

SEMI-AUTOMATED SELECTION AND SCALING OF EARTHQUAKE GROUND MOTIONS

Albert R. KOTTKE¹, Ellen M. RATHJE²

ABSTRACT

Suites of earthquake ground motions are critical parts of the seismic design and analysis process. The selection of these ground motions typically requires extensive trial of combinations by hand. The proposed methodology replaces the trial of combinations by hand with a semi-automated process that selects and scales ground motions to fit a target acceleration response spectrum and standard deviation. The methodology selects motions for the suite based on matching the target spectral shape, and then fits the amplitude and standard deviation of the target by adjusting the scale factors. The trial of combinations is performed through both a rigorous method that tries each possible combination and an iterative approach that selects one motion at a time based on the best fit to the spectral shape. While the rigorous approach may determine the best combination, the time required to compute every combination may be too long to be practical. The developed methods are used to generate a number of possible combinations from which an engineer selects the final suite. An example of the process is carried out for both the rigorous and iterative procedures.

Keywords: ground motion selection, target response spectrum, motion scaling, seismic design

INTRODUCTION

The ground motions produced by earthquakes have the potential to significantly damage the built infrastructure; thus, in the seismic design process engineers use recorded ground motions in numerical simulations to establish the effects of the anticipated seismic loading on structures. The significant natural variability in earthquake ground motions requires engineers to use a suite of ground motions in their analysis.

To select a ground motion suite for a deterministic event, a target response spectrum is defined for the scenario event and a catalog of potential motions is developed. The scenario event is described based on the earthquake magnitude, site-to-source distance, and faulting mechanism, as well as the site conditions. Using a ground motion prediction relationship and the description of the scenario event, the target response spectrum and its associated standard deviation are computed. The catalog of potential motions is selected from a strong motion database from previous earthquakes that approximate the scenario event. A suite of motions is then selected from the catalog and scaled so that the median of the suite best fits the target response spectrum.

Traditionally, the selection of the suite of ground motions is performed by hand and suffers from three problems: (1) the "goodness-of-fit" of the combination is not quantified, (2) there is no direct consideration of the aleatory variability (standard deviation), and (3) the method requires tedious trials of combinations by hand.

¹ Graduate Student, Department of Civil, Architectural, and Environmental Engineering, The University of Texas at Austin. Email: albert@mail.utexas.edu

² Associate Professor, Department of Civil, Architectural, and Environmental Engineering, The University of Texas at Austin. Email: e.rathje@mail.utexas.edu

In the proposed semi-automated methodology, ground motions are selected and scaled to fit both the target response spectrum and its standard deviation. The trial of ground motions is done by an automated algorithm, which avoids the time-consuming trial of combinations by hand, and a quantitative parameter is used to evaluate the goodness-of-fit. The developed method only considers the fit to the target response spectrum, which may fail to capture other important characteristics of the suite. To account for this deficiency the method provides to the user several potential suites of motions from which the final combination must be selected. The reliance on the engineer to make the final decision is important as the experience and judgment of an engineer can not be captured in an algorithm.

SCALING OF RECORDED-TIME HISTORIES

Average Scale Factors

Both the effect of the scale factor and a method to judge the fit of the resulting average response spectrum must be defined before selecting recorded acceleration-time histories to fit a target acceleration response spectrum. An acceleration-time history and its associated acceleration response spectrum, Sa , are scaled by multiplying each value by a single scale factor, s . This multiplication produces a scaled acceleration response spectrum, $Sa^{scaled} = s \cdot Sa$, which also represents the acceleration-time history multiplied by the scale factor, s . To find the median response spectrum of a combination of scaled response spectra, the mean value of the natural logarithm of the spectral accelerations at each period is computed by:

$$\ln Sa_{avg,i}^{scaled} = \frac{1}{n_m} [\ln(s_1 \cdot Sa_{1,i}) + \ln(s_2 \cdot Sa_{2,i}) + \dots + \ln(s_{n_m} \cdot Sa_{n_m,i})] \quad (1)$$

where i represents the i -th period and n_m is the number of motions in the combination. Note that average and median are used interchangeably in this paper, but always represent calculation in natural log space. Using the properties of logarithms—specifically $\ln(a \cdot b) = \ln a + \ln b$ —the average response spectrum can be expanded to:

$$\ln Sa_{avg,i}^{scaled} = \frac{1}{n_m} \underbrace{(\ln s_1 + \ln s_2 + \dots + \ln s_{n_m})}_{\text{Controls Amplitude}} + \frac{1}{n_m} \underbrace{(\ln Sa_{1,i} + \ln Sa_{2,i} + \dots + \ln Sa_{n_m,i})}_{\text{Controls Shape}} \quad (2)$$

Through this expansion, the scale factors can be grouped separately from the original recorded response spectra. The separation of the scale factors from the recorded response spectra indicates that the scale factors control only the amplitude, while the recorded response spectra control the shape. Using this information, motion combinations are selected that best fit a target response spectral shape, while the scale factors are used to adjust the amplitude of the average response spectrum to fit the target response spectrum. The expansion can be simplified to an average scale factor applied to an average response spectrum in natural log space:

$$\ln Sa_{avg,i}^{scaled} = \frac{1}{n_m} \sum_{j=1}^{n_m} \ln s_j + \frac{1}{n_m} \sum_{j=1}^{n_m} \ln Sa_{j,i} = \ln s_{avg} + \ln Sa_{avg,i} \quad (3)$$

For a given group of motions with a computed $\ln Sa_{avg,i}$, the individual scale factors, s_j , can be varied without changing $Sa_{avg,i}^{scaled}$ as long as $\ln s_{avg}$ remains constant. The individual scale factors, s_j , are used to match the amplitude and standard deviation of the target spectrum.

The optimal average scale factor, $\ln s_{avg}$, for the scaled combination is found by calculating the mean difference between the average (unscaled) spectra and target response spectrum:

$$\ln s_{avg} = \frac{1}{n_p} \sum_{i=1}^{n_p} [\ln(Sa_{target,i}) - \ln(Sa_{avg,i})] \quad (4)$$

where n_p is the number of periods in the response spectrum. The optimal scale factor represents the least-squares fit of the average response spectrum to the target response spectrum. The success of the scaled fit can be quantified by computing the sum-of-squares error between the scaled average response spectrum and the target response spectrum:

$$SSE = \sum_{i=1}^{n_p} \left(\ln s_{avg} + \ln Sa_{avg,i} - \ln Sa_{target,i} \right)^2 \quad (5)$$

However, the SSE is dependent on the number of periods used in the response spectrum. To account for the number of periods used, the SSE is normalized by the number of periods and called the mean-of-squares error (MSE):

$$MSE = \frac{1}{n_p} \cdot SSE \quad (6)$$

If too few periods are used to describe the recorded acceleration response spectrum, the computed MSE may be erroneously high or low. A stable value of MSE is obtained if more than 60 periods, equally spaced in log space between 0.01 and 5 seconds, are used. In this study 120 periods are used in all response spectra.

Individual Scale Factors

The motions within a selected combination (suite) could be scaled by applying s_{avg} to each of the motions, but the resulting scaled combination would have the same standard deviation as the unscaled combination. Figure 1a shows seven motions selected to fit a target spectrum. The median response spectrum of the unscaled motions and the median response spectrum of the motions all scaled by s_{avg} are shown. While the median response spectrum of the motions fits the target spectrum well, the variability in the selected motions is greater than the target variability, as shown by the $\pm\sigma$ spectra shown in Figure 1b. To adjust the standard deviation (σ_{ln}) of the combination, each motion requires a separate scale factor such that the standard deviation of the combination increases or decreases to match the target standard deviation without changing the median response spectrum. The scale factors for the individual motions (s_j) are calculated by combining the average scale factor (s_{avg}) with a variability term (σ_j):

$$\ln s_j = \ln s_{avg} + \sigma_j \quad (7)$$

As long as the mean value of the variability term is zero, the mean of the individual scale factors ($\ln s_j$) will be equal to $\ln s_{avg}$. However, for each motion, the variability term, σ_j , must be established.

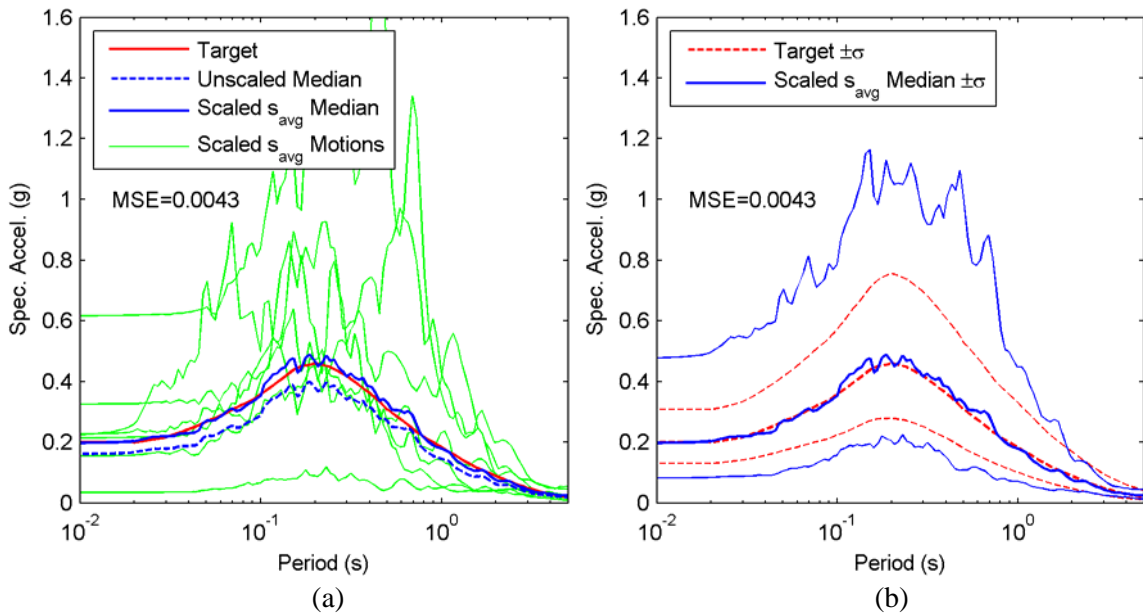


Figure 1. (a) Individual motions scaled by s_{avg} , (b) Median and median $\pm\sigma$ of motions scaled by s_{avg} .

One method of establishing the values of s_j is to use a characteristic of the individual response spectra that, when averaged over all motions in the suite, has zero mean. To do this, the mean spectral response (\overline{Sa}) is defined as the mean $\ln Sa$ over all periods, with the periods equally-spaced in log-space:

$$\overline{Sa} = \frac{1}{n_p} \sum_{i=1}^{n_p} \ln Sa_i \quad (8)$$

The mean spectral response for each motion (\overline{Sa}_j) is computed, as well as the mean spectral response for the combination (\overline{Sa}_{avg}). By subtracting the mean spectral response of the suite (\overline{Sa}_{avg}) from the mean spectral response of each individual motion (\overline{Sa}_j), individual factors, z_j , that have a distribution with zero mean are defined:

$$z_j = \overline{Sa}_j - \overline{Sa}_{avg} \quad (9)$$

The z_j values for the motions shown in Figure 1 are: 0.9859, 0.4903, 0.2454, -0.0442, -0.0992, -0.1671, and -1.411. These z_j values indicate 4 motions are, on average, below the median response spectrum for the suite ($z_j < 0.0$), while three motions are above it ($z_j > 0.0$).

The distribution of z_j about its zero mean retains information regarding the position of the individual response spectra relative to the median, but it still lacks the proper standard deviation. To increase or decrease the standard deviation (σ_{ln}) of the suite, the z_j values are scaled by a constant influence factor (α) and combined with the average scale factor (s_{avg}) to compute the scale factor for each motion (s_j):

$$\ln s_j = \ln s_{avg} + \alpha \cdot z_j \quad (10)$$

The influence factor is adjusted to best fit the σ_{ln} of the target response spectrum. While there may be a more elegant solution to the problem, a simple and effective method involves trial and error. The influence factor is varied from -3 to 3 in steps of 0.01. For every assumed value of α , the individual response spectra in the combination are individually scaled using (10) and the standard deviation of the combination at each period is computed. The computed standard deviation (σ_{ln}) at each period is compared to the target standard deviation ($\sigma_{ln,target}$) at each period and the MSE between the two curves is calculated. The influence factor that results in the lowest MSE is selected, as shown in Figure 2 for the motions considered in Figure 1. The fit of the standard deviation is significantly improved by applying the individual scale factors from (10), however an excellent fit is not obtained because seven motions will never display a smooth curve of σ_{ln} versus period. Using these individual scale factors for the combination presented in Figure 1, the scaled individual spectra, the median spectrum, and the $\pm\sigma$ spectra are shown in Figure 3. The resulting scaled combination fits the target response spectrum and generally fits the target $\pm\sigma$ spectra.

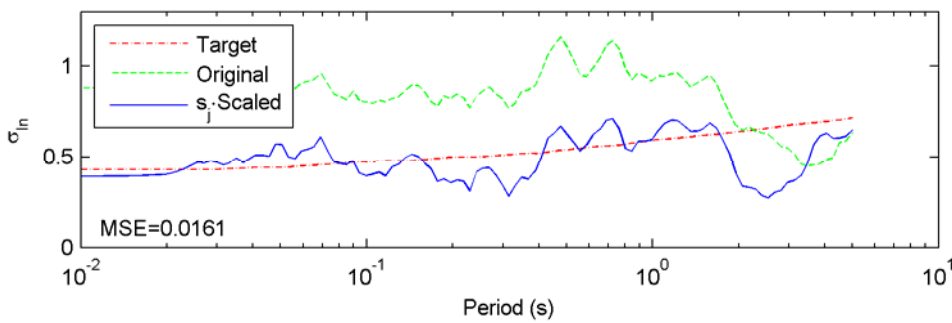


Figure 2. The standard deviation of the unscaled motions and of the motions scaled using s_j scale factors to fit the target standard deviation.

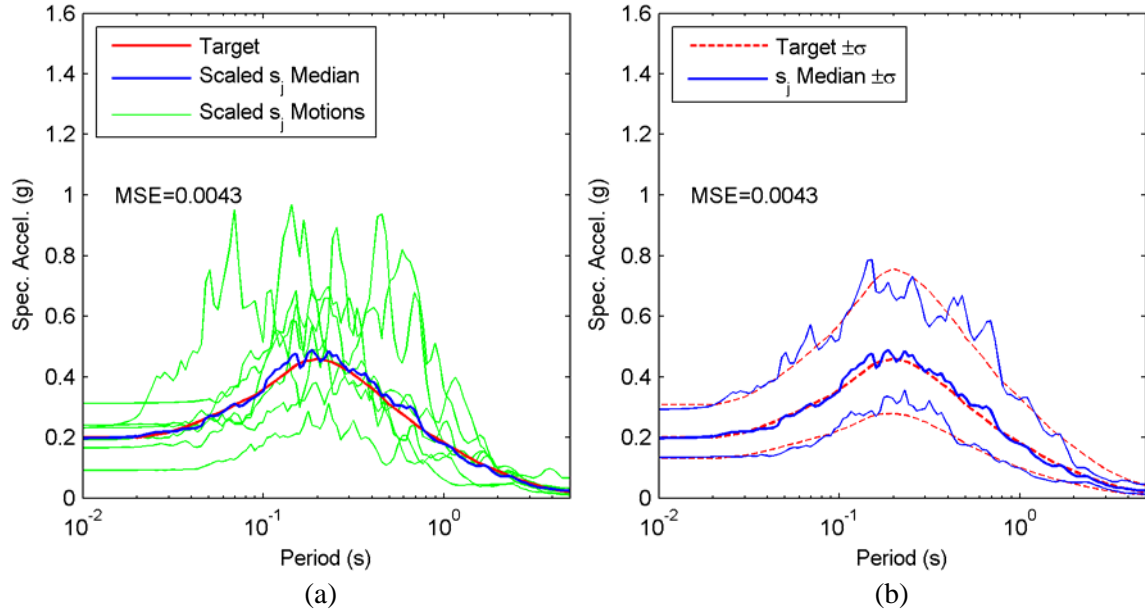


Figure 3. (a) Individual motions scaled by s_j , (b) Median and median $\pm \sigma$ of motions scaled by s_j .

AUTOMATED SELECTION OF GROUND MOTION COMBINATIONS

The previous section focused on scaling a given combination of motions, but did not consider how to select an appropriate combination of motions from a larger catalog of motions. There is one combination of n_m motions from the catalog that will produce the lowest MSE when optimally scaled and compared with the target response spectrum. To find this combination a trial of every possible combination must be performed. The number of required trials is (Devore, 2000):

$$N_{trials} = \frac{k!}{n_m!(k - n_m)!} \quad (11)$$

where k is the number of motions in the catalog and n_m is the number of motions to choose for the suite. This approach is referred to as the *rigorous* approach and the number of trial combinations required becomes very large ($>10^{10}$ trials) as n_m and k increase. The computing time for the large number of trials could take months on a desktop computer. However, if the number of motions to choose is small, the trial of each combination can be accomplished within minutes to hours.

An alternative to trying every combination would be to construct the combination one motion at a time in an iterative approach. Here the motions are selected based on some criteria, instead of randomly checking each possibility. For the *iterative* approach developed in this study, the average response spectrum is constructed by first selecting an initial seed motion and incrementally adding the one complementary motion that results in the lowest *MSE* relative to the target spectrum. The initial seed motion is selected based on the individual motion with the lowest *MSE* when optimally scaled. The process is repeated, adding one motion at a time, until the desired number of motions (n_m) is reached.

The number of trials required by the *iterative* method for a catalog of k motions and n_m motions to choose is:

$$N_{trials} = 1 + \sum_{j=1}^{n_m} (k - j) \quad (12)$$

The number of trials is relatively small compared to the *rigorous* solution, but each trial is selecting the combination based on information rather than randomly generating a combination. Nonetheless, the reduction in trials may result in the combination with the lowest *MSE* being missed or the selected combination having an *MSE* much larger than the minimum *MSE*. To deal with these issues, a

modified-iterative approach was developed that uses the same technique as the *iterative* method but uses every motion as the possible seed motion. The number of trials for the *modified-iterative* approach is:

$$N_{trials} = k \cdot \left[1 + \sum_{j=1}^{n_m} (k - j) \right] \quad (13)$$

The number of trials required to choose a given number of motions (n_m) from catalogs of 25 and 50 motions (i.e. $k = 25$ and $k = 50$) are shown in Figure 4 for the three methods (*rigorous*, *iterative*, and *modified-iterative*). As the number of motions in the catalog increases, the difference in the number of trials between the *rigorous* and *iterative* methods increases. For selecting 7 motions from a catalog of 50 motions, using the *iterative* method requires the trial of 323 combinations, the *modified-iterative* method requires 16,150 trials, and the *rigorous* method requires $9.99 \cdot 10^7$ trials. For both the *iterative* and *modified-iterative* approaches, the combinations can be found within a few seconds. However, the computation of the additional 100 million combinations in the *rigorous* approach would take an estimated 3 days on a typical desktop computer.

To generally compare the results from the different approaches, combinations of seven motions were generated from a 44 motion catalog using both the *rigorous* and *modified-iterative* methods. The *rigorous* computation of every possible combination took about 24 hours to complete and tried $3.83 \cdot 10^7$ unique combinations. The *modified-iterative* procedure tried $1.07 \cdot 10^4$ non-unique combinations of motions in 1.5 minutes. In the *modified-iterative* approach each seed motion produces only one combination and this combination may or may not be unique from previously developed combinations. The *rigorous* procedure generates a unique combination for each trial. A comparison between the *MSE* values of the 44 combinations generated with the *modified-iterative* method (one suite for each seed motion) and the top 100 combinations generated by the *rigorous* method are shown in Figure 5. The *MSE* values for the combinations generated using the *modified-iterative* method are larger than the combinations generated by the *rigorous* method, and the *MSE* increases quickly as one moves from the best-fit combination. This difference is due to the difference in numbers of trials

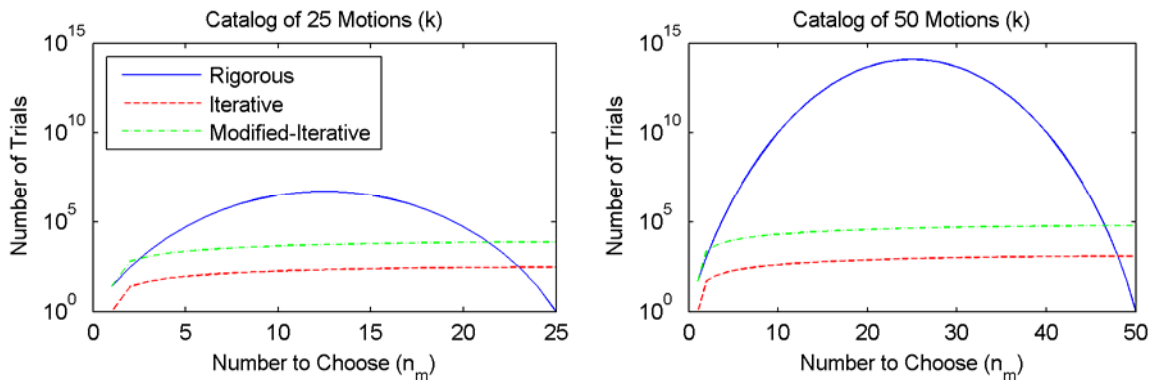


Figure 4. The number of trials for catalogs of 25 and 50 motions

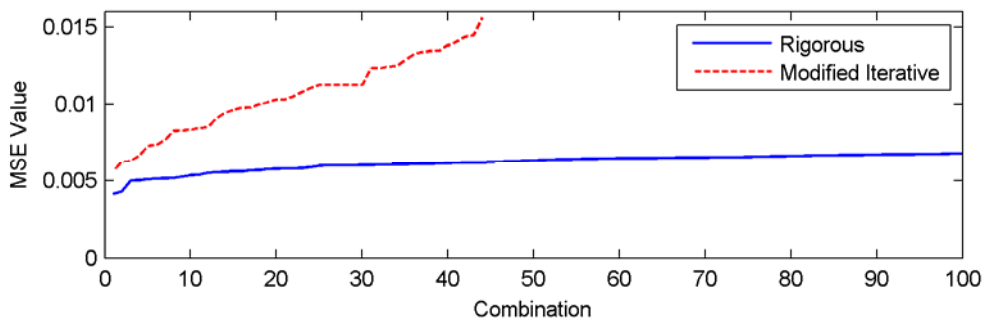


Figure 5. The *MSE* values computed by the rigorous and modified-iterative methods.

between the two methods. However, the best-fit MSE from the *modified-iterative* method is only somewhat larger than for the *rigorous* method.

APPLYING ENGINEERING JUDGMENT

The goal of introducing a semi-automated approach to ground motion selection is to provide the engineer with a tool to help make a better decision on the final selection and scaling of input ground motions. In the selection process it is important that the engineer be active because their judgment and experience cannot be captured in an algorithm. In the developed methodology, engineering judgment is incorporated when selecting the motions in the catalog and when selecting a final combination. These issues are discussed below.

A strong motion record from an earthquake carries more information than can be represented in an acceleration response spectrum (e.g. duration). As a result, these characteristics are not taken into account in the developed methodology. One way in which this issue is addressed is by selecting ground motions for the catalog that have similar source characteristics (magnitude, faulting mechanism, and distance) as the scenario event being considered. Nonetheless, this approach does not guarantee that the ground motion characteristics of importance are captured in the catalog or in each selected suite. Therefore, it is recommended to independently evaluate these additional ground motion characteristics after selection of the combinations.

Selection of motions by human trial and error can take significant time and is not a thorough approach. Using an automated method to select a combination can be very thorough, but can result in neglecting important characteristics (e.g., directivity, variability, etc.) not captured in the algorithm. In the semi-automated procedure, the experience of the engineer is included by generating a number of possible combinations, ranked by their *MSE* fit to the target spectrum, from which the engineer can choose. Additionally, to produce combinations that are significantly different from each other (i.e., to avoid two combinations that only vary by a single motion), the number of repeated motions between combinations is limited to a value specified by the user.

APPLICATION OF DEVELOPED METHODOLOGY

In the application of the semi-automated method for the selection and scaling of earthquake ground motions the following steps are required:

1. Define the scenario event (magnitude, faulting mechanism, distance, site conditions).
2. Using the scenario event, calculate the target response spectrum and variability from a ground motion prediction equation.
3. Select a catalog of motions based on the acceptable event characteristics (magnitude range, faulting mechanisms, distance range, site conditions).
4. Determine the method for generating the combinations, the allowable number of repeat motions between each combination, and the number of ranked combinations to be reported by the program.
5. Run the program to generate the requested number of ranked combinations.
6. Select the final combination.

To illustrate the developed methodologies, suites of seven ground motions were selected for a magnitude 7 strike-slip earthquake at a distance of 20 km, and rock site conditions. Using this scenario event, a target response spectrum and its standard deviation were defined using the Abrahamson and Silva (1997) ground motion prediction equation. To further judge the selected motions, the expected significant duration ($D_{5.95}$) of the scenario event was computed using the Abrahamson and Silva (1996) empirical duration model. The Abrahamson and Silva (1996) duration model predicts a median duration of 15.4 s and $\pm 1\sigma$ values of 9.4 to 25.3 s.

The catalog of potential motions was selected by searching the Next Generation Attenuation (NGA) strong motion database (<http://peer.berkeley.edu/nga/>) using the parameters shown in Table 1. These parameters—magnitude range, distance range, faulting mechanism, and rock site conditions—are specified such that a variety of motions are present in the catalog for the given scenario. The search resulted in a total of 44 motions, as summarized in Table 2. The motions came from 6 different earthquakes with the majority (38 out of 44) being recorded for earthquakes with a magnitude below the scenario event. Because duration is most influenced by earthquake magnitude, the limited number of motions from large magnitude earthquakes may result in combinations with median durations that are smaller than expected for the scenario. If duration were critical in subsequent analyses, the upper limit on the magnitude range could be extended to include more motions from larger earthquakes, and presumably, with longer durations.

Table 1. Parameters used to search for motions for the ground motion catalog.

Parameter	Values
Magnitude (M_w)	6.6 to 7.3
Closest Distance	5 to 30 km
Fault Type	Strike-Slip, Reverse
$V_{s,30}$	600 to 2500 m/s

Table 2. The earthquake events and number of records in the ground motion catalog .

Event	Magnitude	Fault Type	No. of Records
San Fernando	6.61	Reverse	6
Northridge	6.69	Reverse	18
Kobe, Japan	6.9	Strike-Slip	2
Loma Prieta	6.93	Reverse-Oblique	12
Hector Mine	7.13	Reverse	2
Duzce, Turkey	7.14	Strike-Slip	4

In this example, suites of motions are generated using both the *rigorous* and *modified-iterative* methods. For both methods, a procedure is used that updates the best five suites as suites are tried. Two limitations are placed on the ground motions suite. First, the number of motions that are repeated between suites is limited to 4 to produce suites that are significantly different. Second, only one component from a station for a given event is allowed, as each of the two components travel the same path and therefore are not independent.

Two combinations were generated by both the *rigorous* (R-1 and R-2) and *modified-iterative* (M-1 and M-2) methods (Figure 6). The MSE values for the fit of each suite to the target response spectrum ($S_a MSE$) and to the target standard deviation ($\sigma_m MSE$) are presented in Table 3. For each of the ground motion suites, there is good agreement (i.e. low $S_a MSE$) between the target and median response spectra. While the $S_a MSE$ values are all similar, there is a larger range of $\sigma_m MSE$ values due to the algorithm not directly selecting suites based on their fit to the prescribed standard deviation and because it is difficult to get a good fit for $\sigma_m MSE$ when only selecting seven motions.

Both of the suites generated by the *rigorous* method (Figures 6a and 6b) show good agreement with the target response spectrum and both display a balance of scatter about the median. One of the ground motion suites generated through the *modified-iterative* procedure (Figure 6c) contains a motion with excessively large long period energy, which is inappropriate for the scenario event. In comparing the fit of the *rigorous* and *modified-iterative* methods, the similar $S_a MSE$ values indicate that the *modified-iterative* method can do a satisfactory job in selecting motion and provide a good approximation to the best combination.

The scale factors and summary of selected motions for the R-1 combination are presented in Table 4. The scale factors in the ground motions suite fall within the generally acceptable range of 0.5 to 3.

Table 3. The MSE of the generated ground motions suites relative to the target response spectrum and standard deviation.

Suite	S_a MSE	σ_{ln} MSE
<i>Rigorous</i>		
R-1	0.0051	0.0164
R-2	0.0069	0.0093
<i>Modified-Iterative</i>		
M-1	0.0063	0.0215
M-2	0.0066	0.0104

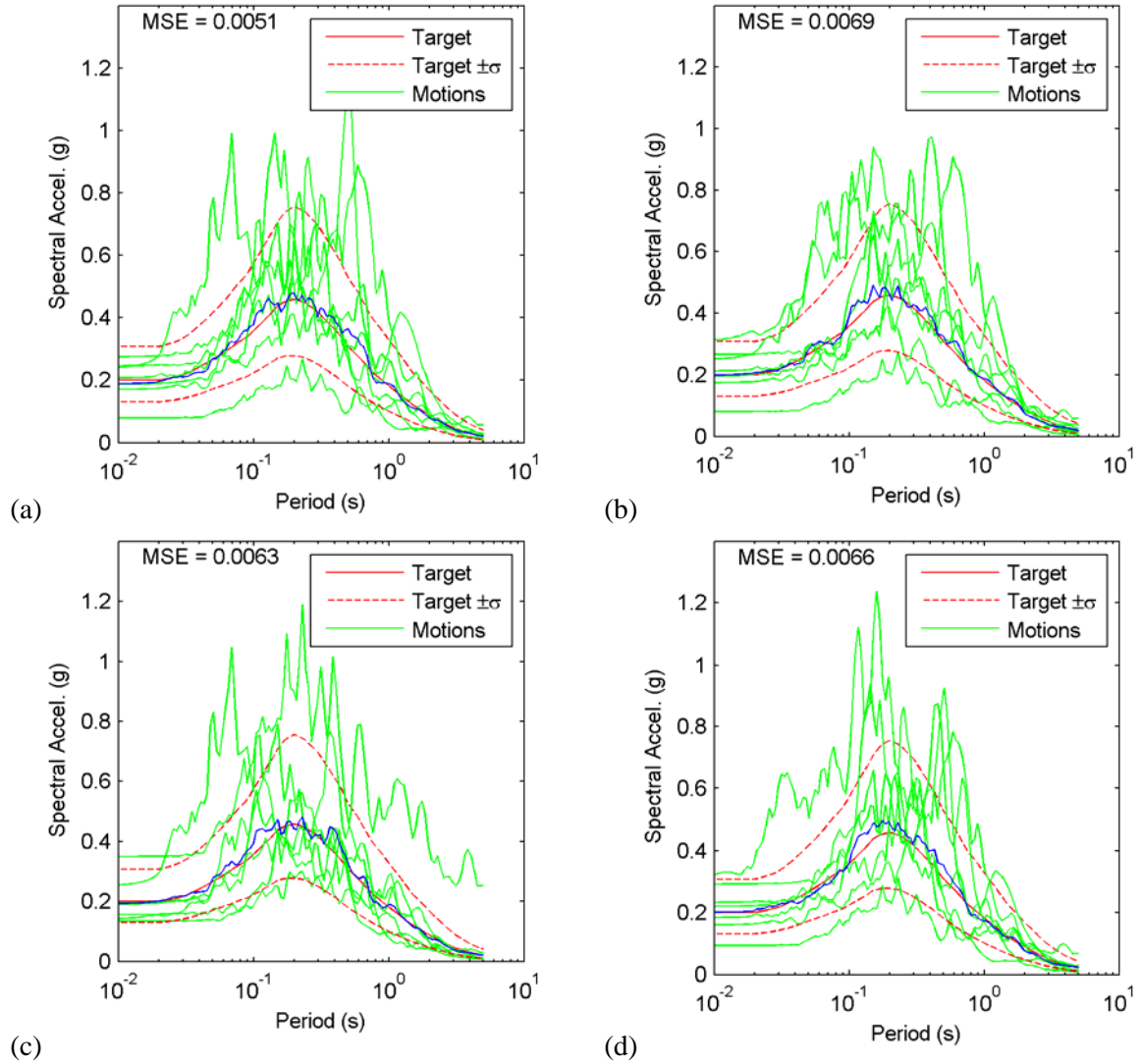


Figure 6. Two ground motion suites generated by the rigorous method (a) and (b), and two generated by the modified iterative method (c) and (d).

Table 4. The scale factors and motions in the R-2 ground motion suite.

Station	Event	M_w	R (km)	Scale	PGA (g)	PGV (cm/s)	D_{5-95} (s)
Gilroy #6, 000	Loma Prieta	6.93	18.33	1.38	0.17	17.65	12.91
Hector, 000	Hector Mine	7.13	11.66	1.00	0.27	28.58	11.66
Burbank-Howard, 060	Northridge	6.69	16.88	1.62	0.19	15.45	11.70
Gilroy Gavilan, 067	Loma Prieta	6.93	9.96	0.88	0.31	25.07	5.00
LA Chalon Rd, 160	Northridge	6.69	20.45	1.15	0.21	31.00	6.90
Lamont 1060, N	Duzce	7.14	25.88	2.85	0.08	31.39	19.06
Lake Hughes #9, 291	San Fernando	6.61	22.57	1.85	0.25	7.12	11.82

The range of closest distances and magnitudes bracket the scenario event ($M_w = 7$, $R = 20$ km). The only concern is that the duration of shaking is shorter than the duration predicted by the empirical model. The median duration is 10.5 seconds which is nearly 5 seconds shorter than the predicted duration from Abrahamson and Silva (1996). The deficiency in the duration of the suite may not be critical because many current design procedures are not sensitive to duration. However, duration would require further consideration if these motions were to be used in subsequent analyses that were sensitive to duration (e.g., Newmark sliding block analyses).

SUMMARY

A new semi-automated method for selection and scaling of acceleration-time histories to fit a target response spectrum and target standard deviation was presented. The new technique replaces the tedious approach of trials of combinations by hand with a semi-automated method. One of the major advantages of the semi-automated method is that a number of potential combinations can be selected; such that the engineer can then decide which combination best represents the scenario event.

The method takes advantage of the fact that the motions, independent of their scale factors, control the spectral shape of the suite, while the scale factors control the amplitude and standard deviation of the suite. Thus, motions can be selected without knowledge of the individual scale factors. For the developed methodology, the motions for the suite are first selected based on the fit to the spectral shape. This step is followed by selection of the individual scale factors for each motion to match both the target response spectrum and target standard deviation. Both rigorous and iterative selection algorithms were introduced and applied to a scenario event. Although the iterative method checked significantly fewer combinations, the resulting suites fit the target spectrum almost as well as the suites from the rigorous method.

Before the proposed algorithm can be used solely for selecting ground motions for a design analysis, the method needs to be validated. In the validation, the algorithm needs to be applied to different earthquake scenarios and the selected motions used in dynamic analyses for various geotechnical and structural applications.

ACKNOWLEDGEMENTS

Financial support for this work was provided by Pacific Gas and Electric, Co. (PG&E) and Dr. Norman Abrahamson. This support is gratefully acknowledged.

REFERENCES

- Abrahamson, N. and Silva, W.J. Empirical ground motion models. Technical report, Brookhaven National Laboratory, 1996.
- N. A. Abrahamson and W. J. Silva. Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seismological Research Letters*, 68(1):94–127, January-February 1997.
- Devore, J.L. *Probability and Statistics for Engineering and the Sciences*, 5th Ed. Duxbury, California, 2000.