CS230

Analysis and Optimization of Ground Motion Prediction Equations Using Machine Learning Architectures

Thomas J. Little
Department of Civil and Environmental Engineering
Stanford University
tjlittle@stanford.edu

1 Introduction

Earthquakes are one of the most lethal naturally occurring hazards, accounting for three of the top five deadliest natural disasters in recent history (3). In addition to being extremely hazardous, earthquakes are also unpredictable, making it difficult to design for their effects. One method that can be used to estimate unknown earthquake behavior is to construct physics-based models of geologic hazards and estimate the impacts of different seismic scenarios. This method, however, is very computationally expensive and requires significant physical knowledge of the fault and surrounding area. Instead a common alternative is to use pre-constructed ground motion prediction equations (GMPEs) to estimate key response parameters. These GMPEs are attenuation relationships that take key characteristics of the earthquake's estimated source, path, and site and output parameters used for structural design and hazard mitigation purposes.

GMPEs are generally difficult to create and require large amounts of knowledge about the underlying mechanics of seismic events and geophysical processes. As a result, there are only a handful of GMPEs that are widely used in practice. As you increase the parameters of a GMPE, generally the range of scenarios and geographic locations that it is applicable to, decreases while its estimated accuracy increases. Because of this, there are different GMPEs currently used with a wide variety of use cases that depend on required accuracy, generalizability and available information.

The goal of this study is to determine the feasibility of using machine learning methods to circumvent this required knowledge and produce new GMPEs with a model that has no prior information on geomechanical principles. One use case of a framework such as this would be to construct a GMPE specific to the available inputs of a given earthquake scenario to create a more accurate hazard prediction than using an existing GMPE with either too many or too few input parameters.

GMPEs are often used to output a response spectrum, which is a tool used by structural engineers and risk analysts to evaluate the potential effects of a seismic event on varying building taxonomies. Therefore, the output of the trained model will also be a response spectrum. In order to quantify the accuracy of this model output, the response spectrum from both the GMPE and from the learned model will be compared to a "ground-truth" response spectrum calculated by running a linear response history analysis on the accelerogram that corresponds to a given earthquake event. Since these response spectra are a series of discrete paired values, the accuracy can be calculated using the mean average percent error of each series. Mean percent error was chosen as a metric in an attempt to limit the over-penalization of response spectra with large ordinate magnitudes, as this is widely variable between seismic events.

CS230: Deep Learning, Fall 2022, Stanford University, CA. (LateX template borrowed from NIPS 2017.)

2 Related work

Currently, a research team from the Department of Civil Engineering at IIT Madras, Chennai is conducting research on the use of various neural network architectures for GMPE estimation, however, the models proposed by this team follow similar classification approaches that are used in traditional GMPE formulation. Specifically, the team's models are predefined to specific tectonic regimes that rely on prior knowledge of the geomechanical makeup of the chosen region. One of the key ideas of this analysis is to identify the feasibility of training a model to overcome these limitations, and become capable of adjusting its prediction to account for varying tectonic conditions.

3 Dataset

The dataset chosen for this analysis was NGA-West2, a collection of ground motion attenuation data covering the western United States. This database contains over 200,000 realizations that include spatial and geotechnical data regarding both the seismic event, and the location where the ground motion was recorded, providing a robust description of each instance.

Each instance in the database also includes a response spectrum associated with the event at the location it was recorded, which will serve as the "ground-truth" for each instance. Although the NGA-West2 database contains ground motions collected across a variety of locations, it is important to note that this database is not comprehensive over all tectonic regimes and largely represents shallow crustal events. Other databases that may be explored in the future include GeoNet, FDSN, and K-Net, all of which contain strong motion data from across the globe.

The primary reason these databases were excluded from this analysis was due to the GMPE used for performance comparison. The GMPE used in this analysis is the 2014 Boore, Stewart, Seyhan, and Atkinson (BSSA) GMPE (1), which is only valid over shallow crustal faults, so using it as a performance metric over all tectonic regimes would not be appropriate. Since the NGA-West2 database contains primarily shallow crustal events, it was determined that this would be a good initial database to use in conjunction with the BSSA GMPE as a metric to gauge new model performance. The few events that do not strictly conform to the shallow crustal regime were left in the dataset to serve as adversarial examples for training, in attempts to increase the robustness of the trained model output.

One key element of the NGA-West2 database is that it contains recordings of the same seismic event from many different geographic locations, all of which may produce different response spectra. Because of this, the data could not be split randomly into test, train, and validation sets. Instead, the data was first clustered to ensure that all recordings of a specific event were contained within only one set. Once this was confirmed, the data was shuffled within each dataset to eliminate any sequential dependencies. The 200,000 events were split roughly into 80% for training, 10% for validation, and 10% for testing.

4 Approach

4.1 Model Inputs

Two scenarios were explored to test the feasibility of using machine learning methods to estimate attenuation relationships. The first scenario exclusively used the inputs required for the BSSA GMPE and evaluated whether a trained model could outperform the GMPE. The second scenario added additional input parameters relating to source and site characteristics and evaluated whether this resulted in increased performance of the trained model. In the following tables the model using only the BSSA input parameters will be referred to as "BSSA Params" and the model using additional parameters will be referred to as "All Params". A full list of parameters included in each of these models can be found in Tables 4 and 5 in the appendix.

4.2 Model Outputs

One of the hyperparameters analyzed during training was the number of model output nodes. The original intuition was to include the natural period of a structure as an input parameter and try to estimate a single spectral acceleration. While this method produced reasonable accuracy (see Results section), its training turned out to be very slow. Instead, if a series of periods were passed as input parameters, the model could output a series of spectral accelerations, thus predicting an entire response spectrum rather than a single value. This second method resulted in lower overall loss and trained on an order of twice the speed as the single-period prediction method. In subsequent testing, it was determined that periods need not be included as input parameters, as long as the range over which the response spectra was being calculated remained constant, which further decreased training time. The hyperparameter search continued by altering the length of the output series, reaching an optimal length of 35 output nodes.

4.3 Model Architectures

Two model architectures were considered in this analysis, a dense neural network and a gradient-boosted decision tree model. For the dense neural network, the number of hidden units and batch size were tuned using randomized search cross validation. Next, the number of layers was tuned in a similar fashion. The resulting architecture is shown in Table 1 below.

Neural Network Architecture		
Layer	Units	Dropout Rate
Dense	# Inputs	-
Dense	64	-
Dropout	-	0.3
Dense	128	-
Dropout	-	0.4
Dense	128	-
Dropout	-	0.6
Dense	64	-
Dropout	-	0.6
Dense	35	-

Table 1: Neural Network Architecture

The gradient-boosted decision tree model was constructed using the XGBoost framework. Several rounds of hyperparameter tuning were conducted using randomized search cross validation, the optimal parameters from this search are shown below in Table 2.

XGBoost Optimal Parameters	
Parameter	Value
Min Child Weight	1
Max Depth	5
Learning Rate	0.1
Gamma	0.1

Table 2: Boosted Trees Architecture

5 Results

While all models were able to outperform the BSSA GMPE on the NGA-West2 dataset, there are some clear trends in the data. An aggregated list of model performance is summarized in the figure below.

Mean Average Percent Error				
Model Type	Train Loss	Validation Loss	Test Loss	BSSA Loss
NN: One Period - BSSA Params	0.2772	0.2564	0.3070	0.3376
NN: BSSA Params	0.2448	0.2207	0.2693	0.3376
NN: All Params	0.2092	0.1893	0.1866	0.2826
XGB: BSSA Params	0.2264	0.2686	0.2962	0.3038
XGB: All Params	0.1028	0.0942	0.0934	0.2548

Table 3: Model Performance Summary

The first trend, as stated previously, is that the model performed better (both in terms of training speed and accuracy) when predicting the entire response spectrum as opposed to predicting spectral accelerations at individual periods. The second trend that emerges is an increase in model performance as additional input parameters are used. Note that in Table 3, the BSSA loss varies slightly, this is due to differing test, train, and validation splits resulting from data being dropped due to different missing input features.

An interesting result was that the XGB: BSSA Params model converged to predict, on average, a response spectra very similar to the BSSA GMPE, with a mean absolute percent error of 0.0757 between the two models. Over the whole dataset, however, both this model and the GMPE had a relatively large error compared to the XGB: All Params model. The model that used additional parameters diverged more from the BSSA GMPE but fit the accelerogram data more accurately. Visualization of these results is shown in the below figure, additional examples are provided in Figures 5 and 6 in the Appendix.

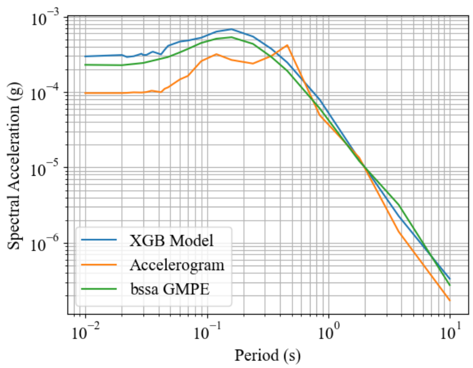

Figure 1: XGB: BSSA Params

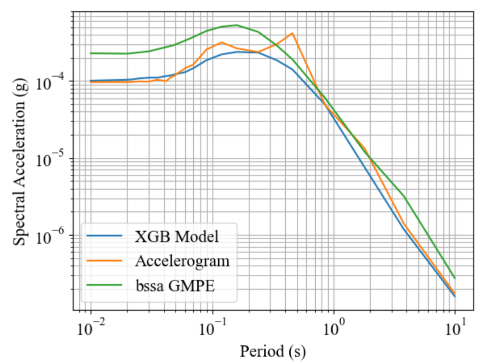

Figure 2: XGB: All Params

Moving forward with the best performing model (XGB: All Param), a feature analysis was conducted to determine which of the provided input features are most responsible for the increase in accuracy. The feature importance was determined using the built-in feature importance method from XGBoost, which computes importance based on how much a feature split improves model performance within the decision tree. The results indicate that there are two primary features responsible for increases in model accuracy, namely PGA and PGV. The full results of this feature analysis can be seen in Figure 4 in the Appendix.

6 Future Work

The purpose of a GMPE is to provide an attenuation relationship without requiring excessive amounts of effort. If the burden of collecting input feature data outweighs the benefit of using the GMPE, it no longer serves its purpose as an efficient calculation tool. For this reason, further investigation should be made into the features chosen for this analysis to conclude whether this is the case for the features in this model. Features such as hypocentral depth, PGA, and PGV may require intensive effort to calculate for unknown faults and therefore negate the usefulness of this model in realistic applications. Additionally, study into the covariance of the input parameters could help to eliminate unnecessary or redundant features. Lastly, alternate model architectures could be explored. The fact that the model performs better when predicting the entire response spectrum series seems to indicate dependence between points on the response spectrum. This information could be used to justify a recurrent neural network architecture, which may provide even better results.

7 Conclusion

While the models presented in this study seem to indicate that using more parameters results in a more accurate model of ground motion attenuation, it is important to consider the tradeoff between collecting the data for these input variables and the desired level of model accuracy. As discussed in the Future Work section, some parameters used in the presented models may be very difficult to collect in practice, making the feasibility of using them potentially unrealistic. However, all models display similar or better performance than the baseline BSSA GMPE, indicating that machine learning frameworks can be used to predict ground motion attenuation without the model requiring any prior knowledge of geomechanical systems. This method shows promise as a feasible alternative to traditional GMPE development and will hopefully continue to be investigated by future researchers.

8 Contributions

This work was developed by Thomas Little.

References

[1] Boore, D. M., Stewart, J. P., Seyhan, E., and Atkinson, G. M. (2014). “NGA-West2 Equations for Predicting PGA, PGV, and 5% Damped PSA for Shallow Crustal Earthquakes.” Earthquake Spectra, 30(3), 1057–1085. http://www.earthquakespectra.org/doi/abs/10.1193/070113EQS184M

[2] Dhanya, J., and S. T. G. Raghukanth. "Ground motion prediction model using artificial neural network." Pure and Applied Geophysics 175.3 (2018): 1035-1064. https://link.springer.com/article/10.1007/s11600-019-00395-yBib1

[3] Lake, T. (n.d.). 6 of the world’s Deadliest Natural disasters. Encyclopædia Britannica. Retrieved December 9, 2022, from https://www.britannica.com/list/6-of-the-worlds-deadliest-natural-disasters

[4] “Welcome to the Peer Ground Motion Database.” PEER Ground Motion Database - PEER Center, https://ngawest2.berkeley.edu/.

Appendix

Features - BSSA Parameters	
Feature Number	Feature Name
f0	Earthquake Magnitude
f1	Joyner-Boore Distance (km)
f2	Rake Angle (degrees)
f3	Northern CA/Southern CA - H11 Z1 (m)
f4	Northern CA/Southern CA - S4 Z1 (m)
f5	Vs30 (m/s)

Table 4: Model Feature - BSSA Param

Features - All Parameters	
Feature Number	Feature Name
f0	Earthquake Magnitude
f1	Strike (degrees)
f2	Dip (degrees)
f3	Rake Angle (degrees)
f4	P-plunge (degrees)
f5	T-plunge (degrees)
f6	Joyner-Boore Distance (km)
f7	Hypocenter Depth (km)
f8	Fault Rupture Length
f9	Fault Rupture Width
f10	Vs30 (m/s)
f11	Northern CA/Southern CA - H11 Z1 (m)
f12	PGA (g)
f13	PGV ((cm/sec)

Table 5: Model Feature - All Params

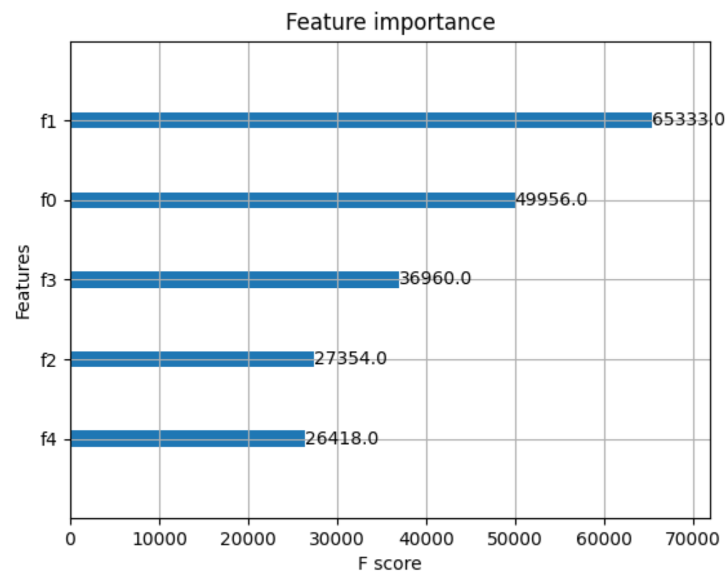

Figure 3: Feature Importance - BSSA Params

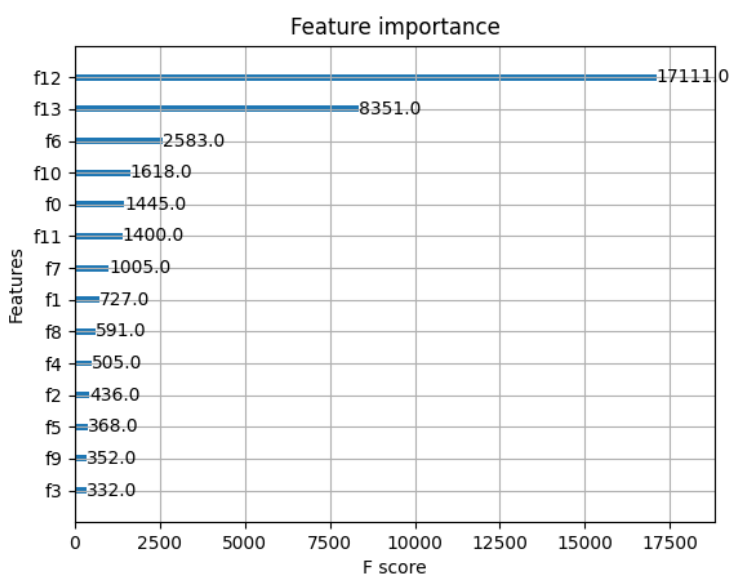

Figure 4: Feature Importance - All Params

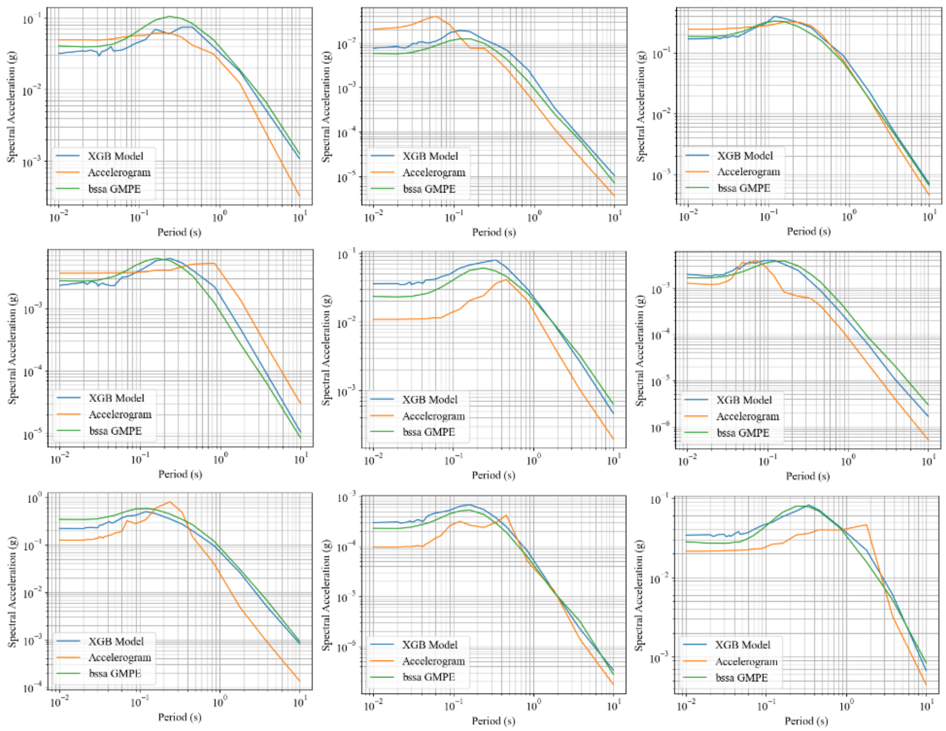

Figure 5: XGB: BSSA Params - Random Examples

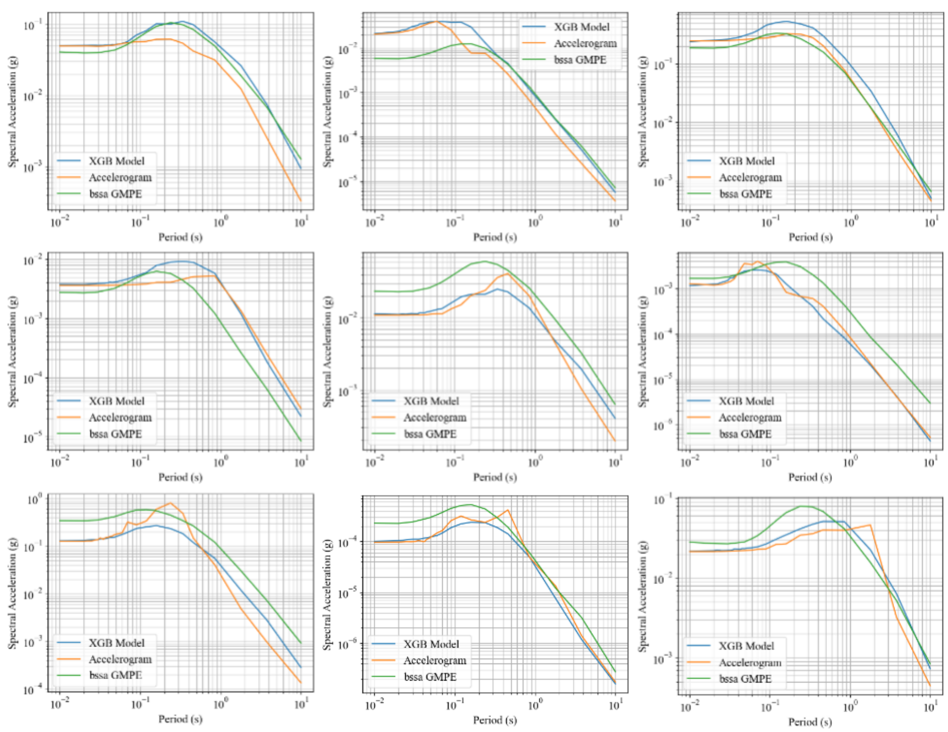

Figure 6: XGB: All Params - Random Examples