Empirical Correlations between Peak Ground Velocity and Spectrum-based Intensity Measures

Brendon A Bradley a) b)

Empirical correlation equations between peak ground velocity (PGV) and several spectrum-based ground motion intensity measures are developed. The intensity measures examined in particular were: peak ground acceleration (PGA), 5% damped pseudo-spectral acceleration (SA), acceleration spectrum intensity (ASI), and spectrum intensity (SI). The computed correlations were obtained using ground motions from active shallow crustal earthquakes and four ground motion prediction equations. Results indicate that PGV is strongly correlated (i.e. a correlation coefficient of $\rho = 0.89$) with SI, moderately correlated with medium to long-period SA (i.e. $\rho \approx 0.8$ for vibration periods 0.5 – 3.0 seconds), and also moderately correlated with short period SA, PGA and ASI ($\rho \approx 0.7 – 0.73$). A simple example is used to illustrate one possible application of the developed correlation equations for ground motion selection.

INTRODUCTION

The peak value of the velocity time history of an earthquake-induced ground motion is a very simple measure of its severity to engineering structures. Despite this, peak ground velocity (PGV) has been demonstrated to provide a significantly better correlation with seismic demand than other simple instrumental ground motion intensity measures such as peak ground acceleration (PGA). Such strong correlations have been observed for damage to structures on a regional scale (e.g. Boatwright et al. 2001, Kaka and Atkinson 2004, Wald et al. 1999), soil liquefaction potential (e.g. Kostadinov and Towhata 2002, Orense 2005), and damage to infrastructure such as pipe networks (e.g. Davis and Bardet 2000, Jeon and O'Rourke 2005). Furthermore, various analytical studies have also shown that the PGV of a ground motion exhibits a strong correlation with seismic demand (and therefore consequent

---

a) Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand
b) Institute of Science and Engineering, Chuo University, Tokyo, Japan
damage) for single-degree-of-freedom, multi-degree-of-freedom and soil-pile-foundation structures (e.g. Akkar and Ozen 2005, Bradley et al. 2009a, Kappos and Kyriakakis 2000), particularly in cases where ground motions contain near-fault directivity effects (e.g. Alavi and Krawinkler 2004, Malhotra 1999, Yang et al. 2009). Empirical ground motion prediction equations are also now available for predicting $PGV$ with arguably the same level of sophistication as other ground motion intensity measures, and with often higher precision (e.g. Power et al. 2008).

The aforementioned references indicate that $PGV$ is an IM which can provide useful information as to the severity of a ground motion. Despite this, $PGV$ is not used in seismic hazard analysis and ground motion selection to the same extent for which other intensity measures, such as $PGA$ or spectral acceleration (SA) have been utilized (e.g. Katsanos et al. (2010) and references therein). This manuscript examines, and develops parametric equations for, the empirical correlation observed between $PGV$ and several other response spectrum-based ground motion intensity measures (IMs), including $PGA$ and $SA$, among others. The developed correlation equations consequently allow estimation of the joint distribution of $PGV$ and these other examined IMs, which will improve the potential use of $PGV$ in seismic hazard analysis and ground motion selection.

This manuscript is laid out as follows: Firstly, the considered IMs, ground motion prediction equations (GMPEs), empirical ground motion dataset, and correlation methodology are presented. Secondly, the observed correlation results are discussed and parametric models developed. Finally, a simple example illustrates one possible use of the correlation equations in developing conditional IM distributions useful for ground motion selection.

CORRELATION OF PGV WITH SPECTRUM-BASED INTENSITY MEASURES

INTENSITY MEASURES EXAMINED

The specific intensity measures considered herein for correlations with $PGV$ are: (i) peak ground acceleration, $PGA$; (ii) (5% damped pseudo) spectral acceleration, $SA$, for periods from 0.01 to 10 seconds; (iii) acceleration spectrum intensity, $ASI$ (Von Thun et al. 1988); and (iv) spectrum intensity, $SI$ (Housner 1952). $ASI$ and $SI$, in particular, are defined as:

$$ASI = \int_{0.1}^{0.5} SA(T; 5\%) dT$$  \hspace{1cm} (1)
\[ SI = \int_{0.1}^{2.5} SV(T, 5\%) dT \]

where \( SA(T, 5\%) \) is the 5\% damped pseudo spectral acceleration at vibration period \( T \) (herein \( SA \) for brevity), and \( SV(T, 5\%) \) is the pseudo-spectral velocity. As can be seen from their definitions, \( ASI \) and \( SI \) can be considered representative of the average intensity of a ground motion for short and moderate periods, respectively (Housner 1952, Von Thun et al. 1988).

The number of ground motion IMs considered herein is limited because of scope and is not intended to imply that other IMs are not important, or that the considered IMs are best used in combination with \( PGV \) in seismic hazard analysis, seismic response analysis, or ground motion selection.

GROUND MOTION PREDICTION EQUATIONS ADOPTED

The predicted distributions of the aforementioned intensity measures for each ground motion record were computed using four of the next generation attenuation (NGA) ground motion prediction equations, specifically: Boore and Atkinson (2008), Chiou and Youngs (2008b), Campbell and Bozorgnia (2008), and Abrahamson and Silva (2008). For brevity, these four GMPEs are herein referred to as BA08, CY08, CB08, and AS08, respectively. The four NGA GMPEs provide explicit predictions for \( PGV \), \( PGA \) and \( SA \). These GMPEs can also be used to predict \( SI \) and \( ASI \) using the analytical equations developed by Bradley et al. (2009b) and Bradley (2009), respectively. Herein, for example, a prediction of \( ASI \) using the Bradley et al. (2009b) analytical equation and the BA08 \( SA \) model is simply referred to as a BA08 \( ASI \) prediction.

GROUND MOTION DATABASE

Ground motion records from active shallow crustal earthquakes in the NGA database (Chiou et al. 2008) (http://peer.berkeley.edu/nga) were used to obtain empirical values of \( PGV \), \( PGA \), \( SA \), \( SI \), and \( ASI \) for use in the subsequent correlation analyses. The specific subset of the NGA database used is that of Chiou and Youngs (2008a), but with the removal of aftershocks from the Chi-Chi earthquake (i.e. a total of 1842 ground motions). While this subset is not the same as the individual subsets used by the various GMPEs developed from the NGA database, all of these GMPEs are applicable for a range of predictor variables (i.e. magnitude, distance, etc.) which encompass all of those records used in the present study (comparison of the distributions of inter- and intra-event residuals obtained using each of the models confirmed this). This ensures that the point estimated correlation coefficients
computed subsequently are unbiased.

Empirical IM values and the predicted IM distribution for each record were only computed if the IM was within the usable period range of the record. For example, if the maximum usable period was 2.0 seconds then no spectral ordinates above 2.0 seconds or spectrum intensity (which is based on $SA$ from 0.1-2.5 seconds) would be computed. Because $PGV$ is also affected by the high pass filter cut-off period (Akkar and Bommer 2007) then those motions with cut-off periods less than 4.0 seconds were also not considered in correlation analyses involving $PGV$.

**POINT-ESTIMATED CORRELATIONS FROM GROUND MOTION DATA**

Prediction equations for ground motion IMs have the general form:

$$lnIM_i = f_i(Rup_k, Site) + \varepsilon_{T_i} \sigma_{T_i}(Rup_k, Site)$$

where $ln$ is the natural logarithm; $f_i(Rup_k, Site) = \mu_{lnIM_i|Rup_k}$ and $\sigma_{T_i}(Rup_k, Site) = \sigma_{T_{lnIM_i|Rup_k}}$ are the predicted mean and standard deviation of $lnIM_i$, which are a function of properties of the earthquake rupture and site of interest; and $\varepsilon_{T_i}$ is the (normalized) total residual for $IM_i$. $\varepsilon_{T_i}$ is also assumed to have a standard normal distribution.

By rearranging Equation (3), the total residual, $\varepsilon_{T_i,n}$, can be interpreted as the number of standard deviations a particular ground motion, $n$, has $lnIM_i$ above the predicted mean:

$$\varepsilon_{T_i,n} = \frac{lnIM_{i,n} - \mu_{lnIM_i|Rup_k}}{\sigma_{T_{lnIM_i|Rup_k}}}$$

Since a single IM represents a significant simplification of a ground motion then the use of multiple IMs (i.e. a vector of IMs) may be desired. The probabilistic characterisation of a ground motion using multiple IMs must account for the fact that the individual IMs are correlated. For example, if a ground motion has a higher than expected 1.0-second spectral acceleration (i.e. $SA(1.0)$), then it may also be expected that $PGV$ will be higher than expected because of the previously perceived similarities between $PGV$ and $SA(1.0)$ (Bommer and Alarcon 2006).

Because of the linearity between $\varepsilon_{T_i,n}$ and $lnIM_{i,n}$ in Equation (4), it follows that the correlation coefficient between two different IM’s (at the same location due to a given rupture $k$) can be estimated based on the correlation of the total residuals, i.e. $\rho_{lnIM_i|Rup_k,lnIM_j|Rup_k} = \rho_{\varepsilon_{T_i},\varepsilon_{T_j}}$. As such, previous studies (Baker and Jayaram 2008, Bradley 2010a, Inoue and
Cornell (1990) have computed $\rho_{\varepsilon_{TI},\varepsilon_{TJ}}$ using the Pearson-product-moment correlation estimator (Ang and Tang 2007):

$$\rho_{x,y} = \frac{\sum_n[(x - \bar{x})(y - \bar{y})]}{\sqrt{\sum_n[(x - \bar{x})^2] \sum_n[(y - \bar{y})^2]}}$$

(5)

where $x$ and $y$ are generic variables (i.e., in this case $\varepsilon_{TI}$ and $\varepsilon_{TJ}$), $\bar{x}$ and $\bar{y}$ are the sample means of $x$ and $y$ and $\sum_n[\ ]$ represents summation over the number of ground motion records (i.e., $n = 1 - N_{\text{record}}$), or the number of earthquakes (in the case of the inter-event correlation). Herein, for brevity, this correlation between two intensity measures, $\rho_{IM_i|\text{Rup}_k,IM_j|\text{Rup}_k}$, will simply be referred to as $\rho_{IM_i,IM_j}$ (where the conditioning on $\text{Rup}_k$ and the same location is implied).

Equation (5) assumes that each pair of observed data is independent of the other observed data pairs (i.e., that $x_n$ and $x_m$ are independent and that $y_n$ and $y_m$ are independent). However, it has long been realised that this is not the case, primarily because of the fact that ground motion datasets used in statistical analyses typically contain multiple records from the same earthquake event. As such, all of the GMPE’s utilized in this study used mixed-effects regression models (Lindstrom and Bates 1990) in which this correlation is accounted for through the use of both inter- and intra-event residuals, representing fixed, and random effects (Abrahamson and Youngs 1992). The relationship between the total (normalized) residual, $\varepsilon_{TI}$, and the normalized inter- and intra-event residuals is given by:

$$\varepsilon_{TI} \sigma_{\varepsilon_{TI}} = \eta_i \tau_i + \varepsilon_i \sigma_i$$

(6)

where $\eta_i$ and $\varepsilon_i$ are the (normalized) inter- and intra-event residuals; and $\tau_i$ and $\sigma_i$ are the inter- and intra-event standard deviations (where the subscript of all terms in Equation (6) pertains to $IM_i$). The GMPE’s utilized in this study all assume that the partitioning of the total residual into inter- and intra-event residuals, as given by Equation (6), makes all of the $\eta_i$ and $\varepsilon_i$ values for the $N_{\text{record}}$ records considered independent. Therefore it is possible to use Equation (5) to estimate the correlation between the inter- and intra-event residuals for two different intensity measures, $IM_i$ and $IM_j$. Using the definition of the correlation coefficient, the correlation between the total residuals can then be found from the inter- and intra-event correlations from:

$$\rho_{\varepsilon_{TI},\varepsilon_{TJ}} = \frac{1}{\sigma_{\varepsilon_{TI}} \sigma_{\varepsilon_{TJ}}} \left[ \rho_{\eta_i\eta_j} \tau_i \tau_j + \rho_{\varepsilon_i\varepsilon_j} \sigma_i \sigma_j \right]$$

(7)
The inter- and intra-event residuals were obtained by performed mixed-effects regression on the total residuals, as also done by Scasserra et al. (2009).

Thus, one can develop correlation models for the inter- and intra-event residuals (i.e. \( \rho_{\eta_i,\eta_j} \) and \( \rho_{\epsilon_i,\epsilon_j} \)), and then with the inter- and intra-event standard deviations for two different intensity measures, \( IM_i \) and \( IM_j \) determine the correlation in the total residuals via Equation (7). This implies that the correlation between the total residuals is a function of the particular rupture scenario considered because of the fact that GMPE standard deviations (i.e. \( \tau_i \) and \( \sigma_i \)) are a function of rupture scenario. While the computation of \( \rho_{\epsilon_i,\epsilon_j} \) via Equation (7) is possible (given that published GMPE’s routinely include both their inter- and intra-event standard deviations), it is undesirable in the sense that many site-specific PSHA programs consider only the total standard deviation (which is all that is needed for a single site). Therefore it is desirable to have explicit predictive equations for \( \rho_{\epsilon_i,\epsilon_j} \) (i.e. the correlation between the total residuals at a single site).

In order to overcome the aforementioned difference, use is made of the fact that the correlation in the total residuals, \( \rho_{\epsilon_i,\epsilon_j} \) is in fact very insensitive to the particular rupture scenario considered. In order to demonstrate this insensitivity it is advantageous to use the fact that \( \sigma_T^2 = \tau_i^2 + \sigma_i^2 \) to re-write Equation (7) as:

\[
\rho_{\epsilon_i,\epsilon_j} = \rho_{\eta_i,\eta_j} \sqrt{(1 - \alpha_i^2)(1 - \alpha_j^2)} + \rho_{\epsilon_i,\epsilon_j} \alpha_i \alpha_j \tag{8}
\]

where \( \alpha_i = \sigma_i / \sigma_T \) is the ratio of the intra-event standard deviation to the total standard deviation for \( IM_i \).

Figure 1 illustrates the ratio, \( \alpha_{PGV} \), between the intra-event and total standard deviations predicted using the CY08 PGV model for all of the ground motions considered in this study. Also shown in Figure 1 are the mean, 16\(^{th}\) and 84\(^{th}\) percentiles of the ratio. It can be seen that the ratio is very stable with a mean, 16\(^{th}\) and 84\(^{th}\) percentiles of 0.905, 0.895 and 0.92, respectively. For the sake of example, if it is given that the inter- and inter-event correlation coefficients between PGV and PGA are \( \rho_{\eta_i,\eta_j} = 0.8 \) and \( \rho_{\epsilon_i,\epsilon_j} = 0.7 \), respectively, then via Equation (8) the mean, 16\(^{th}\) and 84\(^{th}\) percentiles of \( \alpha_{PGV} \) give total correlations between PGV and PGA of \( \rho_{\epsilon_i,\epsilon_j} = 0.718, 0.72, \) and 0.715, respectively. That is, the difference between the 16\(^{th}\) and 84\(^{th}\) percentiles is only 0.05. Furthermore, it is noted that the variation in \( \alpha_i \) is
smaller for the BA08 and CB08 models (which have magnitude independent standard deviations), and similar for the AS08 model. Comparisons with the uncertainties in the correlation coefficients due to finite sample size and the adopted GMPE presented later in the manuscript illustrates that these uncertainties are an order of magnitude larger than that due to the variability in \( \alpha_i \), therefore demonstrating that the correlation between total residuals can be computed adequately from Equation (8) using a constant (mean) value of \( \alpha_i \).

Finally, it is worth noting that while particular attention has been given here to the correlation resulting from multiple ground motion recordings from the same earthquake rupture, there are additional sources of correlation in many ground motion datasets due to multiple recordings at a single location (Chen and Tsai 2002) and spatial correlation of ground motions (Jayaram and Baker 2010). Because these two sources of correlation were not considered in the GMPE’s examined in this study they are not discussed further. These effects should be accounted for in ground motion correlation analyses which use GMPE’s with such effects.

**UNCERTAINTY IN THE CORRELATION COEFFICIENT**

As can be seen from Equation (4), the residual for a particular \( n \)th record, \( \varepsilon_{i,n} \), is a function of the GMPE used to estimate the mean, \( \mu_{\ln IM|Rup_k} \), and standard deviation, \( \sigma_{\ln IM|Rup_k} \). Therefore the correlation coefficient by way of Equation (5) is also a function of the GMPE used. This uncertainty in the correlation coefficient due to the adopted GMPEs is accounted for using the four aforementioned GMPE’s developed for active shallow crustal tectonic regions. It is assumed that each GMPE is equally valid, meaning that each is given a logic tree weight of 0.25 (Kulkarni et al. 1984).

In addition to uncertainty due to the selected GMPEs, Equation (5) provides only the point-estimate of the correlation coefficient, which in reality contains uncertainty due to the finite number of datapoints used in its determination. In order to account for such finite sample uncertainty it is first beneficial to use the approximate variance stabilizing Fisher \( z \) transformation, which is defined by:

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

where \( \rho \) is the pearson correlation coefficient, \( \ln \) is the natural logarithm; \( \tanh^{-1} \) is the inverse hyperbolic tangent function; and \( z \) is the transformed correlation coefficient. It can be shown that \( z \) has approximately a normal distribution with mean given by Equation (9), and
variance given by:

\[ Var[z] = \frac{1}{N - 3} \]  

(10)

where \( N \) is the sample size. An alternative to the use of Equation (10) for estimating the variance in \( z \), which is utilized herein and also applicable for all sample sizes, is the non-parametric bootstrap method (Ang and Tang 2007), in which random bootstrap samples are drawn from the observed data. Given that the transformed correlation coefficient was found to be well represented by a normal distribution (which was also found for other IM combinations in Bradley (2010a)), it follows that the median value of the correlation coefficient, \( \rho_{50} \), is equal to the inverse Fisher transformation of the mean of \( z \), \( \mu_z \), specifically:

\[ \rho_{50} = \frac{e^{2\mu_z} - 1}{e^{2\mu_z} + 1} = \tanh(\mu_z) \]  

(11)

EMPIRICAL CORRELATION RESULTS

EXAMPLE SCATTERPLOTS AND THE MULTIVARIATE NORMALITY ASSUMPTION

Figure 2 illustrates example scatterplots of the normalised inter- and intra-event residuals for \( PGV \) and \( PGA \) obtained from the adopted dataset using the BA08 GMPE and the previously discussed methodology. It can be seen that the correlation coefficient between the inter- and intra-event residuals is similar with values of approximately 0.75. It is also important to note the larger number of data points for the intra-event residuals compared to the inter-event residuals, which consequently leads to the intra-event correlations having less finite sample uncertainty than the inter-event correlations.

The linearity of the scatterplots in Figure 2 provides qualitative evidence that the residuals maybe approximately assumed to follow a bivariate normal distribution. This qualitative observation is examined more formally in Figure 3 which illustrates a Chi Squared quantile-quantile plot (Johnson and Wichern 2007) of the distribution of Mahalanobis distances of the results in Figure 2. The basis of such a plot is the fact that if a distribution is bivariate (or generally multivariate) normal, then a sample of data from such a distribution will have Mahalanobis distances (a normalized Euclidian distance for the difference between an observation and prediction distribution) which have a Chi Square distribution (Johnson and Wichern 2007). Thus, the quantile-quantile plots in Figure 3 examine whether the Mahalanobis distances of the data in Figure 2 have a Chi Square distribution implying that the data themselves have bi-variate normal distributions. The fact that Figure 3 illustrates that both the inter- and intra-event
residuals generally follow the 1:1 line illustrates that the assumption of bivariate normality is sufficiently accurate. Although not shown here for brevity, similar results were also observed for other intensity measure combinations, and also more than two intensity measures (i.e. for the general multivariate normal case). These findings are in agreement with those of Jayaram and Baker (2008), who found that $SA$ values at multiple periods can also be adequately assumed to be multivariate normally distributed.

**INTER-EVENT, INTRA-EVENT AND TOTAL RESIDUAL CORRELATIONS**

Figure 4a-Figure 4c illustrate the correlation coefficients computed between $PGV$ and $PGA$ for the inter-event, intra-event and total residuals, respectively. Figure 4d illustrates the correlation coefficient between the total residuals as computed via Equation (7). In each of Figure 4a-Figure 4d box plots (Ang and Tang 2007) are used to illustrate the distribution of the correlation coefficient due to sample size uncertainty (obtained from bootstrap sampling), while the multiple box plots illustrate the effect of which GMPE is used. It can be seen that the correlation coefficients between the inter-event residuals and intra-event residuals are relatively similar with median values in the range of 0.72-0.76. It can also be clearly seen that the magnitude of the sample size uncertainty for the inter-event correlations is significantly larger than for the intra-event residuals, an obvious side-effect of there being significantly more ground motion records than earthquakes in the empirical dataset. Figure 4d illustrates that the correlation coefficients computed via Equation (7) are relatively similar to the intra-event correlations, as a result of the fact that the intra-event standard deviations are notably larger than inter-event standard deviations. On the other hand, Figure 4c illustrates the empirical correlations of the total residuals obtained by simply performing a correlation analysis (i.e. Equation (5)) of the total residuals directly, rather than the more rigorous approach via Equation (7). Despite its theoretical deficiency, it can be seen that the estimated correlations in Figure 4c are still within the same vicinity as those of Figure 4d (i.e. median correlations in the range 0.66-0.73 in Figure 4c in comparison with 0.73-0.75 in Figure 4d). Another important observation is the notably larger effect of the selected GMPE on the resulting correlation coefficient in Figure 4c, with, as noted above, a range in the median correlation coefficient of 0.02 in Figure 4d, but a range of 0.07 in Figure 4c. These observations suggest that empirical correlation equations which have been previously developed based on total residuals directly are still useful, but are likely to have a small bias and less precision, compared to those estimated by separately accounting for inter- and intra-event correlations.
MEDIAN CORRELATION COEFFICIENT

Figure 5 illustrates the correlation coefficients computed for three of the four different intensity measures examined (i.e. for ASI, SI, and PGA). As one would expect, given that SI represents an average measure of moderate frequency ground motion intensity (Housner 1952), then it has a high correlation with PGV (Figure 5c). On the other hand, as PGA and ASI represent measures of high frequency ground motion intensity (Von Thun et al. 1988), then they have a lower correlation with PGV.

It can be seen that variability in the correlation coefficients (as indicated by the ‘whiskers’ of the box plots) due to sample uncertainty is approximately 0.075, 0.07, and 0.035 for $\rho_{\ln PGV, \ln ASI}$, $\rho_{\ln PGV, \ln PGA}$, and $\rho_{\ln PGV, \ln SI}$, respectively (i.e. as previously mentioned it is a function of the correlation coefficient). It can also be seen that while there is some difference between the correlation coefficients estimated using the four different GMPE’s, that it is not significant, and generally is of the same order of magnitude as the uncertainty in the correlation coefficient due to sample size uncertainty.

Figure 6 illustrates the empirical correlation obtained between PGV and SA at various vibration periods. In Figure 6, the median and 90% confidence interval of the correlation coefficient for a given GMPE (i.e. considering only sample size uncertainty) are indicated by solid and dashed lines, respectively. Different line colors are used to represent the distribution of the correlation coefficient obtained using the four different GMPE’s considered. It can be seen that the correlation between PGV and SA is largest for vibration periods in the range $T = 0.5 - 3.0$ seconds, consistent with the previous use of correlations of PGV and SA(1.0) in response spectrum construction (Bommer and Alarcon 2006, Newmark and Hall 1982). As $T$ increases from moderate to longer periods the correlation reduces, albeit only slightly. On the other hand, as $T$ decreases from moderate to short periods there is an initially drastic reduction to $T = 0.01$ and then a slight increase for very high frequency SA values (i.e. $T < 0.01s$), something which has been observed in other studies also (Baker and Jayaram 2008, Bradley 2010a).

As noted also by Bommer and Alarcon (2006), the correlation coefficient computed here between PGV and moderate period spectral ordinates is not large in an absolute sense (i.e. for SA(1.0) the correlation is in the vicinity of 0.8). Given that the standard deviation in lnSA(1.0) is typically in the order of 0.65 (e.g. Boore and Atkinson 2008), then the uncertainty in lnSA(1.0) given a specific value of PGV is $\sigma_{\ln SA(1.0), \ln PGV} = \sigma_{\ln SA(1.0)} \sqrt{1 - \rho_{\ln SA(1.0), \ln PGV}} = $
A value of \( \sigma_{\ln SA(1.0), \ln PGV} = 0.39 \) implies that there is roughly a 68% probability that the actual value of \( SA(1.0) \) is between 0.68 and 1.48 times the value given by such \( PGV - SA(1.0) \) relationships (e.g. Newmark and Hall 1982). Hence, the results presented here reiterate the recommendation of Bommer and Alarcon (2006) that \( PGV \) should be estimated directly rather than via correlations with \( SA \) values.

**PARAMETRIC CORRELATION MODELS**

**FORMAT FOR PRESENTATION OF CORRELATION EQUATIONS**

Here parametric equations are developed for: (i) the median correlation coefficient, \( \rho_{50} \), and (ii) the standard deviation of the transformed correlation coefficient, \( \sigma_z \). By developing the equations in this manner, if an analyst wishes to neglect correlation coefficient uncertainty they can simply use \( \rho_{50} \) directly. On the other hand, if one wants to explicitly account for correlation coefficient uncertainty then Equation (11) can be used to determine \( \mu_z \) from the equation for \( \rho_{50} \), and then with \( \sigma_z \), the correlation coefficient for a given \( x^{th} \) percentile can be obtained from:

\[
\rho_x = \left( \frac{e^{2z_x} - 1}{e^{2z_x} + 1} \right) = \tanh(z_x)
\]

where \( z_x \) is the \( x^{th} \) percentile of a normal distribution with mean and standard deviation \( \mu_z \) and \( \sigma_z \), respectively (Ang and Tang 2007).

**CORRELATION EQUATIONS BETWEEN PGV AND ASI, PGA, SI, AND SA**

Table 1 provides the median correlation coefficient, \( \rho_{50} \), and the standard deviation of the transformed correlation, \( \sigma_z \), between \( PGV \) and \( ASI, PGA, \) and \( SI \). These values have been obtained directly from the mean and standard deviation of the \( z \) values obtained from the aforementioned correlation analyses (i.e. Figure 5), after combining the results using the four GMPEs in a logic tree. No smoothing of the \( \rho_{50} \) or \( \sigma_z \) values were conducted.

The correlation coefficients computed involving \( PGV \) and \( SA \) have 23 different values for the 23 vibration periods between \( T = [0.01,10] \) for which \( SA \) was computed for. Rather than providing the median and standard deviation values of the correlation coefficient between \( PGV \) and \( SA \) for each of these 23 different vibration periods it is beneficial to represent the correlation coefficient, \( \rho_{\ln PGV, \ln SA(T)} \), as a continuous function of vibration period, \( T \). Such a continuous function offers the benefits of: (i) interpolation of the correlation coefficient at vibration periods different than the 23 periods for which it was explicitly computed; and (ii) a
smooth variation of correlation coefficient as a function of vibration period, rather than the slightly fluctuating empirical values that result due to the finite number of ground motion data, and specific GMPE’s used.

The specific functional form used to represent the variation in the (median) correlation coefficient with vibration period is a piece-wise function with each piece-wise segment having the following functional form:

\[
\rho_{lnPGV,lnSA(T)} = \frac{a_n + b_n}{2} - \frac{a_n - b_n}{2} \tanh[d_n \ln(T/c_n)] \quad e_{n-1} \leq T < e_n
\]  \hspace{1cm} (13)

where \( \tanh \) is the hyperbolic tangent function; \( a_n, b_n, c_n, \) and \( d_n \) are empirical constants used to fit the observed empirical correlation coefficient for piece-wise segment \( n \); and \( e_n \) defines the period range for each of the piece-wise segments. Equation (13) was fit using least squares to minimize the misfit in the mean transformed \( z \) value (i.e. to minimise the error between the Fisher transformation of Equation (13) and the empirical values of \( \mu_z \)). The parameters of each piece-wise segment were also constrained so that the parametric equation is a continuous piece-wise function of \( T \). The obtained parametric equations were then plotted and compared with the empirical correlation values, and adjusted manually if necessary.

It was also found that there was a relatively simple variation of the standard deviation of the transformed correlation coefficient between \( PGV \) and \( SA \) with vibration period. The observed standard deviation was relatively constant, with the exception of spectral accelerations at long periods in which the number of ground motions utilized in the correlation analyses was significantly reduced because of the cut-off filters used in processing. As such, a simple standard deviation model with a constant value for \( SA \)’s with periods less than 2.0 seconds and then an increase with increasing \( T \) was adopted. Specifically,

\[
\sigma_z = 0.037 \cdot \max(1,1 + 0.6 \cdot \ln[T/2])
\]  \hspace{1cm} (14)

As previously mentioned, once Equation (13) and (14) are utilized to compute \( \rho_{z0} \) and \( \sigma_z \), respectively, then any percentile of the correlation coefficient can be obtained using Equations (11) and (12).

Table 2 provides the numerical values of the parameters in Equation (13) for the median correlation between \( PGV \) with \( SA(T) \), in which a total of four piece-wise segments were used. Figure 7a illustrates a comparison between the parametric correlation equation (i.e. Equation (13) and (14)) and the empirical correlation values obtained from the aforementioned analyses. In Figure 7a, the solid line represents the median correlation coefficient, \( \rho_{z0} \), while the dashed lines represent the 16\textsuperscript{th} and 84\textsuperscript{th} percentiles. It can be seen that, in general, the
median correlation coefficient is very well approximated by Equation (13). The only noticeable point of departure between the empirical results and the parametric equation is for periods greater than $T = 5$ seconds. In this region it was decided that the larger correlation of $PGV$ and $SA(10.0)$ compared with $PGV$ and $SA(7.5)$ was likely the result of a lack of ground motion records. Therefore, in the parametric model a smoother variation with vibration period was adopted. Figure 7b compares explicitly the empirical values of $\sigma_z$ with the parametric fit provided by Equation (14). As ascertained also from Figure 7a, the empirical variation of $\sigma_z$ is well captured by Equation (14).

**STANDARD DEVIATION OF TRANSFORMED CORRELATION COEFFICIENT**

Table 1 and Figure 8 illustrated that the standard deviation in the transformed correlation coefficient, $\sigma_z$ (i.e. including both sample size and GMPE uncertainty) is typically in the range $[0.035,0.045]$ with the exception of those correlations involving $SA$ values at long periods. Given that a total of 1842 ground motions were used in this study, Equation (10) can be used to ascertain that the approximate portion of $\sigma_z$ due to sample size uncertainty is $1/\sqrt{1842-3} = 0.023$ (with the exception of long period $SA$ values for which fewer ground motions were used). This indicates that the total standard deviation (comprising sample size and GMPE uncertainty) in the transformed correlation coefficient, is comprised largely by GMPE uncertainty (approximately 0.026-0.039 given a total uncertainty of 0.035-0.045) and to a lesser extent sample size uncertainty (approximately 0.023 as calculated above).

**COMPARISON WITH OTHER CORRELATIONS BETWEEN PGA AND PGV**

While empirical correlation equations for three of the four IM combinations have not, as far as the author is aware, been previously developed, there are two previous instances where correlations between $PGV$ and $PGA$ have been computed. These two instances are in the CB08 and AS08 NGA GMPE’s. Both CB08 and AS08 use the $PGA$ on bedrock in order to determine the non-linear scaling of site response in their $PGV$ GMPEs. As such, when computing the standard deviation of $PGV$ using these equations one needs the correlation between $PGA$ (on rock) and $PGV$. It is worthy of note that in the CB10 and AS08 GMPEs the correlation between $PGV$ and $PGA$ is based on the intra-event residuals only, in comparison to the computation of the correlation between total residuals via inter- and intra-event correlations employed in this study. Figure 8 compares the parametric distribution of the $PGV$-$PGA$ correlation, $\rho_{\ln PGV,\ln PGA}$, model developed here (as given by the parameters in Table 2) with the tabulated point-estimated correlation values in CB08 and AS08. It can be
seen that the point-estimated correlation coefficients computed by CB08 and AS08 are in agreement with the model developed in this study (i.e. the CB08 and AS08 values differ from the median of the model developed here by 6% and 1%, respectively).

APPLICATION: IM DISTRIBUTIONS CONDITIONAL ON PGV

The severity of a specific ground motion for a particular structure is, in general, a function of the amplitude, frequency content, and duration of the ground motion. Various ground motion intensity measures (IMs) capture these three features describing ground motion severity in a simplified and different way. Hence, it can be generally appreciated that the use of a single IM will provide an insufficient representation of the severity of a specific ground motion for a specific structure (Bradley et al. 2009a, Bradley et al. 2009c, Luco and Cornell 2007, Shome et al. 1998). Emerging ground motion selection procedures, such as the generalized conditional intensity measure (GCIM) approach (Bradley 2010b), provide a probabilistic framework in which any number of different ground motion IMs can be considered in ground motion selection, therefore obtaining the benefits of different features of a ground motion that the various IMs capture. The previously examined $PGV$ GMPEs and the developed correlation equations between $PGV$ and other ground motion IMs enable $PGV$ to be considered is such ground motion selection, as the example below illustrates.

Figure 9 illustrates the $PGV$ seismic hazard curve computed using OpenSHA (Field et al. 2003) for an example rock site in California (Lat: 34.053, Long: -118.243, $V_{s,30} = 760m/s$) using the BA08 GMPE and the 1996 USGS earthquake rupture forecast (Frankel et al. 1996). Also shown is the disaggregation of the $PGV$ hazard at the 10% probability of exceedance in 50 years. Based on those ground motions contributing to the seismic hazard at this exceedance probability, and given a $PGV = 27.7$ cm/s, conditional distributions of other ground motion intensity measures can be obtained using the GCIM approach (Bradley 2010b). Figure 10 illustrates these conditional distributions of $PGA$, $SI$, $ASI$ and $SA$, for the given site and $PGV$ with a 10% probability of exceedance in 50 years. The BA08 GMPE and the correlation equations developed in this manuscript were used to predict the conditional distributions of $PGA$, $SI$, $ASI$, and $SA$. Hence, the results in Figure 10 provide the $PGA$, $SI$, $ASI$, and $SA$ distributions of ground motions with a $PGV$ of 27.7 cm/s occurring at the site considered (Bradley 2010b). Hence, for this particular level of seismic hazard ground motions should be selected to match these ‘target’ distributions in order to provide an accurate and precise estimate of seismic demand and seismic performance of structures at this site (e.g.
In addition to the use of the developed correlation equations in the manner of the previous example, they also allow $PGV$ to be used with the other considered intensity measures in vector-based probabilistic seismic hazard analysis (Bazzurro 1998).

CONCLUSIONS

This manuscript has examined, and developed equations for, the empirical correlation between peak ground velocity, $PGV$, and spectrum-based intensity measures, namely: peak ground acceleration ($PGA$), 5% damped pseudo spectral acceleration ($SA$), acceleration spectrum intensity ($ASI$), and spectrum intensity ($SI$). The empirical correlations were based on ground motions from active shallow crustal earthquakes and the use of four different empirical ground motion prediction equations (GMPEs). Particular attention was given to the computation of the empirical correlations based on the combination of the correlations between inter- and intra-event residuals, and the uncertainty in the correlation coefficients due to the finite number of ground motions and various GMPEs considered. The results indicate that $PGV$ is strongly correlated (i.e. a correlation coefficient of $\rho = 0.89$) with $SI$, moderately correlated with medium to long-period $SA$ (i.e. $\rho \approx 0.8$ for vibration periods 0.5 – 2.0 seconds), and also moderately correlated with short period $SA$, $PGA$ and $ASI$ ($\rho \approx 0.7 - 0.73$). Finally, using the developed correlation equations, it was illustrated how distributions of various intensity measures conditioned on the occurrence of a specific value of $PGV$ could be computed for use in the selection of ground motion records for seismic response analysis.

REFERENCES:


Boore, D. M., Atkinson, G. M., 2008. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01s and 10.0s, Earthquake Spectra, 24, 99-138.


Campbell, K. W., Bozorgnia, Y., 2008. NGA Ground Motion Model for the Geometric Mean Horizontal Component of PGA, PGV, PGD and 5% Damped Linear Elastic Response Spectra for Periods Ranging from 0.01 to 10 s Earthquake Spectra, 24, 139-171.


Table 1: Proposed median correlation, $\rho_{50}$, and the standard deviation of the transformed correlation, $\sigma_z$ (shown in brackets) between $ASi$, $SI$, and $PGA$.

<table>
<thead>
<tr>
<th>$\rho_{50}$ ($\sigma_z$)</th>
<th>$ASI$</th>
<th>$PGA$</th>
<th>$SI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PGV$</td>
<td>0.729 (0.046)</td>
<td>0.733 (0.036)</td>
<td>0.890 (0.047)</td>
</tr>
</tbody>
</table>

Table 2: Parameters in Equation (13) defining the piece-wise variation of the median correlation between $PGV$ and $SA(T)$, $\rho_{tr_PGV,lnSA(T)}$.

<table>
<thead>
<tr>
<th>n</th>
<th>$e_n$</th>
<th>$a_n$</th>
<th>$b_n$</th>
<th>$c_n$</th>
<th>$d_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.73</td>
<td>0.54</td>
<td>0.045</td>
<td>1.8</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.54</td>
<td>0.81</td>
<td>0.28</td>
<td>1.5</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td>0.80</td>
<td>0.76</td>
<td>1.1</td>
<td>3.0</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>0.76</td>
<td>0.70</td>
<td>5.0</td>
<td>3.2</td>
</tr>
</tbody>
</table>
Figure 1: Variation in the ratio of the intra-event and total standard deviation predicted using the Chiou and Youngs (2008b) (CY08) GMPE for all of the ground motions examined.
Figure 2: Examples of the correlation obtained between $PGV$ and $PGA$ using the Boore and Atkinson (2008) (BA08) PGA GMPE: (a) inter-event residuals; and (b) intra-event residuals.
Figure 3: Quantile-quantile plots illustrating the adequacy of the assumption of bivariate lognormality for: (a) inter-event residuals; and (b) intra-event residuals given in Figure 2.
Figure 4: Empirical correlation coefficients obtained between $PGV$ and $PGA$ for: (a) inter-event residuals, $\eta$; (b) intra-event residuals, $\varepsilon$; (c) total residuals, $\varepsilon_T$; and (d) total residuals based on Equation (7). In each figure, box plots represent the uncertainty due to the finite sample size, while the different box plots illustrate the uncertainty due to the selected GMPE.
Figure 5: Correlation coefficients obtained between $PGV$ and: (a) $ASI$; (b) $PGA$; and (c) $SI$. In each figure, box plots represent the uncertainty due to the finite sample size, while the different box plots illustrate the uncertainty due to the selected GMPE.
Figure 6: Correlation coefficients obtained between $PGV$ and $SA$. The solid and dashed lines represent the median and 90% confidence interval due to finite sample size, while the different colored lines illustrate the uncertainty due to the selected GMPE.
Figure 7: Comparisons of the empirical correlations and parametric fit of PGV and SA: (a) median, 16th and 84th percentiles; and (b) standard deviation.
Figure 8: Comparison of the parametric correlation model for $PGV$ and $PGA$ developed in this study with previously computed values from Campbell and Bozorgnia (2008) (CB08), and Abrahamson and Silva (2008) (AS08).
Figure 9: (a) Seismic hazard curve for peak ground velocity; and (b) disaggregation of $PGV$ for 10% exceedance in 50 years ($PGV = 27.7$ cm/s).
Figure 10: Conditional (GCIM) distributions of: (a) PGA; (b) SI; (c) ASI; and (d) SA for various periods given PGV with a 10% exceedance probability in 50 years.