UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS
UNECE Machine Learning for Official Statistics Workshop 2023
5-7 June 2023, Geneva Switzerland

04 April 2023

Geospatial Bayesian Methods for Hazard-Impact Modelling

Hamish Patten, David Steinsaltz, Max Anderson Loake (University of Oxford)
hamish.patten@gmail.com

Abstract

Immediately after the occurrence of rapid-onset hazards, such as earthquakes, governmental and humanitarian organisations allocate aid resources to facilitate recovery, minimise displaced populations and reduce the long term effects. Estimates of the disaster impact are noisy, and are often incomplete (especially geospatially). Predictive analytics could play a crucial role in providing key information and guidance to facilitate disaster management. With the recent emergence of a plethora of open-access data relevant to disaster resource modelling, the amount of information available to build models and accurately predict disaster impact has never been so vast and open. The outcome of this research, developed at the Department of Statistics, University of Oxford, has been to build an open-source statistical software product, called the Oxford Disaster Damage Real-time Information Network (ODDRIN), based on R, R-Shiny and Leaflet. Please visit www.github.com/hamishwp/ODDRIN to see the code and the manual.

Having estimates of one impact category can be used to infer another: multiplying the number of destroyed buildings with the average household size can infer long-term population displacement, for example. Using a multivariate-impact model, ODDRIN predicts mortality, human displacement and building damage and destruction, on a global level. Based on historical data from over 180 earthquakes from around the world since 2013, ODDRIN couples different spatial data types into one hierarchical model: aggregated and pointwise ground-based observations (official and non-official), satellite image-based pointwise building damage assessments, and mobile-phone based (CDR or XDR) aggregated spatio-temporal population displacement estimates. The model also provides insight into systemic vulnerability, by integrating variables such as gross income and life expectancy into the calculation. With a strong focus on estimating model uncertainty and ensuring predictive performance, the ODDRIN model is integrated into a Bayesian framework. To parameterise the model, a specific Approximate Bayesian Computation – Sequential Monte Carlo (ABCSMC) algorithm is applied.

The additional benefit of training the model on global historical events is the ability to infer whether a given observed impact value may be an under- or over-estimate. This is shown to be particularly important for population displacement and building damage and destruction estimates, which are shown to have vastly different values when comparing estimates from different sources. A transparent presentation of the predictive performance of ODDRIN, especially per country, highlights the current limits of disaster-impact modelling in general, and the authors propose several potential solutions to address some of these issues.