Prospectively Evaluating the Collaboratory for the Study of Earthquake Predictability: An Evaluation of the UCERF2 and Updated Five-Year RELM Forecasts

Anne Strader (1), Max Schneider (1,2), Danijel Schorlemmer (1,3), and Maria Liukis (3)

(1) GFZ-Potsdam, Potsdam, Germany , (2) Universität Potsdam, Potsdam, Germany , (3) USC, Los Angeles, USA

The Collaboratory for the Study of Earthquake Predictability (CSEP) was developed to rigorously test earthquake forecasts retrospectively and prospectively through reproducible, completely transparent experiments within a controlled environment (Zechar et al., 2010). During 2006-2011, thirteen five-year time-invariant prospective earthquake mainshock forecasts developed by the Regional Earthquake Likelihood Models (RELM) working group were evaluated through the CSEP testing center (Schorlemmer and Gerstenberger, 2007). The number, spatial, and magnitude components of the forecasts were compared to the respective observed seismicity components using a set of consistency tests (Schorlemmer et al., 2007, Zechar et al., 2010). In the initial experiment, all but three forecast models passed every test at the 95% significance level, with all forecasts displaying consistent log-likelihoods (L-test) and magnitude distributions (M-test) with the observed seismicity. In the ten-year RELM experiment update, we reevaluate these earthquake forecasts over an eight-year period from 2008-2016, to determine the consistency of previous likelihood testing results over longer time intervals. Additionally, we test the Uniform California Earthquake Rupture Forecast (UCERF2), developed by the U.S. Geological Survey (USGS), and the earthquake rate model developed by the California Geological Survey (CGS) and the USGS for the National Seismic Hazard Mapping Program (NSHMP) against the RELM forecasts. Both the UCERF2 and NSHMP forecasts pass all consistency tests, though the Helmstetter et al. (2007) and Shen et al. (2007) models exhibit greater information gain per earthquake according to the T- and W- tests (Rhoades et al., 2011). Though all but three RELM forecasts pass the spatial likelihood test (S-test), multiple forecasts fail the M-test due to overprediction of the number of earthquakes during the target period. Though there is no significant difference between the UCERF2 and NSHMP models, residual scores show that the NSHMP model is preferred in locations with earthquake occurrence, due to the lower seismicity rates forecasted by the UCERF2 model.