COVER PAGE

USGS Award Number:	G10AP00036
Title:	Reference site standard deviation for integration of site effects into ground motion prediction equations and seismic hazard analyses
Authors and Affiliations:	Christine Goulet (PI), Jeff Bayless URS Group, Inc. 915 Wilshire Boulevard Los Angeles, CA 90017 tel: 213-996-2200 / fax: 213-996-2290 goulet@ucla.edu jeff.bayless@urs.com
Term covered (start, end dates):	January 1, 2010 – August 31, 2011

TABLE OF CONTENTS

LIST OF ACRONYMS AND ABBREVIATIONS

CAV	cumulative absolute velocity
f_c	low-cut corner filter frequency
GMPE	Ground Motion Prediction Equation
GRA	Ground response analysis
IA	Arias intensity
IM	Intensity measure
km	Kilometers
m	Meters
M_{JMA}	Japan Meteorological Agency magnitude
M_w	Moment magnitude
NGA	Next Generation Attenuation
NIED	National Research Institute for Earth Science and Disaster Prevention
PEER	Pacific Earthquake Engineering Research Center
PGA	Peak ground acceleration
PSHA	Probabilistic Seismic Hazard Analysis
R_{hyp}	Hypo distance (km)
s	Seconds
Sa	Pseudo-Spectral acceleration (5% damping)
SDOF	Single degree of freedom
T	Spectral period
USGS	United States Geological Survey
V_s	Shear wave velocity
V_{S30}^{G}	Average shear-wave velocity in top 30-meters of geomaterial
V_s^B	Shear wave velocity at downhole sensor

Residuals, Amplification and Standard Deviations

δB_e	Between-events residual for earthquake *e*
δW_{es}	Within-event residual at station *s* for earthquake *e*
$\mathrm{Amp_{es}}$	Ground surface to borehole IM amplification factor, earthquake *e*, site *s*
$\overline{\mathrm{Amp_s}}$	Mean amplification at station *s*
$\delta\mathrm{Amp_{es}}$	Residual amplification
τ	Standard deviation of the between-events residuals, δB_e for event *e*
ϕ	Standard deviation of the within-event residuals, δW_{es}, for event *e*, station *s*
σ	Standard deviation (total)
τ^B	Standard deviation of the between-events residuals on baserock (reference rock)
τ^G	Standard deviation of the between-events residuals on ground surface
ϕ^B	Standard deviation of the within-event residuals on baserock (reference rock)
ϕ^G	Standard deviation of the within-event residuals on ground surface
ϕ_{Amp}	Standard deviation of the unexplained site amplification residuals, record equal weight scheme
$\phi_{\mathrm{Amp_s}}$	Standard deviation of the unexplained site amplification residuals, station equal weight scheme

1. ABSTRACT

In probabilistic seismic hazard analyses (PSHA), site effects are accounted for by a modification of the median and standard deviation from ground motion prediction equations (GMPEs) developed for a reference (e.g. rock) condition. The median modification is straight-forward and done in a similar fashion for all GMPEs, but the modification of the standard deviation is not always done in a consistent manner. Because the GMPE standard deviation tends to control the PSHA ground motions for hazard levels of interest in engineering, the issue of the reference standard deviation is an important one.

GMPEs are used in PSHA to quantify seismic ground motions for a given hazard level at a site, which has the potential to be affected by specific seismic sources. GMPEs are usually developed for large tectonic regions, using recordings of many earthquakes at different stations. Such GMPEs therefore reflect the attributes of a large region and not that of the specific site of interest. When using these GMPEs for site-specific analysis, we are assuming that the ground motions at the site follow the same variability as the broader dataset. This is called the ergodic assumption, which consists in effectively "trading space for time" in the ground motion quantification.

It is expected that if one had access to a lot of recordings at a station (same site effects) from earthquakes from a same source region (same wave travel paths), the aleatory variability would be smaller than predicted by models using the ergodic assumption. The median prediction should also be adjusted, but in most cases there is not enough data to develop source-path-site-specific GMPEs so the uncertainty on the "correct" value of the median is treated as epistemic uncertainty, through the use of logic trees.

One common practice is to use a semi-ergodic assumption in which the site effects at a station can be well quantified, through measurements and site response analyses, therefore reducing the variability due to site response that would be predicted for example by a more generic site descriptor (e.g. V_{s30}). This report presents an empirical approach used to quantify the ground motion dispersion due to site effects or site amplification (ϕ_{Amp}).

We use records from the dense Japanese KiK-net seismograph network, which consists of 656 functional sites with collocated surface and downhole sensors. We define data subsets based on magnitude, site-source distance, source depth, recording level, and site class as defined by the soil shear wave velocity. These subsets contain only earthquake events recorded at 5 or more stations, and only stations with recordings from 10 or more different earthquakes. Our methodology includes record subset selection, processing and filtering of raw records, and finally the computing of the contribution of site effects to the ground motion standard deviation. The results show that previously used values for ϕ_{Amp} associated with PGA are reasonable, but that modifications to account for spectral period and site stiffness dependence could be made to improve the assessment of the reference standard deviation. This finding can have important implications for PSHA. We acknowledge critical issues with the data and propose a follow-up study to confirm our preliminary results.

2. INTRODUCTION

2.1. BACKGROUND AND SIGNIFICANCE TO HAZARD REDUCTION

Probabilistic seismic hazard analyses (PSHA) are performed using ground motion prediction equations (GMPEs) in combination with earthquake source models and rupture forecasts. The product of PSHA is hazard curves that apply to the conditions of the site of interest (soil of a certain stiffness). For a proper computation of hazard, the site effects must be accounted for within the GMPE. This is discussed in detail in Goulet and Stewart (2009).

In the context of ground motion hazard assessment, reference (rock) ground motions are often modified to reflect the effects of local site conditions. Often this modification is performed within a GMPE through a site term that is part of the model. Alternatively, reference site ground motions can be modified through separate site-specific analyses, typically using 1D ground response analyses (GRA). In either case, the ground motion modification is applied to both the median and standard deviation of the reference (rock) ground motion prediction. The adjustment of the median is straightforward, consisting of multiplication by a frequency-dependent site factor, but the treatment of the standard deviation is still the subject of discussions. Two of the Next Generation Attenuation (NGA) GMPEs modify the rock standard deviation to reflect soil conditions (Abrahamson and Silva, 2008; Campbell and Bozorgnia, 2008), whereas two others directly evaluate standard deviation for soil without involving the rock value (Boore and Atkinson, 2008; Chiou and Youngs, 2008). When site-specific ground response analyses are performed, similar approaches have been put forth whereby soil standard deviation is modified from the rock value (Bazzurro and Cornell, 2004) or directly estimated (Baturay and Stewart, 2003). In engineering applications, where design is often based on low probabilities of exceedance (long return periods), the standard deviation tends to significantly influence the computed ground motions. It is therefore very important to properly quantify that standard deviation. However, for models in which standard deviation for soil is modified from rock values, there is currently no consensus as to what value of reference site standard deviation to use within these models. There are two distinct schools of thought on this issue.

One of the approaches is to take the dispersion of the within reference (rock) ground motions as equal to the dispersion of the outcropping ground motions on equivalent reference site conditions. The second approach is to reduce the dispersion for within motions relative to outcropping motions, as will be explained below. Two NGA GMPEs (Abrahamson and Silva; Campbell and Bozorgnia, 2008) adopt this approach. Depending on the chosen assumption, one can expect differences in predicted ground motions as large as 25-30% for a ground motion return period of 2475 years (2% in 50 years probability of exceedance). The two different approaches could effectively be applied either within a given GMPE that includes a site effects term or when GRA results are used in combination with rock GMPEs. As it is explained in the next section, the second approach is desirable when using the reference standard deviations from most GMPEs, but a question remains on the level of standard deviation reduction that should be applied. This is therefore an important issue with practical implications.

GMPEs and the Ergodic Assumption

GMPEs are used in PSHA to quantify seismic ground motions for a given hazard level at a site, which has the potential to be affected by specific seismic sources. GMPEs are usually developed for large tectonic regions, using recordings of many earthquakes at different stations. Such GMPEs therefore reflect the attributes of a large region and not that of the specific site of interest. When using these GMPEs for site-

specific analysis, we are assuming that the ground motions at the site follow the same variability as the broader dataset. This is called the ergodic assumption, which consists in effectively "trading space for time" in the ground motion quantification.

It is expected that if one had access to a lot of recordings at a given station (same site effects) from earthquakes from a same source region (same wave travel paths), the aleatory variability would be smaller than predicted by models using the ergodic assumption. The median prediction should also be adjusted, but in most cases there is not enough data to develop source-path-site-specific GMPEs so the uncertainty on the "correct" value of the median is treated as epistemic uncertainty, through the use of logic trees.

One common practice is to use a semi-ergodic assumption in which the site effects at a station can be well quantified, through measurements and site response analyses, therefore reducing the variability due to site response that would be predicted for example by a more generic site descriptor (e.g. V_{s30}, the time-averaged shear wave velocity in the top 30 m of soil column).

This report presents an empirical approach used to quantify the ground motion dispersion due to site effects or site amplification (ϕ_{Amp}). We use records from the dense Japanese KiK-net seismograph network, which consists of 656 functional sites with collocated surface and downhole sensors. We define data subsets based on magnitude, site-source distance, source depth, recording level, and site class as defined by the soil shear wave velocity. These subsets contain only earthquake events recorded at 5 or more stations, and only stations with recordings from 10 or more different earthquakes. Our approach includes record subset selection, processing and filtering of raw records, computation of the standard deviation due to site amplification and analysis of results with site parameter V_{s30}. This quantification requires the clarification of the ground motion standard deviation terminology.

Sigma You Say?

The discussion on GMPE residuals and their standard deviation mostly follows the notation from Al Atik et al. (2010). A summary of the essential parameters relevant to the current study is presented below.

The mixed-effects regression used in modern GMPE development returns a median ground motion prediction μ_{es} along with the between-event (δB_e) and the within-event (δW_{es}) residuals for earthquake event *e* and station (or site) *s* as indicated by the subscripts. The between-events residuals characterize the deviation from the median model due to different earthquakes while the within-event residuals quantify the deviation at different stations for a given earthquake. The residuals are zero-mean, independent, normally distributed random variables with standard deviations τ and ϕ respectively. In the context of most GMPEs, these standard deviation terms usually apply to ground motions at the ground surface. τ and ϕ are assumed to be uncorrelated and the total standard deviation of residuals σ is given by

$$\sigma = \sqrt{\phi^2+\tau^2} \qquad (1.1)$$

For the present work, we are interested in quantifying the part of the standard deviation of the within-event residuals ϕ related to site effects (ϕ_{Amp}). We have collocated recordings both at the ground surface (*G*) and at the bottom of the borehole, which we will call basement rock (*B*) for consistency with the Al Atik et al. (2010) notation.

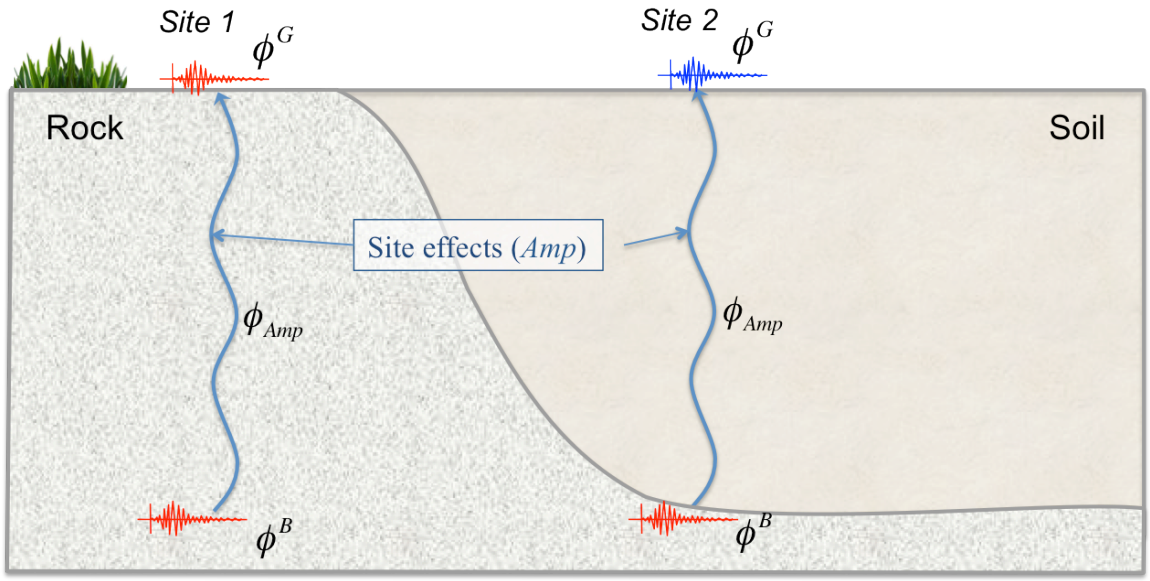


Figure 1: Illustration of baserock (B), ground surface (G) and site effects standard deviation terms.

Figure 1 illustrates these standard deviation terms, which are related through their variances as follow:

$$\left(\phi^{G}\right)^2 = \left(\phi^{B}\right)^2 + \left(\phi_{\mathrm{Amp}}\right)^2 \qquad (1.2)$$

where ϕ_{Amp} represents the dispersion due to wave propagation to the surface (e.g. amplification), generally referred to as *site effects*.

Including Site Effects in GMPEs

As mentioned above, site-specific PSHA are often completed. These site-specific analyses consist in modeling the characterized soil column at the site and completing GRAs to quantify the median site amplification $\mu_{Amp,SiteSpecific}$ and its standard deviation $\phi_{Amp,SiteSpecific}$ directly. Different models of GRA, both empirical and theoretical, exist for integrating the modification of standard deviation into PSHA. These different models have been reviewed by the PIs and are presented in Dr. Goulet's dissertation (Goulet, 2008). The difference between the approaches resides in their treatment of the standard deviation. One key assumption for *all* the models that modify the reference estimates is that the standard deviation for the reference condition (let us temporarily represent this quantity with ϕ_{Rock}) at depth, below the modeled soil column, is known.

For example, in the theoretical model (Bazzurro and Cornell, 2004), the variance on soil is given by:

$$\left(\phi^{G}\right)^2_{\text{Site Specific}} = c * \left(\phi_{\mathrm{Rock}}\right)^2 + \left(\phi_{\mathrm{Amp,Site\ Specific}}\right)^2 \qquad (1.3)$$

where *c* is the slope of the median regression of the site amplification factor (can contain multiple terms, reflecting the regression functional form). In most applications, ϕ_{Rock} is assumed to be the dispersion coming directly from the GMPE. If that GMPE was developed with the ergodic assumption (using a large number of sites), the dispersion that is used actually corresponds to a surface estimate ϕ^G

instead of the desired ϕ^B corresponding to the variability at depth. This misuse of ϕ^G instead of ϕ^B leads to a "double-counting" of the site effects contribution to dispersion ϕ_{Amp}; one contribution coming from the ergodic assumption included in the GMPE and the second part coming from the site-specific analyses.

The correct approach would be to use:

$$\left(\phi^{\mathrm{G}}\right)^2_{\text{Site Specific}} = \mathrm{c} * \left(\phi^{\mathrm{B}}\right)^2 + \left(\phi_{\text{Amp,Site Specific}}\right)^2 \tag{1.4}$$

Which is equivalent to:

$$\left(\phi^{\mathrm{G}}\right)^2_{\text{Site Specific}} = \mathrm{c} * \left[\left(\phi^{\mathrm{G}}\right)^2 - \phi^2_{\text{Amp,GMPE}}\right] + \left(\phi_{\text{Amp,Site Specific}}\right)^2 \tag{1.5}$$

The GMPE subscript in equation 1.5 is used only to differentiate the two components of variability (from the ergodic assumption and from the site-specific GRAs) and will be dropped for the rest of the discussion.

It becomes apparent that if one uses a GMPE developed using the ergodic assumption for reference condition ground motions, a reduction of the within-event standard deviation must be included before applying site effects (either within the GMPE or for use with GRA results). In the NGA GMPEs, both Campbell and Bozorgnia and Abrahamson and Silva reduce $(\phi^B)^2$ by 0.3^2 before applying the site factor. For Campbell and Bozorgnia, the ϕ_{Amp} = 0.3 value comes from the average dispersion from many site categories, as published in the literature (Baturay and Stewart, 2003; Silva et al. 1999, 2000; Bazzurro and Cornell, 2004 and 2005). For Abrahamson and Silva, the ϕ_{Amp} = 0.3 value comes from previous GRA results by Silva (2008, unpublished report) in which input motions, nonlinear modulus reduction and damping curves, and shear wave velocity (V_s) profiles were randomized. We are assuming that the V_s variability considered by Silva is within a reasonable range for rock sites (V_{s30}=760-1100 m/s), based on another unpublished report by Abrahamson and Silva (2007).

However, recent work by the first author (Goulet, 2008) highlighted that the variability of site response for stiff sites should be lower. Although this work by Goulet did not include site profile variability, it raised the question of whether or not reducing the reference rock variance by 0.3^2 is appropriate for site-specific analyses where GRA are performed on soil columns sited above *stiff* site conditions.

2.2. Project Objectives

The objective to this project is to compute ϕ_{Amp} directly using records for collocated stations (surface and downhole) and to quantify the effect of site stiffness on the results. The results from this research can improve seismic hazard assessments by providing guidance on the appropriate value of reference standard deviation to use as input with GMPEs and site-specific models relative to an outcropping standard deviation.

2.3. Approach Overview

The present study utilizes an empirical approach to compute the dispersion (standard deviation) due to site amplification (site effects) for a large dataset. The analyses are done for a suite of intensity measures (IM): peak ground acceleration (PGA) and 5% damped spectral acceleration (S_a) at selected periods. Our approach can be grouped into four major tasks:

1. Gather and process KiK-net acceleration records from surface and borehole sensors along with the associated earthquake events metadata and site properties at the recording stations; select a subset of data for the analyses (Section 3).
2. Compute the standard deviation on the amplification ϕ_{Amp} using records from collocated stations (surface and downhole). Evaluate the effect of site stiffness on the computed ϕ_{Amp} values. This is presented in Section 4.
3. Interpret and summarize results and develop recommendations on the applicability of the results and on future work to be completed.

The original plan presented in the proposal included the analysis of residuals relative to two re-regressed GMPEs to address the issue of the reference standard deviation (quantification of ϕ_{Amp}). The PIs conducted this part of the study as well, but it did not help address the problem of the reference standard deviation. The results from this complementary study are not presented here.

3. GROUND MOTION DATA

3.1. KIK-NET SEISMOGRAPH NETWORK

This study makes use of the KiK-net strong motion observation seismograph network, which is maintained by the Japanese National Research Institute for Earth Science and Disaster Prevention (NIED). KiK-net is a dense network used for monitoring seismic activity in and around Japan, and currently has 656 active recording stations (Figure 2). Each station consists of two sets of three orthogonal accelerometers; one set located at the ground surface and another at the bottom of a borehole of varying depth (100 m or greater). The KiK-net database considered in this project contains nearly 150,000 three-component recordings from 5,894 events recorded in the 1997-2010 period, for both surface and downhole levels.

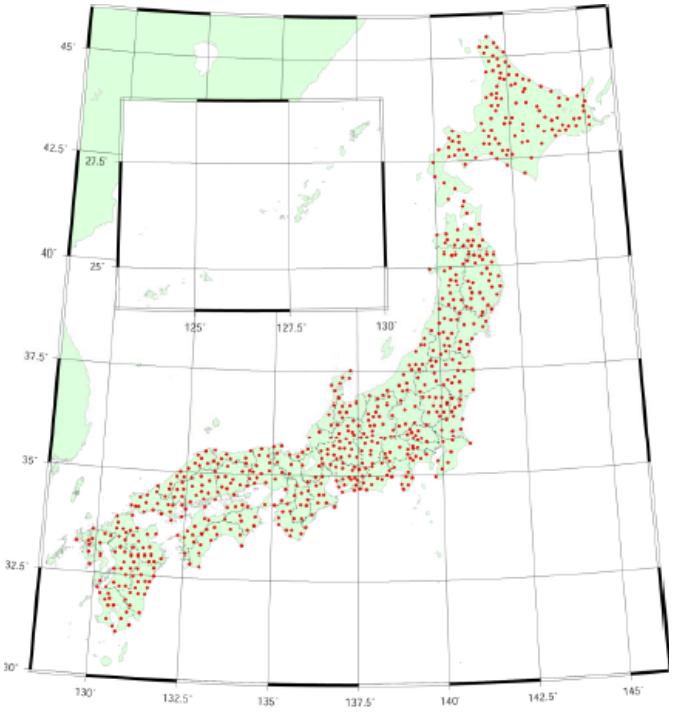


Figure 2: Dense distribution of KiK-net stations (red dots) in Japan (Fujiwara et al., 2004).

During installation of the sensors, the sediment at each site was penetrated until bedrock was reached. For this purpose, bedrock corresponds to NEHRP site categories A (V_s > 1500 m/s) or B (760 < V_s < 1500 m/s). This criterion is what led to varying borehole depths throughout the network. The majority of boreholes are between 100 to 200 m deep, but a few extend to depths beyond 1000 m.

KiK-net is an open data network, with ground motion records available through the Internet just hours after the occurrence of an earthquake. Users can select specific events or stations with a combination of key parameters to download desired records, or access the entire database through an FTP site. Additionally, the shear wave velocity profile for most recording stations is available on the website along with site location maps. Simplified shear wave velocity profiles are available down to, but not below, the downhole sensor (Figure 3). Profile information is stored in comma-separated files containing layer thickness with corresponding shear and compression wave velocities. More detailed rock and soil

condition profiles showing the site lithology, altitude and depth are also available for a subset of stations. Source/path meta-data such as event time, location, depth, Japan Meteorological Agency (JMA) magnitude and station epicentral distance can be compiled from file headers. Aoi et al. (2004) presents further details about the KiK-net network and datasets.

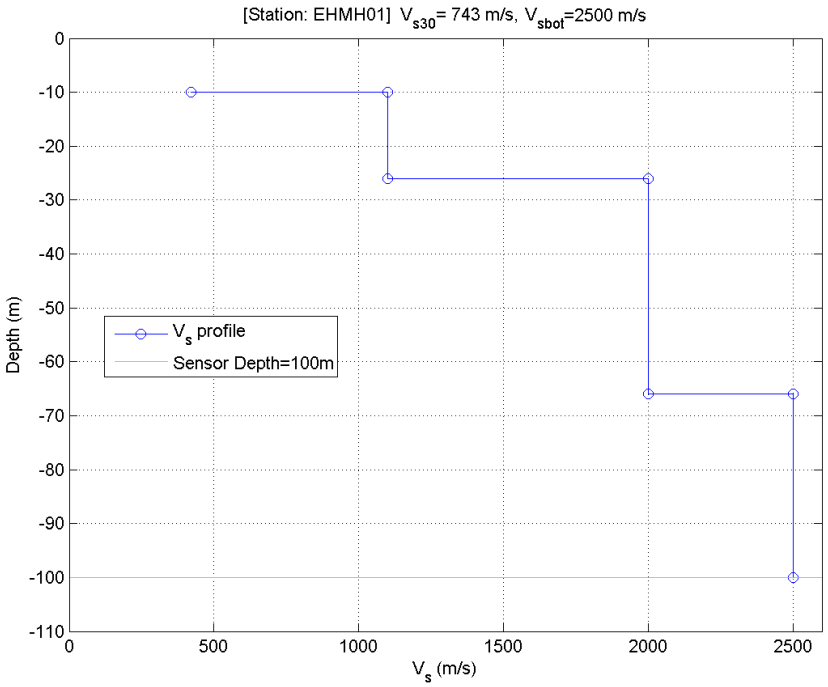


Figure 3: Example of KiK-net shear wave velocity profile data (station EHMH01).

3.2. Data Needs

In order to confidently capture the dispersion of ground motions, it is necessary to obtain a sufficient number of recordings. Individual earthquake events may cause larger or smaller than average ground motions and a certain array might exhibit low dispersion right around its natural period of vibration, but with a large enough sample size the effect of these trends can be quantified and accounted for. We have set a minimum limit of 5 station recordings per event and 5 events per station.

Because the KiK-net network is located in Japan, it records events on faults from different tectonic regimes. The sources are subdivided in three types: subduction intra-plate, subduction inter-plate and shallow crustal sources. Because we want to maximize the portability of results to models currently used in active regions such as California, we are only considering shallow crustal events. The distinction of the tectonic regime for a specific event is a difficult task and is discussed further below.

3.3. Dataset

Our original plan was to use the dataset compiled by Zhao et al. (2006) for the development of their GMPE. The dataset is composed of strong motion records from earthquakes with magnitude 5.0 to 8.3 recorded within 300 km of the source. Zhao et al. (2006) identified 576 Japanese recordings associated to a shallow crustal source type. The authors systematically reviewed the records, also obtained from KiK-net, to detect instrumental problems and to discard unusable data. They sorted the records by

source type and filtered the remaining records, accounting for instrument-specific response. The advantages of this dataset are that it already featured strong ground motions, carefully characterized source and focal mechanisms, stations with multiple records and only usable and processed records.

However, the dataset seemed scarce in light of all the available KiK-net data postdating the Zhao et al. (2006) work, especially for conducting a detailed analysis on site effects. The original dataset contained a limited number of stations with 10 or more recordings for both the surface and borehole sensors (less than 100). This realization prompted us to compile our own dataset directly from the KiK-net database, which also allowed us to include smaller magnitude events. This proved to be a major task and included downloading and selecting the records, compiling meta-data, and processing and filtering the records. Because a large time investment was originally spent on expanding the Zhao et al. (2006) dataset, we have taken a simplified approach to process the records. This process is summarized in the following sections.

3.4. Data Processing

Strong-motion records are always affected by noise or instrumental drift to some degree and careful signal processing needs to be applied to the raw time series. State of the art methods documented by David Boore have been adopted in the present study (Boore, 2002; Boore, 2005; Boore and Bommer, 2005; Douglas and Boore, 2011). Batch, uniform processing was performed on the desired KiK-net records in an attempt to remove baseline errors, high and low frequency distortion, and low signal-to-noise ratios.

The most important decision in the processing scheme is the application of a low-cut (high-pass) filter to remove the low frequency parts of the record contaminated by long period noise (Boore and Bommer, 2005). Low frequency noise is usually expressed as drifts in the displacement time series derived from double integration of acceleration, and a baseline problem can be marked by a non-zero average velocity a sufficiently long time after the quake (Boore, 2002). Figure 4 shows an example record velocity, displacement and Fourier amplitude spectrum before and after the complete filtering process. The determination of the corner frequency is subject to interpretation and its specific value tends to be different for each record. Although we have developed a set of computer codes to optimize the selection of the corner frequency, time and budget constraints were taken into consideration in the choice of a single value to use in the processing which was set to f_c=0.25 Hz. This value was selected carefully by inspection of many records and also by considering the frequencies selected by Zhao et al. (2006) for their own database (personal communication with Dr. Zhao). Future work would include the use of the customized scripts to select the corner frequency on a per-record basis.

Record processing begins with a zeroth-order baseline correction to the raw acceleration file by subtracting the mean of the pre-event portion of the record (up to a second before the first arrival), or the mean of the entire record if the pre-event portion is unavailable (Boore, 2002). The baseline corrected acceleration is tapered along the outer 10% of the record time series to avoid truncation effects using a Tukey (tapered cosine) window. Before filtering, the time series is padded with zeros before and after the record, with total pad length determined by:

$$\mathrm{T_z} = 1.5\mathrm{n}/\mathrm{f_c} \qquad (3.1)$$

where n is the Butterworth filter order (n=5 was used for all the records) and f_c is the filter corner frequency (f_c=0.25). Half of T_z is added before and after the record, therefore 15 zeros have been added

before and after each record before filtering (Boore, 2005). The Nyquist frequency, or one half of the sampling frequency of the sensor, is used as the high-cut (low-pass) corner frequency, f_h. Finally, a fifth order Butterworth band pass acausal (does not introduce a phase shift) filter is applied to the instrument-corrected, tapered and padded acceleration in order to obtain the filtered acceleration time series.

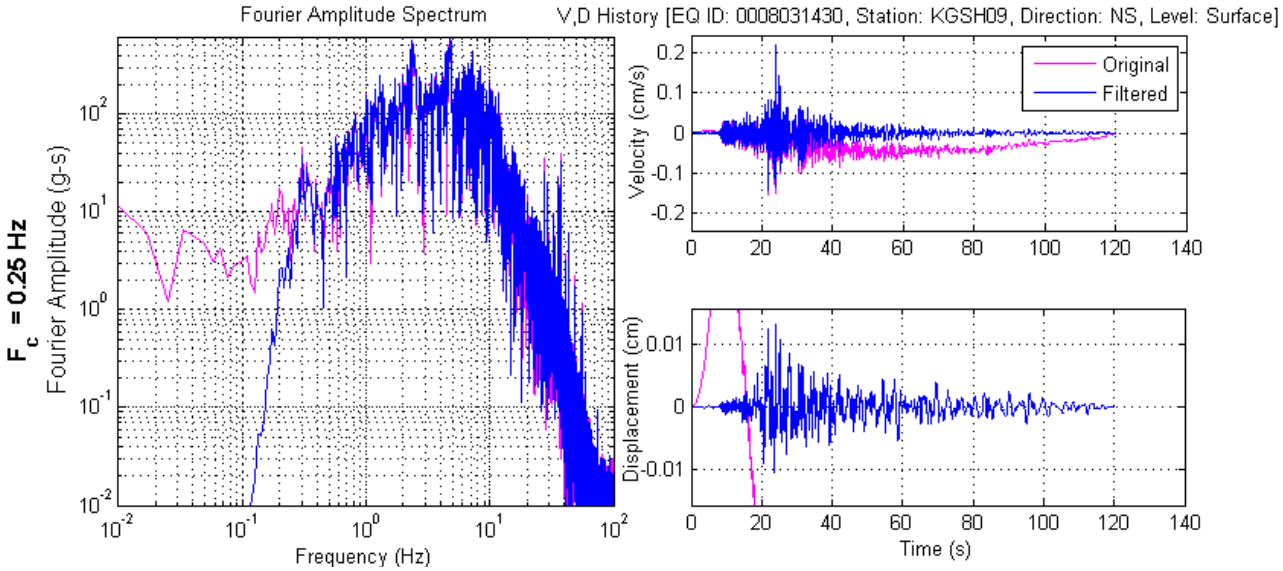


Figure 4: Example of filtered KiK-net record (ID KGSH09008031430). Left panel shows the original and processed Fourier amplitude spectra; right panels show the velocity and displacement time series. The pre-filtered displacement contains drifts which extend beyond the plot boundaries and a non-zero residual displacement at the end of the record (not shown).

3.5. Intensity Measures: PGA and Spectral Acceleration

Pseudo-spectral accelerations of a single degree of freedom (SDOF) oscillator with 5% damping were calculated at 13 selected spectral periods from the processed KiK-net time series. The chosen IMs are: PGA and pseudo-spectral acceleration at the following periods: 0.01, 0.02, 0.03, 0.05, 0.10, 0.20, 0.30, 0.50, 0.60, 1.0, 1.4, 2.0, and 3.0 s. Problems of bias in *Sa* due to filtering are documented in the literature: Akkar and Bommer (2006) and Douglas and Boore (2011) provide summaries. The level of bias introduced by filtering depends on the corner frequency and the signal to noise ratio and becomes more important for increasing spectral periods. Recommendations for the maximum spectral period (T_{max}) to be used range from $0.3/f_c$ to $0.9/f_c$. Akkar and Bommer (2006) developed the most detailed recommendations that take into account the type of data (analog or digital), the site on which the accelerograms are recorded and the filter order. They recommend the use of $T_{max}=0.65/f_c$ for digital records on rock sites to ensure that 90% of *Sa* (or *Sd*, spectral displacement) will be within 10% or the unfiltered Sa with a 95% confidence interval. For the current project, there was no a-priori decision on T_{max} and the analyses were completed up to 3 s. Nonetheless, it is the authors' opinion that ordinates beyond 1.0 s be disregarded (this corresponds roughly to $T_{max} =0.25/ f_c$) until more progress is made on batch KiK-net record-specific processing.

The IM values (PGA and *Sa* at selected periods) are computed in an automated fashion from the processed time series. The records metadata, including the station site properties, and IMs are saved into a flatfile that can be sorted in Excel or MatLab to generate specific subsets of data.

3.6. FINAL DATASETS

3.6.1 General selection criteria

The usable catalogue of processed KiK-net recordings was generated by enforcing the following criteria (see Table 1 for details):

1. Only records for sites with available V_s profiles were kept.
2. Limits on event depth were applied in order to limit the inclusion of subduction events. The tectonic regime (shallow crustal, subduction intra-plate or inter-plate) is not provided for the KiK-net events. The causative tectonic regime is difficult to accurately establish, especially for small magnitude events. Zhao et al. sets the maximum depth for crustal earthquakes to 25 km (Zhao et al., 2006 and Zhao, personal communication) and we used this criterion to exclude events originating at larger depths. Although not perfect, this criterion is a simple one that could be applied broadly and uniformly to the whole dataset.
3. Only events with M_w of 4 or greater were considered. Restrictions on the upper bounds for event magnitude (6.6) and site-source epicentral distance (200 km) were enforced to limit the potential bias from data bins with scarce number of data points.

Table 1 summarizes the selection criteria, along with parameter descriptions. Source type and focal mechanism are not available from KiK-net, so these parameters have not been included as formal criteria (see Section 3.7 for known data issues.)

Table 1 - Event and Site Database Criteria

Parameter	Description	Requirement
V_{s30}^{G}	Time-averaged upper 30 m shear wave velocity, ground surface	> 0
V_s^B	Shear wave velocity at downhole sensor	> 0
H	Downhole sensor depth	> 0
Depth	Rupture hypocenter depth	≤ 25 km
M_w [1]	Moment magnitude[1]	4.0-6.6
R_{epi}	Epicentral distance from recording site	≤ 200 km

[1] - M_{JMA} converted to M_w using Fukushima 1996 relationship (Fukushima, 1996)

3.6.2 Definition of Recording Levels and Site Properties

From the usable catalogue, different data subsets for each 'level' were compiled: surface and downhole. The site parameter for surface stations is V_{s30}^{G}, which is computed directly from the KiK-net site profiles. For borehole records, only the shear wave velocity at the sensor level is available (V_s^B).

Because V_{s30} is used directly in the GMPEs or is used to define site categories, we were hoping to use the same site parameter for both levels in our regression. We have spent a considerable amount of resources early on in the development of a predictive model to infer V_{s30} below the sensor depth since V_s measurements are only reported down to the sensor itself (Bayless and Goulet, 2011). Using all the KiK-net stations with soil profile information, we created a model to predict the lowest *known* V_{s30} (of the 30 m directly above the downhole sensor) based on the properties of the soil column above. Since this value can be calculated from measured V_s, we were able to validate our model before performing an extrapolation to predict the V_{s30} below the downhole sensor.

Several permutations of parameter combinations were considered in the development of this model, including the sensor depth, the average shear wave velocity of the entire column, peak shear wave velocity, the standard V_{s30}, and the V_{s30} of various levels throughout the soil column. We found that the V_s value exactly 30 m above the downhole sensor was the best predictor of the V_{s30} of the bottom known 30 m. We developed two versions of the model; one for profiles with constant shear wave velocity in the 30 m above the sensor, and one for varying profiles down to the sensor depth. Validation with the data showed that both predictions were unbiased, and led to similar results, especially when considering site classes for downhole sensors (Table 3), instead of an exact value of V_{s30}. Based on our study, the best predictor for the V_{s30} below the downhole sensor is V_S^B, the shear-wave velocity measurement just above the sensor.

3.6.3 Final Datasets

Records were only selected for the analyses if they were associated with:

- Stations that recorded a minimum of five events
- Events that were recorded at a minimum of five stations
- Both surface and borehole records were available at a given station for a specific event

These additional selection criteria were imposed after the record processing to ensure that they were strictly enforced. The last requirement never reduced the dataset since each of the selected events was recorded at a minimum of 10 stations. This led to a dataset of 43098 horizontal record pairs from 716 events and 530 stations (Figures 5-7). Figure 5 shows that the dataset is well populated for the selected magnitude and distance ranges. Figure 6 shows the distribution of stations relative to the respective site parameters for surface and borehole basement locations. The surface stations tend to be located on sites with V_{s30} mostly in the 300-700 m/s range (NERHP site categories B-C) while the borehole basement have local shear wave velocities somewhat uniformly distributed in a broad 700-2700 m/s range. Figure 7 shows the distribution of PGA with magnitude, distance and shear wave velocity as well as a scatterplot of PGA at the surface (PGA^G) and borehole depth (PGA^B). The trends are as expected given the site conditions illustrated on Figure 6.

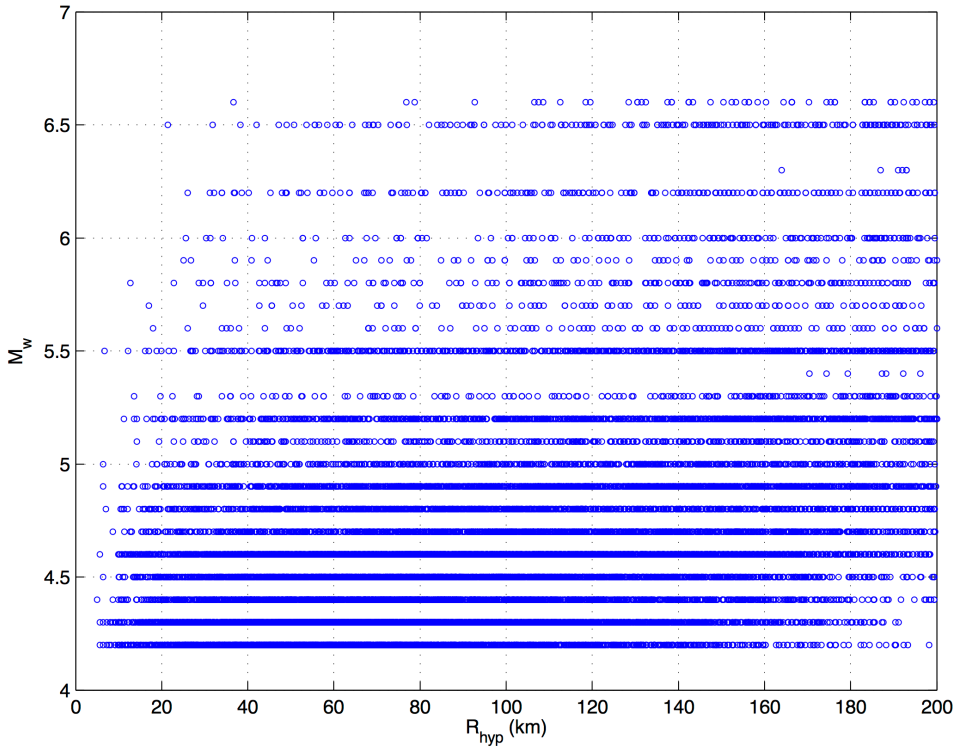


Figure 5: Distribution of moment magnitude (M_w) hypocentral distance (R_{hyp}) for the final complete dataset.

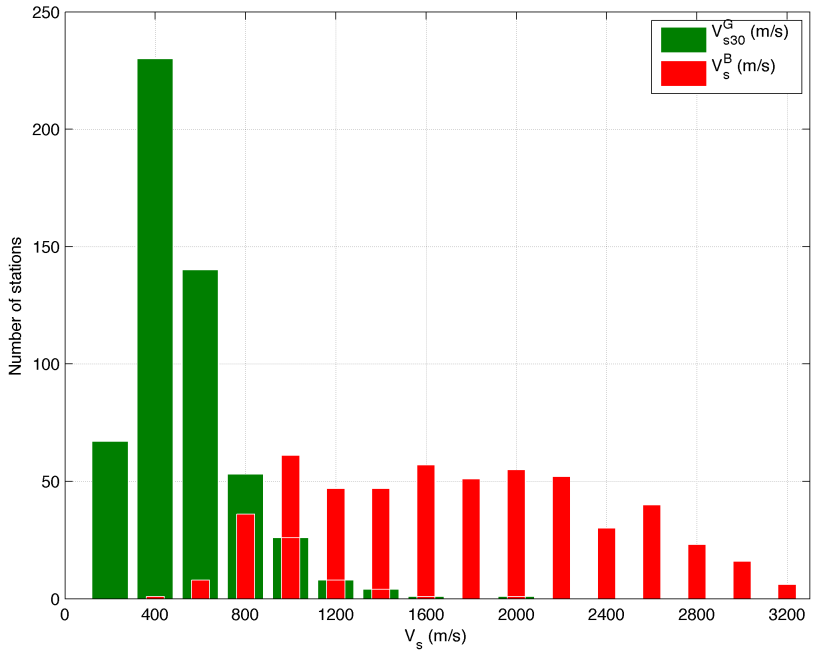


Figure 6: Distribution of stations according to shear wave velocity at the surface ($V_{s30}{}^{G}$) and at the borehole basement ($V_s{}^{B}$) for the final complete dataset.

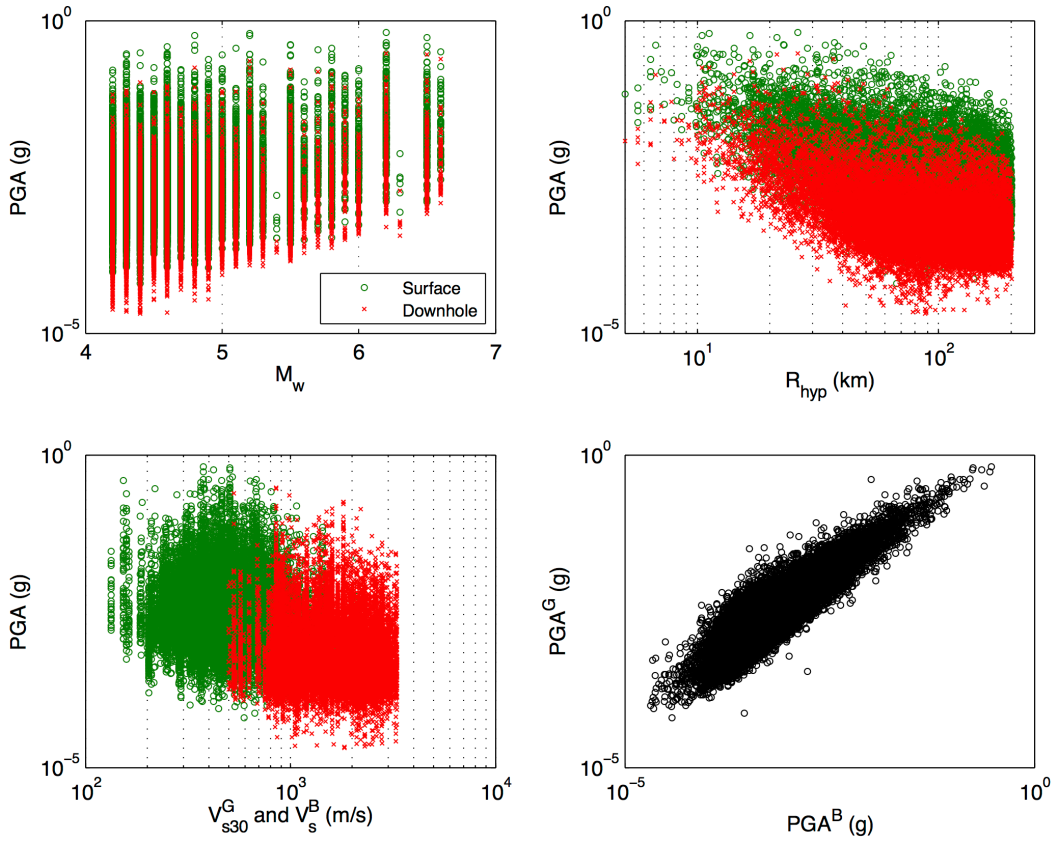


Figure 7: PGA distribution for surface and downhole (superscripts G and B respectively) against moment magnitude (M_w), hypocentral distance (R_{hyp}) and shear wave velocity site parameters for surface and borehole basement locations (V_{s30}^{G} and V_s^{B}).

3.7. Remaining Known Data Issues

As mentioned earlier, in the original scope of the project, we intended to use the record database developed by Zhao et al. (2006). However, further investigation of the dataset showed that it was not adequate for the specific objectives of this project. The development of a new dataset from raw KiK-net data became a major undertaking and choices to simplify the process had to be made. This section provides information on the remaining known data issues we believe should be addressed or formally considered in future work.

KiK-net has the advantage to provide a very large amount of data shortly after each event, but the quality of the data tends to vary. All data files contain raw, unfiltered recordings and many exhibit clipping, low signal-to-noise ratios and/or baseline errors. Therefore it is essential to carefully select records and process them so that they can be used in analyses. The nature of the present study requires analyzing a vast subset of recordings and manual record selection and filtering was not possible; batch data processing was performed instead as detailed in the previous sections. The use of a generic corner frequency is not desirable and can lead to bias for all IMs. Because the records were not systematically reviewed one by one, problem records that are clipped or that have a low signal-to-noise ratio, especially for small events at large distances, where included in the analyses. This may lead to systematic biases in results, especially at long periods. This issue is further discussed in Boore and Atkinson (2008).

Another issue with the KiK-net data is the combination of multiple events within single recordings, or when files associated with a single event contain arrivals from several subsequent sub-events. In some cases the sub-events are visibly separated from the main shock (Figure 8), while in others the sub-events are not distinguishable and therefore cannot be removed manually via truncation (Strasser and Bommer, 2005). Strasser and Bommer (2005) investigated multiple methods to separate the records to retain only the main event but a satisfactory automated procedure was not found. Because of these difficulties combined with the volume of files, records were not screened for sub-events. We believe that peak ground acceleration values (which are taken directly from the acceleration time series) are not likely to be affected by sub-events that have smaller amplitudes than the main shock. Spectral accelerations (which relate to the peak response of a SDOF oscillator subjected to the entire acceleration time series, including sub-events) could be affected by the presence of sub-events, but because *Sa* are also peak values, the effect of smaller amplitude sub-events may not always be critical. Nonetheless, it would be preferable to remove sub-events from the recordings to prevent potential bias of the spectral values. The presence of sub-events would be most important for cumulative intensity measures such as cumulative absolute velocity (CAV) and Arias Intensity (IA). Early on, we have considered CAV as an additional IM, but the presence of sub-events made us reconsider this option.

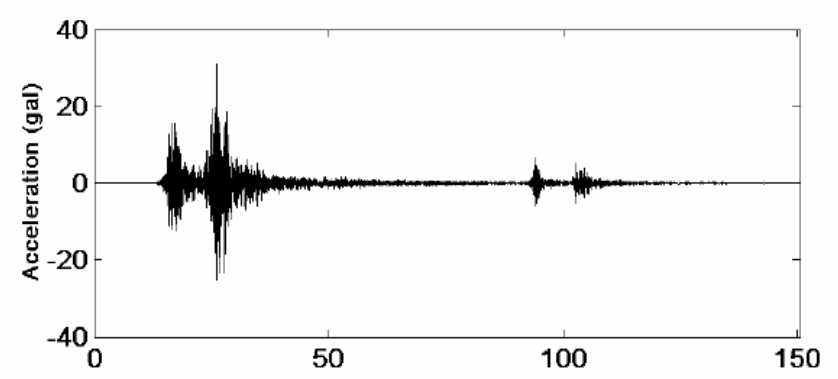


Figure 4: Example acceleration time series showing the arrival of separate sub-events (Source: Figure 3.5 from Strasser and Bommer, 2005)

We developed a semi-automatic code to address the issues or processing and selection described above; we hope to refine the code and use it in a follow-up task (Section 5.2).

Issues involving a lack of meta-data also exist within the KiK-net database. As discussed in Section 3.2, we only consider shallow crustal events to maximize the portability of results to models currently used in California. However, the source type and focal mechanism which require the intervention of a seismologist, are not provided in the database, and the source depth parameter estimates are rough and could be further investigated (John Zhao, personal communication).

Finally, we acknowledge issues in the KiK-net database involving the affect of down-going waves on downhole recordings. Waves reflected from the free surface or from interfaces with strong impedance contrasts are included in the borehole records, as discussed by Safak (1997) and Parolai et al. (2009). The down-going waves exhibit a strong degree of complexity but the effect of their contamination is expected to be of minor amplitude in comparison to the site amplification effects of many stations (Oth et al., 2011). We have not formally investigated these effects.

4. Effect of Site Amplification on Standard Deviation – Direct Computation from Records

4.1 Computation of ϕ_{Amp}

This section summarizes the computation and results of the effect of site amplification on the within-event standard deviation, ϕ_{Amp}. Standard deviations defined in GMPEs are usually computed based on the standard deviation of the residuals of said GMPE. The computation of ϕ_{Amp} is conceptually different and is computed directly from the recorded data, and not in relation to a specific GMPE. This set of computations requires collocated records for each event at each station.

ϕ_{Amp} is an IM-dependent parameter. In this study the IMs are PGA and Sa at various periods (Section 3.5). For each IM, ϕ_{Amp} is computed as follows:

1. Compute the amplification for station *s*, earthquake *e*, Amp_{es}:

$$\mathrm{Amp_{es}} = \left[\ln\left(\mathrm{IM_{es}^{G}}\right) - \ln\left(\mathrm{IM_{es}^{B}}\right)\right] \tag{4.1}$$

for *s*=1:NStations and *e*=1:NEqs

2. Compute the mean amplification $\overline{\mathrm{Amp}}_{\mathrm{s}}$ at each station s:

$$\overline{\mathrm{Amp}}_{\mathrm{s}} = \frac{1}{\mathrm{NEqs_s}} \sum_{\mathrm{e}=1}^{\mathrm{NEqs}_s} \left[\ln\left(\mathrm{IM_{e}^{G}}\right) - \ln\left(\mathrm{IM_{e}^{B}}\right)\right] \tag{4.2}$$

3. Compute the residuals of amplification from each earthquake at each station, $\delta\mathrm{Amp_{es}}$:

$$\delta\mathrm{Amp_{es}} = \mathrm{Amp_{es}} - \overline{\mathrm{Amp}}_{\mathrm{s}} \tag{4.3}$$

4. Compute the standard deviation of the residuals for the sample

$$\phi_{\mathrm{Amp}} = \sqrt{\sum_{\mathrm{s}=1}^{\mathrm{NStations}} \sum_{\mathrm{e}=1}^{\mathrm{NEqs}_s} (\mathrm{Amp_{es}} - \overline{\mathrm{Amp}}_{\mathrm{s}})^2 \Big/ \left[\left(\sum_{s=1}^{NStations} \mathrm{NEqs}_s\right) - 1\right]} \tag{4.4}$$

This computation procedure gives all the records in the dataset equal weights. This may lead to bias due to uneven number of recordings at the different stations, attributing more weight to stations with many recordings relative to stations with a smaller number of recordings.

Alternatively, all the stations can be weighted equally and ϕ_{Amp} can be obtained by replacing step 4 from above by:

4. Compute the standard deviation of the residuals for a given station

$$\phi_{\mathrm{Amp}_s} = \sqrt{\sum_{e=1}^{\mathrm{NEqs}} (\mathrm{Amp}_{es} - \overline{\mathrm{Amp}}_s)^2 / (\mathrm{NEqs}_s - 1)} \tag{4.5}$$

5. Compute the average standard deviation

$$\phi_{\mathrm{Amp}} = \frac{1}{\mathrm{NStations}} \sum_{s=1}^{\mathrm{NStations}} \phi_{\mathrm{Amp}_s} \tag{4.6}$$

If there are many records at each station, both weighting procedures should lead to similar results.

4.2 Summary of Results for the Final Dataset

Figures 9 and 10 illustrate results from both weighting schemes above for PGA. Figure 9 shows the results associated with equations 4.3 (blue circles) and 4.4 (red lines for +/- ϕ_{Amp}) for PGA. Figure 10 shows the results associated with equations 4.5 (blue circles) and 4.6 (red line for ϕ_{Amp}).

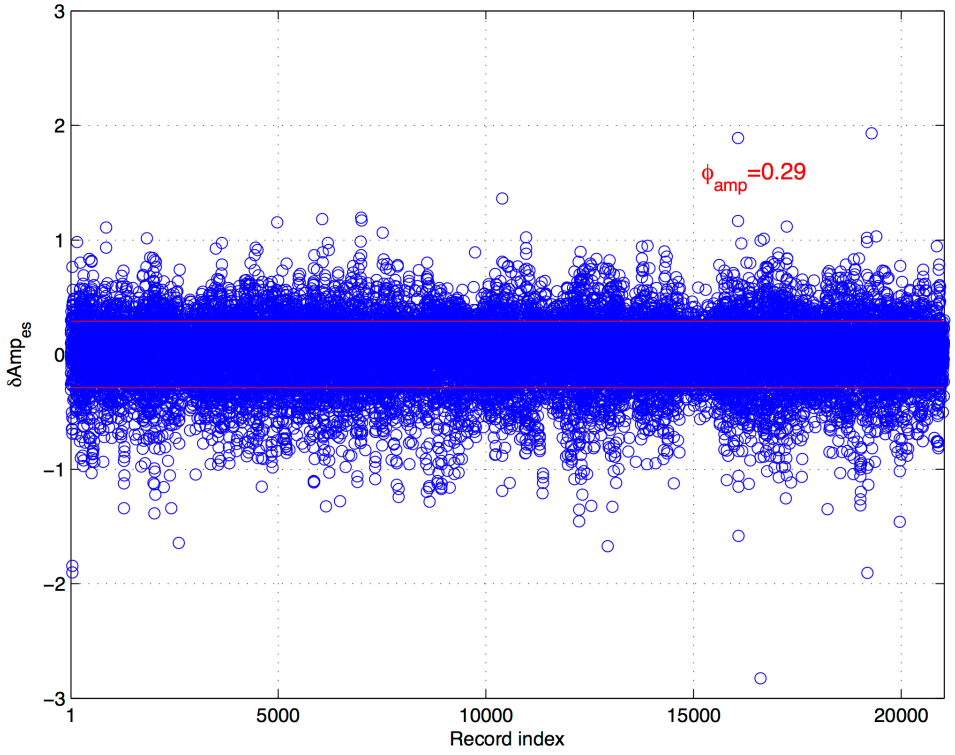


Figure 9: Sample amplification residuals and standard deviation (PGA), reflecting results from equations 4.3 (blue circles) and 4.4 (red lines for +/- ϕ_{Amp}). Each blue circle corresponds to a single record for event *e*, station *s*.

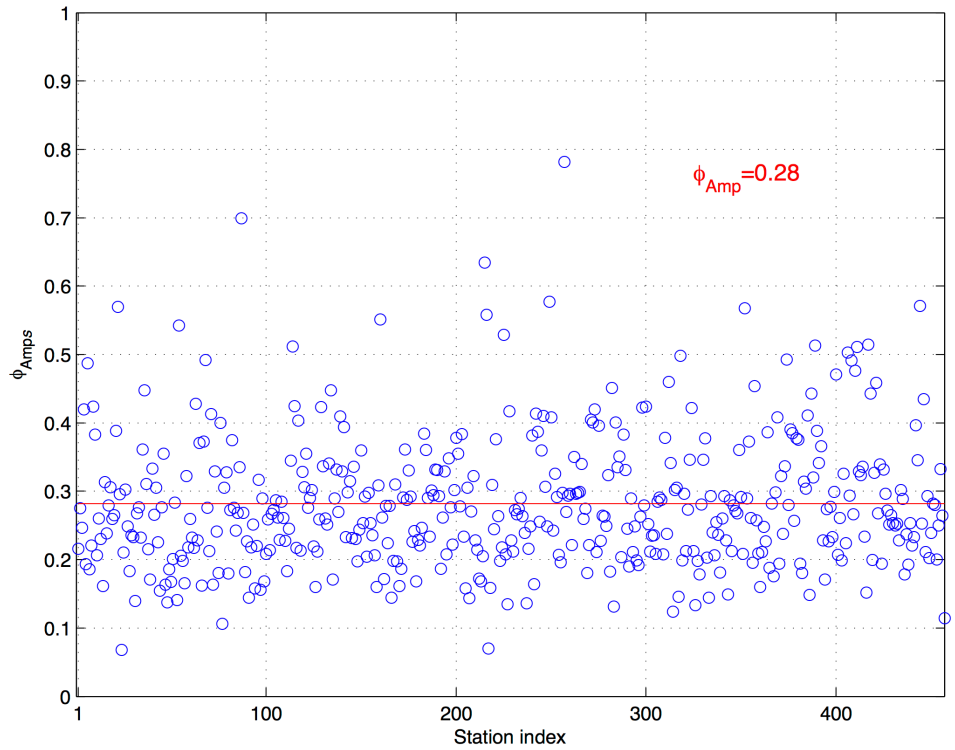


Figure 10: Sample amplification residuals and standard deviation (PGA), reflecting results from equations 4.5 (blue circles) and 4.6 (red line for ϕ_{Amp}). Each blue circle corresponds to ϕ_{Amp_s} for station *s* while the red line is the mean (ϕ_{Amp}).

Results compiled for other IMs are summarized in Table 2 and Figure 11. Results for spectral periods larger than 1.0 s have been greyed out and are not deemed reliable (see discussion in Section 3.5). The results from both weighting schemes (equations 4.4 and 4.6) are found to be statistically equivalent.

Table 2 – Summary of computed ϕ_{Amp} values

IM	ϕ_{Amp} Equation 4.4	ϕ_{Amp} Equation 4.6
PGA	0.29	0.28
Sa(0.01s)	0.29	0.28
Sa(0.02s)	0.3	0.29
Sa(0.03s)	0.3	0.29
Sa(0.05s)	0.3	0.29
Sa(0.1s)	0.25	0.25
Sa(0.2s)	0.23	0.22
Sa(0.3s)	0.22	0.21
Sa(0.5s)	0.21	0.2
Sa(0.6s)	0.21	0.2
Sa(1s)	0.23	0.22
Sa(1.4s)	0.26	0.25
Sa(2s)	0.29	0.27
Sa(3s)	0.31	0.3

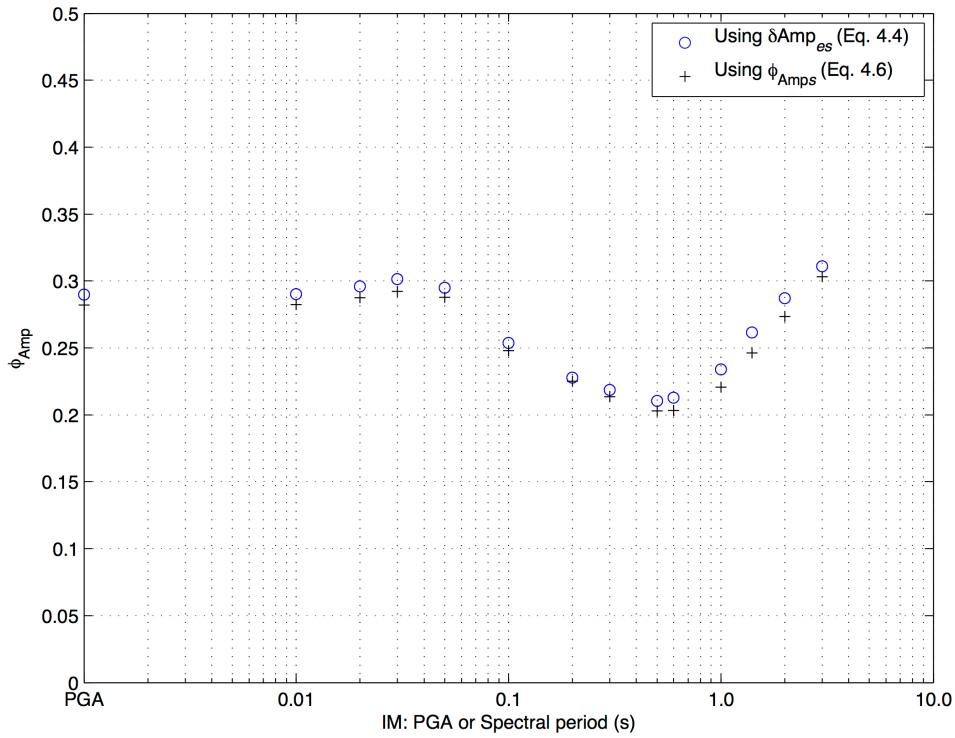


Figure 11: Summary results for ϕ_{Amp} for both weighting schemes described above (blue circles from equation 4.4 and black crosses from equation 4.6).

4.3 V_{s30} DEPENDENCE OF ϕ_{Amp}

Figure 12 shows sample results from equation 4.4 plotted against V_{s30}^{G}. The level of dispersion with the selected IMs is consistent with the ϕ_{Amp} shown on Figure 11. Furthermore, for a given IM, there is an apparent reduction in dispersion with increased site stiffness (V_{s30}^{G}) and this effect is more pronounced for longer period IMs.

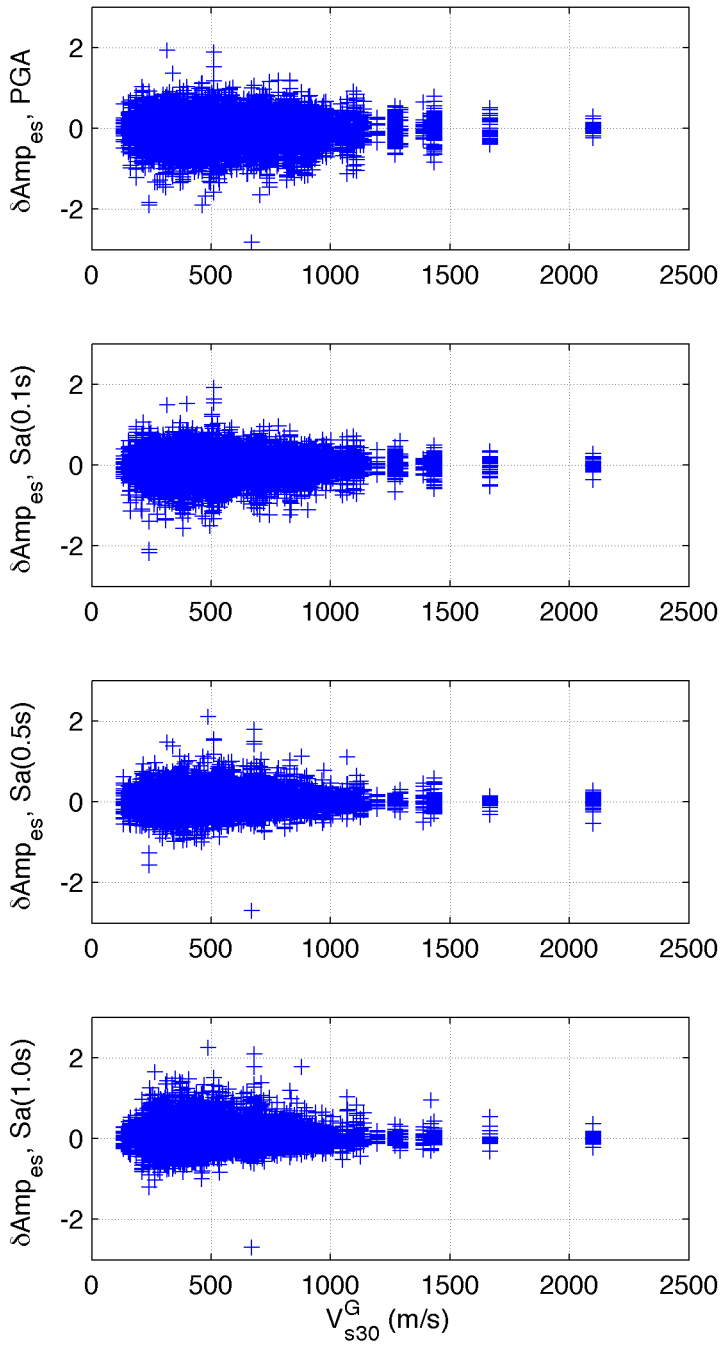


Figure 12: Summary results for ϕ_{Amp} (blue circles from equation 4.4) as a function of V_{s30}^{G} for four sample IMs.

Because of the data distribution (Figure 6) and the relatively poorly/unevenly populated database for V_{s30}^{G} values above 1000 m/s, it seems to make sense to aggregate the results per broad site categories. The NEHRP site categories are a convenient choice (Table 3).

Table 3 – NEHRP site categories (BSSC 2001)

NEHRP Category	Description	Mean V_{s30}^{G} (m/s)
A	Hard rock	> 1500
B	Firm to hard rock	760-1500
C	Dense soil, soft rock	360-760
D	Stiff soil	180-360
E	Soft clays	<180
F	Special study soils	-

Categories E and F are also associated with other criteria that require site-specific knowledge at each station (BSSC, 2001). This type of data is not available for the KiK-net stations and for simplicity, the classification is done solely on the basis of V_{s30}^{G}, for categories A-E. Figure 13 shows a sample plot of the grouped data and the ϕ_{Amp} corresponding to each site category, for PGA.

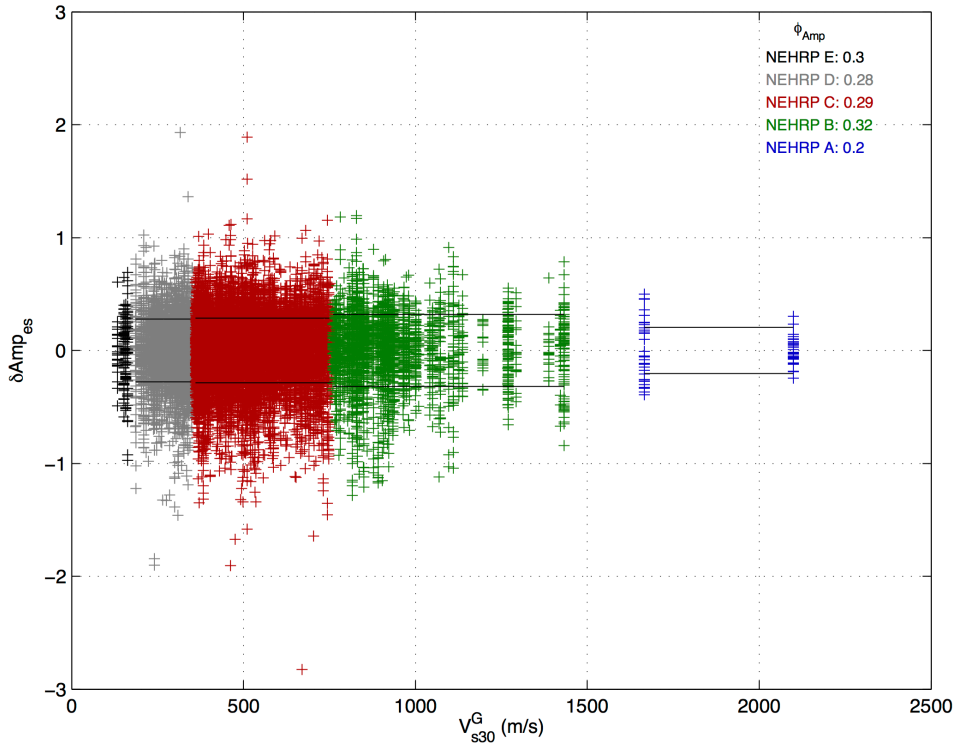


Figure 13: Summary results for ϕ_{Amp} for each NEHRP site class, PGA.

The ϕ_{Amp} for site class A is shown for completeness, but because it is based on only two stations the results are not carried through. The results for all selected IMs are summarized in Table 4 and Figure 14. The most important trend remains with the spectral period, as illustrated in Figure 11. The cross-over of ϕ_{Amp} values for categories C and D around the stiffer category B with increasing spectral period seem to indicate a dependence of ϕ_{Amp} on site non-linearity (Figure 14). The ϕ_{Amp} values associated with site category E appear quite erratic and more insight could be gained with a finer spectral period range definition. There is no practical implication of this behavior for site in category E: the amplification for

these sites is usually assessed through site-specific studies and these sites are not used as reference conditions.

Table 4 – Summary of computed ϕ_{Amp} values per NEHRP category

IM	B	C	D	E
PGA	0.32	0.29	0.28	0.3
Sa(0.01s)	0.32	0.29	0.28	0.31
Sa(0.02s)	0.31	0.29	0.3	0.35
Sa(0.03s)	0.32	0.29	0.32	0.35
Sa(0.05s)	0.3	0.29	0.31	0.4
Sa(0.1s)	0.24	0.26	0.26	0.26
Sa(0.2s)	0.22	0.23	0.23	0.26
Sa(0.3s)	0.19	0.22	0.22	0.27
Sa(0.5s)	0.17	0.21	0.22	0.23
Sa(0.6s)	0.17	0.21	0.24	0.2
Sa(1s)	0.17	0.24	0.26	0.2
Sa(1.4s)	0.19	0.26	0.29	0.25
Sa(2s)	0.21	0.29	0.32	0.34
Sa(3s)	0.24	0.31	0.35	0.4

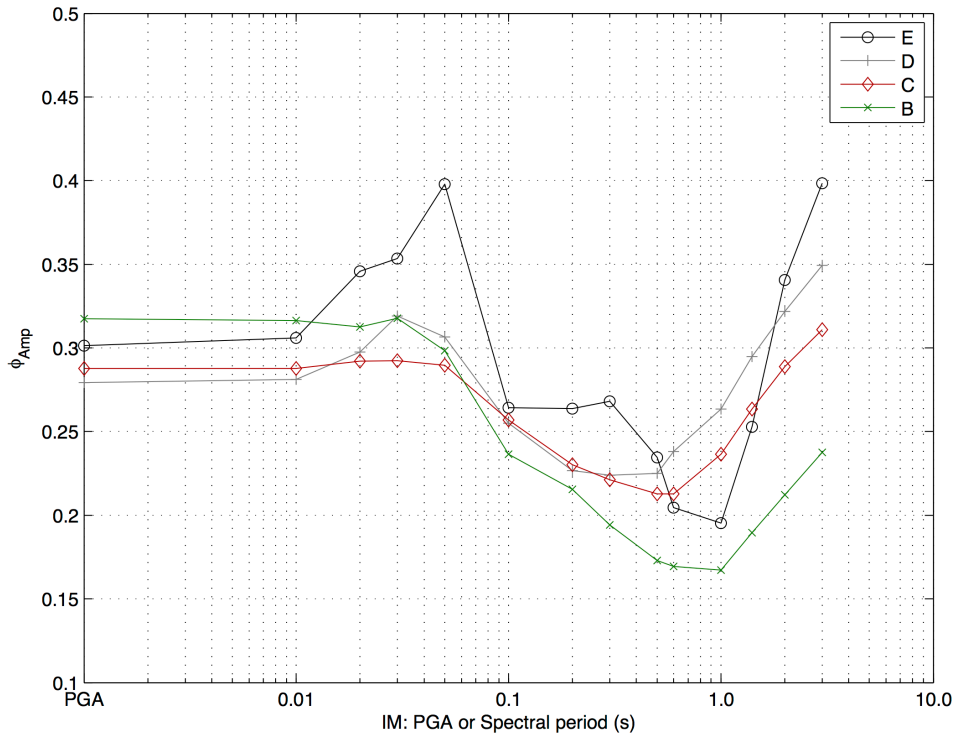


Figure 14: Summary results for ϕ_{Amp} for each NEHRP site class vs. IM.

5. SUMMARY, DISCUSSION AND FUTURE WORK

5.1. SUMMARY

In the present study we have investigated the effect of site amplification on the standard deviation of ground motions, referred to as ϕ_{Amp}. ϕ_{Amp} is the controlling factor in the quantification of dispersion when the site-related ergodic assumption is taken out either through site-specific V_{s30} or with truly site-specific ground motion assessments involving GRAs. We have used the unique opportunity offered by the KiK-net database to directly compute empirical ϕ_{Amp} for different IMs. We have shown that ϕ_{Amp} is IM-dependent and that it also seems to depend on site stiffness, to a certain extent.

Recall from Section 2 that at least two modern GMPEs (Abrahamson and Silva, 2008; Campbell and Bozorgnia, 2008) use a value of 0.3 for ϕ_{Amp}. This 0.3 value comes from site response analyses using PGA to determine the amplitude of motions for quantifying nonlinear effects. The current work indicates that the previously used values of ϕ_{Amp}=0.3 are consistent with ϕ_{Amp} computed directly from the KiK-net data (Table 2) if the site effects are based on reference PGA estimates. However, if site effects are computed from other IMs (e.g. directly from period-dependent *Sa* on reference conditions), then ϕ_{Amp} should include an IM-dependence. In both cases, since ϕ_{Amp} is used to remove the semi-ergodic assumption for different reference conditions (especially for site-specific analyses where the stiffness at the base of the soil columns varies on a site-specific basis), considering the ϕ_{Amp} stiffness-dependence, would provide an improvement over the single value currently used. This could be done at least in a broad sense (e.g. following broad categories of V_{s30}^{G}). We believe that a systematic reprocessing of the data would allow better constraints on ϕ_{Amp} for a wider spectral period range.

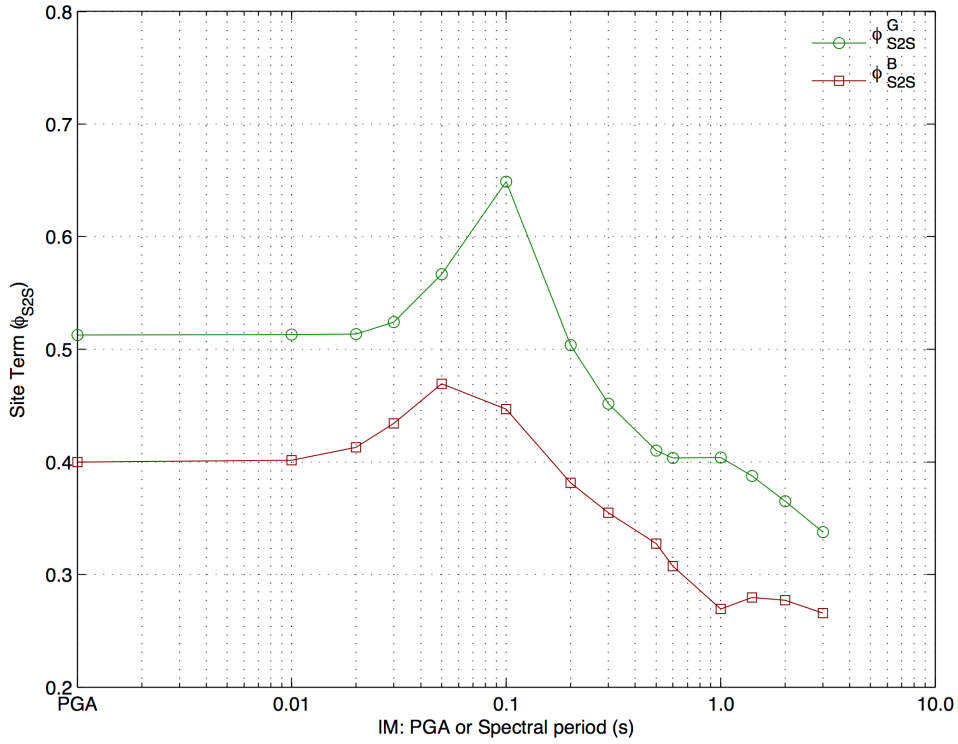
5.2 DISCUSSION AND LIMITATIONS

Many technical issues were identified during the course of the project. We moved forward and applied the whole methodology in spite of these issues and devised a plan to address the problems in future tasks. The preliminary results, due to their potential direct impact on PSHA results (Section 2.1), indicate that there is a value to further pursue the question of the reference rock standard deviation, after some critical issues are properly addressed. These issues are summarized below.

- The preliminary results are subject to data quality. The batch data collection and filtering methods applied most undeniably resulted in the inclusion of several to many poor quality recordings, some of which may include sub-events, and a certain number of poorly processed recordings. Further assumptions had to be made regarding the tectonic environment for each event. These data issues have the potential to affect the validity of the results summarized above. Careful record processing and selection can remediate to many of these issues, and should be done on a record-to-record basis. Unfortunately, due to unexpected complications and limited resources, this was not possible for this project.

- The results presented here serve as a motivation for further study, in particular to further develop a robust semi-automatic processing set of codes. Processed KiK-net records would be useful to revisit the work summarized in this report and to confirm or infirm the preliminary findings. A quality database of processed KiK-net records would be extremely useful beyond the current application and would be a great tool for the engineering and seismology research communities, both nationally and worldwide.

- One issue to consider is the applicability of the Japanese site effects to other active regions such as California. In a related but undocumented task for this project, we have developed GMPEs modified specifically to assess the dispersion of residuals for the KiK-net dataset. Through this work, we have computed site terms from the surface and borehole records, using two different GMPEs. These site terms represent the overall site effects from all the recording stations for the KiK-net database. The site terms were very similar for both GMPEs. We have found that the Japanese data tend to show a large site-response near 0.1s (Figure 15), possibly a first-mode response. This type of peak response appears consistent with the V_s distribution (Figure 6) and the relatively large impedance contrast present at many stations. It is unclear at this point how this type of site response controls the observed ϕ_{Amp} and its dependence on spectral period. It is also unclear if this type of site effect is significant enough to prevent the transportability of results to other active regions (California, for example, is known for generally smoother vertical V_s transitions). The fact that ϕ_{Amp} computed here for PGA was found to be similar with other published values offers somewhat of a validation of the results for this specific IM.



- **Figure 15: Summary of site terms (ϕ_{S2S}) for the KiK-net data vs. IM. Superscripts G for ground surface and B for borehole.**

5.2. FUTURE WORK

Keeping abreast of KiK-net network expansion

We have shown that ϕ_{Amp} is IM-dependent and that it also seems to depend on site stiffness to a certain extent. It would be interesting to see the true effect of hard rock amplification on ϕ_{Amp}, but there are not many stations located on stiff sites. It is important to follow the evolution of the KiK-net network so as to include additional stations located on hard rock.

Systematic data processing

Several data issues summarized in Section 3.6 can be addressed by reprocessing the data systematically using a semi-automatic computer code. Tremendous effort has already gone into the development of a detailed protocol and its associated MatLab code, which is an extension of the one used in the current study. The code utilizes a hybrid approach for processing large quantities of records.

In this semi-automatic procedure, the processing steps described in Section 3.4 are repeated to compute filtered time series over a range of possible low-cut corner frequencies. The optimal corner frequency is determined by a combination of observations that include: a formal assessment of the signal-to-noise ratio using Fourier spectra, the judgment of where the low frequency portion of the FAS deviates from the gradient of the inverse of frequency, and the inspection of velocity and displacement time series (Boore and Bommer, 2005). After sweeping through many frequencies, the code automatically selects a candidate corner frequency and the user can either accept the suggested value or enter another value, based on the visual inspection of the summary plots (similar to Figure 4). Filter parameters can be made consistent for all components from a particular station to ensure that the empirical amplification factor is not affected by the ground motion correction procedure. The processed time series and their associated plots are saved as individual files and the filter information is saved in a database for easy retrieval.

Another issue with the KiK-net data is the occurrence of multiple events within single recordings. Although we are not aware of a method for distinguishing concurrent events, records with sequential events can easily be identified by visual inspection at processing time. In the current version of the code, the records can't be truncated to remove the extra event, but initial flagging of potential problem records could be included with minimal effort.

Other Tasks to Consider

Soil nonlinearity is somewhat included in the results presented above, but it would be important to formally assess the effect of nonlinear response on the computed ϕ_{Amp}.

It would be important to further investigate the impact of reflected waves in the downhole records by a quantitative assessment of the effect (as a percentage of effect on ground motions) and the identification of the frequency range or band at which and the effect is distinguishable.

Finally, we would recommend to repeat the analysis with a finer spectral period spacing to better assess the correlation of trends with spectral period.

6. ACKNOWLEDGEMENTS

We would like to acknowledge the support of the USGS NEHRP External Research Support Program and thank the panel of reviewers who provided useful comments on the original proposal. We would also like to thank the following collaborators who either shared data, opinions or otherwise offered support to this project: Dr. John Zhao, Dr. Kenneth Campbell, Dr. Norman Abrahamson, Dr. Jonathan Stewart, Dr. Adrian Rodriguez-Marek, Dr. Paul Somerville and Dr. David Boore. Finally, we would like to acknowledge the contribution of the Japanese National Research Institute for Earth Science and Disaster Prevention (NIED) for making the KiK-net data available online.

7. REFERENCES

Abrahamson, N. A., and Youngs, R. R, (1992). "A stable algorithm for regression analyses using the random effects model", *Bull. Seism. Soc. Am.*, 82, 505–510.

Abrahamson, N.A. and Silva, W.J. (2008). "Summary of the Abrahamson and Silva NGA ground motion relations", *Earthquake Spectra*, 24, 45-66.

Abrahamson, N.A. and Silva, W.J. (2007). "Abrahamson & Silva NGA Ground Motion Relations for the Geometric Mean Horizontal Component of Peak and Spectral Ground Motion Parameters", Unpublished draft report to be submitted to the Pacific Earthquake Engineering Research Center.

Aoi, S., Kunugi, T., Fujiwara, H. (2004). "Strong-Motion Seismograph Network Operated by NIED: K-Net and KiK-net", *Journal of Japan Association for Earthquake Engineering.* 4 (3)

Bayless and Goulet (2011). "Methods for predicting shear wave velocity below a downhole sensor at KiK-net digital strong motion seismograph stations", Earthquake Engineering Research Institute (EERI) Annual Meeting, La Jolla, California, February 9-12, 2011. *Poster Presentation.*

Baturay, M.B. and Stewart, J.P. (2003). "Uncertainty and bias in ground motion estimates from ground response analyses", *Bull. Seism. Soc. Am.*, 93 (5), 2025-2042.

Bazzurro, P. and C.A. Cornell (2004). "Nonlinear soil-site effects in probabilistic seismic-hazard analysis", *Bull. Seism. Soc. Am*., 94, 2110 -2123.

Boore D.M. and Bommer J.J. (2005). "Processing of strong-motion accelerograms: needs, options and consequences", *Soil Dyn. Earthq. Eng.*,25, 93-115, ISSN:0267-7261

Boore, D. M. (2005). "On pads and filters: Processing strong-motion data", *Bull. Seism. Soc. Am*. 95, 745--750.

Boore, D.M. and Atkinson, G.M. (2008). "Ground motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 and 10.0 s", *Earthquake Spectra*, 24. 99-138.

Boore, D.M., C.D. Stephens, and W.B. Joyner (2002). "Comments on baseline correction of digital strong-motion data: Examples from the 1999 Hector Mine, California, earthquake", *Bull. Seism. Soc. Am*. 92, 1543--1560.

Building Seismic Safety Council (BSSC). (2001). "NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures, Part 1: Provisions and Part 2: Commentary", Federal Emergency Management Agency, FEMA-368 and FEMA-369, Washington D.C.

Campbell, K.W. and Bozorgnia, Y. (2008). "NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD, and 5%-damped linear elastic response spectra for periods ranging from 0.01 to 10 s", *Earthquake Spectra*, 24, 139-171.

Chiou, B.S.-J. and Youngs, R.R. (2008). "Chiou and Youngs PEER-NGA empirical ground motion model for the average horizontal component of peak acceleration and pseudo-spectral acceleration for spectral periods of 0.01 to 10 seconds", *Earthquake Spectra*, 24, 173-215.

Douglas, J. and Boore, D.M. (2011). "High-frequency filtering of strong-motion records", *Bull. Earthquake Engineering,* 9, 395-409.

Fujiwara, H., S. Aoi, T. Kunugi & S. Adachi (2004). "Strong-motion observations of NIED: K-NET and KiK-NET. Proceedings of the COSMOS Workshop on Record Processing Guidelines", Richmond, California, 26-27 May 2004.

Fukushima, Y. (1996). "Scaling relations for Strong Ground Motion Prediction Models with M^2 terms", *Bull. Seism. Soc. Am.* 86 (2): 329-336

Goulet, C.A. (2008). "Improving the Characterization of Seismic Hazard for Performance-Based Earthquake Engineering Design", Doctoral dissertation. Department of Civil and Environmental Engineering, University of California, Los Angeles, CA.

Goulet, C.A. and Stewart J.P. (2009). “Pitfalls of Deterministic Application of Nonlinear Site Factors in Probabilistic Assessment of Ground Motions”, *Earthquake Spectra* 25, 541–555.

Oth, A., S. Parolai and D. Bindi (2011). “Spectral analysis of K-NET and KiK-net data in Japan, Part I: Database compilation and peculiarities”, *Bull. Seismol. Soc. Am.*, 101(2), 652-666.

Parolai, S. (2009). “Denoising of seismograms using the S transform”, *Bull. Seismol. Soc. Am.* 99, 226-234.

Safak, E. (1997). “Models and methods to characterize site amplification from a pair of records”, *Earthquake Spectra* 13, 97-129.

Silva, W. J. (2008). “Site Response Simulations for the NGA project”, Unpublished draft report to be submitted to the Pacific Earthquake Engineering Research Center.

Strasser, F.O. and Bommer, J.J. (2005). “Analysis of intra-event ground-motion residuals from K-NET and KiK-NET data”, Department of Civil and Environmental Engineering, Imperial College London

Zhao J.X., Zhang J., Asano A., Ohno Y., Oouchi T., Takahashi T., Ogawa H., Irikura K., Thio, H.K., Somerville P.G., Fukushima Y. (2006). “Attenuation relations of strong ground motion in Japan using site classification based on predominant period”, *Bull. Seism. Soc. Am.* 96, 914–925