obtained by the rainfall-counting method is considered in this paper, due to its popularity and robustness. The selected measures of cyclic numbers can be applied in most practical cases. The aforementioned empirical equations proposed by Stafford and Bommer [22], which are termed as SB09 model hereafter, will be used to predict \( N_e \) and \( N_R \) in the following section. The SB09 model utilized a subset of the Pacific Earthquake Engineering Research (PEER) NGA-West1 database [24], employing moment magnitude \( M_w \), rupture distance \( R_{rup} \), site parameters and the depth to the top of rupture (Ztor) as indicators. A set of equations has been proposed by Stafford and Bommer [22], while only the basic equations without the consideration of Ztor or directivity effect are used in this study. The median predictions of the SB09 model for \( N_e \) and \( N_R \) with respect to \( M_w \) and \( R_{rup} \) are shown in Fig. 2. It should be noted that the other few prediction equations using different counting methods, e.g., [25], are not considered due to the scope of this paper.

2.2. Other IMs considered

The other IMs considered herein are listed as: (a) peak values of ground motion time histories, including peak ground acceleration (PGA) and peak ground velocity (PGV); (b) SA at multiple periods; (c) cumulative-based intensity measures, including Ia and CAV; and (d) ground motion duration parameters, including significant durations defined as time intervals over which 5–75% and 5–95% of Ia are built.

Fig. 1. A demonstrated example of the use of rainfall-counting approach. This segment consists of several sine waves, which can be counted as five half cycles (1–2; 2–3; 3–4; 4–5; 5–6 and 6–7) and one full cycle (4–4′). The amplitudes of the five half cycles are 0.1, 0.2, 0.15, 0.1, and 0.05 g, respectively; the amplitude of the full cycle is 0.05 g. The segment yields values of \( N_e \) and \( N_R \) at 0.09 and 1.125, respectively.

Fig. 2. Median predictions of the SB09 model for the effective number of cycles \( N_e \) and \( N_R \), respectively. The \( V_{S30} \) value is set as 400 m/s for predicting \( N_R \).
It should be noted that each recorded time history has a usable period range, in order to eliminate low-frequency or high-frequency noises. Therefore, SA at periods larger than the maximum usable period should not be used for subsequent correlation analyses. The number of usable ground motions is expected to decrease as vibration period increases, as is shown in Fig. 3(b).

3. Empirical correlation analyses

3.1. Computational procedures for correlation coefficients

Current GMPEs usually assume that IMs are logarithmically normally distributed, which can be shown as:

$$\ln(IM_i) = \ln(IM_{\text{med}}) + \eta_i + \epsilon_i$$

where $\ln(IM_i)$ and $\ln(IM_{\text{med}})$ denote the measured (geometric mean of two horizontal components of each record) and the predicted logarithmic $i$th IM (e.g., SA, Ia, $N_A$), respectively. $\eta_i$ and $\epsilon_i$ represent the inter-event and intra-event residuals of the $i$th IM (normally distributed with zeros mean and standard deviations $\sigma_i$ and $\sigma_\epsilon$, respectively). The total standard deviation $\sigma_T$ is given as $\sigma_T = \sqrt{\sigma_i^2 + \sigma_\epsilon^2}$. The values of $\sigma_i$ and $\sigma_\epsilon$ for various IMs are generally provided by GMPEs.

The Pearson product-moment correlation coefficient is a widely used measure of linear correlation between two variables [39]. The correlation coefficients between the inter-event or intra-event residuals of different IMs can be estimated as:

$$\rho_{i_1,i_2} = \frac{\sum_{m=1}^{n}(x_{i_1}^{(m)} - \bar{x}_{i_1})(x_{i_2}^{(m)} - \bar{x}_{i_2})}{\left(\sum_{m=1}^{n}(x_{i_1}^{(m)} - \bar{x}_{i_1})^2 \sum_{m=1}^{n}(x_{i_2}^{(m)} - \bar{x}_{i_2})^2\right)^{1/2}}$$

where $x_1$ and $x_2$ are random variables (e.g., $\eta_1$ and $\eta_2$ for the inter-event residuals of IM1 and IM2); $n$ is the total number of the random variables considered (i.e., number of earthquakes for the inter-event correlation, or number of ground motion records for the intra-event correlation); and $\bar{x}_1$ and $\bar{x}_2$ denote the sample mean of variables $x_1$ and $x_2$, respectively. In this paper, IM1 refers to $N_A$ or $N_B$, and IM2 refers to the other aforementioned measures such as PGA, SA, and Ia. For each pair of IMs, $\rho_{i_1,i_2}$ can be computed via Eq. (4).

Under the assumption that $\eta_1$ and $\eta_2$ are independent [40], the correlation between the total residuals can be expressed as:

$$\rho_{T_1,T_2} = \frac{\rho_{i_1,i_2}}{\sqrt{\rho_{\eta_1,\eta_1} \rho_{\epsilon_1,\epsilon_1} + \rho_{i_1,i_2} \rho_{\eta_2,\epsilon_2}}}$$

where $\rho_{\epsilon_1,\epsilon_2}$ and $\rho_{\eta_1,\eta_2}$ ($k = 1, 2$) are the standard deviations of the inter-event, intra-event, and total residuals for the $k$-th IM, respectively. Thus, the correlation between the total residuals for each pair of IMs can be calculated via Eqs. (4) and (5) accordingly.

The above statistical analysis only provides the point-estimate of the correlation coefficient, while the uncertainty of $\rho$ should also be accounted for carefully. Such uncertainty is due to the finite number of sample size, as well as different ground motion models used in its determination. A bootstrap method is often used to construct the confidence intervals of correlation coefficients [39]. The basic idea of this method is to re-sample the observed dataset by random sampling with replacement from the original dataset, and then the correlation coefficients of these bootstrap replicates can be calculated. This process needs to be repeated a certain number of times to accurately estimate the variance of $\rho$.

In addition to the bootstrap method, another widely used method is the Fisher $z$ transformation [41]. This method converts the correlation coefficient $\rho$ into a transformed variable $z$ via:

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

where $\rho$ is the Pearson correlation coefficient; $\tanh^{-1}$ is the inverse hyperbolic tangent function.
hyperbolic tangent function; and \( z \) is the transformed correlation coefficient which is approximately normally distributed. The variance of \( \rho \) usually becomes smaller when it approaches 1 or \(-1\), whereas the variance of the transformed variable \( z \) can keep approximately constant for all \( \rho \) values. Therefore, if the mean and standard deviation of the transformed variable \( z \) are denoted as \( \mu_z \) and \( \sigma_z \), respectively, the corresponding median correlation coefficient \( \rho_{50} \) can be computed as:

\[
\rho_{50} = \frac{e^{\mu_z} - 1}{e^{\mu_z} + 1} = \tanh(\mu_z)
\]

Similarly, a certain percentile of \( \rho \) can be obtained using \( \mu_z \) and \( \sigma_z \) (e.g., \( \rho_{44} = \tanh(\mu_z + \sigma_z) \); \( \rho_{16} = \tanh(\mu_z - \sigma_z) \)).

### 3.2. Correlations of cyclic numbers with PGA, PGV, and SA

Fig. 4 shows the distributions of the computed correlation coefficients of \( N_A \), \( N_R \) with PGA and PGV, respectively. For each pair of the IMs, the correlation coefficient \( \rho \) between the total residuals was first estimated via Eqs. (4) and (5); 1500 bootstrap replicates from the original dataset were then generated for quantifying the uncertainty of \( \rho \). Boxplots are used to show the uncertainty caused by the finite sample size (obtained by the bootstrap method); the results obtained by the

![Fig. 3. (a) Distribution of earthquake recordings with respect to moment magnitude and rupture distance used in this study; (b) the number of usable records to compute the correlation of \( N_A \) and \( N_R \) with SA at different periods.](image)

![Fig. 4. Computed correlation coefficients between (a) \( N_A \) and PGA, (b) \( N_A \) and PGV, (c) \( N_R \) and PGA, (d) \( N_R \) and PGV. In these and subsequent boxplots, the central red line denotes the median of the data (50th percentile), and the edges of the box (blue lines) mark the 25th and 75th percentiles. The ends of the whiskers represent the 0.35th and 99.65th percentiles, respectively, the red plus symbols denote outliers. The results obtained by various GMPEs are shown in each subplot. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image)
four aforementioned GMPEs are shown in each boxplot. Fig. 4(a) and (b) illustrate that the correlations of $N_A$ with PGA and PGV are positive with median values of approximately 0.78 and 0.4, respectively. This is not surprising, since $N_A$ represents the summation of the amplitude for each half cycle, which is expected to be larger if PGA of a ground motion is higher. Fig. 4(c) and (d) show the computed correlations of $N_A$ with PGA and PGV, respectively. Unlike the case of $N_A$, the correlations of $N_{SA}$ with PGA and $N_{SA}$ with PGV are generally negative, with median values of approximately $-0.36$ and $-0.46$, respectively. Hence, as a relative definition, $N_R$ has a much weaker correlation with PGA compared with $N_A$ because of the normalization process. For PGA, $N_R$ exhibits slightly stronger correlation than $N_A$, although the $N_{SA}$-PGV correlation is found to be negative. Besides, it can also be seen that the uncertainty (i.e., 25th and 75th percentiles represented by the edges of the box) due to the finite sample size is generally in the range of 0.02–0.03, whereas the variability from the use of different GMPEs is more notable. There appears to be no specific GMPE which yields systematically higher or lower correlations. The results (a number of $\rho$ based on the bootstrap method) obtained by various GMPEs were combined together in a logic-tree framework with equal weight assigned. The combined results were then transformed via Eq. (6) to evaluate $\mu_\rho$ and $\sigma_\rho$. The calculated median correlation coefficients $\rho_{\text{PGA}}$ as well as $\sigma_\rho$ are listed in Table 2.

Fig. 5 shows the computed empirical correlations [Eq. (5)] between $N_A$, $N_R$ and $SA$ versus vibration period. Similar to the results shown in Fig. 4, the correlations of $N_A$-SA are generally positive, while the $N_{SA}$-SA correlations are negative. As seen in Fig. 5(a), the correlation of $N_A$-SA is generally constant for $T < 0.2$ s, and it shows a decreasing trend as period $T$ increases for $T > 0.2$ s. This implies that $N_A$ is mainly dependent on the high-frequency content of ground motion. On the other hand, the correlation of $N_{SA}$-SA is in the range of $-0.4$ to $-0.1$. It generally increases with increasing vibration period for $T < 0.1$ s and $T > 1$ s, whereas a noticeable decreasing trend of $\rho$ can be observed for periods from 0.1 s to 1 s. The variation of the $N_{SA}$-SA correlation versus period is presently not feasible to explain. Yet, the overall small correlation indicates that $N_A$ can represent additional ground motion characteristics compared to SA.

Based on the aforementioned bootstrap and $z$ transformation approaches, the computed median, 16th and 84th percentiles (blue dashed lines) of the correlation coefficients for $N_{SA}$ and $N_{SA}$-SA are illustrated in Fig. 6(a) and (b), respectively. A smooth parametric function is usually desirable for practical use. The following piecewise linear function is then proposed to fit the median $N_{SA}$ and $N_{SA}$-SA correlations:

$$\rho_{N_{SA}} = a_\rho + \frac{\ln(T/T_0)}{\ln(T_{n+1}/T_0)}(a_{n+1} - a_n) \quad T_0 \leq T \leq T_{n+1}$$

where $N_A$ denotes $N_A$ or $N_R$ in this paper; $a_\rho$, $T_0$ and $T_n$ are regressed parameters in order to capture the overall shape of the correlation coefficients over the whole period range. The values of these parameters are listed in Table 3. For any given $T$, $\rho$ can be predicted by linear interpolation in logarithmic period scale. The solid lines in Fig. 6 are obtained by the proposed parametric function (Eq. (8)), and they compare reasonably well with the empirical data. Besides, it can also be observed that the variations of $\rho$ for the $N_{SA}$-SA and $N_{SA}$-SA correlations versus period $T$ are relatively small. The computed values of $\sigma_\rho$ are in the ranges of 0.03–0.07 and 0.03–0.05 for $N_{SA}$-SA and $N_{SA}$ correlations, respectively. Therefore, it seems not necessary to accurately capture the small variations of $\sigma_\rho$ with periods, and the $\sigma_\rho$ values of 0.05 and 0.04 can be directly used for $N_{SA}$-SA and $N_{SA}$-SA over the whole period range. The predicted 16th and 84th percentiles using the constant $\sigma_\rho$ values are also shown (black dotted line) in Fig. 6.

### 3.3. Correlation of cyclic numbers with $I_a$ and CAV

Fig. 7(a) and (b) display the computed correlations (e.g., the median, 25th, and 75th percentiles) of $N_A$ with $I_a$ and CAV, respectively. It is not surprising that the $N_A$-$I_a$ and $N_A$-CAV correlations are generally moderate, given that both $I_a$ and CAV, as measures of the cumulative intensity of shaking, are highly dependent on the absolute amplitudes of ground motion cycles. The correlation between $N_A$ and $I_a$ is slightly larger than that of $N_A$-CAV. This is possibly due to the fact that both $N_A$ and $I_a$ employ 2 as the exponent coefficient in their definitions. Besides, noticeable differences among the correlations by various GMPEs can be observed for both figures. Such uncertainty could be quantified using the aforementioned bootstrap and $z$ transformation methods.

Fig. 7(c) and (d) display the computed correlations of $N_R$ with $I_a$ and CAV, respectively. As observed previously, there are slight variations in the correlations obtained by using different GMPEs. The median correlation coefficients of $N_R$ with $I_a$ and CAV are approximately $-0.19$ and $-0.04$, respectively. The rather poor correlations imply that, the cumulative-based IMs ($I_a$ and CAV) are relatively independent of the relative definition of the effective cyclic number ($N_R$). It should be noted that the median correlation coefficients $p_{\text{PGA}}$, and the standard deviations $\sigma_\rho$ between $N_A$, $N_R$, PGA, PGV, $I_a$, CAV, and subsequent $D_{5-75}$ and $D_{5-95}$, are summarized in Table 2.

### 3.4. Correlation of cyclic numbers with $D_{5-75}$ and $D_{5-95}$

Fig. 8(a) and (b) show the calculated correlation coefficients of $N_A$ with $D_{5-75}$ and $D_{5-95}$, respectively. It can be seen that the $N_A$-$D_{5-75}$ and $N_A$-$D_{5-95}$ correlations are generally similar, with median values in the range of $-0.35$ to $-0.2$, respectively. The negative correlations imply that a ground motion with a longer than expected $D_{5-75}$ would possibly have a smaller than expected $N_A$ value. Fig. 8(c) and (d) illustrate the calculated correlation coefficients of $N_R$ with $D_{5-75}$ and $D_{5-95}$, respectively. Both the $N_R$-$D_{5-75}$ and $N_R$-$D_{5-95}$ correlations are some degree of positive, while the correlation between $N_A$ and $D_{5-75}$ is slightly larger.

The observed correlations between $N_A$, $N_R$ and $D_{5-75}$, $D_{5-95}$ are consistent with physical intuitions. As discussed previously, $N_A$ is highly correlated with PGA, and it has been studied that $D_{5-75}$ and $D_{5-95}$ are negatively correlated with peak amplitudes (PGA, PGV) of ground motions [14]. Therefore, it is not surprising that the $N_A$-$D_{5-75}$ and $N_A$-$D_{5-95}$ correlations are negative. On the other hand, the normalized $N_R$ shows some degree of positive correlations with significant durations. Since $D_{5-75}$ and $D_{5-95}$ represent the time interval across which a great amount of seismic energy is dissipated, a ground motion with longer $D_{5-75}$ or $D_{5-95}$ is likely to be more numbers of cyclic waves (and hence larger $N_R$). Besides, Bommer et al. [19] also studied the correlations between ground motion durations and the effective cyclic numbers. They computed the correlations directly based on the measured values of the IMs, without the consideration of GMPEs and residuals of these IMs. Therefore, the results presented in this study are different with those provided in Reference [19].

Fig. 9(a) and (b) show the distribution of the inter-event and intra-event residuals between $N_A$ and $N_R$, respectively. The data points of these two figures are almost randomly distributed, with the computed correlation coefficients as $+0.15$ and $-0.04$, respectively. The correlation coefficient between the total residuals using Eq. (5) is calculated

### Table 2

<table>
<thead>
<tr>
<th>IM</th>
<th>PGA</th>
<th>PGV</th>
<th>$I_a$</th>
<th>CAV</th>
<th>$D_{5-75}$</th>
<th>$D_{5-95}$</th>
<th>$N_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_A$</td>
<td>0.78</td>
<td>0.40</td>
<td>0.79</td>
<td>0.64</td>
<td>$-0.27$</td>
<td>$-0.29$</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.043)</td>
<td>(0.15)</td>
<td>(0.061)</td>
<td>(0.045)</td>
<td>(0.035)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>$N_R$</td>
<td>$-0.35$</td>
<td>$-0.46$</td>
<td>$-0.19$</td>
<td>$-0.04$</td>
<td>0.51</td>
<td>0.37</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.057)</td>
<td>(0.032)</td>
<td>(0.029)</td>
<td>(0.056)</td>
<td>(0.064)</td>
<td>(1.00)</td>
</tr>
</tbody>
</table>
as 0.03. Such poor correlation indicates that, although both \( N_A \) and \( N_R \) represent the effective numbers of cycles of a ground motion, the physical interpretations of these two measures are significantly different. Therefore, \( N_A \) and \( N_R \) are better suited for different applications.

### 4. Influence of rupture distance on \( N_A \)-SA and \( N_R \)-SA correlations

It is tempting to examine the influence of causal parameters (i.e., \( M_w, R_{rup} \)) on the \( N_A \)-SA and \( N_R \)-SA correlations. The influence of magnitude \( M_w \) on the correlations is found to be less significant than \( R_{rup} \). Therefore, only the influence of rupture distance on the correlations is investigated in this section.

The empirical residuals are divided into three distance bins, namely \( R_{rup} = 0-30 \text{ km}, 30-80 \text{ km}, \) and \( 80-200 \text{ km} \), respectively. The number of usable records in each distance bin versus spectral period is shown in Fig. 10, from which it can be seen that each distance bin includes an adequate number of data points in order to yield statistically stable results. The procedures introduced in Section 3.1 were used for the correlation calculations based on the binned empirical data. Fig. 11 demonstrates the resulting \( N_A \)-SA and \( N_R \)-SA correlations for the three distance bins. For the \( N_A \)-SA correlations, it can be seen that the far-distance (\( R_{rup} = 80-200 \text{ km} \)) records exhibit the strongest correlations over a wide period range; the correlations for moderate-distance (\( R_{rup} = 30-80 \text{ km} \)) records are noticeably larger than those of short-distance records (\( R_{rup} = 0-30 \text{ km} \)) at periods larger than 2 s. For the \( N_R \)-SA correlations, the moderate-distance ground motions exhibit the strongest correlations for \( T < 1 \text{ s} \), while the far-distance ground motions yield the strongest correlations for \( T > 1 \text{ s} \). Such differences can be attributed to the different ground motion characteristics caused by travelling distances. It has been widely studied that long travelling distance tends to filter out high-frequency components of seismic waves, resulting in (far-distance) ground motions consisting of mainly moderate-to-long period seismic waves.

For practical usage, piecewise linear functions [Eq. (8)] are also developed to fit the correlation data for each distance bin. The regressed \( a_n \) and \( T_n \) parameters for the distance-binned \( N_A \)-SA and \( N_R \)-SA correlations are summarized in Tables 4, 5, respectively. As illustrated in Fig. 11, the proposed piecewise linear curves approximate the empirical data reasonably well. Besides, as listed in Tables 4, 5, a constant \( \sigma \) value is also assigned for each distance bin to quantify the uncertainty of \( N_A \)-SA and \( N_R \)-SA correlations. Compared to the \( \sigma \) values obtained based on the whole ground motion database, the distance-binned \( \sigma \) values are slightly larger, due to the decrease of the available number of data points in each distance bin.

![Fig. 5. Computed correlation coefficients [Eq. (5)] for (a) \( N_A \) and SA; (b) \( N_R \) and SA, respectively.](image)

![Fig. 6. Comparisons of the empirical correlations and the piece-wise linear fitting curves for (a) \( N_A \) and SA, (b) \( N_R \) and SA, respectively.](image)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( N_A )-SA (T)</th>
<th>( N_R )-SA (T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_n )</td>
<td>( T_n )</td>
<td>( a_n )</td>
</tr>
<tr>
<td>1</td>
<td>0.79</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.80</td>
<td>0.15</td>
</tr>
<tr>
<td>3</td>
<td>0.25</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>0.15</td>
<td>2.0</td>
</tr>
<tr>
<td>5</td>
<td>0.02</td>
<td>10.0</td>
</tr>
<tr>
<td>6</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>7</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

Table 3: Coefficients for predicting the \( N_A \)-SA and \( N_R \)-SA correlations.
Fig. 7. Correlation coefficients between (a) $N_a$ and $I_a$, (b) $N_a$ and CAV, (c) $N_R$ and $I_a$, (d) $N_R$ and CAV. The results obtained by various GMPEs are shown in each subplot.

Fig. 8. Correlation coefficients between (a) $N_a$ and $D_{S5-75}$, (b) $N_a$ and $D_{S5-95}$, (c) $N_R$ and $D_{S5-75}$, (d) $N_R$ and $D_{S5-95}$. The results obtained by various GMPEs are shown in each subplot.
5. Discussions

The effective number of cycles is an important measure in geotechnical earthquake engineering, especially in the evaluation of liquefaction potential. The number of ground motion cycles may also influence the degree of seismic damage on structures. Therefore, it would be desirable if the effective number of cycles could be used for a variety of applications. Of the two measures of effective number of cycles considered in this study, \( N_A \) is the absolute definition (calculated based on the absolute amplitude of cycles), while \( N_R \) is the relative definition in which all cycles are normalized by the maximum

\[
\begin{array}{cccc}
\text{Distances (km)} & 0-30 & 30-80 & 80-200 \\
\hline
n & a_n & T_n & a_n & T_n & a_n & T_n \\
1 & 0.75 & 0.01 & 0.79 & 0.01 & 0.82 & 0.01 \\
2 & 0.79 & 0.1 & 0.83 & 0.12 & 0.83 & 0.15 \\
3 & 0.65 & 0.25 & 0.66 & 0.25 & 0.48 & 0.75 \\
4 & 0.25 & 1.0 & 0.24 & 0.75 & 0.19 & 2.0 \\
5 & 0.04 & 4.0 & 0.11 & 2.0 & 0.22 & 4.0 \\
6 & -0.06 & 10 & 0.11 & 10.0 & 0.03 & 10.0 \\
\sigma_t & 0.07 & 0.06 & 0.07 & \\
\end{array}
\]

\[
\begin{array}{cccc}
\text{Distances (km)} & 0-30 & 30-80 & 80-200 \\
\hline
n & a_n & T_n & a_n & T_n & a_n & T_n \\
1 & -0.38 & 0.01 & -0.44 & 0.01 & -0.34 & 0.01 \\
2 & -0.13 & 0.05 & -0.36 & 0.03 & -0.30 & 0.03 \\
3 & -0.09 & 0.13 & -0.12 & 0.1 & -0.11 & 0.1 \\
4 & -0.36 & 0.4 & -0.47 & 0.35 & -0.33 & 0.3 \\
5 & -0.37 & 1.0 & -0.42 & 1.5 & -0.44 & 1.2 \\
6 & -0.25 & 3.0 & -0.09 & 6.0 & -0.19 & 7.0 \\
7 & 0.0 & 10.0 & -0.15 & 10.0 & -0.28 & 10.0 \\
\sigma_t & 0.07 & 0.06 & 0.07 & \\
\end{array}
\]

Fig. 9. Distributions of (a) inter-event and (b) intra-event residuals between \( N_A \) and \( N_R \).

Fig. 10. Number of usable records versus spectral periods in each distance bin.

Table 4

\( N_A \)-SA correlations for different rupture distance bins.

Table 5

\( N_R \)-SA correlations for different rupture distance bins.

Fig. 11. (a) \( N_A \)-SA correlations and (b) \( N_R \)-SA correlations for different rupture distance bins (empirical data and fitting functions are represented as symbols and lines, respectively).
amplitude of cycles. It has been found that $N_A$ is highly correlated with high-frequency IMs such as PGA, Ia, and SA at short periods. Compared to PGA, $N_A$ considers not only the effect of the single cycle with peak amplitude, but also the effect of a number of secondary cycles. Hence, $N_A$ can be regarded as a surrogate of PGA in some cases, in which the estimation of cyclic deformation demand is important [42]. In contrast, $N_R$ exhibits small-to-moderate negative correlations with PGA, PGV, Ia, and SA, indicating that $N_R$ can provide some supplementary information regarding the ground motion characteristics compared with these IMs. Therefore, although $N_R$ alone may not be very useful, $N_R$ in conjunction with primary IMs (e.g., PGA, SA) would be favorable in some engineering applications.

It is not surprising that $N_A$, as an amplitude-based indicator, exhibits negative correlations with $D_{S5-75}$ and $D_{S5-95}$. The correlations of $N_R$ with $D_{S5-75}$ and $D_{S5-95}$ are just moderately positive (with maximum $\rho_{0.5}$ as 0.51), which may contradict the assumption that duration measures can effectively represent the number of cycles of ground motions. These results are some degree of consistent with a previous study which stated that the correlations between ground motion durations and the number of effective cycles are weak [19].

For demonstration purpose, Fig. 12 shows the comparisons of response spectra and time histories for a pair of ground motions. A spectrally equivalent method, which minimizes the sum of the squared effective number of ground motion cycles and other commonly used intensity measures (IMs), including PGA, PGV, spectral accelerations (SA), Ia, CAV, $D_{S5-75}$, and $D_{S5-95}$. The NGA strong motion database and several globally applicable GMPEs have been utilized for calculating the correlation coefficients between these IMs. Two definitions of the effective number of cycles, namely $N_A$ (absolute measure) and $N_R$ (relative measure) are considered in this paper. It was found that $N_A$ has strong positive correlations ($\rho = 0.8$) with high-frequency IMs such as PGA, Ia, and SA at $T < 0.2$ s; moderate positive correlations with PGV, CAV, and SA over a period range of 0.2–1 s; and weak negative correlations ($\rho = −0.3$) with significant durations. The observed correlation results can be explained by the fact that $N_A$ is mainly determined by larger-amplitude ground motion cycles, and it can be classified as a high-frequency amplitude-based IM.

The relative measure $N_R$ generally exhibits small-to-moderate negative correlations with amplitude-based (e.g., PGA, PGV) and cumulative intensity-based IMs (Ia, CAV). This means that $N_R$ can provide additional information regarding the ground motion characteristics compared with these amplitude- and cumulative-based IMs. Besides, it was observed that $N_R$ is moderately correlated ($\rho < 0.6$) with significant duration parameters, indicating that the duration parameters cannot perfectly represent the effective number of ground motion cycles.

The influence of rupture distance on the $N_A$-SA and $N_R$-SA correlations was also examined. It was found that the far-distance ground motions tend to exhibit stronger $N_A$-SA and $N_R$-SA correlations, especially at long spectral periods. A set of piecewise linear functions were proposed to quantify the $N_A$-SA and $N_R$-SA correlations for general and various distance-binned cases. The derived correlation coefficients and the parametric equations can be easily used in vector-based seismic hazard analysis or ground motion selection for scenario earthquakes. A simple example was provided to demonstrate that ground motions selected based on SA only may result in a biased representation of other IMs. Specifically, the correlation results described herein could be used to select a suite of well-representative ground motions, to assess earthquake-induced risks such as liquefaction potential.

Acknowledgments

The work described in this paper was supported by Hong Kong Research Grants Council through Collaborative Research Fund grant No. PolyU8/CRF/13G and General Research Fund grant No. 16213615. The authors thank two anonymous reviewers for their helpful comments to improve this manuscript.

6. Conclusions

This manuscript studied the empirical correlations between the effective number of ground motion cycles and other commonly used intensity measures (IMs), including PGA, PGV, spectral accelerations (SA), Ia, CAV, $D_{S5-75}$, and $D_{S5-95}$. The NGA strong motion database and several globally applicable GMPEs have been utilized for calculating the correlation coefficients between these IMs. Two definitions of the effective number of cycles, namely $N_A$ (absolute measure) and $N_R$ (relative measure) are considered in this paper. It was found that $N_A$ has strong positive correlations ($\rho = 0.8$) with high-frequency IMs such as PGA, Ia, and SA at $T < 0.2$ s; moderate positive correlations with PGV, CAV, and SA over a period range of 0.2–1 s; and weak negative correlations ($\rho = −0.3$) with significant durations. The observed correlation results can be explained by the fact that $N_A$ is mainly determined by larger-amplitude ground motion cycles, and it can be classified as a high-frequency amplitude-based IM.

The relative measure $N_R$ generally exhibits small-to-moderate negative correlations with amplitude-based (e.g., PGA, PGV) and cumulative intensity-based IMs (Ia, CAV). This means that $N_R$ can provide additional information regarding the ground motion characteristics compared with these amplitude- and cumulative-based IMs. Besides, it was observed that $N_R$ is moderately correlated ($\rho < 0.6$) with significant duration parameters, indicating that the duration parameters cannot perfectly represent the effective number of ground motion cycles.

The influence of rupture distance on the $N_A$-SA and $N_R$-SA correlations was also examined. It was found that the far-distance ground motions tend to exhibit stronger $N_A$-SA and $N_R$-SA correlations, especially at long spectral periods. A set of piecewise linear functions were proposed to quantify the $N_A$-SA and $N_R$-SA correlations for general and various distance-binned cases. The derived correlation coefficients and the parametric equations can be easily used in vector-based seismic hazard analysis or ground motion selection for scenario earthquakes. A simple example was provided to demonstrate that ground motions selected based on SA only may result in a biased representation of other IMs. Specifically, the correlation results described herein could be used to select a suite of well-representative ground motions, to assess earthquake-induced risks such as liquefaction potential.
References

[27] Roore DM, Atkinson GM. Ground-motion prediction equations for the average horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.1 to 10 s. Earthq Spectra 2008;24(1):139–71.
[28] Campbell KW, Bozorgnia Y. NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.1 to 10 s. Earthq Spectra 2008;24(1):139–71.