Final Report of 2012 SCEC Proposal # 12022

An Event-Based Earthquake Predictability Experiment

Maximilian J. Werner (Princeton) and Agnès Helmstetter (U. Grenoble)

July 10, 2013

Summary

To investigate the predictive skills of earthquake forecasting models, we developed a suite of event-based, short-term, and intermediate-term forecasting models for California. The event-based forecasts are specified by the conditional intensity function of the ETAS (Epidemic-Type AfterShock) class and evaluated using their exact likelihood functions, rather than the discrete, grid-based, Poisson likelihood function currently used in CSEP experiments. We evaluated the influence of eight popular spatial triggering kernels on the probability gain and found that power-law kernels with scale parameter that grows with mainshock rupture length work best. We also found that lowering the learning catalog threshold to $m2+$ improves forecasts of target earthquakes $m3.95+$, providing further evidence that small quakes improve the predictive skill of clustering models. This event-based experiment has not yet been installed within CSEP because (i) the current prototype code is written in Matlab and should be transferred to open-source software (e.g., Python), and (ii) the likelihood calculation involves numerical integration that requires significant software development at CSEP. We are working with CSEP IT and scientists to develop a strategy for implementation.

A simpler solution to increase probability gains is to reduce the forecast horizon of existing CSEP experiments from one day to one hour or less. This approach leaves some scientific issues unsolved, but requires little software development. To compete in such a class, and to understand the influence of forecast horizon on probability gain, we developed sub-24 hour earthquake forecasts based on the ETAS model [*Werner et al.*, 2011] and two new models K^2 and K^3 [*Helmstetter and Werner*, 2013], which are based entirely on adaptive kernel smoothing of seismicity in time, space and magnitude. 1-hour forecasts reach gains of 200 over time-independent, spatially-varying forecasts – significantly larger than the gains of 55 of 24-hour forecasts. When sub-24hr testing becomes available within CSEP (expected shortly), we will install these models.

In the meantime, we have installed 24-hour versions of our ETAS and K^3 models, as well as 3-month and 1-year versions of a new intermediate-term smoothed seismicity model (CONAN – COupled NeAr-Neighbor smoothed seismicity model), based on adaptive smoothing of space-time seismicity [*Helmstetter and Werner*, 2012]. Since installation of ETAS and K^3 within CSEP in September 2012, 30 earthquakes $m \geq 3.95$ have occurred within the testing region. Although it is too early to judge the practical significance of these results, all models thus far perform better than any other installed one-day model, including the STEP model by *Gerstenberger et al.* [2005] and the critical-branching model by *Kagan and Jackson* [2010a].

1 Event-based (Point-Wise) Earthquake Forecasts for California

1.1 Background

Currently, CSEP conducts earthquake forecasting experiments with multiple forecast horizons ranging from one day to five years. For instance, 24-hour forecasts are generated at midnight for the next day and are based on a learning catalog up to the time when a forecast is issued. The forecast is specified as a time-independent expected rate of earthquakes over the forecast horizon and cannot be modified once issued. Forecasts are thus frozen, irrespective of what occurs during the 24-hour window. Most physical and statistical models, however, predict abrupt and large changes of seismicity right after an earthquake. Thus, allowing models to update more frequently should lead to larger probability gains. Moreover, the assumption of a time-independent Poisson rate during 24-hours is very clearly wrong. Nonetheless, the Poisson assumption allows CSEP to calculate a likelihood score for each forecast according to the simple formula:

$$\log L_{CSEP} = \sum_{i_t} \sum_{i_x} \sum_{i_y} \sum_{i_m} \log p\left(N_p(i_t, i_x, i_y, i_m)|n\right) \tag{1}$$

where n is the number of observed events in each bin that is compared with the expected rate N_p using the Poisson distribution $p(n|N_p)$. This experiment design can lead to biased test results [*Werner and Sornette*, 2008; *Lombardi and Marzocchi*, 2010], but it is a simple design that can be easily evaluated automatically.

The direct calculation of the likelihood function for models specified by an instantaneous conditional intensity function solves the issues introduced by the Poisson assumption (at the cost of a more involved calculation, section 1.4). First, the likelihood function is calculated exactly using the times, locations and magnitudes of earthquakes and incorporating all inter-dependencies between quakes as specified by a model. This circumvents the need for discrete spatial cells, magnitude bins or fixed intervals of time. Second, the likelihood is exact rather than approximated by a Poisson distribution. Therefore, evaluating models that can be specified by conditional intensity functions should lead to higher resolving power between models with more accurate estimates of probability gain.

The conditional intensity $\lambda(t, m, \vec{r}|H_t)$ is defined as the instantaneous probability of an event occurring in the volume dtdmd$\vec{r}$ given the history of events $H_t = \{t_i, m_i, \vec{r_i}\}_{t_i<t}$ [*Daley and Vere-Jones*, 2003]:

$$\lambda(t, m, \vec{r}|H_t) = \Pr(\text{one event occurs in } [t, t+dt), [m, m+dm), [\vec{r}, \vec{r}+d\vec{r})) \tag{2}$$

In particular, ETAS is defined by

$$\lambda(t, m, \vec{r}|H_t) = p(m)\left[\mu(\vec{r}) + \sum_{t_i<t} k e^{\alpha(m_i - m_d)} \frac{(1-p)c^{1-p}}{(t-t_i+c)^p} S_{\theta_s}(\vec{r} - \vec{r_i}, m_i)\right] \tag{3}$$

Here, $\mu(\vec{r})$ is a time-independent, spatially-variable background density and $p(m)$ is the (separable) magnitude density. The sum calculates contributions to the current hazard from all prior events i, which decay in time according to the Omori-Utsu law, increase with magnitude m_i according to the productivity law and decay in space according to some spatial kernel $S_{\theta_S}(\vec{r} - \vec{r_i}, m_i)$. The magnitude threshold m_d defines an arbitrary magnitude above which the model is applied and the parameters $\theta = \{k, \alpha, p, c, \theta_s\}$ need to be estimated.

The rate equation (3) is continuous in time and space until it is updated by the next earth-

quake. The predictive skill of this model class can therefore be evaluated exactly at a set of target observations $\{t_i, m_i, \vec{r}_i\}_{i=1,N}$ over a test volume of duration $[S, T)$, region R and magnitude interval $[m_d, \infty)$, by calculating the conditional log-likelihood function [*Daley and Vere-Jones*, 2003]

$$\log L = \sum_{S<t_i<T} \log \lambda(t_i, m_i, \vec{r}_i|H_{t_i}) - \int_{m_d}^{\infty} \int_R \int_S^T \lambda(t, m, \vec{r}|H_t) \mathrm{dt}\ \mathrm{d}\vec{r}\ \mathrm{dm} \tag{4}$$

To generate current discrete-interval forecasts that are evaluated using $\log L_{CSEP}$, most modelers simply integrated the rate (3) over each space-time-magnitude bin to obtain the expected number N_p per bin [*Werner et al.*, 2011], while others may assume $N_p = \lambda$ [*Kagan and Jackson*, 2010b]. This is clearly suboptimal, as discussed above.

1.2 Method

We developed a set of computer codes for such an event-based earthquake predictability experiment that uses (4) to evaluate models. We employed the popular ETAS class [*Ogata*, 1988] and analyzed the influence of different spatial kernels on the probability gain.

For event-based forecasts, we have thus far considered two isotropic kernels (2D Gaussian and power-law distributions), each with a scale or variance that is parameterized in four different ways, according to popular choices in the literature, leading to a total of eight different isotropic kernels. The Gaussian functions with standard deviation $\sigma(m)$ are defined by

$$\text{Gaussian}: \qquad f(x,y) = \frac{1}{2\pi\sigma(m)^2} e^{-\frac{x^2+y^2}{2\sigma(m)^2}}, \tag{5}$$

where $\sigma(m)$ either is a constant σ_0 (denoted 'Gauss'), or scales with magnitude according to either $\sigma(m)^2 = \sigma_0 e^{\log(10)\alpha(m-m_c)}$ ('Gauss Ogata' [*Ogata*, 1998]) or $\sigma(m) = 0.5 + f_d 0.01 \times 10^{0.5m}$ ('Gauss Helmstetter' [*Helmstetter et al.*, 2006; *Werner et al.*, 2011]) or $\sigma(m)^2 = \epsilon^2 + s_r^2 10^{m-4}$ ('Gauss Kagan' [*Kagan*, 1991]). The power-law functions with scale $d(m)$ are defined by

$$\text{Power}-\text{law}: \qquad f(x,y) = \frac{q-1}{\pi d(m)^2} \left(1 + \frac{x^2+y^2}{d(m)^2}\right)^{-q}, \tag{6}$$

where $d(m)$ either is a constant d_0 ('Power') or scales according to either $d(m)^2 = d_0 e^{\log(10)\alpha(m-m_c)}$ ('Power Ogata' [*Ogata*, 1998]), $d(m) = 0.5 + f_d 0.01 \times 10^{0.5m}$ ('Power Helmstetter' [*Helmstetter et al.*, 2006; *Werner et al.*, 2011]) or $d(m)^2 = d_0 e^{\gamma(m-m_c)}$ ('Power Zhuang' [*Zhuang et al.*, 2004]).

1.3 Results

In Figure 1, we compare the information gains of event-based forecasts from the eight kernels based on Californian target earthquakes $m \geq 3.95$ between 1992 and 2012. We used the ANSS catalog from 1981 as learning catalog. The information gain $I = (LL - LL_{TI})/N$ measures the average log-likelihood ratio per earthquake of the forecast compared with a time-independent, spatially-variably forecast. In this case, we used a reference model (named CONAN) based on adaptive space-time smoothing [*Helmstetter and Werner*, 2012], as detailed below. We also use this model as the background model for short-term forecasts. The results show that scaling the aftershock kernel with mainshock rupture length clearly improves the forecasts; power-laws outperform Gaussians; and Zhuang's power-law formulation appears best. Additionally, using small $m \geq 2$ quakes as triggering sources clearly improves the forecasts.

1.4 CSEP Testing Center Software Design Considerations

The evaluation of the event-based log-likelihood (4) requires the calculation of two terms $\log L = LL_1 - LL_2$. The first term in (4) is easy, as the evaluation software can simply ask a modeler's code to provide the value of λ at the target location $t_i, m_i, \vec{r_i}$ after providing the full learning catalog of past earthquakes t_j, where $t_j < t_i$. This can be done repeatedly for all target quakes.

The second term in (4), however, requires a multi-dimensional integration of the conditional intensity over time, space and magnitude. Multi-dimensional integration usually involves Monte Carlo integration, which we have not yet tested. Instead, we used properties of the conditional intensity function of ETAS to simplify the integral. Because ETAS defined by (3) has a separable magnitude density, the magnitude dimension integrates to one. Similarly, the sum and integral can be interchanged, and the space and time integrals separated, thus simplifying the integral to:

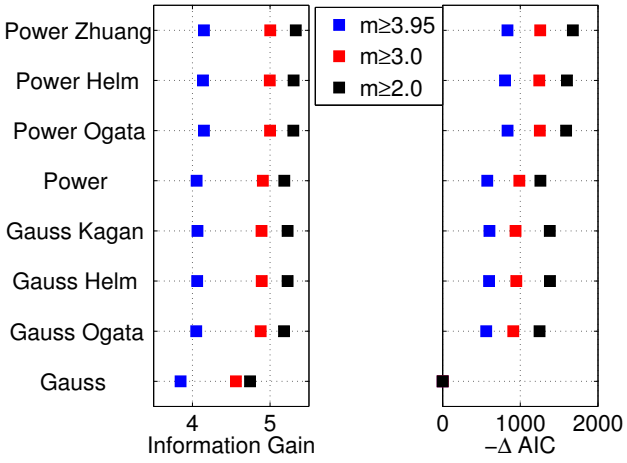

Figure 1: *(Left) Information gains of ETAS models for Californian targets m*3.95+ *between 1992 and 2012, using sources [3.95+, 3.0+, 2.0+] since 1985. (Right) Negative AIC referenced to the simple Gaussian kernel, i.e., better fitting models, penalized for extra parameters, to the right.*

$$LL_2 = \int_R \int_S^T \mu(\vec{r})\mathrm{dt}\ \mathrm{d}\vec{r} + \sum_{t_i<t} k e^{\alpha(m_i - m_d)} \int_S^T \frac{(1-p)c^{1-p}}{(t-t_i+c)^p}\mathrm{dt} \int_R \mathrm{S}_{\theta_s}(\tilde{\mathrm{r}} - \tilde{\mathrm{r}}_\mathrm{i}, \mathrm{m_i})\ \mathrm{d}\vec{r} \tag{7}$$

Furthermore, the background rate can be specified as a normalized spatial density (that integrates to one over the testing region) multiplied by the expected rate $\mu_s(T-S)$ and Omori's law analytically integrated, leaving:

$$LL_2 = \mu_s(T-S) + \sum_{t_i<t} k e^{\alpha(m_i - m_d)} \Phi(T-t_i) \int_R S_{\theta_s}(\vec{r} - \vec{r_i}, m_i)\ d\vec{r} \tag{8}$$

Therefore, the sole remaining integral that may need to be evaluated numerically is the spatial integral I_S over the triggering kernel. Usually, this integral involves integrating the spatial aftershock footprint of small quakes that is relatively small compared to the large testing region, and the integral is essentially one. However, this may not be the case for large earthquakes, especially close to the boundary of the testing region (both inside and outside). Numerically calculating spatial kernels for all earthquakes takes a long time. Instead, we implemented a scheme to test whether epicenters are sufficiently close to a boundary such that the integral might differ significantly from one. This cut-off distance, beyond which we assumed $I_S = 1$, should depend on the spatial triggering kernel. We used $d_{max} = 2 * 5.92\sigma$ [*Zechar and Jordan*, 2010] for the Gaussian kernels with standard deviation σ (which might depend on the magnitude m of the triggering event), and $d_{max} = 100 * d_0$ for the power-law kernels, where d_0 is the scale parameter. For the remaining earthquakes for which calculating the integral I_S was necessary, we used Matlab's integral2 function to integrate the spatial densities in 0.1 by 0.1 degree spatial cells within the testing region. We only

used spatial cells within d_{max} to calculate the integral. Figure 2 shows an example of the cells that were integrated to calculate the integral of the aftershock kernel of the 2010 El Mayor Cucapah earthquake.

This solution has advantages and disadvantages. First, it was written in Matlab for simplicity and the author's preference, but the CSEP testing center prefers open-source software, such as Python. The codes we developed would thus need to be transferred. Secondly, the scheme we developed makes heavy use of the simplifying assumptions of the ETAS model. The reduction of the multi-dimensional integral to a simple 2D spatial integration requires (i) a separable magnitude density, (ii) a separable spatial and temporal aftershock triggering kernel, and (iii) a linear formulation of the triggering contributions from past quakes. The class of models that fit these requirements is clearly narrow. At the same time, we note that most short-term models, including most ETAS flavors and the critical branching model by Kagan and Jackson, fit into this class. We have not explored to what extent more complex models would worsen the computation of LL_2, in part because there are no obvious candidates available.

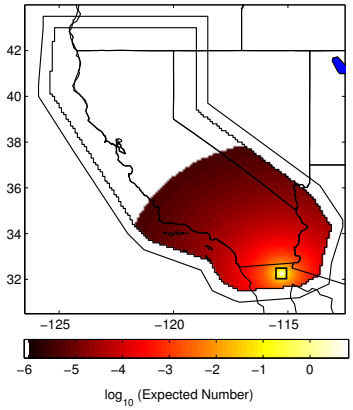

Figure 2: *Illustration of the 2D spatial aftershock kernel that needs to be integrated numerically to compute the event-based log-likelihood score. Here, the aftershock kernel of the 2010 El Mayor Cucapah earthquake is shown as an example of a large earthquakes near the testing region boundary. Our method sums the numerical integrals from each 0.1 by 0.1 degree cell that contributes significantly.*

Based on these considerations, we recommend the following for CSEP software development. First, the simpler solution to achieving larger probability gains and allowing models to update after quakes is to simply reduce the testing interval from the current 24-hours to one hour or less (section 2.2). Second, we recommend building up capabilities for event-based testing by starting with the kind of simple ETAS model we described above, and using the simplified log-likelihood calculation that involves only the 2D spatial integral. Third, more complex models will eventually drive the software requirements, and at some point Monte Carlo integration may become necessary. Finally, an interesting alternative may be the use of residual analysis, which diagnoses models using the conditional intensity, but without requiring any integration. At a recent CSEP workshop in June 2013 at USC, the global CSEP community ranked this expansion of CSEP activities as important but, because of the additional required software development, endorsed the recommendation of concentrating on a reduction of the testing interval to one hour or less before exploring event-based testing.

2 Fixed-Interval Earthquake Forecasts for California

2.1 3-month and 1-year Forecasts: Model CONAN

We developed an intermediate-term, time-independent earthquake forecasting model that employs spacetime kernels to smooth seismicity [*Helmstetter and Werner*, 2012]. The major advantage of the method is that it does not require prior declustering of the catalog, circumventing the relatively subjective choice of a declustering algorithm. Past earthquakes are smoothed in space and time

using adaptive Gaussian kernels. The bandwidths in space and time associated with each event are a decreasing function of the seismicity rate at the time and location of each earthquake. This yields a better resolution in spacetime volumes of intense seismicity and a smoother density in volumes of sparse seismicity. The long-term rate in each spatial cell is then defined as the median value of the temporal history of the smoothed seismicity rate in this cell. To calibrate the model, the earthquake catalog is divided into two parts: the early part (the learning catalog) is used to estimate the model, and the latter one (the target catalog) is used to compute the likelihood of the models forecast. We optimized the models parameters by maximizing the likelihood of the target catalog. To estimate the kernel bandwidths in space and time, we used a coupled near-neighbor method, after which the model is named: CONAN (COupled NeAr-Neighbor model). We applied these methods to Californian seismicity and compared the resulting forecasts with our previous method based on spatially smoothing a declustered catalog [*Werner et al.*, 2011]. All models use small $m \geq 2$ earthquakes to forecast the rate of larger earthquakes and use the same learning catalog. Our new preferred model slightly outperforms our previous forecast, providing a probability gain per earthquake of about 5 relative to a spatially uniform forecast. We have installed 3-month and 1-year versions of CONAN within the SCEC Testing Center in September 2012. CONAN also serves as the background model for our short-term and event-based forecasts.

2.2 Sub-24 hour ETAS Forecasts

To explore the influence of the forecast horizon on the predictive skill of Omori-clustering models, we developed a suite of ETAS forecasts and optimized the model's parameters for horizons ranging from 1 hour to 1 day. In Figure 3, we show the information gains per earthquake for each forecast horizon, for a variety of learning catalog thresholds. We used the ANSS catalog in Calfornia from 1981 as learning catalog, and earthquakes $m \geq 3.95$ during 1986-2013 as targets. The probability gain of 1-hour forecasts reaches 200, while that of 24-hour forecasts is closer to 55. The results also show that including small quakes $m \geq 2$ improves forecasts.

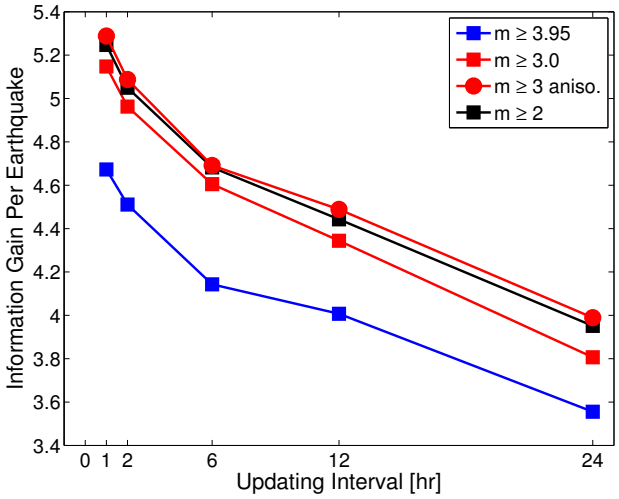

Figure 3: *Information gains of the ETAS model with isotropic (squares) and anisotropic (circles) 'power Helm' kernel for different forecast horizons and learning catalog thresholds.*

We also investigated the influence of an anisotropic spatial aftershock triggering kernel [*Helmstetter et al.*, 2006; *Werner et al.*, 2011]. Following [*Helmstetter et al.*, 2006; *Werner et al.*, 2011], we replaced the isotropic kernel of large $m5.5+$ quakes by the sum of fixed-bandwidth kernels from early aftershocks. For instance, a large strike slip earthquake would initially cast an isotropic aftershock kernel; after the first few hours, however, many aftershocks would have occurred along the rupture plane of the mainshock. At the next forecast update, these early aftershocks are smoothed using fixed-bandwidth kernels, and their normalized sum captures both the extent of the rupture and its orientation. In Figure 3, we show that this method leads to information gain increases for the fixed-interval forecasts tested so far (using $m3+$ as sources, $m5.5+$ treated anisotropically) that are comparable to those obtained by adding an order of magnitude more triggering sources ($m2+$).

Sub-24-hour experiments are currently under software development at the CSEP testing center and expected to complete in the coming months. We intend to submit this ETAS model to the forecast group when it is available.

2.3 24-hour Forecasts: Non-Parametric Kernel Models K^2 and K^3

We developed new methods for short-term earthquake forecasting that employ space, time and magnitude kernels to smooth seismicity [*Helmstetter and Werner*, 2013]. These methods are purely statistical and rely on very few assumptions about seismicity. In particular, we do not use Omori's law, and only one of our two new models assumes a Gutenberg-Richter law to model the magnitude distribution; the second model estimates the magnitude distribution non-parametrically with kernels. We employ adaptive kernels of variable bandwidths to estimate seismicity in space, time and magnitude bins. To project rates over short time-scales into the future, we simply assume persistence, that is, a constant rate over short time windows. The resulting forecasts from the two new kernel models are compared with those of the ETAS model generated by *Werner et al.* [2011]. In Figure 4, we show 24-hour forecasts generated by the three models for 1992/6/30, two days after the M7.3 Landers earthquake. Also shown is the common background model based on the intermediate-term CONAN model described in section 2.1. While our new methods are simpler and require fewer parameters than ETAS, the obtained probability gains are surprisingly close. Nonetheless, ETAS performs significantly better in most comparisons, and the kernel model with a Gutenberg-Richter law attains larger gains than the kernel model that non-parametrically estimates the magnitude distribution. Finally, we found that combining ETAS and kernel model forecasts, by simply averaging the expected rate in each bin, can provide greater predictive skill than ETAS or the kernel models can achieve individually.

We submitted the ETAS and K^3 models to the CSEP testing center in September 2012. There, our models compete with the models submitted by other researchers in the forecast group of daily forecasts of earthquakes with target magnitudes $M_t \geq 3.95$. We submitted two versions of each model, one uses all prior earthquakes $M_d \geq 2$ as the learning catalog, while the other uses $M_d \geq 3$. For each learning catalog threshold, we also submitted a simple ensemble model that averages the forecasts of the ETAS and K^3 models.

Between the date of installation at CSEP and the date of the latest available test results (31 May 2013 as of the time of writing, 1 July 2013), 30 earthquakes $M \geq 3.95$ occurred within the testing region. Thus far, our models appear to perform well and as intended. Although it is too early to judge the practical significance of these results, all three models thus far perform better than any other installed one-day model. Within both the $M_d \geq 2$ and the $M_d \geq 3$ forecast group, K^3 achieved the highest gain, followed by the ensemble model and ETAS. According to the T and W-tests, all three models currently achieve significantly larger gains than any other model, including the STEP model by *Gerstenberger et al.* [2005] and the critical-branching model by *Kagan and Jackson* [2010a].

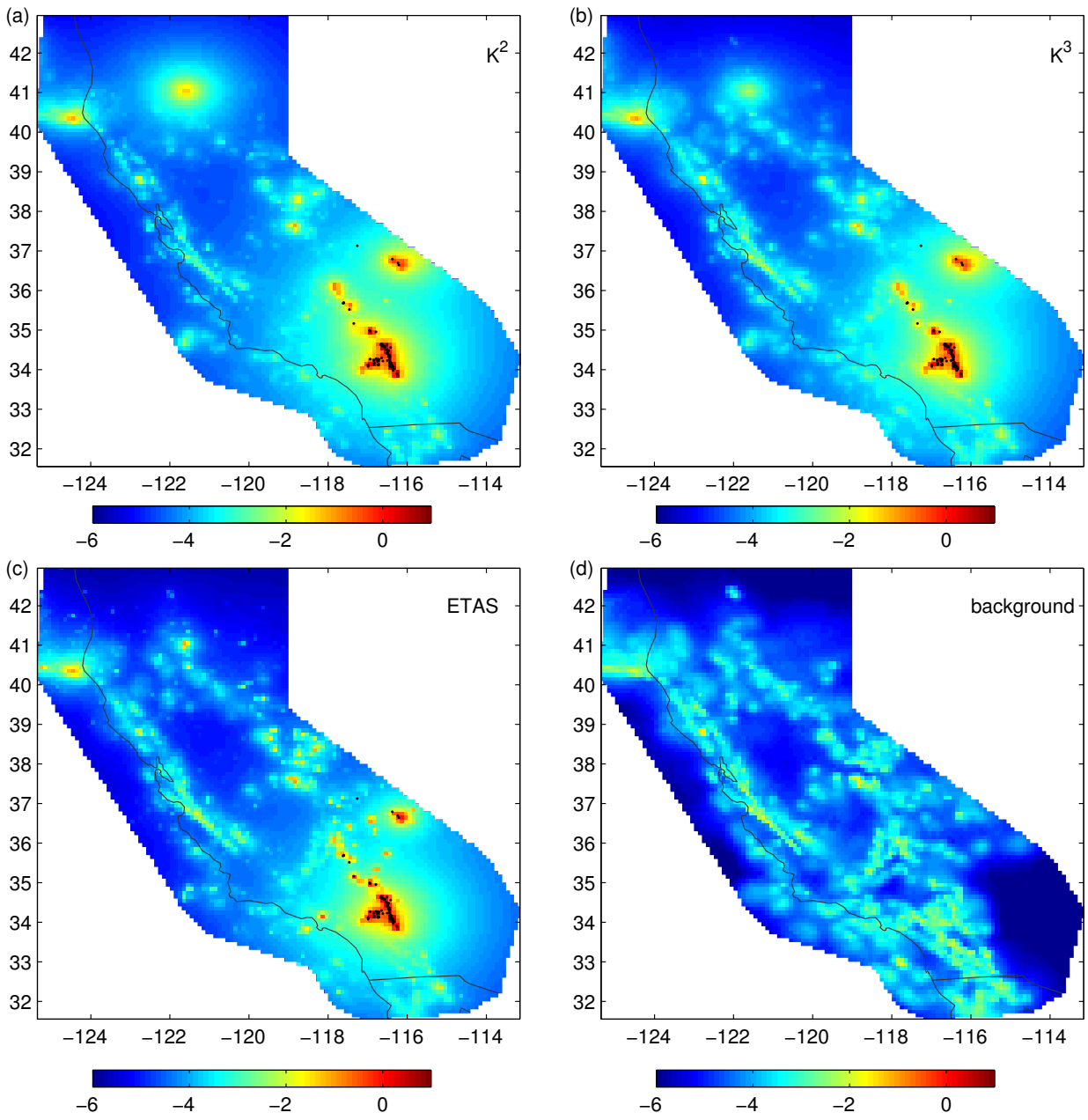

Figure 4: Map of predicted rate of $M \geq 3$ earthquakes for day 1992/6/30, two days after the Landers $M = 7.3$ mainshock, for (a) K^2, (b) K^3 and (c) ETAS models, in logarithmic scale. Black dots represent $M \geq 3$ target earthquakes which occurred on day 1992/6/30. The background rate used by ETAS model is shown in (d) and is based on a new intermediate-term CONAN model [*Helmstetter and Werner*, 2012], described in section 2.1. Kernel models use the same background rate as ETAS but with a slightly different amplitude (μ_0=0.55 background events per day for ETAS, 0.35 for K^2 and 0.39 for K^3).

References

Daley, D. J., and D. Vere-Jones (2003), *An Introduction to the Theory of Point Processes*, vol. I, Springer, New York, USA.

Gerstenberger, M. C., S. Wiemer, L. M. Jones, and P. A. Reasenberg (2005), Real-time forecasts of tomorrow's earthquakes in California, *Nature*, *435*(7040), 328–331.

Helmstetter, A., and M. J. Werner (2012), Adaptive spatiotemporal smoothing of seismicity for long-term earthquake forecasts in California, *Bulletin of the Seismological Society of America*, *102*(6), 2518–2529.

Helmstetter, A., and M. J. Werner (2013), Adaptive smoothing of seismicity in time, space and magnitude for time-dependent earthquake forecasts in California, *Bulletin of the Seismological Society of America*, in revision.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2006), Comparison of short-term and time-independent earthquake forecast models for southern California, *Bull. Seismol. Soc. Am.*, *96*(1), doi:10.1785/0120050067.

Kagan, Y., and D. Jackson (2010a), Short- and long-term earthquake forecasts for California and Nevada, *Pure and Applied Geophysics*, *167*(6-7), 685–692, doi:10.1007/s00024-010-0073-5.

Kagan, Y. Y. (1991), Likelihood analysis of earthquake catalogs, *Geophys. J. Intern.*, *106*, 135–148.

Kagan, Y. Y., and D. D. Jackson (2010b), Short- and long-term earthquake forecasts for california and nevada, *Pure and Appl. Geophys.: The Frank Evison Volume*, *167*(6/7), 685–692, doi:10.1007/s00024-010-0073-5.

Lombardi, A., and W. Marzocchi (2010), The assumption of poisson seismic-rate variability in CSEP/RELM experiments, *Bulletin of the Seismological Society of America*, *100*(5A), 2293–2300, doi:10.1785/0120100012.

Ogata, Y. (1988), Statistical models for earthquake occurrence and residual analysis for point processes, *J. Am. Stat. Assoc.*, *83*, 9–27.

Ogata, Y. (1998), Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, *5*(2), 379–402.

Werner, M. J., and D. Sornette (2008), Magnitude uncertainties impact seismic rate estimates, forecasts and predictability experiments, *J. Geophys. Res. Solid Earth*, *113*(B8), doi:10.1029/2007JB005427.

Werner, M. J., A. Helmstetter, D. D. Jackson, and Y. Y. Kagan (2011), High resolution long-term and short-term earthquake forecasts for California, *Bull. Seismol. Soc. Am.*, *101*(4), doi:10.1785/0120090340.

Zechar, J. D., and T. H. Jordan (2010), The area skill score statistic for evaluating earthquake predictability experiments, *Pure and Appl. Geophys.*, *167*(8/9), doi:10.1007/s00024-010-0086-0.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2004), Analyzing earthquake clustering features by using stochastic reconstruction, *J. Geophys. Res.*, *109*(B05301), doi:10.1029/2003JB002879.