Selection of ground motion time series and limits on scaling

Jennie Watson-Lamprey a,*, Norman Abrahamson b

a Department of Civil and Environmental Engineering, University of California at Berkeley, Berkeley, CA 94720, USA
b Pacific Gas & Electric, 245 Market St., San Francisco, CA 94129, USA

Accepted 20 July 2005

Abstract

A procedure to select time series for use in non-linear analyses that are intended to result in an average response of the non-linear system is proposed that is not based simply on magnitude, distance, and spectral shape. A simple model of a yielding system is used as a proxy for the non-linear behavior of a more complicated yielding system. As an example, Newmark displacements are used as a proxy for more complex slope-stability models. The candidate scaled time series are evaluated to find those that yield a response of the simple non-linear system that is near the expected response for the design event. Those scaled time series with responses near the expected value are selected as the optimum time series for defining average response even if the scale factors are larger than commonly accepted (e.g. scale factors > factor of 2).

Keywords: Newmark displacement; Time series; Selection; Non-linear; Slope-stability; Scaling

1. Introduction

In many seismic analyses, ground motion time series are required in addition to the design response spectrum. It is common practice to select empirical recordings of ground motion and scale these ground motions to the level of the design spectrum. The selection of the records and the amount of scaling that can be applied remains controversial.

Typically, the time series is selected from recorded ground motions with similar magnitudes and similar distances to the design earthquake. Other factors such as the site condition, style-of-faulting, and spectral content may also be considered. Additionally, the scale factor required to scale the time series to the design spectrum may also be considered. In general, scale factors closer to unity are preferred and many ground motion experts recommend a limit on the amount of scaling applied. Recommended limits on scaling typically range from factors of 2 to 4 ([1]). These limits are generally based on the comfort level of the engineer and not on quantitative evaluations of scaling.

In this paper, we propose a method for selecting time series such that after scaling, they will lead to a near average response of a non-linear system. The approach is to first find a simple non-linear model that can serve as a proxy for a more complicated non-linear model and then using the simple non-linear model, find the time series that have properties that lead to a average response of the simple model. The main concept is that using a simple non-linear model, we can evaluate a large number of candidate time series and from this large set identify those that lead to average response. It also allows the evaluation of the limits of scaling of ground motions.

First, we evaluate the scaling using traditional selection of ground motion time series based only on the magnitude, distance, and site condition. We then develop an improved procedure for selecting records that is less sensitive to scaling.

In addition to scaling a time series to a design spectrum, another common approach is to modify the ground motion recordings to match the entire response spectrum ([1]). Selection of records for use in developing spectrum compatible ground motion is not considered in this study.

2. Data base

The Pacific Earthquake Engineering Research (PEER) Center, NGA strong motion database ([2]) was used for this study. This database contains 7075 records from 175 earthquakes. For this study we reject earthquakes from subduction zones and Northeastern California and use only freefield records. The subset of data includes magnitudes from 4.5 to 7.9 and distances from 0 to 300 km from 6158 records.
3. Scaling for selection by Mag–Dist bins

Before developing the new approach, the traditional approach of selecting time series based on magnitude, distance, and site condition is considered. The limits of scaling records are evaluated for just a single design case. In this example 103 deep soil site recordings were selected that lie in the magnitude range 6.5–7.0 and the rupture distance (Rrup) range 0–15 km. Three scale parameters are considered: PGA, PGV, and Arias Intensity. Travasarou showed that Arias intensity is a strong predictor of displacement response of earth structures ([3]). The average values for PGV, PGA and Arias Intensity were determined for the group, the individual recordings were then scaled to match each of these average values: PGA $Z_{0.44 \text{ g}}$, PGV $Z_{50 \text{ cm/s}}$, Arias Intensity $Z_{0.18 \text{ g}^2 \text{ s}}$. The Newmark displacements were calculated for a yield acceleration of 0.1 g. To evaluate the limits for which the scaling is valid, the resulting displacements are plotted against the scale factor in Figs. 1–3 for PGA, PGV, and Arias Intensity, respectively.

For all three scaling parameters, the displacement is correlated with the scale factor. This indicates that the computed displacement will be biased if large (or small) scale factors are used. For example, using Arias Intensity scaling, there is a 67 % under-prediction of the displacement for a scale factor of 3. Similarly, using PGV scaling, there is a 24 % over-prediction of displacement for a scale factor of 3.

A second observation is that there is a large variability in the computed displacements for records with similar scale factors, so even if a recording is selected that requires only a small amount of scaling (factor close to unity), that recording may not give a good estimate of the average displacement. Since it is common in practice to use a small number of recordings (typically 1–3), the computed response will be sensitive to the recording selected. Lastly, there are some records with large scale factors that produce displacements close to the average, so it is possible to get unbiased results even for large scale factors, but we need to be smarter in selecting records. To address this issue, we first develop a model for the Newmark displacement based only on the properties of the time series. We then use this model to guide the selection of records.

4. Model of Newmark displacements

A regression analysis was conducted to develop a model for the Newmark displacement based solely on characteristics of the ground motion time series. The 6158 recordings from the data set were scaled by seven different scale factors: 0.5, 1, 2, 4, 8, 16 and 20. The Newmark displacements of these scaled recordings were then computed for three yield accelerations: 0.1 g, 0.2 g, and 0.3 g. This results in a total set of 129,318 displacements.

This set of displacements were fit to the following model

$$\ln(\text{Newmark Disp (cm)})$$

$$= \left(a_1 + b_1(\ln(Sa_T=1) - 0.45) + b_2(\ln(Sa_T=1) - 0.45)^2 + c_1(\ln(A_{RMS}) - 1.0) + d_1(\ln(Sa_T=1)/\text{PGA}) + d_2(\ln(Sa_T=1)/\text{PGA})^2 + e_1(\ln(Dur_k) - 0.74) + e_2(\ln(Dur_k) - 0.74)^2 + \frac{1}{f_1(\ln(\text{PGA}/k_y) - f_2)} \right)$$

(1)
where $S_a$ is spectral acceleration with 5% damping at 1 s in $g$, 
PQA is the peak ground acceleration in $g$, $A_{RMS}$ is the root mean square of acceleration in $g$, and the $Dur_{k_y}$ is the duration for which the acceleration is greater than the yield acceleration in s. The $Dur_{k_y}$ is equal to the total time for which the acceleration exceeds the yield acceleration (in the direction of maximum displacement).

The coefficients estimated using least-squares are given in Table 1. The residuals from this model are shown as a function of scale factor in Fig. 4. This figure shows that the model for the median displacement is unbiased even for scale factors up to 20. The model is slightly biased for scale factors less than one, but these cases are not generally of engineering concern. Therefore, scaling itself does not lead to biased results. This indicates that we can scale by large factors if we select the appropriate recordings.

The standard deviations computed for subsets of the data are compared to the Eq. (1) model in Fig. 5. As with the median, the variability of the Newmark displacement is not dependent on the scale factor therefore there is no fundamental problem with scaling. System response can be predicted based on PGA, $k_y$, $S_a$, $A_{RMS}$ and $Dur_{k_y}$ therefore the appropriate recording can be chosen based on these earthquake characteristics.

### 5. Selection of time series

In practice, only a small number of time series are used in an engineering analysis. If only a small number of time series are used, then the analysis is implicitly focused on the average response of the system, not the variability of the response. Therefore, we want to select the time series to give us the best estimate of the average response of a non-linear system. Traditionally, time series are selected simply based on magnitude, distance, and site condition. Often, the response spectral shape is considered in the selection to avoid unusual spectral shapes ([4]), but it is difficult to anticipate which time series will give an average response to a non-linear system.

The evaluation of the Newmark displacement in the preceding section showed that the Newmark displacement can be modeled with a relatively small variability if four properties of the time series are known: PGA, $S_a$, $A_{RMS}$, and $Dur_{k_y}$. If we select the time series such that after scaling, the time series parameters lead to a median Newmark displacement similar to that expected for the design event, then that time series can be expected to give a near average response for a more complicated slope deformation evaluation. That is, we use the Newmark displacement as a proxy for the non-linear behavior of a more complicated yielding system (in this case, slope stability) in guiding the selection of the time series.

In a typical case, the design ground motion is specified by a response spectrum and an earthquake scenario ($M, R$). In addition, the site classification, in terms of $V_{S30}$, is also known. To use the Newmark displacement model, we need three additional parameters: PGV, $A_{RMS}$, and $Dur_{k_y}$. Using the PEER strong motion data base, models are derived for these parameters given that the typically design ground motion information is available. The $A_{RMS}$ is given by

$$
\ln(A_{RMS}(g)) = -1.167 + 1.02 \ln(\text{PGA})
$$

with a standard deviation of 0.20 natural log units. Finally, the $Dur_{k_y}$ is given by

$$
\ln(Dur_{k_y}(s)) = -2.775 + 0.956(\ln(\text{PGA} / k_y))
$$

$$
- \frac{1.554}{\ln(\text{PGA} / k_y) + 0.390} - 0.597(\ln(\text{PGA}))
$$

$$
+ 0.381 \ln(S_aT=1s) + 0.334 M
$$

![Fig. 4. Newmark displacement model residuals versus scale factor.](image)

![Fig. 5. Newmark displacement model standard deviation residuals versus scale factor.](image)

### Table 1: Newmark displacement model

<table>
<thead>
<tr>
<th>Parameter Estimate Std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
</tr>
<tr>
<td>$b_1$</td>
</tr>
<tr>
<td>$b_2$</td>
</tr>
<tr>
<td>$c_1$</td>
</tr>
<tr>
<td>$d_1$</td>
</tr>
<tr>
<td>$e_1$</td>
</tr>
<tr>
<td>$e_2$</td>
</tr>
<tr>
<td>$f_1$</td>
</tr>
<tr>
<td>$f_2$</td>
</tr>
<tr>
<td>$f_3$</td>
</tr>
</tbody>
</table>
with a standard deviation of 0.52 natural log units. For cases with directivity the Dur\textsubscript{txy} is given by

\[
\ln(\text{Dur}_{txy}(s)) = -2.801 + 0.946(\ln(\text{PGA}/k_y) - 0.392(\ln(\text{PGA})) + 0.84 \ln(S_{A_{T=1s}}) + 0.346M - 0.112X)
\]

\times \text{Taper}

(4)

with a standard deviation of 0.52 natural log units. Where the Taper function is given by:

\[
\text{Taper} = \begin{cases} 
M \leq 6.5 & \to 0 \\
6.5 < M < 7 & \to 2(M - 6.5) \\
M \geq 7 & \to 1 \\
\end{cases}
\]

\[
R_{\text{rup}} \leq 30 \to 1 \\
30 < R_{\text{rup}} < 50 \to 0.05(50 - R_{\text{rup}}) \\
R_{\text{rup}} \geq 50 \to 0 \\
\]

(5)

Given this information, the following procedure can be used to select a time series:

1. Compute the median \(A_{\text{RMS}}\) and Dur\textsubscript{txy} given the design event parameters using the models given above in Eqs. (2) and (3), respectively.
2. Compute the median Newmark displacement for the design event given the PGA, \(S_{A_{T=1s}}\), \(A_{\text{RMS}}\) and Dur\textsubscript{txy} using the model in Eq. (1).
3. Select candidate ground motion recordings for scaling. This can be based on the traditional magnitude and distance bins approach (but the bins can be wider).
4. Scale all acceleration time series to the design event ground motion. This may be done by scaling to the PGA, to the PGV, to the Arias Intensity, or to the spectral acceleration averaged over a period band. Equations for estimating the design event PGV and Arias Intensity are given below.
5. Reject records where \(A_{\text{RMS}}\) and Dur\textsubscript{txy} are not within one-half of a standard deviation of their median values (from step 1).
6. Calculate the difference between the estimated logarithm of the median Newmark displacement for the design event (from step 2) and the logarithm of the median Newmark displacement expected for each scaled ground motion (from step 4) and square.
7. Repeat steps 1–6 for \(k_y\) values of 1.5\(k_y\), \(k_y/1.5\), and \(k_y/2\). This will generate four lists of scaled candidate records.
8. For records that appear on all lists, calculate the root mean square of the differences for the four yield acceleration values. If both components of a record appear on the list then the component with the lower ranking is rejected.
9. Select the record(s), which have the smallest root mean square of differences.

This process allows prioritization of the available recordings for scaling, but it does not determine if any of the records are appropriate. Some maximum misfit could be defined to prevent the selection of an inappropriate recording (which is still the best available).

Travasarou [3] showed that, when parameters are considered individually, Arias Intensity is the most important when predicting Newmark displacement. Our analysis of the data confirms this result. However, when more than one parameter is used, the dependence on Arias Intensity disappears, as seen in Fig. 6. The function for Arias Intensity is given by

\[
\ln(\text{Arias Intensity} (g^2/s)) = [-1.991 + 0.497M - 0.355 \ln(V_{S30}) - 0.081(\ln(R_{\text{rup}}) - 3.4) - 0.018(\ln(R_{\text{rup}}) - 3.4)^2 + 1.296 \ln(\text{PGA}) + 0.328 \ln(S_{A_{T=1s}})]
\]

(6)

with a standard deviation of 0.34 natural log units. The standard deviation of this model is in contrast to Travasarou [3] who provides an attenuation relationship for Arias Intensity with a
The addition of the ground motion level provides significant information for calculating Arias Intensity.

An additional parameter for scaling is PGV. The function for PGV is given by

$$\ln(\text{PGV}(\text{cm}/\text{s})) = \left[4.987 + 0.0429(\ln(\text{Sa}_{T=1s})) - 0.0392(\ln(\text{Sa}_{T=1s}))^2\right]$$

$$+ 0.624(\ln(\text{Sa}_{T=3s})) + 0.0363(\ln(\text{Sa}_{T=3s}))^2$$

$$+ 0.0876M + 0.496(\ln(\text{PGA}))$$

$$+ 0.00871(\ln(\text{PGA}))^2$$

with a standard deviation of 0.25 natural log units.

6. Example application

As an example, consider the design event listed below:

- PGA: 0.6g
- SA$_{T=1s}$: 1.1g.
- M: 7
- V$_{30}$: 400 m/s
- X: 0.5

The response spectrum is shown in Fig. 7.

For this site, assume that the yield acceleration is 0.1 g.

**Step 1.** Using the design event parameters, the Durky and $A_{RMS}$ are:

- Dur$_{ky}$ = 2.411 s
- $A_{RMS}$ = 0.185 g

**Step 2.** Using the parameters from step 1 and the PGA from the design spectrum, the median Newmark displacement is computed using Eq. (1):

$$\text{Design event Newmark displacement} = 57.3 \text{ cm}$$

**Step 3.** Select candidate recordings. For this example, we used a very wide range of records to show that a wider range than typically considered is appropriate.

- Mag range: M6.0–7.9
- Distance: 0–50 km

### Table 2

<table>
<thead>
<tr>
<th>Event</th>
<th>Station name</th>
<th>EQ ID#</th>
<th>Mag.</th>
<th>$R_{rup}$ (km)</th>
<th>$V_{30}$ (cm/s)</th>
<th>Scale factor</th>
<th>Scaled PGA</th>
<th>Scaled $\text{Sa}_{T=1s}$</th>
<th>Scaled Durky</th>
<th>$A_{RMS}$</th>
<th>Expected Newmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design event</td>
<td></td>
<td>–</td>
<td>7</td>
<td>5</td>
<td>400</td>
<td>–</td>
<td>0.6</td>
<td>1.1</td>
<td>2.41</td>
<td>0.19</td>
<td>57</td>
</tr>
<tr>
<td>Chalfant valley—02 1986</td>
<td>Benton</td>
<td>103</td>
<td>6.19</td>
<td>22</td>
<td>271</td>
<td>3.57</td>
<td>0.63</td>
<td>1.1</td>
<td>2.59</td>
<td>0.18</td>
<td>61</td>
</tr>
<tr>
<td>Loma prieta 1989</td>
<td>APEEL—Pulgas</td>
<td>118</td>
<td>6.93</td>
<td>42</td>
<td>415</td>
<td>3.87</td>
<td>0.61</td>
<td>1.1</td>
<td>2.81</td>
<td>0.17</td>
<td>64</td>
</tr>
<tr>
<td>North ridge—01 1994</td>
<td>Leona Valley #2</td>
<td>127</td>
<td>6.69</td>
<td>37</td>
<td>446</td>
<td>7.10</td>
<td>0.65</td>
<td>1.1</td>
<td>2.60</td>
<td>0.18</td>
<td>61</td>
</tr>
<tr>
<td>North ridge—01 1994</td>
<td>Leona Valley #3</td>
<td>127</td>
<td>6.69</td>
<td>37</td>
<td>685</td>
<td>6.15</td>
<td>0.52</td>
<td>1.1</td>
<td>2.34</td>
<td>0.18</td>
<td>55</td>
</tr>
<tr>
<td>North ridge—01 1994</td>
<td>Littlerock—Braider canyon</td>
<td>127</td>
<td>6.69</td>
<td>47</td>
<td>822</td>
<td>7.45</td>
<td>0.54</td>
<td>1.1</td>
<td>2.29</td>
<td>0.18</td>
<td>54</td>
</tr>
<tr>
<td>Coalinga—01 1983</td>
<td>Parkfield—VC4W</td>
<td>76</td>
<td>6.36</td>
<td>35</td>
<td>376</td>
<td>10.91</td>
<td>0.50</td>
<td>1.1</td>
<td>2.68</td>
<td>0.17</td>
<td>62</td>
</tr>
<tr>
<td>Coalinga—01 1983</td>
<td>Parkfield—FZ 16</td>
<td>76</td>
<td>6.36</td>
<td>28</td>
<td>339</td>
<td>3.53</td>
<td>0.69</td>
<td>1.1</td>
<td>2.67</td>
<td>0.19</td>
<td>64</td>
</tr>
<tr>
<td>Chi—03 1999</td>
<td>TCU054</td>
<td>172</td>
<td>6.2</td>
<td>37</td>
<td>461</td>
<td>12.59</td>
<td>0.50</td>
<td>1.1</td>
<td>2.13</td>
<td>0.20</td>
<td>54</td>
</tr>
<tr>
<td>North ridge—01 1994</td>
<td>Wrightwood—Jackson Flat</td>
<td>172</td>
<td>6.69</td>
<td>49</td>
<td>376</td>
<td>8.65</td>
<td>0.77</td>
<td>1.1</td>
<td>2.66</td>
<td>0.19</td>
<td>63</td>
</tr>
<tr>
<td>Chi—04 1999</td>
<td>CHY02</td>
<td>173</td>
<td>6.2</td>
<td>20</td>
<td>428</td>
<td>7.44</td>
<td>0.53</td>
<td>1.1</td>
<td>2.83</td>
<td>0.17</td>
<td>66</td>
</tr>
</tbody>
</table>
All soil types
In all, 1532 records are considered.

Step 4. In this example, we scale the records to the design $S_{aT=1s} = 1.1g$.

Step 5. Reject records whose $A_{RMS}$ and $Dur_{\xi}$ lie outside of the one-half standard deviation bounds. The standard deviations for $A_{RMS}$ and $Dur_{\xi}$ are 0.221 and 0.552, respectively. For this example, records considered lie within the range:

- $A_{RMS}: 0.167-0.204$
- $Dur_{\xi}: 1.86 - 3.13$

In all, 38 records are considered.

Step 6. Using the model in Eq. (1), the expected displacement for each recording is computed. The difference in the logarithm of the expected displacement and the logarithm of the expected median displacement for the design event is computed and then squared. These are shown in Fig. 8 as a function of magnitude. Note that there are good agreements to the design case for a wide range of magnitudes.

Step 7. Steps 1–7 are repeated for yield acceleration values of 0.15g, 0.0667g and 0.05g. Four lists of candidate recordings are generated.

Step 8. For each recording that appears on the four candidate lists the root mean square of the differences for each yield acceleration is calculated. The recordings are then ranked based on these differences. The ten records with the smallest difference from step 6 are the least are shown in Table 2 and Fig. 9.

Step 9. The records with the smallest difference from step 8 are chosen. The magnitudes of the earthquakes range from 6.19 to 6.93 and the distances range from 31 to 62 km. This is a wider range than is typically considered. The scale factors range from 3.5 to 12.6 indicating that some time series requiring large scale factors may still provide accurate estimates of the average response. A common feature of the time histories is that after scaling, they have similar durations above the yield acceleration.

7. Conclusion

Limits to scaling are appropriate if the candidate time series are selected based only on the magnitude, distance, and site condition; however, if a selection criteria based on the properties of the scaled ground motion is used rather than just selecting time histories based on magnitude, distance, and site then time series can be scaled by large factors and still lead to an average response. Using the record properties also allows us to broaden the range of time series typically considered (broader magnitude, distance, and site condition range).

References