Using Bayesian regression to develop joint Fourier domain GMPEs for different tectonic regimes

Mahdi Bahrampouri, PhD Student, Virginia Tech
Adrian Rodriguez-Marek, Professor, Virginia Tech

Background and motivation

The scaling of ground motion intensities with magnitude, distance, and site parameters may not be the same for different tectonic regimes. In the literature of developing GMPEs, the difference in tectonic regime is dealt with one of two options:

- **A scaling factor** (e.g., Foulser-Piggott and Goda 2015). This method assumes different tectonic regimes have similar scaling for magnitude, distance, and site parameters, and only differ by a constant term (i.e., different intercept in logarithmic domain). This assumptions have not been rigorously proven.
- **Separate set of coefficients for each tectonic regime** (e.g., Bullock et al. 2017). While estimating separate set of coefficients for each tectonic regime is theoretically robust, it splits the database into different groups and reduces stability of some parameters. For instance, the number of ground motions that cause nonlinear site response is limited and splitting the database into tectonic regimes reduces the data even further.

In this study, we use Bayesian regression to develop a joint Fourier domain GMPEs which has different scaling with magnitude, distance, and site parameters for each tectonic regime, if the difference is statistically significant. We choose modeling Fourier amplitude over pseudo spectral acceleration because there are several advantages to using Fourier spectra in place of response spectra:

1. The scaling of Fourier spectra in the GMPE is easier to constrain using seismological theory.
2. The ratio of Fourier amplitude at surface over bedrock is only dependent on site response and it does not depend on magnitude and distance, as does spectral acceleration. Therefore, we postulate that site terms in Fourier domain are not dependent on tectonic regime, magnitude, or distance.
3. Transferring GMPEs from host to target regions can be done using theoretical models.
4. The ability to do hazard in Fourier domain, which we could achieve only with correlations of duration.

Database

We have developed a database for ground motions recorded in Japan by the Kiban Kyoshin network (KiKnet). The database has all the earthquake in KiK-net website with magnitude larger than 3 since 1996 until the end of 2017. In addition to the information KiK-net provides for each event, we have enriched the database using moment tensor solution provided by F-net or, for some earthquake, literature source models. A variety of distance measures are computed for each ground motion including rupture distances for the event with F-net or literature source models. Each ground motion is processed using an automated algorithm. Several intensity measures (i.e., spectral acceleration, smoothed and down sampled Fourier amplitude, Arias Intensity, and duration) of the processed ground motions are presented in the database. Finally, the database has a site catalog based on shear wave velocity profiles provided by KiK-net. The events for each tectonic regime is presented in the Figure 1. The database is available at https://doi.org/10.17603/ds2-e0ts-c070.

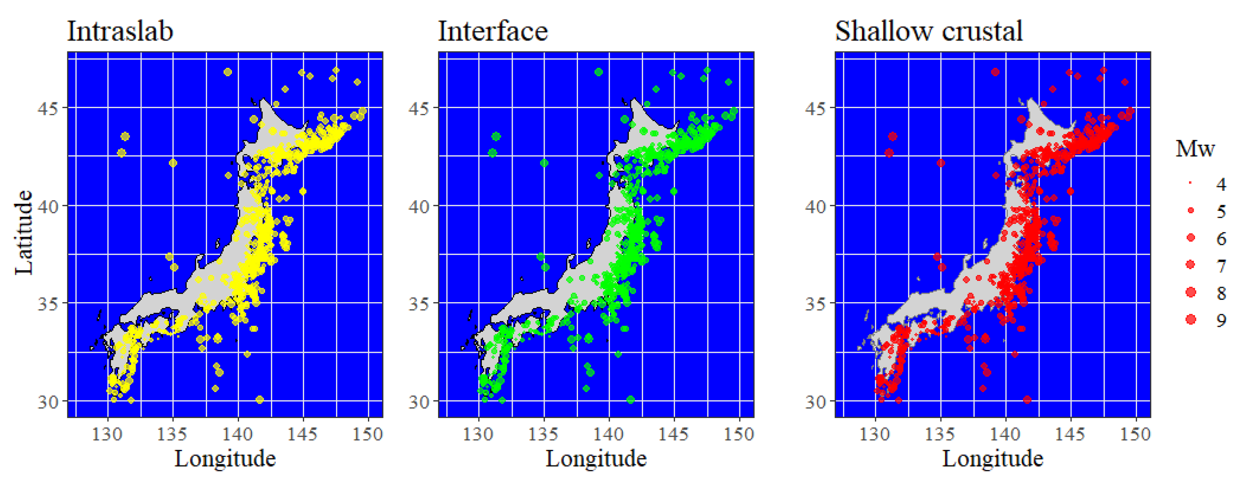

Figure 1. distribution of events in the database grouped by tectonic regime

Bayesian Regression

Generally, GMPEs are developed by finding the parameters or a set functional form that maximize the likelihood. In a Bayesian regression, instead of maximizing the likelihood function alone, we would assume prior distributions for the parameters and use Bayes theorem:

$$\boldsymbol{posterior} \propto \boldsymbol{likelihood} \times \boldsymbol{prior}$$

in which prior is the prior distribution for the parameters. Using a noninformative prior is the same as maximum likelihood regression. Using a prior with zero centered distribution will shrink the unnecessary parameters to zero. Table 1 shows alternative priors and justifies our choice of a horseshoe prior for the parameters that indicate the difference between tectonic regimes.

Table 1. property of each candidate prior and the effect on the regressed parameters

	Property	Consequence
Normal distribution	Large amount of probability around zero	unnecessary parameters become zero ✓
	Thin tail	Necessary parameters would be shrunken. Bias ✗
Laplace	Low amount of probability around zero	Unnecessary parameters will not be zero ✗
	Fat tail	Not biased
Horseshoe	Mass probability at zero	unnecessary parameters become zero ✓
	Fat tail	Not bias ✓

Functional Form

The functional form we use is:

Sh: 1 if shallow crustal and zero otherwise

Int: 1 if interface and zero otherwise

Red parameters have horseshow prior and black parameters have noninformative prior

$$f_M(f) = c_1 + (c_2 + c_{2sh} * sh + c_{2int} * int)\,(M-6) + (c_3 + c_{3sh} * sh + c_{3int} * int)\ln\left(1 + e^{10(c_M + c_{Msh} * sh + c_{Mint} * int - M)}\right)$$

$$f_{ztor} = (c_9 + c_{9sh} * \text{sh} + c_{9int} * \text{int}) Z_{tor\,d}$$

$$f_P = (c_4 + c_{4sh} * sh + c_{4int} * int)\log\left\{\sqrt{R_{RUP}^2 + h^2}\right\} + (c_7 + c_{7sh} * sh + c_{7int} * int) R_{RUP}$$

$$f_S = (c_8 + c_{8sh} * sh + c_{8int} * int)\frac{log\left(1 + \left(\frac{V_{sN}}{V_{s30}}\right)^5\right)}{5}$$

$$f_{Z1} = (c_{11} + c_{11sh} * sh + c_{11int} * int)\ln\left[\frac{\min(Z_1, 2) + 0.01}{Z_{1,ref} + 0.01}\right]$$

This functional form is adopted from Bayless and Abrahamson (2019) with three main modifications:

1. The site term scaling with V_{S30} saturation is frequency dependent. In Figure 5 we show why the saturation point should increase with frequency and how the data confirms that.

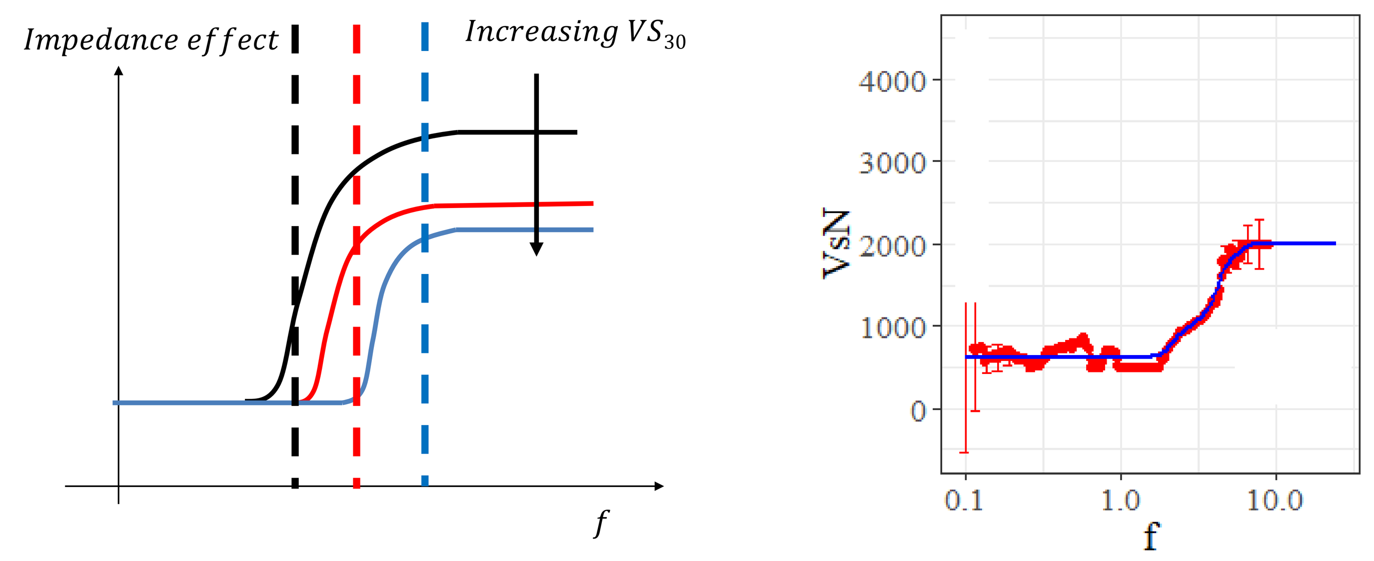

Figure 2. illustration of site term scaling with VS30 saturation. a) Schematic impedance effect versus frequency for different sites. b) estimated saturation point versus f

2. We don't assume the geometric attenuation is frequency independent. Although based on a point source model the geometric attenuation is frequency independent, the data shows lower geometric attenuation for lower frequencies. In Figure 3, the difference between shallow crustal model residuals with geometric frequency independent term and with geometric frequency dependent term.

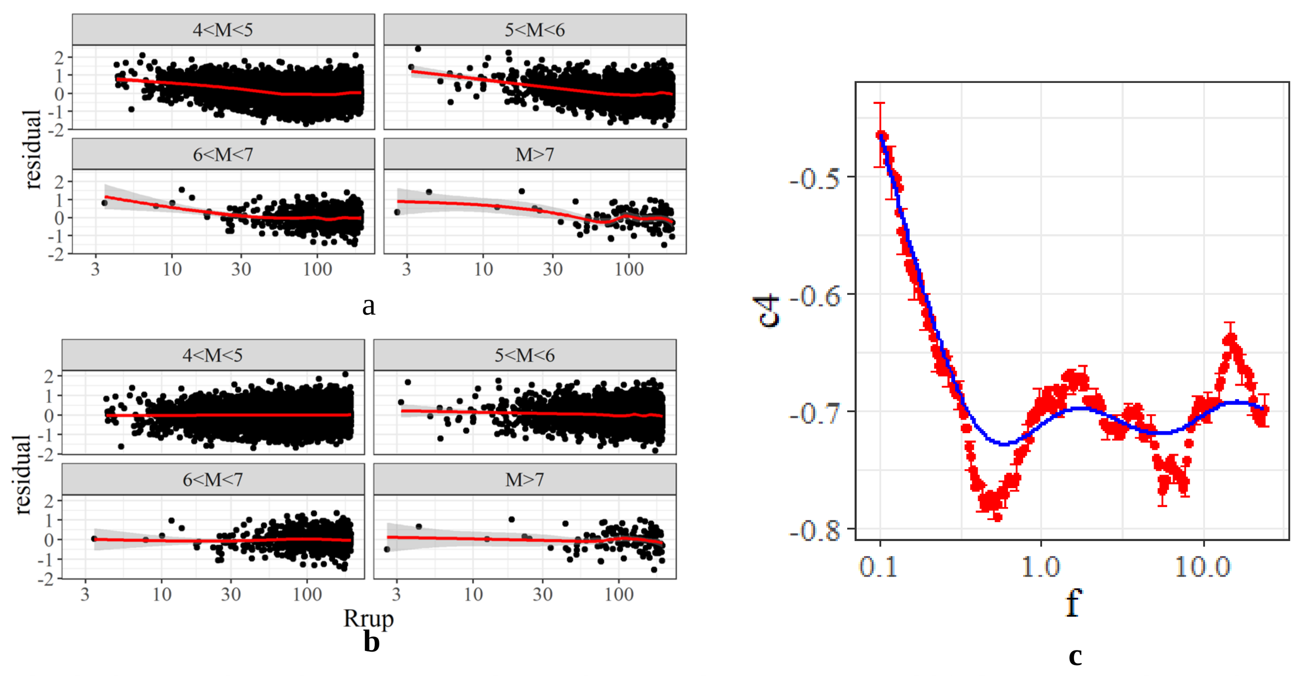

Figure 3: residual plots with a) frequency independent anelastic term and b) frequency dependent anelastic term for 0.6 Hz. Estimated anelastic attenuation term versus frequency for shallow crustal earthquakes.

3. We have fixed the κ_0 as a function of V_{S30} in the regression model to have a robust extrapolation after 25 Hz and before 0.1 is illustrated in Figure 4.

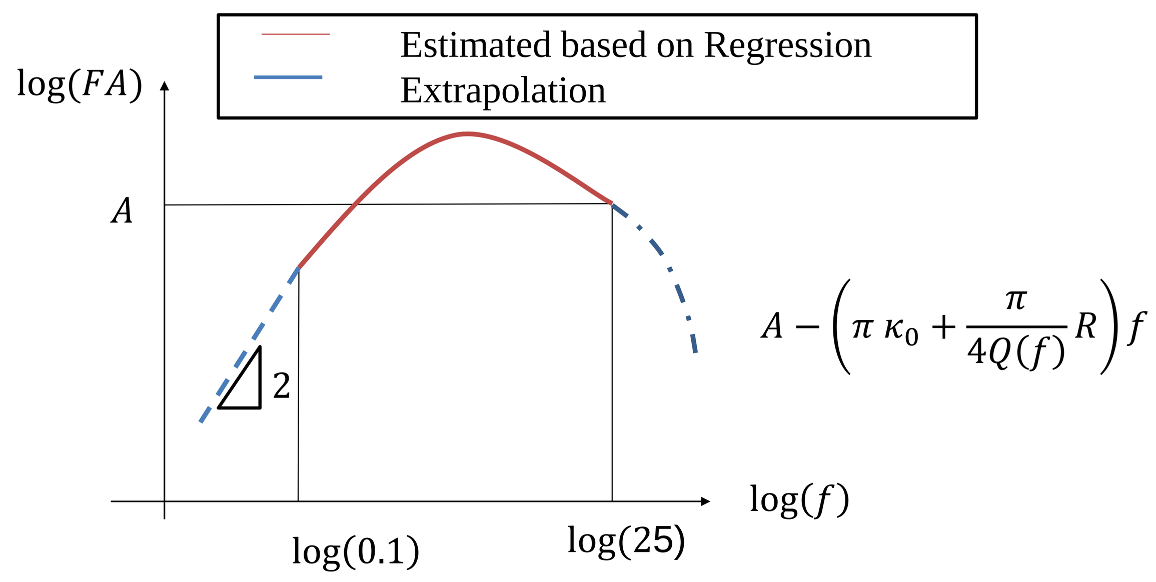

Figure 4. Extrapolation of FA for frequencies higher than 25 Hz and lower than 0.1

Conclusion

In this study, we present an approach to develop joint regression GMPEs for different tectonic regimes using Bayesian regression. In this approach, if the difference between magnitude, distance and site is statistically significant, different scaling parameters are generated. Otherwise, the database is used to estimate the a single scaling parameter. The functional form used in this study is adopted from Bayeless and Abrahamson (2019) with few modifications.

References

Bayless, Jeff, and Norman A. Abrahamson. "Summary of the BA18 Ground-Motion Model for Fourier Amplitude Spectra for Crustal Earthquakes in California." Bulletin of the Seismological Society of America 109.5 (2019): 2088-2105.
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.
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.