Correlation between Ground Motion Based Shaking Intensity Estimates and Actual Building Damage

A. Tan & A. Irfanoglu
School of Civil Engineering, Purdue University, USA

SUMMARY:
Inspection records from 104,025 buildings surveyed in the aftermath of the 17 January 1994 Northridge, California earthquake (Mw=6.7) along with publicly available ground motion data from that earthquake are used to investigate correlation between ground motion intensity and damage levels, and to find reliable building damage indicators. Damage is represented using three damage levels as prescribed by the inspectors. Each structure is associated with the ground motion parameters obtained from the closest ground motion recording station or the closest geographic grid point provided by the United States Geological Survey (USGS). As the nature of the dependent (damage levels) and some of the independent variables are ordered and integer-valued, besides the regular statistical correlation analysis a random parameter ordered probit statistical model is considered in the study. A critical evaluation of parameters that have strong influence on building damage is provided. The impact of distance to ground motion recording station on observed correlations is also presented.

Keywords: ground shaking intensity, building damage, correlation, ordered probit, Northridge earthquake

1. INTRODUCTION

Estimating spatial distribution and level of building damage is important to design better buildings and assist emergency response effectively following strong earthquakes in urban areas.

Historically, extent and variation of building damage in an area are expressed using earthquake ground shaking intensity scales, such as the Modified Mercalli Intensity scale (Wood and Neumann, 1931). Unfortunately, shaking intensity scales are ambiguous as one’s judgment about damage may differ with that of others’ and with what the structure actually experiences (Irfanoglu and Freeman, 2006). Furthermore, the degree of damage in populations of buildings is considered only qualitatively in shaking intensity scales.

The current approach employed by the USGS to estimate potential for building damage is carried out by estimating the earthquake ground shaking intensity using the automated Instrumental Modified Mercalli Intensity \( I_{mm} \), which is originally based on the MMI measure (Wald et al., 1999). \( I_{mm} \) uses the peak values of ground motion, namely, the peak ground acceleration (PGA) and peak ground velocity (PGV), to assign a ground shaking level as in the Modified Mercalli Intensity (MMI). As rapid as it is, this instrument-only based approach does not account for the structural characteristics of buildings and, therefore, may not provide useful information about the damage state of the built environment following an earthquake. In fact, as with the MMI, \( I_{mm} \) comes with the caveat that “[…] locations within the same intensity area will not necessarily experience the same level of damage since damage depends heavily on the type of structure, the nature of the construction, and the details of the ground motion at that site”
(USGS, 2012). This caveat indicates the fundamental barrier in making $I_{mm}$ an effective measure of extent and variation of damage in a building population.

Studies have shown that the current method in estimating damaged-percentage from the shaking intensity ratings is not accurate (Irfanoglu and Freeman, 2006). The inaccuracy can be attributed mainly to the following: (1) shaking intensity MMI, the scale that forms the basis for $I_{mm}$, is an ambiguous representation of building damage, and (2) structural damage does not solely depend on the ground motion but also on the characteristics of the buildings. Extensive reference to current approach, despite the caveat that accompanies it, heightens the need to search for building damage indicators that take into account both the structural and ground motion characteristics.

2. BUILDING INSPECTION AND GROUND MOTION RECORDS FROM THE 1994 NORTHRIDGE EARTHQUAKE

2.1. Ground Motion Records

Ground shaking from the $M_w=6.7$ 1994 Northridge earthquake was recorded by 185 ground motion stations in the greater Los Angeles area. USGS published PGA, PGV, $I_{mm}$, and 0.3-sec, 1.0-sec and 3.0-sec 5%-damped pseudo spectral accelerations (PSA) of these 185 free-field stations. USGS processed and interpolated the aforementioned ground motion parameters to provide them in uniformly-spaced grid locations (USGS, 2009b), which will be referred to as pseudo stations.

2.2. Building Inspection Records

In the aftermath of the 1994 Northridge earthquake, structures were inspected and ranked according to severity of damage they had sustained. Three damage levels were used: severely damaged or collapsed (no-entry; red-tagged), partially damaged (restricted entry; yellow-tagged), and not damaged (green-tagged) (OES, 1995). The records from the Governor’s Office of Emergency Services (OES) of the state of California indicated that 104,025 buildings in the Los Angeles metropolitan area were inspected. 82,291 of these buildings were green-tagged, 10,440 were yellow-tagged, and 2,633 building were given red tags. 8,661 of the listed buildings were not assigned any damage level. The OES report also provided further information about the buildings, such as location coordinates of the buildings, building year of construction, and type of construction, such as steel, reinforced concrete, masonry or wood-frame. Detailed information about the building database is provided in the referenced OES report (OES, 1995).

The following are parameters that are considered in the OES 1994 Northridge earthquake building damage database and which are of interest for this study: (1) construction material, (2) year of construction, (3) number of stories, (4) the distance from the earthquake epicenter, and (5) shape classifications and quality ratings as judged by the inspectors, (6) damage level. AH-531 (AH, 2010) and ATC-20 (ATC, 2001) provide full descriptions of these parameters.

3. DATA ANALYSIS APPROACH

The building damage severity levels are ordered. One of the suitable statistical approaches to analyze ordered outcomes is the ordered probit model (Greene, 2008). The central idea of the ordered probit model is that there is a latent continuous metric underlying the ordinal responses observed by the analyst. Estimable parameters (referred to as thresholds) partition the continuous variable values into a series of regions corresponding to the various ordinal categories. The latent continuous variable $y^*$ is a linear combination of some predictors $x$ plus a disturbance term that has a standard Gaussian distribution. Like
the models for binary data, one is concerned with how changes in the predictors translate into the probability of observing a particular ordinal outcome (Jackman, 2000).

For this study, ordered probit model is especially appropriate compared to ordinary least square (OLS) regression mainly because of two reasons: (1) in ordered data, conventional discrete outcome models such as standard multinomial probit can result in the loss of estimation efficiency (Shafizadeh and Mannering, 2006; Washington et al., 2011); and (2) an ordered scale may actually represent the alternatives which a decision point may have in a non-temporal hierarchy of independent variables (McKelvey and Zavoina, 1975; Jackman, 2000; Greene, 2008)—for example, the difference between 3 and 2 on the coded building damage severity level (moving from red-tagged to yellow-tagged) can be very different from the difference between 3 and 1 (moving from red-tagged to green-tagged). As a result, using OLS regression might be misleading and its error will be heteroscedastic, i.e. with sub-populations that have different variabilities than others (Duncan et al., 1998). For data visualization purposes, correlation analysis will be conducted along with the random parameter ordered probit model.

The ordered probit model uses the following form:

\[ y^* = \beta^T x + \varepsilon \]  
(3.1)

where \( y^* \) is the dependent variable (damage severity level; coded as 1, 2 and 3 for green-, yellow- and red-tag, respectively), \( \beta \) is the vector of estimated coefficients, \( x \) is the vector of explanatory variables (predictors), and \( \varepsilon \) is the random disturbance. The building damage severity level, \( y \), is related to the underlying latent variable \( y^* \). Using Eqn. (3.1), the observed ordinal data, i.e. damage severity level, \( y \), is defined as:

\[
\begin{align*}
    y &= 1 \text{ if } y^* \leq \mu_0 \\
    y &= 2 \text{ if } \mu_0 < y^* \leq \mu_1 \\
    y &= 3 \text{ if } y^* \geq \mu_1
\end{align*}
\]  
(3.2)

where \( \mu_0 \) an \( \mu_1 \) are the estimable parameters (“thresholds”) that define \( y \). Without loss of generality, in this study \( \mu_0 \) is normalized to zero. Accordingly, from (3.2), the following probabilities are obtained:

\[
\begin{align*}
    P(y = 1) &= \Phi (-\beta^T x) \\
    P(y = 2) &= \Phi (\mu_1 - \beta^T x) - \Phi (-\beta^T x) \\
    P(y = 3) &= 1 - \Phi (\mu_1 - \beta^T x)
\end{align*}
\]  
(3.3)

Where \( P(y = n) \) is the probability of occurrence of response category \( n \), and \( \Phi \) is the standard Gaussian probability distribution function. The interpretation of the primary parameter set of the model, defined as \( \beta \), is as follows: positive signs indicate higher probability of the highest-ordered category, i.e. red-tag \( (y = 3) \), and unambiguously decrease the probability of the lowest-ordered discrete category, green-tag \( (y = 1) \), as the value of the associated variables increase. The negative signs, on the other hand, suggest the opposite. In order to understand the effect of the other damage severity classifications for a particular variable, such as yellow-tag \( (y = 2) \), one must observe their marginal effects. Marginal effects will provide a clearer indication of how changes in specific explanatory variables affect the probabilities of intermediate ordered categories such as yellow-tag \( (y = 2) \). For continuous variables, the marginal effects are computed from the partial derivatives
\[
\begin{align*}
\frac{\partial P(y = 1)}{\partial x} &= -\Phi \left(-\beta^T x\right) \beta \\
\frac{\partial P(y = 2)}{\partial x} &= -\left[\Phi \left(\mu_i - \beta^T x\right) - \Phi \left(-\beta^T x\right)\right] \beta \\
\frac{\partial P(y = 3)}{\partial x} &= \Phi \left(\mu_i - \beta^T x\right) \beta
\end{align*}
\] (3.4)

The interpretation of the coefficients on the marginal effects is similar to the interpretation of the primary coefficients of the model. Marginal effects measure the change in the outcome probability of each threshold category \(P(y = n)\), given a unit change in a continuous variable \(x\). Positive and negative signs of the marginal effects suggest an increase and a decrease in probability, respectively. For dummy variables (variables that only have values of 0 or 1), the marginal effects are computed by changing the dummy variable from 0 to 1 while keeping the other variables constant. In the context of building damage severity levels, a large marginal effect indicates that the coefficient has a relatively large effect on damage severity level, while a small marginal effect indicates a minimal effect on damage severity level.

Ground motion measurements vary across the observations (Zerva and Zervas, 2002). In order to account for any unobserved heterogeneity (unobserved factors that may vary across observations) in the data, the random parameters model is applied to the ground motion measurement. In particular, if the parameters are known to be varying across the observations, constraining the parameters to be constant when they actually vary across observations can lead to inconsistent, inefficient, and biased parameter estimates (Greene, 2008; Washington et al., 2011). Greene (2008) developed estimation procedures for incorporating random parameters in the ordered probit modeling scheme by letting

\[
\beta_i = \beta + u_i
\] (3.5)

where \(\beta_i\) is the vector of individual-specific parameters, and \(u_i\) is the randomly distributed term. Estimation of the random parameters model is done by simulated maximum likelihood estimation using a Halton sequence approach (Washington et al., 2011).

4. OBSERVED RELATIONSHIPS

Ground motion and building damage information in the cities of Santa Monica and Newhall are used to illustrate various observations.

4.1. Correlation between PGA, PGV and 0.3-sec 5%-damped PSA and Spatial Distribution of Building Damage in Santa Monica and Newhall Areas

As seen in Fig. 1, \(I_{mm}\) fails to show sensible trend in predicting damaged-percentage (percentage of buildings damaged in a given population of buildings) with the increase in the shaking intensity in Santa Monica, which was 25 km to the south of the epicentral region. The western part of Santa Monica, where high incidence of damaged buildings was observed, experienced lower \(I_{mm}\) compared to the eastern part of Santa Monica where lower incidence of damaged buildings was observed. The PGV (Fig. 2) experienced in the Santa Monica area was fairly uniform. In this case, PGV was unable to predict the spatial distribution of building damage. PGA (Fig. 3), on the other hand, was more consistent in predicting damaged-percentage with increase in the PGA in Santa Monica. Similar to the case with PGA, 0.3-sec 5%-damped PSA (Fig. 4) was also consistent with the spatial distribution of building damage in the area. 0.3-sec 5%-damped PSA was observed to be better predictor than \(I_{mm}\), PGA, and PGV.
Figure 1. $I_{\text{sh}}$ contours overlaid building damage distribution in Newhall (left) and Santa Monica (right) areas. Note: red dots represent red- and yellow-tagged buildings while blue dots represent green-tagged buildings.

Figure 2. PGV (cm/sec) contours overlaid building damage distribution in Newhall (left) and Santa Monica (right). Note: red dots represent red- and yellow-tagged buildings while blue dots represent green-tagged buildings.

Figure 3. PGA (%g) contours overlaid building damage distribution in Newhall (left) and Santa Monica (right) areas. Note: red dots represent red- and yellow-tagged buildings while blue dots represent green-tagged buildings.
Newhall area, 21 km to the north of the epicentral region, on the other hand, experienced fairly uniform Imm, PGV, and PGA. Subsequently, these parameters were unable to predict the spatial distribution of building damage in the Newhall area. Among the four ground motion parameters, 0.3-sec 5%-damped PSA was the only ground motion parameter that was able to predict the damaged-percentage with increase in the corresponding PSA (Fig. 4).

Although 0.3-sec 5%-damped PSA was able to predict the spatial distribution of building damage in local areas, it cannot predict the distribution of damaged-percentage regionally. Newhall area experienced higher Imm, PGA, PGV, and 0.3-sec 5%-damped PSA compared to the Santa Monica area. However, the damaged-percentage in the Santa Monica area was higher compared to that in the Newhall area. It is clear that there are distinct characteristics retained within the two local areas. In fact, if the damaged-fractions of all inspected structures were plotted against Imm and 0.3-sec 5%-damped PSA, the trend was found to be negative and flat, respectively (Fig. 5). Consequently, it is clear that these characteristics must be taken into account when producing building damage likelihood maps following strong earthquakes.

Figure 4. 0.3-sec 5%-damped PSA (%g) contours overlaid building damage distribution in Newhall (left) and Santa Monica (right). Red dots represent red- and yellow-tagged buildings; blue dots represent green-tagged buildings

Figure 5. Damaged-fractions (red- and yellow-tagged combined) plotted against Imm (left) and 0.3-sec 5%-damped PSA (right) for all inspected structures

4.2. Predicting Spatial Distribution of Building Damage using both Ground Motion Parameters and Building Characteristics

Random parameter ordered probit model was used to find the parameters that can predict the spatial distribution of building damage given ground motion parameters and building characteristics. In order to
account for the site-specific characteristics and the spatial variability of the ground motions, the random parameters chosen were the constant and the ground motion parameters. Due to limited information provided in the inspection records for steel, reinforced concrete and masonry structures, the ordered probit model was applied only to wood-frame structures. Table 4.1 shows the results from the random parameters ordered probit model of I\text{mm} and 0.3-sec 5%-damped PSA. Detailed correlation analyses for various parameters are provided in Tan (2012). Results from those analyses are also consistent with the results obtained from the random parameter ordered probit model.

The results from the ordered probit model indicate that among all ground motion parameters studied, 0.3-sec 5%-damped PSA is the ground motion parameter that is most significant as a damage indicator. PGA and PGV are found to be insignificant. I\text{mm} is found to be slightly more significant that PGA and PGV but less reliable compared to 0.3-sec 5%-damped PSA.

Structures that experience higher 0.3-sec 5%-damped PSA are more likely to be severely damaged. Marginal effects show that a unit increase of 0.3-sec 5%-damped PSA (in %g) increases the probability of being severely damaged (red-tagged) and the probability of being partially damaged (yellow-tagged). Subsequently, the probability of having no damage, i.e. green-tagged, decreases, when the 0.3-sec 5%-damped PSA increases. Details of marginal effects of the random parameter ordered probit model are provided in Tan (2012).

Regarding the physical characteristics of the buildings, taller buildings (up to three stories) are found to be more vulnerable to strong motions. Two- and three-story buildings have higher probability of being damaged compared to single-story buildings. There is insufficient data to model four-story or taller buildings. Shape classifications, quality ratings as judged by the inspectors, and the distance from the earthquake epicenter are not reliable building damage potential indicators.

The results further indicate that year of construction is a significant damage predictor. On average, the probability of being severely damaged is higher in structures built between 1910 and 1930, while lower for structures built between 1940 and 1960 is lower on average. Structures built between 1900 and 1910, between 1930 and 1940, and between 1960 and 1970 performed at about average. Structures built after 1970 have a slightly lower probability of being severely damaged, on average.

The results indicate that although the log-likelihood of the 0.3-sec 5%-damped PSA was observed to be superior compared to the log-likelihood of the I\text{mm}, the two log-likelihoods do not differ by a significant amount. The results further indicate that the weights, i.e. the estimated coefficients, of the independent variables from the building characteristics between the two models do not differ significantly.

Fig. 6 shows the damaged-fractions plotted against the probability of damage estimated using I\text{mm} or 0.3-sec 5%-damped PSA. The probability of damage was estimated using the estimated coefficients from the random constant and random I\text{mm} or random 0.3-sec 5%-damped PSA ordered probit model, taken at 350-meter distance from the pseudo stations. In this plot, the model is applied to all inspected buildings located within 350-meter distance from the pseudo stations that had inspection records listing the parameters considered in this study. The visualization of the results indicates that 0.3-sec 5%-damped PSA is superior to I\text{mm} as a parameter to predict potential building damage. This model is much better in predicting the spatial distribution of building damage compared to those solely using I\text{mm} or 0.3-sec 5%-damped PSA as potential building damage indicators (Fig. 5).

The imperfections of the probability plot of 0.3-sec 5%-damped PSA, such as the mismatches between the predicted probability and the damaged fractions, are expected. These imperfections were caused by ignoring any known or unknown parameters that may play role in predicting building damage, such as possible aftershock damage or information about the building retrofits. The building characteristics that
were taken into account and found to be statistically significant in predicting the spatial distribution of building damage are only (1) the number of stories and (2) the year of construction. Other parameters should be searched in order to obtain a better estimate of building damage likelihood for the Greater Los Angeles area.

### Table 4.1. Random Constant and Random PSA Ordered Probit Model Results for Wood-frame Buildings Located within 350 m from the Pseudo Stations

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>0.3-sec 5%-damped PSA (%g)</th>
<th>( I_{\text{mm}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>Estimated coefficient</td>
<td>t-statistic</td>
</tr>
<tr>
<td></td>
<td>-1.88</td>
<td>-33.17</td>
</tr>
<tr>
<td><em>Standard deviation of parameter distribution</em></td>
<td>0.78</td>
<td>40.10</td>
</tr>
<tr>
<td>Ground Motion Parameters</td>
<td>0.0038</td>
<td>9.21</td>
</tr>
<tr>
<td><em>Standard deviation of parameter distribution</em></td>
<td>0.0011</td>
<td>7.70</td>
</tr>
<tr>
<td>Number of stories</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-story</td>
<td>0.24</td>
<td>8.12</td>
</tr>
<tr>
<td>Three-story</td>
<td>0.69</td>
<td>11.20</td>
</tr>
<tr>
<td>Year of construction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between 1910-1920</td>
<td>0.44</td>
<td>6.49</td>
</tr>
<tr>
<td>Between 1920-1930</td>
<td>0.41</td>
<td>9.98</td>
</tr>
<tr>
<td>Between 1950-1960</td>
<td>-0.28</td>
<td>-7.48</td>
</tr>
<tr>
<td>After 1970</td>
<td>-0.15</td>
<td>-3.35</td>
</tr>
<tr>
<td>Model parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold parameter ( \mu_1 )</td>
<td>1.15</td>
<td>44.86</td>
</tr>
<tr>
<td>Number of observations</td>
<td>5468</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-2631.3</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 6.** Red-tagged fractions plotted against the probability of the building being severely damaged. The numbers next to the corresponding markers represent the number of buildings fallen within a probability interval.

#### 4.3. Spatial Variability of 0.3-sec 5%-damped PSA

Spatial variation of seismic ground motion describes the differences in ground motion between various locations. The variability can be mainly attributed to the wave passage, loss of coherence, and different local soil conditions (Saxena et al., 2000). There are a minimum number of samples necessary to capture the relationships between damage, ground motion and structural characteristics confidently. Given the
density of the recording stations and the dwellings in an area, there are trade-offs associated with the choice of the distance to a ground motion from a recording station. Greater distance will give us more data; however, the accuracy of the ground motion will decrease as the samples are located further away from the recording stations.

As shown in Fig. 7, the coefficient of variation for spatial variability of 0.3-sec 5%-damped PSA coefficient changes from 20% to 220% between points located at 250 meters and 450 meters away from the pseudo stations, respectively. From this study, it is observed that the reasonable cut-off distance for the greater Los Angeles area is determined to be no more than 350 meters. At 350 m, the COV of the spatial variability of the 0.3-sec 5%-damped PSA coefficient was found to be about 30%.

![Figure 7](image_url)

**Figure 7.** Estimated coefficients and standard deviations of the 0.3-sec 5%-damped PSA at various distances from the pseudo stations of wood-frame structures. The number of samples within each interval is also shown.

5. CONCLUSIONS

This study focused on identifying building and ground motion characteristics which could be used as reliable estimators of building damage. The 1994 Northridge, California earthquake building inspection dataset and publicly available ground motion data from 137 recording stations (out of 185 reported by USGS) as well as USGS-generated interpolated ground motion parameter maps (referred as to pseudo stations, or grid) were used in the study. Assumptions made are: (1) possible additional damage caused by the aftershocks from the main 1994 Northridge earthquake is neglected, (2) uncertainties associated with the bias and inconsistency of the inspectors when they inspected the buildings are neglected, and (3) error distribution of the independent variables are assumed to follow a Gaussian distribution. A simple correlation analysis and random-parameter ordered probit models were used to analyze the data. The following are the conclusions from this study:

1. Instrumental Modified Mercalli ($I_{mm}$) intensity, peak ground acceleration (PGA), peak ground velocity (PGV), and 0.3-sec 5%-damped pseudo spectral acceleration (PSA), by themselves, are not reliable damage indicators.
2. When used in conjunction with building characteristics, 0.3-sec 5%-damped PSA was observed to be the most reasonable damage indicator. PSA is most useful if the period considered for the spectral response is the median or average of the fundamental periods of the chosen group of structures in the region of interest. In this study, 0.3 sec was the appropriate period. The equivalent viscous damping ratio, on the other hand, was taken to be 5% as PSA values for that damping ratio were readily available. It must be noted that the spectral response ordinates alone are not enough to predict the spatial distribution of building damage accurately.
3. During the 1994 Northridge earthquake, dwellings built between 1910 and 1930 were more likely to be damaged. Dwellings built between 1940 and 1960 were less likely to be damaged. On average, structures built after 1960 performed worse than those built between 1940 and 1960, but performed better compared to those built between 1910 and 1930. Overall, there is a trend of better building performance from early 1900s structures to modern structures.

4. Two- and three-story dwellings were more likely to be damaged compared to one-story dwellings. Due to sample set limitations, no judgment can be made for structures taller than three stories.

5. It was observed that distance from earthquake epicenter as well as building shape classifications and quality ratings as judged by the inspectors are not reliable building damage indicators.

6. This study has shown that the coefficient of variation for spatial variability of 0.3-sec 5%-damped PSA coefficient can reach 220% for points located at 450 meters away from the pseudo stations. A reasonable cut-off distance for the greater Los Angeles area was observed to be no more than 350 meters.

REFERENCES


Joint OES-FEMA Disaster Field Office. (1994). Building and Safety Damage Assessment (Northridge Earthquake Disaster DR-1008), California Governor’s Office of Emergency Services, Sacramento, CA.


