ABSTRACT

Developing an understanding of duration of strong-shaking is important when characterizing engineering ground motion because the duration of strong shaking can be correlated to critical damage indices in the performance-based design of civil engineering structures. Most existing empirical models for duration are for crustal earthquakes. Using the Next Generation Attenuation Subduction Zone Database, we derived a regional empirical model for duration for subduction zone earthquakes. In evaluating the subduction data from large magnitudes (M>8), we found that the commonly used measure of duration based on the time between 5% and 75% or the Arias intensity does not work well if the ground motion has multiple sections of strong shaking that are separated in time. In crustal earthquakes, the slope of the Arias intensity is approximately constant over the 5% - 75% range. With the concept that the parameter of interest is the actual duration for which the shaking has significant energy, we propose a revised algorithm to computing significant duration that computes the slope of the Arias intensity in 5% increments and then summed the time increments with the largest slopes to reach 75% range, effectively skipping the time increments with small slopes (low energy). This new method provides a measure of duration that captures the duration of strong shaking in large subduction earthquakes but remains consistent with the current definition for typical ground motions without large separations in the strong shaking in the recording. Using this new definition, we derive empirical models for the duration for subduction earthquakes in Japan, Taiwan, South America, Cascade, Alaska, Central America & Mexico and New Zealand. Compared to the duration models for crustal earthquakes, the subduction duration model shows stronger distance dependence and weaker magnitude dependence.

1Seismologist engineer, GeoEngineers Inc., Redmond, WA 98052
2Researcher, Pacific Earthquake Engineering Research Center, University of Berkeley, CA
3Professor, Civil and Environmental Engineering Department, University of Berkeley, CA
4Researcher, Pacific Earthquake Engineering Research Center, University of Berkeley, CA
REGIONAL GROUND MOTION DURATION PREDICTION MODEL FOR SUBDUCTION REGIONS

M. Walling, N. Kuehn, N. Abrahamson, and S. Mazzoni

ABSTRACT

Developing an understanding of duration of strong-shaking is important when characterizing engineering ground motion because the duration of strong shaking can be correlated to critical damage indices in the performance-based design of civil engineering structures. Most existing empirical models for duration are for crustal earthquakes. Using the Next Generation Attenuation Subduction Zone Database, we derived a regional empirical model for duration for subduction zone earthquakes. In evaluating the subduction data from large magnitudes (M>8), we found that the commonly used measure of duration based on the time between 5% and 75% or the Arias intensity does not work well if the ground motion has multiple sections of strong shaking that are separated in time. In crustal earthquakes, the slope of the Arias intensity is approximately constant over the 5%-75% range. With the concept that the parameter of interest is the actual duration for which the shaking has significant energy, we propose a revised algorithm to computing significant duration that computes the slope of the Arias intensity in 5% increments and then summed the time increments with the largest slopes to reach 75% range, effectively skipping the time increments with small slopes (low energy). This new method provides a measure of duration that captures the duration of strong shaking in large subduction earthquakes but remains consistent with the current definition for typical ground motions without large separations in the strong shaking in the recording. Using this new definition, we derive empirical models for the duration for subduction earthquakes in Japan, Taiwan, South America, Cascade, Alaska, Central America & Mexico and New Zealand. Compared to the duration models for crustal earthquakes, the subduction duration model shows stronger distance dependence and weaker magnitude dependence.

Introduction

The duration of shaking occurring at a site during an earthquake is an important characteristic of ground motion and has been shown to be a useful parameter in engineering design. Duration has been shown to influence liquefaction, slope-stability and structural damage [1]. In seismic hazard, duration is often used when selecting ground-motions for matching to Conditional Mean Spectra (CMS) and Uniform Hazard Spectra (UHS), by checking that the duration of the sample records is within a desired range of empirical duration prediction models (DPMs).

Examples of DPMs developed in the past 25 years include [2, 3, 4, 5, 6, 7, 8]. These DPM

---

1Seismologist engineer, GeoEngineers Inc., Redmond, WA 98052
2Researcher, Pacific Earthquake Engineering Research Center, University of Berkeley, CA
3Professor, Civil and Environmental Engineering Department, University of Berkeley, CA

characterize duration from crustal regions, apart from [8]. Although duration has been shown to be a useful parameter, little attention has been given to developing DPMs for subduction regions, presumably because comprehensive datasets of duration have not existed until now with the Next Generation Attenuation-Subduction (NGA-Sub) Project database [9]. Instead it is common for DPMs from crustal regions to be used as analogs for subduction regions, although there have been no studies performed to support this assumption that crustal earthquake duration models are applicable to subduction earthquakes.

The primary purpose of this article is to present an empirical regional DPM for interface and intraslab source-types developed using the comprehensive NGA-Sub Project database. We also introduce a new algorithm to compute ground motion duration that we think better quantifies the duration of strong shaking than the traditional duration metrics, and we also propose a functional form for the duration model that we feel better characterizes the physical process than other functional forms used in previous studies.

### Dataset

A ground motion duration dataset of recordings from subduction earthquakes was compiled as part of the NGA-Subduction Project [9]. The complete dataset consists of 71,343 recordings from seven global regions, Alaska, Cascade, Central America and Mexico, New Zealand, Japan, South America and Taiwan. The metadata with the source, site and path parameters used in the analysis are the same metadata distributed to the NGA-Subduction ground motion modelers for the ground motion prediction equation (GMPE).

A subset of data was selected based on source-type, distance, depth, and availability of metadata. The selection criteria is summarized as follows: records with source-types identified as interface and intraslab source; records that had $R_{jb}$ distances less than $R_{\text{max}}$ and $R_{jb}$ distances less than 300 kilometers; records with available $V_s30$ values; and intraslab sources with a hypocentral depth greater than 30 kilometers. The final dataset consisted of 10,037 interface records from 124 earthquake events, and 8,137 intraslab records from 86 earthquake events. Table 1 summarizes the number of records from each region by source-type.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Number of Intraslab Records</th>
<th>Number of Interface Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>353</td>
<td>96</td>
</tr>
<tr>
<td>Cascade</td>
<td>117</td>
<td>0</td>
</tr>
<tr>
<td>Japan</td>
<td>4836</td>
<td>8842</td>
</tr>
<tr>
<td>New Zealand</td>
<td>416</td>
<td>46</td>
</tr>
<tr>
<td>Taiwan</td>
<td>2084</td>
<td>397</td>
</tr>
<tr>
<td>South America</td>
<td>282</td>
<td>526</td>
</tr>
<tr>
<td>Central America &amp; Mexico</td>
<td>49</td>
<td>130</td>
</tr>
</tbody>
</table>

### Ground Motion Duration Defined

There are multiple duration metrics that include significant duration, bracketed duration, arias intensity, and cumulative absolute velocity. The definition that we use in this paper is the duration
based on the normalized arias intensity of the acceleration shown by Eq. 1, which measures the percentage of total energy released in an earthquake. The duration, $D_{I1-I2}$, is the time interval between $I(t)$ reaches two values given by Eq.2:

$$I(t) = \frac{\int_{0}^{t} a(\tau)^2 d\tau}{\int_{0}^{\infty} a(\tau)^2 d\tau}$$ (1)

$$D_{I1-I2} = t(I_2) - t(I_1)$$ (2)

where $I(t)$ is the normalized arias intensity and ranges from 0 to 1; $a(\tau)$ is the acceleration time history; $t$ is time; $D_{I1-I2}$ is commonly referred to as the significant duration because it represents the time interval that contains the most significant shaking at the site. The most commonly used definition of significant duration by engineering seismologists and earthquake engineers is when $I_1 = 0.05$ and $I_2 = 0.95$, and when $I_1 = 0.05$ and $I_2 = 0.75$, denoted as $D_{0.05-0.95}$ [10] and $D_{0.05-0.75}$, [2], respectively.

The definition shown in Eq. 1 and Eq. 2 works well when the largest energy within the earthquake arrives at the station all together, as shown by Fig.1a, which is typical of earthquakes with only one rupture; however, this definition does not work well when there is more than one rupture on the fault causing the energy from each rupture to arrive at the station separated in time. Having multiple ruptures occurring during an earthquake event is not uncommon in subduction regions which tend to have large fault dimensions with more area to rupture, such as the M9 Tohoku event in Japan. Fig.1a and Fig.1b show two records from this event with very different signatures. In Fig.1a the largest shaking occurs together in the record, whereas in Fig. 1b the largest shaking occurs separated in time.

For records that have more than one full body wave visibly present, like Fig. 1b, the time interval covering the most significant shaking in the record also includes some of the least significant shaking in the record (e.g., the duration between the first and second wave of energy in the record). Because we are only interested in duration of the most significant shaking in the record, we propose a revised algorithm to compute significant duration that filters out the duration from
the less significant parts of the record. The revised algorithm is described as follows:

1. Compute the normalized arias intensity in Eq. 1.
2. Compute the derivative of the normalized arias intensity in time increments \( t_1 \) and \( t_2 \).
3. Rank the derivative from high to low in Step 2.
4. Resort the normalized arias intensity from Step 1 according to the rank from Step 3.
5. Compute the duration for \( D_{0.05-0.75} \).
6. For Husid plots, the duration from Step 4 should be placed back in original time order.

Our proposed revised algorithm to compute significant duration is superior to the commonly used algorithm (Eq. 1 and Eq. 2) because only the duration of the strongest shaking within the record is measured, which is not a guarantee with the predecessor algorithm. Furthermore, when only one body wave is visibly present, as in Fig.1a, the predecessor algorithm and proposed algorithm converge to the same duration values. To demonstrate this point, Husid plots comparing the predecessor algorithm (shown as a blue dashed line and labeled ‘old method’) and our new proposed algorithm (shown as a solid red line and labeled ‘new method’) are presented in Fig.2a and Fig.2b for the example Tohoku records. In Fig.2a, the \( D_{0.05-0.75} \) for record shown in Fig.1a is approximately 114 seconds for both methods, whereas in Fig.2b the \( D_{0.05-0.75} \) for record shown in Fig.1b is approximately 53 and 32 seconds for the predecessor and proposed algorithm, respectively. The proposed algorithm suggests that the duration of significant energy release is much shorter than the duration computed using the predecessor algorithm.

![Figure 2a](image1.png)  Husid plot showing the duration with the old (blue-dashed line) and new method (red line).

![Figure 2b](image2.png)  Husid plot showing the duration with the old method (blue-dashed line) and new method (red line).

**Ground Motion Duration Analysis**

A \( D_{0.05-0.75} \) DPM was developed using the NGA-Subduction database, which was later used to develop a \( D_{0.05-0.95} \) model. For this paper, we only discuss the \( D_{0.05-0.75} \) model and leave the development of \( D_{0.05-0.95} \) model to a later paper. To come up with the functional form used in the regression analysis, we initially reviewed the functional forms used in the past studies, such as functional forms by [2, 3, 4, 5, 7, 8]. Our review showed that the correlation between the source and path terms have been incorrectly handled in the regression analysis in previous duration models, by either treating these effects as multiplicative instead of additive, or improper placement of the source-event terms and path terms in the regression model.
Ground motion duration is the sum of the source duration and path duration [6,10] and is modeled by Eq. 3. [6] noted that it makes sense physically to treat these terms as additive because these effects should contribute independently to the total duration, like the separation of the source and path effects on the duration as implemented in the point-source stochastic method.

\[ Y_{dur} = D_{path} + D_{source} \]  

in which \( Y_{dur} \) is the total duration, \( D_{source} \) is the source duration, and \( D_{path} \) is the path duration. Based on empirical observations, the ground-motion duration, \( Y_{dur} \), is lognormally distributed. Both the path duration \( D_{path} \) and source duration \( D_{source} \) are modeled as independent lognormally distributed:

\[ D_{path} \sim LN(\mu_{path}, \sigma_{path}) \]  
\[ D_{source} \sim LN(\mu_{source}, \sigma_{source}) \]

in which \( \mu_{path} \) and \( \mu_{source} \) denote the median predictions for the path and source duration, respectively, which are functions of distance and magnitude, and \( \sigma_{path} \) and \( \sigma_{source} \) are the standard deviations. The functional forms are modified from [9]:

\[ \mu_{path} = c_2 * R_{jb} \quad \text{for} \quad R_{jb} \geq 40 \]  
\[ \mu_{source} = \frac{1}{3.2 * 4.9 * 10^6} * 10^{(0.5 * M + 5.35)} * e^{(-\frac{b_1 + b_2}{3} * (M - 6))} \]

where \( M \) is the earthquake magnitude, and \( R_{jb} \) is the Joyner-Boore distance in kilometers, and \( \mu_{path} = c_2 * 40 \) for \( R_{jb} < 40 \). Eq. 4 to 7 are equivalent to Eq. 8, shown by:

\[ Y_{dur} = c_2 * R_{jb} \epsilon_{ij} + \eta_i \frac{1}{3.2 * 4.9 * 10^6} * 10^{(0.5 * M + 5.35)} * e^{(-\frac{b_1 + b_2}{3} * (M - 6))} \]

where \( \eta_i \) and \( \epsilon_{ij} \) are the event-specific and record-specific residual from the \( i \)th earthquake and \( j \)th station, which are distributed according to a lognormal distribution with median 1 and standard deviation \( \sigma_{source} \) and \( \sigma_{path} \), respectively.

The parameters of the model are estimated via Bayesian inference using the program Stan [12, 13]. To account for regional differences in the scaling of the duration with magnitude and distance, the coefficients \( c_2, b_1 \), and \( b_2 \) are regional dependent. To avoid the problem of having too few records in a particular region, the model is cast as a hierarchical/multi-level model, which allows sharing statistical strength between regions. The multi-level nature of the model means that the regional coefficients share a common global prior distribution, shown by Eq. 9, 10 and 11 as follows:

\[ c_2 \sim N(\mu_{c2}, \sigma_{c2}) \]
\[ b_1 \sim N(\mu_{b_1}, \sigma_{b_1}) \]  
\[ b_2 \sim N(\mu_{b_2}, \sigma_{b_2}) \]

where \( \mu_{c_2}, \mu_{b_1}, \) and \( \mu_{b_2} \) are the global mean coefficients of \( c_2, b_1, \) and \( b_2. \) The prior distributions, shown below in Eq. 12 to 14, for the global mean coefficients are set based on the values of the Brookhaven model [9]:

\[ \mu_{c_2} \sim N(0.063,0.1) \]  
\[ \mu_{b_1} \sim N(0.85,2) \]  
\[ \mu_{b_2} \sim N(5,10) \]

The standard deviations \( \sigma_{c_2}, \sigma_{b_1}, \) and \( \sigma_{b_2} \) describe how much the coefficients \( c_2, b_1, \) and \( b_2 \) can vary across regions. Their prior distributions are exponential distributions described by Eq. 15, 16, and 17.

\[ \sigma_{c_2} \sim Exp(6) \]  
\[ \sigma_{b_1} \sim Exp(6) \]  
\[ \sigma_{b_2} \sim Exp(30) \]

The prior distributions for the event and record specific standard deviation are half-Cauchy distributions with parameter 0.5, as recommended by [14]. Due to the hierarchical nature of the model, coefficients in regions with only few data points are automatically shrunk towards the global mean. The use of Bayesian inference allows us to assess the uncertainties of each parameter in a probabilistic way. Fig. 3a and 3b, show the model residuals versus distance for interface and intraslab source-type, respectively.
Duration Model

The mean regional ground motion duration prediction model for six subduction regions is presented versus distance in Fig. 4 and 5 for earthquake magnitude 7 and 8, respectively, and versus earthquake magnitude in Fig. 6 and 7, for Rjb equal to 10 km and 100 km, respectively. The regions shown are Alaska in red, Japan in dark green, New Zealand in Green, Taiwan in blue, South America in purple and Central America & Mexico in pink. The interface source-type models are shown as a solid line, and the intraslab source-type models are shown as a dashed line. The solid black line represents the Brookhaven DPM from [2] for crustal regions.

Regional differences in duration are larger for interface source-types than intraslab source-types. Regional differences tend to go away at the large magnitudes where there are few records from regions outside Japan, thus this maybe more an artifact of the dataset than an agreement between regions. Comparing the duration from the subduction source-type, interface durations are greater than the intraslab durations, but with comparable distance scaling (similar distance slopes). The magnitude scaling for magnitudes less than 7 is also comparable, but for magnitudes greater than 7, the intraslab durations show a greater magnitude scaling (steeper magnitude slope). Comparing the crustal duration (solid black line) to the subduction duration, the crustal sources show a larger magnitude scaling and lower distance scaling with duration.

Figure 4  Duration model versus distance for earthquake M7. Solid lines are interface source-type, dashed lines are intraslab source-type, black line is Brookhaven (crustal model).
Figure 5  Duration model versus distance for M8. Solid lines are interface source-type, dashed lines are intraslab source-type, black line is Brookhaven(crustal model).

Figure 6  Duration model versus earthquake magnitude at $R_{jb} = 10$ km. Solid lines are interface source-type, dashed lines are intraslab source-type, black line is Brookhaven(crustal model).
Visual inspection of the NGA-Subduction dataset records show that many records have large excitations separated in time. Defining the duration as one continuous time interval is not meaningful for these records, motivating us to propose a new algorithm to compute duration that only uses the strongest excitation in the record. The effect of using this algorithm on the predicted duration value was not evaluated in this paper but will be evaluated in a future study. Additionally, more studies on the effect of ground motion duration on structural response need to be performed to determine the value of this revised definition of duration.

A functional form of the duration model used in this study properly models source and path effects as additive and properly treats the random effects as multiplicative on source and path effects separately. This is the first proper statistical treatment of the random effects in a duration model. The framework to develop the regional ground motion duration prediction model using a Bayesian inference with hierarchical nature of the model is also a novel approach for developing a ground-motion duration prediction model. This framework has the advantage that the regions with little data, such as Cascadia, trend toward the global mean and that uncertainties of each parameter is propagated in a probabilistic way. The results of this analysis show the subduction duration model shows stronger distance dependence and weaker magnitude dependence than crustal duration model. A comparison of the interface and intraslab duration show a larger magnitude and distance scaling with interface sources.
**Acknowledgments**

We would like to acknowledge the members and supporters of the NGA-Subduction Project.

**References**


