NGA-Subduction Global Ground-Motion Models with Regional Adjustment Factors

Grace A. Parker*
Jonathan P. Stewart
Department of Civil and Environmental Engineering
B. John Garrick Institute for Risk Sciences
University of California, Los Angeles
*Now at the U.S. Geological Survey, Moffett Field, California

David M. Boore
Geophysicist, Los Altos, California

Gail M. Atkinson
Department of Earth Science
Western University
London, Ontario, Canada

Behzad Hassani
Civil Design, BC Hydro
Burnaby, British Columbia, Canada

PEER Report 2020/03
Pacific Earthquake Engineering Research Center
Headquarters, University of California at Berkeley

August 2020
Disclaimer

The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the study sponsor(s), the Pacific Earthquake Engineering Research Center, or the Regents of the University of California.
NGA-Subduction Global Ground-Motion Models with Regional Adjustment Factors

Grace A. Parker*
Jonathan P. Stewart
Department of Civil and Environmental Engineering
B. John Garrick Institute for Risk Sciences
University of California, Los Angeles
*Now at the U.S. Geological Survey, Moffett Field, California

David M. Boore
Geophysicist, Los Altos, California

Gail M. Atkinson
Department of Earth Science
Western University
London, Ontario, Canada

Behzad Hassani
Civil Design, BC Hydro
Burnaby, British Columbia, Canada

PEER Report 2020/03
Pacific Earthquake Engineering Research Center
Headquarters at the University of California, Berkeley
August 2020
Next Generation Attenuation Subduction (NGA-Sub) is a multi-year, multidisciplinary project with the goal of developing an earthquake ground-motion database of processed time series and ground-motion intensity measures (IMs), as well as a suite of ground-motion models (GMMs) for global subduction zone earthquakes. The project considers interface and intraslab earthquakes that have occurred in Japan, Taiwan, New Zealand, Mexico, Central America, South America, Alaska, and Cascadia. This report describes one of the resulting GMMs, one important feature of which is its ability to describe differences in ground motions for different event types and regions.

We use a combination of data inspection, regression techniques, ground-motion simulations, and geometrical constraints to develop regionalized models for IMs for peak ground acceleration, peak ground velocity, and 5%-damped pseudo-spectral acceleration at 26 oscillator periods from 0.01 to 10 sec. We observe significant differences in ground-motion scaling for interface and intraslab events; therefore, the model terms for source and path effects are developed separately. There are complex distance-scaling effects in the data, including regional variations and forearc and backarc effects. No differences in site effects between the event types were observed; therefore, a combined site term is developed that is taken as the sum (in natural log units) of a linear term conditioned on the time-averaged shear-wave velocity in the upper 30 m ($V_{S30}$), and an empirically constrained nonlinear term. Basin sediment depth terms are developed for Cascadia and Japan that are conditioned on the depth to the 2.5 km/sec shear-wave velocity horizon ($Z_{2.5}$).

Our approach to model development was to first constrain a path term capturing the observed effects, then to subsequently investigate magnitude scaling, source-depth scaling, and site effects. Regionalized components of the GMM include the model amplitude, anelastic attenuation, magnitude-scaling corner, $V_{S30}$-scaling, and sediment depth terms.

Aleatory variability models are developed that encompass both event types, with different coefficients for each IM. Models are provided for four components of ground-motion variability: (1) between-event variability, $\tau$; (2) within-event variability, $\phi$; (3) single-station within-event variability, $\phi_{SS}$; and (4) site-to-site variability, $\phi_{S2S}$. The aleatory variability models are magnitude independent. The within-event variability increases with distances beyond 200 km due to complexities in path effects at larger distances. Within-event variability is $V_{S30}$-dependent for distances less than 200 km, decreasing for softer soils with $V_{S30}$ less than 500 m/sec. These reductions are attributed to soil nonlinearity. An ergodic analysis should use the median GMM and aleatory variability computed using the between-event and within-event variability models. An analysis incorporating non-ergodic site response (i.e., partially non-ergodic) should use the median GMM at the reference-rock shear-wave velocity (760 m/sec), a site-specific site amplification model, and aleatory variability computed using the between-event and single-station within-event variability models. Epistemic uncertainty in the median model is represented by standard deviation terms on region-dependent model constant terms, which facilitates scaled-backbone representations of model uncertainty in hazard analyses.

Model coefficients are available in the electronic supplement to this report (Tables E1–E4), and coded versions of the model are available in Excel, MatLab, R, and Python from Mazzoni et al. [2020(b)].
ACKNOWLEDGMENTS

This project was partially sponsored by the U.S. Geological Survey (USGS) contract G16AP00181, as well as the UCLA Civil & Environmental Engineering Department. The support of these organizations is gratefully appreciated. The opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the study sponsors, the Pacific Earthquake Engineering Research Center (PEER), or the Regents of the University of California.

We thank Yousef Bozorgnia and the Pacific Earthquake Engineering Research Center (PEER) for organizing the Next Generation Attenuation – Subduction (NGA-Sub) project. We thank the broader NGA-Sub team and, in particular, Silvia Mazzoni, Sean AWD, Victor Contreras, and Tadahiro Kishida for their efforts in producing the NGA-Sub database. We thank the other NGA-Sub model developers, including Yousef Bozorgnia, Nick Gregor, Nico Kuehn, Norm Abrahamson, Brian Chiou, Kenneth Campbell, and Linda Al Atik, for their helpful feedback during ground-motion model development and verification processes. We appreciate the careful editorial review of this report by Claire Johnson.
# CONTENTS

ABSTRACT .................................................................................................................................. iii  
ACKNOWLEDGMENTS ............................................................................................................. v  
TABLE OF CONTENTS ........................................................................................................... vii  
LIST OF FIGURES ................................................................................................................... xi  

1 INTRODUCTION .................................................................................................................. 1  
1.1 NGA-Sub Project ............................................................................................................. 1  
1.2 Literature Review ......................................................................................................... 2  

2 DATA SELECTION .............................................................................................................. 5  

3 NEAR-SOURCE SATURATION ............................................................................................ 9  
3.1 Motivation ..................................................................................................................... 9  
3.2 Available Empirical Constraints .................................................................................. 11  
3.3 Finite-Fault Simulations ............................................................................................... 12  
3.4 Modification for Slab Events ....................................................................................... 14  
3.5 Comparisons to Existing Models ................................................................................ 15  

4 DEVELOPMENT OF MEDIAN MODEL ............................................................................ 17  
4.1 Median Model Functional Form ................................................................................... 17  
4.1.1 Global and Regional Models ................................................................................... 17  
4.1.2 Regional Terms and Coefficients .......................................................................... 19  
4.2 Overview of Regression Procedure ............................................................................. 20  
4.3 Development of Model Terms ..................................................................................... 21  
4.3.1 Distance-Scaling ..................................................................................................... 21  
4.3.2 Magnitude Scaling ................................................................................................. 25  
4.3.3 Source-Depth-Scaling ............................................................................................ 27  
4.3.4 Regional and Global Constant Calibration ............................................................. 31  
4.3.5 Site Response ......................................................................................................... 33  
4.4 Model Residuals ......................................................................................................... 33  

5 SITE AMPLIFICATION ........................................................................................................ 41
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>5.2</td>
<td>Global and Regional Models</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Global Site-Amplification Model</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Regional Parameters</td>
</tr>
<tr>
<td>5.3</td>
<td>Model Development and Results</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Linear Site Amplification</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Nonlinear Site Amplification</td>
</tr>
<tr>
<td>5.4</td>
<td>Basin Amplification</td>
</tr>
<tr>
<td>5.5</td>
<td>Model Performance</td>
</tr>
<tr>
<td>5.6</td>
<td>Model Limitations and Recommended Use</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Aleatory Variability</td>
</tr>
<tr>
<td>6.1</td>
<td>Model Functional Form</td>
</tr>
<tr>
<td>6.2</td>
<td>Model Development</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Between-Event Variability, ( \tau )</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Within-Event Variability, ( \phi )</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Site-to-Site Variability, ( \phi_{SS} )</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Within-Event Single-Station Variability, ( \phi_{SS} )</td>
</tr>
<tr>
<td>6.2.5</td>
<td>Comparison of Alternate Within-Event Variance Models</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Model-to-Model Comparisons</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Summary and Future Work</td>
</tr>
<tr>
<td>8.1</td>
<td>Summary</td>
</tr>
<tr>
<td>8.2</td>
<td>Recommendations for Implementation</td>
</tr>
<tr>
<td>8.2.1</td>
<td>Model Limitations</td>
</tr>
<tr>
<td>8.2.2</td>
<td>Recommended Application to Cascadia</td>
</tr>
<tr>
<td>8.2.3</td>
<td>Epistemic Uncertainty outside of Cascadia</td>
</tr>
<tr>
<td>8.3</td>
<td>Future Work</td>
</tr>
<tr>
<td>8.3.1</td>
<td>NGA-Subduction Model Validation</td>
</tr>
<tr>
<td>8.3.2</td>
<td>Residuals Analyses in Backarc Complexes</td>
</tr>
<tr>
<td>8.3.3</td>
<td>Epistemic Uncertainty</td>
</tr>
</tbody>
</table>

REFERENCES
ELECTRONIC SUPPLEMENT

Table E.1: Coefficients for global and regional interface models
Table E.2: Coefficients for global and regional intraslab models
Table E.3: Coefficients for aleatory variability models
Table E.4: Coefficients for epistemic uncertainty models
## LIST OF FIGURES

| Figure 1.1 | Global map of plate boundaries from Stern [2002]. Convergent boundaries shown as solid lines with black teeth. | 2 |
| Figure 2.1 | Magnitude-distance distribution of recordings from interface (left) and intraslab (right) events, color-coded by region. | 6 |
| Figure 2.2 | Number of events and number of recordings selected for model development according to the criteria in Chapter 2 for interface (circles) and intraslab (triangles) events. Combined values are shown as diamonds. The left panel shows metrics for PGV (-1) and PGA (0). | 7 |
| Figure 3.1 | Example of the underdetermined problem of fitting $h$ using NGA-Sub data. Two simple path models [Equation (3.3)] fit to an interface event from Japan with $M = 8.29$, one with $h = 10$, and the other with $h = 30$ km. Both models produce a fit with negligible differences in the residual standard error (0.6138 vs. 0.61, respectively). | 10 |
| Figure 3.2 | A comparison of event-specific empirical estimates of near-source saturation term $h$ (see text for description of studies). Also shown are models for $h$ as a function of earthquake moment magnitude from Yenier and Atkinson [2014; 2015] (YA14, YA15, respectively), and Atkinson and Silva [2000] (AS00). | 11 |
| Figure 3.3 | Schematic showing fault and EXSIM finite-fault simulation station array. Stations are located at distances of 10, 20, 30, 40, 50, 70, 100, 200, 300, 500, 700, and 1000 km and azimuths of 90°, 60°, and 45° (not to scale). | 12 |
| Figure 3.4 | Comparison of the Yenier and Atkinson [2014; 2015] near-source saturation models (labeled YA14 and YA15), $h$ estimates from EXSIM simulations run to emulate subduction interface events, and the near-source saturation model developed in this study, given by Equation (3.2) with $h_1 = -0.82$ and $h_2 = 0.252$. | 13 |
| Figure 3.5 | Simulation results at $T = 0.2$ sec for an $M8$ intraslab earthquake with a high dip angle (73°; upward triangles) and a low dip angle (17°; downward triangles) compared to distance-scaling models using varying values of the near-source saturation value $h$. | 15 |
Figure 3.6  Comparison of published near-source saturation models. The models proposed in this work for interface and slab earthquakes are shown as red lines. The suite of models shown in blue are constrained using large earthquakes from active tectonic regions [Yenier and Atkinson 2014] (YA14); Yenier and Atkinson [2015] (YA15); and Atkinson and Silva [2000] (AS00). The models shown in black are from published subduction zone GMMs [Abrahamson et al. 2016] (BC Hydro) and Atkinson and Boore [2003] (AB03).

Figure 4.1  Schematic of path model, $F_p$ [Equation (4.2)], showing the near-source saturation model, $h$, the geometrical spreading slope, and the curvature due to anelastic attenuation.

Figure 4.2  Comparison of coefficients controlling the magnitude dependence of geometrical spreading, $b_4$, between the simulation-based HA18 model and the present work.

Figure 4.3  Example of PGV data for magnitude bin 6.5–7 over the distance range ($R_{rup} \leq 100$ km); decay is controlled by geometrical spreading.

Figure 4.4  Example of PGV data for magnitude bin 6.5–7 over the entire model distance domain ($R_{rup} \leq 1000$ km) with the geometrical spreading terms from Figure 3.6, plus the best fit global anelastic attenuation term. Results shown for interface and intraslab events.

Figure 4.5  Anelastic attenuation coefficient, $a_0$, as a function of oscillator period for interface events (top) and intraslab events (bottom). Lack of data for interface events in Cascadia means there is no regional value of $a_0$; instead, the global value is recommended.

Figure 4.6  Global interface (left) and intraslab (right) magnitude-scaling models [$F_M$, Equation (4.5)] and path model event terms as a function of $M$ for 0.2 and 2.0-sec PSA. For plotting purposes, the recommended global $m_c$ values were used for the intraslab and interface model ($M_{7.6}$ and 7.9, respectively).

Figure 4.7  Variation of event terms as a function of hypocentral depth at PGA, 0.2-sec PSA, and 1.0-sec PSA. Interface events are shown on the left and intraslab events are shown on the right. Binned means with standard errors and best-fit depth scaling model [Equation (4.6)] shown for intraslab events.

Figure 4.8  (a) Schematic showing fault plane geometry used to derive relationship between $Z_{tor}$ and $Z_{hyp}$ for forward use in hazard analyses; and (b) normalized down-dip hypocentral depth ($Z_{dip}/W$) for intraslab events in the NGA-Subduction database, color-coded by region.
Figure 4.9  A comparison of the global model constant ($c_0$) with regional constants for (a) interface and (b) intraslab earthquakes.................................................................33

Figure 4.10  Model bias terms $c$ for interface and intraslab regional models. .........................34

Figure 4.11  Event terms, $\eta_{E,i}$, from interface events as a function of moment magnitude for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars)....................................................................................................35

Figure 4.12  Event terms, $\eta_{E,i}$ from intraslab events as a function of moment magnitude for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars)....................................................................................................36

Figure 4.13  Event terms, $\eta_{E,i}$ from interface events as a function of hypocentral depth for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars)..............................................................................................................37

Figure 4.14  Event terms, $\eta_{E,i}$ from intraslab events shown as a function of hypocentral depth for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars)....................................................................................................38

Figure 4.15  Within-event residuals, $\delta W_{ij}$, from interface events for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Residuals are color-coded by subduction zone region. .................................................................................................................................39

Figure 4.16  Within-event residuals, $\delta W_{ij}$, from intraslab events for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Residuals are color-coded by subduction zone region. .................................................................................................................................40

Figure 5.1  Global $V_{S30}$ scaling model, $F_{lns}$, for PGV, PGA, and a range of PSA oscillator periods 0.1–10.0 sec.........................................................................................................................46

Figure 5.2  Comparison of $V_{S30}$-scaling slope between the global NGA-Subduction model, the Seyhan and Stewart [2014] (SS14) slope for active tectonic regions, and the Parker et al. [2019] (Pea19) slope for CENA.................................47

Figure 5.3  Regional $V_{S30}$-scaling model, $F_{lns}$, for 0.2-sec PSA shown in dashed line, compared to global model shown in solid line. .................................................................48
Figure 5.4  Regional $V_{S30}$-scaling model, $F_{lin}$, for 1.0-sec PSA shown in dashed line, compared to global model shown in solid line. .................................................................49

Figure 5.5  Global and regional values of $V_{S30}$-scaling slopes $s_1$ and $s_2$ [Equation (5.2)]. Regions and periods for which only $s_2$ is shown indicate that the value of $s_1$ is equivalent to that of $s_2$. ..................................................................................50

Figure 5.6  Nonlinear site model $F_{nl}$ for PGA, 0.2 sec, 1.0 sec, and 5.0 sec shown as a function of $PGA_r$ for $V_{S30}$ bins. The corresponding model from SS14 is shown for comparison using the same $f_1$ coefficient. ..................................................51

Figure 5.7  Values of parameter $f_2$ estimated using the NGA-Subduction dataset shown with the proposed relation in Equation (5.4), along with the model from SS14, empirical values of $f_2$ from NGA-West2, and simulation-based values of $f_2$ from NGA-West2. .......................................................................52

Figure 5.8  Relationship between $Z_{2.5}$ with $Z_1$ for ground-motion recording sites in Japan used in model development. The correlation coefficient between these depth parameters is 0.694. .................................................................................................53

Figure 5.9  Mean $Z_{2.5}$ ($\mu_{Z_{2.5}}$) as a function of $V_{S30}$ for (a) Japan and (b) Cascadia, used to center the basin sediment depth scaling model through the use of differential sediment depth $\delta Z_{2.5}$ ..................................................................................................................54

Figure 5.10 Relationship between within event residuals and differential depth for sites in Japan for PGA, PGV, and 5.0-sec PSA [data and model from Equation (5.5)]. Dashed line is reference at 0 residual. ..................................................55

Figure 5.11 Basin-edge outlines used in the definition of regional basin sediment-depth effects for the Pacific Northwest region of the U.S. Basin outlines modified from McPhee et al. [2014], Ramachandran et al. [2006], and Lowe et al. [2003]. Dashed line around Seattle Basin shows updated outline from Wirth et al. [2018(b)], published after the conclusion of NGA-Subduction database development. Figure modified from Ahdi et al. [2020]. ....................................................................................................................56

Figure 5.12 Trends of mean within-event residuals in Cascadia calculated without basin term ($F_b$) with period for all sites (red), basin sites or sites with known sediment depths ($Z_{2.5} > 0$; blue), and non-basin sites ($Z_{2.5} = 0$; black). Standard errors of the mean are also shown. ..........................................................................................57

Figure 5.13 Trends of Cascadia within-event residuals calculated without basin term ($F_b$) with differential depth for general Cascadia basin sites (left column) and Seattle basin sites (right column) for 0.1 sec (top), 2.0 sec (middle), and 5.0 sec (bottom). Dashed line is a reference at 0 residual. ..................................................58
Figure 5.14  Predictions of response spectra computed using the global interface reference rock GMM from Chapter 4 with the global site-amplification model described herein for $V_{S30} = 200, 400, 700, \text{ and } 1000 \text{ m/sec}$: (a) M8 event at $R_{rup} = 30 \text{ km}$, representing a condition producing nonlinear site response; and (b) M7 event at $R_{rup} = 200 \text{ km}$, representing a condition producing linear site response.

Figure 5.15  Site terms for the global GMM presented in Chapter 4 in combination with the site-amplification model presented herein, for 0.2- and 1.0-sec PSA. Data from interface (top) and intraslab events (bottom) are shown separately.

Figure 6.1  Values of $\tau$ from event terms binned in 0.5 magnitude units for PGA and 1.0-sec PSA, shown with 95% confidence intervals. Dashed line represents the value of $\tau$ from a mixed effects analysis over the full magnitude range.

Figure 6.2  Period-dependence of between-event variability as computed for each period shown with 95% confidence intervals, and as represented in aleatory variability model. Black line represents the smoothed modeled $\tau$ for forward applications.

Figure 6.3  Values of within-event variance, $\phi_2$, for PGA and 1.0-sec PSA shown with 95% confidence intervals as a function of earthquake magnitude. Dashed line represents value of $\phi_2$ from a mixed-effects regression over all data.

Figure 6.4  Values of $\phi_2$ for PGA and 1.0-sec PSA shown with 95% confidence intervals for within-event residuals binned by $R_{rup}$. Dashed line represents value of $\phi_2$ from a mixed-effects regression over all data.

Figure 6.5  Values of $\phi_2$ for PGA and 1.0-sec PSA shown with confidence intervals for within-event residuals binned by $V_{S30}$.

Figure 6.6  Model predictions for within-event standard deviation [Equation (6.3)] at 0.2-sec PSA as a function of rupture distance (left) and $V_{S30}$ (right).

Figure 6.7  Values of $\phi_{2S2S}$ for PGA (left) and 1.0-sec PSA (right) shown with 95% confidence intervals for within-event residuals binned by M. Dashed line represents value of $\phi_{2S2S}$ from a mixed-effects regression over all data.
Figure 6.8 Values of $\phi_{SS}^2$ for PGA and 1.0-sec PSA shown with 95% confidence intervals for within-event residuals binned by $R_{rup}$. Solid line shows value of $\phi_{SS,0}^2$ [Equation (6.6)].

Figure 6.9 Values of $\phi_{SS}^2$ for PGA and 1.0-sec PSA shown with 95% confidence intervals for within-event residuals binned by $V_{S30}$. Solid line shows model for $\Delta Var_{SS} (V_{S30})$ [Equations (6.6–6.7)].

Figure 6.10 Values of single-station variance $\phi_{SS}^2$ computed using magnitude-binned records for PGA and 1.0-sec PSA shown with 95% confidence intervals. Dashed line represents value of $\phi_{SS}^2$ from a mixed-effects regression over all data.

Figure 6.11 Values of single-station variance ($\phi_{SS}^2$) computed using distance-binned records for PGA and 1.0-sec PSA shown with 95% confidence intervals. Also shown is the best fit model for $\phi_{SS}^2$.

Figure 6.12 Values of single-station variance ($\phi_{SS}^2$) computed using $V_{S30}$-binned records for PGA and 1.0-sec PSA, adjusted by centering on the median $V_{S30}$ value at each period, $V_M$, with 95% confidence intervals shown as red bars. Also shown is the best fit model for $\Delta Var_{SS}$ [Equation (6.10)].

Figure 6.13 Model predictions for within-event standard deviation at 0.2-sec PSA as a function of rupture distance (left) and $V_{S30}$ (right). Black lines show predictions using the total $\phi$ model [Equation (6.3)], and red lines show predictions using the partitioned $\phi$ model [Equation (6.2)].

Figure 6.14 Model predictions for within-event standard deviation at 1.0-sec PSA as a function of rupture distance (left) and $V_{S30}$ (right). Black lines show predictions using the total $\phi$ model [Equation (6.3)], and red lines show predictions using the partitioned $\phi$ model [Equation (6.2)].

Figure 7.1 Comparison of 1.0-sec PSA distance-scaling between global and regional models for interface M9 events and 760 m/sec site condition. Cascadia model not shown as it matches the global model at M9.

Figure 7.2 Comparison of 1.0-sec PSA distance-scaling between global model and Atkinson and Boore [2003] (AB03) model for interface M9 events and 760 m/sec site condition (AB03 shown for NEHRP Site Classes B and C).
Figure 7.3  Comparison of 1.0-sec PSA distance-scaling between global / Japan models and Zhao et al. [2006] and Zhao et al. [2016(a)] models for interface $M_9$ events and 760 m/sec site condition ..................................................75

Figure 7.4  Comparison of 1.0-sec PSA distance-scaling between global model and Abrahamson et al. [2016] / BC Hydro [2012] and Abrahamson et al. [2018] (updated BCH) models for interface $M_9$ events and 760 m/sec site condition ..........................................................76

Figure 7.5  Comparison of 1.0-sec PSA distance-scaling between global model and simulation-based models from Atkinson and Macias [2009] and Gregor et al. [2006] for interface $M_9$ events and 760 m/sec site condition ..........................................................76

Figure 7.6  Comparison of 1.0-sec PSA distance-scaling between global and regional models for intraslab $M_8$ events and 760 m/sec site condition ..........................................................77

Figure 7.7  Comparison of 1.0-sec PSA distance-scaling between global / Cascadia models and Atkinson and Boore [2003] (AB03) Cascadia model for intraslab $M_8$ events and 760 m/sec site condition (AB03 shown for NEHRP Site Classes B and C) ..........................................................78

Figure 7.8  Comparison of 1.0-sec PSA distance-scaling between global / Japan models and Zhao et al. [2006] and Zhao et al. [2016(a)] models for intraslab $M_8$ events and 760 m/sec site condition ..........................................................78

Figure 7.9  Comparison of 1.0-sec PSA distance-scaling between global / Cascadia models and Abrahamson et al. [2016]/BC Hydro [2012] and Abrahamson et al. [2018] (updated BCH) models for intraslab $M_8$ events and 760 m/sec site condition ..........................................................79

Figure 8.1  Regional variations of mean estimates of interface constants ($c_{0,k}^{adj}$) adjusted for applicability with Cascadia corner magnitudes for PGA. Adjusted constants shown with their estimation errors. The recommended mean (red dashed line), weighted standard error of the mean (grey fill), and range represented by the mean ±1 weighted standard deviation (black dashed lines) are shown. .............................................................................84

Figure 8.2  Regional variations of mean estimates of intraslab constants ($c_{0,k}^{adj}$) adjusted for applicability with Cascadia corner magnitudes for PGA. Adjusted constants shown with their estimation errors. The recommended mean (red dashed line), weighted standard error of the mean (grey fill), and range represented by the mean ±1 weighted standard deviation (black dashed lines) are shown. .............................................................................84
Figure 8.3  Epistemic uncertainty in PGA from regional $c_0$ values shown in magnitude space for two Cascadia scenarios: (a) interface events with $R_{rup} = 50$ km, $V_{S30} = 760$ m/sec, and (b) intraslab events with $R_{rup} = 75$ km, $Z_{hyp} = 50$ km, and $V_{S30} = 760$ m/sec. Solid lines represent median Cascadia models, dotted lines represent ±1 weighted standard deviation of regional constants, corresponding to dotted lines in Figures 8.1 and 8.2. .........................85

Figure 8.4  Epistemic uncertainty from regional $c_0$ values about the median global model for PGA taken as ±1 weighted standard deviation (black solid and dashed lines). The median PGA model for Cascadia is shown for comparison with dashed red lines. .................................................................85

Figure 8.5  Epistemic uncertainty in regional constants (open circles) and model fit recommended for application (solid lines) for interface (left) and intraslab events (right). ........................................................................................................87
1 Introduction

1.1 NGA-SUB PROJECT

Next Generation Attenuation-Subduction (NGA-Sub) was a large, multi-year, multi-disciplinary project with the goal of producing a relational database of uniformly processed ground-motion data with supporting metadata, and a suite of global and regional ground-motion models (GMMs) for subduction zone earthquakes. This project was organized by the Pacific Earthquake Engineering Research Center (PEER) and encompasses subduction zones around the world, including those in Japan, Taiwan, the Pacific Northwest (United States and British Columbia, Canada), Alaska, the Aleutian Islands, New Zealand, Mexico, Chile, Colombia, Ecuador, and Peru; see Figure 1.1.

The NGA-Sub project had two phases: (1) database development, including compilation of uniformly processed time series, computation of ground-motion intensity measures (IMs), and development of metadata from global subduction zone earthquake events and recording sites (e.g., Kishida et al. [2020]; Contreras et al. [2020]; and Ahdi et al. [2020], respectively), and (2) model development, in which multiple teams developed models for predicting IMs using the NGA-Sub database and auxiliary materials such as ground-motion simulations.

As a part of NGA-Sub, a database of over 214,000 individual ground-motion components recorded during interface and intraslab subduction events was developed [Kishida et al. 2020]. This rich database is the largest database developed as a part of an NGA project. Typical of NGA research programs, database developers and ground-motion modelers have interacted extensively to make a vetted and quality-assured database, and teams of GMM developers interacted during model development in a manner that helped to identify problems with individual models, yet allowed for independent approaches.

The work presented in this report is part of the model development phase of NGA-Sub. The first chapter includes a literature review of the existing state of ground-motion characterization for global subduction zones. Chapter 2 is focused on database selection. Chapter 3 describes how we have constrained near-source saturation of ground motion for subduction zone events. Chapter 4 presents the GMM for the reference-site velocity condition of 760 m/sec. Chapter 5 presents the accompanying site-amplification model. Chapter 6 describes the aleatory variability model. Chapter 7 presents model verification using existing subduction zone GMMs. Chapter 8 presents implementation recommendations, model limitations, and proposed future work.
1.2 LITERATURE REVIEW

Subduction zones are the descending limbs of mantle convection cells at convergent plate boundaries [Stern 2002], where one piece of lithosphere overrides a second, less buoyant section. Two types of earthquakes are generated in these regimes: (1) interface events, which occur due to the coupling of the subducting and overriding plate; and (2) intraslab earthquakes, which occur within the subducting plate. Interface earthquakes are influenced by the age of the down-going plate; fast subduction of young and buoyant lithosphere causes stronger coupling between the plates compared to the slow subduction of old, colder plates that leads to weaker coupling. Intraslab events are influenced by the thermal state of the interior of the subducting slab [Stern 2002] and tend to have normal faulting mechanism due to extension in the plate during subduction. As shown in Figure 1.1, this type of tectonic environment occurs in many regions globally, making the resulting seismic hazard relevant for many populated areas including the Cascadia region of the U.S.

Because of the seismic hazard presented by subduction zones, the ground motions they produce have been the subject of much study, both empirical and simulation based. Early studies of empirical ground motions from subduction zone regions did not investigate regional differences in subduction ground motions nor did it distinguish between event types or did so only through adjustment of a constant term; see Atkinson [1997]; Youngs et al. [1988; 1997]; and Crouse et al. [1988]. For example, Youngs et al. [1997] presented an ergodic GMM developed using mixed effects regression of 350 recordings from Alaska, Cascadia, Japan, Mexico, Peru, and the Solomon Islands. Random effects on regional parameters were not used, and event type was considered only
As the size and reliability of ground-motion databases increased, the ergodic assumption—that ground motion should behave similarly across global regions—was disproven. Atkinson and Boore [2003] used 1200 recordings from global events within the magnitude-distance range of engineering interest to develop a GMM for subduction zones. They found significant regional differences. For example, ground-motion amplitudes in Cascadia were found to be reduced at short oscillator periods by up to a factor of two relative to those in Japan for the same event type, magnitude, source-to-site distance, and NEHRP site category. They also found that intraslab events produce larger ground motions than interface events within 100 km of the fault, but they decay faster with distance leading to smaller intraslab motions at larger distances.

Due to these global differences in ground motions, many regional GMMs for subduction zones have been developed, in particular for data-rich regions such as Japan (Si and Midorikawa [1999]; Zhao et al. [2006]; and Zhao et al. [2016(a) and (b)]) and Taiwan (Lin and Lee [2008]). In regions without much available data, such as Cascadia, simulations have been used to further inform our understanding of the earthquake ground motions and the hazard they pose.

Gregor et al. [2002] performed a suite of stochastic finite-fault simulations for $M_{8.0}$, $M_{8.5}$, and $M_{9.0}$ interface earthquakes in Cascadia. The stochastic finite-fault model was validated against the 1985 $M_{8.0}$ Michoacán, Mexico, earthquake and the 1985 $M_{8.0}$ Valparaiso, Chile, earthquake. A simple GMM was then fit to the computed intensity measures. The model predicts similar peak ground acceleration (PGA) values at short distances ($\leq 150$ km) to the Youngs et al. [1997] model, but larger PGA values at long distances due to slower distance attenuation. Atkinson and Macias [2009] also performed a suite of stochastic finite-fault simulations for Cascadia interface events with $M_{7.5}$–$9.0$. They validated their simulations using the $M_{8.1}$ Tokachi-Oki, Japan, earthquake sequence and then adjusted the simulations accounting for the average source, attenuation, and site parameters of the Cascadia region. They found that variations of input parameters due to regional differences such as distance attenuation produce large uncertainties in the resulting simulated ground motions by up to a factor of two at 100 km.

Prior to the NGA-Sub project, models were developed for application to high-impact infrastructure projects in the Cascadia region such as dams; see BC Hydro [2012] and Abrahamson et al. [2016]). These models were developed using mixed-effects regressions of an expanded dataset from Atkinson and Boore [2003], consisting of 9946 horizontal time series pairs from 292 earthquakes. Their analyses found that the same magnitude-scaling slope can be used for interface and intraslab events, and different distance-scaling slopes are needed in the forearc and backarc regions of subduction zones. Comparisons to previous models show that at short distances ($\leq 100$ km), Abrahamson et al. [2016] predictions fall within the range of existing GMMs; however, at longer distances, the model predicts lower ground motions due to faster distance attenuation. The Abrahamson et al. [2016] model is meant to be global, with a range of epistemic uncertainty in the constant term that can be used to represent regional variation in ground-motion amplitudes. The model does not have regionalized anelastic attenuation or $V_{S30}$-scaling terms.

Frankel et al. [2018] and Wirth et al. [2018(a)] produced a set of broadband (0–10 Hz) synthetic seismograms for $M_{9}$ Cascadia interface events by combining synthetic seismograms derived from three-dimensional (3D) finite-difference simulations ($\leq 1$ Hz) with finite-source, stochastic synthetics ($\geq 1$ Hz), informed by the $M_{9.1}$ Tohoku, Japan, and $M_{8.8}$ Maule, Chile,
earthquakes. For sites not in sedimentary basins, the simulated ground motions match predictions from Abrahamson et al. [2016] at 0.1–6.0 sec but are larger at longer periods. They also found that sites in Cascadia-area basins, such as the Tacoma and Seattle Basins, show site-amplification factors of 2–5 for periods 1.0–10.0 sec, which are much larger than that predicted by the NGA-W2 GMMs [Bozorgnia et al. 2014].

The Next Generation Attenuation-Subduction project began in 2014 with the goal of producing a uniformly processed ground-motion database and a suite of improved GMMs to represent epistemic uncertainties in predicted median ground motions. The remainder of this report describes the development of a semi-empirical global GMM with regional adjustment factors for interface and intraslab subduction events using the NGA-Subduction ground-motion database [PEER 2020]. The model presented herein improves upon prior work as follows:

- Uses a much larger dataset;
- Considers regionalization in the ground-motion amplitude, anelastic attenuation, magnitude-scaling, $V_{30}$-scaling, and sediment depth terms;
- Treats the amplitude and distance-, magnitude-, and depth-scaling terms differently between interface and intraslab event types; and
- Includes within the aleatory variability model important dependencies of within-event variability on rupture distance and site condition, and distinguishes the single-station and site-to site variances from the total within-event variance.
2 Data Selection

The NGA-Sub database contains more than 71,000 three-component time series from 1883 earthquakes acquired from subduction zone regions around the world. The overall relational database combines three individual databases: an earthquake source database, an earthquake recording database, and a recording station database. The combined relational database can be used to extract a single summary “flatfile” with one line per recording for use in the development of GMMs [Mazzoni et al. 2020(a)].

The portions of the database pertaining to ground motions are discussed in detail by Kishida et al. [2020]. A subset of records from the NGA-Sub database was used for model development. The records selected satisfy all of the following criteria:

1. Metadata necessary for model development are available [M, rupture distance (Rrup), hypocentral depth (Zhyp), V30, etc.];
2. Earthquake classified with high confidence as being interface, intraslab, or in the lower double seismic zone;
3. Earthquake is a class 1 event (i.e., a mainshock) according to Wooddell [2018] method 2 using an 80-km cutoff distance;
4. Rrup ≤ min (Rmax, 1000 km), where Rmax is a maximum distance limit set based on seismic network properties [Contreras et al. 2020].
5. Sensor depth ≤ 2m;
6. Interface events with hypocentral depths (Zhyp) ≤ 40 km and intraslab events with Zhyp ≤ 200 km;
7. Pseudo-spectral acceleration at oscillator periods T ≤ TLU, where TLU refers to the longest useable period based on the corner frequencies used to process the record;
8. Earthquake epicenter and the station are both located in the forearc region;
9. Earthquakes without multiple event flags; these are events for which the recordings appear to indicate that more than one seismic source affected the ground motions;
10. Earthquakes with source review flags = 0, 1, 2 or 4, which indicate earthquakes that underwent quality control checks and meet metadata quality standards;
11. Records without a late P-wave trigger flag, which indicate recordings where the P-wave arrival was not missed;
12. After applying criteria 1–11, we only used records from events with at least three recordings.

Screening for instrument housing was not performed because the NGA-Sub data collection sought to collect data from only those instruments judged to be effectively free field (instrument huts and small buildings without foundation embedment [Ahdi et al. 2020]).

The magnitude-distance distribution of records for PGA from interface and intraslab events after application of the above screening is shown in Figure 2.1. The data in Figure 2.1 are plotted with identification of the major regions that contribute data to NGA-Sub and for which regional effects were considered in model development. The number of events and recordings used for model development varied as a function of period, with a range of 3215–6374 records and 90–122 events for combined data from both event types; see Figure 2.2.

![Figure 2.1 Magnitude-distance distribution of recordings from interface (left) and intraslab (right) events, color-coded by region.](image)
Figure 2.2 Number of events and number of recordings selected for model development according to the criteria in Chapter 2 for interface (circles) and intraslab (triangles) events. Combined values are shown as diamonds. The left panel shows metrics for PGV (-1) and PGA (0).
3 Near-Source Saturation

3.1 MOTIVATION

It is typical for ground-motion path models, including those used herein, to express the decay of the natural log of ground-motion intensity (\(\ln IM\)) with the natural log of a distance metric (\(\ln R\)) that combines the rupture distance (closest distance from a point to the source, \(R_{rup}\)) with a near-source saturation term, \(h\) as follows:

\[
R = \sqrt{R_{rup}^2 + h^2}
\]

(3.1)

The use of \(h\) in the expression for \(R\) causes ground motions to saturate (reach an approximately constant level) at close rupture distances where \(R_{rup} < h\). This term—also sometimes called the finite-fault term or fictitious depth—is necessary due to two potential geometric effects [Yenier and Atkinson 2014; Rogers and Perkins 1996]:

1. Whereas path models are based on the distance to the closest part of the finite fault (at distance \(R_{rup}\)), other portions of the fault at greater distance also contribute to the observed ground-motion amplitudes. This can be accommodated by an added distance term such as with \(h\) in Equation (3.1).

2. To the extent that ground motions are controlled by slip on the closest part of the fault, the slip at that location would, in isolation from other parts of the fault, correspond to a smaller seismic event than the full earthquake rupture; essentially, a nearby site can only see part of the fault. This effect can also be accommodated using Equation (3.1).

To account for such effects, a near-source saturation term is typically included in the distance metrics used in GMMs and in equivalent point-source simulations [Boore 2015; Yenier and Atkinson 2014]. Initially we considered using the subduction data to constrain the near-source saturation; however, due to the typical offshore or deep locations of subduction earthquakes and the lack of recording stations at close source-to-site distances (Figure 2.2), there are not enough data close to the source to constrain this feature; in other words, the near-source saturation term is under-determined.

This is shown by an example in Figure 3.1, in which data from NGA-Sub earthquake ID = 4000068 (\(M8.29\)) are plotted with various values of \(h\) (using data adjusted to a 760 m/sec site condition using the Seyhan and Stewart [2014] model, hereafter referred to as SS14). The path models are determined using two different values of \(h\) but provide equally good fits. Additionally, because models for \(h\) are magnitude-dependent, attempts to regress them from data are difficult
because of trade-offs with other magnitude-dependencies in the data, such as magnitude scaling and magnitude-dependent geometrical spreading. This is not a new problem in GMM development for subduction regions; previous empirical determination of this portion of GMMs has been primarily for active tectonic regions (e.g., Boore et al. [2014(b)]; Yenier and Atkinson [2014]; and Abrahamson et al. [2014]); GMMs for subduction zones have mostly borrowed this portion of the model from other regions or used simulations to constrain it.

Therefore, to develop a model for $h$, we used a combination of empirical estimates from active tectonic regions at small-to-large magnitudes [Atkinson et al. 2016; Yenier and Atkinson 2014], and a suite of EXSIM simulations of moderate-to-large subduction interface events performed as part of the present work; EXSIM is an open-source stochastic finite-source simulation algorithm [Motazedian and Atkinson 2005; Boore 2009; and Assatourians and Atkinson 2012]. Details on both constraints are given in the following sections. As shown subsequently, $h$ can generally be described using a log-linear model over a certain magnitude range:

$$\log_{10} h = h_1 + h_2 M$$  \hspace{1cm} (3.2)

**Figure 3.1**  Example of the underdetermined problem of fitting $h$ using NGA-Sub data. Two simple path models [Equation (3.3)] fit to an interface event from Japan with $M = 8.29$, one with $h = 10$, and the other with $h = 30$ km. Both models produce a fit with negligible differences in the residual standard error (0.6138 vs. 0.61, respectively).
3.2 AVAILABLE EMPIRICAL CONSTRAINTS

Yenier and Atkinson [2014] (YA14) used records from 11 shallow earthquakes in global active tectonic regions with $M \geq 6$ to study near-source saturation, with the goal of determining if the equivalent point-source framework is able to describe observed ground motions at short distances. They fit distance-scaling functions to each event, and observed a magnitude dependence in their best-estimate values of $h$, something that was not widely incorporated into GMMs at the time (e.g., Boore et al. [2014(a)], hereafter referred to as BSSA14), Campbell and Bozorgnia [2014], and Chiou and Youngs [2014] do not have magnitude-dependent near-source saturation). They produced a model for $h$ in the form of Equation (3.2), with $h_1 = -1.72$ and $h_2 = 0.43$, which is plotted in Figure 3.2.

Yenier and Atkinson [2015(a)] examine the NGA-W2 database [Ancheta et al. 2014] to estimate the best-fit source parameters for each California earthquake using matching between empirical and simulated response spectra. As part of this work, they considered $h$ models from Yenier and Atkinson [2014] and Atkinson and Silva [2000], along with event-specific empirical estimates of $h$ from YA14, Boore et al. [2014(b)], and earthquakes in the Christchurch, New Zealand, sequence; see Figure 3.2. They propose a parameterization using Equation (3.2), with $h_1 = -0.405$ and $h_2 = 0.235$; see Figure 3.2. This parameterization was used by Yenier and Atkinson [2015(b)] in the development of their equivalent point-source based GMM for central and eastern North America (CENA).

Figure 3.2 A comparison of event-specific empirical estimates of near-source saturation term $h$ (see text for description of studies). Also shown are models for $h$ as a function of earthquake moment magnitude from Yenier and Atkinson [2014; 2015] (YA14, YA15, respectively), and Atkinson and Silva [2000] (AS00).
Atkinson et al. [2016] examined a number of small induced events that are well recorded at short source-to-site distances from the Geysers region of California [Sharma et al. 2013] to better constrain near-source saturation effects for small magnitude earthquakes ($M_{1.5}$–3.6). They applied a similar method of fitting event-specific distance-scaling as was applied in Yenier and Atkinson [2014]. As shown in Figure 3.2, their results appear to be consistent with the near-source saturation model of Yenier and Atkinson [2015(b)].

### 3.3 FINITE-FAULT SIMULATIONS

To constrain $h$ at the large magnitudes necessary for subduction zone GMMs, we ran a suite of ground motion simulations using EXSIM. Input parameters were chosen to be compatible with properties of interface events in the NGA-Sub database; this work was initiated by Nicolas Kuehn [Kuehn, Personal Communication, 2018] and modified for the present application. Simulations were performed for earthquakes with $M = 3.75$–9.5 in 0.25-magnitude unit intervals, with five runs per magnitude. For each run, the fault length and width were generated randomly using fault-source scaling relationships developed by Strasser et al. [2010], and hypocenter locations on the fault plane were randomly sampled with a uniform distribution over the fault plane. Stress drop was taken as 150 bars. Depth to the top of rupture was taken as 5 km to maintain a small source-to-site distance, and a fault dip between $15^\circ$–$28^\circ$ was randomly assigned, which is in line with estimates of dip for interface events in the database. As shown in Figure 3.3, ground motions were generated at 36 sites located at 12 distances between 10 and 1000 km along three azimuths ($45^\circ$, $60^\circ$, and $90^\circ$). The simulations used a simple distance attenuation model, with a geometrical spreading coefficient of -1.3 (i.e., $\ln IM \propto 1.3 \ln R$) based on initial observations of empirical distance scaling.

![Figure 3.3 Schematic showing fault and EXSIM finite-fault simulation station array. Stations are located at distances of 10, 20, 30, 40, 50, 70, 100, 200, 300, 500, 700, and 1000 km and azimuths of $90^\circ$, $60^\circ$, and $45^\circ$ (not to scale).](image-url)
Once the simulated ground motions were generated, PSA values from five runs per magnitude were combined, and a simple path model was fit for each magnitude bin,

\[ F_{P, \text{EXSIM}} = c_0 + c_1 \ln R + c_2 R \]  

(3.3)

where \( R \) is defined by Equation (3.1). Coefficients \( c_1 \) and \( c_2 \), representing geometrical spreading and anelastic attenuation effects, respectively, were initially fit using simulated ground motions at \( R_{rup} \geq 40 \text{ km} \) to avoid the influence of near-source saturation effects at closer distances. Then, with \( c_1 \) and \( c_2 \) fixed, \( h \) and \( c_0 \) were fit using the simulated ground motions over the entire distance domain.

Figure 3.4 shows the resulting median \( h \) values along with their 95% confidence intervals as a function of magnitude for PGA, peak ground velocity (PGV), and pseudo-spectral acceleration (PSA) at \( T = 0.5 \) and 3.0-sec PSA. Also shown in Figure 3.4 are the YA14 and YA15 models, which are empirically constrained for crustal earthquakes to \( M \sim 7.5 \). The estimates of \( h \) from EXSIM fall significantly below those for the YA14 and YA15 models for events of \( M > 6.5 \). The EXSIM estimates do not exhibit systematic variations with period.

Figure 3.4 shows our proposed near-source saturation model for interface events [Equation (3.2)], which is period-independent \( (h_1 = -0.82, h_2 = 0.252) \). We constrained our model to be similar to the values of YA14 and YA15 for magnitudes up to 5.5. However, we departed from those models and followed the trend of the EXSIM results for 5.5 < \( M \) ≤ 9.5; see Figure 3.5. In this magnitude range, the proposed model is more similar to the results of Atkinson and Silva [2000].

![Figure 3.4](image)

**Figure 3.4** Comparison of the Yenier and Atkinson [2014; 2015] near-source saturation models (labeled YA14 and YA15), \( h \) estimates from EXSIM simulations run to emulate subduction interface events, and the near-source saturation model developed in this study, given by Equation (3.2) with \( h_1 = -0.82 \) and \( h_2 = 0.252 \).
3.4 MODIFICATION FOR SLAB EVENTS

Originally we used the same near-source saturation model for both interface and intraslab events, assuming that the paucity of data at small distances would not allow for discrimination between event types. However, residuals analyses revealed that the slab GMM over predicted median ground motions by a factor of ~1.3 for large slab events (M6.5+) at short distances (≤ 75 km), indicating that the $h$ model given by Equation (3.2) (with the recommended coefficients of $h_1 = -0.82$ and $h_2 = 0.252$) may be saturating at ruptures distances that are too small.

Removing this bias required increasing $h$ in the M6.5–7 range. Once such adjustments were made, very large values of $h$ were given by the exponential function Equation (3.2) for larger values of $M$, which we considered to be non-physical if $h$ is related to fault-plane dimensions. Having no data to constrain a maximum value of $h$ for intraslab events in the M7+ region (e.g., Figure 2.2), we examined existing simulation results for intraslab events from Ji and Archuleta [2018].

As shown in Figure 3.5, we fit Equation (3.3) to the Ji and Archuleta [2018] simulated-response spectral values for a M8 slab earthquakes in Japan using alternate fixed values of $h = 29$ km (value from YA15 for M8), 35 km, 40 km, and 50 km. Based on visual inspection, the simulation results reject the model fit with $h = 50$ km. Advised by these findings, we adopted an upper limit of $h = 35$ km as it seems to be in the best agreement with the simulation results at the closest distances. We enforced this upper limit at the regional corner magnitude ($m_c$) used in the magnitude-scaling model (described in Chapter 4). We re-fit Equation (3.2) over the magnitude range of M4–$m_c$ such that the resulting model predicts similar values of $h$ as YA15 at M4 and $h = 35$ km at $m_c$. This is achieved with the following expression for $h$ (in units of km) for use with slab events:

$$h = \begin{cases} 10^{\left(\frac{1.050}{m_c-M}\right)\left(M-m_c\right)+1.544} & M \leq m_c \\ 35 & M > m_c \end{cases}$$

(3.4)

The regional corner magnitude $m_c$ in the function for $h$ was chosen in part to produce a smooth model in ground motion–magnitude space, consistent with the magnitude-scaling corner. We argue that this choice also has a physical basis. Values of intraslab $m_c$ are based on the seismogenic thickness of subducting slabs, which depend on the temperature gradient in the slab and are closely related to crustal age [Stern 2002; Ji and Archuleta 2018]. For subduction zones with a larger seismogenic thickness, the relatively large values of $m_c$ increase the likelihood of a large fraction of the rupture occurring directly beneath a site rather than most of the rupture being located some distance down-strike. This would tend to increase the magnitude at which magnitude-saturation occurs (for use in the M-scaling model) and increase ground motions within the distance range controlled by saturation. The latter effect (increased ground motions within the saturation region) is provided by Equation (3.4) when $m_c$ is increased because $h$ is reduced for $M < m_c$; hence the connection of the break magnitude in the $h$ model to $m_c$ has physical justification. This behavior is captured in the slab $h$ term, as subduction zones with larger values of $m_c$ have smaller $h$ values.
3.5 COMPARISONS TO EXISTING MODELS

Figure 3.6 compares the following models: (1) the proposed near-source saturation models for interface [Equation (3.2)] and slab [Equation (3.4)] earthquakes; (2) the YA14 and YA15 models described in Section 3.2; (3) the Atkinson and Silva [2000] model (also described in Section 3.2); (4) the model used in the BC Hydro subduction-zone GMM [Abrahamson et al. 2016]; and (5) the model used in the Atkinson and Boore [2003] subduction-zone GMM. The Abrahamson et al. [2016] near-source saturation model consists of a magnitude-dependent fictitious depth constrained using the finite-fault simulations from Gregor et al. [2002]. The Atkinson and Boore [2003] near-source saturation term is constrained based on fault geometry. First, they defined a magnitude-dependent fault with length and width given by Wells and Coppersmith [1994]. Then, for a given $R_{rup}$, they computed a series of distances between the site and gridded locations on the plane; the average of these is taken as $R$ [Equation (3.1)]. Using the values of $R$ and $R_{rup}$, they solved for $h$.

The near-source saturation models produce similar results for $M \lesssim 6$, with the range at $M6$ being about 7 km. Above $M6$, the Atkinson and Boore [2003] model (YA14) and proposed slab models increase more rapidly with magnitude than the remaining models. The Atkinson and Silva [2000] model and proposed interface models have the shallowest slope, with the YA15 and Abrahamson et al. [2016] models being somewhat intermediate. Given the lack of near-source empirical constraints for subduction zone earthquakes in particular, the spread at large magnitudes is not surprising. None of the models shown in Figure 3.6 are empirically based from subduction-zone ground motions. Rather, the models are constrained by a combination of data from active
tectonic regions, finite-fault simulations, and empirical fault geometry relationships; the spread of the models is indicative of epistemic uncertainty in this model parameter. It is noteworthy that the reason Yenier and Atkinson revised their model from 2014 to 2015 was to prevent over-saturation of ground motions at close distances for very large events (i.e., to prevent near-source amplitudes from decreasing with increasing magnitude).

Figure 3.6 Comparison of published near-source saturation models. The models proposed in this work for interface and slab earthquakes are shown as red lines. The suite of models shown in blue are constrained using large earthquakes from active tectonic regions [Yenier and Atkinson 2014] (YA14); Yenier and Atkinson [2015] (YA15); and Atkinson and Silva [2000] (AS00). The models shown in black are from published subduction zone GMMs [Abrahamson et al. 2016] (BC Hydro) and Atkinson and Boore [2003] (AB03)].
4 Development of Median Model

The first section of this chapter presents the suite of median model functional forms in their entirety for implementation purposes, including the interface, intraslab, and regional models. Section 4.2 provides a high-level overview of the regression procedure, and Section 4.3 gives details on the development of each model term, including the distance-, magnitude-, and source-depth-scaling. Although the site term is included in Sections 4.1–4.2 for completeness, details on development are given in Chapter 5. The aleatory variability model is given in Chapter 6.

4.1 MEDIAN MODEL FUNCTIONAL FORM

4.1.1 Global and Regional Models

Due to differences in path and source-scaling attributes, separate global GMMs are provided for interface and intraslab earthquakes. Both models share a common functional form, with some coefficients being the same for both event types, while other coefficients vary. Each median model has five terms: a constant \( c_0 \) that controls the overall amplitude of the predicted ground motion; a path term \( F_P \) that describes the decay of ground motion with distance; a magnitude-scaling term \( F_M \); a source depth-scaling term \( F_D \); and a site-amplification term \( F_S \). These terms are additive in natural log space:

\[
\mu_{ln,y} = c_0 + F_P + F_M + F_D + F_S \quad (4.1)
\]

The path term \([\text{Equation (4.2)}]\) consists of a magnitude-dependent geometrical spreading term that represents the purely geometrical effect of the spreading of energy as seismic waves propagate from a point source along a spherical wave front, and an anelastic attenuation term that represents the per-cycle damping as seismic waves pass through the earth. The \( F_P \) term accounts for near-source saturation using parameter \( h \) \([\text{Equation (4.4)}]\), which is combined with site-to-source distance metric \( R_{rup} \) as described in Section 3 \([\text{Equations (3.1)}/(4.3)]\). The near-source saturation term \( h \) depends on event type.

\[
F_P = c_1 \ln R + b_4 M \ln R + a_0 R \quad (4.2)
\]

\[
R = \sqrt{R_{rup}^2 + h^2} \quad (3.1/4.3)
\]

\[
h = 10^{-0.82+0.252M} \quad \text{(interface events)} \quad (4.4a)
\]
The magnitude-scaling term is a piecewise function with parabolic and linear segments, transitioning at a corner magnitude $m_c$:

$$F_M = \begin{cases} 
  c_4(M - m_c) + c_5(M - m_c)^2 & M \leq m_c \\
  c_6(M - m_c) & M > m_c
\end{cases} \quad (4.5)$$

The source-depth scaling term is a tri-linear function conditioned on hypocentral depth, with two corner depths $d_{b1}$ and $d_{b2}$:

$$F_D = \begin{cases} 
  m (d_{b1} - d_{b2}) + d & Z_{hyp} < d_{b1} \\
  m (Z_{hyp} - d_{b2}) + d & d_{b1} < Z_{hyp} \leq d_{b2} \\
  d & Z_{hyp} > d_{b2}
\end{cases} \quad (4.6)$$

where $d_{b1} = 20$ km and $d_{b2} = 67$ km for intraslab events. There is no source-depth scaling for interface events; i.e., $F_D = 0$. The model is conditioned on the hypocentral depth ($Z_{hyp}$), which as described in Section 4.3.3 can be replaced with a mean depth ($\bar{Z}_{hyp}$) that depends on the depth to top of rupture ($Z_{tor}$), fault width, and fault dip angle, and is easier to apply than $Z_{hyp}$ in seismic hazard analyses. This mean depth can be used in place of hypocentral depth without change to the model coefficients or appreciable change to the between-event variability $\tau$.

The ergodic site term, $F_S$, is comprised of three components: (1) a linear term ($F_{lin}$) that represents the site amplification at small strains; (2) a nonlinear term ($F_{nl}$) that accounts for attenuation of high-frequency components of ground motion from the additional soil damping that occurs under strong shaking conditions at soil sites; and (3) a basin-depth term ($F_b$) that approximately represents the site-amplification effects that can occur in deep sedimentary basins, such as basin edge-generated surface waves and focusing. The three terms are summed in natural logarithm space:

$$F_S = F_{lin} + F_{nl} + F_b \quad (4.7)$$

The functional forms for each of the terms are given in Equations (4.8) to (4.13). The linear term is tri-linear in $V_{S30}$ space.

$$F_{lin} = \begin{cases} 
  s_1 \ln \left( \frac{V_{S30}}{V_1} \right) + s_2 \ln \left( \frac{V_1}{V_{ref}} \right) & V_{S30} \leq V_1 \\
  s_2 \ln \left( \frac{V_{S30}}{V_{ref}} \right) & V_1 < V_{S30} \leq V_2 \\
  s_2 \ln \left( \frac{V_{S30}}{V_{ref}} \right) & V_{S30} > V_2
\end{cases} \quad (4.8)$$

Site amplification scales linearly with the natural log of $V_{S30}$ between corner velocities $V_1$ and $V_2$, going through zero amplification (in natural log units) at $V_{ref} = 760$ m/sec. Only data from Taiwan show a break in slope at $V_1$ similar to that observed previously in Japan [Campbell and Bozorgnia 2014] and CENA [Hassani and Atkinson 2017; Parker et al. 2019]. In other words, for most regions
\[ F_{nl} = f_1 + f_2 \ln \left( \frac{PGA_r + f_3}{f_3} \right) \]  
(4.9)

where \( f_i \) is zero, meaning that the effect of nonlinearity disappears as \( PGA_r \) goes to zero, \( f_3 \) is taken as 0.05g across all periods, and \( f_2 \) is defined as in Chiou and Youngs [2008], with some coefficient changes:

\[ f_2 = f_4 \left[ \exp \left( f_5 (\min(V_{S30}, 760) - 200) \right) - \exp \left( f_5 (760 - 200) \right) \right] \]  
(4.10)

For Japan and Cascadia, we define a period-dependent basin-depth term, \( F_b \), that is given in Equations (4.11)–(4.13). Elsewhere, the term is zero. \( F_b \) is conditioned not on the depth itself, but on the depth differential relative to a regionally appropriate mean depth for a given \( V_{S30} (\mu_{Z2.5}) \); the rationale for this choice is that the \( V_{S30} \)-conditioned models (\( F_{lin} \) and \( F_{nl} \)) reflect an implicit sediment depth effect. The \( F_b \) term is intended to capture mean differences in basin effects for \( Z_{2.5} \neq \mu_{Z2.5} \) (i.e., \( \delta Z_{2.5} \neq 0 \)), and is represented by a tri-linear equation in \( \delta Z_{2.5} \) space, controlled by slope \( e_3 \), and the values of \( F_b \) at \( \delta Z_{2.5} = e_1/e_3 \) and \( e_2/e_3 \), which are \( e_1 \) and \( e_2 \), respectively.

\[ F_b = \begin{cases} 
  e_1 & \delta Z_{2.5} \leq \frac{e_1}{e_3} \\
  e_2 & \frac{e_1}{e_3} < \delta Z_{2.5} < \frac{e_2}{e_3} \\
  e_3 & \delta Z_{2.5} \geq \frac{e_2}{e_3}
\end{cases} \]  
(4.11)

\[ \delta Z_{2.5} = \ln(Z_{2.5}) - \ln(\mu_{Z2.5}(V_{S30})) \]  
(4.12)

When \( e_3 = 0 \), \( F_b \) should be taken as zero. The centering relationship is a variation on the Nweke et al. [2018] \( \mu_{Z2} \) model: an error function with coefficients \( \nu_\mu, \nu_\sigma, \theta_0, \) and \( \theta_1 \) adjusted using the NGA-Subduction database:

\[ \ln(\mu_{Z2.5}) = \ln(10) \times \theta_1 \left[ 1 + \operatorname{erf} \left( \frac{\log_{10}(V_{S30}) - \log_{10}(\nu_\mu)}{\nu_\sigma \sqrt{2}} \right) \right] + \ln(10) \times \theta_0 \]  
(4.13)

For Cascadia, the coefficients in Equation (4.13) are: \( \theta_0 = 3.94, \theta_1 = -0.42, \nu_\mu = 200 \text{ m/sec}, \) and \( \nu_\sigma = 0.2 \). For Japan, the coefficients in Equation (4.13) are: \( \theta_0 = 3.05, \theta_1 = -0.8, \nu_\mu = 500 \text{ m/sec}, \) and \( \nu_\sigma = 0.33 \).

The basis for this model and coefficients used for Japan and Cascadia are given in Chapter 5. Coefficients for the global interface model are given in Table E1 of the electronic supplement, and coefficients for the global intraslab are given in Table E2. The development of the constant, distance-scaling, magnitude-scaling, and source depth-scaling terms are discussed in Section 4.3. The development of the site term is described in Chapter 5.

### 4.1.2 Regional Terms and Coefficients

Several of the model coefficients in the global GMM are regionalized: the constant \( c_0 \), the anelastic attenuation coefficient \( c_0 \), the magnitude corner \( m_c \), the \( V_{S30} \)-scaling model coefficients \( s_1, s_2, V_1, \)
and $V_2$, and all of the basin-depth term coefficients. The basin-depth term coefficients $e_2$ and $e_3$ are further regionalized for different Pacific Northwest basin structures (e.g., Seattle versus Tacoma basins).

Table E1 in the electronic supplement gives regional coefficients for the interface GMMs for use in Alaska and the Aleutian Islands, Cascadia, Central America and Mexico, Japan, South America, and Taiwan, respectively. Table E2 in the electronic supplement gives regional coefficients for the intraslab GMMs for use in Alaska, the Aleutian Islands, Cascadia, Central America and Mexico, Japan, South America, and Taiwan, respectively.

For forward use in regions other than those directly considered in model development, we recommend using the global model, with a range of epistemic uncertainty that represents, at a minimum, the effects of regional variations in the anelastic attenuation and $V_{S30}$-scaling coefficients, as well as the constant term and corner magnitude (which are correlated). The correlation of the constant and $m_c$ means that as $m_c$ increases, the constant must increase to maintain the same ground motion level for $M < m_c$. For this reason, we recommend identifying a regionally appropriate range of $m_c$ for the application region, and then defining corresponding values of the constant that provide the desired level of epistemic variability in median ground motion space. Additional details and recommendations on this issue are given in Section 8.2.

### 4.2 OVERVIEW OF REGRESSION PROCEDURE

In order to capture physically meaningful trends in the data due to source, path, and site processes and avoid significant trade-offs in model coefficients, model development took place in a series of steps in which we progressively constrained more aspects of the model. The generalized scheme is described below, and details are provided in Section 4.3.

1. Adjust ground-motion data to a single reference-site condition, $V_{S30} = 760$ m/sec, using the Seyhan and Stewart [2014] (SS14) site-amplification model from NGA-West2 (denoted $F_{str}^{S}$);

2. Constrain the near-source saturation model using empirical findings from active tectonic regions, and simulations of larger interface and slab events (Sections 3.3–3.4);

3. Fix the magnitude-dependent geometrical spreading [coefficient $b_4$, Equation (4.2)] based on analyses of simulated ground motions, and then use a mixed-effects regression of data with $R_{rup} \leq 100$ km to estimate the magnitude-independent geometrical spreading [coefficient $c_1$, Equation (4.2)]. The distance limit was applied to avoid effects of anelastic attenuation at larger distances. The constraint on $b_4$ was applied to avoid tradeoffs with the magnitude-scaling function;

4. Compute event terms relative to the path model [Equations (4.2–4.4)] to identify source-scaling trends, which led to the selected function [Equation (4.5)]. Fit the coefficients in Equation (4.5) and in the source-depth scaling function [Equation (4.6)] based on features of these event terms. These analyses account for regional differences in $m_c$ as given by Ji and Archuleta [2018] and Campbell [2020];

5. Estimate regional constants $c_0$ [Equation (4.1)] as random effects through a mixed-effects regression on residuals computed using the sum of Equations (4.2–4.6). Derive
constants for the global model to represent the average of regional models for \( M < m_c \) in ground motion space at the approximate center of the data in magnitude-distance space;

6. Use the NGA-Subduction database to check and re-calibrate the linear and nonlinear site-amplification terms [Equations (4.8–4.10)], which produces modifications to SS14;

7. Perform residuals analyses using the revised site-amplification model, in combination with other model components, to adjust the constant term, \( c_0 \), as needed; and

8. Visually inspect coefficients and, if necessary, smooth in the oscillator-period domain using a five-point triangular-weighted running mean. This has the effect of smoothing response spectral shapes provided by the median model.

4.3 DEVELOPMENT OF MODEL TERMS

4.3.1 Distance-Scaling

The development of the near-source saturation model \([h; \text{Equations (4.3–4.4)]}\) is described in Chapter 3. This section describes the fitting of the remaining elements of the path model, \( F_P \) [Equation (4.2)].

The path model has two components: a geometrical spreading term and an anelastic attenuation term. The geometrical spreading term represents the decay of energy as it moves from a point source along a spherical wave front. In an idealized homogeneous elastic half-space, the energy at any point on the radius of the sphere will decay as \( R^{-1} \); however, heterogeneities in the earth produce scattering, reflections, refractions, and wave-type conversions. As a result, the empirical exponent [i.e., \( c_1 \) in Equation (4.2)] is not unity. The transition from Fourier amplitude spectra (FAS) to response spectra (RS) introduces a magnitude-dependence in this term [Yenier and Atkinson 2014; Hassani and Atkinson 2018(a)], which can be represented as \( (b_3 + b_4 M) \ln R \). The \( b_3 \) term is not used in Equation (4.2) as it is combined with \( c_1 \). All three of these terms affect the shape of decay of ground motion with the natural logarithm of rupture distance; see Figure 4.1.

The anelastic attenuation term represents the per-cycle energy dissipation; it is a property of the material through which the seismic wave is traveling. This term controls curvature in the decay of natural log ground motion with the natural log of rupture distance, which strongly influences the rate of attenuation at large distances; see Figure 4.1.
Figure 4.1  Schematic of path model, $F_P$ [Equation (4.2)], showing the near-source saturation model, $h$, the geometrical spreading slope, and the curvature due to anelastic attenuation.

To fit the path model independently from the source terms, we use a two-step regression of $F_P$ terms in Equation (4.2) in the style of Joyner and Boore [1993; 1994]. The path coefficients derived from this process ($c_1, b_4$) are the same across all magnitudes, but a preliminary event term, $\eta e'$, was determined for each individual event and IM. These event terms, which have a trend with $M$, were considered in the subsequent derivation of source terms (Sections 4.3.2-3). Originally, we adopted values of $b_4$ from the simulation-based GMM of Hassani and Atkinson [2018(a)] (hereafter referred to as HA18). HA18 took the generic point-source simulation-based GMM of Yenier and Atkinson [2015(a)] and modified it to enable adjustments of ground motions to reflect the impact of variable site attenuation parameter $\kappa$. However, our analyses indicated that the values of $b_4$ from HA18 were too small to adequately capture the magnitude-dependent component of geometrical spreading observed in the data, especially for intraslab events. We set $b_4 = 0.1$ for both event types, which improved the model fit to the observations. A comparison of the HA18 $b_4$ coefficients and the selected value is shown in Figure 4.2. The $c_1$ coefficient, which represents the magnitude-independent portion of geometrical spreading, is set empirically via regression.
Despite the large size of the NGA-Sub database, it is not possible to constrain both the slope and curvature of the path model simultaneously due to substantial trade-offs between these two model components. We address this by fitting $c_1$ to the subset of data with $R_{rup} \leq 100$ km to avoid the portion of the data with the most curvature at large distances; see Figure 4.3. These analyses were performed using $b_4 = 0.1$. Different values of $c_1$ were derived for interface and intraslab events. As shown in Figure 4.3, data from intraslab events show steeper geometrical spreading in comparison to data from interface events, consistent with the results of previous studies (e.g., Atkinson and Boore [2003] and Abrahamson et al. [2016]).

With the geometrical spreading coefficients fixed, the anelastic attenuation coefficient, $a_0$, was fit as a regional effect and both a global value and regional values were produced. All values of $a_0$ were smoothed with respect to period and constrained to go to zero at 10 sec, as the per-cycle damping at long oscillator periods is negligible. As shown in Figure 4.4, the anelastic attenuation rate is slower for intraslab events than for interface events, indicating that although the interface data show less overall distance attenuation, there is more curvature in the intraslab data at large distances. The smoothed global and regional values of $a_0$ for interface and intraslab events are shown in Figure 4.5. In general, the anelastic attenuation in Central America and Mexico is less than that given by the global value (absolute value of $a_0$ is smaller), the anelastic attenuation in South America and Alaska is close to the global value, and the anelastic attenuation in Cascadia, Japan, and Taiwan is larger than the global value.
**Figure 4.3** Example of PGV data for magnitude bin 6.5–7 over the distance range ($R_{rup} \leq 100$ km); decay is controlled by geometrical spreading.

**Figure 4.4** Example of PGV data for magnitude bin 6.5–7 over the entire model distance domain ($R_{rup} \leq 1000$ km) with the geometrical spreading terms from Figure 3.6, plus the best fit global anelastic attenuation term. Results shown for interface and intraslab events.
Figure 4.5: Anelastic attenuation coefficient, $a_0$, as a function of oscillator period for interface events (top) and intraslab events (bottom). Lack of data for interface events in Cascadia means there is no regional value of $a_0$; instead, the global value is recommended.

4.3.2 Magnitude Scaling

Once the path model was set [Equations (4.2–4.4)], we used event terms, $\eta_E'$ to visualize trends in the data with respect to magnitude, which informed the formulation of the magnitude-scaling model. Event terms represent the average bias over all ground-motion recordings for one event relative to a particular model; following the initial regression stage, magnitude dependence is present in the event terms. To model this dependence, Equation (4.5) was first fit to the data with
the parabolic term $c_5 = 0$, allowing the two linear slopes ($c_4$, $c_6$) to be set by regression. Values of $m_c$ were constrained based on geometrical considerations specific to each subduction zone region. In the case of intraslab earthquakes, down-dip width of the event is limited by slab thickness. Events that rupture through the full slab thickness are expected to saturate (i.e., ground-motion scaling with magnitude slows) when the rupture aspect ratio (i.e., ratio of the along-strike length to down-dip width) exceeds unity. This occurs because increasing magnitude produces increasing rupture far from the site, which would be expected to have little impact on high-frequency ground motions (i.e., saturation). This feature is presented in simulated ground motions by Ji and Archuleta [2018], who also provided saturation magnitudes for slab earthquakes specific to each region considered in NGA-Sub. We take these saturation magnitudes as $m_c$ for use in Equation (4.5) for intraslab events. Saturation magnitudes for interface events are given by Campbell [2020], who used estimates of the seismogenic interface width to constrain the saturation magnitude. Global values of $m_c$ were taken as weighted averages over the values for regions considered. Table 4.1 provides these saturation magnitudes for each considered region.

With $m_c$ fixed in this manner, the linear magnitude-scaling coefficients were treated as fixed effects, and the constant $c_0$ was treated as a random effect conditioned on region and NGA-Sub earthquake ID. These coefficients were constrained to enforce $c_6 \leq c_4$, which ensures slowing of the magnitude-scaling for $M > m_c$. Lastly, the parameter that controls the parabolic behavior of the model below the break point, $c_5$, was fit to the event terms with all other coefficients fixed to their values from the first regression iteration; see Figure 4.6.

<table>
<thead>
<tr>
<th>Region</th>
<th>Interface $m_c$ [Campbell 2020]</th>
<th>Intraslab $m_c$ [Ji and Archuleta 2018]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>7.9</td>
<td>7.6</td>
</tr>
<tr>
<td>Alaska</td>
<td>8.6</td>
<td>7.20</td>
</tr>
<tr>
<td>Aleutian Islands</td>
<td>8.0</td>
<td>7.98</td>
</tr>
<tr>
<td>Cascadia</td>
<td>7.7</td>
<td>7.20</td>
</tr>
<tr>
<td>Northern Central America &amp; Mexico</td>
<td>7.4(^1)</td>
<td>7.40</td>
</tr>
<tr>
<td>Southern Central America &amp; Mexico</td>
<td>7.4(^1)</td>
<td>7.60</td>
</tr>
<tr>
<td>Japan – Kuril-Kamchatka Trench (Pacific Plate)</td>
<td>8.5</td>
<td>7.65</td>
</tr>
<tr>
<td>Japan – Nankai-Ryukyu Trench (Philippine Sea Plate)</td>
<td>7.7</td>
<td>7.55</td>
</tr>
<tr>
<td>Northern South America</td>
<td>8.5</td>
<td>7.30</td>
</tr>
<tr>
<td>Southern South America</td>
<td>8.6</td>
<td>7.25</td>
</tr>
<tr>
<td>Taiwan</td>
<td>7.1</td>
<td>7.70</td>
</tr>
</tbody>
</table>

\(^1\) For central America and Mexico, the interface $m_c$ is not varied for northern and southern regions, but instead it is taken as the average for the whole margin.
4.3.3 Source-Depth-Scaling

At any stage of the model development process, it is possible to compute residuals, which can be used to examine model performance relative to predictor variables. Total residuals between an observed IM for event $i$ and site $j$ ($Y_{ij}$) are computed as:

$$ R_{ij} = \ln(Y_{ij}) - \mu_{\lnY}(M_i, R_{rup,i,j}, V_{530,j}) $$  \hspace{1cm} (4.14)

where $\mu_{\lnY}$ is the natural log mean from the GMM at a particular step of model development. Total residuals $R_{ij}$ can be partitioned into mean bias term ($c$), between-event residuals ($\eta_{E,i}$), and within-event residual ($\delta W_{ij}$) using mixed effects analysis [R Core Team 2019; Bates et al. 2015] as follows:

$$ R_{ij} = c + \eta_{E,i} + \delta W_{ij} $$  \hspace{1cm} (4.15)
Event term is indicated here using typical notation ($\eta_E$) because it is evaluated relative to a model that has both magnitude- and depth-terms rather than $\eta_E'$, which is computed relative to a model without a source term; see Section 4.3.1.

We developed the depth-scaling model based on event terms computed using site-adjusted data and the source and path models described in previous sections (i.e., $\mu_{inY} = c_0 + F_p + F_M + F_s^{atr}$). Those event terms were examined for trends with earthquake-source depth. Two measures of depth were considered: hypocentral depth ($Z_{hyp}$) and depth to top of rupture ($Z_{tor}$). There are two general considerations in selecting an appropriate depth metric: (1) predictive power and (2) convenience for forward application.

From a technical point of view, we consider $Z_{hyp}$ to be preferred to $Z_{tor}$ for two reasons: (i) it is a more fundamental parameter that may be related to earthquake stress drop, especially for subduction zone earthquakes that tend to be deeper than earthquakes in other tectonic regimes [Bilek and Lay 1998; 1999]; and (ii) we expect less uncertainty in estimates of $Z_{hyp}$ than of $Z_{tor}$ because the majority of events in the NGA-Subduction database do not have published finite-fault models available and thus have estimates of $Z_{tor}$ from randomized simulations using existing empirical fault geometry relationships (i.e., Contreras et al. [2020]). Regarding the first reason, we recognize that $Z_{tor}$ was used in GMMs for active tectonic regions (e.g., Bozorgnia et al. [2014]) to capture the depletion of high-frequency radiation of surface-rupturing events compared to buried events (potentially due to a shallow weakened zone in the crust, e.g., Graves and Pitarka [2010]); however, we do not expect intraslab subduction events to be surface rupturing.

On the other hand, $Z_{hyp}$ is cumbersome in application because it requires randomization of source location on the fault, which involves an additional loop in the hazard integral. Where possible, we recommend considering hypocentral location in an additional hazard analysis integral as it represents realistic variability in earthquake source processes. However, it is possible to replace the event-specific hypocentral depth with the mean depth expected for a given fault plane, $\bar{Z}_{hyp}$. The value of $\bar{Z}_{hyp}$ is equally as convenient for hazard analyses as $Z_{tor}$ because they are both determined once a fault rupture plane is defined (which is already part of hazard analysis for distance calculation). Hence, no additional hazard integral loops are required in forward analysis. For the present analysis, we use $Z_{hyp}$ to formulate the model functional form as we believe it has more predictive power, but we subsequently develop relations for $\bar{Z}_{hyp}$ for forward use in applications. The estimated value of $\bar{Z}_{hyp}$ can be used in Equation (4.6) without further modification to either the model coefficients or the between-event variability.

When examining trends in the event terms with source-depth metrics, we observed no statistically significant trends for interface events; see Figure 4.7. Therefore, the tri-linear model term given in Equation (4.6) is applicable only to intraslab events. We initially fit Equation (4.6) to the event terms using a nonlinear least-squares regression with all parameters free. Based on these results, a single corner depth, $d_b$, was chosen for all periods: 67 km. Then the regression was repeated with the corner depth constrained, and the slope $m$ and coefficient $d$ were fit. This process was iterated again by constraining $m$ and smoothing $d$. The model slope $m$ goes to zero at the lower end of the depth range populated with data, which is 20 km. The model goes to zero at 2.0 sec, as the increase in ground-motion amplitudes due to increased stress drop is only observed at short periods. Figure 4.7 shows the event terms as a function of $Z_{hyp}$ for PGA, 0.2 sec, and 1.0 sec for both event types, and the model term $F_D$ for intraslab events.
Figure 4.7  Variation of event terms as a function of hypocentral depth at PGA, 0.2-sec PSA, and 1.0-sec PSA. Interface events are shown on the left and intraslab events are shown on the right. Binned means with standard errors and best-fit depth scaling model [Equation (4.6)] shown for intraslab events.

The trend in Figure 4.7 can be interpreted as a consequence of the stress drop increasing with increasing depth in the equivalent point source framework, which has been observed previously for shallow events in active tectonic regions [Yenier and Atkinson 2015(b); Hassani and Atkinson 2018(a)], for events in stable continental regions such as CENA [Yenier and
Atkinson 2015(a)], and for induced earthquakes in Oklahoma [Novakovic et al. 2018]. Atkinson and Boore [2003] also has a linear source-depth scaling term in their GMM for subduction zones.

Mai et al. [2005] used a database of 80 finite-fault models [Mai 2004] to examine hypocenter positions within fault planes. They computed the down-dip depth of hypocenters ($Z_{dd}$) for each earthquake, compare the mean of $Z_{dd}$ normalized by fault width ($W$) for strike–slip and dip–slip events, and looked for any trends with magnitude. We used the finite-fault models for events in the NGA-Sub database to validate the findings of Mai et al. [2005] and to develop a relationship to predict mean values of $Z_{hyp}$ given $Z_{tor}$ and fault geometry, as shown in Figure 4.8(a) and given by,

$$Z_{dd} = \frac{Z_{hyp} - Z_{TOR}}{\sin(dip)} \quad (4.16)$$

$$\bar{Z}_{hyp} = Z_{TOR} + \theta_W W \sin(dip) \quad (4.17)$$

To constrain the average normalized down-dip depth (parameter $\theta_W$), we examined data for 25 intraslab earthquakes (interface events are not considered because our model has no source-depth scaling for interface events). Figure 4.8(b) shows that $\theta_W = Z_{dd}/W$ does not have a trend with magnitude; therefore, we used the average value of 0.48. This indicates that on average the earthquakes are nucleating near the center of their subsequent rupture planes. This compares favorably to the Mai et al. [2005] value of 0.43 for dip–slip events (not only subduction). Using this value of $\theta_W$, the mean hypocentral depth ($\bar{Z}_{hyp}$) computed using fault geometry [Equation (4.17)] can be used in place of $Z_{hyp}$ in Equation (4.6) for forward applications without changes to model coefficients if the true hypocentral depth is unknown.

![Figure 4.8](image_url)

(a) Schematic showing fault plane geometry used to derive relationship between $Z_{tor}$ and $Z_{hyp}$ for forward use in hazard analyses; and (b) normalized down-dip hypocentral depth ($Z_{dd}/W$) for intraslab events in the NGA-Subduction database, color-coded by region.
4.3.4 Regional and Global Constant Calibration

The last step in model development for the reference-rock GMM was the determination of the global and regional model constants, \( c_0 \) [Equation (4.1)], and final event terms, \( \eta_E \), through a mixed-effects residuals analysis. Total residuals were computed using Equation (4.14) with the mean GMM taken as:

\[
\ln \mu_{i,j} = F_p(R_{np,i,j}, M_i) + F_M(M_i) + F_D(Z_{hyp,i}) + F_S(V_{S30})
\]  

(4.18)

Then, total residuals were partitioned into constants, \( c_{0,k} \), for each region \( k \), and event terms \( \eta_{E,i} \), using linear mixed effects in R [(R Core Team, 2019; Bates et al. 2015):

\[
R_{ijk} = c_{0,k} + \eta_{E,i} + \delta W_{ij}
\]

(4.19)

Equation (4.19) is equivalent to Equation (4.15), but with the general bias term \( c \) replaced with region-specific constants \( c_{0,k} \). Where adequate data in regions and/or sub-regions exist, \( c_{0,k} \) is set from data. Otherwise, constraints were applied in setting \( c_{0,k} \) when data are sparse for a region or subregion.

In the case of intraslab events, \( c_{0,k} \) was generally set from data. Through residual analyses, we found that some regions (namely, Alaska and South America) had large sub-regional variations in event terms. These variations corresponded to geography; the earthquakes in the Aleutian Islands on average have larger ground motions than earthquakes in mainland Alaska, and the earthquakes in the southern part of South America (namely, Chile) have larger short-period ground motions than earthquakes in the northern section of the subduction zone (e.g., Ecuador and Colombia). In South America, this could be due to the subduction of different tectonic plates (e.g., the Nazca versus the Caribbean). Because of these regional variations, we allow our intraslab model to have different sub-regional constants for these two regions; see Figure 4.9. In other regions with subregions (Central America and Mexico, Japan, and Taiwan), variations of the constant between subregions were checked and were found to not be required.

In the case of interface events, \( c_{0,k} \) was set from data where possible and with constraints in certain cases. In particular:

- Aleutian Islands, Alaska: Data are relatively abundant for the Aleutian Islands (10 events) but sparse for Alaska (1 event). The Aleutian constant is set empirically, and the Alaska constant is set such the Aleutian median ground motion is matched for \( M < m_c \).

- Central America: Whereas Campbell [2020] defines subregions, the \( m_c \) values are similar, and we do not observe significant sub-regional differences. Accordingly, we adopt a common \( m_c = 7.4 \) value and common constants for both sub-regions. The constant is set empirically.

- Japan Pacific Plate, Japan Philippine Sea Plate: Data are relatively abundant for the Pacific Plate sub-region (13 events) but sparse for the Philippine Sea Plate (1 event). The Pacific constant is set empirically, and the Philippine Sea Plate constant is set such that the Pacific Plate median ground motion is matched for \( M < m_c \).
South America (north and south sub-regions): Data are relatively abundant for the south sub-region (13 events) but sparse for the north (2 events). The south constant is set empirically, and the north constant is set such that the southern South America median ground motion is matched for $M < m_c$.

Taiwan: Campbell [2020] defines subregions to the northeast and southwest of Taiwan that have similar $m_c$ values. However, there are no data for the southwest subregion. We set the constant empirically for the northeast subregion and apply it for the full region.

Cascadia: Due to lack of data, the constant was set such that the global median ground motion is matched for $M > m_c$. The match at larger magnitudes was applied because such events are more hazard-critical than events with $M < m_c$.

The global constant was not set by mixed effects analysis for two reasons: (1) the constant is strongly correlated to $m_c$ and should be set for compatibility with the global $m_c$; and (2) this constant would be an unweighted average of the regional constants, which would give too much weight to regions with small data populations. Instead, we set the global constant to be compatible with the weighted average of regional median reference-rock ($V_{S30} = 760$ m/sec) ground motions at the center of the data in the distance range $\leq 100$ km to avoid effects of regional anelastic attenuation differences. For interface events, this was at $M7.0$ and $R_{rup} = 65$ km; for intraslab this was $M6.0$, $R_{rup} = 75$ km. The weights were taken as the inverse of the constant parameter variances. The resulting global and regional constants are shown as a function of period for interface and intraslab events in Figure 4.9.

As shown in Figure 4.9, constant terms for interface events are lower than those for intraslab events. In general, there is more region-to-region variability in the constants for interface events than for intraslab events. Given the correlation between regional constants and $m_c$, the increased spread of interface constants can be understood to largely reflect the larger range of $m_c$ (about 1.5 magnitude units for interface, 0.8 for intraslab). Intraslab events in Japan, the Aleutian Islands, and the southern section of South America have larger-than-average constants, and the northern section of South America and Alaska have lower-than-average constants. The constants for the two South American sub-regions converge for periods larger than about 0.2 sec. For interface events, Japan and South America have the largest constants and the largest $m_c = 8.5–8.6$. Taiwan and Central America and Mexico have the lowest constants and the lowest $m_c = 7.1–7.4$.

Due to lack of data for interface events, we do not have an empirically derived regional constant for Cascadia. The empirically determined intraslab constant for Cascadia is much lower than other regions; therefore, it is not recommended for use and is not shown in Figure 4.9. Rather, the constant values for Cascadia shown in Figure 4.9 are developed to match the global model prediction at $M > m_c$, accounting for the differences in $m_c$ between the global and Cascadia models. Further recommendations for applications in Cascadia are given in Section 8.2.2.
4.3.5 Site Response

The development of the GMM, from path effects through to the constant term, requires iteration with respect to the site response model. As described previously, we initially used the $V_{S30}$-based site response model for active tectonic regions ($F_{Sattr}$) from SS14 to develop all source- and path-related model coefficients. The resulting model is used in Chapter 5 to develop a site-response model that is specific to subduction regions, including regional effects as appropriate. This site response model includes:

- Global period-dependent $V_{S30}$-scaling specific to subduction zones. This scaling is the same for both source types (interface and intraslab);
- Regional adjustment factors for the global $V_{S30}$-scaling coefficients (applies to Alaska, Cascadia, Japan, South America, and Taiwan);
- Nonlinear site response model that has been adjusted from that in SS14; and
- Models for site amplification in sedimentary basins, specifically developed for applications in Japan and Cascadia.

With the site-response model updated from $F_{Sattr}$ (SS14) to $F_S$ (Chapter 5), the global and regional constant terms were re-computed. The values shown in Figure 4.9 are based on the final site-response model, and match the values tabulated in the electronic supplement.

4.4 MODEL RESIDUALS

Residuals analyses were performed to check model performance with respect to predictor variables. Three types of model residuals were considered (modified from Al Atik et al. [2010]): within-event residuals ($\delta W_{ij}$), the site-to-site component of within-event residuals (also known as site terms, $\eta_S$) and between-event residuals (also known as event terms, $\eta_{E_i}$). Here we focus on
\( \delta W_{ij} \) and \( \eta_E \), using the residuals partitioning in Equation (4.15), to evaluate overall model performance with respect to source and path parameters and constant term \( c \) to evaluate overall model bias. Site terms are considered further in Chapter 5. The event terms and within-event residuals are computed using Equations (4.14–4.15), with the final subduction-specific site-amplification model \( F_s \).

Figure 4.10 shows the overall model bias when region-specific constants and other terms are used. We expect a small but non-zero bias because of the manual adjustments to constant terms to improve model performance; see Section 4.3.4. Bias terms are relatively low, generally \( \pm 0.1 \) ln units.

Event terms are shown as a function of moment magnitude for PGA, 0.2, 1.0, and 5.0 sec for the interface model in Figure 4.11 and the intraslab model in Figure 4.12. The event terms were computed using regional terms where applicable and are color-coded by region in each plot. Binned means are shown for the global dataset in the figures. Not all regional data have a sufficient number of events over a wide enough \( M \) range (> ~ 2 \( M \) units) to judge model effectiveness, for example, interface (Taiwan and Central America and Mexico) and slab (Alaska and Central America and Mexico). For the other regions, the event terms do not appear to trend with magnitude. Similarly, the event terms do not trend appreciably with hypocentral depth for PGA, 0.2, 1.0, and 5.0 sec for the interface and slab models (Figures 4.13 and 4.14, respectively).

Within-event residuals are shown as a function of distance for PGA, 0.2, 1.0, and 5.0 sec for the interface model in Figure 4.15, and for the intraslab model in Figure 4.16. Residuals were computed using regional terms where applicable and are color-coded by region. For both the overall dataset and regional datasets, the trend of residuals with distance are reasonably flat.

![Figure 4.10](image.png)

Figure 4.10  Model bias terms \( c \) for interface and intraslab regional models.
Figure 4.11 Event terms, $\eta_{E,i}$, from interface events as a function of moment magnitude for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars).
Figure 4.12 Event terms, $\eta_{E,i}$, from intraslab events as a function of moment magnitude for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars).
Figure 4.13  Event terms, $\eta_{EI}$, from interface events as a function of hypocentral depth for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars).
Figure 4.14  Event terms, $\eta_{E,i}$, from intraslab events shown as a function of hypocentral depth for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Event terms are color-coded by subduction zone region and plotted with their standard errors (gray bars).
Figure 4.15 Within-event residuals, $\delta W_{ij}$, from interface events for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Residuals are color-coded by subduction zone region.
Figure 4.16 Within-event residuals, $\delta W_{ij}$, from intraslab events for PGA and 0.2-sec, 1.0-sec, and 5.0-sec PSA. Residuals are color-coded by subduction zone region.
5 Site Amplification

5.1 INTRODUCTION

The ground-motion model described in Chapter 4 is one of four models being developed as a part of the NGA-Subduction Project. The model was developed using a combination of empirical data analysis, finite-fault simulations, and geometrical constraints (e.g., Ji and Archuleta [2018] and Campbell [2020]) to predict PGA, PGV, and PSA at oscillator periods between 0.01–10 sec. The model as given in Chapter 4 applies only to the reference-rock condition of $V_{S30} = 760$ m/sec (i.e., the NEHRP B/C boundary condition; see Frankel et al. [1996]). To use the model for other site conditions such as soil or weathered rock, a site-amplification model is necessary, which is the subject of this chapter.

Site-amplification models developed empirically for data-rich areas typically have three components:

- A linear site-amplification term that expresses the effect of the shallow site condition on the ground-motion IM. Typically, this is a $V_{S30}$-scaling term, although in some cases fundamental frequency of a site ($f_0$; Hassani and Atkinson [2016]) or site categories (e.g., Zhao et al. [2006]; [2016a and b]) are used in place of $V_{S30}$;
- A nonlinear term that decreases the amplitude of the ground-motion IM as the strength of shaking increases; and
- Secondary terms beyond $V_{S30}$, which approximately account for resonance effects within soil columns or 3D wave-propagation effects within sedimentary basins. Basin effects impacting ground motion can include wave focusing [Baher and Davis 2003; Stephenson et al. 2000] or body-to-surface wave conversion [Graves 1993; Graves et al. 1998; Kawase 1996; and Pitarka et al. 1998]. Secondary parameters that have been used in combination with $V_{S30}$ include site frequency [Kwak et al. 2017; Hassani and Atkinson 2018(b)] for resonance effects and depth to a firm layer (represented by a high shear-wave velocity isosurface; Day et al. [2008]) for basin effects.

In some cases, equivalent-linear or nonlinear ground response simulations are used to derive site-amplification models (e.g., Darragh et al. [2015] and Harmon et al. [2019]) or the nonlinear components of semi-empirical models (e.g., Kamai et al. [2014] and Seyhan and Stewart [2014]).
In the development of subduction zone GMMs, a common past practice has been to borrow some or all of the above site-amplification model components from active tectonic regions (including simulations performed for active regions). For example, the Atkinson and Macias [2009] GMM for the Cascadia subduction zone uses the Choi and Stewart [2005] nonlinear site-amplification model for active regions. Abrahamson et al. [2016] developed a global $V_{S30}$-scaling term using subduction zone data but used the simulation-based nonlinear site-amplification terms of Walling et al. [2008], which were derived using dynamic soil properties for geologic conditions in California. The Abrahamson et al. [2016] model does not consider basin depth or regionalization. Moreover, in the Pacific Northwest region of the U.S., site factors used in current building code applications are derived using data from active tectonic regions (i.e., the model of Seyhan and Stewart [2014]). Another previous practice has been to use ground response simulations conducted for a particular subduction region to constrain site terms. Zhao et al. [2016(a) and (b)] use site classes defined from site period, and for each site category they apply a site-amplification model derived from one-dimensional equivalent-linear ground response simulations for Japan [Zhao et al. 2015].

This chapter presents an empirical subduction-specific site-amplification model to be paired with the reference-rock conditioned GMM in Chapter 4, developed using the database described in Chapter 2. The large dataset compiled in the NGA-Subduction project [PEER 2020] allows for an empirical global model for $V_{S30}$-scaling to be developed along with regional adjustments. The applicability of the nonlinear term in SS14 to the global subduction data is also investigated, and some modifications are proposed. A model for basin depth effects is provided for Japan and the Cascadia region.

5.2 GLOBAL AND REGIONAL MODELS

5.2.1 Global Site-Amplification Model

The site-amplification model is given as the sum of three terms in natural logarithmic units:

$$F_s = F_{lin} + F_{nl} + F_b$$  \hspace{1cm} (4.7/5.1)

where $F_{lin}$ is the linear site-amplification term conditioned on $V_{S30}$, $F_{nl}$ is the nonlinear site-amplification term, and $F_b$ is the basin-depth term. The functional form for the linear $V_{S30}$-scaling term is given as:

$$F_{lin} = \begin{cases} 
  s_1 \ln \left( \frac{V_{S30}}{V_1} \right) + s_2 \ln \left( \frac{V_1}{V_{ref}} \right) & V_{S30} \leq V_1 \\
  s_2 \ln \left( \frac{V_{S30}}{V_{ref}} \right) & V_1 < V_{S30} \leq V_2 \\
  s_2 \ln \left( \frac{V_2}{V_{ref}} \right) & V_{S30} > V_2 
\end{cases}$$ \hspace{1cm} (4.8/5.2)

$F_{lin}$ is tri-linear in $V_{S30}$ space, but only data from Japan and Taiwan show a break in slope at $V_1 = 270$ m/sec, similar to that observed previously in Japan [Campbell and Bozorgnia 2014] and CENA [Parker et al. 2019; Hassani and Atkinson 2017]. In other words, for most regions and for the global model, $s_1 = s_2$. 
The nonlinear term has the same form as the NGA-West2 Seyhan and Stewart [2014] model (SS14):

\[ F_{nl} = f_1 + f_2 \ln \left( \frac{PGA_r + f_3}{f_3} \right) \]  

\( f_1 \) to \( f_3 \) are model parameters, and \( PGA_r \) represents the peak acceleration expected for the reference site condition of 760 m/sec. The parameter \( f_1 \) represents the level of amplification that is independent of \( PGA_r \), which is accommodated by \( F_{lin} \). As a result, \( f_1 \) is not needed in \( F_{nl} \) and is taken as zero in Equation (5.3). The parameter \( f_3 \) represents a transition level of \( PGA_r \), whereby for \( PGA_r \ll f_3 \), \( F_{nl} \) goes to zero; for \( PGA_r \gg f_3 \), \( F_{nl} \) approaches a constant slope of \( f_2 \) with respect to the log of \( PGA_r/f_3 \). We related \( f_2 \) to \( V_{S30} \) as (modified from Chiou and Youngs [2008]):

\[ f_2 = f_4 \left( \exp \left( f_5(\min(V_{S30}, 760) - 200) \right) - \exp \left( f_5(760 - 200) \right) \right) \]  

The basin-depth term \( (F_b) \) is conditioned on region, differential depth \( \delta Z \), and oscillator period. Region options are Japan, the Seattle basin, general Pacific Northwest basin (PNW-basin; e.g. Tacoma, Everett), and the Pacific Northwest outside of defined basin boundaries (PNW-no basin). Basin boundaries are modified from McPhee et al [2014], Ramachandran et al. [2006], and Lowe et al. [2003], as described in Ahdi et al. [2020]. For all other locations, \( F_b = 0 \). For the four specified regions, the basin depth function is:

\[
F_b = \begin{cases} 
  e_1 & \delta Z_{2.5} \leq \frac{e_1}{e_3} \\
  e_3 \delta Z_{2.5} & \frac{e_1}{e_3} < \delta Z_{2.5} < \frac{e_2}{e_3} \\
  e_2 & \delta Z_{2.5} \geq \frac{e_2}{e_3} 
\end{cases} 
\]  

\( e_1 \) to \( e_3 \) are region- and period-specific model coefficients, \( e_3 \) has units of 1/km, and \( e_1 \) and \( e_2 \) are dimensionless. Differential depth is defined as:

\[ \delta Z_{2.5} = \ln(Z_{2.5}) - \ln(\mu_{Z2.5}(V_{S30})) \]  

\( \mu_{Z2.5}, \) is conditioned on \( V_{S30}, \) and is given in Equation (5.7), modified from Nweke et al. [2018]. The centering model coefficients are distinct for the Pacific Northwest and for Japan, as described in Section 5.4.

\[ \ln(\mu_{Z2.5}) = \ln(10) \times \theta_1 \left[ 1 + \text{erf} \left( \frac{\log_{10}(V_{S30}) - \log_{10}(\nu_{\mu})}{\nu_\sigma \sqrt{2}} \right) \right] + \ln(10) \times \theta_0 \]  

For Cascadia, the coefficients in Equation (5.7) are: \( \theta_0 = 3.94, \theta_1 = -0.42, \nu_{\mu} = 200 \text{ m/sec}, \) and \( \nu_\sigma = 0.2. \) For Japan, the coefficients in Equation (5.7) are: \( \theta_0 = 3.05, \theta_1 = -0.8, \nu_{\mu} = 500 \text{ m/sec}, \) and \( \nu_\sigma = 0.33. \)

Coefficients for the global and regional site-amplification models are independent of event-type. These coefficients are given in Tables E1 and E2 in the electronic supplement.
5.2.2 Regional Parameters

Regional slope \( s_1 \) is given for Japan and Taiwan, and regional slope \( s_2 \) is given for Alaska, Cascadia, Japan, South America, and Taiwan in the electronic supplement. For all models, \( V_1 = 270 \text{ m/sec} \). Due to sparsity of data, the global site-amplification model is recommended for Central America and Mexico. There are no regional variations in the nonlinear model, \( F_{nl} \), due to lack of data to constrain coefficients for each region.

Regional basin-depth model parameters \( e_1 \) to \( e_5 \) for Japan and the Pacific Northwest are given in Tables E1 and E2 in the electronic supplement. Regional centering \( V_{330} \)-conditioned depth models (for \( \mu_{Z2.5} \)) are described in Section 5.4.

5.3 MODEL DEVELOPMENT AND RESULTS

5.3.1 Linear Site Amplification

The first step in computing the linear \( V_{330} \)-scaling implied by the NGA-Subduction data is computing within-event rock residuals using the reference rock conditioned GMM, \( \mu_{r,ij} \), given by Equation (4.1), the event terms \( \eta_{E,i} \) given by Equation (4.15), and a nonlinear model, \( F_{nl} \):

\[
\delta W^r_{ij} = \ln(Y_{ij}) - \left[ \mu_{r,ij} + F_{nl,ij} + \eta_{E,i} \right]
\]  

(5.8)

Subscripts \( i \) and \( j \) refer to event and station, respectively. Superscript \( r \) indicates the term is for the reference-rock velocity condition of 760 m/sec. For this step of the model development process, data \((Y_{ij})\) from both interface and intraslab events are combined as we do not expect differences in the source to affect amplification due to site properties. This expectation is tested subsequently using residuals analyses.

Within-event rock residuals \( \delta W^r \) are not expected to average to zero because they represent the difference between data for soil site conditions and model predictions for a reference-rock condition. As such, when taken in aggregate, these residuals provide an estimate of site response per the non-reference site approach [Field and Jacob 1995]. Ideally, the differences between \( \ln(Y_{ij}) \) and the quantity in brackets would be due to site response only, although in reality other factors contribute to non-zero realizations of \( \delta W^r \). The event term is included in the sum within the brackets to remove bias in total residuals that is related to source and hence unrelated to site. There can be biases associated with particular source-to-site paths that are not accounted for in Equation (5.8). An essential element of the non-reference site approach is that the path model should be unbiased in a broad sense, even if it may be biased for a particular realization. If this is the case, then many samples of path errors (over many observations) would average to zero, which in turn would leave site as the remaining source of non-zero mean \( \delta W^r \) values. The \( F_{nl} \) term is included within the brackets in Equation (5.8) to remove nonlinear site effects because the initial focus is on the linear site response. This adjustment is small for most data points, only being appreciable for relatively near-fault (strong shaking) conditions, soft soils, and high-frequency IMs.
The within-event rock residuals are partitioned into reference-rock site terms ($\eta_{S,j}$), which represent the average site amplification observed over many events for each recording station, and the remaining residual ($\varepsilon_{ij}$),

$$\delta W_{ij} = \eta_{S,j} + \varepsilon_{ij}$$ (5.9)

The partitioning is done using mixed-effects analysis in R [R Core Team, 2019; Bates et al. 2015]. The $\varepsilon_{i}$ term represents variation in ground motion due to event-to-event variations in site response and path errors.

The $\eta_{S,j}$ terms are examined for trends with $V_{S30}$, to develop the $V_{S30}$-scaling model ($F_{lin}$). The model development is iterative because of the use of a nonlinear model ($F_{nl}$) in Equation (5.8). The first iteration used an available $F_{nl}$ term in the literature from SS14. Subsequent iterations used a modified $F_{nl}$ term derived in the next section. The results shown here reflect the final outcome once the $F_{nl}$ term was set.

Figure 5.1 shows the variation of $\eta_{S,j}$ with $V_{S30}$ using results from all regions together, along with the model from Equation (5.2). The model fit was performed using nonlinear least-squares regression in R [R Core Team 2019]. Each gray symbol in the figure represents a reference-rock site term for a single site. The scatter of these terms is appreciable. Data trends can be more readily appreciated by examining the variation with $V_{S30}$ of binned means, which are shown along with their standard errors. The results indicate a steady increase in site amplification as $V_{S30}$ decreases and a flattening of the relationship for stiff sites, which is captured in the model by a flat trend for $V_{S30} > V_2$. As found previously in active regions, the trend is significant for all IMs considered, but is strongest for periods of 0.5 to 5.0 sec (along with PGV), and weakest at shorter and longer periods. The wavelengths associated with these intermediate periods are much longer than 30 m, so the strength of this trend is expected to be a result of correlation between $V_{S30}$ and the average velocity structure at greater depths (such correlations have been documented for a combined dataset from Japan, California, Turkey, and Europe; see Boore et al. [2011]).

Figure 5.2 compares slope parameter $s_2$ as derived for the global model in this study to comparable parameters in the SS14 model for active tectonic regions and the Parker et al. [2019] model for CENA. The $V_{S30}$-scaling in the global subduction model is comparable to that in active regions for PSA oscillator periods up to 0.2 is slightly weaker at periods 0.2–1.0 sec, and significantly weaker (about 0.2 units) at periods longer than 1.0 sec. The global subduction slope is stronger than that in CENA for PSA oscillator periods greater than 0.15 sec. The shape of the $s_2$ parameter in the global subduction model is similar to that for active tectonic regions, with weaker scaling for periods greater than 0.15 sec. In contrast, the $V_{S30}$-scaling for CENA is relatively constant with respect to period.

Once the global model was set, additional plots as in Figures 5.3 and 5.4 were prepared for each region. Figure 5.3 compares the 0.2-sec PSA data for each region to the global model [Equation (5.2)] and a regional model reflecting regional coefficients. Figure 5.4 shows the same information for 1.0-sec PSA. The global value of $V_2$ is used for each period as the sparsity of data when split by region causes $V_2$ to be under-determined. Lastly, $V_1$ and $s_1$ were fit to the Taiwan and Japanese datasets to allow a break in slope at slow $V_{S30}$.
Figure 5.1  Global $V_{S30}$ scaling model, $F_{lin}$, for PGV, PGA, and a range of PSA oscillator periods 0.1–10.0 sec.
Figure 5.2 Comparison of $V_{S30}$-scaling slope between the global NGA-Subduction model, the Seyhan and Stewart [2014] (SS14) slope for active tectonic regions, and the Parker et al. [2019] (Pea19) slope for CENA.

Figure 5.5 shows the variation with period of the global and regional $V_{S30}$-scaling slopes $s_1$ and $s_2$. The regional variations from the global model $s_2$ are modest for most regions (e.g., Japan, Cascadia, and South America) but are large enough to be statistically significant. Alaska shows the largest deviation from the global model, with a steeper $V_{S30}$-scaling slope at most periods. At short periods, South America has slopes shallower than the global average. Only Taiwan and Japan have distinct estimates of $s_1$ due to an observed break in $V_{S30}$-scaling slope at $V_1 = 270$ m/sec. Values of $s_1$ are shown as filled triangles in Figure 5.5 for Japan from 0.075–0.75 sec, and for Taiwan from 0.25–10 sec. At other periods, the values of $s_1$ are equivalent to $s_2$ and therefore not plotted in Figure 5.5. Due to lack of data, we recommend that the global model be used in Central America and Mexico, and no regional slope values are shown.

The plots in Figures 5.1, 5.3, and 5.4 pass through zero at $V_{S30}$ values lower than the desired reference condition of 760 m/sec. This indicates that adjustment of the constant term is needed to shift the residuals such that $F_{lin} = 0$ at 760 m/sec, which was performed iteratively after the determination of site-amplification model coefficients.
Figure 5.3  Regional $V_{S30}$-scaling model, $F_{lin}$, for 0.2-sec PSA shown in dashed line, compared to global model shown in solid line.
Figure 5.4 Regional $V_{S30}$-scaling model, $F_{lin}$, for 1.0-sec PSA shown in dashed line, compared to global model shown in solid line.
Figure 5.5 Global and regional values of $V_{50}-$scaling slopes $s_1$ and $s_2$ [Equation (5.2)]. Regions and periods for which only $s_2$ is shown indicate that the value of $s_1$ is equivalent to that of $s_2$.

5.3.2 Nonlinear Site Amplification

To investigate the nonlinear component of site amplification implied by the NGA-Subduction data, the residuals between the data and a GMM with a linear site term are computed:

$$\delta W_{ij}^{\text{lin}} = \ln(Y_{ij}) - \left[ \mu_{ij}^{\text{lin}} + F_{5,i} + \eta_{E,i} \right]$$

(5.10)

The GMM is exercised for the reference-rock condition [Equation (5.1)]; event terms are as given by Equation (4.15). The linear site term is as given in Equation (5.2). The mean of within-event residuals given by Equation (5.10) should be zero if site response is linear. As a result, we look for conditions where the mean trend departs from zero to identify conditions giving rise to nonlinear site response.

The computed within-event residuals for linear site response ($\delta W_{ij}^{\text{lin}}$) are plotted in Figure 5.6 against the expected median PGA for the reference-rock condition (760 m/sec). This median PGA is computed from the GMM reference-rock mean and PGA event term,

$$PGA_{r,ij} = \exp\left(\mu_{r,ij}^{PGA} + \eta_{E,i}^{PGA} \right)$$

(5.11)

where the PGA superscript indicates that the mean model and event term are taken for the IM of PGA. The $PGA_r$ in Equation (5.11) represents the expected shaking intensity that would have occurred at the site had the site condition been the reference condition. $PGA_r$ affects the extent to
which nonlinear soil behavior is expected [Equation (5.3)]. As in SS14, these plots were made for $V_{S30}$ bins ≤ 200, 200-310, 310-520, 520–760, and ≥ 760 m/sec.

The data in Figure 5.6 typically have a mean across all values of $PGA_r$ of zero but in most cases have a variation with $PGA_r$, which causes under prediction bias ($\delta W_{ij}^{lin} > 0$) for low $PGA_r$ and over prediction bias $\delta W_{ij}^{lin} < 0$ for high $PGA_r$. Such features indicate nonlinearity and are most clearly evident for $V_{S30}$ bins ≤ 200 and 200–310 m/sec. For each of the $V_{S30}$ bins considered in Figure 5.6, the data trend was fit using Equation (5.3), from which discrete values of $f_1$ and $f_2$ were obtained for each period. The value of $f_3$ is period-independent and set to 0.05g based on visual inspection. This is smaller than the SS14 value of $f_3 = 0.1g$. 

![Figure 5.6 Nonlinear site model $F_{nl}$ for PGA, 0.2 sec, 1.0 sec, and 5.0 sec shown as a function of $PGA_r$ for $V_{S30}$ bins. The corresponding model from SS14 is shown for comparison using the same $f_1$ coefficient.](image-url)
The $f_2$ results were plotted as a function of $V_{S30}$ to evaluate the applicable coefficients for the model in Equation (5.4). Coefficients $f_4$ and $f_5$ were fit through this process as were the velocities that appear in Equation (5.4), which have been modified relative to those given in Chiou and Youngs [2008]. The fit of the selected model to the $f_2$ values from individual bins is shown in Figure 5.7, which also shows $f_2$ values from NGA-West2 data and simulations, and the SS14 $F_{nl}$ model. Overall, the nonlinear site amplification inferred from the NGA-Subduction data agrees with what was found in NGA-W2 for $V_{S30}$ bins > 310 m/sec but shows less nonlinearity for the 100–200 m/sec and 200–310 m/sec bins. There is no significant nonlinearity observed in the data for $V_{S30} \geq 760$ m/sec.

**Figure 5.7** Values of parameter $f_2$ estimated using the NGA-Subduction dataset shown with the proposed relation in Equation (5.4), along with the model from SS14, empirical values of $f_2$ from NGA-West2, and simulation-based values of $f_2$ from NGA-West2.
Non-zero values of $f_i$ were obtained for data in velocity bins, as shown in Figure 5.6. These values are not used in model development because they are merely a by-product of the residuals analysis in Equation (5.10), assuming linear site response. When the full, nonlinear site term is used in residuals analysis, any biases are accommodated in constant terms [$c_0$ in Equation (4.1)].

5.4 BASIN AMPLIFICATION

For Japan and the Pacific Northwest region of North America, we have developed period-dependent basin sediment-depth models, $F_b$, that are additive in the site-amplification model, $F_S$ [Equation (5.1)]. The basin-depth model is conditioned on $\delta Z_{2.5}$ [Equation (5.6)], which represents the difference between $Z_{2.5}$ at a site, and the $Z_{2.5}$ predicted by an empirical $V_{S30} - Z_{2.5}$ centering relationship, $\mu_{Z_{2.5}}(V_{S30})$, where $Z_{2.5}$ is the depth to the 2.5 km/sec shear-wave velocity horizon. Sites in Japan have basin depths $Z_{2.5}$ and $Z_1$, which are correlated; see Figure 5.8. As a result, either parameter could be applied. However, we chose to use $Z_{2.5}$ for compatibility with the Cascadia model, as described further below.

The parameterization of differential sediment depth, $\delta Z_{2.5}$, is meant to center the basin-depth term relative to the $V_{S30}$-based site-amplification model components ($F_{in} + F_{nt}$), which implicitly include mean basin effects for a given $V_{S30}$ because they are derived in part from ground motions recorded in basins. As shown in Figure 5.9, we fit Equation (5.7) to the $V_{S30}$ and $Z_{2.5}$ data for stations used in GMM development for Japan and Cascadia to develop the models for $\mu_{Z_{2.5}}(V_{S30})$ given in Section 5.2.1.

![Figure 5.8](image)

**Figure 5.8** Relationship between $Z_{2.5}$ with $Z_1$ for ground-motion recording sites in Japan used in model development. The correlation coefficient between these depth parameters is 0.694.
The functional form for the basin model [Equation (5.5)] was developed to capture trends from residuals of data relative to models consisting of the rock GMM (Chapter 4) combined with the $V_{S30}$-related related components of $F_s$. These residuals were computed as,

$$\delta W_{ij}^{nb} = \ln(Y_{ij}) - \left[ \mu_{ij} + \eta_{E,i} + F_{lin,j} + F_{nl,i,j} \right]$$  \hfill (5.12)

where $nb$ indicates that no basin component ($F_b$) is included in the residuals computation. The $F_{lin}$ and $F_{nl}$ models are as given in Sections 5.3.1 and 5.3.2.

Residuals were computed with Equation (5.12) using data from Japan and Cascadia. Figure 5.10 shows trends of Japan residuals [Equation (5.12)] with differential depth for PGA, PGV, and 5.0-sec PSA. Features of the data evident in these plots include:

- Positive values of $\delta Z_{2.5}^b$ (larger-than-average depths) correlate with positive residuals at long periods ($T \geq 0.5$ sec in Japan and $T \geq 0.75$ sec in Cascadia). This has been observed in prior models for active tectonic regions;

- Positive values of $\delta Z_{2.5}^b$ correlate with negative residuals at short periods (0.075–0.15 sec in Japan and 0.075–0.2 sec in Cascadia). While not incorporated in previous models, this effect has physical meaning, as deeper sites would be expected to have relatively large effects of material damping thereby decreasing short-period ground motions and increasing the magnitude of negative residuals for these conditions; and

- Negative $\delta Z_{2.5}$ correlates with negative residuals at long periods ($T \geq 0.5$ sec in Japan and $T \geq 0.75$ sec in Cascadia), and positive residuals at short periods (0.075–0.2 sec in Japan and 0.075–0.3 sec in Cascadia), which is consistent with aforementioned phenomena in an inverse sense.

The basin model in Equation (5.5) is formulated to capture these features using a piecewise linear function. Model predictions are plotted with the data in Figure 5.10 for Japan.

![Figure 5.9](image)

**Figure 5.9** Mean $Z_{2.5}$ ($\mu Z_{2.5}$) as a function of $V_{S30}$ for (a) Japan and (b) Cascadia, used to center the basin sediment depth scaling model through the use of differential sediment depth $\delta Z_{2.5}$. 

54
In Cascadia, we choose to parameterize basin sediment depth models on $Z_{2.5}$ due to the regional geologic conditions in the Pacific Northwest. The presence of surficial or near-surface glacial till, with high shear-wave velocities around 1.0 km/sec, means that depth parameters should be measured to higher shear-wave velocity isosurfaces to represent the crystalline basement underlying basin fill. As described in Ahdi et al. [2020], $Z_{2.5}$ estimates are derived from regional basin models (e.g., Stephenson [2007] and Stephenson et al. [2017]). Outlines of particular basins (Everett, Portland, Seattle, Tacoma, Tuanatin, Georgia, and North Willamette) are provided in Figure 5.11; these outlines are modified from McPhee et al. [2014], Ramachandran et al. [2006], and Lowe et al. [2003], as described in Ahdi et al. [2020] (Section 5.5.2). Some areas have $Z_{2.5}$ estimates from the Stephenson models but are not within the outline of a basin according to our developed classification scheme. There are also areas outside of the domain of Stephenson models populated with $Z_{2.5} = 0$. These depths are considered not physical, and we consider such locations as non-basin sites for the analyses that follow.

Figure 5.10  Relationship between within event residuals and differential depth for sites in Japan for PGA, PGV, and 5.0-sec PSA [data and model from Equation (5.5)]. Dashed line is reference at 0 residual.
Figure 5.11  Basin-edge outlines used in the definition of regional basin sediment-depth effects for the Pacific Northwest region of the U.S. Basin outlines modified from McPhee et al. [2014], Ramachandran et al. [2006], and Lowe et al. [2003]. Dashed line around Seattle Basin shows updated outline from Wirth et al. [2018(b)], published after the conclusion of NGA-Subduction database development. Figure modified from Ahdi et al. [2020].

Figure 5.12 shows residuals computed without an $F_b$ term [Equation (5.11)] for all sites, sites with sediment depth estimates (most within sedimentary basins as indicated by $(Z_{2.5} > 0)$ and sites outside of sedimentary basins and without sediment depth estimates $(Z_{2.5} = 0)$. The results for all sites (Figure 5.12, red) show that the rock GMM with regional $V_{S30}$-based model is unbiased for all periods. However, that model is biased both for sites with an estimate of sediment depth—generally in mapped basins (Figure 5.12, blue)—and sites without estimates of sediment depth—generally outside of basins (Figure 5.12, black). Figure 5.13 shows trends of residuals as a function of differential depth for PGA, PGV, and 5.0-sec PSA. Figure 5.13(a) and (b) shows these residuals for the overall population and the Seattle basin, respectively. In both cases, trends are evident that are similar to those in Japan for $\xi Z_{2.5} > 0$ (positive residuals at long periods, occasionally negative residuals at short periods). In most cases, no bias is evident in Cascadia for $\xi Z_{2.5} < 0$. The proposed model has one option for sites in the Seattle basin, one option for sites within basin boundaries other than the Seattle basin (as defined in Figure 5.11), and one option for sites that have an estimate of sediment depth but are outside of defined basin boundaries, for which a modification is made to coefficients $e_2$ and $e_3$ through a $\Delta e$ term (Tables E1–E2). The available information is not sufficient to distinguish amplification effects for individual basins other than Seattle due to sparsity of station coverage.
Figure 5.12  Trends of mean within-event residuals in Cascadia calculated without basin term ($F_0$) with period for all sites (red), basin sites or sites with known sediment depths ($Z_{2.5} > 0$; blue), and non-basin sites ($Z_{2.5} = 0$; black). Standard errors of the mean are also shown.
Figure 5.13 Trends of Cascadia within-event residuals calculated without basin term \( (F_b) \) with differential depth for general Cascadia basin sites (left column) and Seattle basin sites (right column) for 0.1 sec (top), 2.0 sec (middle), and 5.0 sec (bottom). Dashed line is a reference at 0 residual.

5.5 MODEL PERFORMANCE

Figure 5.14 shows predictions of response spectra obtained by combining the interface reference-rock GMM from Chapter 4 with the site-amplification model described here. Global parameters are used in the site-amplification model. Figure 5.14(a) applies for an interface event of \( M_8 \) and \( R_{rup} = 30 \) km that produces strong shaking conditions. Figure 5.14(b) applies for an interface \( M_7 \) earthquake at \( R_{rup} = 200 \) km that produces much weaker shaking. In both cases, median spectra are shown for site conditions of \( V_{S30} = 200, 400, 700, 1000 \) m/sec.

The relatively weak shaking condition [Figure 5.14(b)] shows steady increases in spectral ordinates as site conditions become softer, with the strongest changes in the period range of 0.2–
2.0 sec. The strong shaking condition [Figure 5.14(a)] shows a more modest increase in long period spectral ordinates and a complex pattern at short periods that is affected by differing amounts of nonlinearity, such that predictions for the softest site conditions have the lowest ground-motion amplitudes.

An assumption implicit to the model development is that site response is not affected by event type, meaning that the model applies equally to interface and slab events. This is checked by plotting site terms as a function of $V_{S30}$ for both event types (Figure 5.15). The lack of bias and trends demonstrates that the assumption is valid.

![Figure 5.14 Predictions of response spectra computed using the global interface reference rock GMM from Chapter 4 with the global site-amplification model described herein for $V_{S30} = 200, 400, 700, \text{ and } 1000 \text{ m/sec}$: (a) M8 event at $R_{rup} = 30 \text{ km}$, representing a condition producing nonlinear site response; and (b) M7 event at $R_{rup} = 200 \text{ km}$, representing a condition producing linear site response.]

### 5.6 MODEL LIMITATIONS AND RECOMMENDED USE

The seismic site-amplification model presented in this chapter is for use in conjunction with the NGA-Subduction GMM presented in Chapter 4. The model could also be used with other GMMs conditioned at 760 m/sec, but a check should be performed for bias against ground-motion data, which if present, would require adjustment of the GMM constant term.

The site-amplification model, $F_S$, is applicable to PGA, PGV, and PSA between 0.01–10 sec oscillator periods. It should not be used outside of the range of $V_{S30}$ used in model building, 150–2000 m/sec. Regional coefficients are recommended for Alaska, Cascadia, Japan, South America, and Taiwan, and basin models are provided for Japan and individual basin structures in the Pacific Northwest. The global model is recommended in Central America and Mexico. We do not consider basin amplification in several notable basins subjected to subduction-zone ground
motions (Mexico City, Santiago, and Taipei), although such models could be developed in future work. For forward use in regions not included in model development, we recommend using a range of $s_2$ values that captures the range of regional epistemic uncertainty.

The site-response model presented here is ergodic. It will not produce a site-specific amplification factor, even with a measured $V_{S30}$ from a site of interest. Site-specific (or non-ergodic) site response can be evaluated separately using recordings at or near the site of interest, or via ground-response analysis simulations using a measured $V_s$ profile [Stewart et al. 2017].

![Figure 5.15](image)

**Figure 5.15** Site terms for the global GMM presented in Chapter 4 in combination with the site-amplification model presented herein, for 0.2- and 1.0-sec PSA. Data from interface (top) and intraslab events (bottom) are shown separately.
6 Aleatory Variability

The aleatory variability in the model represents the natural variations in earthquake ground motions relative to the median model predictions. For a given set of model input parameters, variations between realized ground motions and the model are possible due to differences in the earthquake source (represented by non-zero event term, $\eta_E$), and variations related both to path and site response (non-zero remaining within-event residual, $\delta W_{ij}$). As shown in Figures 4.11–4.16 and Figure 5.15, event terms and within-event residuals have zero mean and no remaining systematic trends with predictor variables. Each also has an accompanying standard deviation (modified from Al Atik et al. [2010]):

- Standard deviation of event terms $\eta_E$, also known as between-event variability, is denoted as $\tau$; and
- Standard deviation of within-event residuals $\delta W_{ij}$ is denoted as $\phi$.

The total aleatory variability ($\sigma$) is obtained by summing the variances as:

$$\sigma = \sqrt{\tau^2 + \phi^2}$$  \hspace{1cm} (6.1)

The value of $\phi$ can be further partitioned into the standard deviation of site terms, $\eta_S$, referred to as site-to-site variability ($\phi_{S2S}$), and the remaining within-event variability ($\epsilon_{ij}$), the standard deviation of which is referred to as the single-station within-event variability, $\phi_{SS}$.

$$\phi = \sqrt{\phi_{S2S}^2 + \phi_{SS}^2}$$  \hspace{1cm} (6.2)

Contemporary probabilistic seismic hazard analyses (PSHAs) require models for $\tau$, $\phi_{S2S}$, and $\phi_{SS}$. In this chapter we investigate dependencies of these standard deviation terms on a series of independent variables and propose models to capture observed trends.

6.1 MODEL FUNCTIONAL FORM

We found that between-event variability, $\tau$, is not a function of model input parameters, does not vary significantly with event type, but does vary significantly with PSA oscillator period. For within-event variability $\phi$, we provide a direct model that is a function of $R_{rup}$ and $V_{S30}$, and then provide a partitioned version of the model as per Equation (6.2) that depends on the same input parameters.

The direct model for total $\phi$ is given in Equations (6.3–6.5):
\[
\phi = \sqrt{\phi^2(R_{rup}) + \Delta \text{Var}(V_{S30})}
\]  

\[
\phi^2(R_{rup}) = \begin{cases} 
\frac{\phi_1^2}{\ln(R_{rup}/R_1)} R_{rup} \leq R_1 \\
\phi_1^2 R_1 < R_{rup} < R_2 \\
\frac{\phi_2^2}{\ln(R_{rup}/R_1)} R_{rup} \geq R_2
\end{cases}
\]  

\[
\Delta \text{Var}(V_{S30}) = \begin{cases} 
\phi_1^2 \left( \frac{\ln(R_2/\ln(m(R_1,min(R_2,R_{rup})))))}{\ln(R_2/\ln(R_1))} \right) V_{S30} \leq V_1 \\
\phi_1^2 \left( \frac{\ln(V_2/V_{S30})}{\ln(V_2/V_1)} \right) \left( \frac{\ln(R_2/\ln(m(R_1,min(R_2,R_{rup})))))}{\ln(R_2/\ln(R_1))} \right) V_1 < V_{S30} < V_2 \\
0 V_{S30} \geq V_2
\end{cases}
\]

Equation (6.4) describes a flat-ramp-flat trilinear relationship for variance as a function of distance that has period-independent corner distances of $R_1 = 200$ km and $R_2 = 500$ km. Equation (6.5) reduces $\phi^2$ in the initial flat portion ($R_{rup} < R_1$) and the sloped portion (between $R_1$ and $R_2$), but not at larger distances, which reflects the conditions under which site response nonlinearity is most prevalent. The reduction is maximized for slow sites $V_{S30} \leq V_1 = 200$ m/sec, is null for fast sites ($V_{S30} \geq V_2 = 500$ m/sec), and has a linear transition between $V_1$ and $V_2$ controlled by slope $\phi_1^2$. The model as given in Equations (6.2–6.5) is applicable to both interface and intraslab events. Values of model coefficients are given in the electronic supplement to this report.

Lastly, the models for $\phi_{S2S}^2$ and $\phi_{S3S}^2$ are given in Equations (6.6–6.7) and (6.8–6.10), respectively. The $\phi_{S2S}^2$ model is tri-linear as a function of $V_{S30}$ and also scales with $R_{rup}$ for soft sites, and the $\phi_{S3S}^2$ model is tri-linear in both $V_{S30}$ and $R_{rup}$. The model components are formulated such that their sum is similar to the total within-event variance $\phi^2$ [e.g., Equation (6.2)].

The model for $\phi_{S2S}^2$ is given as:

\[
\phi_{S2S}^2(V_{S30}, R_{rup}) = \phi_{S2S,0}^2 + \Delta \text{Var}_{S2S}(V_{S30}, R_{rup}) 
\]  

\[
\Delta \text{Var}_{S2S}(V_{S30}, R_{rup}) = \begin{cases} 
\frac{a_1 \ln \left( \frac{V_3}{V_M} \right)}{\ln \left( \frac{R_4}{\max(R_3, \min(R_4, R_{rup}))} \right)} V_{S30} \leq V_3 \\
\frac{a_1 \ln \left( \frac{V_{S30}}{V_M} \right)}{\ln \left( \frac{R_4}{\max(R_3, \min(R_4, R_{rup}))} \right)} V_3 < V_{S30} < V_M \\
\frac{a_1 \ln \left( \frac{V_{S30}}{V_M} \right)}{\ln \left( \frac{R_4}{\max(R_3, \min(R_4, R_{rup}))} \right)} V_M < V_{S30} < V_4 \\
\frac{a_1 \ln \left( \frac{V_{S30}}{V_M} \right)}{\ln \left( \frac{R_4}{\max(R_3, \min(R_4, R_{rup}))} \right)} V_{S30} \geq V_4
\end{cases}
\]
where $\phi_{SS,0}^2$, $a_1$, and $V_M$ are period-dependent model coefficients, and $V_3 = 200$ m/sec, $V_4 = 800$ m/sec, $R_3 = 200$ km, and $R_4 = 500$ km are period-independent coefficients.

The model for $\phi_{SS}$ is given as:

$$\phi_{SS}(R_{rup}, V_{S30}) = \phi_{SS}^2(R_{rup}) + \Delta Var_{SS}(V_{S30})$$

(6.8)

$$\phi_{SS}^2(R_{rup}) = \begin{cases} \phi_{SS,1}^2 + \phi_{SS,2}^2 \ln \left( \frac{R_4}{R_5} \right) + \phi_{SS,1}^2 & R_{rup} \leq R_5 \\ \phi_{SS,2}^2 & R_{rup} > R_6 \end{cases}$$

(6.9)

$$\Delta Var_{SS}(V_{S30}) = \begin{cases} a_2 \ln \left( \frac{V_3}{V_M} \right) \left( \ln \left( \frac{R_4}{\max(R_5, \min(R_4, R_{rup}))} \right) \right) & V_{S30} \leq V_3 \\ a_2 \ln \left( \frac{V_{S30}}{V_M} \right) \left( \ln \left( \frac{R_4}{\max(R_5, \min(R_4, R_{rup}))} \right) \right) & V_3 < V_{S30} < V_M \\ a_2 \ln \left( \frac{V_{S30}}{V_M} \right) & V_M \leq V_{S30} < V_4 \\ a_2 \ln \left( \frac{V_4}{V_M} \right) & V_{S30} \geq V_4 \end{cases}$$

(6.10)

where $\phi_{SS,1}^2$, $\phi_{SS,2}^2$, $a_2$, and $V_M$ are period-dependent model coefficients, and $R_3 = 200$ km, $R_4 = 500$ km, $R_5 = 500$ km, $R_6 = 800$ km, $V_3 = 200$ m/sec, and $V_4 = 800$ m/sec are period-independent coefficients. All period-dependent aleatory variability model coefficients are given in Table E3 of the electronic supplement to this report.

### 6.2 MODEL DEVELOPMENT

The standard deviation models in Section 6.1 were developed to capture dispersion trends in between- ($\eta_E$) and within-event ($\delta W$) residuals. In this section, we bin residuals by predictor variables to illustrate these features and show model fits to the data.

#### 6.2.1 Between-Event Variability, $\tau$

Prior studies for active tectonic regions [Bozorgnia et al. 2014; Gregor et al. 2014] have shown that between-event variability tends to decrease as magnitude increases. We investigated this for subduction zones by grouping event terms (i.e., between-event residuals) between $M_{4.5}$ and 9.5 in 0.5 magnitude unit bins. Figure 6.1 shows standard deviations and their 95% confidence intervals computed for each bin for PGA and 1.0-sec PSA. Also shown for reference is the value of $\tau$ across all data as established from a mixed-effects regression. Visual inspection of many such plots for these and other periods do not reveal appreciable trends in $\tau$ with magnitude. Moreover,
\( \tau \) does not show appreciable differences across event-types. As a result, our \( \tau \) model is independent of \( M \) and event type. Figure 6.2 shows the period-dependence of computed \( \tau \) values, their 95% confidence intervals, and the smoothed representation recommended for application. The peak in \( \tau \) near 0.1-sec period is observed across our considered subduction zone regions and has also been observed for at small magnitudes in active regions (e.g., see Figure 16 in BSSA14 and Figure 11 in Campbell and Bozorgnia [2014]).

![Figure 6.1](image1.png)

**Figure 6.1** Values of \( \tau \) from event terms binned in 0.5 magnitude units for PGA and 1.0-sec PSA, shown with 95% confidence intervals. Dashed line represents the value of \( \tau \) from a mixed effects analysis over the full magnitude range.

![Figure 6.2](image2.png)

**Figure 6.2** Period-dependence of between-event variability as computed for each period shown with 95% confidence intervals, and as represented in aleatory variability model. Black line represents the smoothed modeled \( \tau \) for forward applications.
### 6.2.2 Within-Event Variability, $\phi$

Using the same binning and inspection approach as used for between-event variability, we look for trends in total within-event variance ($\phi^2$) with magnitude, rupture distance, and $V_{S30}$. At short periods (e.g. PGA), no trend in $\phi^2$ with magnitude is apparent; see left-hand side of Figure 6.3. At 1.0-sec PSA, $\phi^2$ for $M < 7$ and at $M=9$ are approximately equivalent, with an increase for intermediate magnitudes (approximately $M=7-8.75$), but with large uncertainty compared to values at lower magnitudes and at $M=9$; see right-hand side of Figure 6.3. As a result of the lack of a consistent trend with $M$ and the large uncertainties of binned $\phi^2$ values in the range of $M=7-8.75$, we do not include this trend in our model [Equations 6.3–6.5]. The lack of a clear trend of $\phi^2$ with magnitude departs from previous findings for active tectonic regions (e.g., BSSA14).

Within-event dispersion increases for distances beyond 200 km (Figure 6.4) for PGA and other short-period parameters. We anticipate this is caused by complexities related to ground-motion attenuation that are not fully captured by regional terms in the path model, perhaps due to scattering and wave-type conversions at large distances. Similar features have been observed previously for active tectonic regions (e.g., BSSA14).

Within-event dispersion for PGA and other short-period parameters decreases for sites with $V_{S30}$ below 500 m/sec (Figures 6.5). This is thought to be related to site-response nonlinearity, which reduces the dispersion of site response. These effects are not observed at long periods (1.0-sec PSA).

We model trends in $\phi^2$ using a piecewise function for within-event variance conditioned on $R_{rup}$ and $V_{S30}$ [Equations (6.3–6.5)]. First, a minimum value of $\phi^2$ equal to 0.30 is established by computing the variance of within-event residuals for records with both $R_{rup} \leq 200$ km and $V_{S30} \leq 200$ m/sec. This minimum value is relatively period independent. Then values of $\phi_1^2$ and $\phi_2^2$ [Equation (6.4)] are estimated via a weighted least-squares regression using the distance-binned values of variance at each oscillator period, where only data with $V_{S30} \leq 500$ m/sec were considered. Weights are taken as the inverse of the standard error of binned $\phi^2$ values. Lastly, $\phi_0^2$ [Equation (6.5)] is estimated via a least-squares regression on $V_{S30}$-binned variance at each oscillator period, with only data having $R_{rup} \leq 200$ km considered. Figure 6.6 shows the principle features of the total within-event aleatory variability model at 0.2-sec PSA. Figure 6.6 shows the tri-linear variation of $\phi$ for 0.2-sec PSA with distance for $V_{S30} = 760, 400,$ and $200$ m/sec (see left), and the $V_{S30}$-dependence of $\phi$ for $R_{rup} = 1000, 400,$ and $100$ km (see right).
Figure 6.3  Values of within-event variance, $\phi^2$, for PGA and 1.0-sec PSA shown with 95% confidence intervals as a function of earthquake magnitude. Dashed line represents value of $\phi^2$ from a mixed-effects regression over all data.

Figure 6.4  Values of $\phi^2$ for PGA and 1.0-sec PSA shown with 95% confidence intervals for within-event residuals binned by $R_{rup}$. Dashed line represents value of $\phi^2$ from a mixed-effects regression over all data.

Figure 6.5  Values of $\phi^2$ for PGA and 1.0-sec PSA shown with confidence intervals for within-event residuals binned by $V_{S30}$. 
6.2.3 Site-to-Site Variability, $\phi_{S2S}$

We approach the development of the partitioned within-event variability models [Equations 6.6–6.10] in much the same way as the total $\phi$ model. We start by visual inspection of binned site-to-site variance ($\phi_{S2S}$) with earthquake magnitude, rupture distance, and $V_{S30}$. This is done by performing a mixed-effects analysis on the binned datasets to estimate site terms and their variances.

As for the total $\phi^2$ values, we do not observe a significant trend in $\phi_{S2S}^2$ with magnitude; see Figure 6.7. Although we observe a slight trend in $\phi_{S2S}^2$ at large $R_{rup}$ for PGA and other short-period parameters (Figure 6.8), we do not model this dependence. We do not believe there is a physical basis for distance-dependent site-to-site variability and want to avoid non-essential model complexity. The distance trend may be an artifact of path-to-path variability that is mapped into site terms $\eta_{S}$, and hence into their variability ($\phi_{S2S}^2$). We establish the value of $\phi_{S2S,0}^2$ based on the weighted average of binned values between 50–200 km; see Equation (6.6) and solid line in Figure 6.8.

Next, we examined the dependence of $\phi_{S2S}^2$ on $V_{S30}$ by subtracting $\phi_{S2S,0}^2$ from $V_{S30}$-binned values of $\phi_{S2S}^2$, and plotted these differential variances at the median $V_{S30}$ for each bin; see Figure 6.9. The model in Equation (6.7) was fit to pass through zero at the median $V_{S30}$ for the population, (i.e., $V_M$). We see an increase in site-to-site variability for stiff sites relative to $\phi_{S2S,0}^2$ and a decrease for soft sites. This is consistent with our understanding of the effects of soil nonlinearity on site-response variability. We fit this trend with a tri-linear function with respect to $V_{S30}$, as given in Equation (6.7) and shown in Figure 6.9. The value of $\Delta Var_{S2S}$ goes to zero at long periods where nonlinear effects in site response diminish. Because $\Delta Var_{S2S}$ is associated with soil nonlinearity—and thus shaking intensity—we applied its full effect for $R_{rup} < R_3$ and scaled this effect to zero for $R_{rup} > R_4$. 
Figure 6.7 Values of $\phi^2_{S25}$ for PGA (left) and 1.0-sec PSA (right) shown with 95% confidence intervals for within-event residuals binned by M. Dashed line represents value of $\phi^2_{S25}$ from a mixed-effects regression over all data.

Figure 6.8 Values of $\phi^2_{S25}$ for PGA and 1.0-sec PSA shown with 95% confidence intervals for within-event residuals binned by Rrup. Solid line shows value of $\phi^2_{S25,0}$ [Equation (6.6)].

Figure 6.9 Values of $\phi^2_{S25}$ for PGA and 1.0-sec PSA shown with 95% confidence intervals for within-event residuals binned by $V_{S30}$. Solid line shows model for $\Delta Var_{S25}(V_{S30})$ [Equations (6.6–6.7)].
6.2.4 Within-Event Single-Station Variability, $\phi_{SS}$

Within-event single-station variance $\phi_{SS}^2$ is computed from the residuals remaining after fixed source and site effects are removed. Accordingly, it reflects the impact of path-to-path variability and event-to-event variability in site response for a given site. As such, dependencies on both rupture distance and site condition may be anticipated.

We examined trends in binned single-station variance ($\phi_{SS}^2$) with earthquake magnitude, rupture distance, and $V_{S30}$. While there may be an increase in $\phi_{SS}^2$ with magnitude (Figure 6.10), we do not incorporate this feature into the $\phi_{SS}^2$ model for the reasons as given in Section 6.2.2.

Figures 6.11–6.12 show dependencies of $\phi_{SS}^2$ on $R_{rup}$ and $V_{S30}$, respectively. Using data for all velocities, we observe no appreciable distance-dependence in binned values of $\phi_{SS}^2$ up to 500 km, and then a sharp increase occurs. We investigated whether this increase is a result of significant changes of regional contributions to data; in particular, Japanese data makes up 43% of recordings with $R_{rup} < 500$ km and 20% with $R_{rup} > 500$ km. While we cannot exclude the possibility that the increase is at least in part regional, we nonetheless retained this feature in the model. Accordingly, we fit a piecewise linear function to the binned values, with corner distances at $R_5 = 500$ km, below which the variance is equal to $\phi_{SS1}^2$, and $R_6 = 800$ km, above which the variance is equal to $\phi_{SS2}^2$. Figure 6.11 also shows the model [Equation (6.9)] fit to the binned $\phi_{SS}^2$ values. The physical basis for distance-dependent $\phi_{SS}^2$ is path-to-path variability, as described previously for $\phi$.

As performed for the site-to-site variance, we evaluated the $V_{S30}$-dependence of single-station variance by taking differences between $V_{S30}$-binned variances and the global average variance for sites with $R_{rup} < R_5$ ($\Delta Var_{SS}$). A trilinear model [Equation (6.10)] was fit to the results, which is forced to go through zero at the median $V_{S30}$ value ($V_M$) used in determining $\phi_{SS1}^2$ and $\phi_{SS2}^2$. We observed an increase in $\phi_{SS}^2$ for fast $V_{S30}$ and a decrease for slow $V_{S30}$. The tri-linear function has corner velocities of 200 and 800 m/sec. As in the site-to-site variability model, because $\Delta Var_{SS}$ is associated with soil nonlinearity, we applied its full effect for $R_{rup} < R_3$ and scaled this effect to zero for $R_{rup} > R_4$.

![Figure 6.10](image-url) Values of single-station variance $\phi_{SS}^2$ computed using magnitude-binned records for PGA and 1.0-sec PSA shown with 95% confidence intervals. Dashed line represents value of $\phi_{SS}^2$ from a mixed-effects regression over all data.
6.2.5 Comparison of Alternate Within-Event Variance Models

As given in Equation (6.2), the sum of the within-event single-station ($\phi_{SS}^2$) and site-to-site ($\phi_{S2S}^2$) variances is equivalent to the total within-event variance ($\phi^2$). Because we have a direct model for $\phi$ [Section 6.2.2, Equations (6.3–6.5)] and an indirect estimate based on the component models [Sections 6.2.3–6.2.4, Equations (6.6–6.10)], we compare these estimates in Figures 6.13 (0.2-sec PSA) and 6.14 (1.0-sec PSA).

For 0.2-sec PSA (Figure 6.13), the two models provide similar estimates of $\phi$ for close distances ($R_{rup} < 200$ km) and $V_{S30} < 400$ m/sec. For other conditions, the partitioned model shows more $V_{S30}$-dependence (including at large distances and stiff sites) and less distance-dependence than in the direct $\phi$ model. At longer periods (e.g., 1.0-sec PSA, see Figure 6.14) the models are similar for intermediate to large distances and ($R_{rup} > 300$ km) and stiff-soil to rock site conditions ($V_{S30} > 300$ m/sec). At short distances and soft-site conditions ($V_{S30} = 200$ m/sec, $R_{rup} = 100$ km), the partitioned model has larger estimates of variability by about 0.1 natural log unit.
For application purposes, we recommend using the $\phi_{S5}^2$ model in partially non-ergodic seismic hazard analyses in which the site term is site specific, in conjunction with a range of epistemic uncertainty about the site-specific amplification factor (this epistemic uncertainty should be less than $\phi_{S2.5}^2$). For ergodic analyses, we recommend using the total $\phi$ model; see Section 6.2.2 and Equations (6.3–6.5).

**Figure 6.13** Model predictions for within-event standard deviation at 0.2-sec PSA as a function of rupture distance (left) and $V_{S30}$ (right). Black lines show predictions using the total $\phi$ model [Equation (6.3)], and red lines show predictions using the partitioned $\phi$ model [Equation (6.2)].

**Figure 6.14** Model predictions for within-event standard deviation at 1.0-sec PSA as a function of rupture distance (left) and $V_{S30}$ (right). Black lines show predictions using the total $\phi$ model [Equation (6.3)], and red lines show predictions using the partitioned $\phi$ model [Equation (6.2)].
7 Model-to-Model Comparisons

Model-to-model comparisons are performed in this chapter as a form of model verification. Part of the NGA-Subduction Project includes model-to-model comparisons between the five NGA-Subduction models, including this report: Kuehn et al. [2020]; Chiou et al. [2020]; Abrahamson and Gulerce 2020; and Si et al. [2020], and with existing models for subduction zones. The additional models that are considered are Atkinson and Boore [2003], Zhao et al. [2006], Zhao et al. [2016(a) and (b)], Gregor et al. [2006], Atkinson and Macias [2009], and Abrahamson et al. [2016; 2018]. In this chapter we compare the present model to existing models. We do not show comparisons with other NGA-Subduction models, which are provided in Gregor et al. [2020].

We first show regional comparisons in distance-scaling for an $M_9$ interface event with $V_{S30} = 760$ m/sec at 1.0-sec PSA; see Figure 7.1. Most of the regional variations are relatively modest for short distances (< 200 km) and increase at larger distances (> 600 km). At large distance, Central America and Mexico, South America, Alaska, and the Aleutian Islands have the strongest motions due to slower decay with distance than Japan and Taiwan; see Figure 7.1. The global model largely tracks that for Japan (Pacific Plate) but is slightly higher at large distances due to a difference in anelastic attenuation. For $M=9$ and $V_{S30} = 760$, the Cascadia interface model predictions are equivalent to the global model predictions, and thus the Cascadia model is not shown separately in Figure 7.1.

Comparisons in distance-scaling for the global GMM to those of existing subduction-zone GMMs are given in Figures 7.2–7.5 for the same $M_9$ interface scenario at 1.0-sec PSA. Figures 7.6–7.9 show distance-scaling comparisons for an $M_8$ intraslab event at 70 km hypocentral depth, with $V_{S30} = 760$ m/sec, for 1.0-sec PSA.

Figure 7.2 compares the global model to the Atkinson and Boore [2003] GMM for NEHRP Site Classes B and C. The near-source saturation occurs at a much larger distance for Atkinson and Boore [2003], the geometrical spreading is slower, and the anelastic attenuation introduces a similar level of curvature to the path model at large distances (~500 km). Figure 7.3 compares the global model to the Zhao et al. [2006; 2016(a)] models for interface events. For both models, the near-source saturation distance is similar as is the geometrical spreading. The anelastic attenuation is less in the Zhao et al. [2006] model but similar in the global model to Zhao et al. [2016(a)], perhaps due to similarities in the two datasets (i.e., large contributions of data for Japan). However, the overall predicted ground-motion amplitude is slightly less for Zhao et al. [2016(a)] than the global model for distances greater than 40 km. Figure 7.4 compares the distance scaling of the global model to Abrahamson et al. [2016; BC Hydro] and Abrahamson et al. [2018] (updated BCH). In both cases, the models are quite similar except for an overall ground-motion amplitude
that is slightly larger for Abrahamson et al. [2018]. Lastly, Figure 7.5 compares the global model to simulation-based models of Atkinson and Macias [2009] and Gregor et al. [2006]. The Gregor et al. [2006] path model does not agree with the distance-scaling implied by the NGA-Subduction dataset. Relative to our global model, the Atkinson and Macias [2009] model has a similar near-source saturation distance, flatter geometrical spreading, and a similar rate of anelastic attenuation.

Figure 7.1  Comparison of 1.0-sec PSA distance-scaling between global and regional models for interface M9 events and 760 m/sec site condition. Cascadia model not shown as it matches the global model at M9.
Figure 7.2  Comparison of 1.0-sec PSA distance-scaling between global model and Atkinson and Boore [2003] (AB03) model for interface M9 events and 760 m/sec site condition (AB03 shown for NEHRP Site Classes B and C).

Figure 7.3  Comparison of 1.0-sec PSA distance-scaling between global / Japan models and Zhao et al. [2006] and Zhao et al. [2016(a)] models for interface M9 events and 760 m/sec site condition.
Figure 7.4  Comparison of 1.0-sec PSA distance-scaling between global model and Abrahamson et al. [2016] / BC Hydro [2012] and Abrahamson et al. [2018] (updated BCH) models for interface M9 events and 760 m/sec site condition.

Figure 7.5  Comparison of 1.0-sec PSA distance-scaling between global model and simulation-based models from Atkinson and Macias [2009] and Gregor et al. [2006] for interface M9 events and 760 m/sec site condition.
Figure 7.6 shows the same comparison as Figure 7.1, but for an M8 intraslab event. Regional variations between the source and path models are substantially reduced. A factor that may contribute to this reduced intraslab M8 regional variability relative to the large interface M9 variability is that the intraslab scenario falls within the range of observations for many regions, whereas the interface scenario is an extrapolation for all regions except Japan.

Figure 7.7 repeats the comparison from Figure 7.2 but for M8 intraslab events. In this case, the Atkinson and Boore [2003] models have a slightly larger near-source saturation distance than the proposed global model and a similar distance-scaling slope in the intermediate distance range (100–200 km). The anelastic attenuation for the Atkinson and Boore [2003] model exceeds that for the proposed global model but is consistent with that for the Cascadia model. Figure 7.8 compares the global model to the Zhao et al. [2006 and 2016(b)] models for Japan. The Zhao et al. [2006] model is very similar to the proposed global model and Japan models for this scenario, whereas the Zhao et al. [2016(b)] model has significantly slower distance-scaling and almost no anelastic attenuation (i.e., curvature). Lastly, the proposed global and Cascadia intraslab models are compared to the Abrahamson et al. [2016] and Abrahamson et al. [2018] Cascadia models in Figure 7.9. Both of the Abrahamson et al. models predict lower ground-motion amplitudes than the proposed model, have more curvature at long distances (> 300 km), and have comparable slopes at intermediate distances (100–300 km).
Figure 7.7  Comparison of 1.0-sec PSA distance-scaling between global / Cascadia models and Atkinson and Boore [2003] (AB03) Cascadia model for intraslab M8 events and 760 m/sec site condition (AB03 shown for NEHRP Site Classes B and C).

Figure 7.8  Comparison of 1.0-sec PSA distance-scaling between global / Japan models and Zhao et al. [2006] and Zhao et al. [2016(a)] models for intraslab M8 events and 760 m/sec site condition.
Figure 7.9  Comparison of 1.0-sec PSA distance-scaling between global / Cascadia models and Abrahamson et al. [2016]/BC Hydro [2012] and Abrahamson et al. [2018] (updated BCH) models for intraslab M8 events and 760 m/sec site condition.
8 Summary and Future Work

8.1 SUMMARY

This report describes the development of a ground-motion model (GMM) applicable to subduction zone regions world-wide and an accompanying site-amplification model that considered shallow shear-wave velocity ($V_{S30}$) and basin depth. Both the GMM and site-amplification model are formulated with regional adjustment factors that can be used to customize the models for regional conditions. For locations where regional factors are not defined, a global version of the model can be applied, although with larger epistemic uncertainty.

The median GMM for a reference site condition of 760 m/sec is described by Equations (4.1–4.6) with the coefficients in the electronic supplement to this report. The site-amplification model is defined by Equations (4.7–4.13), with coefficients given in the electronic supplement. The uncertainty model is given by Equations (6.1–6.10), with coefficients given in the electronic supplement. Special considerations related to application of the model to the Cascadia region are provided in Section 8.2.2. A coded version of the median and aleatory variability models for forward applications is available from Mazzoni et al. [2020(b)] in Excel, MatLab, R, and Python.

8.2 RECOMMENDATIONS FOR IMPLEMENTATION

8.2.1 Model Limitations

The GMM presented in this report can be used to predict PGA, PGV, and PSA at 24 oscillator periods between 0.01–10.0 sec for interface and intraslab subduction-zone events. The interface model is valid for $M_{4.5–9.5}$, $R_{rup} = 20–1000$ km, $Z_{hyp} \leq 40$ km, and $V_{S30} = 150–2000$ m/sec. The intraslab model is valid over $M_{4.5–8.5}$, $R_{rup} = 35–1000$ km, $Z_{hyp} = 0–200$ km, and $V_{S30} = 150–2000$ m/sec. Both models are applicable only to sites in the forearc region of subduction zones. Future work is planned to evaluate model performance in regional backarc regimes and to create additional anelastic attenuation terms where necessary.

Regional modifications to the global models are provided for Alaska, the Aleutian Islands, Cascadia, Central America and Mexico, Japan–Pacific Plate, Japan–Philippine Sea Plate, South America (including separate consideration for the northern and southern portions of South America), and Taiwan. The regional modifications may apply to the constant term, anelastic attenuation term, corner magnitude, and the $V_{S30}$-scaling and basin-depth components of the site-amplification model. For forward applications to regions not considered during model
development, we recommend that the logic tree include branches that account for epistemic uncertainties in regionalized terms. For a given coefficient, the range considered on alternate branches should—at a minimum—incorporate the global variations of mean estimates; coefficient estimation errors produce additional uncertainties that may also be considered. In the construction of a logic tree, variations applied to constant terms should be linked to those for corner magnitude \( m_c \) due to the correlation of these parameters. An approach that considers these correlations is illustrated below for the example of Cascadia.

### 8.2.2 Recommended Application to Cascadia

Cascadia is a region is of practical importance for hazard applications in the U.S. and Canada, but we lack data to support model development for interface events (Figure 2.1). The situation is somewhat better for intraslab events, where some data are available although mostly limited to small magnitudes. We have developed a model that we consider technically defensible for hazard applications, but it must be recognized that: (1) epistemic uncertainties are larger than in other relatively data-rich regions; and (2) the model development was aimed at hazard-critical large magnitude scenarios, which may lead to the overprediction of data from small-magnitude intraslab events. This section discusses the recommended regional constants and their epistemic uncertainties for hazard applications in Cascadia. In addition, at the end of the section we summarize recommendations for all regionalized terms for Cascadia, including anelastic attenuation, corner magnitudes, and site response.

For both interface and intraslab events, we formulated Cascadia constants so that model predictions match the global model at large magnitudes. We recommend applying epistemic uncertainty to constant terms \( c_0 \), which for a fixed value of corner magnitude \( m_c \), affects ground motions at all magnitudes [Equations (4.1) and (4.5)]. We posit that this epistemic uncertainty should be based on the distribution of global constants; however, because those constants are correlated with corner magnitudes \( m_c \), adjustments are needed so that all of the constants are applicable to the application-region \( m_c \). Consider a particular application region (in this case Cascadia; denoted \( cas \)) and constants derived for the other considered regions (region numbers denoted with index \( k \)). The region \( k \) constants \( (c_{0,k}) \) can be adjusted to equivalent Cascadia constants \( (c_{0,\text{adj}}) \) as,

\[
c_{0,\text{adj}} = c_{0,k} + c_4 \Delta m_{c,k} + c_5 \Delta m_{c,k}^2
\]

where \( \Delta m_{c,k} = m_{c,\text{cas}} - m_{c,k} \), in which \( m_{c,\text{cas}} \) is 7.7 for interface and 7.2 for intraslab events; see Table 4.1. For the interface case, Figure 8.1 shows these adjusted regional constants for PGA, each of which is compatible with the Cascadia-specific \( m_c \) value. The mean Cascadia interface constant shown in Figure 8.1 (red line) essentially matches the global value. A match is not required because the Cascadia constant was not set to match the global model at all magnitudes, but rather for \( M > m_c \); the similarity in this case occurs because the global and Cascadia \( m_c \) values are similar. Each of the adjusted regional constants is shown with its corresponding standard error.

For forward application, we recommend using the Cascadia constant as the mean of a distribution representing its epistemic uncertainty. The weighted standard deviation (weight inversely related to the standard errors of \( c_{0,k} \)) and weighted standard error are shown in Figure 8.1 as dashed and solid lines, respectively. The weighted standard deviation of \( c_{0,k}^{\text{adj}} \) values for
interface PGA is 0.43, and the recommended epistemic uncertainty range, taken as the Cascadia $c_0 \pm$ the weighted standard deviation of $c_{0,k}^{adj}$, is 3.42–4.28. We encourage the use of logic trees that consider this uncertainty, although we reserve the assignment of branch weights to future work; see Section 8.3.3.

For the intraslab case, the limited available data for small magnitude events (Figure 2.1) exhibits weaker ground motions than the global average. Exceptions are the 2001 M6.8 Nisqually and 2010 M6.5 Ferndale events, which have event terms relative to the global model that are only slightly negative at long periods; see Figure 4.12. As a result, the question is whether to set the model constant to accommodate the average of the Cascadia event terms, which is negative. Our recommendation, which was formulated in group discussion with other NGA-Sub modelers, is to not allow this reduction. Instead, as mentioned previously, our Cascadia constant was formulated so that the Cascadia model matches the global model at large magnitudes, rather than being set empirically. This recommended Cascadia constant is shown in Figure 8.2, along with adjusted regional constants computed in the manner described above; see Equation (8.1). The Cascadia constant exceeds the global value because the Cascadia corner magnitude $m_c^{cas} = 7.2$ is lower than the global value (7.6). The weighted standard deviation of $c_{0,k}^{adj}$ values is 0.35. One approach for representing epistemic uncertainty in the constant is to use the Cascadia $c_0$ value $\pm$ the weighted standard deviation of $c_{0,k}^{adj}$ (ranging from 9.13–9.84), as shown in Figure 8.2. Alternatively, a left-skewed distribution could be considered to encompass smaller ground motions compatible with the Cascadia data (not shown in Figure 8.2).

An example of the epistemic uncertainty ranges in PGA for Cascadia introduced by the distribution of $c_{0,k}^{adj}$ values is shown in Figure 8.3. For both event types, the distribution of $c_{0,k}^{adj}$ values represented by the dotted lines in Figure 8.3 reflects the weighted standard deviation. This translates to about a factor of 2.3 in ground motion for interface PGA and about a factor of 2.0 for intraslab PGA.

Figure 8.4 shows epistemic uncertainty ranges in PGA for the global model, computed in a similar manner to the Cascadia cases in Figure 8.3. The Cascadia median is superimposed upon this range with the red dashed line. The results show the compatibility of the Cascadia and global median models over the full magnitude range for interface. For intraslab events, Cascadia and global are compatible at large magnitudes, but Cascadia exceeds the global model at small magnitudes because of Cascadia’s lower $m_c$.

In summary, both the interface and intraslab models have separate constant values formulated such that model predictions match the global model at $M > m_c$. Epistemic uncertainty in those constants should be considered as described in this section. Aside from constant terms, other regional coefficients in our GMM are the anelastic attenuation coefficients, the site-amplification model, and corner magnitudes. We recommend using the Cascadia-specific $V_{S30}$-scaling slope and basin term for both event types. We recommend taking $m_c$ as the value recommended in Campbell [2020] for Cascadia interface events and the value recommended in Ji and Archuleta [2018] for intraslab events; see Table 4.1. We recommend the global anelastic attenuation coefficient for interface events and the Cascadia specific value for intraslab events. All recommended coefficients are provided in the electronic supplement to this report (Tables E1–E2). The coded versions of the models provided by Mazzoni et al. [2020(b)] implement the recommendations provided in this section for the Cascadia median model.
Figure 8.1 Regional variations of mean estimates of interface constants ($c_{0,adj}^k$) adjusted for applicability with Cascadia corner magnitudes for PGA. Adjusted constants shown with their estimation errors. The recommended mean (red dashed line), weighted standard error of the mean (grey fill), and range represented by the mean ±1 weighted standard deviation (black dashed lines) are shown.

Figure 8.2 Regional variations of mean estimates of intraslab constants ($c_{0,adj}^{intr}$) adjusted for applicability with Cascadia corner magnitudes for PGA. Adjusted constants shown with their estimation errors. The recommended mean (red dashed line), weighted standard error of the mean (grey fill), and range represented by the mean ±1 weighted standard deviation (black dashed lines) are shown.
Figure 8.3  Epistemic uncertainty in PGA from regional $c_0$ values shown in magnitude space for two Cascadia scenarios: (a) interface events with $R_{rup} = 50$ km, $V_{S30} = 760$ m/sec, and (b) intraslab events with $R_{rup} = 75$ km, $Z_{hyp} = 50$ km, and $V_{S30} = 760$ m/sec. Solid lines represents median Cascadia models, dotted lines represent ±1 weighted standard deviation of regional constants, corresponding to dotted lines in Figures 8.1 and 8.2.

Figure 8.4  Epistemic uncertainty from regional $c_0$ values about the median global model for PGA taken as ±1 weighted standard deviation (black solid and dashed lines). The median PGA model for Cascadia is shows for comparison with dashed red lines.

8.2.3 Epistemic Uncertainty outside of Cascadia

An approximate procedure for considering epistemic uncertainties in a median GMM is the scaled backbone approach [Atkinson et al. 2014]. In this approach, the median model is scaled up and down in ground-motion space uniformly with respect to all independent variables. To facilitate the application of this approach, we provide epistemic uncertainties on constant terms by region. These uncertainties are represented by epistemic standard deviation $\sigma_e$.

For Cascadia, $\sigma_e$ is computed using the method given in Section 8.2.2 for various oscillator periods. The resulting values are plotted in Figure 8.5 (top frame) and fit with the following function:
\[
\sigma_e(T) = \begin{cases} 
\sigma_{e1} & T < T_1 \\
\sigma_{e1} - (\sigma_{e1} - \sigma_{e2}) \frac{\ln(T/T_1)}{\ln(T_2/T_1)} & T_1 < T < T_2 \\
\sigma_{e2} & T > T_2
\end{cases}
\] (8.2)

Different \(\sigma_e\) values are estimated for interface and intraslab events (Figure 8.5, left and right columns, respectively). The global value of \(\sigma_e\) is taken as the standard deviation of the regional spread of median GMMs in the same parameter range used to define the global constant. This is \(M_{7.0}\) and \(R_{rup} = 65\) km for interface, and \(M_{6.0}\) and \(R_{rup} = 75\) km for intraslab events; see Section 4.3.4 and Figure 8.5.

For all other regions, \(\sigma_e\) was taken as the standard error of the regional constant value, which is relatively small in data rich regions (e.g., Japan; Figure 8.5, bottom) and relatively large in data-sparse regions (e.g., Central America and Mexico; Figure 8.5, middle). Model coefficients for the epistemic uncertainty about the constants are given in Table E4 of the electronic supplement to this report.
8.3 FUTURE WORK

8.3.1 NGA-Subduction Model Validation

The process of developing the NGA-Subduction database had two main phases: an initial phase in which data was gathered from various sources, and a second phase where the collected data were processed, metadata was compiled, and the relevant source, site, and ground-motion parameters
were assembled into a relational database [Mazzoni et al. 2020(a)]. This workflow required that a
cutoff date be applied to the collection of new data, which was in 2016.

A number of significant data sources either were not developed during the NGA-
Subduction project due to limited resources or became available to the project after the cutoff date.
These data, which were not included in the development of the model within this report, provide
opportunities for model validation in parameter ranges that are of practical interest, particularly at
large magnitudes. Sources that could be considered include:

- Data from New Zealand in the NGA-Subduction database (adopted from Van
Houtte et al. [2016]). This dataset was excluded from model development due
to a lack of Class 1/Class 2 classifications for the events (Contreras et al. [2020];
Section 4.6); we did not want to include ground motions from earthquakes that
were potentially aftershock events in our analysis. This dataset includes 2687
recordings from 104 subduction interface and intraslab events with \( M_{4.5-7.8} \);

- Data from subduction zone earthquakes that have occurred since 2016,
including the 2017 Chiapas and Puebla, Mexico earthquakes (e.g., the database
published in Sahakian et al. [2018]), the 2018 Anchorage earthquake; sequence
(e.g., the Rekoske et al. 2019 data release), and the 2020 \( M_{7.8} \) Perryville
earthquake sequence;

- Data from the Hellenic arc in Greece (e.g., data used in Skarlatoudis et al.
[2013]);

- Data from the Calabrian arc in Italy; and

- The Frankel et al. [2018] and Wirth et al. [2018(b)] ground motions for
Cascadia interface events computed from broadband simulations. These
simulations are potentially useful to validate scaling relations (with magnitude,
distance, \( V_{530} \), and basin depth).

8.3.2 Residuals Analyses in Backarc Complexes

The NGA-Subduction GMM presented in Chapter 4 is only applicable to sites in the forearc
regions of subduction zones. Data recorded at sites in backarc regions were excluded to simplify
the path modeling process. Using the completed GMM and site-amplification models, we plan to
undertake a residuals analysis to determine if there are any attenuation differences between forearc
and backarc zones. Stronger attenuation in the backarc has been observed in Japan as part of NGA-
Subduction and in prior work (e.g., Ghofrani and Atkinson [2011], Skarlatoudis and Papazachos
[2012], and Cramer and Jambo [2020]), although we expect this to vary by region. Some other
NGA-Subduction model developers have found an absence of differences between forearc and
backarc attenuation rates for Alaska and Cascadia [Abrahamson, Personal Communication
[2018]].

Where differences are observed, we plan to implement an additional anelastic attenuation
term [Equation (8.1)] that modifies the distance-scaling for the fraction of the path inside the
backarc. In this case, the path model \( F_P \) may look like:

\[
F_P = c_1 \ln R + (b_3 + b_4 M) \ln R + a_0 R_{FA} + a_{BA} R_{BA}
\]  

(8.2)
where the first two terms are as defined in Equations (4.2–4.4), \( a_0 \) is the same coefficient derived in Section 4.3.1, and \( R_{FA} \) and \( R_{BA} \) are defined as the components of the site-to-source distance \( R \) [Equation (4.3)] in the forearc and backarc regions, respectively (\( R = R_{FA} + R_{BA} \)). Coefficient \( a_{BA} \), to be evaluated by regression, characterizes anelastic attenuation in backarc regions and is likely regionally variable.

### 8.3.3 Epistemic Uncertainty

The last component of future work that is needed to support PSHA applications is to develop quantitative epistemic uncertainty recommendations of two kinds:

1. Regional epistemic uncertainty based on the GMM presented herein; and
2. Model-to-model epistemic uncertainty for a particular application region based on the suite of available subduction zone GMMs, including other NGA-Subduction GMMs [Abrahamson and Gulerce [2020]; Kuehn et al. [2020]; Chiou et al. [2020]; and Si et al. [2020]).

The first component is described in some detail in Sections 8.2.1–8.2.3; what is still needed is to carry out the procedures described there for the range of periods of interest and for other regional coefficients: \( a_0 \), and \( s_2 \). Additionally, logic tree branch weights should be considered. The second component will be addressed collectively by all of the NGA-Subduction GMM developers (e.g., Gregor et al. [2020]).
REFERENCES


Ji C., Archuleta R. (2018). Scaling of PGA and PGV deduced from numerical simulations of intraslab earthquakes, Department of Earth Science, University of California, Santa Barbara, CA.


Kuehn N., Bozorgnia Y., Campbell K.W., Gregor N. (2020); Partially non-ergodic ground-motion model for subduction regions using NGA-Subduction database, *PEER Report No. 2020/03*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.


PEER REPORTS

PEER reports are available as a free PDF download from https://peer.berkeley.edu/peer-reports. In addition, printed hard copies of PEER reports can be ordered directly from our printer by following the instructions at https://peer.berkeley.edu/peer-reports. For other related questions about the PEER Report Series, contact the Pacific Earthquake Engineering Research Center, 325 Davis Hall, Mail Code 1792, Berkeley, CA 94720. Tel.: (510) 642-3437; and Email: peer_center@berkeley.edu.

PEER 2020/02  Data Resources for NGA-Subduction Project. Yousef Bozorgnia (PI) and Jonathan P. Stewart (Editor). March 2020.


<table>
<thead>
<tr>
<th>Year</th>
<th>Title of the Document</th>
<th>Authors</th>
<th>Publication Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006/02</td>
<td>Pacific Earthquake Engineering Research Center Highway Demonstration Project.</td>
<td>Anne Kiremidjian, James Moore, Yue Yue Fan, Nesrin Basoz, Ozgur Yazali, and Meredith Williams.</td>
<td>April 2006.</td>
</tr>
</tbody>
</table>


PEER REPORTS: ONE HUNDRED SERIES


PEER 2010/102  Analysis of Cumulative Absolute Velocity (CAV) and JMA Instrumental Seismic Intensity (Ijma) Using the PEER–NGA Strong Motion Database. Kenneth W. Campbell and Yousef Bozorgnia. February 2010.


The Pacific Earthquake Engineering Research Center (PEER) is a multi-institutional research and education center with headquarters at the University of California, Berkeley. Investigators from over 20 universities, several consulting companies, and researchers at various state and federal government agencies contribute to research programs focused on performance-based earthquake engineering.

These research programs aim to identify and reduce the risks from major earthquakes to life safety and to the economy by including research in a wide variety of disciplines including structural and geotechnical engineering, geology/seismology, lifelines, transportation, architecture, economics, risk management, and public policy.

PEER is supported by federal, state, local, and regional agencies, together with industry partners.

PEER reports can be ordered at https://peer.berkeley.edu/peer-reports or by contacting

Pacific Earthquake Engineering Research Center
University of California, Berkeley
325 Davis Hall, Mail Code 1792
Berkeley, CA 94720-1792
Tel: 510-642-3437
Email: peer_center@berkeley.edu

ISSN 1547-0587X