



## Information Sharing Between Ground Motion Models from Different Regions via Dirichlet Process Priors

M. Hermkes, N. Kuehn, C. Riggelsen, and K. Vogel

Institute of Earth and Environmental Science, University of Potsdam, Potsdam, Germany

In probabilistic seismic hazard analysis (PSHA) seismic ground motion data, induced by earthquakes, are collected at different geographical regions. Instead of building ground motion models, which estimates intensity parameters, e.g. peak ground acceleration or spectral acceleration, given earthquake and site related parameters, for each region individually, it is preferable to share information across the regions to increase the overall prediction performance.

One of the most important methods to share information correlation between models is Hierarchical Bayesian modeling, where parameters of the region-specific models are coupled by a common prior. As a result of learning the parameters of the model and the hyperparameters of the common prior jointly, the function estimation of a specific region is affected by its own training data and by data from the other region related through the coupled prior. Generally, the common prior is specified in a parametric form with unknown hyperparameters. A drawback of such a prior by reason of its modality is that the relationship between all ground motion models are treated equally, but it is desirable that only similar models share information to permit negative transfer.

To deal with these issues we propose a nonparametric hierarchical Bayesian model where the common prior is drawn from a Dirichlet Process (DP). Such a nonparametric prior has the ability to fit the model well with respect to the data without restriction about the functional form of the prior distribution. Furthermore, the employed DP prior induces a partition of region-specific models, so that models within each cluster share the same parameterization. First of all, we present a linear regression model, for which the weights of the covariates and the model variance are drawn from a DP prior. As base distribution for the DP we have chosen a normal inverse-Gamma prior which is the natural conjugate prior to the normal likelihood of the applied regression model. In addition, we extend this model by replacing the linear regression model by Gaussian Processes. The resulting models can be seen as a DP Mixture (DPM) of linear regression functions, respectively DPM of Gaussian Processes. By choosing conjugate priors, the base distribution can be analytically marginalized, but the sum over all latent partitions makes exact Bayesian inference intractable. Instead of using MCMC sampling machinery which may be slow to convergence, we apply the Bayesian Hierarchical Clustering (BHC) algorithm (Heller and Ghahramani, Proceedings of ICML'05) to make approximative inference.

The experiments are performed on the Next Generation Attenuation (NGA) and Allen & Wald data set. For comparison we also consider the performance of a single task learning method (training a separate model for each task) and a complete pooling approach (train a model on the complete data) as baseline methods. The results show improved prediction performance of the DPM models compared to these baseline methods.