Final Report for USGS Award G11AP20209

Project Title: Collaboratory for the Study of Earthquake Predictability: Research and Development to Support USGS Operational Earthquake Forecasting

Principal Investigator: Thomas H. Jordan
University of Southern California
3651 Trousdale Parkway, ZHS 169
Los Angeles, CA 90089-0742
tjordan@usc.edu | 213-740-5843 | 213-740-0011 fax

February 8, 2014

Introduction

Operational earthquake forecasting (OEF) is the dissemination of authoritative information about the time dependence of seismic hazards to help communities prepare for potentially destructive earthquakes (ICEF, 2009, 2011; Jordan and Jones, 2010). Forecasting models considered by the USGS for operational purposes will have to demonstrate reliability and skill with respect to established reference forecasts, such as the long-term, time-independent models used in probabilistic seismic hazard analysis (PSHA). This demonstration requires objective evaluation of how well forecasting models correspond to data collected after the forecast has been made (prospective testing), as well as checks against data previously recorded (retrospective testing).

The Collaboratory for the Study of Earthquake Predictability (CSEP) is an international partnership to develop and maintain a cyberinfrastructure for the rigorous prospective testing of earthquake forecasting models. CSEP is led by the Southern California Earthquake Center (SCEC), which operates its U.S. testing center and is responsible for most of the CSEP software development.

The goal of this 18-month project, which began on Sept 1, 2011, was to develop a SCEC and USGS CSEP working group and begin to adapt the existing CSEP infrastructure to the needs of the USGS OEF program. In Sept 2008, the National Earthquake Prediction Evaluation Council (NEPEC) conducted a review of CSEP at the request on the USGS Earthquake Hazards Program (EHP), and the recommendations in their 2009 report have guided the development of this project and the prioritization of its activities.

A. Key Accomplishments Related to Proposed Tasks

In Section A of this final report, we review the progress made under each of the six major tasks specified in the original project proposal. In Section B, we outline other accomplishments related to OEF testing.

1. SCEC-USGS CSEP Working Group

The first task was to establish a SCEC-USGS CSEP Working Group to guide project activities. This working group was configured in consultation with USGS leadership. Dr. Andrew Michael of EHP's Menlo Park office, who was designated as the USGS lead for this project, co-chairs the Working Group. The initial membership is listed in Table 1.

Discussions of the CSEP Working Group activities were conducted at a CSEP session convened at the 2011 SCEC Annual Meeting. This session was followed by a CSEP workshop at SCEC on January 23,

2012, with key USGS personnel from Pasadena, Menlo Park, and Golden to coordinate CSEP plans in support of USGS OEF development. One outcome of this workshop was the recommendation that strict versioning be introduced into the new USGS catalog, COMCAT, and this recommendation has since been adopted by NEIC. A third CSEP workshop was held by SCEC in Rancho Mirage on June 6-7, 2012, which brought together researchers from CSEP testing centers at SCEC (U.S.), ETHZ (Europe), ERI (Japan), and GNS Science (New Zealand).

SCEC hosted two special CSEP workshops on the problem of evaluating earthquake forecasting and prediction methods developed and tested outside of CSEP, including those that incorporate electromagnetic and other non-seismic data: "Evaluating Ground-Based and Space-Based Methods of Earthquake Forecasting" (July 26-27, 2011), and "Testing External Forecasts and Predictions" (May 7-8, 2013). A significant accomplishment was to bring the M8 forecasts into CSEP testing, an explicit objective for CSEP recommended in the 2008 NEPEC review.

Table 1. SCEC-USGS CSEP Working Group

Co-Chairs: Andrew Michael (USGS, Menlo Park) and Thomas H. Jordan (SCEC, USC)

USGS members: Harley Benz (Golden), Michael Blanpied (Reston), Karen Felzer (Pasadena), Ned Field (Golden), Douglas Given (Pasadena), Robert Graves (Pasadena), Jeanne Hardebeck (Menlo Park), Kenneth Hudnut (Pasadena), Lucy Jones (Pasadena), David Oppenheimer (Menlo Park), Morgan Page (Pasadena)

SCEC members: Gregory Beroza (Stanford), Matthew Gerstenberger (GNS, New Zealand), Egill Hauksson (Caltech), David Jackson (UCLA), Yan Kagan (UCLA), Philip Maechling (USC), Warner Marzocchi (INGV, Italy), David Rhoades (GNS, New Zealand), Danijel Schorlemmer (GFZ, Germany), Terry Tullis (Brown), Max Werner (Princeton), Ilya Zaliapin (UNR), Jeremy Zechar (ETHZ, Switzerland)

Coordination and Administration: Tran Huynh (SCEC, USC)

Technical support: Maria Liukis (SCEC, USC), John Yu (SCEC, USC)

2. Reduction of Testing Latency

The second task was to reduce the updating interval of the short-term forecasting models from 1 day to 1 hour or less, in order to increase the information gain in forecasting aftershock sequences.

To accomplish this task, CSEP introduced a 30-minute forecast group that reduces the testing latency from 1 day to 30 minutes. Currently, two models are generating forecasts in this group: an ETAS model and a smoothed seismicity model (PPE), both by David Rhoades. In Figure 1, we compare 30-minute forecasts with a 24-hour forecast during the Brawley Swarm, here during the evening of August 26, 2012. The 30-minute forecasts reflected changes in occurrence probability by a factor of more than ten as a result of moderate M5+ earthquakes, whereas the 24-hour forecast was frozen until midnight.

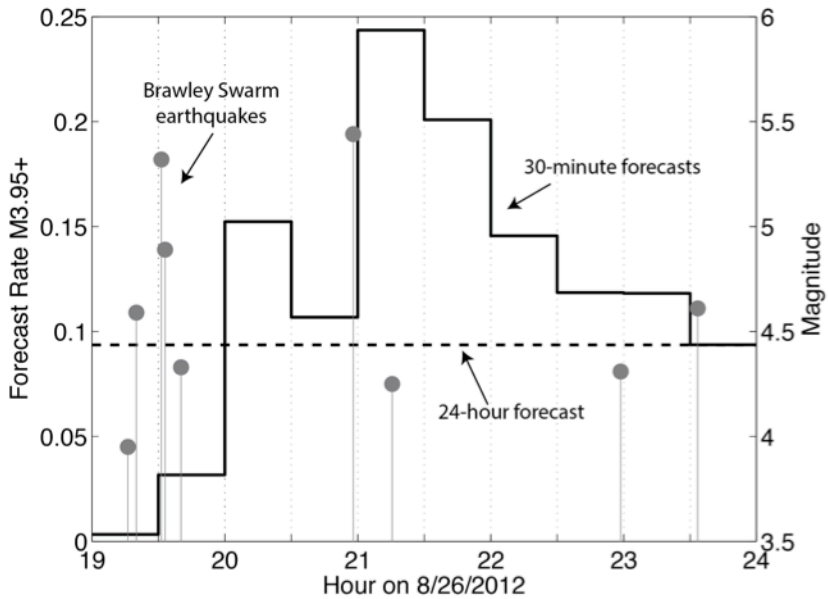

Figure 1. Illustration of 30-minute forecasts compared to a 24-hour forecast during the Brawley Swarm on August 26, 2012. The 30-minute forecasts reflect changes in seismic activity; the 24-hour forecasts are fixed until midnight. Both forecasts are based on an ETAS model by David Rhoades.

Retrospective testing of ETAS model forecasts by Werner and Helmstetter demonstrated that the information gain per earthquake increases dramatically as the updating interval is reduced from 1 day to 30 minutes (Figure 2). The information gain from 1992 until 2012 for earthquakes with magnitudes greater than 3.95 in the California testing region increases from 4 per earthquake to about 5.5 per earthquake using their preferred version of the model.

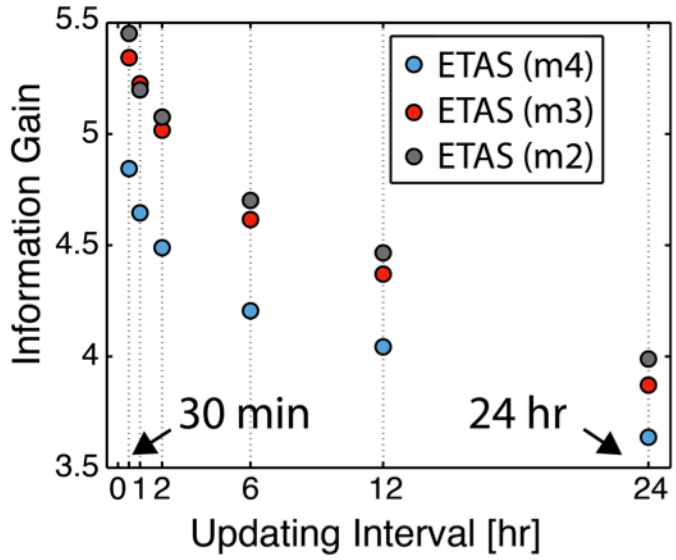

Figure 2. Illustration of the rapid decrease of the information gain per earthquake as a function of the updating interval. Shown is the average information gain per earthquake as a function of updating interval for three different ETAS models by Werner and Helmstetter. Forecasts are evaluated retrospectively on M3.95+ earthquakes from 1992 until 2012 in the California testing region.

3. Reference Models

The 2008 NEPEC review recommended the adoption of simple seismicity-based models that provide reference forecasts against which the skill of candidate models can be evaluated. Two reference models were installed at the SCEC CSEP testing center to address this task:

- A global 1-day seismicity forecast with high spatial resolution (0.1º), developed from the critical branching model by Kagan and Jackson (2012). This model uses the Global CMT catalog as input.
- UCERF2, the long-term model by Field et al. (2008), which is built on the California component of the current NSHMP hazard model.

This project also sponsored the development of a version of the Kagan-Jackson model based on the PDE catalog, which offers a lower magnitude threshold. The PDE catalog is now being automatically imported into CSEP, and testing will begin once the CSEP software has been adapted to recent changes in the PDE format.

4. Reduction of Catalog Latency

The authoritative data source used in most CSEP experiments has been a mature earthquake catalog, such as the ANSS catalog, requiring a testing latency of about one month. However, OEF must be based on immature, real-time catalogs that are incomplete and have larger errors in source parameters. Under this project, CSEP has begun the development of testing procedures that can account for the deficiencies of the real-time catalogs.

This research entails real-time imports of the ANSS catalog for California and generation of real-time, 1-day forecasts from all models currently installed within the 1-day forecast group for California. Using Subversion, the real-time catalog is differenced with the 1-month-delayed catalog in general CSEP use, which quantifies the catalog differences.

The SCEC CSEP testing center is now generating short-term forecasts based on real-time catalogs. The preliminary experiments have thus far been hosted on the CSEP certification server and are not accessible to the public. Their purpose is to evaluate the difference in quality of the real-time forecasts compared with forecasts generated with the delayed catalog.

5. UCERF Testing

The Uniform California Earthquake Rupture Forecast describes a sequence of a time-dependent, fault-based models created by the Working Group on California Earthquake Probabilities (WGCEP). UCERF2 was released in 2008, and this model is now under prospective testing in the CSEP five-year forecast group for California. UCERF2 has also been evaluated against the RELM models.

UCERF3 was scheduled for release in 2012. Owing to delays in this release, we have not yet begun the prospective testing of the UCERF3 model.

6. Retrospective Testing

Although CSEP is focused on prospective testing, its standardized procedures make it an excellent platform for model calibration and retrospective testing. Under this task, we have expanded CSEP activities to include retrospective testing of earthquake forecasting models. This has led to a notable retrospective experiment that is testing physics-based models against the Canterbury earthquake sequence.

Canterbury Retrospective Experiment. The 2010 M7.1 Darfield (Canterbury), New Zealand, earthquake triggered an interesting and complicated earthquake cascade that provides a wealth of new scientific data to study the physical properties of earthquake interaction and triggering. A global group of earthquake scientists interested in the physical mechanisms and predictability of earthquake cascades have agreed,

as a result of increased collaboration between CSEP groups in the US, New Zealand, Japan, and Europe, to conduct a joint retrospective evaluation of short-term earthquake forecasting models during the Canterbury sequence. The scientific goals are to (i) improve our understanding of the physical mechanisms governing earthquake interaction and triggering using the data collected from the Darfield sequence; (ii) improve short-term earthquake forecasting models and time-dependent hazard assessment for Canterbury and to apply lessons learned elsewhere; and (iii) understand the influence of poor-quality, real-time data on the skill of forecasts that might be issued in real time in an operational setting. To achieve these goals, eight different research groups from the US, New Zealand, Japan and Europe are developing more than 20 physics-based and statistical forecasting models that will be compared using the CSEP infrastructure for retrospective testing. CSEP software developers will install these models in the New Zealand testing center by early 2014. The physics-based forecasting models will be run pseudo-prospectively at updating intervals of various durations (1-day, 1-month, 1-year), starting from the Darfield mainshock. All models will remain under CSEP testing after the retrospective experiment, providing truly prospective evaluations.

Other Accomplishments

In addition to addressing the six main tasks outlined above, this project supported general development of CSEP testing procedures and sponsored SCEC participation in several important international collaborations aimed at improving OEF capabilities.

1. New CSEP Testing Procedures

We have added the T-test into the battery of tests employed by the CSEP testing centers. This type of forecast evaluation test was selected as a priority for inclusion into CSEP because we have found that models may fail other tests (N-test, L-test, R-test) because of short-term clustering of earthquakes based on the Poissonian assumption used in the forecast, and the limitation that models can only be updated periodically. As a result, the comparison of two models subjected to the same restrictions, by the T-test, may be easier to interpret than the previous CSEP test results.

A comparison between T-test and R-test results on retrospective analyses of models submitted to the New Zealand testing center and on the published half-time results of the prospective five year RELM experiment in California, indicates that the T-test rejects models less frequently than the R-test. In particular, it is less likely to reject one model in favor of another when the difference in their likelihoods is mostly attributable to only one or two of the target earthquakes.

The T and W tests implemented as part of this project have been used to evaluate forecast models in both California and New Zealand. Results from evaluation of STEP forecast model in New Zeland show significant probability gain for ETAS as compared to the existence reference forecast, based on existing New Zealand National seismic hazard maps for the Darfield region. These tests have also been applied to existing short-term forecasts for California. Current paired T-test results show that STEPJAVA is a winning model in CA one-day forecasts, EEPAS-0F is a winner in 3-months forecasts, Helmstetter et al. model is a winner in 5-year RELM Mainshock and Mainshock/Aftershock forecasts groups.

2. Support of Operational Earthquake Forecasting in New Zealand

The damaging Canterbury earthquake sequence commenced with the September 2010 M7.1 Darfield earthquake and continues to this day. Following the damaging aftershocks in Christchurch, GNS Science coordinated the development of a hybrid operational earthquake forecasting model to be used for setting building standards and urban planning for the rebuilding of the city. The model estimates occurrence probabilities of magnitude M ≥ 5.0 for the Canterbury region for each of the next 50 years. It combines short-term, medium-term and long-term forecasting models. Short-term models include the STEP (Short-Term Earthquake Probability) and ETAS (Epidemic-Type Aftershock Sequence) models, which incorporate the Utsu-Omori inverse power law for decay of aftershock activity. Medium-term models

include two versions of the EEPAS (Every Earthquake a Precursor According to Scale) model with different weighting strategies. These models are based on the precursory scale increase phenomenon (an increase in the rate of minor earthquake activity which typically precedes major earthquakes) and associated predictive relations for the magnitude, precursor time and precursor area. The long-term models include several different smoothed seismicity models, some designed to forecast main shocks only and others designed to forecast all earthquakes, including aftershocks. The weights accorded to the individual models in the operational model (the EE hybrid model) were determined by an expert elicitation process. Another hybrid model (the AVMAX hybrid model) involving only one model from each class (long-term, medium, and short-term) was used prior to the expert elicitation process. For both hybrid models, the annual rate of earthquake occurrence in a particular spatial cell was defined as the maximum of a long-term rate and a time varying rate. The difference was that, in the EE model, a weighted average of more individual models was used to compute the long-term and time-varying rates.

Although an optimal hybrid model for New Zealand as a whole would have a higher contribution to the time-varying component from medium-term models, the same does not necessarily apply to the Canterbury region. Tests of one-day, three-month and five-year models installed in the New Zealand testing center show that, after the Darfield earthquake, the Canterbury earthquakes were well described by the ETAS short-term one-day model, which has very large information gains over smoothed seismicity models calculated before the Darfield, or updated at regular intervals during the aftershock sequence.

The aftershock decay captured by the short-term models is the most predictable component of future Canterbury seismicity. The medium-term component is less predictable, because the precursor time parameters are not well established for low seismicity regions. There is considerable variation in the estimates from the individual long-term models contributing to the EE Hybrid model. But, based on the tests carried out in our special study, the EE hybrid, which gives appreciable weight to four different long-term models, is likely to outperform most of the individual models in the next 50 years.

Coulomb stress modeling of the Canterbury earthquakes also supports the notion that hybrid forecasts models outperform models based on a single method. A further retrospective Canterbury experiment is now being set up within CSEP. This experiment is designed to test the performance of forecasts based on Coulomb stress calculations, updated following each of the larger earthquakes in the Canterbury sequence. A number of Coulomb modelers have agreed to submit their codes for independent testing. This experiment is being coordinated by Warner Marzocchi of INGV and Matt Gerstenberger of GNS Science, as described in the previous section.

At the request of GNS Science, SCEC assisted with the testing of the individual and hybrid models by carrying out a specially designed retrospective forecasting experiment. Masha Liukis visited GNS Science in August 2012 to install the models in the NZ testing center and to initiate the test runs. The performance of the various models in the NZ CSEP testing region was compared over 26 years (the period during which the earthquake catalogue is adequate for the purpose). Retrospective annual forecasts of earthquakes with magnitude M > 4.95 from 1986 on were made for time-lags ranging from zero up to 25 years. The number of target earthquakes decreases with increasing time lag, from 303 at a time lag of zero down to 19 at a time lag of 25 years. Two standard CSEP tests were used in the evaluation: the N-test, which compares the observed number of earthquakes with the number forecast by the model, and the T-test, which measures the information gain per earthquake (IGPE) of one model over another.

The N-tests show that all models tend to under-predict the number of earthquakes in the test period (Figure 3). This was found to be mainly due to the unusually large number of earthquakes with M > 4.95 that have occurred in NZ since the M7.8 Dusky Sound earthquake of 15 July 2009, including the Canterbury earthquakes.

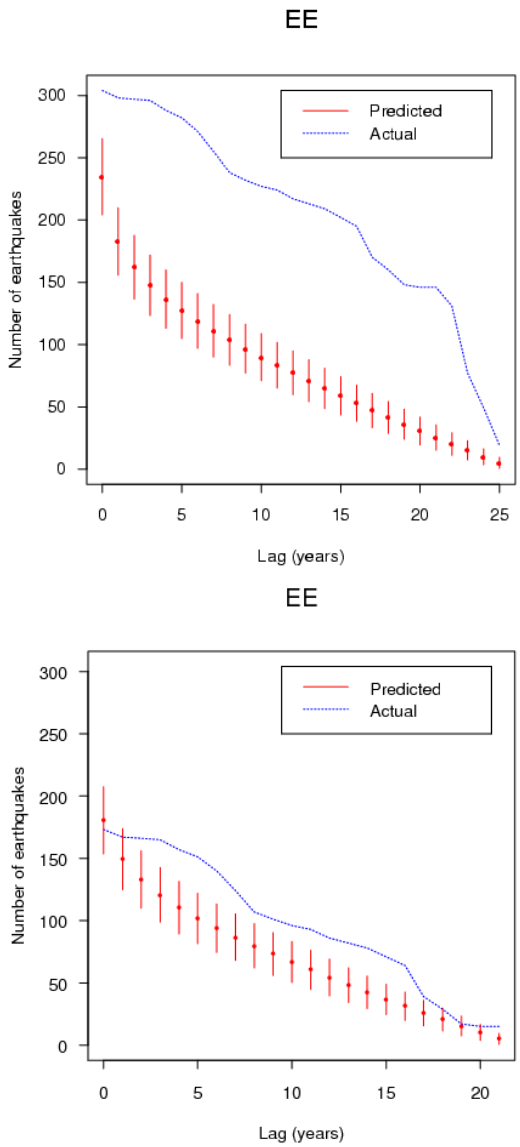

Figure 3. N-tests of EE hybrid model. Upper: for period 1982-2012. Lower: for period 1986-2008.

The T-tests show that both hybrid models are more informative than most of the individual models for all time lags (Figures 4-6). Using data form the full 26-year test period, the IGPE relative to a stationary and spatially uniform reference model (a model of “least information”) drops off steadily as the time lag increases, to become zero at a time-lag of about 20 years (Figure 5). When the unusual period since the Dusky Sound earthquake is removed from the tests, the hybrid models both show a significant positive IGPE over the model of least information at all lags, but do not outperform all the individual long-term models at long time lags. The results are therefore seen to be sensitive to unusual features of the test catalogue, and a longer catalogue would be needed to obtain robust results.

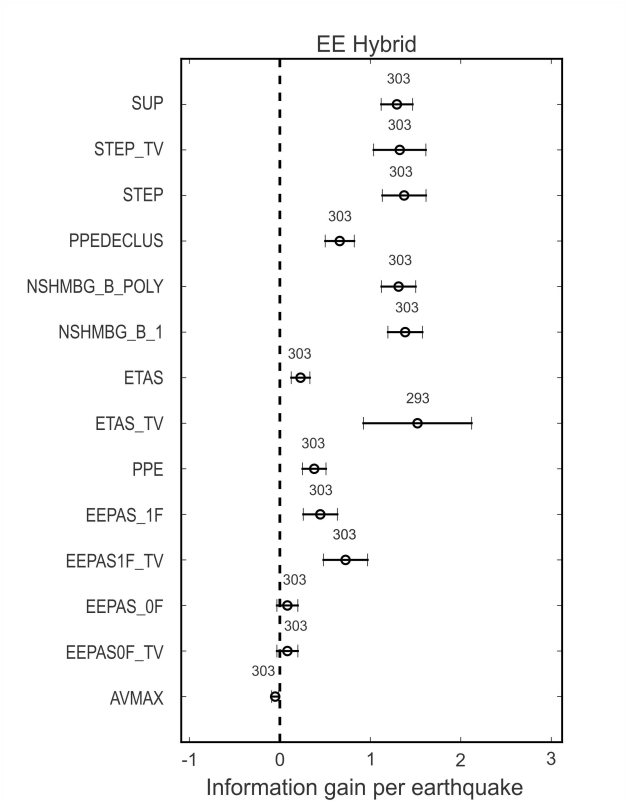

Figure 4. Information gain per earthquake of the EE hybrid models over other models in one year forecasts with zero time lag in the New Zealand CSEP testing region from 1986 to 2012. Error bars are 95% confidence intervals according to the T-test. The number of target earthquakes is given above the error bars. This test shows that the two hybrid models outperformed all individual models.

An optimal hybrid model with the same general form as the EE hybrid model was computed for each time lag from the 26-year test period. In the optimal hybrid, the time-varying component is dominated by the medium-term models, with hardly any contribution from the short-term models for time lags up to 12 years. The short-term and medium-term model rates diminish with increasing time lag, with the result that the time-varying component as a whole has hardly any impact on the optimal hybrid model for time lags greater the 12 years.

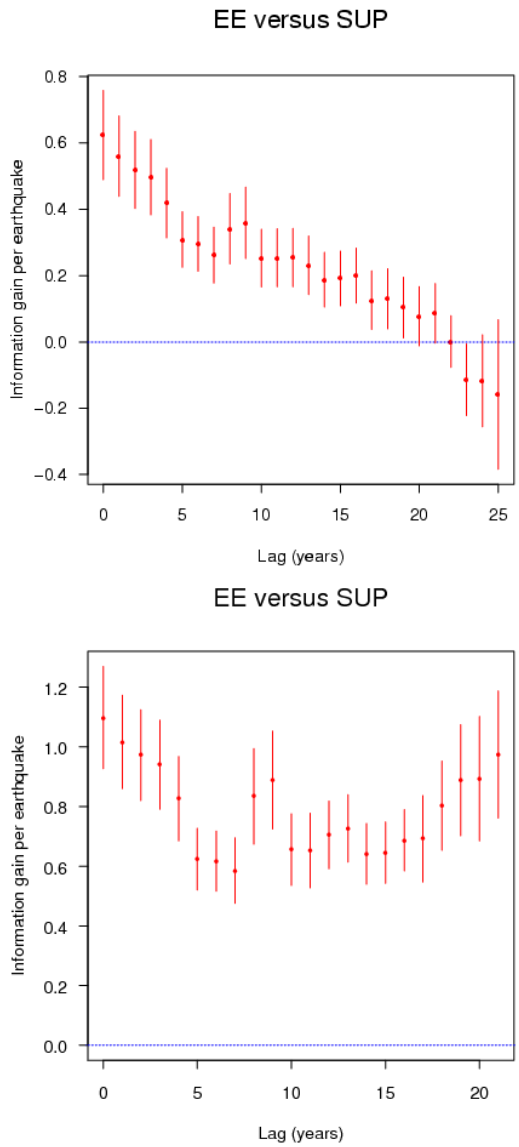

Figure 5. Information gain per earthquake of the EE hybrid model over the spatially uniform Poisson (SUP) model, as a function of the time lag. Error bars are 95% confidence limits. Upper: for period 1986-2012. Lower: for period 1986-2008.

3. Testing of Tohoku Models

Clustering-based models aiming to forecast earthquakes in Japan have been put under test in the CSEP testing center operated by the Earthquake Research Institute of Tokyo. The 2011 March 11 magnitude 9.0 Tohoku-Oki earthquake in Japan provided a unique opportunity to test the existing 1-day CSEP models against its unprecedentedly active aftershock sequence. The original CSEP experiment performed tests after the catalogue was finalized to avoid bias due to poor data quality. However, our study differed from this tradition by using the preliminary catalogue revised and updated by the Japan Meteorological Agency (JMA), which is often incomplete but is immediately available. This study was a first step towards opertional earthquake forecasting in Japan. At least one model passed the test in most combinations of the target day and the testing method, although the models could not take account of the megaquake in advance and the catalogue used for forecast generation was incomplete. However, all models had only limited forecasting power for the period immediately after the quake, and this conclusion does not change when the preliminary JMA catalogue is replaced by the finalized one. The models

perform stably with respect to the catalogue replacement. However, further research on model improvement is needed to ensure the reliability of forecasts for the days immediately after the mainshock. Seismicity is expected to remain high in all parts of Japan over the coming years.

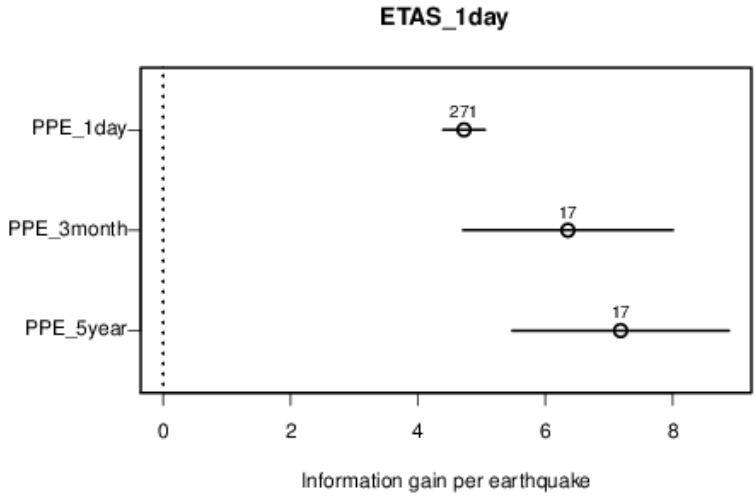

Figure 6. Information gain of the ETAS one-day model over the PPE month with updating at intervals of one day, three months and five years, for earthquakes with magnitude M > 3.95 in the CSEP New Zealand test region during the period from 4 September 2010 to 8 March 2011. About 75% of the earthquakes in this period were in the Darfield aftershock region.

4. Testing of GEM Models

The Global Earthquake Model (GEM) project is developing the first homogeneous hazard and risk model for the entire world. The GEM is being assembled from data, known physical properties, statistical descriptions of physical phenomena, and expert opinion. The GEM Testing Center at GFZ Potsdam, led by Danijel Schorlemmer, is collaborating with CSEP in its software development and testing programs. For seismicity rate model testing, T&E uses existing CSEP testing capabilities at the SCEC CSEP testing center and is creating new testing experiments that combine GEM models with other CSEP-registered models (e.g. RELM models). The joint tasks include the analysis of the test results for the 3-month model experiment in California and comparative tests with UCERF2.

Because the CSEP software distribution implements the complete work-flow for unbiased and prospective tests, it is the perfect basis for extensions covering tests of other phenomena. Ground Motion T&E has developed procedures for testing intensity prediction equations (IPEs) based on the CSEP testing center software, and a similar development is underway for ground-motion prediction equations (GMPEs). T&E is also developing testing procedures for testing complete earthquake hazard and risk models. The first target is the testing of the 1996 national seismic hazard map against 17 years of ground motion observations, which will be done in close collaboration with SCEC and USGS Golden.

Publications

Listed below are publications of research sponsored wholly or in part by this project:

Field, E. H., G. P. Biasi, P. Bird, T. E. Dawson, K. R. Felzer, D. D. Jackson, K. M. Johnson, T. H. Jordan, C. Madden, A. J. Michael, K. R. Milner, M. T. Page, Y. Parsons, T., Powers, P.M., Shaw, B.E., Thatcher, W.R., Weldon, R.J., II, and Zeng (2013), Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model: U.S. Geological Survey Open-File Report 2013–1165, 96 p., California Geological Survey Special Report 228, and Southern California Earthquake Center Publication 1792, http://pubs.usgs.gov/of/2013/1165/, submitted for publication to *Bull. Seismol. Soc. Am.*, August 2013.

Gerstenberger, M. C., S. Steacy, C. Williams, D. A. Rhoades and A. Christophersen (2011). Development of static Coulomb modelling to forecast aftershocks using a multi-model approach, Abstract S13A-2260 presented at 2011 Fall Meeting, AGU, San Francisco, Calif., 5-9 Dec.

Jordan, T. H. (2012), Lessons of L'Aquila for Operational Earthquake Forecasting, *Seismol. Res. Lett.*, **84**, 1-4, doi:10.1785/0220120167.

Marzocchi, W., J. D. Zechar, and T. H. Jordan (2012), Bayesian forecast evaluation and ensemble earthquake forecasting, *Bull. Seismol. Soc. Am.*, **102**, 2574-2584, doi:10.1785/0120110327.

Nanjo, K. Z., H. Tsuruoka, S. Yokoi, Y. Ogata, G. Falcone, N. Hirata, Y. Ishigaki, T. H. Jordan, K. Kasahara, K. Obara, D. Schorlemmer, K. Shiomi, and J. Zhuang (2012), Predictability study on the aftershock sequence following the 2011 Off the Pacific Coast of Tohoku, Japan, earthquake: First results, *Geophys. J. Int.*, 191, 653-658, doi:10.1111/j.1365-246X.2012.05626.x.

Rhoades, D. A., M. C. Gerstenberger, A. Christophersen and M. Liukis (2013). Utilising short-term and medium-term forecasting models for earthquake hazard estimation in the wake of the Canterbury earthquakes, *GNS Science Consultancy Report* 2013/141.

Rhoades, D. A., M. C. Gerstenberger, M. Liukis and A. Christophersen (2013). Retrospective Testing and CSEP Model Development in New Zealand, Oral presentation at Collaboratory for the Study of Earthquake Predictability Technical Workshop, Los Angeles, 8-9 May 2013.

Rhoades, D. A., M. Liukis, M. C. Gerstenberger and A. Christophersen (2013). Retrospective test of an operational earthquake forecasting model for Canterbury. Oral presentation to Symposium S301, IASPEI 2013 Assembly, Gothenburg, 22-26 July 2013.

Rhoades, D. A., M. C. Gerstenberger, A. Christophersen, J. D. Zechar, D. Schorlemmer, M. J. Werner, and T.H. Jordan (2014), Regional Earthquake Likelihood Models: Information gains of multiplicative hybrids, *Bull. Seismol. Soc. Am.*, submitted.

Steacy, S., M. C. Gerstenberger, C. Williams, D. A. Rhoades, and A. Christophersen (2013). A new hybrid Coulomb/statistical model for forecasting aftershock rates. *Geophys. J. Int.*, doi:10.1093/gji/ggt404.

Tsuruoka, H., N. Hirata, D. Schorlemmer, F. Euchner, K. Z. Nanjo, and T. H. Jordan (2012), CSEP Testing Center and the first results of the earthquake forecast testing experiment in Japan, *Earth Planets Space*, **64**(8), 661-671, doi:10.5047/eps.2012.06.007.

Williams, C.A. S. Steacy, M. C. Gerstenberger, and D. A. Rhoades (2012). Using static Coulomb models with rate- and state-friction models to estimate seismicity rates for the Canterbury, New Zealand, Earthquake sequence. Poster presentation at SCEC 2012 Annual Meeting, Palm Springs.

Zechar, J. D., Schorlemmer, D., Werner, M. J., Gerstenberger, M. C., Rhoades, D. A., & Jordan, T. H. (2013). Regional earthquake likelihood models I: First-order results. *Bull. Seismol. Soc. Am.*, **103**(2A), 787-798.

References

International Commission on Earthquake Forecasting for Civil Protection (2009). *Operational Earthquake Forecasting: State of Knowledge and Guidelines for Implementation, Findings and*

Recommendations, 12 pp., released by the Dipartimento della Protezione Civile, Rome, Italy, October, 2009.

International Commission on Earthquake Forecasting for Civil Protection (2011). *Operational Earthquake Forecasting: State of Knowledge and Guidelines for Implementation, Final Report*, 84 pp. *Ann. Geophys.*, **54**(4), 315-391, doi:10.4401/ag-5350.

Jordan, T. H., and L. M. Jones (2010). Operational earthquake forecasting: Some thoughts on why and how, *Seismol. Res. Lett.*, **81**, 571-574, doi:10.1785/gssrl.81.4.571.