Ground Motion Prediction Equations for Non-Spectral Parameters using the KiK-net Database

Mahdi Bahrampouri

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Civil Engineering

Adrian Rodriguez-Marek (Chair)
Russell Green
Katerina Ziotopoulou

July 14th 2017
Blacksburg, VA

Keywords: Ground motion prediction equation, Arias Intensity, Duration, single station analyses

Ground Motion Prediction Equations for Non-Spectral Parameters using the KiK-net Database

Mahdi Bahrampouri

ABSTRACT

The KiK-net ground motion database is used to develop ground motion prediction equations for Arias Intensity (I_a), 5-95% Significant Duration (Ds_{5-95}), and 5-75% Significant Duration (Ds_{5-75}). Relationships are developed both for shallow crustal earthquakes and subduction zone earthquakes (hypocentral depth less than 45 km). The models developed consider site amplification using V_{S30} and the depth to a layer with V_S=800 m/s (h_{800}). We observe that the site effect for I_a is magnitude dependent. For Ds_{5-95} and Ds_{5-75}, we also observe strong magnitude dependency in distance attenuation. We compare the results with previous GMPEs for Japanese earthquakes and observe that the relationships are similar. The results of this study also allow a comparison between earthquakes in shallow-crustal regions, and subduction regions. This comparison shows that Arias Intensity has similar magnitude and distance scaling between both regions and generally Arias Intensity of shallow crustal motions are higher than subduction motions. On the other hand, the duration of shallow crustal motions are longer than subduction earthquakes except for records with large distance and small magnitude causative earthquakes. Because small shallow crustal events saturate with distance, ground motions with large distances and small magnitudes have shorter duration for shallow crustal events than subduction earthquakes.

Ground Motion Prediction Equations for Non-Spectral Parameters using the KiK-net Database

Mahdi Bahrampouri

General Audience Abstract

This thesis presents the development of new Ground Motion Prediction Equations (GMPEs) for the prediction of the duration and the Arias Intensity of earthquake strong motions. . Arias Intensity is an index for the energy in the ground motion. The GMPEs are based on the Japanese KiK-net database. Based on the causative earthquake source, source to site path, and site properties, GMPEs give estimation of the mean and standard deviation of the parameters. This information is necessary for conducting probabilistic seismic hazard analyses.

The characteristics of the ground motions with the same magnitude and source to site distance vary amongst different tectonic regimes. For this reason, we develop different GMPEs for earthquakes from different tectonic regimes (subduction zone and shallow crustal earthquakes). The primary motivation for this research is that no existing GMPEs for duration are directly applicable to subduction-zone earthquakes. In addition, because the same stations recorded both types of events, we can directly compare the effect of tectonic environment on the selected ground motion parameters. The estimation of mean duration and mean Arias intensity made by this study show while magnitude and distance scaling of Arias Intensity is the same for shallow crustal and subduction earthquakes, the tectonic regime has a significant effect on duration of ground motion.

Acknowledgment

First, I would like to acknowledge the contribution that Professor Adrian Rodriguez-Marek has made to all aspects of this work. He not only guided me through research, but also provided me with the chance to learn a lot from him. What I learned from him varies from Geotechnical Earthquake Engineering to statistics and probability, from the way I can approach a problem to the importance of validating the results, and from writing an email to a colleague to coding in Matlab. His intelligence and his support are what make me optimist about the upcoming years. During research, I have benefited from discussions with Professor Russell Green and Professor Kim Olsen several times. I also want to acknowledge my committee members Professor Adrian Rodriguez-Marek, Professor Russell Green, and Professor Katerina Ziotopoulou for reading thesis and giving comments about this work. This work is built on the database collected and processed by Haitham Dawood. Haitham answered my questions about the database several times and his help is appreciated. The funding of this project was provided by Electricité de France under cooperation contract No. 3000-5910144023.

I was lucky to find fantastic friends in Blacksburg who not only made my life away from my family possible but also fun. Listing all of them is hard but I want to let them know their friendship means a lot to me. The last but surely not the least, I appreciate my family for their support, especially my parents and my oldest brother Mohsen. I would not be here without their support.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2

Chapter 3

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 1
Introduction

1.1 Problem description

Non-spectral ground motion parameters, such as Arias Intensity and duration, are commonly used in engineering and seismological applications. The usefulness of these parameters relies on the ability to predict their values for future earthquakes. Various ground motion prediction equations (GMPEs) for this type of parameters exist for shallow crustal tectonic regions, but not for subduction regions. The primary objective of the research presented in this thesis is to fill this gap by developing GMPEs for non-spectral ground motion parameters using data from the Japanese KiK-net ground motion network. These data have been uniformly processed (Dawood et al. 2016) and contain recorded ground motions from earthquakes occurring in shallow-crustal regions and in the subduction zones located to the east of the Japanese islands. We develop separate equations for each of these two tectonic regimes and compare the resulting models. Because the same stations recorded both types of events, we can directly compare the effect of tectonic environment on the selected ground motion parameters.

The primary motivation for this research is that no existing GMPEs for duration are directly applicable to subduction-zone earthquakes. There is one model applicable for predicting Arias Intensity for subduction earthquakes (Foulser-Piggott and Goda 2015). However, this model considers the difference in tectonic regime by a constant offset. We demonstrate that distance- and magnitude-scaling are also different between the two tectonic regimes. The proposed GMPEs are developed using nested- and crossed-mixed effect regression, which enabled us to consider random effects for event and site terms simultaneously. Although this manner of calculating parameters is more complex than

the conventional approach, which considers only the event terms as the random effect, it results in unbiased GMPEs. We also compare the effects of considering site variability on the resulting GMPEs.

1.2 Organization of thesis

The first chapter of present thesis explains the motivation for developing new GMPEs for Arias Intensity and duration. The second and third chapters are self-contained papers that will be submitted for publication in peer-reviewed journals. Each chapter provides GMPEs for one non-spectral parameter. The two chapters and their authorship are:

Chapter Two - Mahdi Bahrampouri, Adrian Rodriguez-Marek, Russell Green (to be submitted) Ground Motion Prediction Equations for Arias Intensity Using the Kik-Net Database

Chapter Three - Mahdi Bahrampouri, Adrian Rodriguez-Marek, Russell Green (to be submitted) Ground Motion Prediction Equations for Significant Duration Using the Kik-Net Database

The final section of this thesis is a chapter summarizing the conclusions and the engineering significance of this study.

1.3 References

Dawood, H. M., Rodriguez-Marek, A., Bayless, J., Goulet, C., and Thompson, E. (2016). "A Flatfile for the KiK-net Database Processed Using an Automated Protocol." *Earthquake Spectra*, 32(2), 1281-1302.

Foulser-Piggott, R., and Goda, K. (2015). "Ground-Motion Prediction Models for Arias Intensity and Cumulative Absolute Velocity for Japanese Earthquakes Considering Single-Station Sigma and Within-Event Spatial Correlation." *Bulletin of the Seismological Society of America*, 105(4), 1903-1918.

Chapter 2
Prediction Equations for Arias Intensity Using the Kik-Net Database

Abstract

In seismic design, intensity measures are chosen based on how well these parameters correlate with the damage caused by earthquakes and on the ability to predict these intensity measures for a given earthquake scenario. As a an index for the energy content of ground motion, Arias intensity has proved to be efficient in several applications including earthquake-induced slope failure, liquefaction, and damage in structures. In this paper, the KiK-net database is used to present Ground Motion Prediction Equations (GMPEs) for Arias Intensity of shallow crustal and subduction zone earthquakes. The proposed GMPEs are applicable for **M** 4-9. The predictive models incorporates both V_{S30} and depth to a stiff horizon (V_S=800 m/s) in the prediction of site effects. In addition, we observed that the site amplification of Arias Intensity is dependent on the earthquake scenario. We capture this dependence in the proposed functional form. Moreover, the effect of the volcanic belt in Japan on Arias Intensity is also included in the proposed relationships.

2.1 Introduction

In this paper we present ground motion prediction equations (GMPEs) for Arias Intensity of shallow crustal and subduction zone earthquakes developed using the KiK-net database. Arias Intensity (Arias 1970) is an index forthe energy content of ground motions and incorporates both the duration and amplitude of the entire ground motion time history. Arias Intensity has been correlated to liquefaction triggering (Kayen and Mitchell 1997), earthquake induced-landslides (Wilson and Keefer 1985; Harp and Wilson 1995; Lee et al. 2008; Jibson 2007), damage levels in adobe and clay structures (Benito

and Herraiz 1997), and the response of stiff bridge structures to earthquake ground motions (Mackie and Stojadinovic 2002).

Kayen and Mitchell (1997) developed a relationship for predicting Arias Intensity as a function of magnitude and closest distance to the fault rupture plane. This relationship is based on 66 earthquake records from the western United States, primarily California. Travasarou et al. (2003) developed a GMPE for Arias Intensity using a much larger data set (1208 records from 78 earthquakes). This work, however, also uses a database composed exclusively of shallow crustal earthquakes in active tectonic regions. Based on almost the same database used by Travasarou et al. (2003), Foulser-Piggott and Stafford (2012) proposed GMPEs that introduced a non-linear site amplification term incorporating site properties through the Vs_{30} parameter. Foulser-Piggott and Goda (2015) developed a new GMPE for Arias Intensity using motions recorded in Japan. The database used by these authors includes motions for earthquakes from both shallow crustal tectonic and subduction-zone regimes. The resulting GMPEs differentiate the two tectonic regimes by proposing a linear scaling. This GMPE also uses different anelastic attenuation rates for the forearc and backarc regions of Japan. More recently, Bullock et al. (2017) used ground motion records from stations on rock from around the world to develop a GMPE for different tectonic regimes (i.e., interface, intraslab, intraplate, and shallow crustal). The work presented in this paper adds to the existing GMPEs of Arias Intensity for shallow crustal and subduction earthquakes. Since models for both types of earthquakes are developed from a common set of stations, the proposed models allow a comparison of magnitude and distance scaling of Arias Intensity between both types of earthquakes.

In the remainder of this paper, we first explain the database of records used in this study and the protocol for computing Arias Intensity of each record. We then present the functional forms adopted for the GMPE followed by the residual plots as a measure of the performance of the proposed models.

Finally, we present a comparison of the resulting model for shallow crustal and subduction environments with those from other published GMPEs.

2.2 Ground motion data and computation of Arias Intensity

We use a subset of motions recorded by the Kiban-Kyoshin network (KiK-net) in Japan, gathered and processed by Dawood et al. (2016). To compile the database, motions with unknown Vs_{30}, tectonic regime, and sampling frequency are excluded. In addition, we exclude records with **M**<4 and shallow crustal motions with rupture distance greater than 200 km and subduction-zone motions with rupture distances greater than 1000 km because they do not have engineering significance. Our database has 13,966 recorded motions from 976 shallow crustal events and 30,606 recorded motions from 2,332 subduction events. Figure 1 shows the magnitude and distance distribution of ground motions and the geographical distribution of earthquake sources.

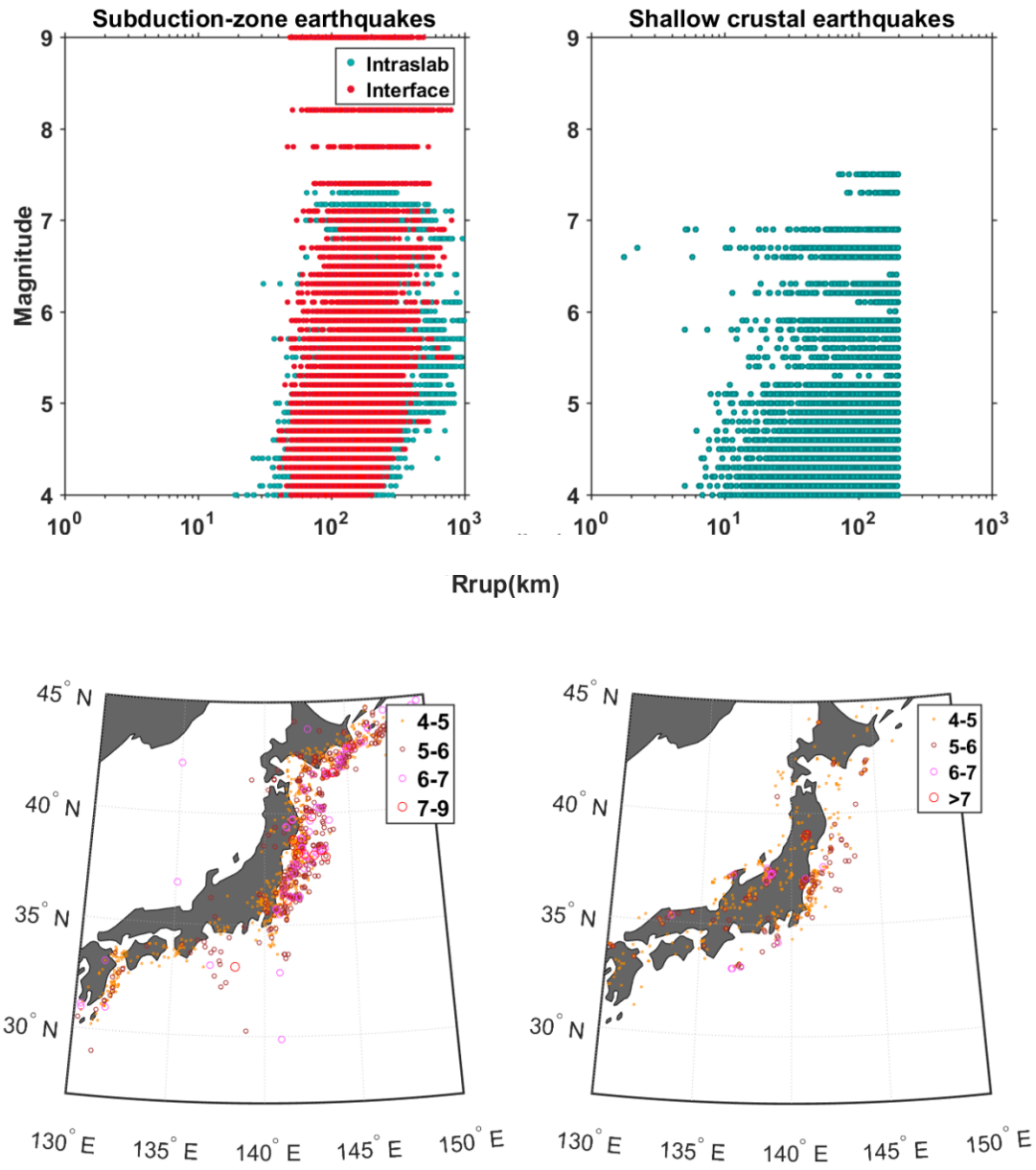


Figure 1. Upper figures show magnitude and distance distributions for each tectonic regime. Lower figures are illustrations of the event source distributions. Plots on the left are for subduction earthquakes, and on the right for shallow crustal earthquakes.

The ground motion time histories processed by Dawood et al. (2016) were used to compute Arias Intensity of the ground motions. Arias Intensity is defined as (Arias 1970):

$$I_a = \frac{\pi}{2g}\int_o^{t_{max}} a(t)^2 dt \qquad \mathbf{1}$$

where I_a is Arias Intensity in units of m/s^2, $a(t)$ is the ground acceleration, g is the acceleration of gravity and t_{max} is the total duration of the recorded time history. While t_{max} is rather arbitrary and a function of the processing protocol, the latter part of a record does not contribute significantly to the integral in Equation 1 and hence does not affect significantly the values of I_a. The definition of I_a is index of the energy absorbed during an earthquake by a set of undamped SDOFs with natural frequencies uniformly distributed between 0 and infinity (Trifunac and Brady 1975). There are several approaches to combine two horizontal components of Arias Intensity, including using a random component, the maximum component, the geometric mean, or the arithmetic mean from the as-recorded components, and the vector sum of the two as-recorded components. Models developed for geometric or arithmetic mean of components have lower variability than models developed for the random component. We use the arithmetic mean because this value is independent of the directions along which horizontal motions are recorded.

2.3 GMPE development

2.3.1 Regression Approach

The regression analyses have to account for an uneven distribution of recordings per event and recordings per station. The random effects algorithm of Abrahamson and Youngs (1992) was proposed to account systematically for the particular effects of each earthquake. This is captured by introducing a random effect variable for each earthquake (i.e., the *event term*). The event terms capture the systematic deviations from the median behavior for each earthquake. When the ground motion database used in the regression also includes multiple recordings at every station, the systematic deviation from the median behavior for each station can also be captured via the introduction of a *site term*. Examples of past studies that accounted for systematic deviations in the site term include Rodriguez-Marek et al. (2013), Rodriguez-Marek et al. (2011), and Lin et al. (2011). These studies

used a multiple-stage regression, where an initial regression captures the effects of event terms and a secondary analysis of residual is performed to compute the site terms. Stafford (2014) pointed out that the two-stage approach has some shortcomings and proposed using crossed and nested mixed-effect regression, such that multiple random effects (e.g., site and event terms) are calculated at the same time as the coefficients of the median ground motion equation are computed. In this study, we use the approach suggested by Stafford (2014). For terminology and notation, we adopt the one proposed by Al Atik et al. (2010):

$$\ln y_{eS} = \ln \widehat{y_{eS}} + \delta S2S_S + \delta B_e + \delta W_{es} \quad \textbf{2}$$

$$\ln \widehat{y_{eS}} = f(\boldsymbol{x},\boldsymbol{\theta}), \quad \textbf{3}$$

where y_{eS} is the observed value of intensity measure for earthquake e and station s; $\widehat{y_{eS}}$ is the median predicted value of intensity measure y_{eS}, which in turn is a function of the explanatory variables $\boldsymbol{x}$ (e.g. magnitude, distance) and fixed effect parameters $\boldsymbol{\theta}$ (model coefficients); δB_e is the event term calculated for the event e; and $\delta S2S_s$ is the site term calculated for station s. Both δB_e and $\delta S2S_S$ are zero-mean random variables with standard deviation τ and ϕ_{S2S}, respectively. Finally, δW_{es} is a zero-mean random variable that represents the residual variability and has a standard deviation ϕ_{SS}. Assuming that these three variables are statistically independent, the total standard deviation, σ, is given by:

$$\sigma = \sqrt{\phi_{SS}^2 + \tau^2 + \phi_{S2S}^2} \quad 4$$

Note that ϕ_{SS} corresponds to the standard deviation of the residuals that have been corrected by event- and site-terms, and is also called the single-station standard deviation; τ is the between-event or inter-event standard deviation; and ϕ_{S2S} is the site-to-site standard deviation.

The regression analyses are conducted with the lme4 package in R (Bates et al. 2006). The functional forms (e.g., the equation for $ln\,\widehat{y_{eS}}$) is a key step in the regression analyses. We use equations from the literature as a starting point and then evaluate the need for modifications in the functional form based on an analysis of residuals. Effects due to the source, path, and the site effect are investigated separately. The regression model adopted for this study is described below.

2.3.2 GMPE functional form

As is the case for most ground motion parameters, the Arias Intensity of a ground motion is affected by the characteristics of the source, the path, and the site. Stafford et al. (2009) derived a theoretical predictive equation for Arias Intensity based on Brune's point source model. By making some simplifying assumptions, such as constant stress drop, Stafford et al. (2009) showed that the logarithm of Arias Intensity correlates linearly with magnitude and the logarithm of distance. Stafford et al. (2009) also indicated that using a theoretical model with linear scaling of Arias Intensity with magnitude may not hold true for large earthquakes. Also, Travasarou et al. (2003) found a decrease in stress drop for large magnitude earthquakes which results in saturation of Arias Intensity. We use these observations as the foundation of our functional form for Arias Intensity, but we modify it based on empirical observations.

Figure 2 plots an estimate of the source term of Arias Intensity obtained by removing a preliminary estimate of path effects from the measured Arias Intensity. These data show a linear trend with magnitude with a break in linear scaling at large magnitudes. Observe that for subduction events (Figure 2a), a hinge point where the magnitude scaling levels out occurs at **M**=7. This hinge point is

adopted for our functional form. Figure 2b shows the path-corrected Arias Intensity values for crustal earthquakes. Observe that the hinge point for shallow crustal occurs at **M**=6 and is included in the functional form.

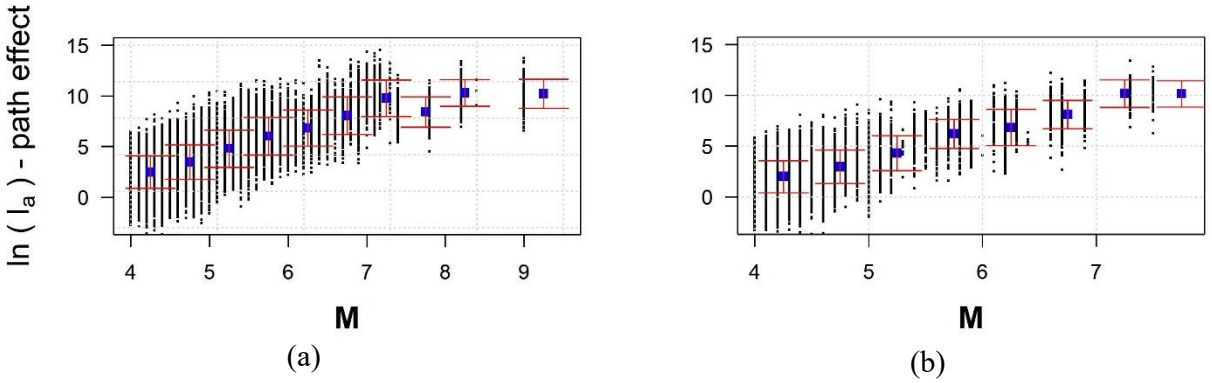


Figure 2 Scatter plot of Arias Intensity corrected for path effects plotted versus magnitude for (a) subduction earthquakes, and (b) shallow crustal earthquakes. The blue dots represent the average of each bin and the red error bars show ± 1 standard deviation.

Foulser-Piggott and Goda (2015), in their regression model for Arias Intensity, included different anelastic attenuation rates for different regions of Japan. They divided the recorded motions of Japan into backarc and forearc motions based on the location of the station relative to the volcanic belt and considered different coefficients for the attenuation rate for these two groups of stations. However, the location of the site alone cannot capture the path between source and site. Dawood and Rodriguez-Marek (2013), in a regression model for spectral accelerations, showed that the attenuation rates are different for regions along the volcanic belt in Japan than for regions in the forearc and backarc. Moreover, paths entirely within the forearc or backarc regions have similar attenuation rates. A similar observation was made by McVerry et al. (2006) for attenuation of spectral acceleration in New Zealand.

In this study, we adopt a functional form that considers an additional attenuation term for source-station paths that cross the volcanic belt. This is achieved by adding a flag (V) that takes a value equal to one if the path crosses the volcanic belt, and zero otherwise. Therefore, the coefficient of V represents the

energy lost due to crossing the volcanic belt. This functional form for path effects of Arias Intensity performs better statistically than a model that considers forearc or backarc location alone. In addition, we deviate from a linear model for path effects. The measured Arias Intensity values, corrected by a preliminary model for source effects, are shown in Figure 3. Observe that the scaling with distance deviates from linearity, possibly due to finite fault effects. Many different functional forms were tested to capture the trends shown in Figure 3. For subduction zone earthquakes, a saturation parameter captures the decrease in slope for close distances. For shallow crustal earthquakes, a hinge point in the path effect function models the difference in slope for distances greater than 100 km (see Equations 12 and 13 below).

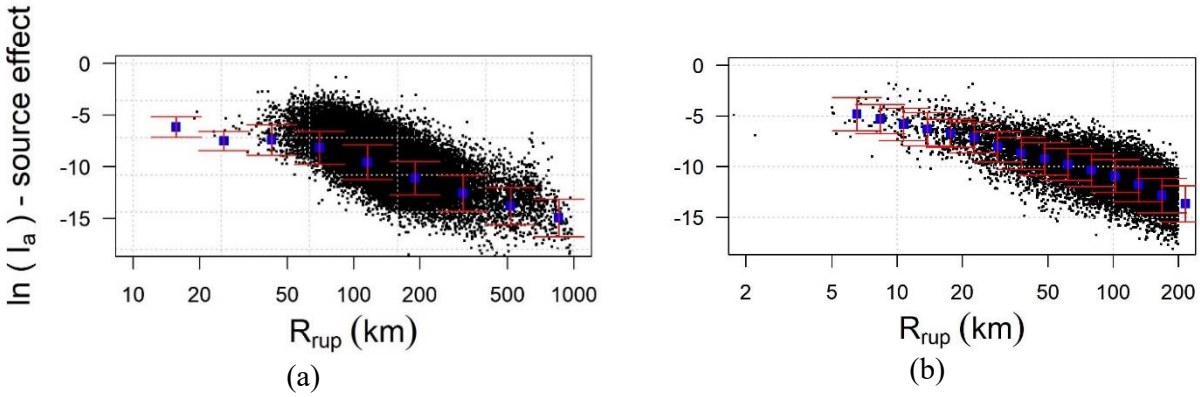


Figure 3 Scatter plot of Arias Intensity minus the effect of source, plotted versus distance for (a) subduction earthquakes and (b) shallow crustal earthquakes. The blue dots represent the average of each bin and the red arrows show ± 1 standard deviation

A site amplification function is proposed, which incorporates both the average site velocity over the upper 30 meters (V_{S30}), depth to shear wave velocity 800 m/s (h_{800}), and magnitude. Figure 4 shows the logarithm of Arias Intensity minus source and path effects $[I_a - f(M) - f(R)]$ plotted versus V_{S30}. Observe that the trends are generally linear (in log-log space), with decreasing Arias Intensity with increasing V_{S30}. However, for small magnitude earthquakes (Figures 5a and 5c), the trend deviates from linearity for low values of V_{S30}.

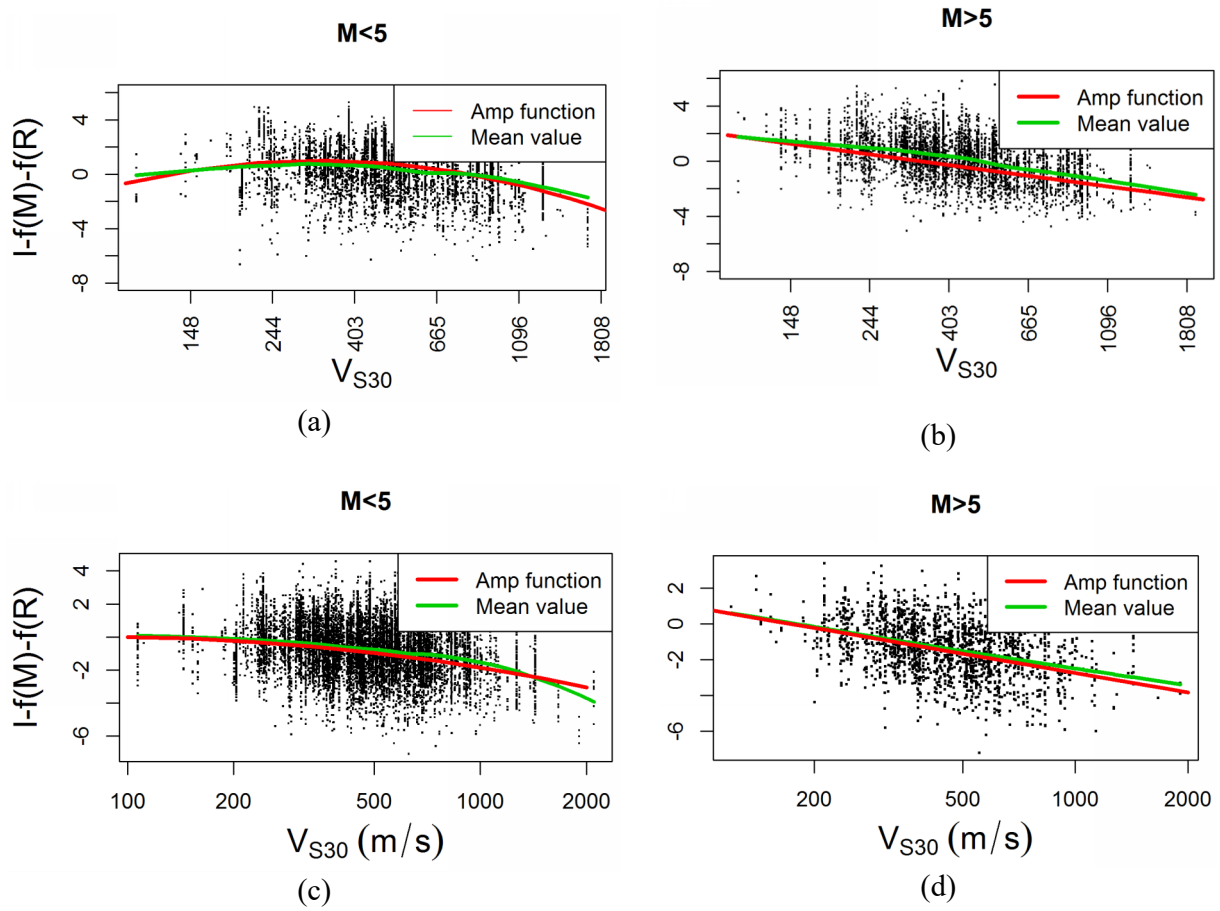


Figure 4. Logarithm of Arias Intensity minus the effect of source and path versus. Upper figures (a and b) show subduction records and lower figures (c and d) show shallow crustal records. The green line shows the mean value of observed points and the red line shows the predicted amplification function.

The observed trends for low magnitude earthquakes can be explained by considering the frequency content of ground motions. The amplification of I_a can be expressed using Parseval's law as:

$$Amp = \frac{\int a_s(t)^2 dt}{\int a_r(t)^2 dt} = \frac{\int a_s(\omega)^2 dw}{\int a_r(\omega)^2 dw} = \frac{\int TF(\omega)^2 a_r(\omega)^2 d\omega}{\int a_r(\omega)^2 d\omega} \quad \mathbf{5}$$

where, a_s and a_r denote acceleration of ground motion in soil and rock, respectively, the arguments t and w indicate time and frequency domains, respectively, and $TF(\omega)$ is the transfer function for site

amplification (i.e., $TF(\omega) = a_s(\omega)/a_r(\omega)$). Equation 5 indicates that the amplification of Arias Intensity is the average amplification across all frequencies weighted by the Fourier amplitude of rock motions. Since the source frequency of low-magnitude earthquakes contains less low-frequency waves, the result is that amplification reduces as magnitude decreases, in particular for soft soils where resonant frequencies are lower. This scaling occurs only at lower magnitudes, because for larger magnitudes the corner frequency is low enough that it encompasses the full range of frequencies where the TF for soft soils is high. This trend is captured with a V_{S30} scaling term. The red lines in Figure 4 are the prediction of the site amplification, which capture well the observed trends.

The effect of depth to a shear wave velocity 800 m/s (h_{800}) is included in the site effects term. Since V_{S30} and h_{800} are correlated, we use a functional form that includes only the difference between the observed h_{800} at a site minus the h_{800} predicted using V_{S30} (this is labelled δh_{800}). The predicted h_{800} is given by:

$$h_{800,p} = exp(-5.23/2 * ln(\frac{V_{S30}^2 + 412^2}{1360^2 + 412^2}) - 0.9) \quad \textbf{6}$$

where $h_{800,p}$ is the h_{800} (in meters) predicted value based on the site's V_{S30} . The difference between h_{800} observed at a site minus the h_{800} predicted based on V_{S30} is:

$$\delta h_{800} = h_{800} - h_{800,p} \quad \textbf{7}$$

Finally, the depth to the top of rupture has a significant effect on Arias Intensity and is included in the functional form. In addition, we observe earthquakes with Z_{tor} more than 200 km are different from other earthquake. This difference is captured buy a dummy variable (i.e. D_{200}). Conversely, the residual analysis shows that the Focal rupture mechanism is not significant enough to be included in

the regression. With these considerations, the proposed functional form for predicting Arias Intensity is given by the following set of equations:

$$\ln\left(I_a\right) = F_{source}\left(M_w\right) + F_{path}\left(R_{rup},V\right) + F_{site}(\boldsymbol{M},V_{s30}) \quad \textbf{8}$$

$$\text{Subduction:} \quad F_{Source} = \begin{cases} a_1 + a_2 * \mathbf{M} & M_w < 7 \\ a_1 + a_2 * 7 + a_3 * (\boldsymbol{M} - 7) & M_w \geq 7 \end{cases} + a_4 * \ln\left(\frac{Z_{tor} + 0.1}{a_5}\right) + a_6 * D_{200} \quad \textbf{9}$$

$$deep = \begin{cases} 0 & Z_{tor} < 200 \\ 1 & Z_{tor} \geq 200 \end{cases} \quad 10$$

$$\text{Shallow crustal: } F_{Source} = \begin{cases} a_1 + a_2 M_w & M_w < 6 \\ a_1 + a_2.6 + a_3(\boldsymbol{M} - 6) & M_w \geq 6 \end{cases} + a_4 * \ln\left(\frac{Z_{tor} + 0.1}{a_5}\right) \quad \textbf{11}$$

$$\text{Subduction: } F_{path} = b_1 V + b_2 \log\left(R + b_3\right) \quad \textbf{12}$$

$$\text{Shallow crustal: } F_{path} = b_1 V + \begin{cases} b_2 \ln\left(R\right) & R < 100 \\ b_3 \ln\left(\frac{R}{100}\right) + b_2 \ln\left(100\right) & R \geq 100 \end{cases} \quad \textbf{13}$$

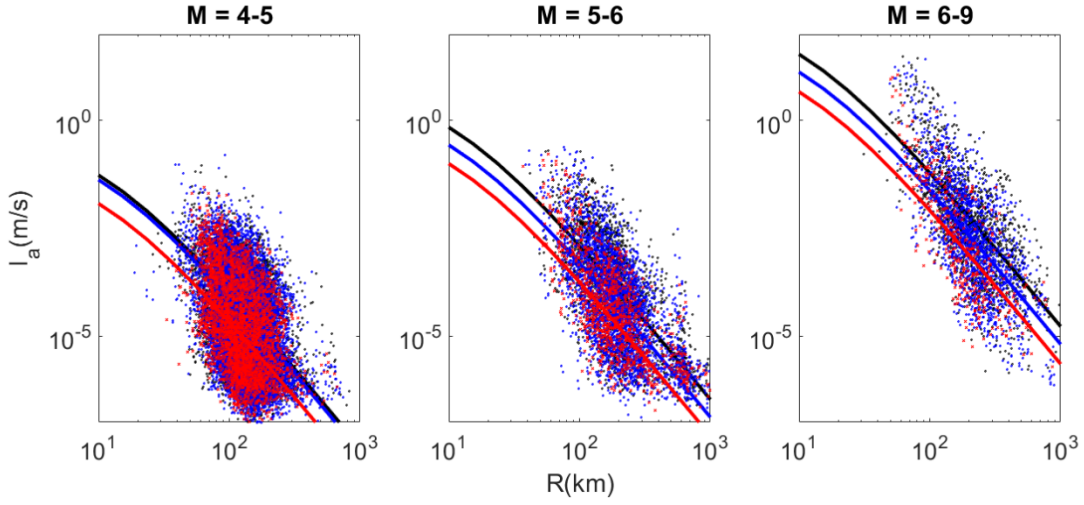
$$flag = \begin{cases} 1 & \boldsymbol{M} \leq c_1 \\ \frac{Mw - c_2}{c_1 - c_2} & c_1 \leq \boldsymbol{M} \leq c_2 \\ 0 & \boldsymbol{M} \geq c_2 \end{cases} \quad \textbf{14}$$

$$F_{site} = c_3 flag * \left[ln(Vs30) - c_4\right]^2 + c_5(1 - flag).\left(ln(Vs30) - c_6\right) + c_7\, \delta h_{800} \quad \textbf{15}$$

where a_1, a_2, a_3, b_1, b_2, b_3, and c_1 to c_4 are model parameters, M_w is moment magnitude, R is the closest distance to the ruptured fault, V is zero if the site-to-source path has crossed a volcanic belt and

1 otherwise. Figure 5 shows the prediction of subduction and shallow crustal Arias Intensity, respectively, alongside observed values of Arias Intensity for different magnitude and V_{S30} bins.

All the parameters for Equations 7 to 13, with the exception of the hinge points in the path attenuation (c_1 and c_2), are calculated using mixed-effect nonlinear regression. Table 1 shows the output of the regression for both subduction-zone and shallow crustal earthquakes. Ergodic coefficients correspond to the models for which the random effect for the site (e.g., the site term) is not included in their development. They are presented here only for studying the effect of adding site variability in GMPE development. Since observations and trial and error were used to determine the location of the hinge points (c_1 and c_2), they do not have an associated standard deviation. Table 2 shows the standard deviation of each component of the total variability (Equation 5).



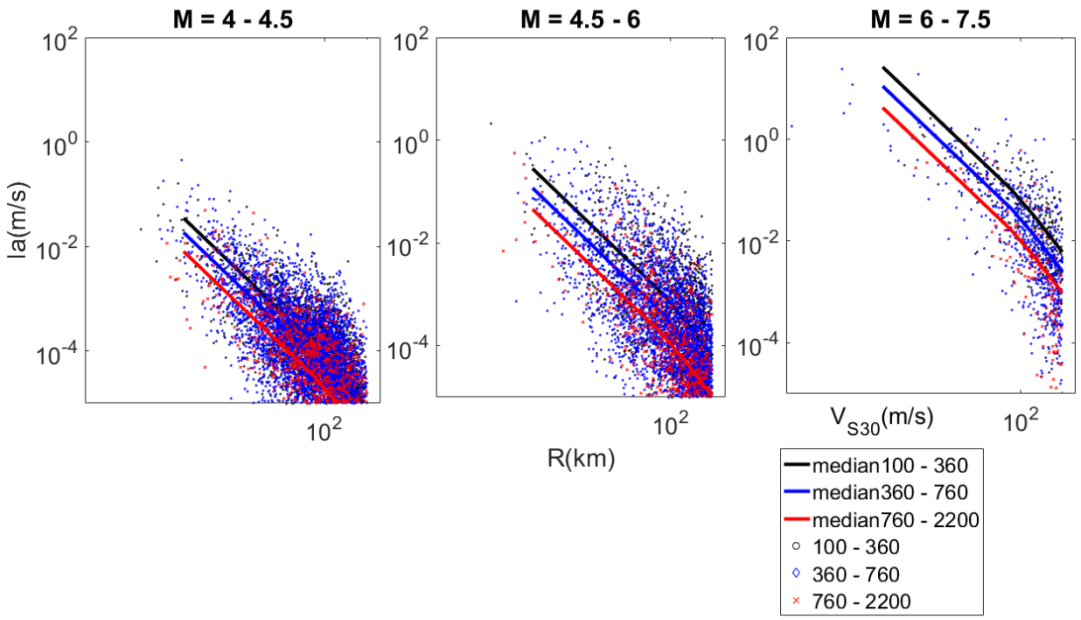


Figure 5: Comparison between Arias Intensity predictions for each magnitude and V_{S30} bin and observed values for shallow crustal earthquakes.

Table 1. Coefficients of the Arias Intensity models

	Subduction				Shallow crustal			
	Ergodic		Non-ergodic		Ergodic		Non-ergodic	
Parameter	Value	Standard deviation	Value	Standard deviation	Value	Standard deviation	Value	Standard deviation
a_1	0.08327	0.36572	-3.13473	0.68582	-11.419	0.79753	-10.85737	0.004177
a_2	2.35041	0.07890	2.531434	0.12607	3.19401	0.19988	3.228600	0.003935
a_3	1.95902	0.39168	1.994401	0.46158	2.74499	0.45547	2.87285	0.004188
a_4	0.65946	0.002553	0.724039	0.06619	0.45604	0.03006	0.404979	0.004140
a_5	27	-	27	-	10	-	10	-
a_6	1.74065	0.24382	1.110710	0.003241	-	-	-	-
b_1	-0.99229	0.01804	-0.76055	0.002546	-0.2234	0.03064	-0.325876	0.003837
b_2	-4.06505	0.02924	-3.70003	0.07158	-2.5275	0.02697	-2.604178	0.004173
b_3	26.88409	0.002365	9.81090	2.47057	-3.8917	0.08526	-3.572257	0.003995
c_1	4.5	-	4.5	-	4	-	4	-
c_2	5	-	5	-	5	-	5	-
c_3	-0.56500	0.03544	-1.37	0.004175	-0.71908	0.06066	-0.324002	0.003995
c_4	5.73276	0.03207	5.784335	0.002543	5.76759	0.04237	4.596193	0.004175
c_5	-1.31821	0.02736	-1.63000	0.001094	1.36197	0.04567	1.553334	0.004167
c_6	6.03858	0.08327	5.784301	0.001095	5.40656	0.20004	5.117818	0.004171
c_7	-0.00191	0.0000979	-0.00213	0.0002	-0.0010	0.000107	0.0011298	0.000108

Table 2. Standard deviation values of Arias Intensity models

	Subduction		Crustal	
Parameters	Ergodic	Non-ergodic	Ergodic	Non-ergodic
ϕ_{SS}	1.040993	0.7229769	0.8691	0.7582826
τ	1.163843	0.8513596	1.1848	0.8101719
ϕ_{S2S}	-	1.117143	-	0.8849222
σ	1.561473	1.579721	1.468202	1.387207

Scatter plot of event terms, site terms, and the residuals versus different variables are used to check performance of the models. Plots of the non-ergodic model for subduction and shallow crustal earthquakes are shown in Figures 6and 7. The red line is output of a LOWESS function, which is a non-parametric local average that indicates the general trend of the data; when the value of the red line is near zero the function is unbiased.

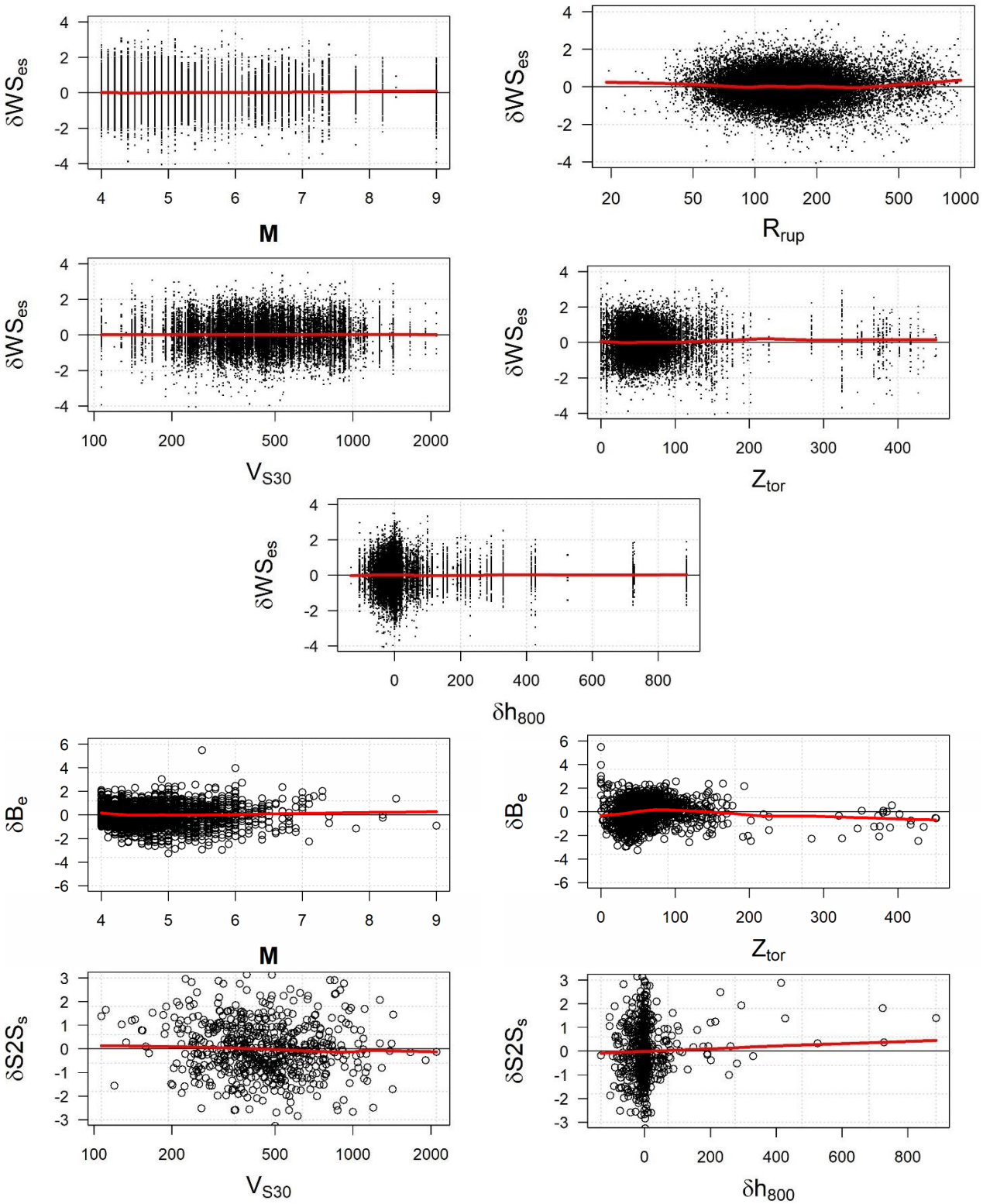


Figure 6. Scatter plots of residual component of Arias Intensity model for shallow crustal earthquakes versus various variables listed on the plots. $\boldsymbol{\delta S2S_S}$ and $\boldsymbol{\delta B_e}$ are site term and event term respectively

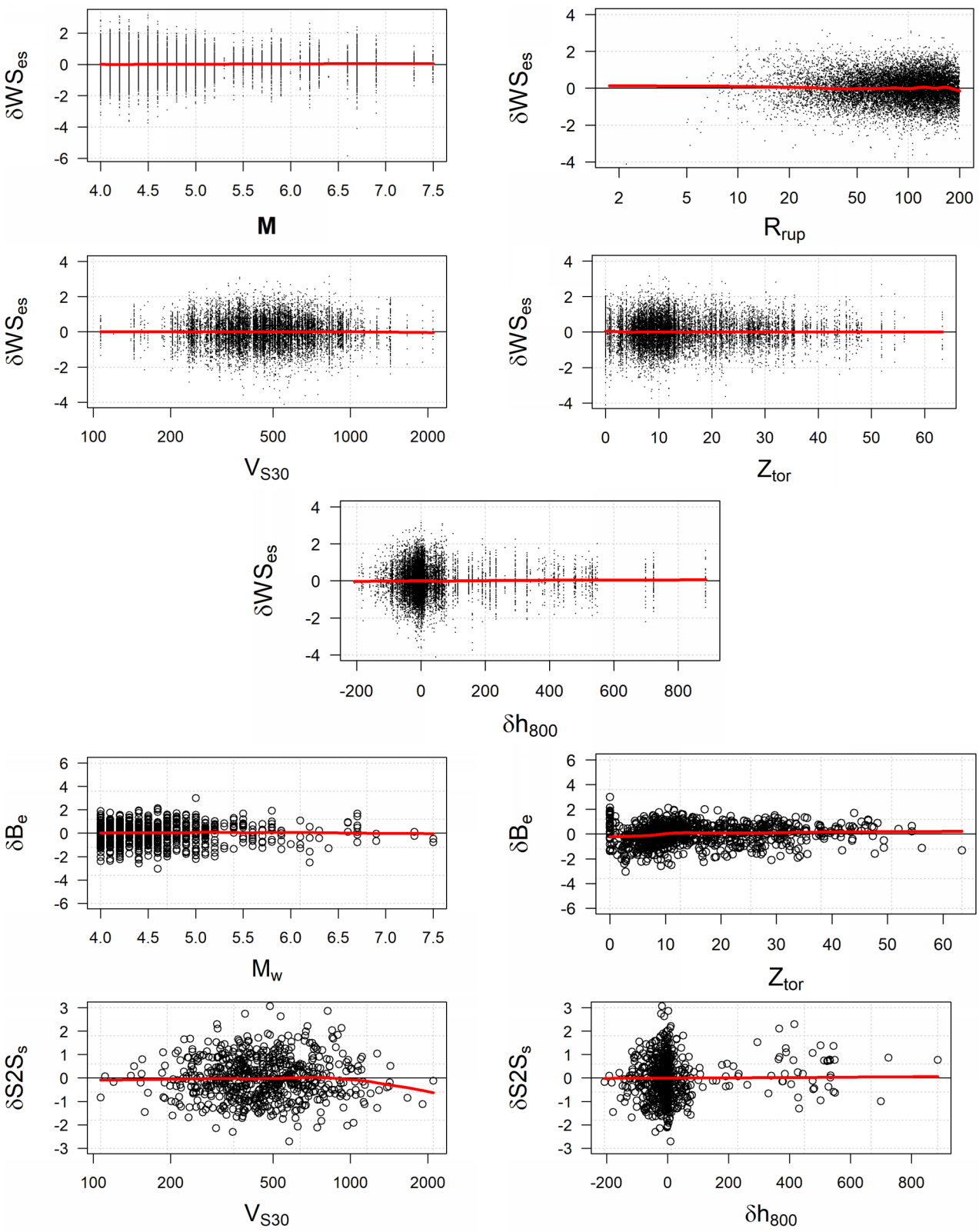


Figure7. Scatter plots of the Arias Intensity model for subduction earthquakes versus various variables listed on the plots. $\boldsymbol{\delta S2S_S}$ and $\boldsymbol{\delta B_e}$ are site term and event term respectively

2.4 Discussion

The derived models for Arias Intensity have two main differences with respect to the models in the literature: they do not consider the effect of focal mechanism nor the decrease in site amplification for high intensity motions. In this section first we discuss these issues further and then we compare the proposed relationships with those in the literature.

2.4.1 Effect of focal mechanism

Analyses of residuals indicate that the effect of the focal mechanism on Arias Intensity is not statistically significant and hence is not included in our relations. This counters the GMPEs proposed by other researchers (e.g., Foulser-Piggott and Goda 2015, Travasarou et al. 2003). Moreover, the models in the literature do not show the same trends for the effect of focal mechanism. In the Foulser-Piggott and Goda (2015) relations, the Arias Intensity of motions caused by normal faulting are higher than those of reverse faulting and the prediction of Arias Intensity for intraslab subduction earthquakes is higher than for interface earthquakes. In Foulser-Piggott and Stafford (2012), the predictions of Arias Intensity of motions caused by reverse faulting are higher than other types of focal mechanism. In the Travasarou et al. (2003) relationship, the lowest prediction of Arias Intensity corresponds to normal faulting and then strike slip and the highest is reverse or oblique reverse. Figure 8 illustrates the effect of focal mechanisms on the Arias Intensity of motions in our database. This plots shows that the effects of focal mechanisms on the Arias Intensity is not statistically significant.

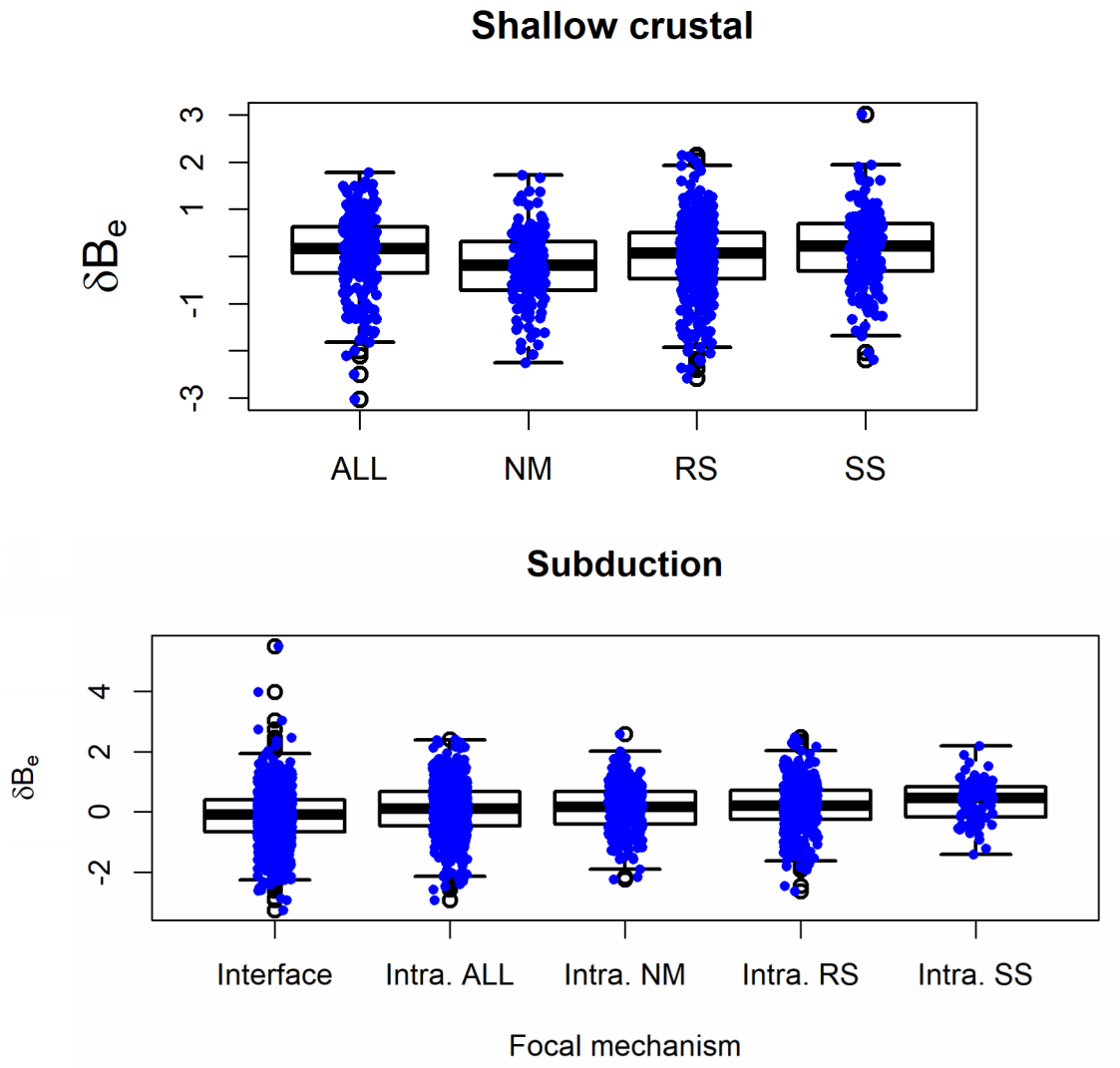


Figure 8. Box plots of event terms versus focal mechanisms for shallow crustal and subduction earthquakes. In the x-axis All, NM, RS, SS, and Intra. stand for unknown faulting type, normal slip, reverse slip, strike slip, and intraslab, respectively. Box plots show the median of each group with a thick line and the edge of each box represents the first and third quintile. In addition the gaps represents the inner and outer fence meaning any data passing them are outliers based on the Tukey (1977) criterion

2.4.2 Effect of site effects nonlinearity

The second difference between the models in this study and some other published GMPEs is we have not considered the nonlinear behavior for high intensity ground motions. Based on the site response analysis conducted by Papaspiliou (2010), Foulser-Piggott and Stafford (2012) showed that site amplification of Arias Intensity starts to decrease when the Arias intensity at rock exceeds some limits.

The Arias Intensity value that triggers this behavior and the slope of site amplification versus Arias Intensity at rock after triggering nonlinear behavior depend on the stiffness of sites. The stiffness of sites can be represented by V_{S30}. Foulser-Piggott and Goda (2015) and Foulser-Piggott and Stafford (2012) have included the nonlinear site amplification of Arias Intensity at high intensity motions. Figure 9 shows the residuals of Arias Intensity models versus the predicted Arias Intensity for each record at rock for sites having V_{S30} less than 400 m/s. Based on the mechanism just explained, we expected to see negative residuals at high intensities. This plot shows that while this behavior exists in our database for Arias Intensity values higher than 1 m/s (shown by red circles), the number of records that trigger nonlinear behavior of soil is too small to enable us model this behavior.

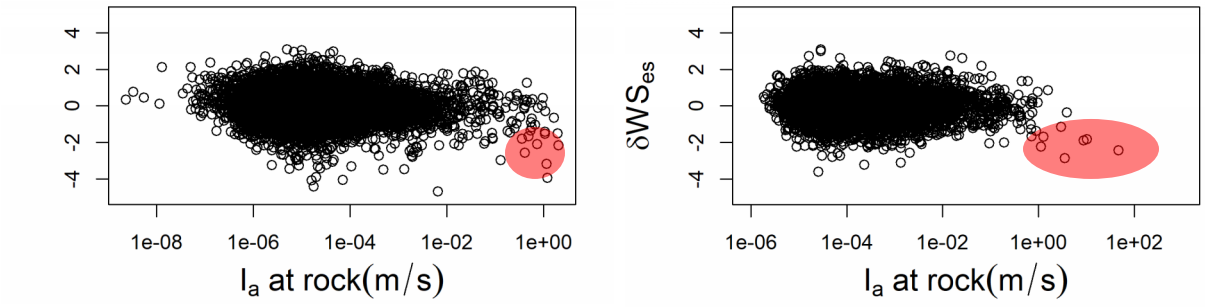


Figure 9. Scatter plot of model residual versus predicted Arias Intensity at rock(V_{S30} 400 m/s) for shallow crustal on left and subduction earthquakes on the right

2.4.3 Comparison with previous GMPEs

Figure 10 shows a comparison of the proposed relationship for Arias Intensity with other published relationships. The figure in the left column compares the predictions from this study with previous ones for a range of magnitudes and the figure in the right column show a similar comparison for a range of distances. As the Bullock et al.(2017) and Lee and Green (2010) relations are valid for rock, we use V_{S30} equal to 760(m/s) for models that require V_{S30}, and site condition B for Travasarou et al.

(2003) model. Also, the tectonic regime used for Bullock et al.(2017) and Foulser-Piggott and Goda (2015) are intraslab. The comparisons suggest that the attenuation rate of Arias Intensity in Japan is higher than other active tectonic regions. The two models that are developed based on Japanese data (this study, Foulser-Piggott and Stafford 2012, and the subduction relations of Bullock et al. 2017) show noticeably higher attenuation of Arias Intensity with distance when compared with the other models.

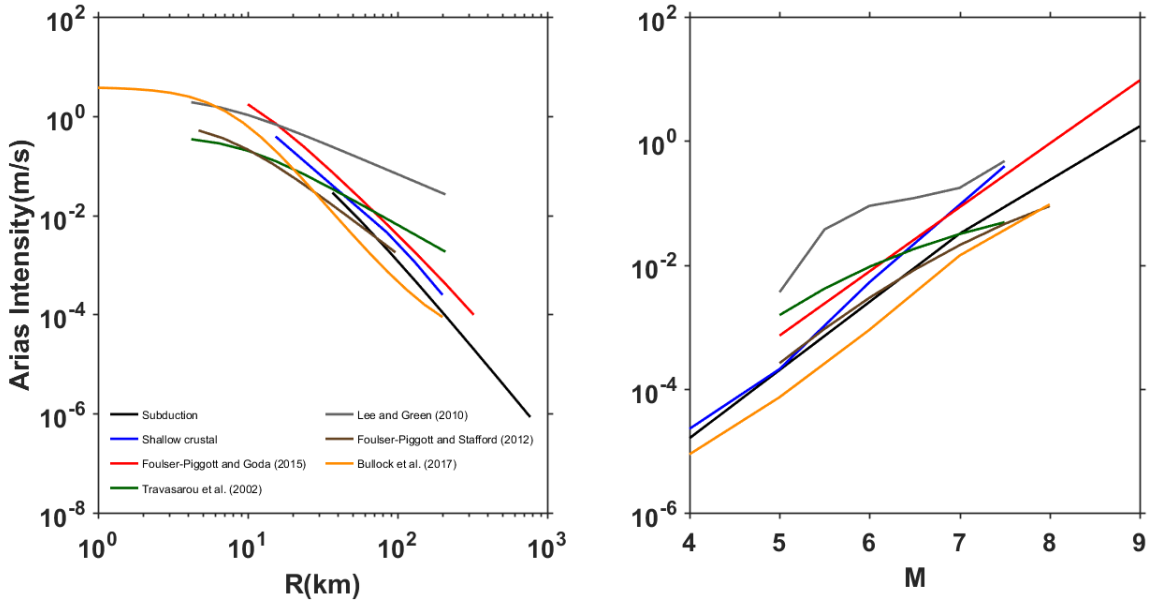


Figure 10. Comparison of of this study relations for Arias Intensity with literature. For the left plot the magnitude is fixed at 6 and for the right plot the distance is fixed at 70 km. Intraslab models are shown for Foulser-Piggott and Goda (2015) and for Bullock et al. (2017). For Travarsou et al. (2002) and Foulser-Piggot and Stafford (2012) normal faulting relations are presented. 760 m/s is used for the models that use V_{S30} and rock condition is used for the models that include site classification

The comparison shown in Figure 10 is not comprehensive because magnitude or distance are fixed in each plot. In addition, some studies (Foulser-Piggott and Goda 2015, and Foulser-Piggott and Stafford 2012, and Bullock et al. 2017) proposed models for different tectonic regimes that we have not presented here to avoid a crowded plot. Based on Scherbaum et al.'s (2010) suggestion, we make a comprehensive comparison of proposed relationships for Arias Intensity using Sammon's maps. Sammon's map is a visualization technique that gives a representation of points with higher dimension

in a 2D plot. In the following paragraphs, we explain the relations visualized in Figure 11 and then we explain the steps of plotting Sammon's map to clarify what it means and how the plot can be interpreted. Then, we discuss our interpretation of Figure 11.

To produce Sammon's map, first a set of values for distance and magnitude that all models are applicable for are chosen. For rupture distance, we use ten values ranging from 40 to 200 equally spaced in logarithmic domain. For magnitude, we use 5.25, 5.75, 6.25, 6.75, and 7.25 . Hence, for each relationship 50 values of Arias Intensity are calculated which corresponds to 50 different scenarios. Each relationship can be represented algebraically by a vector with 50 components, or by a point in a 50D space. The normalized distance between each point can be calculated by:

$$\overline{\Delta \mathrm{GM}_{\mathrm{i,j}}} = \sqrt{\frac{1}{N}\sum_{k=1}^{k=N}\left(\ln\left(I_a\right)_{i'k} - \ln\left(I_a\right)_{j'k}\right)^2} \qquad \textbf{14}$$

where $\ln\left(I_a\right)_{i'k}$ is the prediction of Arias Intensity by i^{th} GMPE for k^{th} scenario and N is the number of scenarios. Equation 14 gives a matrix of distances with zeros on the diagonal. Then, we search for a set of points in 2D space that have a distance matrix similar to the distance matrix computed for points 50D space. To accomplish that, a set of random points is chosen in 2D space and the difference between the distance matrix of points in 50D space and 2D space (quantified with error map that is explained in Sammon 1969) is minimized with moving the points in 2D. The minimization is done by the MASS package in R (Ripley 2014). Sammon's map is the plot of points in 2D.

Figure 11 compares the GMPEs for Arias Intensity of previous studies with the ones proposed here. This plot includes predictions of different tectonic regimes, and focal depths for the models that considers these factors. In addition, the plot includes the median of models and the median of models modified with different magnitude and distance scaling. To make the plot clearer we use abbreviations for studies, tectonic regime and scaling relations. The first part in each label of Figure 11 indicates the

study, the second part is the tectonic regime of relationship, and the third part indicates the focal depth used in calculating Arias Intensity. The abbreviations are described in Table 3.

Table 3: Description of the abbreviations used in Figure 11

Group	Abbreviation	description
Studies	T	This study relation ships
	FG	Foulser-Piggott and Goda 2015
	TBA	Travasarou et al. 2003
	FS	Foulser-Piggott and Stafford 2012
	B	Bullock et al. 2017
	LG	Lee and Green 2010
Tectonic regime	CR	Shallow crustal
	SUB	Subduction zone
	IS	Intraslab
	IF	Interface
	SC	Stable continental
Depth	10 or 27	The focal depth of earthquake in kilometers
Scaled relations	Mix	$\frac{1}{N}\sum_{relations} \ln(I_a)$
	S--/S++	$Mix \mp 1$
	S-/S+	$Mix \mp 0.5$
	M--/M++	$Mix \mp 0.8 * (\mathbf{M} - 6)$
	M-/M+	$Mix \mp 0.4 * (\mathbf{M} - 6)$
	R--/R++	$Mix \mp 1.2 * ln(\frac{Rrup}{70})$
	R-/R+	$Mix \mp 0.6 * ln(\frac{Rrup}{70})$

Figure 11 implies that the subduction relationships proposed in this study are close to the subduction relationships of Bullock et al. (2017) and the shallow crustal relationships proposed in this study are close to the relationships of Foulser-Piggott and Goda (2015). Considering the fact that the database of Foulser-Piggott and Goda (2015) and the subduction database of Bullock et al. (2017) come from Japan, this observation implies that regional effects are the primary effects on Arias Intensity of earthquakes. In fact, with the exception of the Bullock et al. (2017), the Japan-based relationships and

the relationships based on other databases separate well in Figure 11 (relationships based on Japanese data are plotted as triangles while others are plotted as circles).

Figure 11 also enables us to identify the difference between shallow crustal and subduction earthquakes and visualize the effect of focal depth. We observe that the predictions of shallow crustal motions are higher than subduction motions (i.e., predictiosn move along the axes of the scaled models S--, S-, Mix, S+ and S++). This is observed both for the models proposed in this study and for the models proposed by Bullock et al. (2017). Similarly, the Arias Intensity of earthquakes with depth equal to 27 km is higher than Arias Intensity of earthquakes with depth equal to 10 km. This is seen both for the relationships in this study and in the Bullock et al. (2017) study.

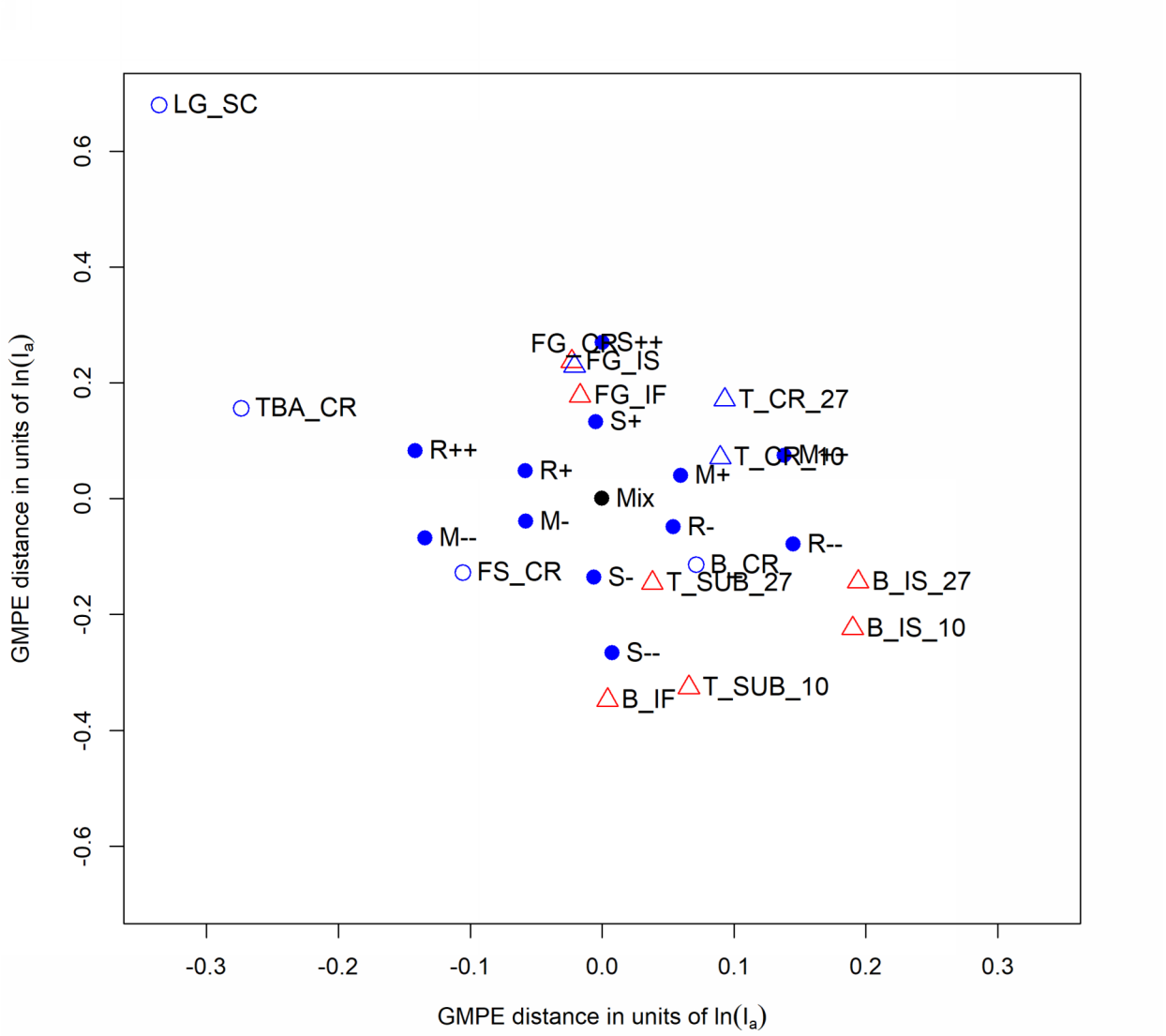


Figure 11. Sammon plot of GMPEs for Arias Intensity. Red color is used for subduction relationshisp and blue color for other regimes. Triangles are used for relationships based on Japanese data; otherwise, circles are used.

2.5 Conclusion

We have developed GMPEs for Arias Intensity of shallow crustal and subduction zone earthquakes. The shallow crustal GMPE is applicable for earthquakes with magnitudes ranging from 4 to 7.5 and the subduction GMPE is applicable for magnitudes ranging from 4 to 9. The shallow crustal GMPE is applicable for predicting ground motions with maximum rupture distance of 200 km, and the subduction GMPE is applicable for rupture distances less than 1000 km. The minimum R_{rup} for which

the equations are applicable depends on the magnitude. For shallow crustal equations, the minimum R_{rup} is 10 km for earthquakes with magnitude ranging from 4 to 5, 30 km for magnitudes ranging from 5 to 6, and 50 km for magnitudes larger than 6. For subduction zone equations, the minimum R_{rup} is 30 km for earthquakes with magnitudes ranging from 4 to 5, 60 km for magnitudes 5 to 6, and 100 km for magnitudes larger than 6. The input parameters for the models are moment magnitude (**M**), rupture distance (R_{rup}), average shear wave velocity over the upper thirty meters (V_{S30}), a binary (1 or 0) variable that describes whether the site-to-source path crosses the volcanic belt, and another binary variable that describes whether the event depth is larger than 200 km. this latter variable only applies for intraslab subduction earthquakes. The effect of tectonic regime is considered by developing separate models for shallow crustal and subduction earthquakes. The equations should not be used for sites with V_{S30} higher than 1500 m/s because the residuals are biased for this range of V_{S30}.

We observed that the amplification of I_a is magnitude dependent. This observation cannot be explained by the nonlinearity of site response because for soft soil sites the amplification decreases with decreasing magnitude. This behavior can be explained by considering the differences in frequency content between earthquakes of different magnitudes. Moreover, when clustering data into magnitude bins, we observe different slopes of duration versus R_{rup} for different magnitude bins. This behavior was captured by using a magnitude dependent path effect for duration equations.

When comparing shallow crustal and subduction earthquakes, we observe that I_a of shallow crustal earthquakes are higher than subduction earthquakes for similar magnitudes and distances. Other relations that include shallow crustal and subduction earthquakes show the same behavior. The relationships developed here are compared with relationships from literature. We observed that differences in the predictions of models using Japanese data suggest that there is a strong regional

effect in Arias Intensity. In addition, rate of decay of Arias Intensity with distance is higher in Japan than other regions of the world.

2.6 Reference

Abrahamson, N., and Youngs, R. (1992). "A stable algorithm for regression analyses using the random effects model." *Bulletin of the Seismological Society of America*, 82(1), 505-510.

Al Atik, L., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., and Kuehn, N. (2010). "The variability of ground-motion prediction models and its components." *Seismological Research Letters*, 81(5), 794-801.

Arias, A. (1970). "A measure of earthquake intensity." *R.J. Hansen, ed. Seismic Design for Nuclear Power Plant, MIT Press, Cambridge, Massachusetts,* pp. 438-483

Atkinson, G. M., and Boore, D. M. (2003). "Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions." *Bulletin of the Seismological Society of America*, 93(4), 1703-1729.

Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2006). "The lme4 package." URL http://cran r-project. org/740src/contrib/Descriptions/lme4. html.

Benito, B., and Herraiz, M. (1997). "An approach to the measurement of the potential structural damage of earthquake ground motions." *Earthquake Engineering and Struct Dynamics*, 26, 79-92.

Bullock, Z., Dashti, S., Liel, A., Porter, K., Karimi, Z., and Bradley, B. (2017). "Ground Motion Prediction Equations for Arias Intensity, Cumulative Absolute Velocity, and Peak Incremental Ground Velocity for Rock Sites in Different Tectonic Environments." Accepted for Publication in the *Bulletin of the Seismological Society of America*.

Dawood, H. M., and Rodriguez-Marek, A. (2013). "A Method for Including Path Effects in Ground-Motion Prediction Equations: An Example Using the Mw 9.0 Tohoku Earthquake Aftershocks." *Bulletin of the Seismological Society of America*, 103(2B), 1360-1372.

Dawood, H. M., Rodriguez-Marek, A., Bayless, J., Goulet, C., and Thompson, E. (2016). "A Flatfile for the KiK-net Database Processed Using an Automated Protocol." *Earthquake Spectra*, 32(2), 1281-1302.

Foulser-Piggott, R., and Goda, K. (2015). "Ground-motion prediction models for Arias intensity and cumulative absolute velocity for Japanese earthquakes considering single-station sigma and within-event spatial correlation." *Bulletin of the Seismological Society of America*, 105(4), 1903-1918.

Foulser-Piggott, R., and Stafford, P. J. (2012). "A predictive model for Arias intensity at multiple sites and consideration of spatial correlations." *Earthquake Engineering & Structural Dynamics*, 41(3), 431-451.

Frankel, A., McGarr, A., Bicknell, J., Mori, J., Seeber, L., and Cranswick, E. (1990). "Attenuation of high-frequency shear waves in the crust: Measurements from New York state, South Africa, and southern California." *Journal of Geophysical Research: Solid Earth*, 95(B11), 17441-17457.

Harp, E. L., and Wilson, R. C. (1995). "Shaking intensity thresholds for rock falls and slides: Evidence from 1987 Whittier Narrows and superstition hills earthquake strong-motion records." *Bulletin of the Seismological Society of America*, 85(6), 1739-1757.

Jibson, R. W. (2007). "Regression models for estimating coseismic landslide displacement." *Engineering Geology*, 91(2), 209-218.

Kayen, R. E., and Mitchell, J. K. (1997). "Assessment of liquefaction potential during earthquakes by Arias intensity." *Journal of Geotechnical and Geoenvironmental Engineering*, 123(12), 1162-1174.

Lee, C.-T., Huang, C.-C., Lee, J.-F., Pan, K.-L., Lin, M.-L., and Dong, J.-J. (2008). "Statistical approach to earthquake-induced landslide susceptibility." *Engineering Geology*, 100(1), 43-58.

Lee, J., and Green, R. A. (2010) "An empirical arias intensity relationship for rock sites in stable continental regions." *Proceedings of the 7th International Conference on Urban Earthquake Engineering (7CUEE), 2010, Tokyo, Japan*

Lin, P.-S., Chiou, B., Abrahamson, N., Walling, M., Lee, C.-T., and Cheng, C.-T. (2011). "Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models." *Bulletin of the Seismological Society of America*, 101(5), 2281-2295.

Mackie, K., and Stojadinovic, B(2003). "Optimal probabilistic seismic demand model for typical highway overpass bridges." *Proc., 12th European Conference on Earthquake Engineering*, 9-13.

McVerry, G. H., J. X. Zhao, N. A. Abrahamson, and P. G. Somerville (2000). "Crustal and subduction zones attenuation relations for New Zealand earthquakes." *Proceedings of the 12th World Conference of Earthquake Engineering 2000, Auckland, New Zealand*.

Papaspiliou, M. I. (2010). "On the incorporation of site response in probabilistic seismic hazard analyses." Ph.D. Thesis, Imperial College London, 2010.

Ripley, B. (2014). "Support Functions and Datasets for Venables and Ripley's MASS. R Package Version 7.3-30."

Rodriguez-Marek, A., Montalva, G. A., Cotton, F., and Bonilla, F. (2011). "Analysis of single-station standard deviation using the KiK-net data." *Bulletin of the Seismological Society of America*, 101(3), 1242-1258.

Rodriguez-Marek, A., Cotton, F., Abrahamson, N. A., Akkar, S., Al Atik, L., Edwards, B., Montalva, G. A., and Dawood, H. M. (2013). "A model for single-station standard deviation using data from various tectonic regions." *Bulletin of the Seismological Society of America*, 103(6), 3149-3163.

Sammon, J. W. (1969). "A nonlinear mapping for data structure analysis." *IEEE Transactions on computers*, 100(5), 401-409.

Scherbaum, F., Kuehn, N. M., Ohrnberger, M., and Koehler, A. (2010). "Exploring the proximity of ground-motion models using high-dimensional visualization techniques." *Earthquake Spectra*, 26(4), 1117-1138.

Stafford, P. J. (2014). "Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models." *Bulletin of the Seismological Society of America*, 104(2), 702-719.

Stafford, P. J., Berrill, J. B., and Pettinga, J. R. (2009). "New predictive equations for Arias intensity from crustal earthquakes in New Zealand." *Journal of Seismology*, 13(1), 31-52.

Travasarou, T., Bray, J. D., and Abrahamson, N. A. (2003). "Empirical attenuation relationship for Arias intensity." *Earthquake Engineering & Structural Dynamics*, 32(7), 1133-1155.

Trifunac, M. D., and Brady, A. G. (1975). "A study on the duration of strong earthquake ground motion." *Bulletin of the Seismological Society of America*, 65(3), 581-626.

Tukey, J. W. (1977). "Exploratory data analysis." *Addison-Wesley, Boston, MA*

Wilson, R. C., and Keefer, D. K. (1985). "Predicting areal limits of earthquake-induced landsliding." *Ziony, J.I. (Ed.), Evaluating Earthquake Hazards in the Los Angeles Region — An Earth-science Perspective. US Geol. Surv. Prof. Paper* 1360, 316–345.

Chapter 3
Ground Motion Prediction Equations for Significant Duration Using the Kik-Net Database

Abstract

Significant duration of strong shaking quantifies the length of time in which earthquake energy arrives at a site. Significant duration has multiple applications in Geotechnical and Structural Engineering. At the time of this publication, there are no published ground motion prediction equations for duration applicable to subduction earthquakes. To address this need and to identify the difference between significant duration of motions resulting from earthquakes from different tectonic regimes, we develop predictive equations for significant duration applicable to interface and intraslab subduction earthquakes and shallow crustal earthquakes using the KiK-net database. The GMPEs are applicable to earthquakes with **M** 4-9. In the proposed relations, we capture the effect of earthquake size on path duration. Based on the relationships proposed in this study, we note that the duration of shallow crustal earthquakes are longer than the duration of subduction earthquakes. This study's predictions of duration are consistent with previous studies for small to moderate shallow crustal earthquakes and are generally longer for large ones.

3.1 Introduction

In this paper we develop a set of significant duration ground motion prediction equations (GMPEs) applicable to three earthquake types: interface and intraslab subduction earthquakes and shallow crustal earthquakes. The proposed GMPEs are developed using data recorded by the Kiban-Kyoshin network

(KiK-net) database. The primary contribution of this work is to fill the gap in existing GMPEs for significant duration, since there are no GMPEs available for significant duration for subduction earthquakes. In addition, because we use data from the same geographical region; the results of this study can help to identify the effect of earthquake type on the significant duration of ground motions.

The duration of shaking is a fundamental ground motion characteristic that is important for engineering and seismological applications. In the latter, ground motion duration is used for the generation of time histories in stochastic simulations (Boore 2003; Boore and Thompson 2012; Boore and Thompson 2014). The empirical duration models typically employed in stochastic simulations have been simple functions with only two terms: magnitude and distance. More elaborate models (e.g., Bommer et al. 2009) now include terms for the depth to top of rupture (Z_{tor}), the average shear wave velocity over the upper 30 m (V_{S30}), and a more complex magnitude distance functional form. Duration is particularly important for the stochastic simulation method because if all parameters are held constant, the response spectral ordinates, particularly for short periods, decrease as duration increases (Boore 2003). Stochastic point source or stochastic finite fault models have been used to model subduction earthquakes (e.g., Atkinson and Boore 1997; Gregor et al. 2002; Atkinson and Macias 2009; Ghofrani et al. 2013), typically by developing simple event-specific duration models or using simple duration models developed from correlations of time-domain and spectral domain parameters (e.g., Atkinson 1995). In general, stochastic modeling of subduction-zone earthquakes will benefit from the proposed duration model for subduction-zone earthquakes.

Duration of shaking is also important for engineering problems, such as damage evaluation of buildings (Hancock and Bommer 2006; Kunnath and Chai 2004; Chai 2005; Uang and Bertero 1988) and the evaluation of liquefaction triggering. Liquefaction of loose saturated sand has resulted in large amounts of damage in various earthquakes, including subduction earthquakes (e.g., the 1964 Good Friday earthquake in Alaska, the 2001 Nisqually earthquake in Washington State, and most recently the 2011

Tohoku, Japan, and 2011 Maule, Chile, earthquakes). Liquefaction triggering is a phenomenon that occurs in saturated sandy soils involving the transfer of overburden stress from the soil skeleton to the pore fluid, with the commensurate increase in pore water pressure and reduction in effective stress and soil strength. Several laboratory studies have shown that pore water pressure rise is related to cumulative strain energy during cyclic testing (Ishac and Heidebrecht 1982; Law et al. 1990; Green et al. 2000; Polito et al. 2008). For a given soil profile, strain energy (and hence, pore water pressure rise) increases not only with the amplitude of ground motion, but also with duration of shaking. In the traditional liquefaction triggering evaluation approach, the correlation of liquefaction potential to ground motion duration is accounted for by means of the magnitude scaling factor (MSF; e.g., Youd et al. 2001). These factors have been largely constrained using shallow crustal earthquakes, and their application to subduction regions implies a degree of uncertainty. This is particularly relevant for several sites on the Pacific Northwest and Alaska, where the evaluation of liquefaction potential has to account both for earthquakes in shallow crustal faults with magnitude typically smaller than M_w 7.5, but capable of producing significant accelerations due to proximity of the fault, and for mega thrust events occurring in the subduction-zone off the Pacific Coast. These larger events (e.g., M_w9+ for the Cascadia subduction zone) may not control the hazard in the high frequency range, but they could control the liquefaction hazard due to the large duration of strong ground motion.

There are several definitions of ground motion duration (Bommer and Martinez-Pereira 1999). The *significant duration* used in this work is defined as the interval between the times when the Arias Intensity (Arias 1970) of an earthquake reaches two relative thresholds. In this work, we develop GMPEs for $Ds_{5\text{-}95}$ and $Ds_{5\text{-}75}$, which corresponds to the time interval between Arias Intensity equaling 5% and 95% of its final value and the time interval between Arias Intensity equaling 5% and 75% of its final value. We use the geometric mean of the duration of the two as-recorded horizontal components of motion. In this work, any references to “duration” are therefore references to the

geometric mean of both Ds_{5-95} and Ds_{5-75}. Several researchers have developed GMPEs for significant duration of ground motions. Bommer et al. (2009; heretofore referred to as BSA09) used the Next Generation Attenuation (NGA) database (Chiou et al. 2008) to develop predictive equations for various ground motion parameters including significant duration. The BSA09 equations, however, are only applicable to shallow crustal earthquakes for rupture distances less than 100 km and magnitudes from 4.8 to 7.9. Du and Wang (2016; heretofore referred to as DW16) and Afshari and Stewart (2016; heretofore to as AS16) used the NGA-West2 database (Ancheta et al. 2014) to develop prediction equations for significant duration. These equations, again, are applicable for shallow crustal active tectonic regions but they cover a wider range of magnitude and distance. Lee and Green (2014; heretofore LG14) developed predictive equations for significant duration applicable for western North America and central and eastern North America. Finally, Yaghmaei-Sabegh et al. (2014) used the ground motion data recorded in Iran to develop a duration GMPE

In this paper we first explain the database used in developing the GMPEs and then describe the regression approach used in developing the predictive models. We then introduce GMPEs for significant duration for different tectonic regimes and discuss the scaling of these parameters with the selected predictive variables. Finally, we present a comparison of the proposed model for shallow crustal and subduction environments with results from other published GMPEs.

3.2 Database

The ground motions used in this work are a subset of the motions recorded by the Kiban-Kyoshin network (KiK-net) in Japan, gathered and processed by Dawood et al. (2016). The ground motion data available from the KiK-net website was processed using an automated processing protocol. The station and event metadata was also collected and is available at NEEShub (https://nees.org/resources/7849).

This metadata includes a classification of each event as either a subduction-zone or shallow crustal event using the criteria of Garcia et al. (2012).

The development of the GMPEs presented in this work use a subset of the Dawood et al. (2016) dataset. Motions for which V_{S30}, tectonic regime, or sampling frequency are unknown are excluded. We exclude all records with **M**<4. While most engineering applications use records from $\mathbf{M} \geq 5$, we use a lower magnitude threshold to ensure that the GMPE is well constrained for all earthquakes of engineering significance. This follows a suggestion by Bommer et al. (2007), who showed that current GMPEs are not applicable at the bounds of the magnitude range of the data underlying the GMPE. A final exclusion criterion is based on distance. Shallow crustal motions with rupture distances greater than 200 km and subduction-zone motions with rupture distances greater than 1000 km are excluded from the dataset. The resulting ground motion database for shallow crustal earthquakes includes 13,966 motions from 976 events, and the ground motion database for subduction earthquakes includes 30,606 motions from 2,332 events. Figure 1 shows the distribution of magnitude versus rupture distance and the geographical distribution of earthquake sources. As sources generating subduction-zone earthquakes are mostly offshore, our models for subduction-zone earthquakes are not well constrained for distances below 40 km. Similarly, there is little data for shallow crustal earthquakes for distances lower than about 10 km. We will state stronger limits on applicability of our GMPEs based on residual plots in the conclusions. Figure 2 shows the number of stations that have multiple recordings from shallow crustal and subduction tectonic regimes. As there are not adequate stations with V_{S30} above 1000 m/s and below 200 m/s, the developed GMPEs should be used with caution for such stations.

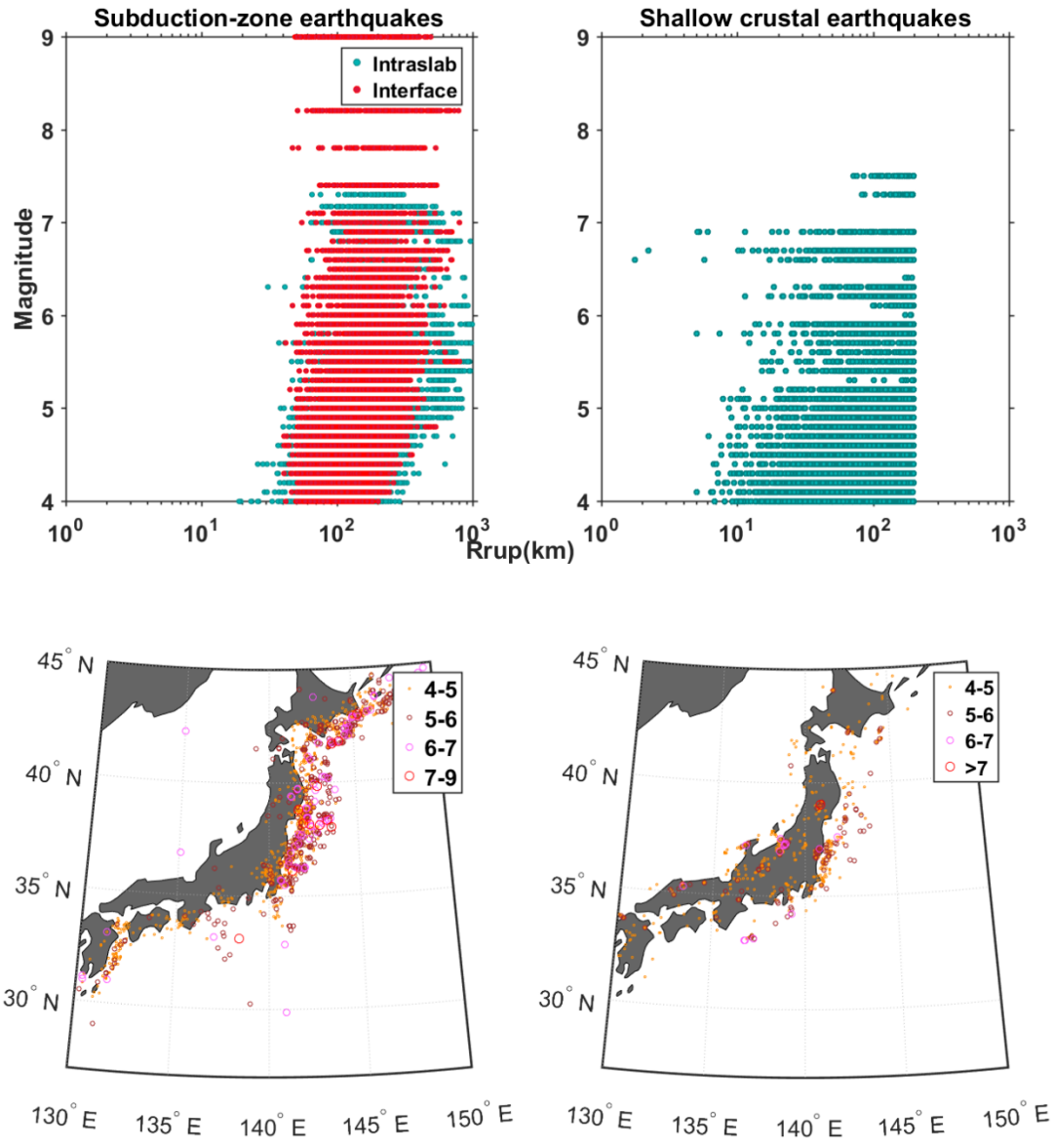


Figure 1: Upper figures show magnitude and distance distributions for each tectonic regime. Lower figures are illustrations of the event source distributions. Plots on the left are for subduction earthquakes, and on the right for shallow crustal earthquakes.

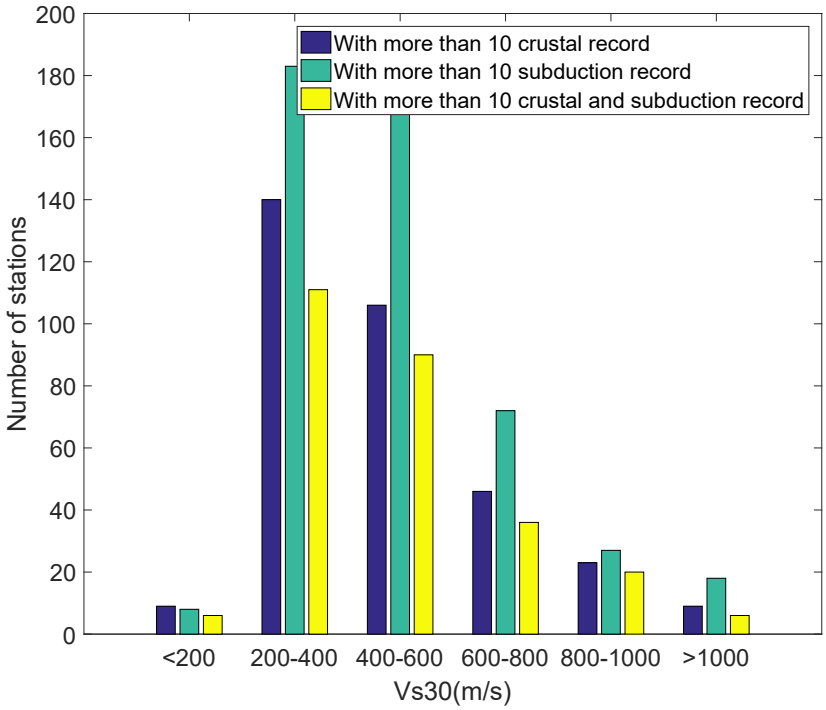


Figure 2: Number of station that have multiple recordings from shallow crustal and subduction tectonic regions

3.3 Model development

The model adopted herein for prediction of duration includes a source, a path, and a site term. This section presents first the regression approach, followed by the development of the functional form and the model for aleatory variability. Finally, the performance of the model, in terms of the observed residual, is evaluated and discussed.

3.3.1 Regression approach

As discussed in the next subsection, the functional form considers source and path duration terms that are added in arithmetic domain (e.g., these are additive terms). However, the residuals are assumed to follow a lognormal distribution. That makes estimating the model parameters simultaneously not possible because of convergence issues. To overcome this problem, the model parameters are estimated

in three stages; in each stage, we estimate the parameters of the source, the path, or the site effect terms and fix the other two. These steps are repeated to ensure consistency.

Abrahamson and Youngs (1992) noticed that an uneven distribution of recordings per event may introduce biases in the regression. To correct for this, Abrahamson and Youngs (1993) proposed a random effects algorithm. In the Abrahamson and Youngs (1992) method, each event is associated with an event term (δB_e), which is the estimated difference between the recorded intensity measures for a specific event e and the median model. The same issues can arise if the number of recordings per stations in a dataset are uneven. Following Stafford (2014), we address this issue by including random effects for station and event terms in each stage of the regression analyses. Hence, the predicted intensity measure $y_{e,S}$ is given by:

$$\ln y_{e,S} = \ln \widehat{y_{e,S}} + \delta S2S_S + \delta B_e + \delta W_{es} \quad \textbf{1}$$

$$\ln \widehat{y_{eS}} = f(\boldsymbol{x},\boldsymbol{\theta}), \quad \textbf{2}$$

where y_{eS} is the observed value of intensity measure for earthquake e and station s, $\widehat{y_{eS}}$ is the median predicted value of intensity measure $y_{e,S}$, which in turn is a function of the explanatory variables $\boldsymbol{x}$ (e.g. magnitude, distance) and fixed effect parameter $\boldsymbol{\theta}$ (model coefficients); δB_e is the event term calculated for the event e, and $\delta S2S_s$ is the site term calculated for station s. The terminology follows that proposed by Al Atik et al. (2010).

The event term (δB_e), the site term ($\delta S2S_s$), and the residual (δW_{es}, which are called site- and event-corrected residual in Al Atik et al. 2010) are assumed to be zero-mean random variables with standard deviations denoted by τ, ϕ_{S2S}, and ϕ_{SS}, respectively . Moreover, these random variables are assumed to be statistically independent, hence the total standard deviation (σ) is computed by:

$$\sigma = \sqrt{\phi_{SS}^2 + \tau^2 + \phi_{S2S}^2} \quad \textbf{3}$$

3.3.2 Functional form for the median duration

It is common in the development of duration models to use a functional form with a term for the significant duration at the source which is modified to account for path and site effects. Two mechanisms cause the increase in duration of ground motion when it propagates from source to site. The first mechanism is the time difference between arrivals of waves as shear and compression waves propagate with different velocity in soil. The interval between arrivals of different waves is only a function of distance; hence, it can be calculated independent of earthquake size and its effect is to add to the source duration. The second mechanism is the scattering and spreading of waves when they propagate from source to site. This path effect can be dependent on earthquake size as the amount of scattering is a function of frequency content and fault dimensions. Site effects, on the other hand, tend to stretch the duration of an earthquake due to wave trapping and the difference in the amplification of different frequency components. This effect is captured by multiplying the duration (e.g., an addition in logarithmic domain).

The models presented in AS16, KS06, and LG14 include a term for path duration that is added to the source duration in the arithmetic domain. Hence, this functional form assumes that only the first mechanism described above affects durations. As will be shown later, our data does show magnitude dependence of the path effect. To capture this effect, we add the source and path duration in the arithmetic domain, and we capture the effect of earthquake size on the path duration model by making the distance slope magnitude dependent. In addition, we consider the effect of site response on duration by adding a site effect term in logarithmic domain, which implies a multiplication.

Based on an evaluation of the observed durations, we selected a modified version of the functional form KS06 to model the significant duration at source. The KS06 equation for significant duration at the source is:

$$ln(Ds\ at\ R=0)\ = \frac{\left(\frac{\Delta\sigma}{10^{1.5\boldsymbol{M}\ +\ 16.05}}\right)^{-1/3}}{4.9.10^{6}} \qquad 4$$

where $\Delta\sigma$ is the stress drop in dyne-cm and $\boldsymbol{M}$ is moment magnitude. KS06 considered the logarithm of stress drop as a linear function of magnitude:

$$\Delta\sigma = \exp\left(b_1 + b_2\boldsymbol{M}\right) \qquad \mathbf{5}$$

Equations 4 and 5 can be rewritten as:

$$ln(Ds\ \ at\ R=0)\ =\ 10^{(1.5\ -\ \log(e)\cdot b_2)\left[\boldsymbol{M}\ -\left(\frac{\log(e)\cdot b_1\ -\ 16.05}{1.5\ -\ \log(e)\cdot b_2}\ +\ \log\left(4.9\cdot 10^{6}\right)\right)\right]} \qquad \mathbf{6}$$

where e is Euler's number. For regression purposes, the first and the second parentheses are replaced with two model parameters. To evaluate the performance of this functional form, we fix the path model parameters and calculate source duration of all the earthquakes using a random effects regression. Figure 3 presents the resulting source durations as a function of magnitude for $Ds_{5\text{-}95}$. We have not presented the plots of $Ds_{5\text{-}75}$, but the same observations hold true for $Ds_{5\text{-}75}$. This figure suggests the need to add an intercept to Equation 6. The intercept is dependent on fault mechanism for shallow crustal earthquakes and intraslab earthquakes. Observe that reverse slip earthquakes are longer than reverse slip (Figure 3).

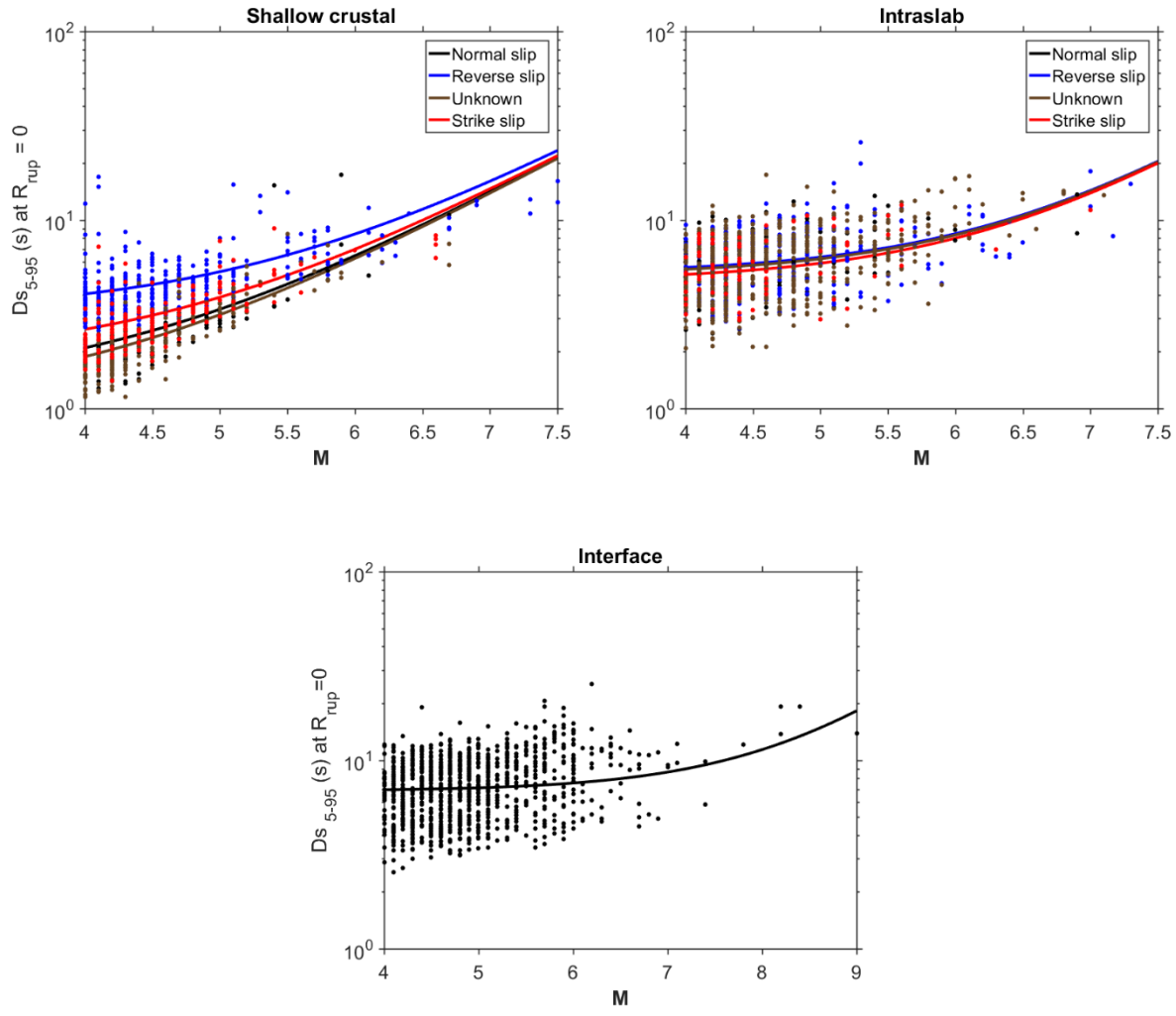


Figure 3: Source duration of earthquakes versus magnitude alongside with prediction of source duration for different tectonic regims

Figure 4 shows the duration of shallow crustal earthquakes plotted as a function of distance for different magnitude bins. The data shown in the plot indicates that for shallow crustal earthquakes the duration of ground motions saturates completely after a distance of 60 km when magnitude is smaller than 5. The effect of this saturation decreases with earthquake size. For earthquakes larger than 7, we observe no saturation. To account for this, distance-scaling is modelled with a bilinear plot. At short distances, the slope of duration versus distance is independent of magnitude (equal to r_1 in Equation 11). At large

distances, the distance slope varies linearly from 0 at M=5 to r_1 at M=7. The blue lines in Figure 4 show the prediction of duration for the center of each magnitude bin.

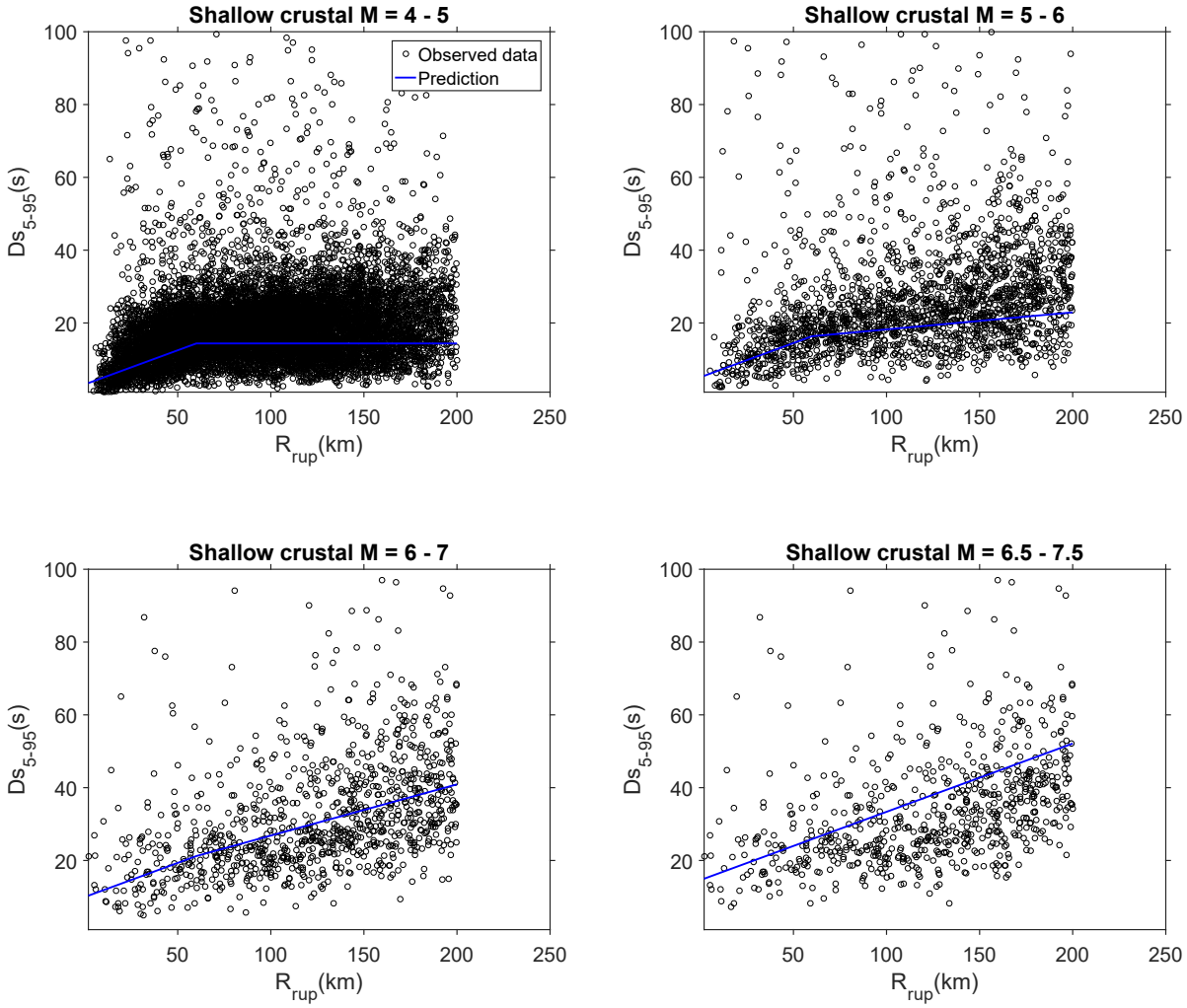


Figure 4: Duration of shallow crustal earthquakes for different magnitude bins versus distance

Figure 5 presents the duration of Intraslab and Interface earthquakes versus distance for different magnitude bins. Based on the data shown in these figures, we set the slope of duration versus distance to zero for distances smaller than 40 km and for distances larger than 325 km. To develop the functional form for path effect, we fit a line to each magnitude bin and plot the slope of each line as a function of magnitude in Figure 6. The slope duration versus distance is constant for magnitudes lower than 5.5

and increases with magnitude for larger magnitudes. We model this behavior in our functional form with the solid lines shown in Figures 5 and 6. As discussed before, we expect, based on physical considerations, the slope of duration against distance being dependent on magnitude. This effect is clearly seen in Figure 6 as the slope of duration against distance increases with magnitude.

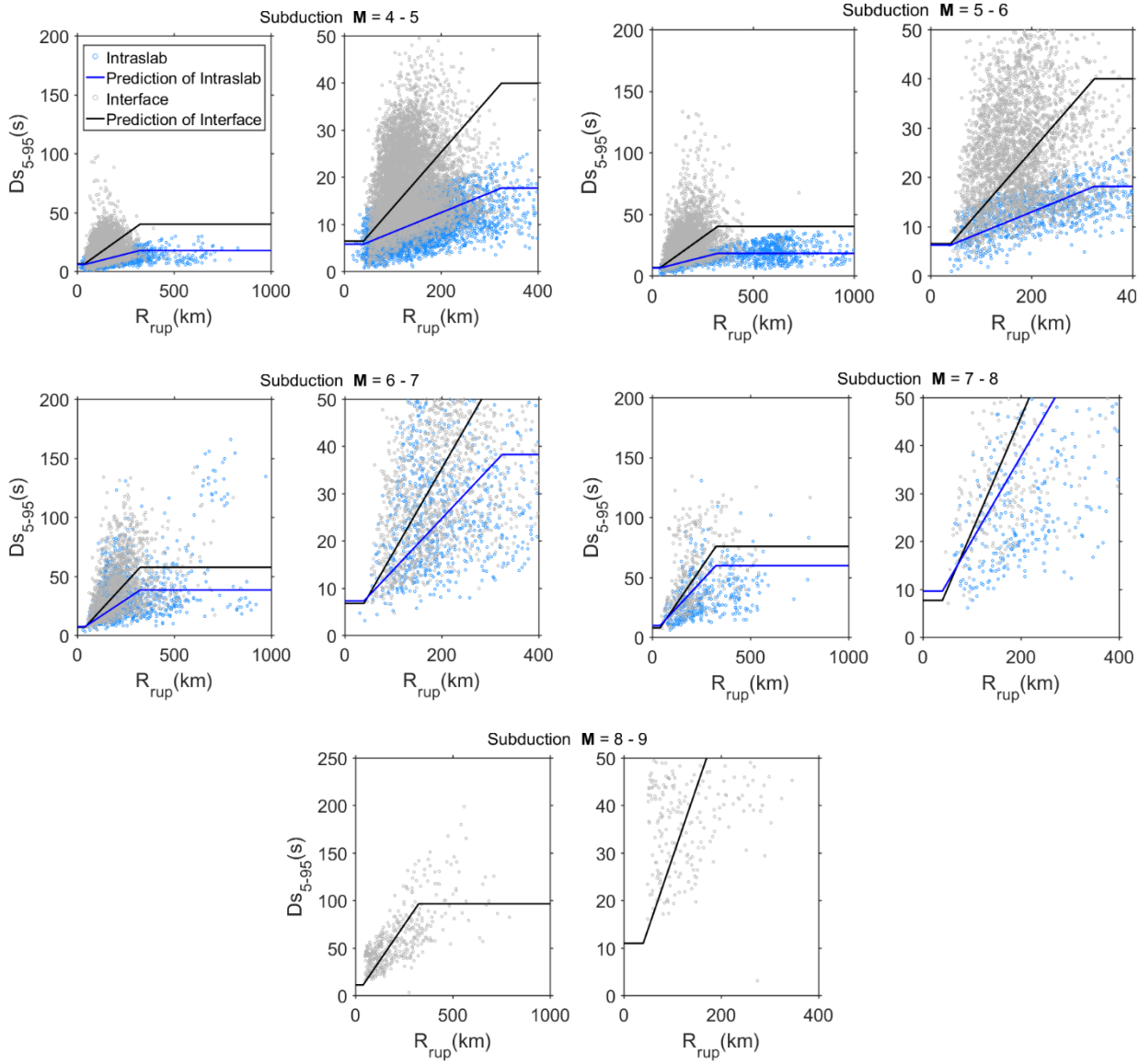


Figure 5: Duration of subduction earthquakes for different magnitude bins versus distance. For each magnitude bin, the plot on left and right show the data with different x and y axis limits

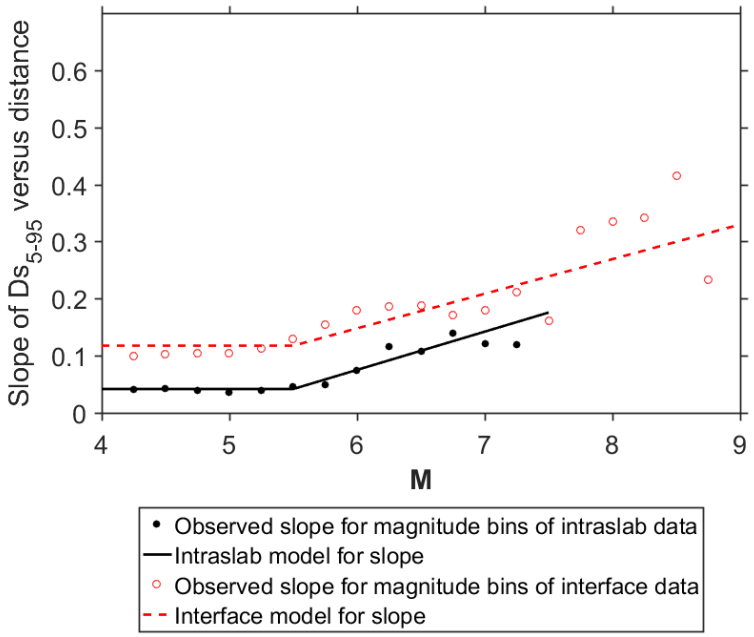


Figure 6: Path duration slop versus distance for each magnitude bin and the model fitted to them

We model site effects on Ds using two site parameters: V_{S30} and the depth to a shear wave velocity of 800 m/s (h_{800}). Since these two parameters are correlated, we use a functional form with the differences between the observed h_{800} at a site minus the predicted h_{800} using V_{S30} (this is labelled δh_{800}) in the site effect functional form. The predicted h_{800} is

$$h_{800,p} = exp(-5.23/2 * ln(\frac{V_{S30}^2 + 412^2}{1360^2 + 412^2}) - 0.9) \quad \textbf{7}$$

where $h_{800,p}$ is the h_{800} predicted based on V_{S30} of site in meters. The difference between h_{800} observed at a site minus the h_{800} predicted based on V_{S30} is:

$$\delta h_{800} = h_{800} - h_{800,p} \quad \textbf{8}$$

Figure 7 presents the residuals of the source and path duration models (in natural log scale) plotted as a function of V_{S30} and δh_{800} for $Ds_{5\text{-}95}$. Note that the residuals do not show any clear trend for V_{S30}

greater than 600 m/s and for δh_{800} greater than 250 m. To reflect this observation, the proposed site term scales with these parameters only up to these values.

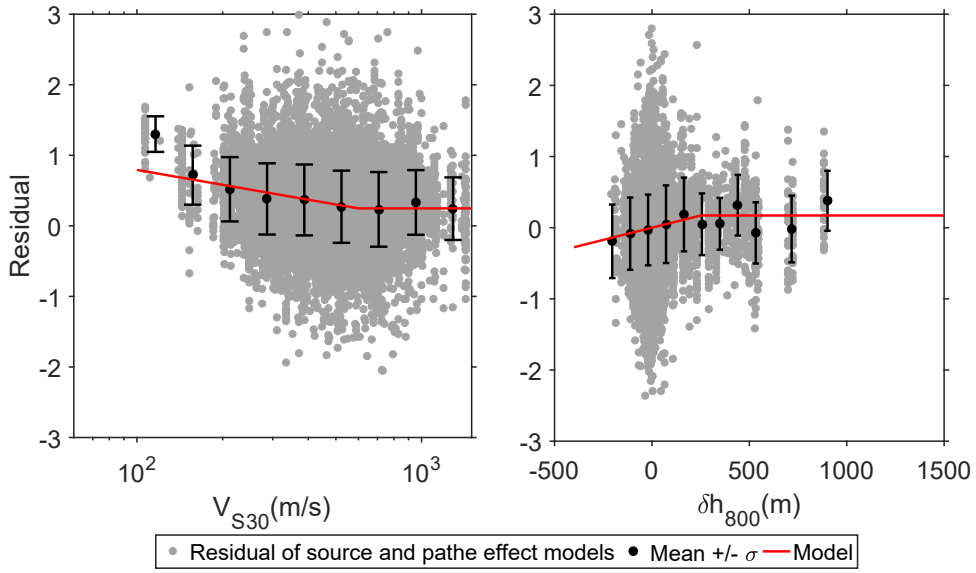


Figure 7: Residual of source and path duration model versus distance

Based on the previous discussion, the proposed functional form for predicting significant duration is:

$$\ln(Ds) = \ln\left(F_{source}(\boldsymbol{M}, mechanism) + F_{path}(R_{rup}, \boldsymbol{M})\right) + F_{site}(V_{S30}, \delta h_{800}) \quad \textbf{9}$$

The predicted duration at the source is:

$$F_{Source} = 10^{m_1(M_w - m_2)} + m_3(mechanism) \quad \textbf{10}$$

For shallow crustal earthquakes, the path term is given by:

$$F_{path}(R) = \begin{cases} r_1 * R & R_{rup} \le R_1 \\ r_1 * R_1 + MES * r_1 * (R_{rup} - R_1) & R_{rup} > R_1 \end{cases} \quad \textbf{11}$$

$$MES = \begin{cases} 0 & \boldsymbol{M} < M_1 \\ \frac{\boldsymbol{M} - M_1}{M_2 - M_1} & M_1 < \boldsymbol{M} < M_2 \\ 1 & \boldsymbol{M} > M_2 \end{cases} \quad \textbf{12}$$

And for subduction motions, the path term is given by:

$$F_{path}(R) = \begin{cases} 0 & R_{rup} \leq R_1 \\ \left(\mathrm{r}_2 * \max(\mathbf{M},5) + \mathrm{r}_1\right) * \left(R_{rup} - R_1\right) & R_1 < R_{rup} \leq R_2 \\ \left(\mathrm{r}_2 * \max(\mathbf{M},5) + \mathrm{r}_1\right) * \left(R_2 - R_1\right) & R_{rup} > R_2 \end{cases} \quad \textbf{13}$$

Finally, the site term is given by :

$$F_{site}(Vs_{30}) = s_1 * \ln\left(\frac{\min\left(Vs_{30}, V_1\right)}{\mathrm{V}_1}\right) + s_2 * \min\left(\delta h_{800}, \delta h_{800,1}\right) + s_3 \quad \textbf{14}$$

In these equations, m_1, m_2, m_3, r_1, r_2, s_1, s_2, s_3 R_1, R_2, M_1, M_2, V_1, and $\delta h_{800,1}$are model parameters, $\mathrm{M_w}$ is the moment magnitude of earthquake, R_{rup} is the closest distance between fault plane and site, V_{S30} is the average shear wave velocity of the top 30 meters of soil, and δh_{800} is defined above (Equations 7 and 8).

The model parameters for source duration are estimated by nonlinear mixed effect regression and the model parameters for path duration and site effect are calculated by linear mixed effect regression. The calculation is done using the lme4 package (Bates et al. 2006) in R. The results are presented in Table 1. In addition, the standard deviation of residual components and the total standard deviation is presented in Table 2 to 4. Since observations and trial and error are used to fix R_1, R_2, M_1, M_2, V_1, and $\delta h_{800,1}$, these parameters are not associated with a standard deviation.

Table 1. Coefficients of Ds5-95 models. NS, SS, and RS stand for normal strike, strike slip, and reverse slip, respectively

	Crustal		Intraslab		Interface	
Parameter	Value	Standard deviation	Value	Standard deviation	Value	Standard deviation
m_1	0.39044	0.09053	0.46849	0.01127	0.4019741	0.01872
m_2	4.16	0.002341	5.88217	0.27388	6.3702	0.1866
m_3	1.231456(NM) 1.762263(SS) 3.185811(RS) 1.016(unknown)	0.003757 0.003757 0.003757 0.003756	5.07585(NM) 4.72137(SS) 5.1771(RS) 5.04754(unknown)	0.10964 0.13076 0.105 0.15329	6.8479	0.07792
M_1	5	-	5.5	-	5.5	-
M_2	7	-	-	-	-	-
r_1	0.16388	0.0109	-0.32702593	0.037946	-0.2159	0.06346
r_2	-	-	0.067037956	0.006365452	0.060634	0.009495
R_1	60	-	40	-	40	-
R_2	-	-	325	-	325	-
s_1	-0.28131	0.026020	-0.2527317	0.024477	-0.260	0.033253
s_2	0.000568	0.000113	0.0016555	0.000148	0.0011	0.000201
s_3	-0.09483	0.017813	0.0683161	0.015568	-0.0198	0.022146
V_1	600	-	600	-	600	-
$\delta h_{800,1}$	250	-	250	-	250	-

Table 2. Standard deviation values of Ds5-95 models

Parameters	Crustal	Intraslab	Interface
ϕ_{SS}	0.38481	0.2867625	0.28752
τ	0.3346	0.1941758	0.36169
ϕ_{S2S}	0.13558	0.1232922	0.17918
σ	0.527654	0.367611125	0.495573

Table 3. Coefficients of $Ds_{5\text{-}75}$ models. NS, SS, and RS stand for normal strike, strike slip, and reverse slip, respectively

	Crustal		Intraslab		Interface	
Parameter	Value	Standard deviation	Value	Standard deviation	Value	Standard deviation
m_1	0.30651	0.04029	0.21421	0.00552	0.19913	0.002808
m_2	4.59132	0.31472	4.19153	0.04783	3	0.042812
m_3	0.51747(NM) 1.79028(SS) 1.85878(RS) 0.96191(unknown)	0.06703 0.08287 0.07074 0.30952	1.474566(NM) 1.056999(SS) 1.68073(RS) 1.584(unknown)	0.0763 0.090337 0.071048 0.052525	0.92648	0.042812
M_1	5	-	5.5	-	5.5	-
M_2	7	-	-	-	-	-
r_1	0.06162	0.00205	-0.046076489	0.012424	-0.05289	0.012424
r_2	-	-	0.012424037	0.001688	0.016137	0.001688
R_1	60	-	40	-	40	-
R_2	-	-	325	-	325	-
s_1	-0.1894367	0.0339384	-0.1200882	0.035021	-0.1464	0.03797
s_2	0.0003362	0.0001471	0.0010966	0.000219	0.00075	0.00023
s_3	-0.0397935	0.0212835	0.4106921	0.02281	0.35697	0.0246
V_1	600	-	600	-	600	-
$\delta h_{800,1}$	250	-	250	-	250	-

Table 4. Standard deviation values of $Ds_{5\text{-}75}$ models

Parameters	Crustal	Intraslab	Interface
ϕ_{SS}	0.54408	0.4699898	0.386565
τ	0.33639	0.3727782	0.3540803
ϕ_{S2S}	0.16631	0.1989792	0.1901495
σ	0.660935	0.632017975	0.557639839

3.3.3 Aleatory variability model

Previous researchers (e.g., AS16 and DW16) have observed decrease in the ϕ_{SS} and τ with magnitude. We generate Figure 8 to understand if there is such trend in our dataset. Figure 8 shows the standard deviation of the event terms and the residuals computed for various magnitude bins (each magnitude bin is 0.5 units of magnitude). This figure shows that the standard deviation components do not change

significantly with magnitude. For this reason, we adopt a constant model for each standard deviation component.

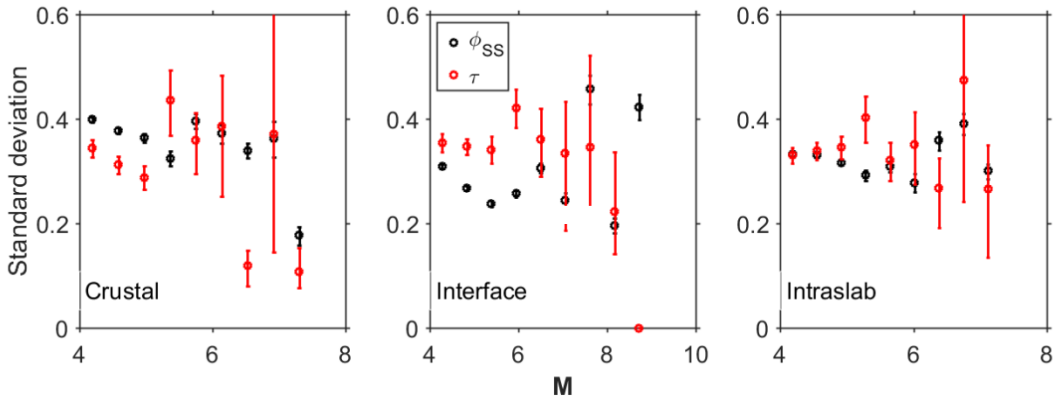


Figure 8: The standard deviation of event terms and model residuals for various magnitude bins. The standard deviations are plotted at the center of each magnitude bin. The lines shows the 5 and 95 percentile of standard deviation distribution

3.3.4 Model performance

Scatter plot for the event term, site terms, and residuals versus explanatory variables are used to check the performance of the models. Plots of the $Ds_{5\text{-}95}$ model for shallow crustal earthquakes are shown in Figure 9. The plots of the other models (Interface subduction and Intraslab subduction $Ds_{5\text{-}95}$ models and shallow crustal, interface subduction, and intraslab subduction $Ds_{5\text{-}75}$ models) are shown in Appendix A. The red circles in Figure 9 are the mean of the residuals for each bin and, the red line corresponds to plus and minus one standard deviation. When the mean is close to zero the model is unbiased for that bin. In general, the model is unbiased for all considered magnitude and distance ranges. A similar observation can be made for subduction earthquakes.

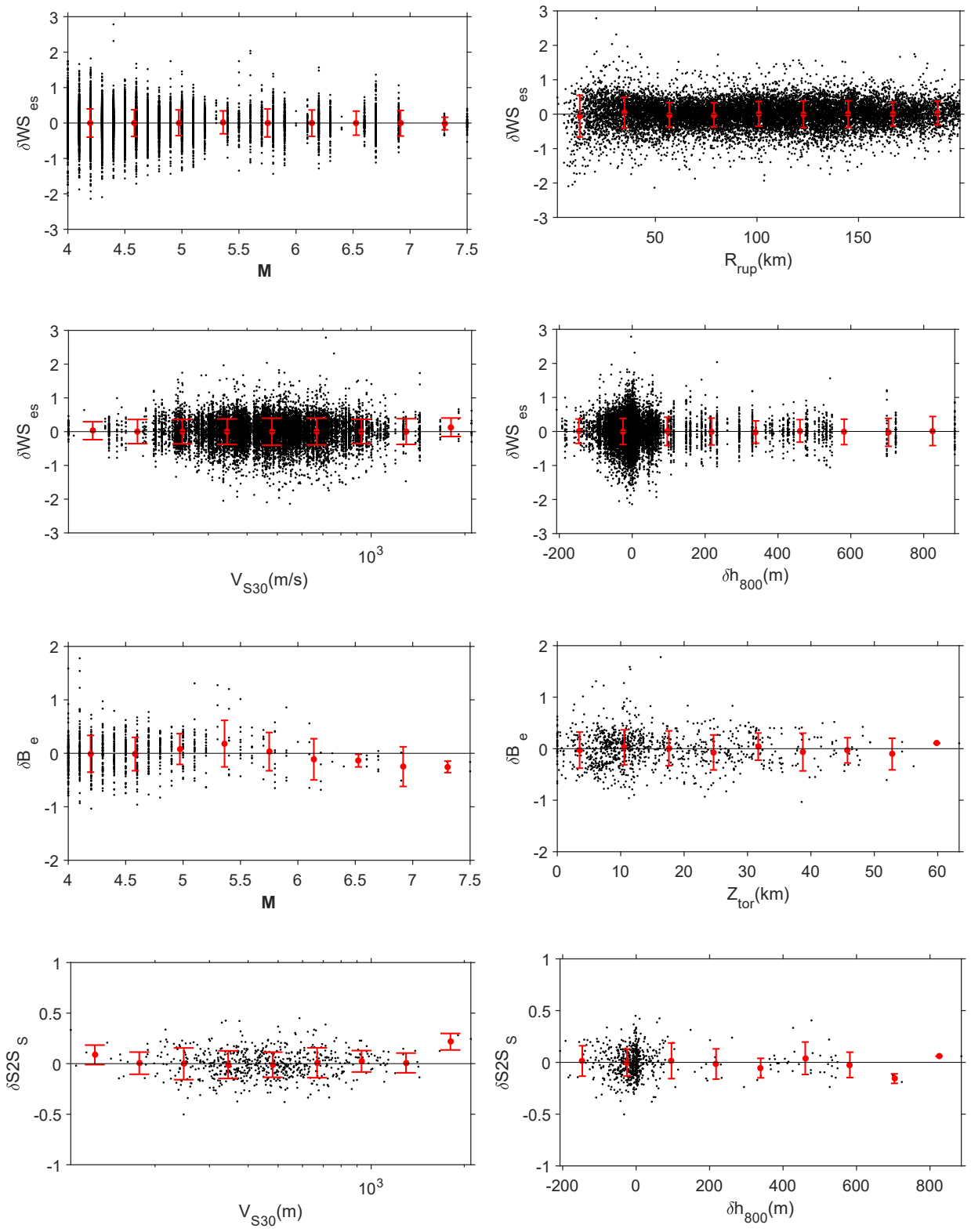


Figure 9: Residuals of model for shallow crustal earthquakes versus explanatory variables showing mean values and plus and minus standard deviation of each bin

3.4 Comparison with previous models

Figure 10 shows a comparison of the proposed relationship for duration with other published relationships. The figures in the left column compare the predictions from this study with previous ones for a range of magnitudes and the figures in the right column show a similar comparison for a range of distances. Note that all models from literature except LG14 are applicable for global shallow crustal earthquakes. The LG14 model is applicable for stable continental tectonic regimes and the database comes from central and eastern United States. The comparison of shallow crustal models implies that while the duration of small magnitude earthquakes are close to each other, the prediction of duration for large magnitude earthquakes generally are higher in Japan than other places. The comparison between the prediction for shallow crustal, interface, and intraslab models shows significant difference between shallow crustal motion duration and subduction motion. Hence, using GMPEs developed for shallow crustal earthquakes to predict subduction earthquakes duration results in significant over prediction.

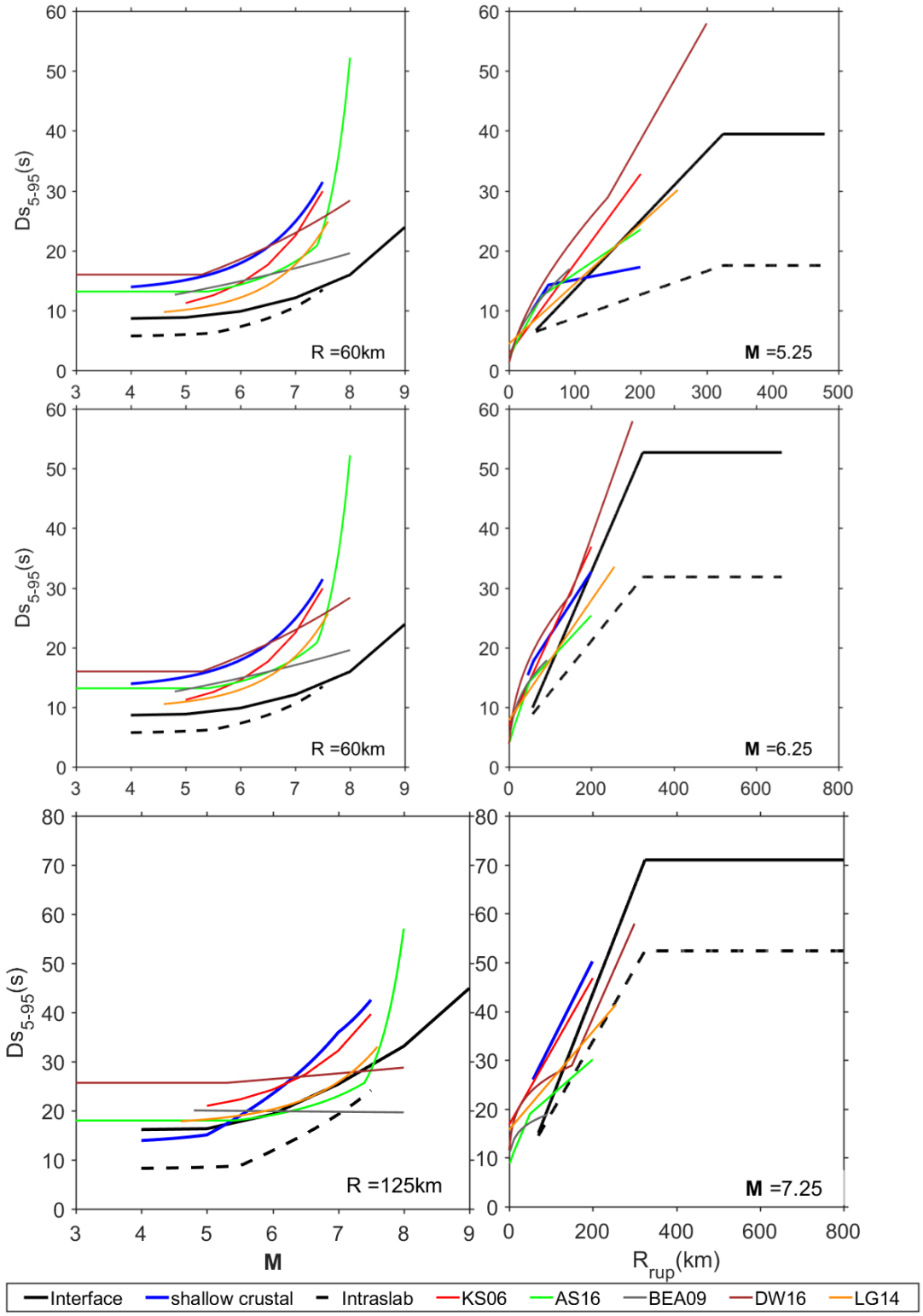


Figure 10: Comparison of predictions of different models of duration with results from this study. In all plots, V_{S30} is 760 m/s. Site class B is used for models that included site classification, and soil sites are used for models that had different models for soil and rock.

3.5 Conclusion

In this paper, we present new GMPEs for two definition of significant duration (i.e. $Ds_{5\text{-}95}$ and $Ds_{5\text{-}75}$) applicable to interface and intraslab subduction earthquakes and shallow crustal earthquakes. The explanatory variables of our models are moment magnitude **M**, rupture distance, V_{S30}, h_{800} (the depth to shear wave velocity of 800 m/s), and fault mechanism. The effects of source and path are modelled as additive terms, while site effects are modelled as multiplicative. We also observed dependence of path duration on magnitude, and this dependence is included in the proposed model. The effect of depth to top of rupture was not significant enough to be included in the model. The equations for shallow crustal earthquakes are applicable for earthquakes with magnitudes ranging from 4 to 7.5, and the subduction equations are applicable for earthquakes with magnitudes 4 to 9. The equations should not be used for sites with V_{S30} higher than 1500 m/s because the residuals are biased for this range of V_{S30}. The maximum distance for shallow crustal is set to 200 km and the maximum distance for subduction earthquake is set to 1000 km. The minimum R_{rup} for which the equations are applicable depends on the magnitude. For shallow crustal equations, the minimum R_{rup} is 10 km for earthquakes with magnitude ranging from 4 to 5, 30 km for magnitudes ranging from 5 to 6 and 50 km for magnitudes larger than 6. For subduction zone equations, the minimum R_{rup} is 30 km for earthquakes with magnitudes ranging from 4 to 5, 60 km for magnitudes 5 to 6 and 100 km for magnitudes larger than 6 .

We compare the models with previous models in the literature. While the predictions are consistent with models from the literature at small magnitudes, the predicted duration for shallow crustal earthquakes are longer than those published models for large earthquakes. This implies that the duration of motions in Japan are longer than for other geographical regions of the world. Moreover,

comparison of models developed for different tectonic regimes implies that shallow crustal earthquakes have longer durations than subduction earthquakes in general.

3.6 Reference

Abrahamson, N., and Youngs, R. (1992). "A stable algorithm for regression analyses using the random effects model." *Bulletin of the Seismological Society of America*, 82(1), 505-510.

Afshari, K., and Stewart, J. P. (2016). "Physically parameterized prediction equations for significant duration in active crustal regions." *Earthquake Spectra*, 32(4), 2057-2081.

Al Atik, L., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., and Kuehn, N. (2010). "The variability of ground-motion prediction models and its components." *Seismological Research Letters*, 81(5), 794-801.

Ancheta, T. D., Darragh, R. B., Stewart, J. P., Seyhan, E., Silva, W. J., Chiou, B. S.-J., Wooddell, K. E., Graves, R. W., Kottke, A. R., and Boore, D. M. (2014). "NGA-West2 database." *Earthquake Spectra*, 30(3), 989-1005.

Arias, A. (1970). "A measure of earthquake intensity."*R.J. Hansen, ed. Seismic Design for Nuclear Power Plant, MIT Press, Cambridge, Massachusetts,* pp. 438-483.

Atkinson, G. M. (1995). "Attenuation and source parameters of earthquakes in the Cascadia region." *Bulletin of the Seismological Society of America*, 85(5), 1327-1342.

Atkinson, G. M., and Boore, D. M. (1997). "Stochastic point-source modeling of ground motions in the Cascadia region." *Seismological Research Letters*, 68(1), 74-85.

Atkinson, G. M., and Macias, M. (2009). "Predicted ground motions for great interface earthquakes in the Cascadia subduction zone." *Bulletin of the Seismological Society of America*, 99(3), 1552-1578.

Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2006). "The lme4 package." URL http://cran r-project. org/740src/contrib/Descriptions/lme4. html.

Bommer, J. J., and Martinez-Pereira, A. (1999). "The effective duration of earthquake strong motion." *Journal of Earthquake Engineering*, 3(02), 127-172.

Bommer, J. J., Stafford, P. J., and Alarcón, J. E. (2009). "Empirical Equations for the Prediction of the Significant, Bracketed, and Uniform Duration of Earthquake Ground Motion." *Bulletin of the Seismological Society of America*, 99(6), 3217-3233.

Bommer, J. J., Stafford, P. J., Alarcón, J. E., and Akkar, S. (2007). "The influence of magnitude range on empirical ground-motion prediction." *Bulletin of the Seismological Society of America*, 97(6), 2152-2170.

Boore, D. M. (2003). "Simulation of ground motion using the stochastic method." *Pure and applied geophysics*, 160(3-4), 635-676.

Boore, D. M., and Thompson, E. M. (2012). "Empirical improvements for estimating earthquake response spectra with random-vibration theory." *Bulletin of the Seismological Society of America*, 102(2), 761-772.

Boore, D. M., and Thompson, E. M. (2014). "Path durations for use in the stochastic-method simulation of ground motions." *Bulletin of the Seismological Society of America*, 104, 2541–2552

Chai, Y. (2005). "Incorporating low-cycle fatigue model into duration-dependent inelastic design spectra." *Earthquake Engineering & Etructural Dynamics*, 34(1), 83-96.

Chiou, B., Darragh, R., Gregor, N., and Silva, W. (2008). "NGA project strong-motion database." *Earthquake Spectra*, 24(1), 23-44.

Dawood, H. M., Rodriguez-Marek, A., Bayless, J., Goulet, C., and Thompson, E. (2016). "A Flatfile for the KiK-net Database Processed Using an Automated Protocol." *Earthquake Spectra*, 32(2), 1281-1302.

Du, W., and Wang, G. (2016). "Prediction Equations for Ground-Motion Significant Durations Using the NGA-West2 Database." *Bulletin of the Seismological Society of America*.

Garcia, D., Wald, D. J., and Hearne, M. (2012). "A global earthquake discrimination scheme to optimize ground-motion prediction equation selection." *Bulletin of the Seismological Society of America*, 102(1), 185-203.

Ghofrani, H., Atkinson, G. M., Goda, K., and Assatourians, K. (2013). "Stochastic finite-fault simulations of the 2011 Tohoku, Japan, earthquake." *Bulletin of the Seismological Society of America*, 103(2B), 1307-1320.

Green, R., Mitchell, J., and Polito, C. "An energy-based excess pore pressure generation model for cohesionless soils." Proc., John Booker Memorial Symp.—Developments in Theoretical Geomechanics, D. W. Smith and J. P. Carter, eds., Balkema, Rotterdam, The Netherlands, 383–390.
Gregor, N. J., Silva, W. J., Wong, I. G., and Youngs, R. R. (2002). "Ground-motion attenuation relationships for Cascadia subduction zone megathrust earthquakes based on a stochastic finite-fault model." *Bulletin of the Seismological Society of America*, 92(5), 1923-1932.

Hancock, J., and Bommer, J. J. (2006). "A state-of-knowledge review of the influence of strong-motion duration on structural damage." *Earthquake Spectra*, 22(3), 827-845.

Ishac, M. F., and Heidebrecht, A. C. (1982). "Energy dissipation and seismic liquefaction in sands." *Earthquake Engineering & Structural Dynamics*, 10(1), 59-68.

Kunnath, S., and Chai, Y. (2004). "Cumulative damage-based inelastic cyclic demand spectrum." *Earthquake engineering & Structural Dynamics*, 33(4), 499-520.

Law, K. T., Cao, Y., and He, G. (1990). "An energy approach for assessing seismic liquefaction potential." *Canadian Geotechnical Journal*, 27(3), 320-329.

Lee, J., and Green, R. A. (2014). "An empirical significant duration relationship for stable continental regions." *Bulletin of Earthquake Engineering*, 12(1), 217-235.

Polito, C. P., Green, R. A., and Lee, J. (2008). "Pore pressure generation models for sands and silty soils subjected to cyclic loading." *Journal of Geotechnical and Geoenvironmental Engineering*, 134(10), 1490-1500.

Stafford, P. J. (2014). "Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models." *Bulletin of the Seismological Society of America*, 104(2), 702-719.

Uang, C.-M., and Bertero, V. V. (1988). *"*Use of energy as a design criterion in earthquake-resistant design*"*, *Earthquake Engineering Research Center, University of California Berkeley, USA.*

Yaghmaei-Sabegh, S., Shoghian, Z., and Neaz Sheikh, M. (2014). "A new model for the prediction of earthquake ground-motion duration in Iran." *Natural hazards*, 70(1), 1-24.

Youd, T., Idriss, I., Andrus, R. D., Arango, I., Castro, G., Christian, J. T., Dobry, R., Finn, W. L., Harder Jr, L. F., and Hynes, M. E. (2001). "Liquefaction resistance of soils: summary report from the

1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils." *Journal of geotechnical and geoenvironmental engineering*, 127(10), 817-833.

Chapter 4
Conclusion

4.1 Summary

I have developed GMPEs for two non-spectral parameters (Arias Intensity, I_a, and Significant Duratino,0 Ds_{5-95}) using on the Japanese Kik-net database. The input parameters for the models are moment magnitude ($\boldsymbol{M}$), rupture distance (R_{rup}), average shear wave velocity over the upper thirty meters (V_{S30}), and a binary (1 or 0) variable that describes whether the site-to-source path crosses the volcanic belt. The effect of tectonic regime is considered by developing separate models for shallow crustal and subduction earthquakes. The equations for shallow crustal earthquakes are applicable for earthquakes with magnitudes ranging from 4 to 7.5, and the subduction equations are applicable for earthquakes with magnitudes 4 to 9. The equations should not be used for sites with V_{S30} higher than 1500 m/s because the residuals are biased for this range of V_{S30}. The maximum distance for shallow crustal is set to 200 km, and the maximum distance for subduction earthquake is set to 1000 km. The minimum R_{rup} for which the equations are applicable depends on the magnitude. For shallow crustal equations, the minimum R_{rup} is 10 km for earthquakes with magnitude ranging from 4 to 5, 30 km for magnitudes ranging from 5 to 6, and 50 km for magnitudes larger than 6. For subduction zone equations, the minimum R_{rup} is 30 km for earthquakes with magnitudes ranging from 4 to 5, 60 km for magnitudes 5 to 6, and 100 km for magnitudes larger than 6 .

4.2 Findings

We observed that the amplification of I_a is magnitude dependent. This observation cannot be explained by the nonlinearity of site response because for soft soil sites the amplification decreases with

decreasing magnitude. This behavior can be explained by considering the differences in frequency content between earthquakes with different magnitudes. Moreover, when clustering data into magnitude bins, we observe different slopes of duration versus R_{rup} for different magnitude bins. This behavior was captured by using a magnitude dependent path effect for duration equations. Based on our data, the increase in duration with distance saturates by distance especially for small shallow crustal earthquakes. Finally, we define a dummy variable, V, to consider the effect of volcanic belt. This variable is equal to one when the source to site path has crossed volcanic belt and zero otherwise. While the effect of V on the Arias Intensity earthquakes was clear, it was not statistically significant in the duration model. This observation implies the volcanic belt of Japan does not affect duration of earthquakes significantly.

When comparing shallow crustal and subduction earthquakes, I observe that the magnitude and distance scaling for I_a are similar for both types of earthquakes. However, this is not the case for duration where the predictions for the two types of earthquakes differ significantly.

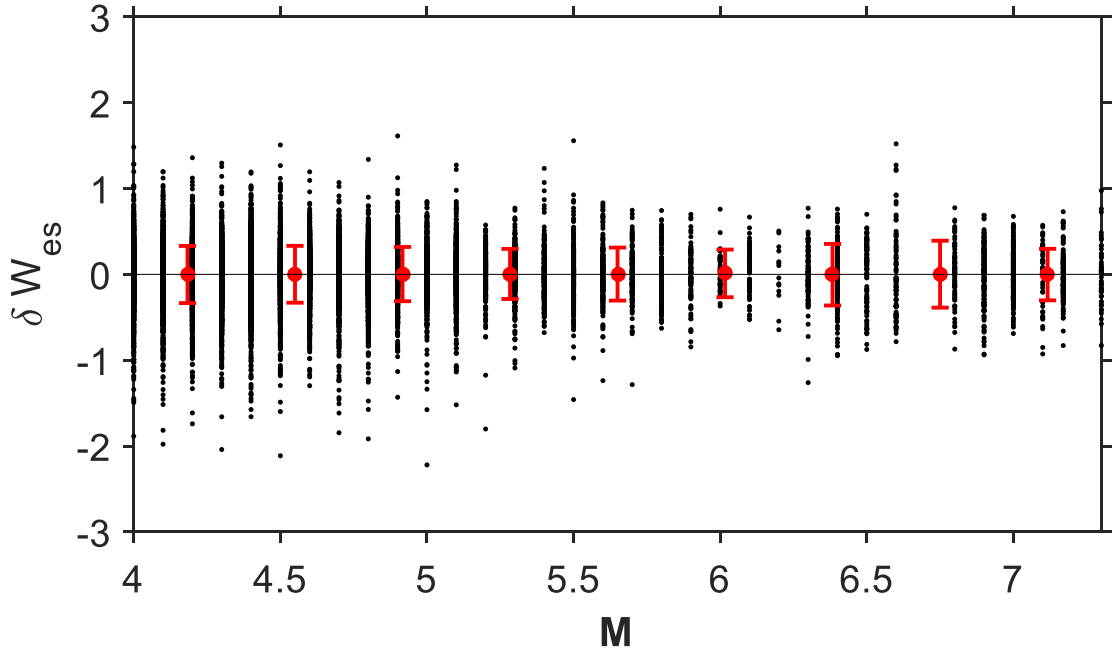
The predictions of this study for Ds_{5-95} and I_a are compared with other regions of the world. The comparisons make use of Sammon's maps to facilitate the comparison for multiple magnitudes and distances. The comparisons shows that I_a is highly region-dependent. The I_a rate of decay in Japan is higher than other regions of the world; GMPEs that are developed based on Japanese data plot closer to each other in Sammon's maps. The comparisons also show Ds_{5-95} of small earthquakes are close to each other for Japan and for the western United States since, our database comes from Japan and the other models were developed form global earthquake motions or from California earthquakes . The prediction of Ds_{5-95} of large earthquakes differs from this study and other studies, which implies the duration of large earthquakes are regional dependent. Moreover, shallow crustal motions caused by large earthquakes have longer durations than subduction earthquakes based on our GMPEs. This

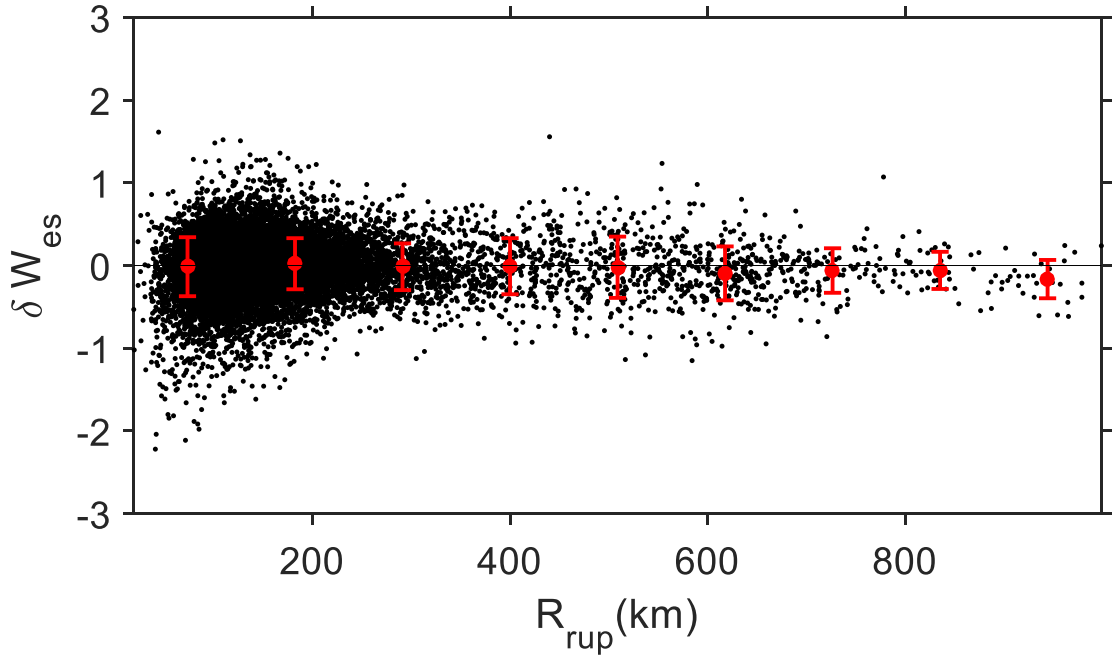
observation implies that using GMPEs developed for shallow crustal earthquakes for subduction earthquakes would result in considerable overestimation of ground motion duration.

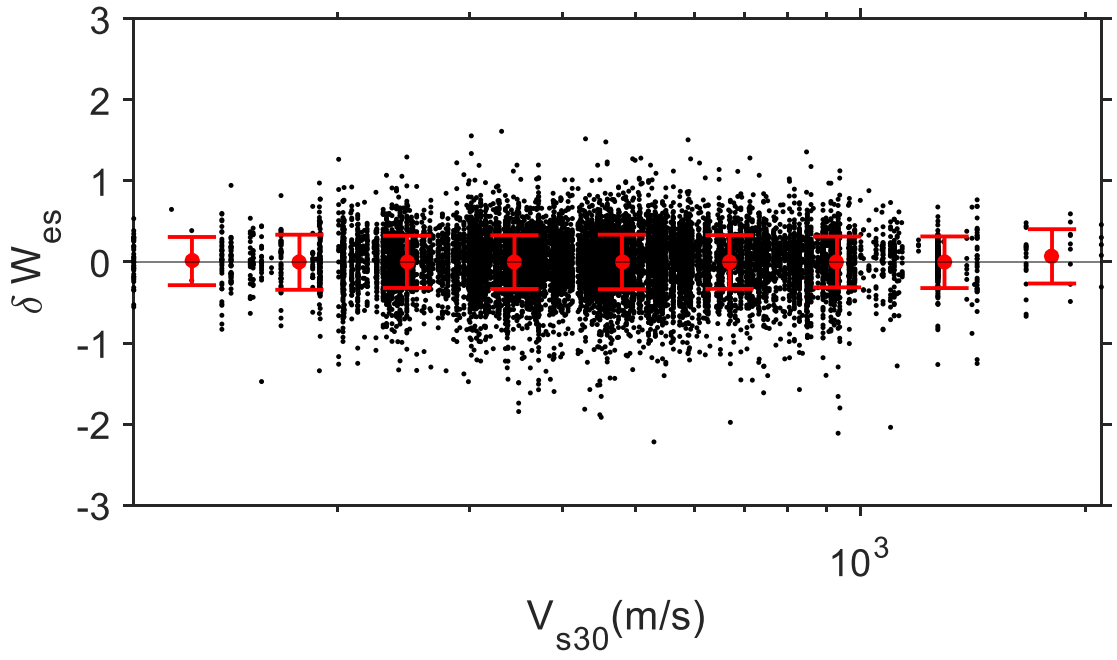
Appendix A

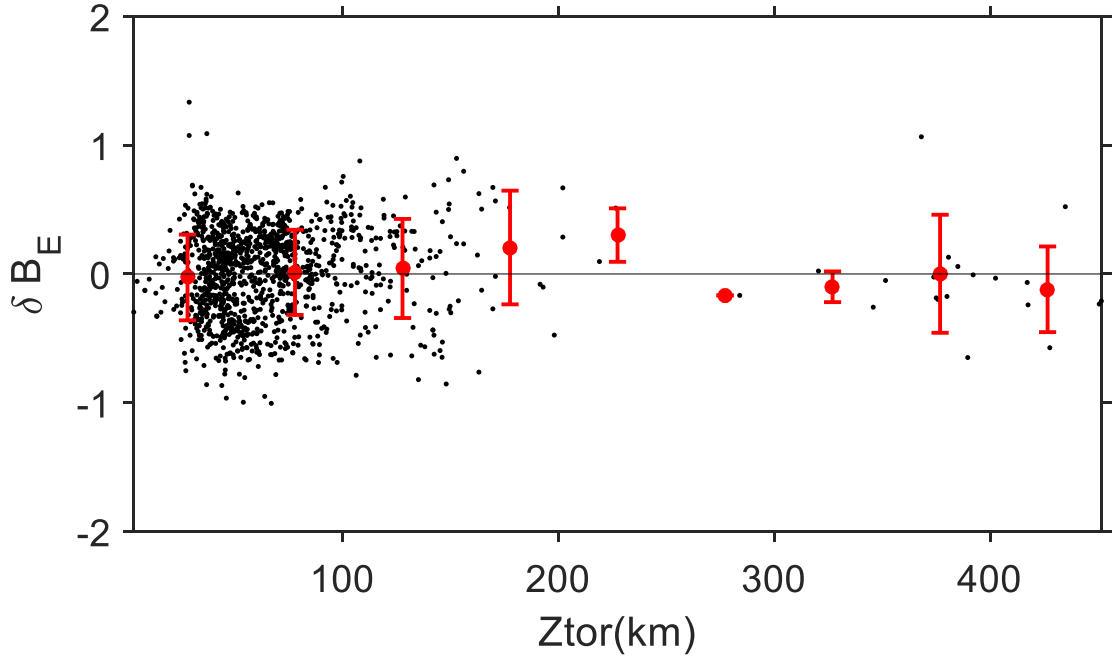
Residual plots of Interface subduction and intraslab subduction $Ds_{5\text{-}95}$ models and residual plots of shallow crustal and Interface subduction and intraslab subduction $Ds_{5\text{-}75}$ models

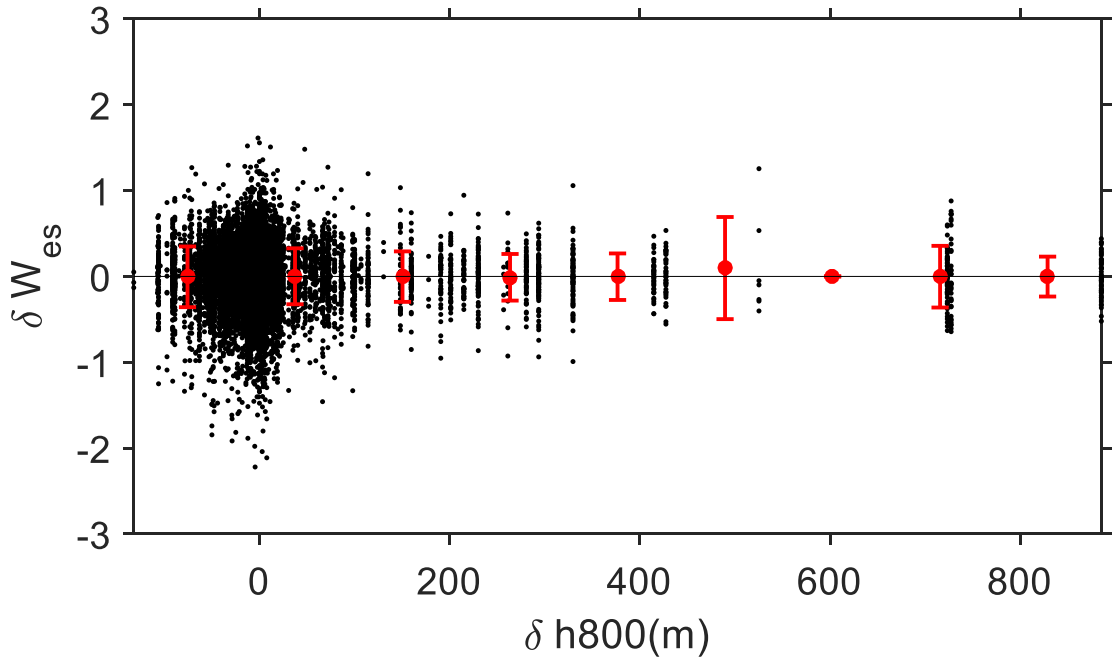
Intraslab subduction $Ds_{5\text{-}95}$ model

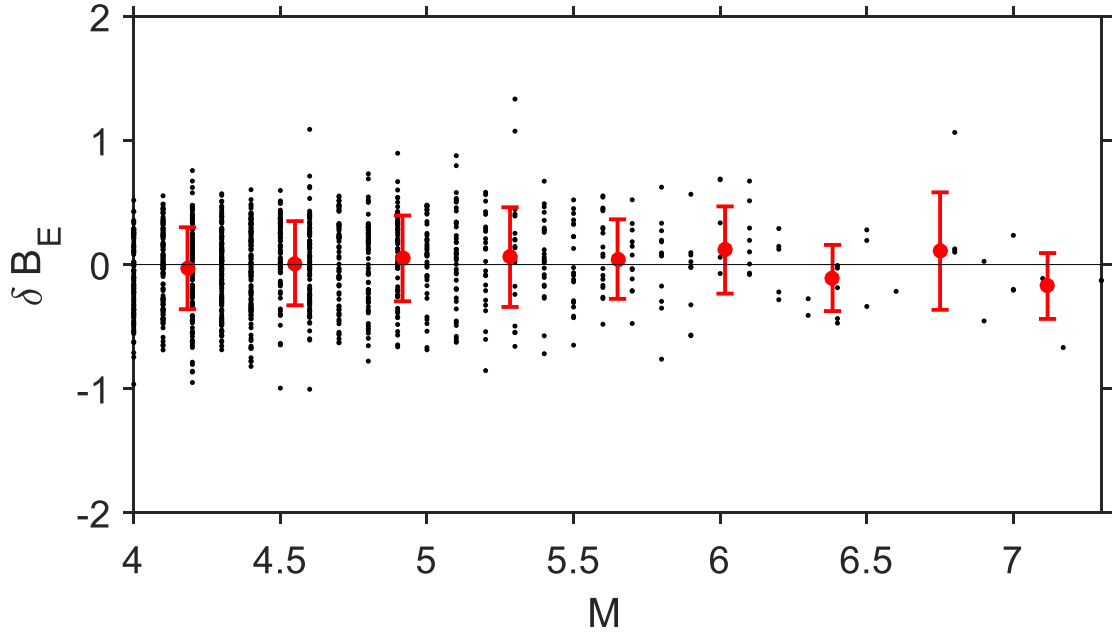


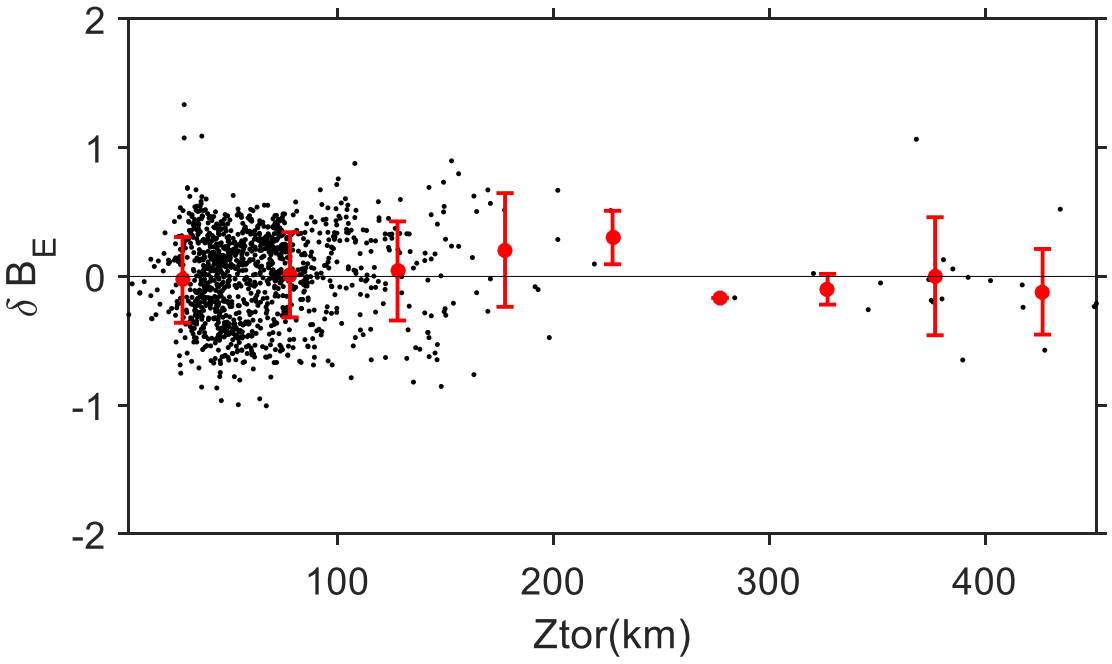


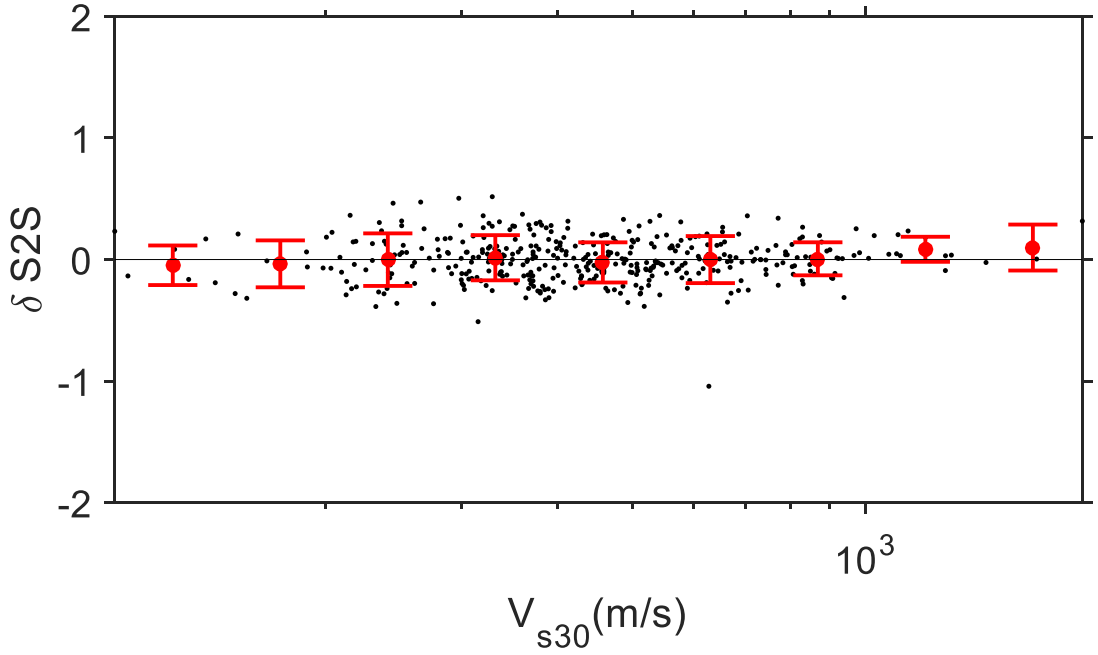

3
2
1
0
-1
-2
-3
δW_{es}
10^3
V_{s30}(m/s)

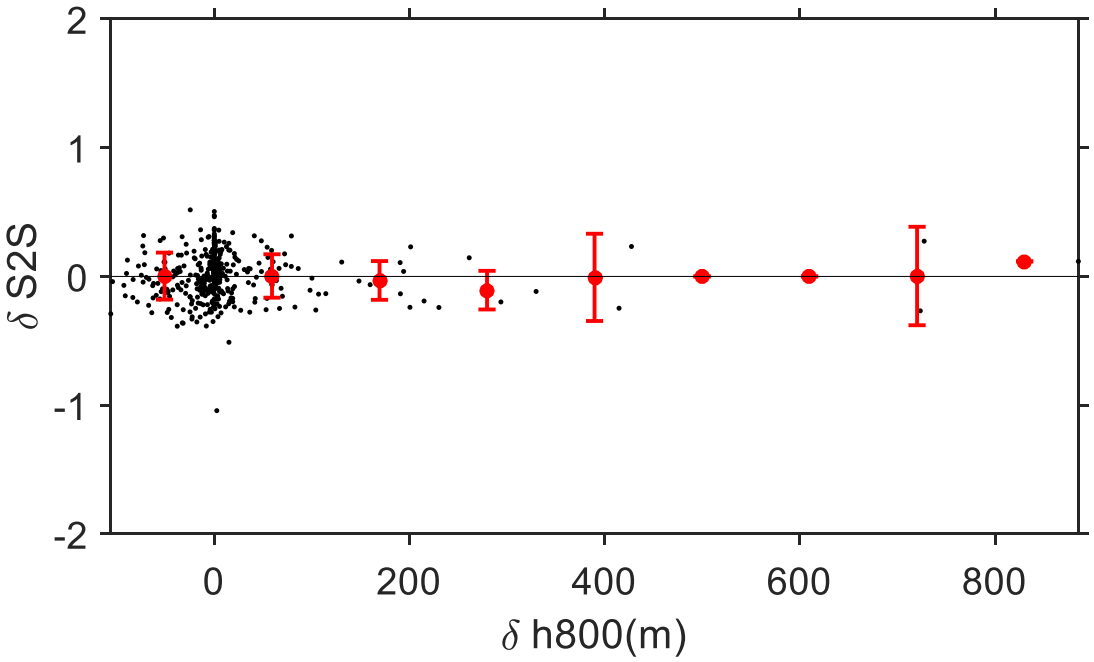

2
1
0
-1
-2
δB_E
100
200
300
400
Ztor(km)


3
2
1
0
-1
-2
-3
δ W_es
0
200
400
600
800
δ h800(m)

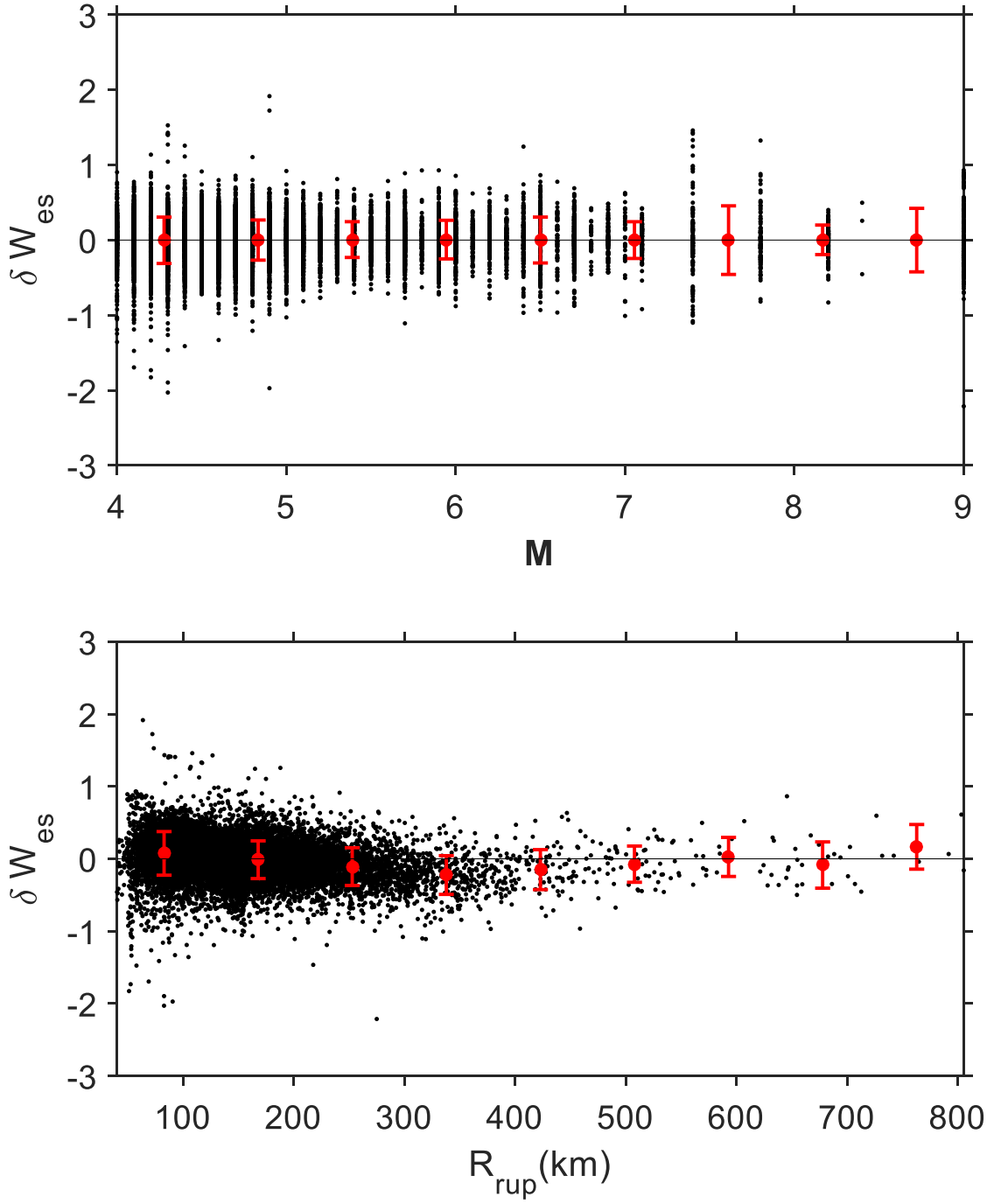

2
1
0
-1
-2
δ B_E
4
4.5
5
5.5
6
6.5
7
M

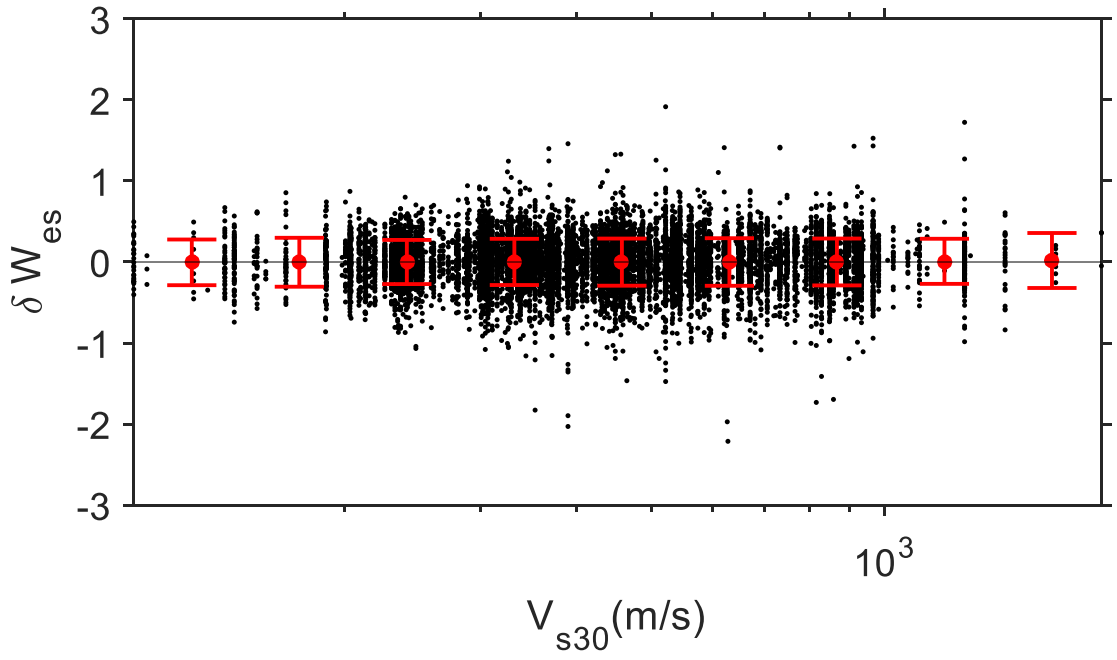

2
1
0
-1
-2
δB_E
100
200
300
400
Ztor(km)

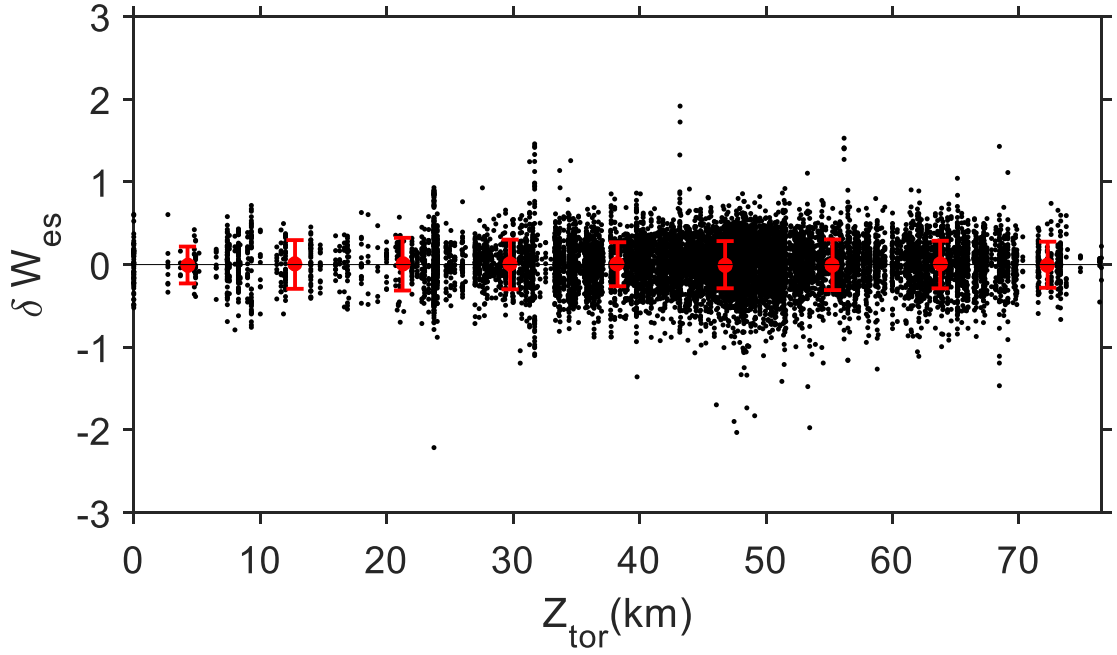

2
1
0
-1
-2
δ S2S
10^3
V_{s30}(m/s)

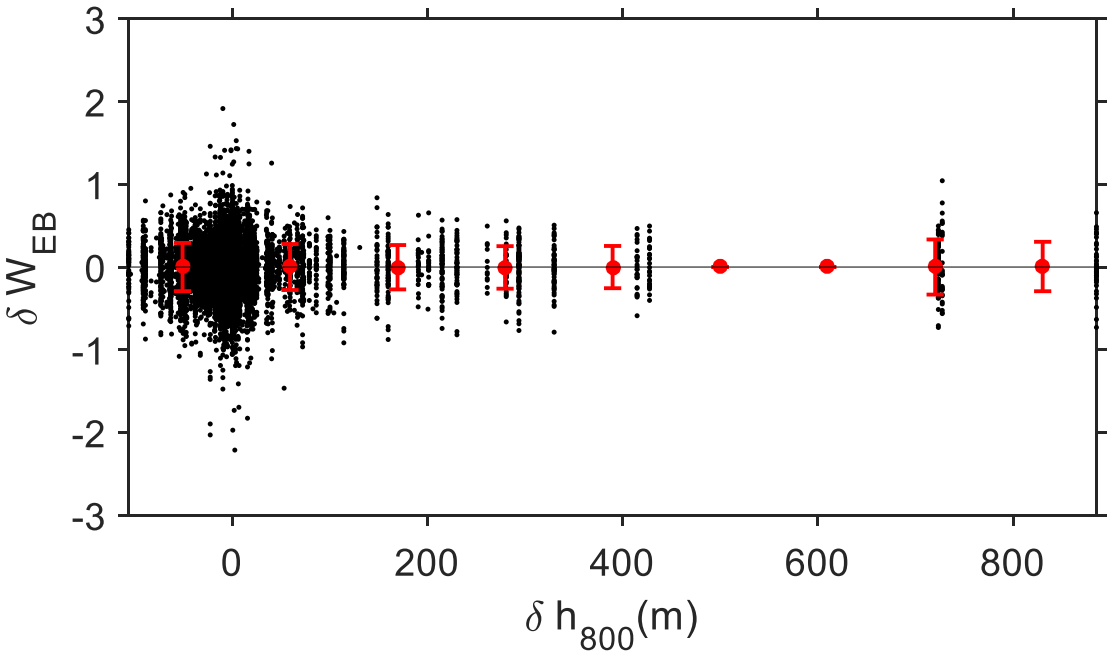

2
1
0
-1
-2
δ S2S
0
200
400
600
800
δ h800(m)

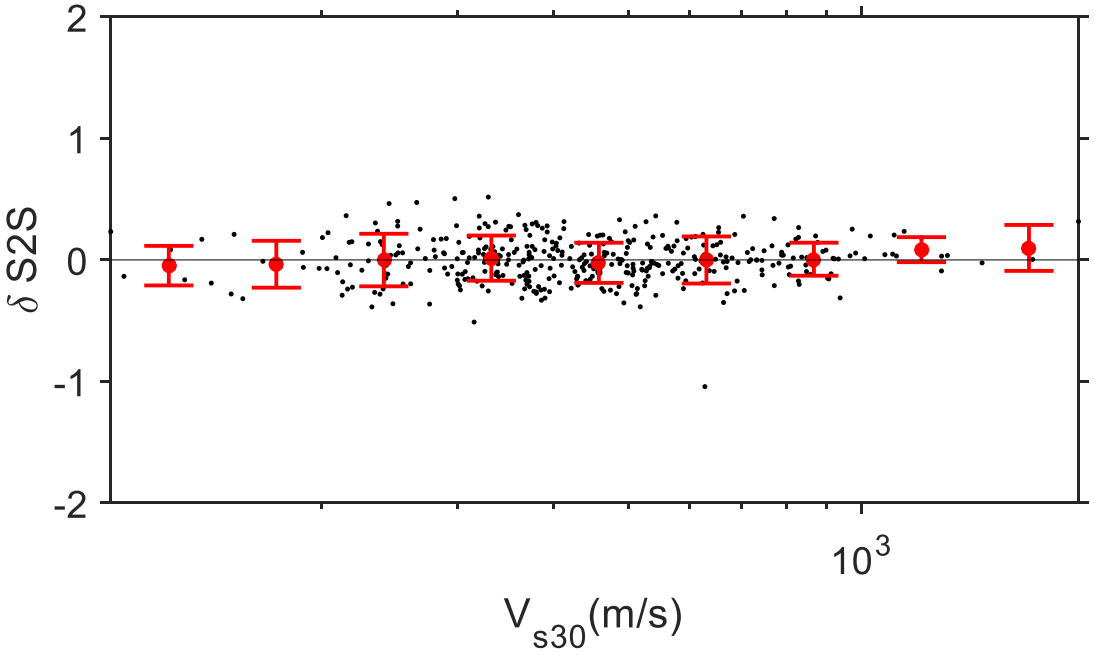
Interface subduction $Ds_{5\text{-}95}$ model

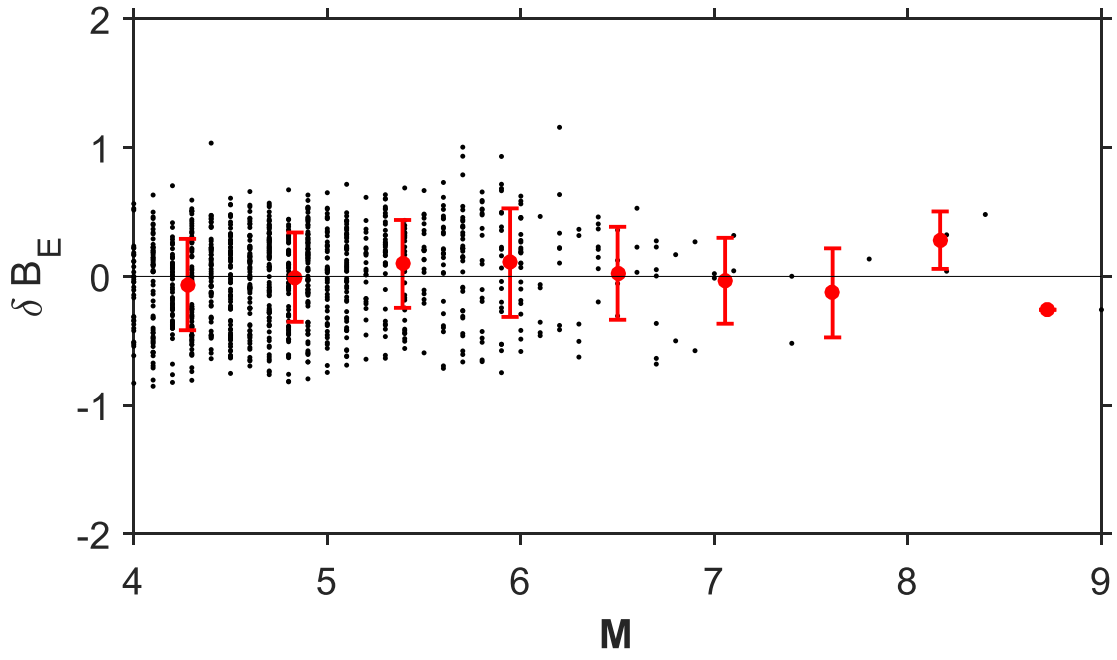


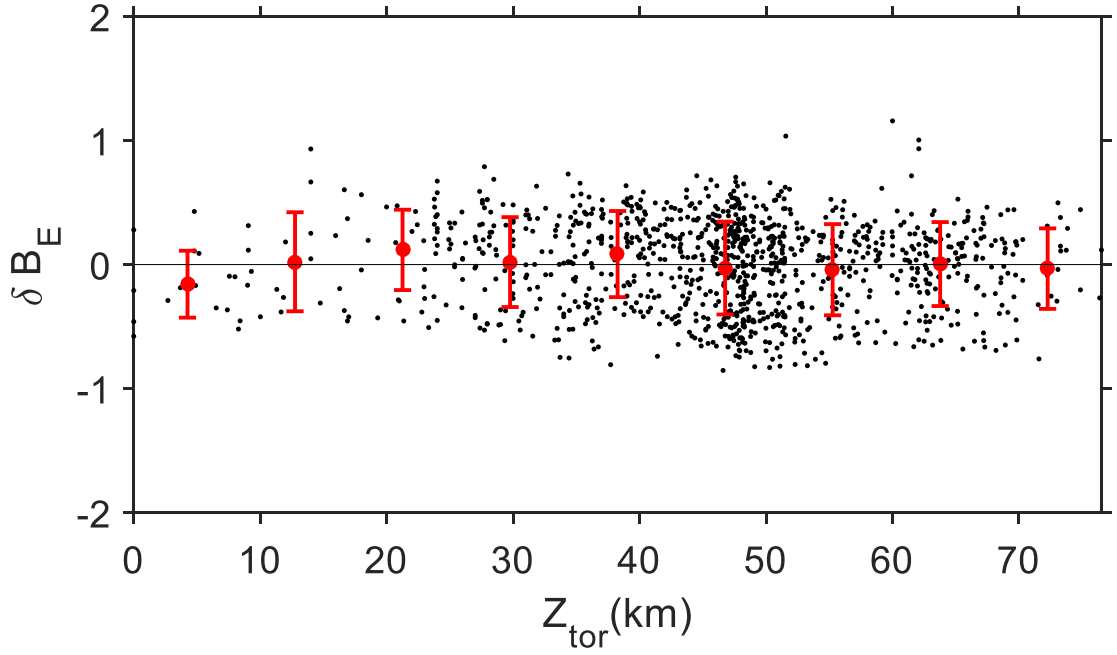

3
2
1
0
-1
-2
-3
δW_{es}
10^3
V_{s30}(m/s)

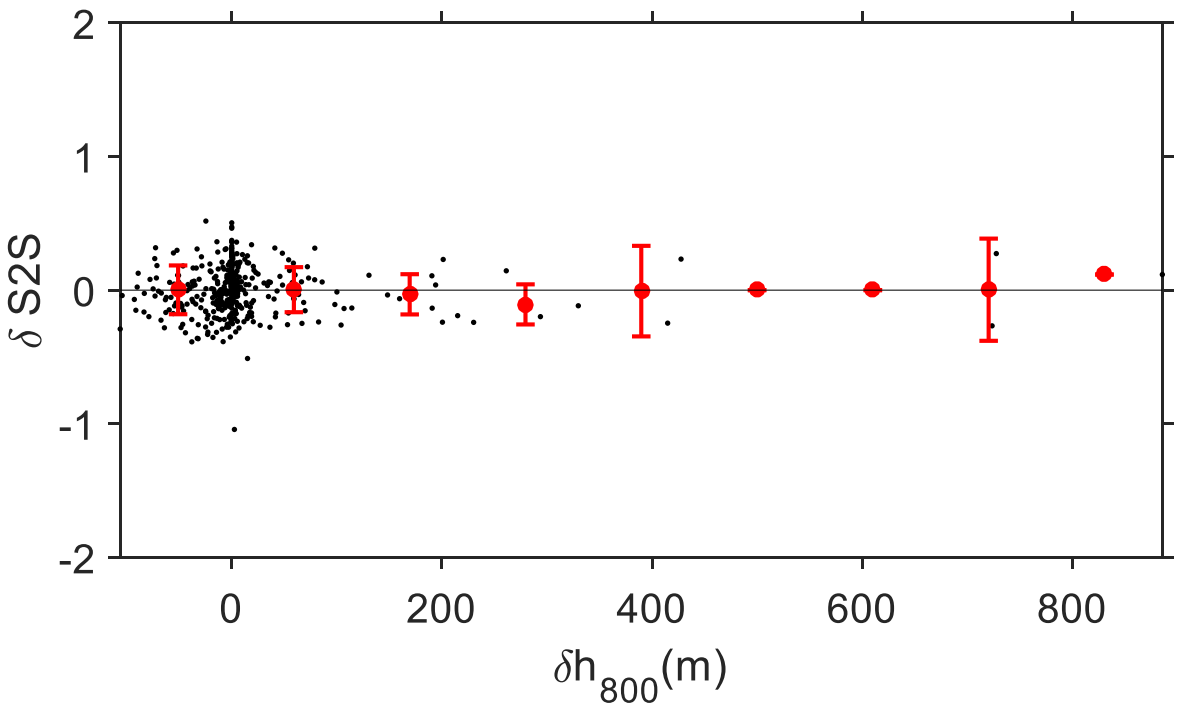

3
2
1
0
-1
-2
-3
δW_{es}
0
10
20
30
40
50
60
70
Z_{tor}(km)


3
2
1
0
-1
-2
-3
δW_{EB}
0
200
400
600
800
δh_{800}(m)

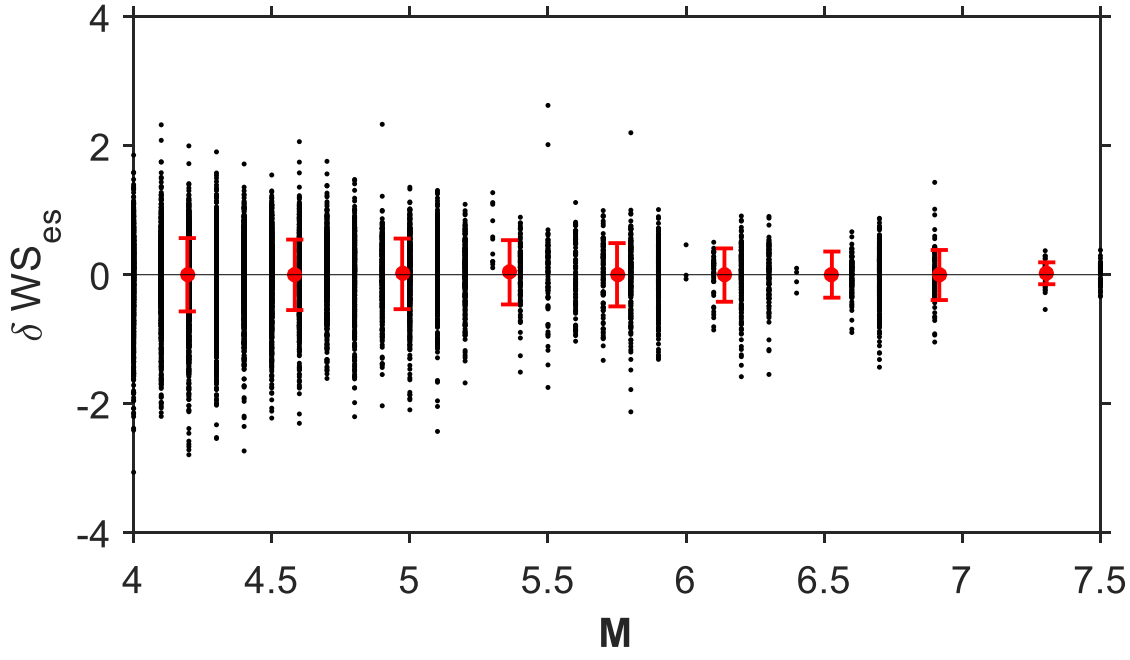

2
1
0
-1
-2
δ S2S
10^3
V_{s30}(m/s)

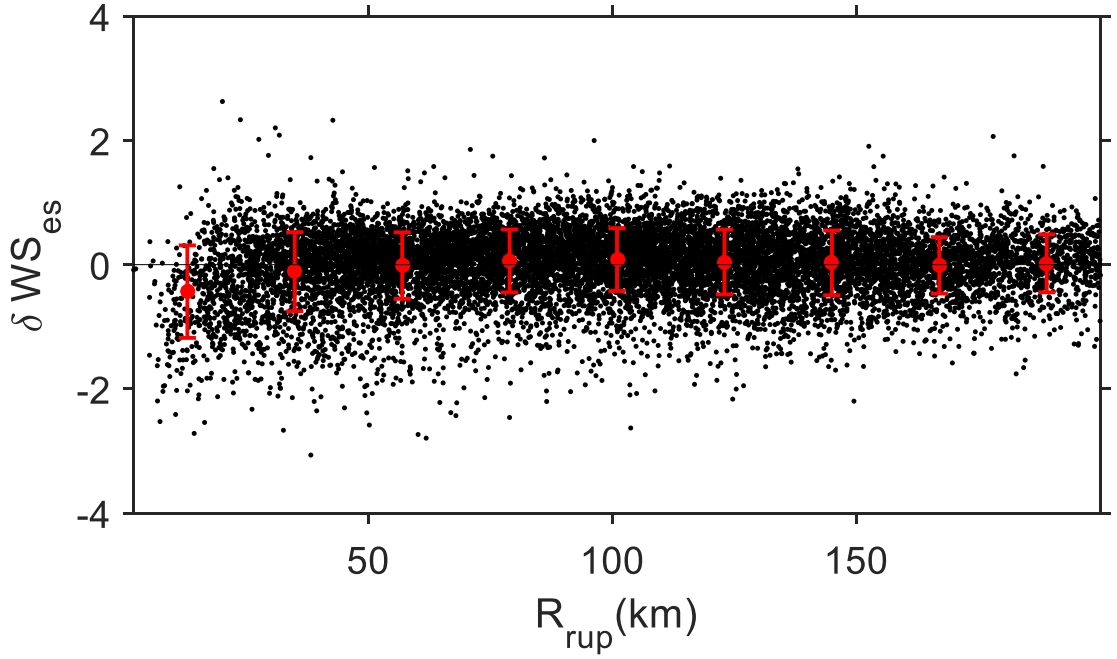

2
1
0
-1
-2
δB_E
4
5
6
7
8
9
M

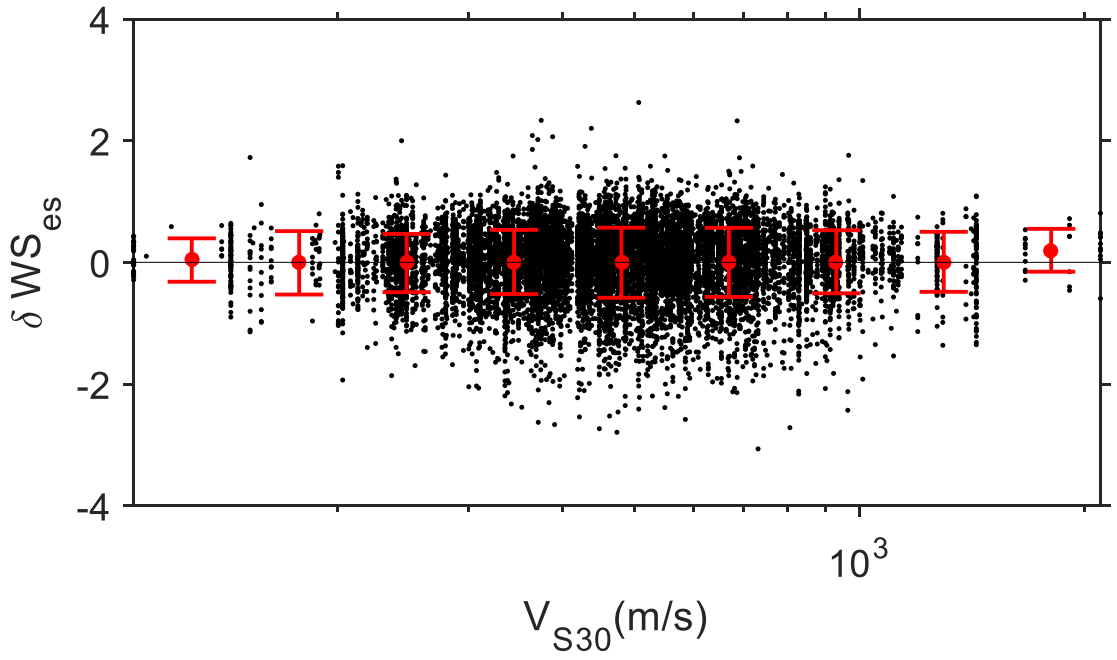

2
1
0
-1
-2
δB_E
0
10
20
30
40
50
60
70
Z_{tor}(km)

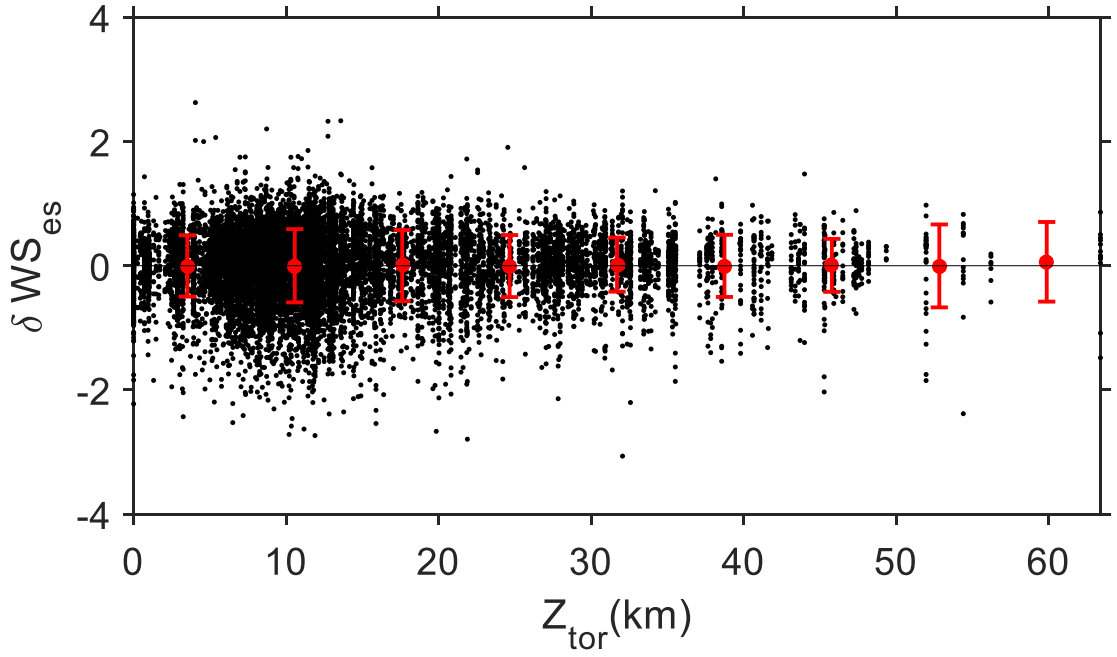

2
1
0
-1
-2
δ S2S
0
200
400
600
800
δh800(m)

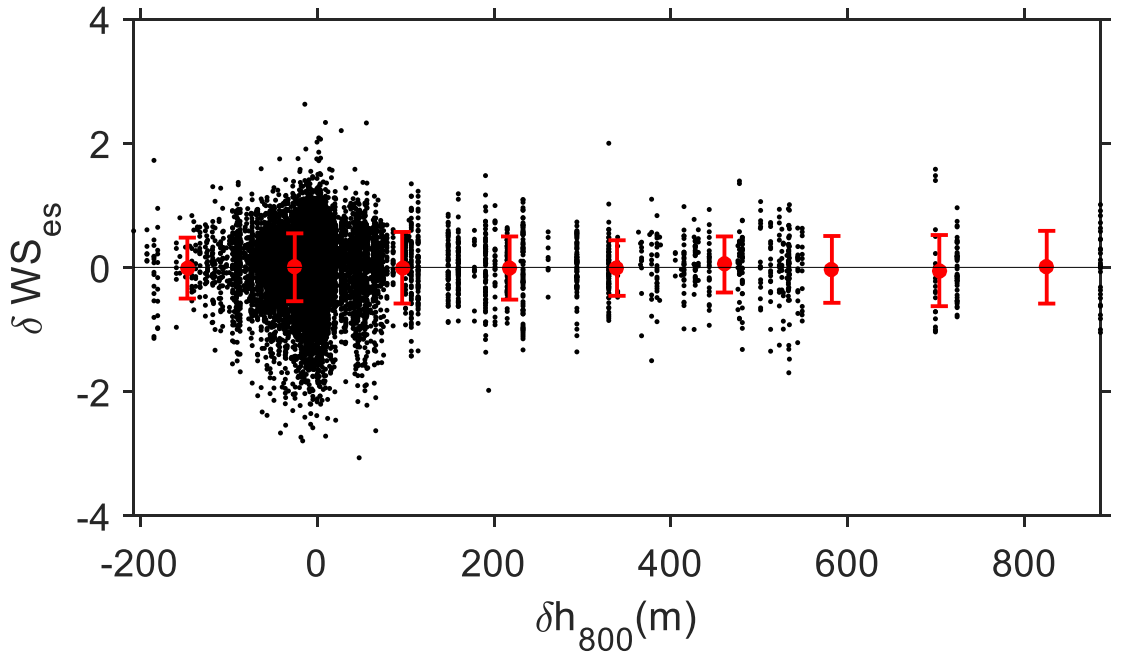
Shallow crustal $Ds_{5\text{-}75}$ model

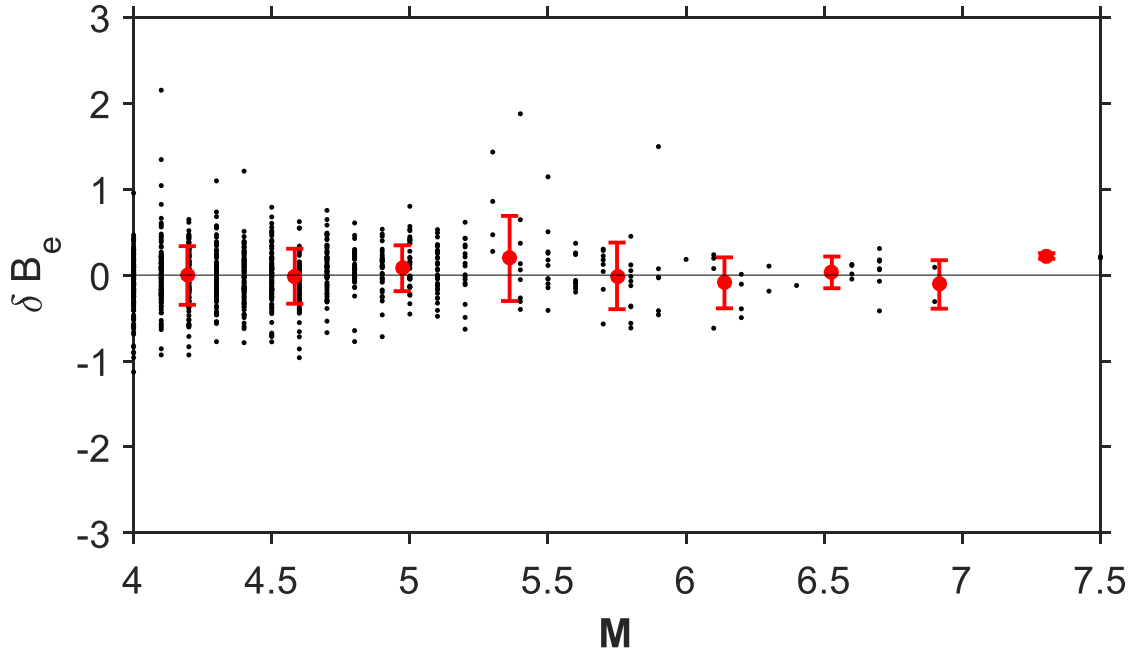


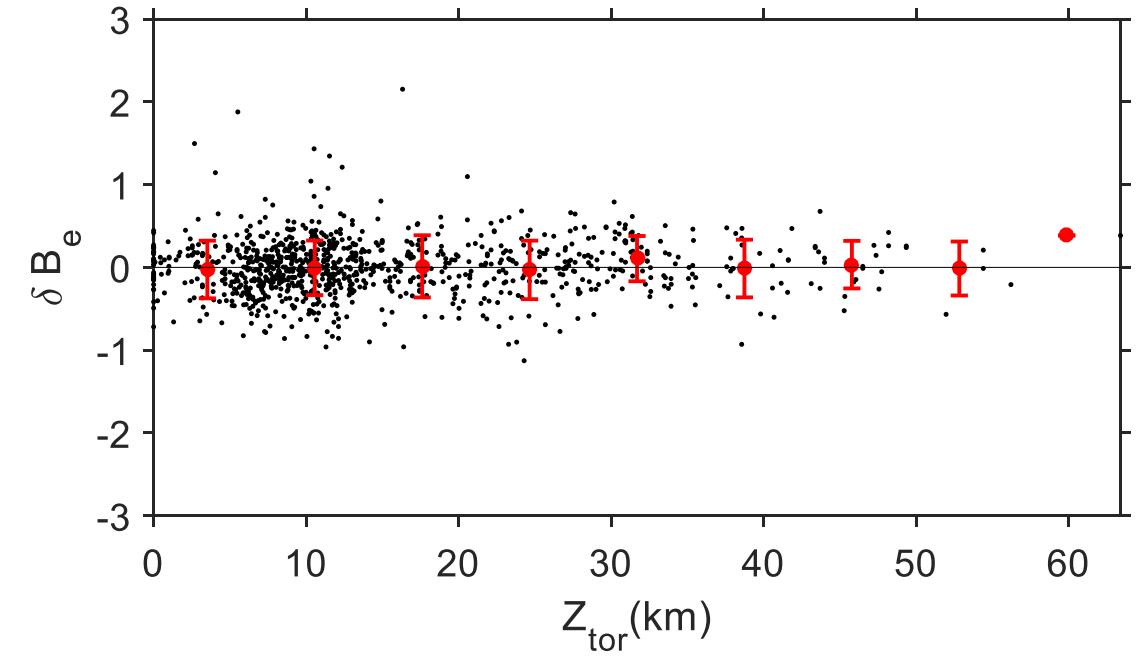


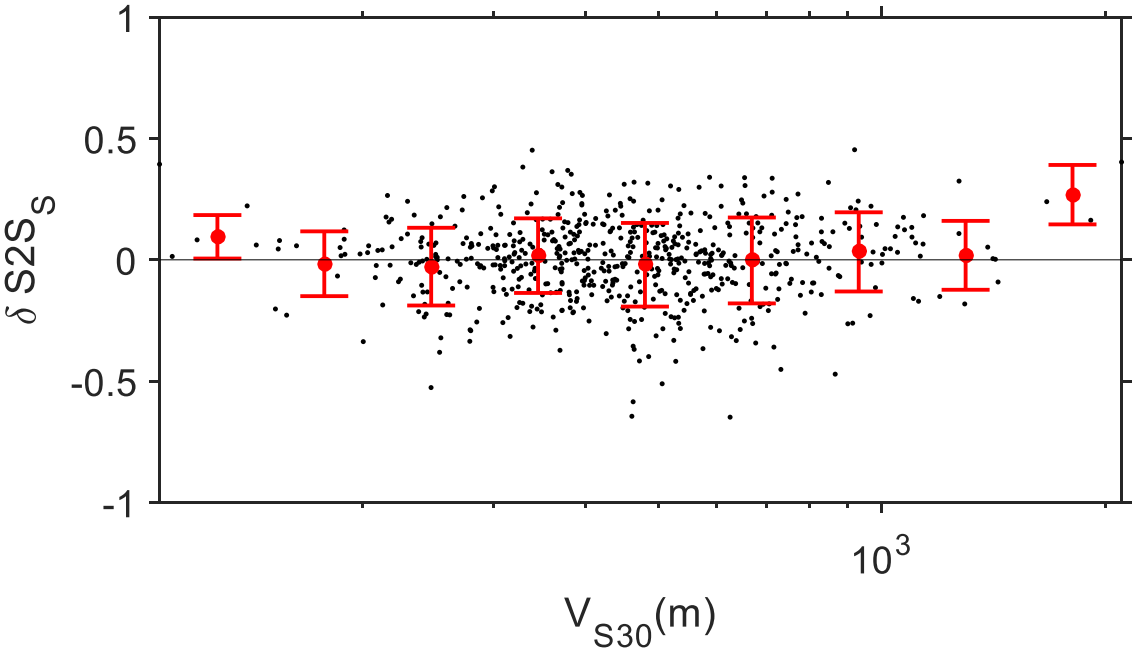

4
2
0
-2
-4
δWS_{es}
10^3
V_{S30}(m/s)

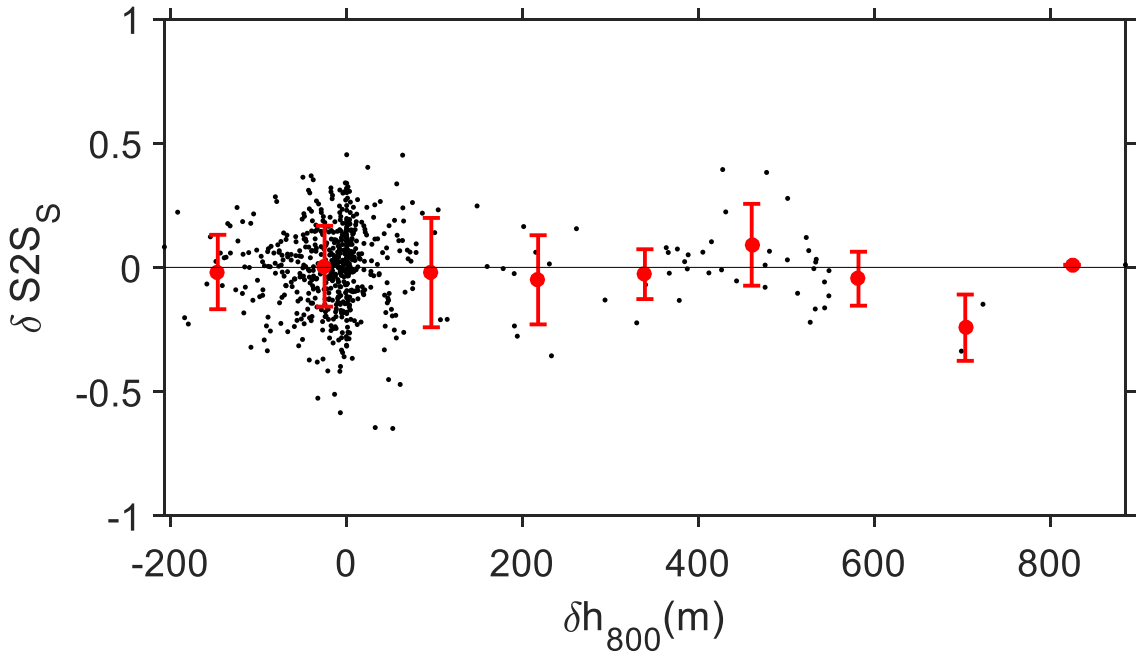

4
2
0
-2
-4
δWS_{es}
0
10
20
30
40
50
60
Z_{tor}(km)


4
2
0
-2
-4
-200
0
200
400
600
800
δWS_{es}
δh_{800}(m)

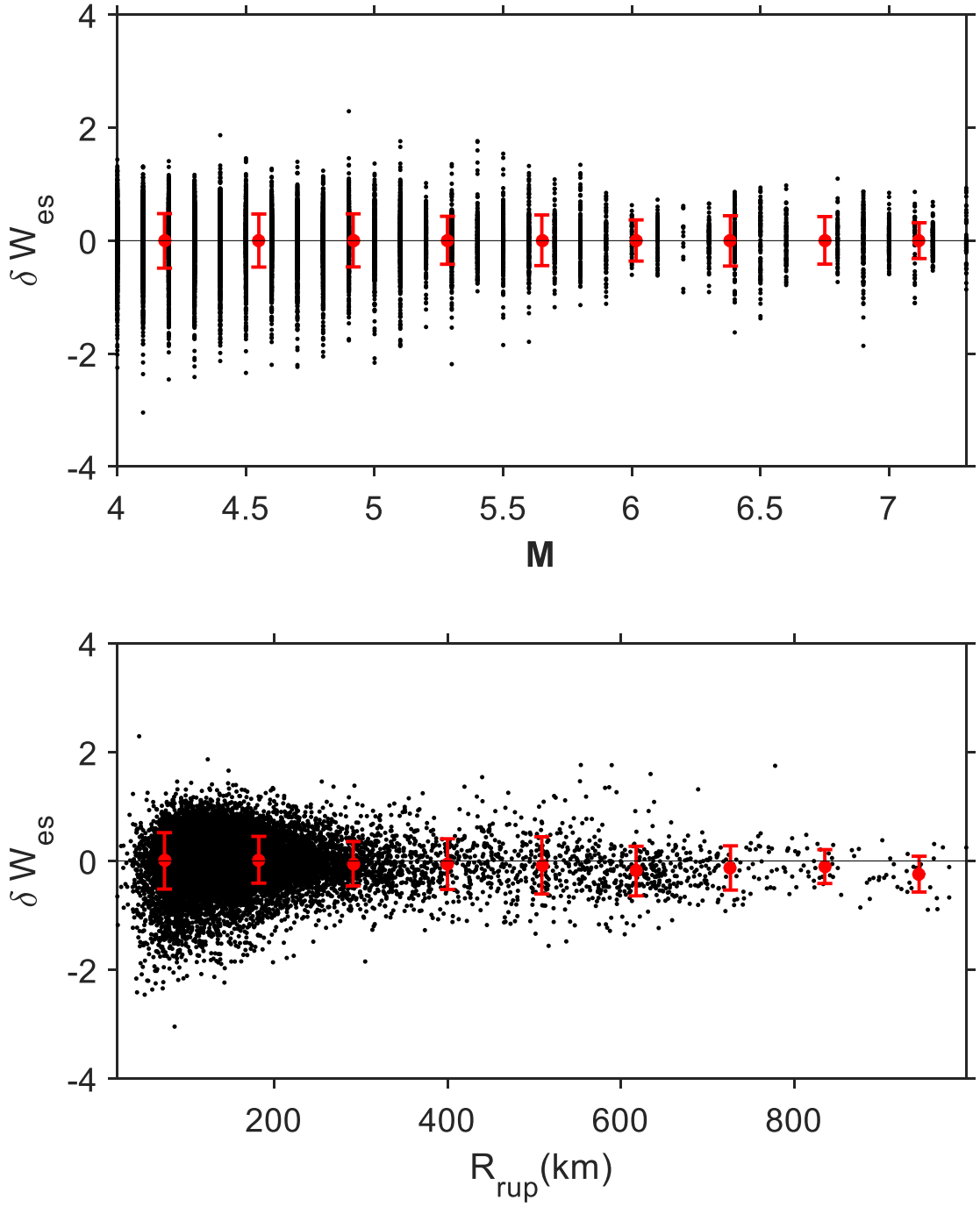

3
2
1
0
-1
-2
-3
4
4.5
5
5.5
6
6.5
7
7.5
δB_e
M

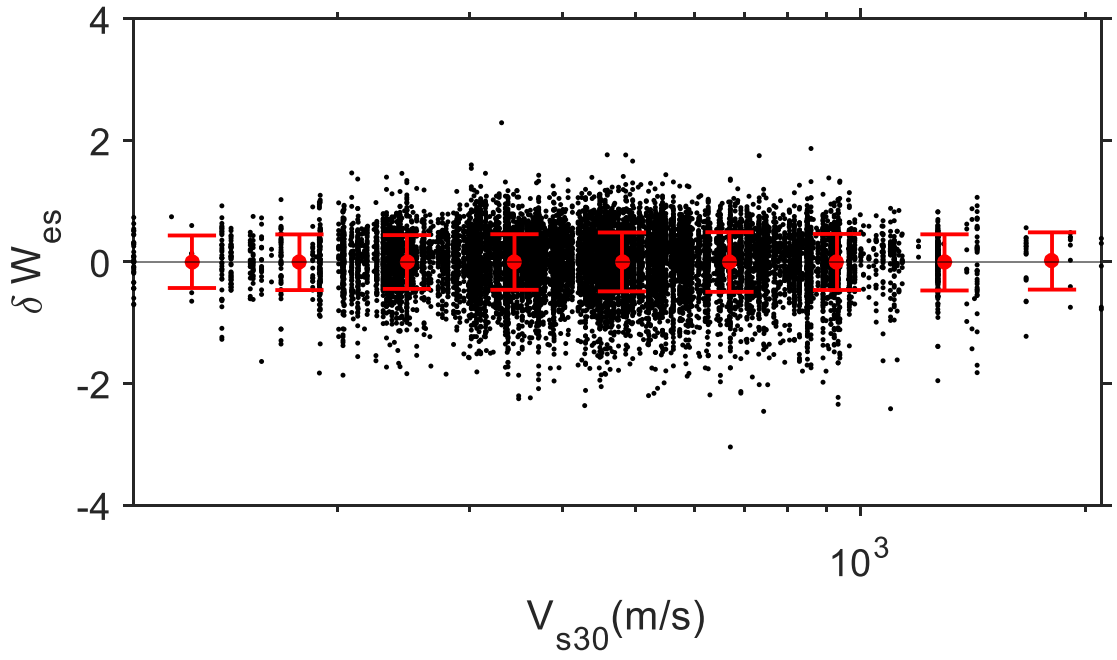

3
2
1
0
-1
-2
-3
0
10
20
30
40
50
60
δB_e
Z_{tor}(km)

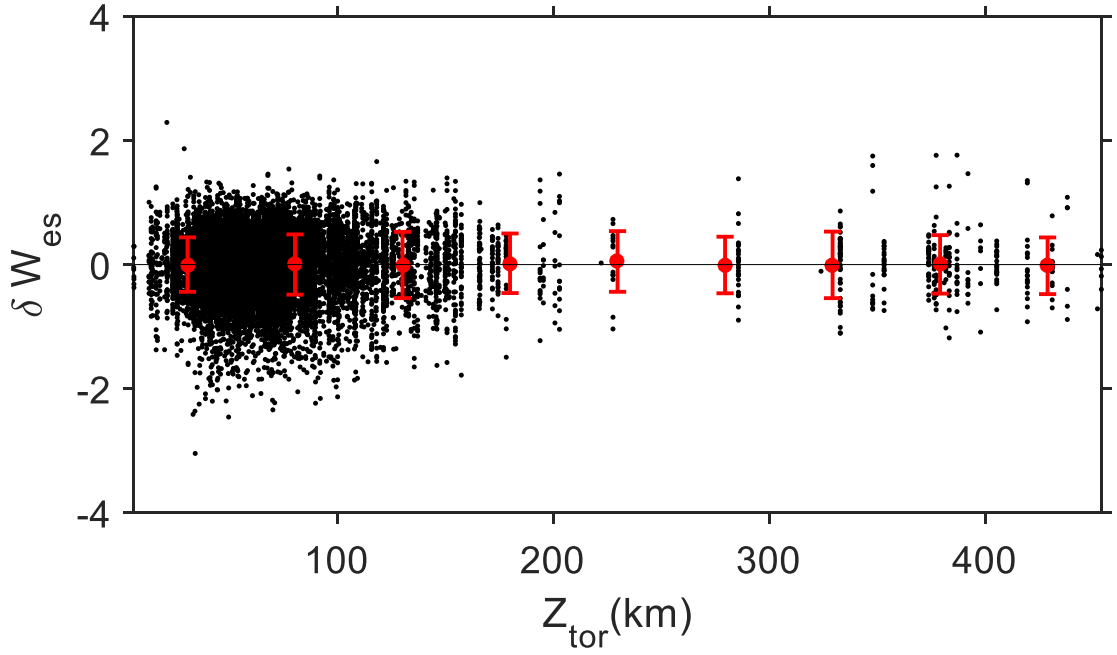

1
0.5
0
-0.5
-1
10^3
$\delta S2S_S$
V_{S30}(m)

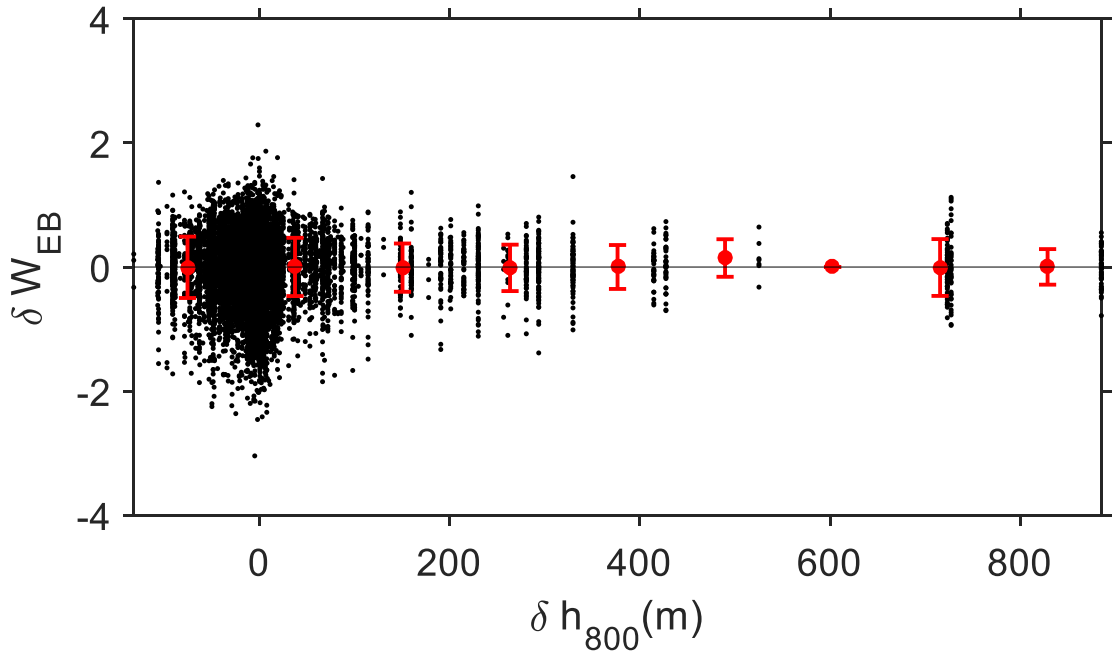

1
0.5
0
-0.5
-1
-200
0
200
400
600
800
$\delta S2S_S$
δh_{800}(m)

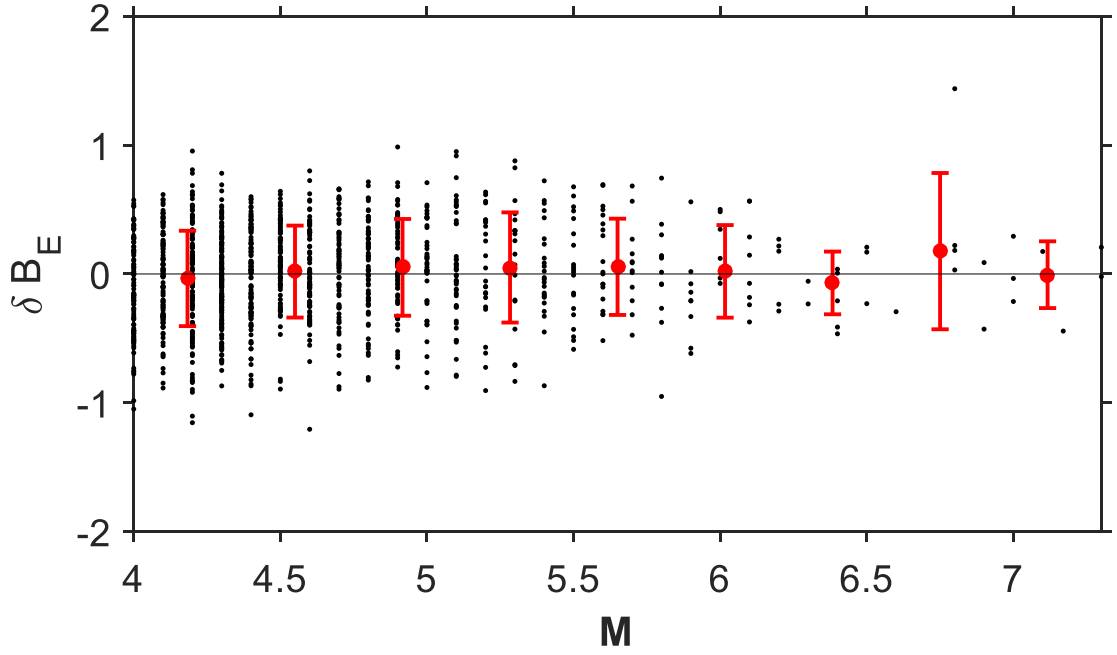
Interface subduction $Ds_{5\text{-}75}$ model

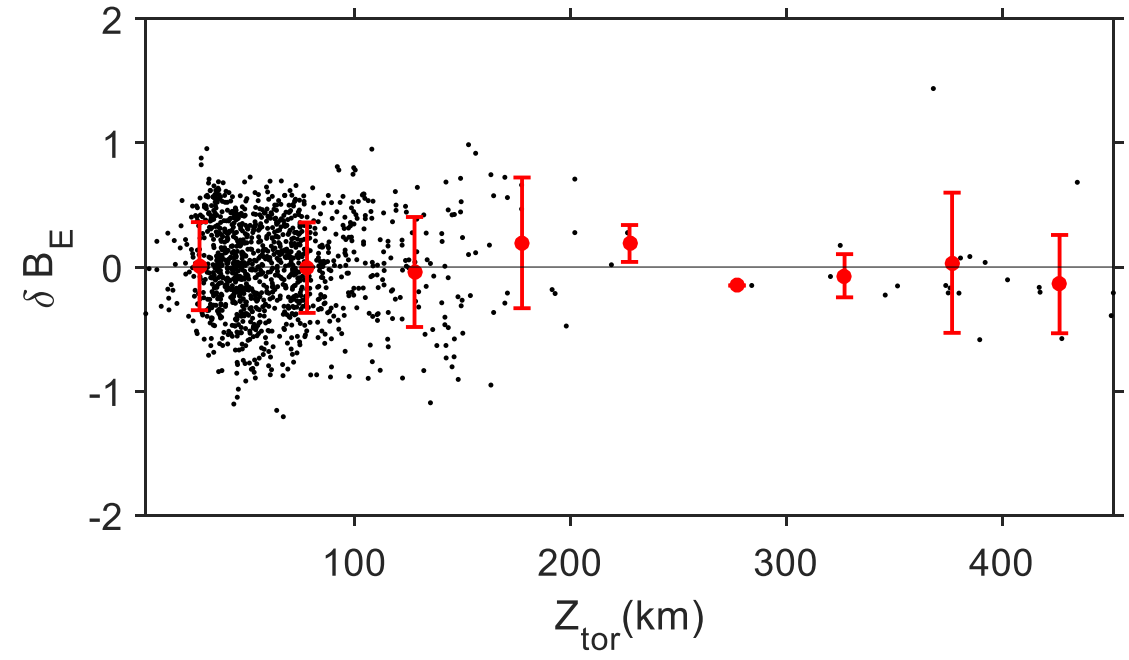


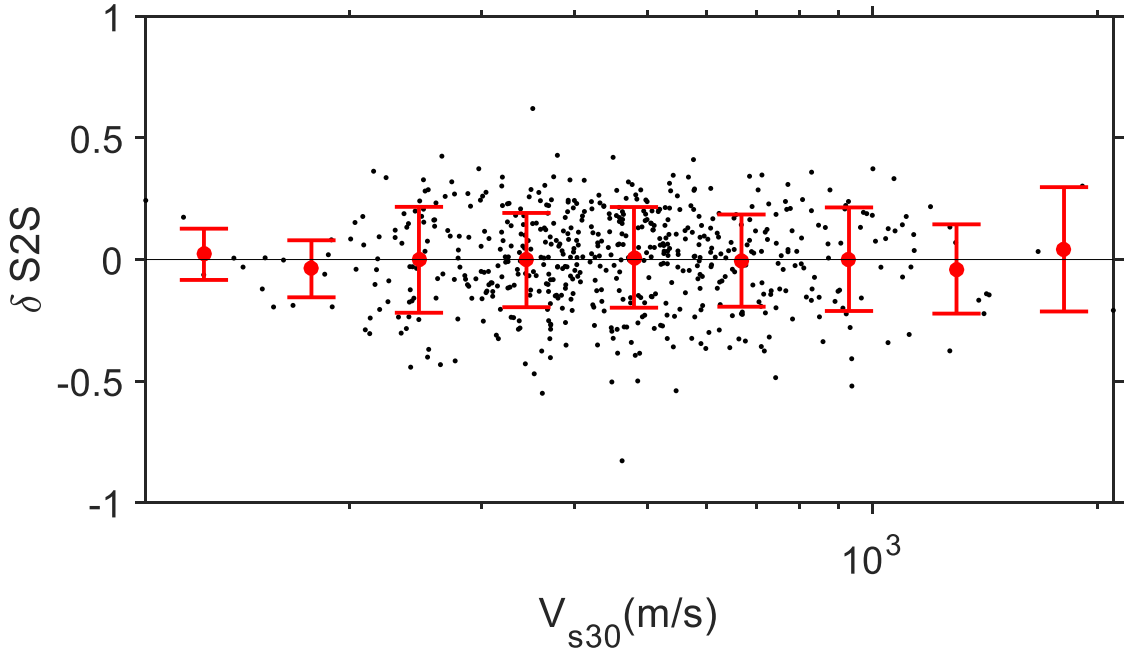

4
2
0
-2
-4
δW_{es}
10^3
V_{s30}(m/s)

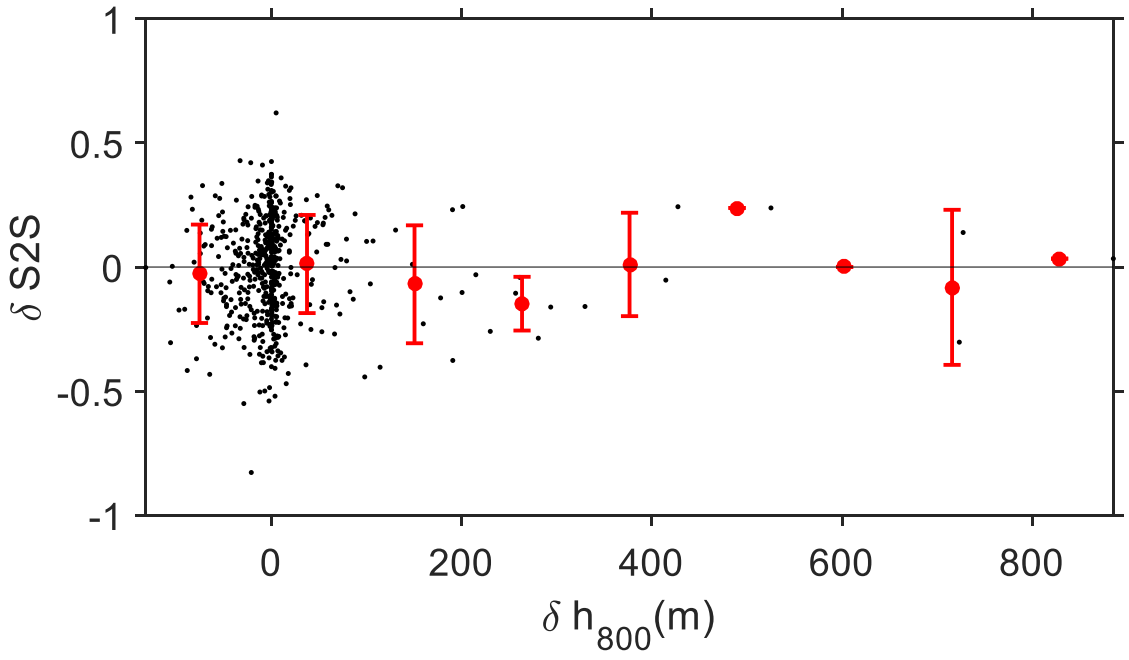

4
2
0
-2
-4
δW_{es}
100
200
300
400
Z_{tor}(km)


4
2
0
-2
-4
δW_{EB}
0
200
400
600
800
δh_{800}(m)

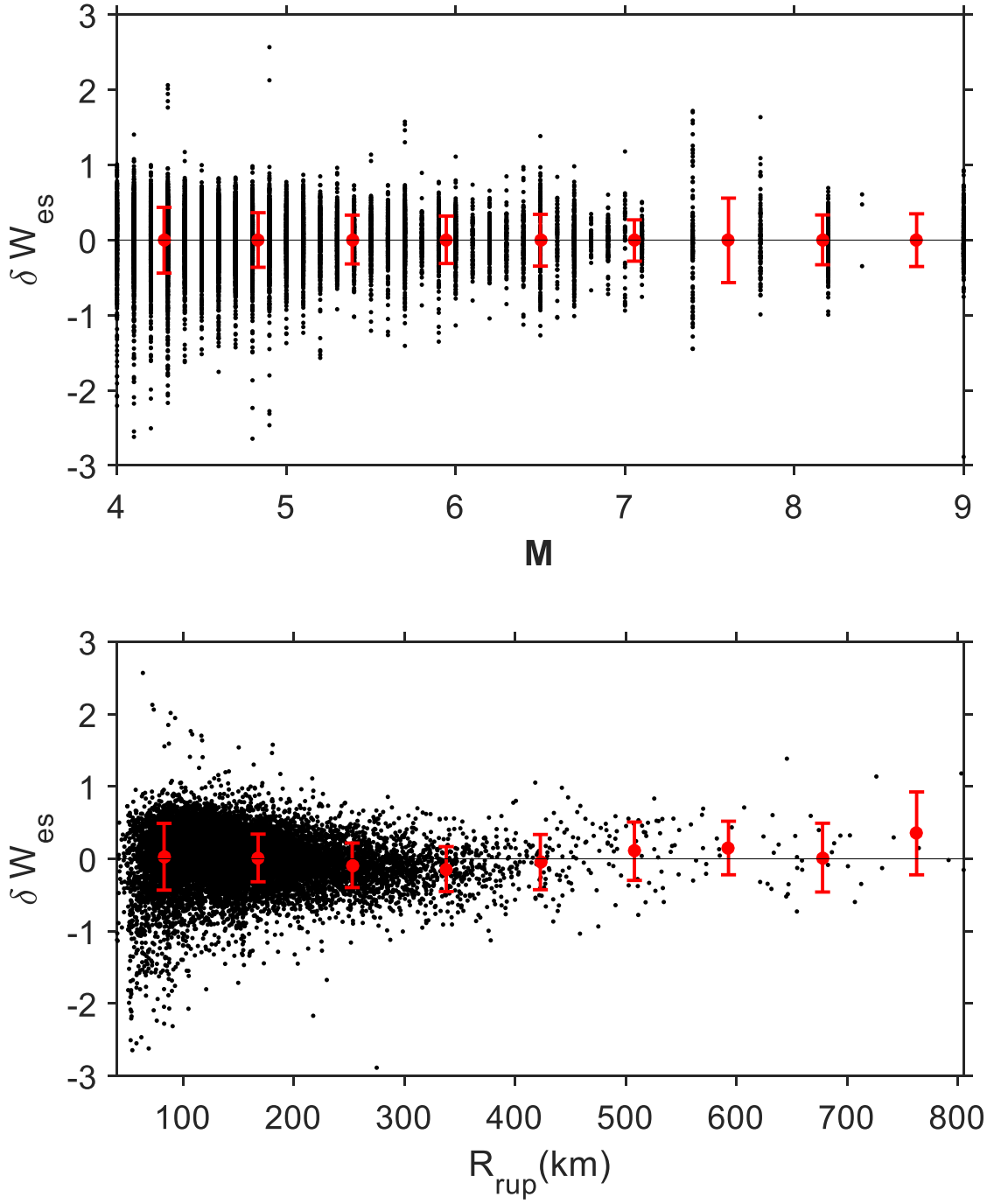

2
1
0
-1
-2
δB_E
4
4.5
5
5.5
6
6.5
7
M

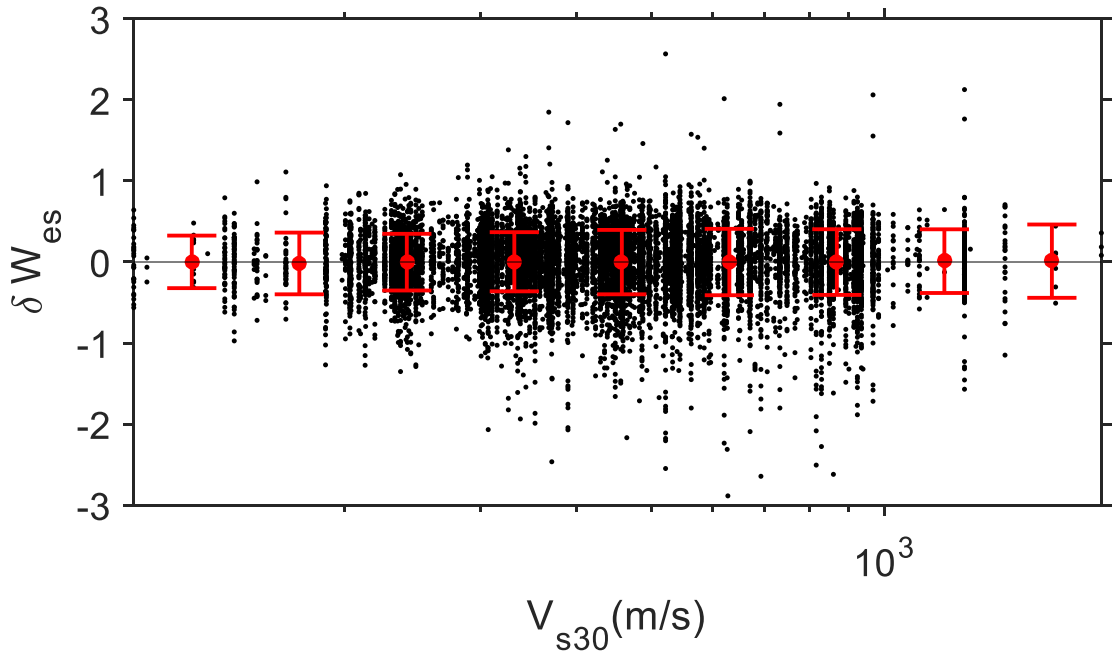

2
1
0
-1
-2
δ B_E
100
200
300
400
Z_tor(km)

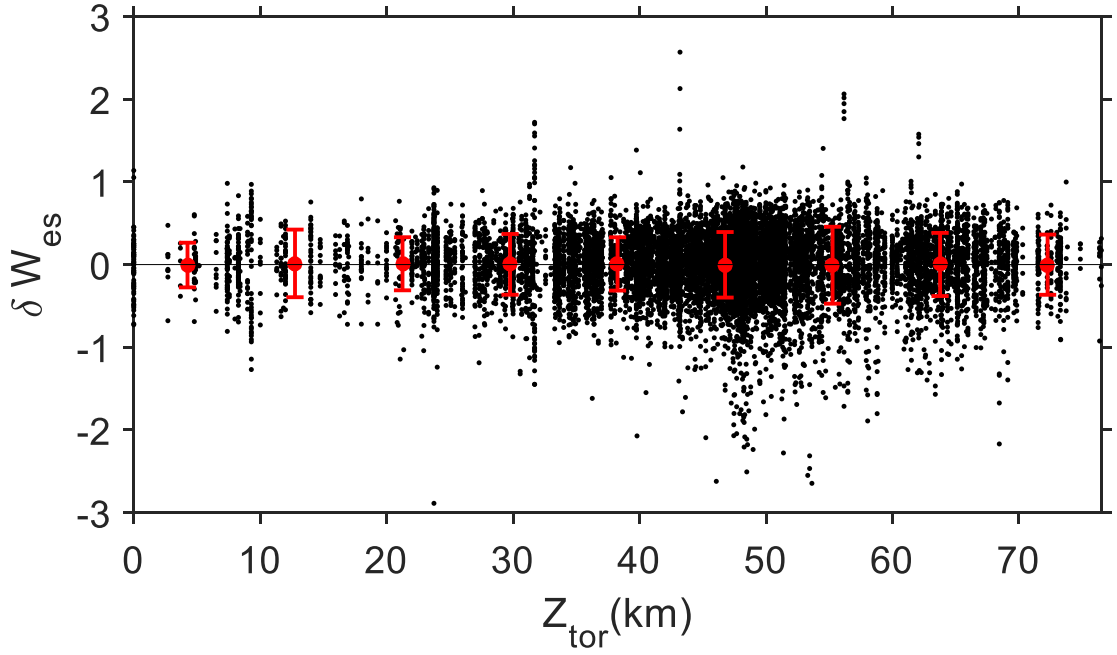

1
0.5
0
-0.5
-1
δ S2S
10^3
V_s30(m/s)

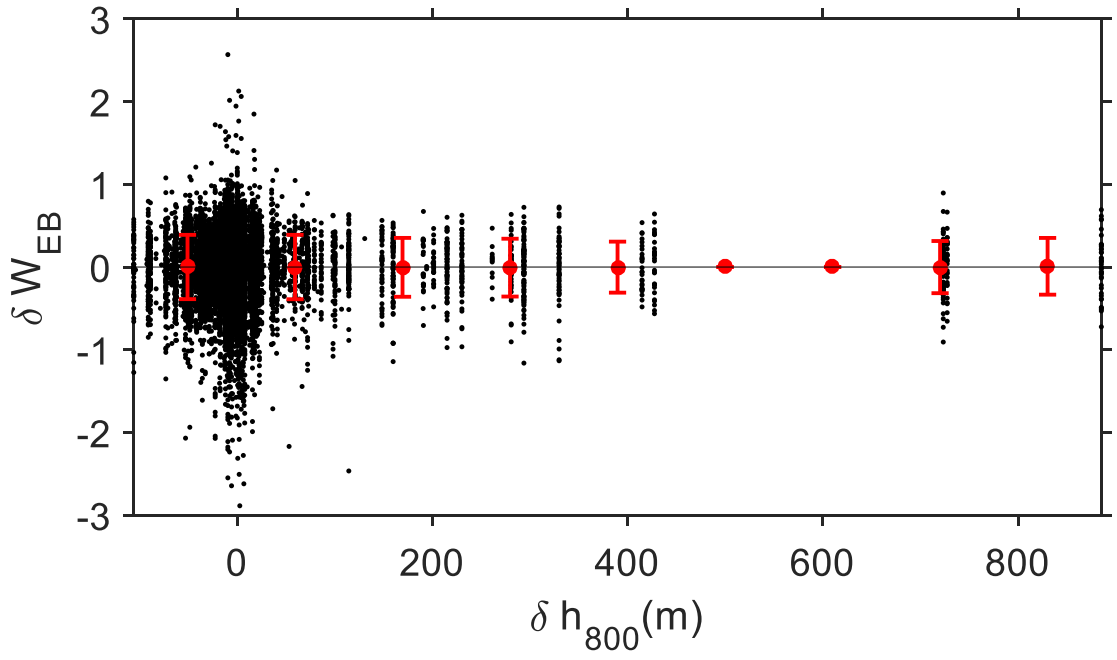

1
0.5
0
-0.5
-1
δ S2S
0
200
400
600
800
δ h800(m)

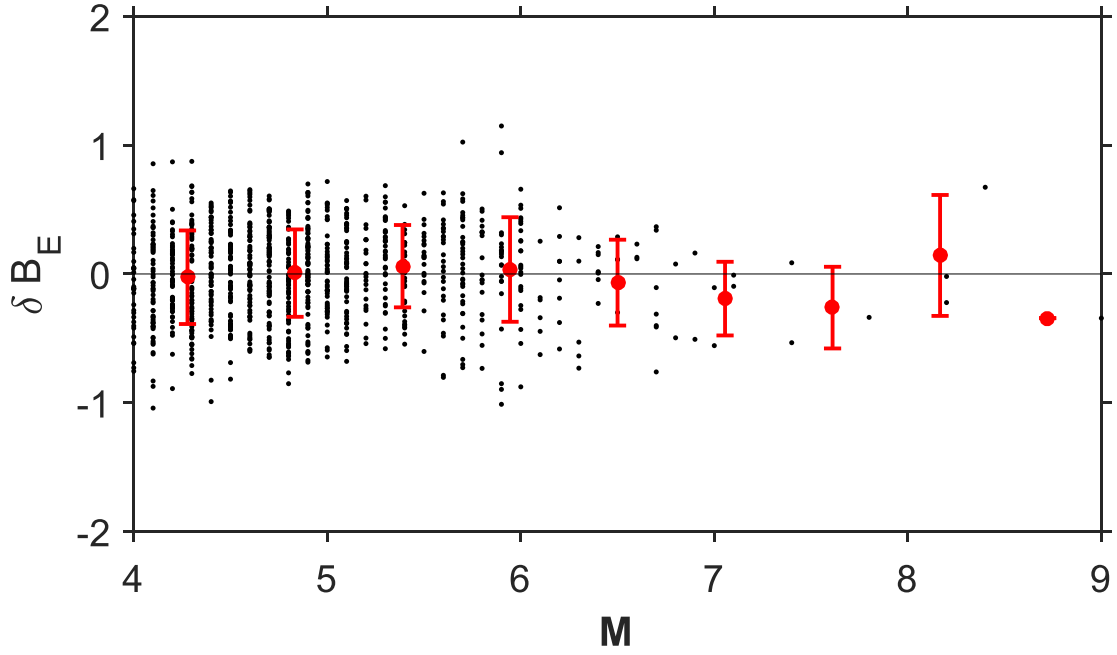
Intraslab subduction $Ds_{5\text{-}75}$ model

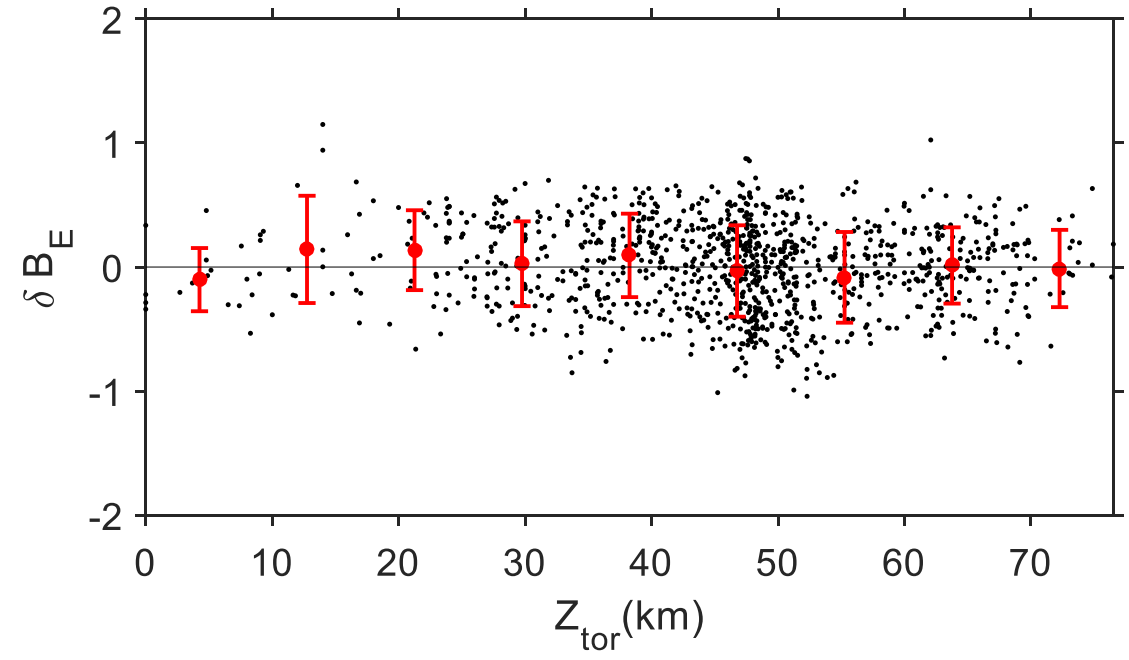


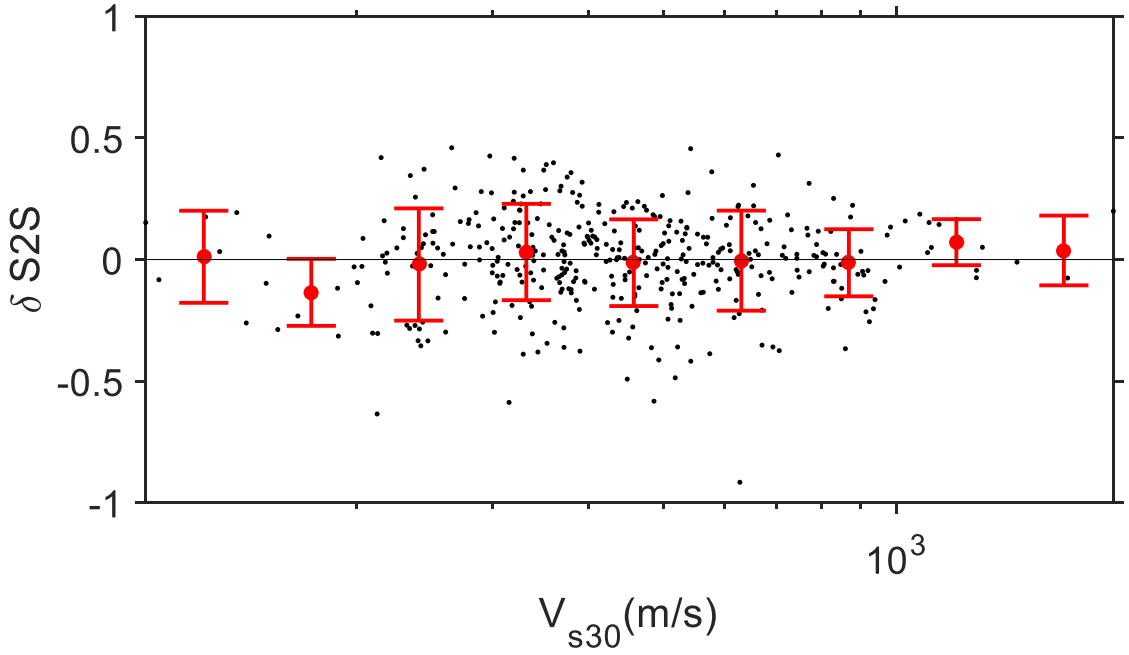

3
2
1
0
-1
-2
-3
δW_{es}
10^3
V_{s30}(m/s)

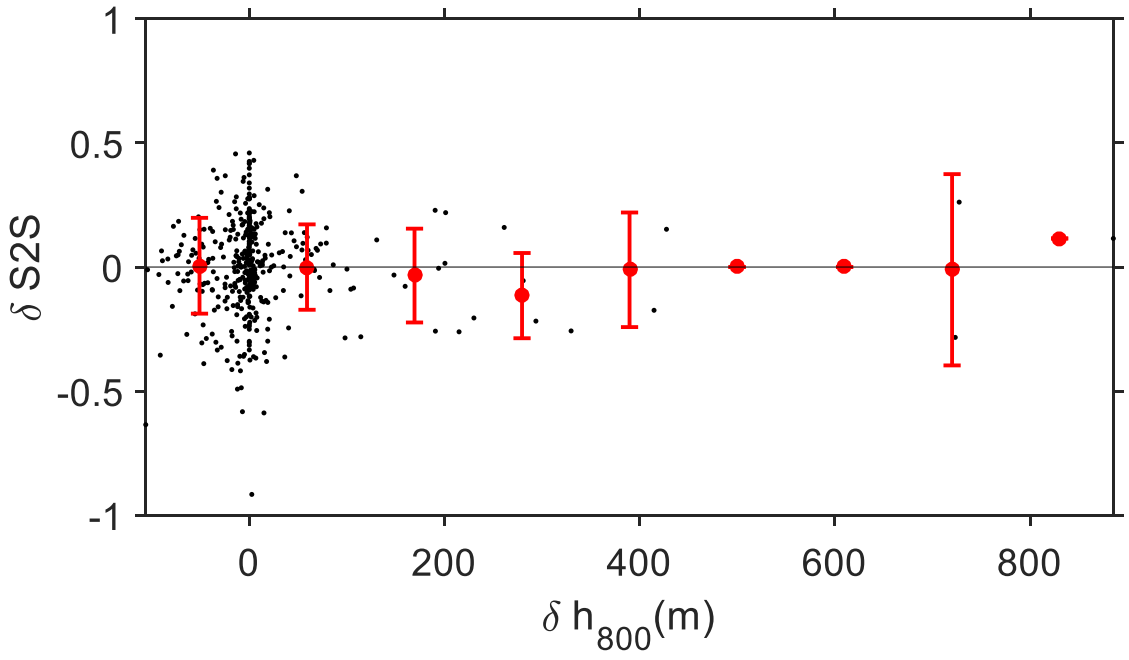

3
2
1
0
-1
-2
-3
δW_{es}
0
10
20
30
40
50
60
70
Z_{tor}(km)


3
2
1
0
-1
-2
-3
δW_{EB}
0
200
400
600
800
δh_{800}(m)


2
1
0
-1
-2
δB_E
4
5
6
7
8
9
M


2
1
0
-1
-2
δ B_E
0
10
20
30
40
50
60
70
Z_tor(km)


1
0.5
0
-0.5
-1
δ S2S
10^3
V_s30(m/s)


1
0.5
0
-0.5
-1
δ S2S
0
200
400
600
800
δ h800(m)