Geophysical Research Abstracts
Vol. 14, EGU2012-2951, 2012
EGU General Assembly 2012

Graphical Models as Surrogates for Complex Ground Motion Models

K. Vogel, C. Riggelsen, N. Kuehn, and F. Scherbaum
University of Potsdam, Institute of Earth and Environmental Science, Germany (kvog@geo.uni-potsdam.de)

An essential part of the probabilistic seismic hazard analysis (PSHA) is the ground motion model, which estimates the conditional probability of a ground motion parameter, such as (horizontal) peak ground acceleration or spectral acceleration, given earthquake and site related predictor variables. For a reliable seismic hazard estimation the ground motion model has to keep the epistemic uncertainty small, while the aleatory uncertainty of the ground motion is covered by the model.

In regions of well recorded seismicity the most popular modeling approach is to fit a regression function to the observed data, where the functional form is determined by expert knowledge. In regions, where we lack a sufficient amount of data, it is popular to fit the regression function to a data set generated by a so-called stochastic model, which distorts the shape of a random time series according to physical principles to obtain a time series with properties that match ground-motion characteristics. The stochastic model does not have nice analytical properties nor does it come in a form amenable for easy analytical handling and evaluation as needed for PSHA. Therefore a surrogate model, which describes the stochastic model in a more abstract sense (e.g. regression) is often used instead.

We show how Directed Graphical Models (DGM) may be seen as a viable alternative to the classical regression approach. They describe a joint probability distribution of a set of variables, decomposing it into a product of (local) conditional probability distributions according to a directed acyclic graph. Graphical models have proven to be a all-round pre/descriptive probabilistic framework for many problems. Their transparent nature is attractive from a domain perspective allowing for a better understanding and gives direct insight into the relationships and workings of a system. DGMs learn the dependency structure of the parameters from the data and do not need, but can include prior expert knowledge.

We investigate DGMs admitting to different decompositions/factorizations of the joint distribution, that is, the restrictions that are imposed by the graph: Bayesian Networks, Naive Bayes and Tree Augmented Naive Bayes. In order to use ground motion data for what we call a distribution-free learning of DGMs, we need to discretize the variables. The number of intervals and their boundaries have to be chosen carefully, since essential information about the distributions and dependencies of the variables may be lost otherwise. Depending on the DGM, we use different established discretization methods and extend them for the usage of a continuous target variable. The learned networks are compared to a regression approach and show equally good results in prediction of the target variable. Moreover, the entirely data-driven approach of learning the BN enables for a correct interpretation of the (in)dependences between the variables, as opposed to imposed algebraic interaction effects of the regression model.