

Short-Term Earthquake Forecasting Using Early Aftershock Statistics

by Peter Shebalin, Clément Narteau, Matthias Holschneider, and Danijel Schorlemmer

Abstract We present an alarm-based earthquake forecast model that uses the early aftershock statistics (EAST). This model is based on the hypothesis that the time delay before the onset of the power-law aftershock decay rate decreases as the level of stress and the seismogenic potential increase. Here, we estimate this time delay from $\langle t_g \rangle$, the time constant of the Omori–Utsu law. To isolate space–time regions with a relative high level of stress, the single local variable of our forecast model is the E_a value, the ratio between the long-term and short-term estimations of $\langle t_g \rangle$. When and where the E_a value exceeds a given threshold (i.e., the c value is abnormally small), an alarm is issued, and an earthquake is expected to occur during the next time step. Retrospective tests show that the EAST model has better predictive power than a stationary reference model based on smoothed extrapolation of past seismicity. The official prospective test for California started on 1 July 2009 in the testing center of the Collaboratory for the Study of Earthquake Predictability (CSEP). During the first nine months, 44 $M \geq 4$ earthquakes occurred in the testing area. For this time period, the EAST model has better predictive power than the reference model at a 1% level of significance. Because the EAST model has also a better predictive power than several time-varying clustering models tested in CSEP at a 1% level of significance, we suggest that our successful prospective results are not due only to the space–time clustering of aftershocks.

Online Material: The EAST model; retrospective test and nine-month prospective evaluation versus time-dependent rate-based models.

Introduction

The aftershock rate, Λ , is often well-described by the Omori–Utsu law,

$$\Lambda(t) = \frac{K}{(t+c)^p}, \quad (1)$$

where t is the elapsed time since the mainshock, K the aftershock productivity, p a power-law exponent, and c the time delay before the onset of the power-law aftershock decay rate (Utsu, 1961). The c value is often interpreted, at least partially, as an artifact (Kagan, 2004; Lolli and Gasperini, 2006) related to difficulties in detecting events shortly after a mainshock. The possible reasons are aftershocks buried in coda waves, overlapping aftershock records, catalog compiler overload, absence or malfunction of seismic stations close to the source zone, and other technical or administrative factors. For example, the analysis of high-resolution aftershock records at Parkfield, California, and in Japan shows that a careful identification of microearthquakes directly after mainshocks leads to lower c values than the ones obtained from standard earthquake catalogs (Vidale *et al.*, 2004; Peng *et al.*, 2006, 2007; Enescu *et al.*, 2007, 2009). Nevertheless, even in these studies, the c value is never equal to zero (Nanjo *et al.*, 2007). Furthermore, after analyzing catalogs of seismicity

above completeness magnitude thresholds, several studies have also pointed out that the c value may change according to the magnitude range of aftershocks (Shcherbakov *et al.*, 2004) and/or the recurrence time of major earthquakes in the area (Narteau *et al.*, 2002). From these findings, it can be argued that many properties of the aftershock decay rate over short times remain partly unknown, despite evidence that they can provide a useful source of information about the physical mechanisms of earthquake triggering.

Theoretical models for aftershock production may explain systematic variations of the c value. Considering that aftershocks result from a steplike perturbation of the stress field in the neighborhood of a triggering event, the vast majority of these models assume that the amplitude of this perturbation is inversely proportional to the time delay before the onset of the power-law aftershock decay rate. For example, this is the case in rate-and-state friction models (Dieterich, 1994), in damage mechanics models (Shcherbakov and Turcotte, 2004; Ben-Zion and Lyakhovsky, 2006), and in static fatigue models (Scholz, 1968; Narteau *et al.*, 2002). Then, assuming that the amplitude of the near-field stress redistribution is directly proportional to the level of stress imposed by the large-scale tectonic forcing (Atkinson,

1987), a decreasing c value may be directly related to the accumulation of stress within the seismogenic crust.

To test this idea, Narteau *et al.* (2005) studied stacks of aftershocks sequences in southern California from the previous 20 years. To avoid artifacts arising from overlapping records, they did not consider large earthquakes and their aftershock sequences (Shebalin, 2004); they analyzed only the largest aftershocks of mainshocks in the magnitude range of $2.5 < M < 4.5$ and disregarded data during the first minutes after the mainshocks. Then, they observed that the c value suddenly increases after large earthquakes (e.g., Landers, Hector Mine) and slowly decreases at a constant rate during periods of low seismicity. Using the same approach along the San Andreas fault (SAF) system in central California, Narteau *et al.* (2008) also analyzed the evolution of the c value during aseismic deformation and transient slip events. Combining all these observations, they showed that in creep-slip and stick-slip zones the rate of change of the c value is proportional to the deficit of slip rate along the SAF. For a linear elastic rheology, it follows that the rate of change of the c value can be directly related to the level of stress.

Inspired by this result and the systematic variations of b , the slope of the earthquake-size distribution (Schorlemmer *et al.*, 2005), Narteau *et al.* (2009) used Californian and Japanese earthquake catalogs with focal mechanisms to study the evolution of the c value across different faulting styles. They observed that c values for thrust earthquakes are on average three times smaller than c values for normal ones, taking intermediate values for strike-slip events. Assuming that thrust, strike-slip, and normal faulting are associated with a decreasing value of differential shear stress (Sibson, 1974), the time delay before the onset of the power-law aftershock decay rate can be related once again to the level of stress in the seismogenic crust.

With these observations, we see a growing body of theoretical and observational evidence that suggests a dependency of the early aftershock decay rate on stress. Here, we build on this dependence for developing an alarm-based forecast model that uses early aftershock statistics at much more local scales. This alarm-based model, called EAST for its use of early aftershock statistics, is designed to detect places more prone to moderate and large earthquakes within an active fault zone. It satisfies all the requirements for being tested in testing center of the Collaboratory for the Study of Earthquake Predictability (CSEP; Jordan, 2006; Zechar *et al.*, 2010): forecasts are provided on a regular predefined spatial grid and for a predefined time interval and magnitude ranges.

In this paper, we introduce the alarm function of the EAST model, compare it to the c value, and show how these two quantities are related to each other. Then, we present the EAST model (The East Forecast Model section) and a retrospective evaluation of its performance using Molchan diagrams (Retrospective Evaluation of the EAST Model in California section). In the The EAST Model in the CSEP Testing Center for California section, we present the results of the first nine months of official CSEP testing, which started

on 1 July 2009. Finally, we discuss the stability of our results and, with respect to the magnitude of the target earthquakes, the duration for which prospective tests have to be conducted to determine the predictive power of the model.

Early Aftershocks Statistics: Geometric Mean of Aftershock Times versus c Value

To compute the time delay before the onset of the power-law aftershock decay rate, the standard approach is the maximum-likelihood estimate of the c value in equation (1). If only short time series of aftershocks are available, the maximum-likelihood estimator has a large variance. Moreover, an implicit equation has to be solved for each estimation. Finally, it is not clear that the Omori–Utsu law is a good description of the aftershock decay rate over short time. Therefore, given the difficulty in precisely estimating the aftershock decay rate over a short time, we propose the geometric mean of elapsed times from the mainshock to aftershocks, $\langle t_g \rangle$, as a simplified proxy for c . Considering a single aftershock sequence, we define

$$\langle t_g \rangle := \sqrt[n]{\prod_{i=1}^n t_i} = \exp \left[\frac{1}{n} \sum_{i=1}^n \ln(t_i) \right], \quad (2)$$

where t_i , $i \in [1, 2, \dots, n]$, is the time between the i th aftershock in a fixed time interval $[t_{\text{start}}, t_{\text{stop}}]$ and the mainshock.

Considering that the aftershock decay rate follows the Omori–Utsu law, there is a relation between the $\langle t_g \rangle$ value and the c value for a given p value and a given time interval after a mainshock. In more general terms, let us suppose that aftershock production follows a nonhomogeneous Poisson process with rate $\lambda(t)$. In this case, considering a time interval $[t_{\text{start}}, t_{\text{stop}}]$ after a mainshock, the distribution density of the elapsed time from mainshock to aftershocks is

$$f(t) = \frac{\lambda(t)}{\int_{t_{\text{start}}}^{t_{\text{stop}}} \lambda(u) du}, \quad t \in [t_{\text{start}}, t_{\text{stop}}]. \quad (3)$$

In the same time interval, the expectation value of the logarithm of the elapsed time from mainshock to aftershocks is

$$\mathbb{E}(\ln(t)) = \int_{t_{\text{start}}}^{t_{\text{stop}}} \ln(t) f(t) dt. \quad (4)$$

We now suppose that λ obeys the Omori–Utsu law. Because the observed p values are always close to 1, we fix the value to $p = 1$ in equation (1). Then, we have

$$\lambda(t) = \frac{K}{(c+t)}. \quad (5)$$

It follows that

$$\mathbb{E}(\ln(t)) = \frac{\int_{t_{\text{start}}}^{t_{\text{stop}}} \frac{\ln(u)}{u+c} du}{\ln(t_{\text{stop}} + c) - \ln(t_{\text{start}} + c)} =: h(c). \quad (6)$$

For a specific aftershock sequence, an estimator of this expectation value is simply the arithmetic mean of the

logarithms of the elapsed times from the mainshock to aftershocks

$$\ln(\langle t_g \rangle) = \frac{1}{n} \sum_{i=1}^n \ln(t_i). \quad (7)$$

Therefore, we have an estimator, \hat{c} , for c by solving the implicit equation

$$h(\hat{c}) = \ln(\langle t_g \rangle). \quad (8)$$

Note that because the function h does not depend on the aftershock productivity K , this estimator does not depend on K either. An explicit expression for equation (6) can be given in terms of a polylogarithm special function

$$h(c) = \frac{\ln(t_{\text{stop}}) \ln(1 + \frac{t_{\text{stop}}}{c}) - \ln(t_{\text{start}}) \ln(1 + \frac{t_{\text{start}}}{c}) + \text{Li}_2(-\frac{t_{\text{stop}}}{c}) - \text{Li}_2(-\frac{t_{\text{start}}}{c})}{\ln(1 + \frac{t_{\text{stop}}}{c}) - \ln(1 + \frac{t_{\text{start}}}{c})}, \quad (9)$$

where

$$\text{Li}_2(z) = - \int_0^z \frac{\ln(1-u)}{u} du = \sum_{k=1}^{\infty} \frac{z^k}{k^2}. \quad (10)$$

Figure 1 shows two numerical solutions for $t_{\text{stop}} = 10^{-1}$ days: $t_{\text{start}} = 10^{-7}$ days (solid line) and $t_{\text{start}} = 10^{-3}$ days (dotted line). In both cases, we observe that $\langle t_g \rangle$ is a monotonic function of c with slopes that are slightly

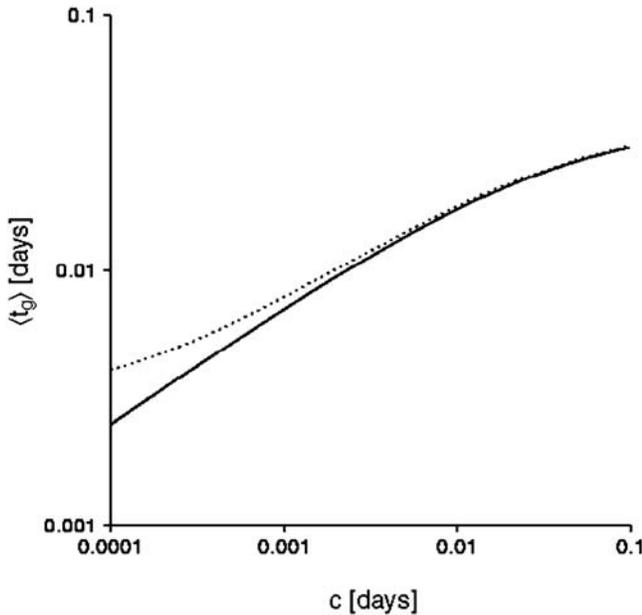


Figure 1. The geometric mean of elapsed times from mainshocks to aftershocks, $\langle t_g \rangle$, with respect to the c value. Numerical solutions of equation (9) for $t_{\text{stop}} = 10^{-1}$ days with $t_{\text{start}} = 10^{-6}$ days (solid line) and $t_{\text{start}} = 10^{-4}$ days (dotted line).

lower than 0.45 in bilogarithmic scale. In what follows, we only use the $\langle t_g \rangle$ value to quantify the characteristic time delay before the onset of the power-law aftershock decay rate, keeping in mind this relation to the c value.

The EAST Forecast Model

The EAST model is an alarm-based model with a single alarm function derived from $\langle t_g \rangle$, the geometric mean of elapsed times from mainshocks to aftershocks (equation 2). This alarm function is estimated locally in space and time to determine whether or not an alarm is issued. Where the alarm function exceeds a given threshold, an alarm is issued, and at least one target earthquake with a magnitude $M \geq M_{\text{target}}$ is expected to occur during the next time step δt . Here, the

magnitude threshold M_{target} is a free parameter that can be adjusted to account for different factors (e.g., duration of the analysis, level of seismic activity). To follow the rules of the CSEP testing center (Schorlemmer and Gerstenberger, 2007), we choose a time step $\delta t = 3$ months, and we discretize the area covered by the test into a square grid of side length $\delta = 0.1^\circ$. Thus, we decompose the entire space–time region into a three-dimensional grid of space–time cells noted $c(x, y, t)$. For each of these cells, we identify a set of mainshocks to stack the corresponding aftershock sequences with respect to the mainshock times.

Mainshock and Aftershock Selection

Our declustering method uses a set of parameters $\{M_{\text{target}}, r_1, r_2, T, D, T_{\text{short}}, T_{\text{long}}, t_{\text{start}}, t_{\text{stop}}, R_0, N_{\text{min}}, M_{\text{min}}^M, M_{\text{max}}^M, M_{\text{min}}^A\}$ that is described in Table 1. First, we eliminate all earthquakes that are aftershocks of events with a magnitude $M \geq M_{\text{target}}$ by removing earthquakes of magnitude smaller than M that are within a $r_1 10^{r_2 M}$ -km radius during the first T days after a magnitude $M \geq M_{\text{target}}$ earthquake. Second, an event is not considered a mainshock if there is at least one earthquake of the same or higher magnitude in the time interval $[-t_{\text{stop}}, t_{\text{stop}}]$. All remaining events are considered mainshocks. Their respective aftershocks are selected within R_0 -km radius in the time interval $[t_{\text{start}}, t_{\text{stop}}]$ after the mainshock's focal time. In the model, we always take $t_{\text{stop}} < 1$ days to concentrate on early aftershocks. Furthermore, we always take $t_{\text{start}} > 10$ s to limit artifacts related to aftershock catalog incompleteness immediately after mainshocks.

Another critical issue of our declustering method is that the selected earthquakes are classified in ranges of

Table 1
Parameter Values of the EAST Model and Their Ranges*

Parameter	Units	Value in the Main Test	Minimum Value	Maximum Value
M_{\min}^M		2.5	1.8	3.0
M_{\max}^M		4.5	3.4	8.0
M_{\min}^A		1.8	0	1.9
r_1	km	1×10^{-2}	0	3×10^{-2}
r_2		0.725	0	0.8
T	d	40	0	1×10^3
D	km	25	21	35
T_{short}	a	5	4.1	7.9
T_{long}	a	25	8	50
R_0	km	4	1.3	16
t_{start}	d	1×10^{-4}	0	1×10^{-3}
t_{stop}	d	1×10^{-1}	3×10^{-2}	3×10^{-1}
N_{\min}		3	2	5

*To determine these ranges, one parameter is changed at a time, the results are compared to the main retrospective results presented in Figure 3a ($M_{\text{target}} = 5$), and the difference of the integral of the Molchan trajectory from $\tau_{\text{RI}} = 0$ to $\tau_{\text{RI}} = 0.3$ should not be larger than 5%. Because the r_1 , r_2 , and T values can be equal to 0, it is not always necessary to eliminate mainshocks that are themselves aftershocks of large earthquakes, as this improves the predicting accuracy of the EAST model only slightly. The value of M_{\min}^A and t_{start} depend only on aftershock catalog completeness, and smaller values may be used. This is also the case for higher M_{\max}^M and T_{long} values. The M_{\max}^A value is given by the magnitude of the mainshock. Finally, only six parameters cannot be eliminated from the procedure: M_{\min}^M , D , T_{short} , R_0 , t_{stop} , and N_{\min} .

mainshocks, $[M_{\min}^M, M_{\max}^M]$, and ranges of aftershocks, $[M_{\min}^A, M_{\max}^A]$. Thus, we investigate only earthquakes from particular magnitude ranges that can be shown to be complete even in early times of the aftershock sequence. Mainshocks should be sufficiently small to have shorter coda waves such that subsequent aftershocks can be reliably detected. Aftershocks should be sufficiently large to ensure completeness at that particular magnitude level from the early times of an aftershock sequence. These magnitude ranges are determined from seismological and statistical constraints following Narteau *et al.* (2009) (Table 1): M_{\min}^M is determined by the global catalog completeness; M_{\max}^M is the mainshock magnitude below which the mean magnitude of a given range of aftershocks remain stable; M_{\min}^A is the aftershock magnitude for which the estimation of the c value remains stable as we increase the time at which we start the fit. On the other hand, M_{\max}^A is not a parameter of the model any more because here we simply consider the magnitude of the mainshock to use the maximum number of events for each aftershock sequence.

The Alarm Function

Within the study area, each space–time cell $c(x, y, t)$ is associated with two space–time volumes with the same spatial extent, a circle with diameter D , but two different time intervals $T_{\text{short}} \ll T_{\text{long}}$. The first volume covers the recent period from $t - T_{\text{short}}$ to t . The second one covers the preceding period from $t - T_{\text{long}}$ to $t - T_{\text{short}}$. In these two

space–time volumes, we identify mainshocks and stack their aftershocks by sorting them according to the elapsed time from their respective mainshocks. Let us consider that N_{short} and N_{long} are the numbers of aftershocks in the two stacks. If $N_{\text{short}} < N_{\min}$, we do not define the value of the alarm function of the corresponding space–time cells $c(x, y, t)$. If $N_{\text{short}} \geq N_{\min}$ and $N_{\text{long}} \geq N_{\min}$, we use the stacks of aftershocks as individual sequences to estimate the $\langle t_g \rangle_{\text{short}}$ and $\langle t_g \rangle_{\text{long}}$ values from equation (2). If $N_{\text{short}} \geq N_{\min} > N_{\text{long}}$, we take the same equation to calculate the $\langle t_g \rangle_{\text{short}}$ value but consider that $\langle t_g \rangle_{\text{long}} = t_{\text{stop}}$. Then, we use these short-term and long-term estimations of $\langle t_g \rangle$ to calculate E_a , the alarm function of the EAST model:

$$E_a(x, y, t) = \frac{\langle t_g \rangle_{\text{long}}}{\langle t_g \rangle_{\text{short}}}. \quad (11)$$

The $\langle t_g \rangle_{\text{long}}$ value is used as a reference level to identify abnormally small values of $\langle t_g \rangle_{\text{short}}$ (i.e., abnormally small c values). Thus, we try to isolate space–time regions with a relative high level of stress that are currently more prone to earthquakes. In practice, E_a values above a given threshold, E_a^0 , define the space–time cells occupied by alarms. In these cells, earthquakes with magnitude $M \geq M_{\text{target}}$ are expected to occur during the next time step δt . In addition, larger E_a values should correspond to sites that are currently more vulnerable to target earthquakes.

Retrospective Evaluation of the EAST Model in California

Before starting prospective testing of the EAST model in the CSEP testing center, we evaluated the model’s performance in retrospective tests for the testing region of California (Schorlemmer and Gerstenberger, 2007) for the period from 1984 to 2008. Thus, we test our underlying hypothesis and analyze the stability of the EAST model with respect to the model parameters and the choice of the space–time region under consideration. According to the CSEP experiment, we always use the same time step $\delta t = 3$ months and the same spatial mesh of $0.1^\circ \times 0.1^\circ$.

Table 1 shows the values of the parameters of the EAST model and, for each of them, the range of values for which the results are stable. During the testing period, approximately 50% of $M \geq 5.0$ earthquakes occurred near Cape Mendocino (mostly offshore), in the California–Nevada border zone, and in Mexico. In these zones, the density of seismic stations is lower and the number of aftershocks in the stacks is smaller. Hence, we prefer not to consider the entire CSEP California testing region (dashed line in Fig. 2) in the retrospective analysis. Instead, we concentrate on a more central region where the catalogs of early $M > M_{\min}^A$ aftershocks are likely to be more complete (solid line in Fig. 2). Nevertheless, even in this case, the E_a value cannot be defined everywhere because of an insufficient number of aftershocks, and some target earthquakes occurred in cells

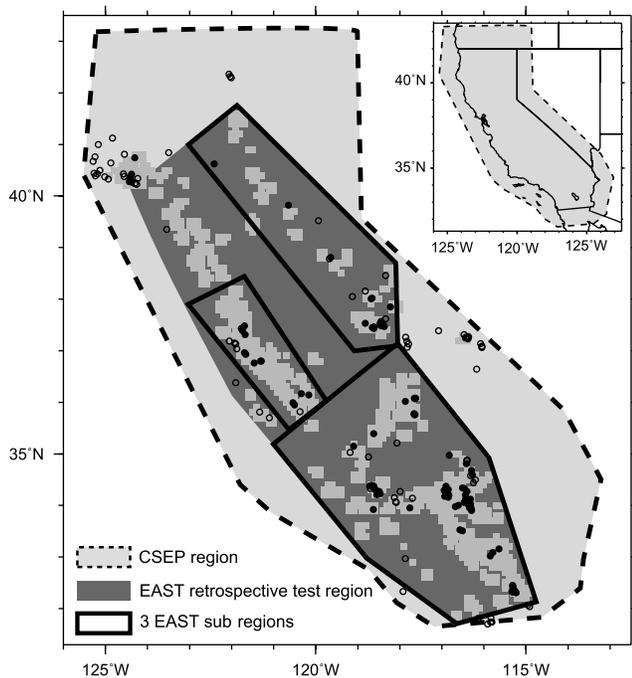


Figure 2. Testing regions and past seismicity in California. Dashed lines limit the CSEP California testing region. The dark gray zone shows the region in which we perform the retrospective analysis of the EAST model. Solid lines limit three subregions chosen for additional tests: southern, central, and north-west California. Gray cells $c(x, y)$ show where the E_a value is defined on 1 January 2009. For a period from 1984 to 2008, black circles show epicenters of $M \geq 5$ earthquakes that occur in zones where the E_a value can be defined. For the same time period, open circles show epicenters of $M \geq 5$ earthquakes that occur in zones where the E_a value cannot be defined.

with undefined E_a value (see the distributions of gray cells and open circles in Fig. 2, respectively).

CSEP testing centers host various classes of forecast models. Most of them are rate based. Their outputs are estimates of the expected rate of earthquakes in prespecified space–time–magnitude bins. In contrast, alarm-based forecast models do not try to quantify earthquake rates. Instead, they concentrate on the identification of space–time regions with high seismogenic potential. Alarm-based models have at least one control parameter, the so-called alarm function (Zechar and Jordan, 2008). This alarm function is evaluated at every point in space and time. Its higher values correspond to a higher, but usually unknown, probability of earthquakes. Then, because it is impossible to estimate this probability, a threshold value is used to determine if an alarm is issued locally. By adjusting the amplitude of this threshold, it is possible to cover any fraction between 0 and 1 of the space–time region with alarms. Then, the quality of the prediction has to be quantified with respect to the entire range of threshold values.

Molchan diagrams (Molchan, 1990, 1991) are usually used to evaluate the performance of the space–time prediction of an alarm-based model with respect to a chosen reference-model or null hypothesis (Molchan and Keilis-Borok,

2008, Zechar and Jordan, 2008). A binary outcome of the forecasts is considered. If a target earthquake occurs within an alarm, then a successful prediction is scored. Otherwise it is a failure to predict or a miss. Molchan diagrams plot the miss rate ν with respect to τ_{RI} , a variable that is often described as the fraction of the space–time region occupied by alarms. More exactly, it is the ratio of the Poissonian earthquake frequency of the region occupied by alarms to the Poissonian earthquake frequency of the entire space–time region. These frequencies are given by the chosen reference model. Therefore, $1 - \tau_{RI}$ is the expected miss rate of the reference model. In a Molchan diagram, each threshold value of the alarm function corresponds to a specific point. If the corresponding miss rate of the model is smaller than $1 - \tau_{RI}$, the model has better predictive power than the reference model. This comparison can be systematically implemented from $\tau_{RI} = 0$ (no alarm) to $\tau_{RI} = 1$ (permanent alarm) by decreasing the threshold value of the alarm function. Then, we can visually estimate from the so-called Molchan trajectory the predictive power of the model with respect to the reference model. An important property of this approach is the higher weight given to a successful prediction in a zone of low seismicity (small earthquake frequency) than in a zone of high seismicity (high earthquake frequency). Another property is that zones without seismicity are not taken into account.

As a null hypothesis, we here employ a relative intensity (RI) reference model, a time-independent model of earthquake frequencies. This reference model is obtained by smoothing the location of earthquakes in the past (Kossobokov and Shebalin, 2003; Helmstetter *et al.*, 2006; Molchan and Keilis-Borok, 2008; Zechar and Jordan, 2008). Obviously, it would be preferable to use a time-dependent reference model of seismicity, especially in a prospective test. Such a comparison of models will be done in the CSEP testing center during a regular experiment to which our model is submitted. Hence, at this stage, we prefer to keep the most commonly accepted RI-type time-independent model. Different models may have different spatial resolution, different smoothing length, and different time scales. Here, we use the same spatial resolution and smoothing length as in the EAST model in order to assess the gain in prediction power that may be attributed to the temporal aspect of our forecast model. Furthermore, mainshocks and aftershocks are identified using the declustering method of Gardner and Knopoff (1974), and we use the longest time period available with respect to catalog completeness in California before the beginning of the retrospective test. In practice, we use $M \geq 3$ earthquakes from 1960 to 1984. We suppose that this magnitude threshold is sufficiently low to guarantee a correct evaluation of the RI reference model in all spatial cells in which an earthquake may occur. Then, for each spatial cell $c(x, y)$ of the EAST model, the RI reference model is obtained by calculating the rate of $M > 3$ mainshocks in circles of diameter $D = 25$ km.

In what follows, Molchan diagrams are always constructed according to the same convention:

- The solid line is the Molchan trajectory calculated from the highest ($\tau_{\text{RI}} = 0$) to the smallest ($\tau_{\text{RI}} = 1$) threshold value E_a^0 of the alarm function.
- The dotted line is the Molchan trajectory that incorporates zones in which the E_a value cannot be defined because of an insufficient number of aftershocks in the stacks. To obtain this curve, we replace the alarm function of the EAST model by the seismic rate of the reference model in all the cells $c(x, y, t)$ in which the E_a value has not been defined. Then, when the threshold E_a^0 of the alarm function reaches its lowest value, we complement the space–time region occupied by the alarm with the cell $c(x, y, t)$, which has a decreasing seismic rate in the reference model. Thus, we have an increasing τ_{RI} value, and we can calculate the corresponding miss rate ν .
- The dashed diagonal line $\nu = 1 - \tau_{\text{RI}}$ corresponds to an unskilled forecast model with respect to the reference model, considering an infinite number of target earthquakes (see Zechar and Jordan, 2010, for a finite number of target earthquakes).
- The shaded area is the zone of the Molchan diagram in which the prediction of the EAST model is better than the prediction of the RI reference model at a level of significance of $\alpha = 1\%$ (Kossobokov and Shebalin, 2003; Shebalin et al. 2006; Zechar et al., 2007). In other words, only 1% of the best predictions of the RI reference model fall in the shaded region. To find the limit of this area at a given τ_{RI} value, we consider that the number of earthquakes predicted by the reference model is a random variable that follows a binomial distribution with parameters n and P . n is the number of $M > M_{\text{target}}$ earthquakes in the space–time region under consideration. P is the probability to predict a $M > M_{\text{target}}$ in this region according to the RI reference model. From this binomial distribution, we can

calculate the miss rate distribution of mean value $1 - \tau_{\text{RI}}$. Then, the upper limit of the shaded region is simply the α quantiles of this miss-rate distribution.

Figure 3 shows a comparison between the prediction of the EAST model and the prediction of the RI reference model for three different M_{target} values, $M_{\text{target}} \in \{5, 5.5, 6\}$. (Ⓔ The same comparisons for the entire CSEP testing region are available as an electronic supplement to this paper.)

In all three cases, the Molchan trajectories show that the EAST model has better predictive power than the RI reference model at a significance level of 1%. In addition, steeper Molchan trajectories with increasing M_{target} values suggest that the EAST model works better in zones where large events did occur. Equivalently, we may say that the apparent slope of the expected earthquake-size distribution is likely to decrease in the space–time region occupied by alarms.

In Figure 4 and Figure 5, we analyze in more detail the space–time structures of alarms for two threshold values of the alarm function: $E_a^0 = 2.0$ ($\tau_{\text{RI}} = 0.1$ in Fig. 3) and $E_a^0 = 1.0$ ($\tau_{\text{RI}} = 0.22$ in Fig. 3). Figure 4 shows the histogram of the duration of successful alarms for four different M_{target} values, $M_{\text{target}} \in \{4, 5, 5.5, 6\}$. Using the time step $\delta t = 3$ months, we obtain the duration of a successful alarm by counting the number of consecutive times the E_a value exceeded the threshold value E_a^0 before a predicted $M > M_{\text{target}}$ earthquake. The median of all distributions is approximately one year; and, in about 80% of the cases, the alarm duration is less than two years. These results suggest a characteristic time scale of less than a year for the forecast of the EAST model. In this case, the time step $\delta t = 3$ months used in the CSEP testing center is small enough to be implemented in the EAST model.

Figure 5 shows the evolution of τ_{RI} with respect to time for the entire region of the retrospective test (solid lines in Fig. 2) and three subregions (dotted lines in Fig. 2).

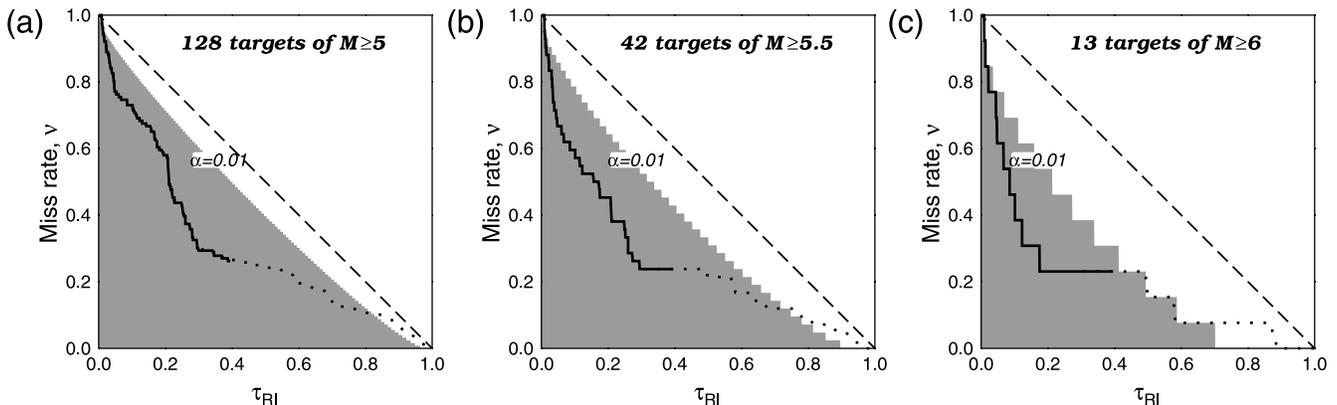


Figure 3. Retrospective evaluation of the EAST model in California from 1984 to 2008 for three M_{target} values: (a) $M_{\text{target}} = 5$, (b) $M_{\text{target}} = 5.5$, (c) $M_{\text{target}} = 6$. Using a Molchan diagram, we compare the prediction of the EAST model to the prediction of the RI reference model. The solid line is the Molchan trajectory calculated from the highest to the lowest threshold value E_a^0 of the alarm function. The dotted line is the Molchan trajectory that incorporates zones where the E_a value cannot be defined (see text). The dashed diagonal line corresponds to an unskilled forecast model with respect to the reference model. The shaded area indicates the zone of the Molchan diagram in which the prediction of the EAST model is better than the prediction of the RI reference model at a level of significance $\alpha = 1\%$.

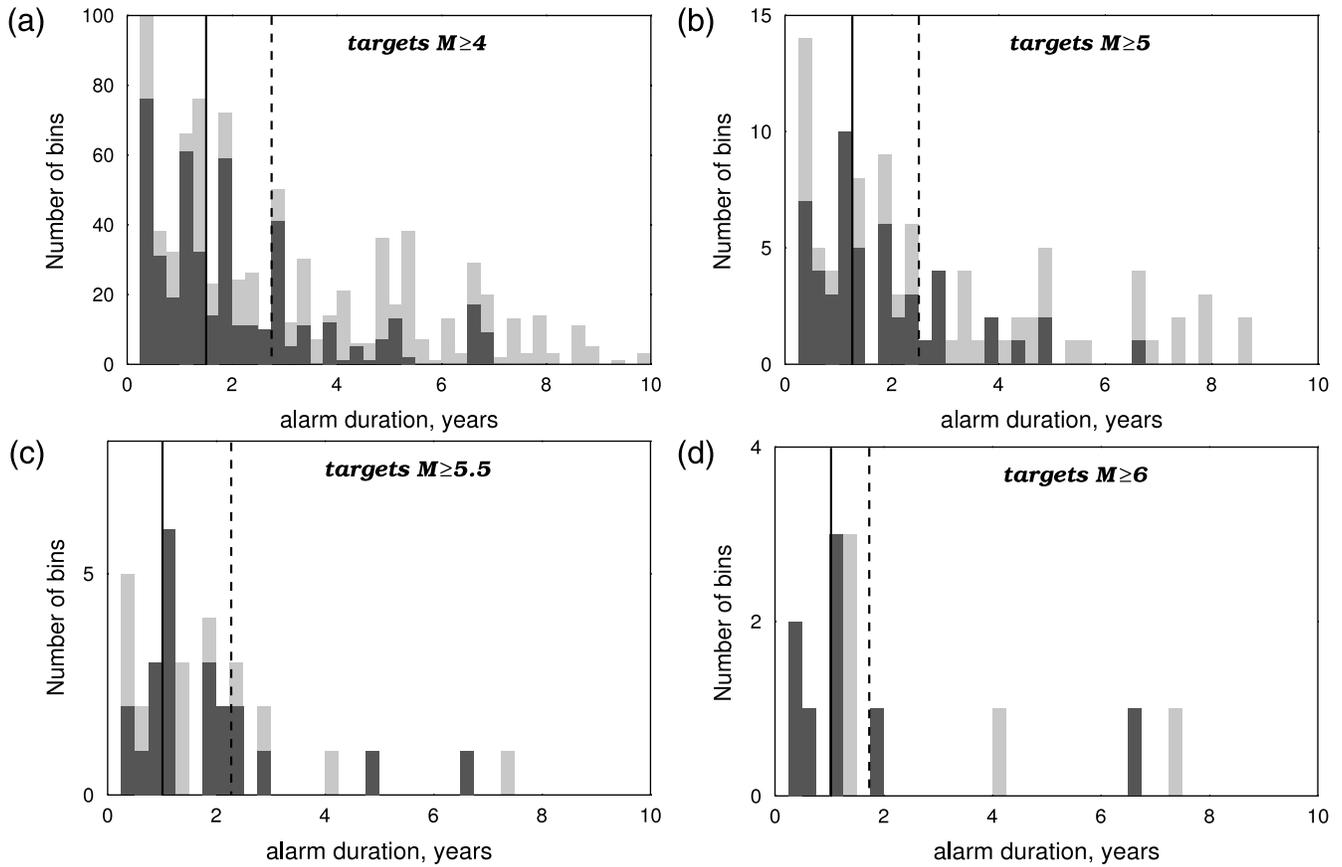


Figure 4. Distributions of the duration of successful alarms: (a) $M_{\text{target}} = 4$, (b) $M_{\text{target}} = 5$, (c) $M_{\text{target}} = 5.5$, and (d) $M_{\text{target}} = 6$. Histograms in light and dark gray correspond to $E_a^0 = 1$ and $E_a^0 = 2$, respectively. Solid and dashed lines mark the 50% and 80% quantiles, respectively.

The τ_{RI} values are systematically larger at the beginning of the retrospective test. Interestingly, the number of $M > M_{\text{target}}$ earthquakes is also larger in this earlier period. To check the stability of the prediction algorithm on two periods

with a different level of seismicity, we perform two independent analyses from 1 January 1984 to 30 June 1992 and from 1 July 1992 to 31 December 2008 (Fig. 6). In both cases, Molchan trajectories are quite similar to each other and

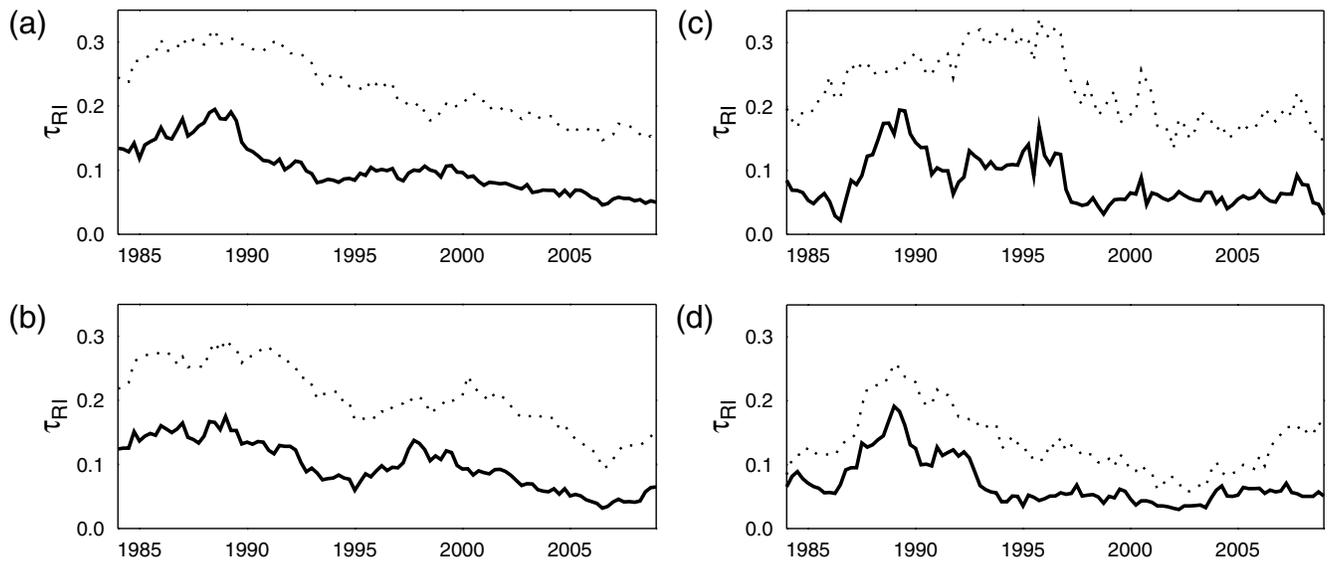


Figure 5. Temporal evolution of τ_{RI} in (a) the entire region covered by the test and in (b) southern, (c) central, and (d) northwest California (see the solid and dotted lines in Fig. 2). Dashed and solid lines correspond to $E_a^0 = 1$ and $E_a^0 = 2$, respectively.

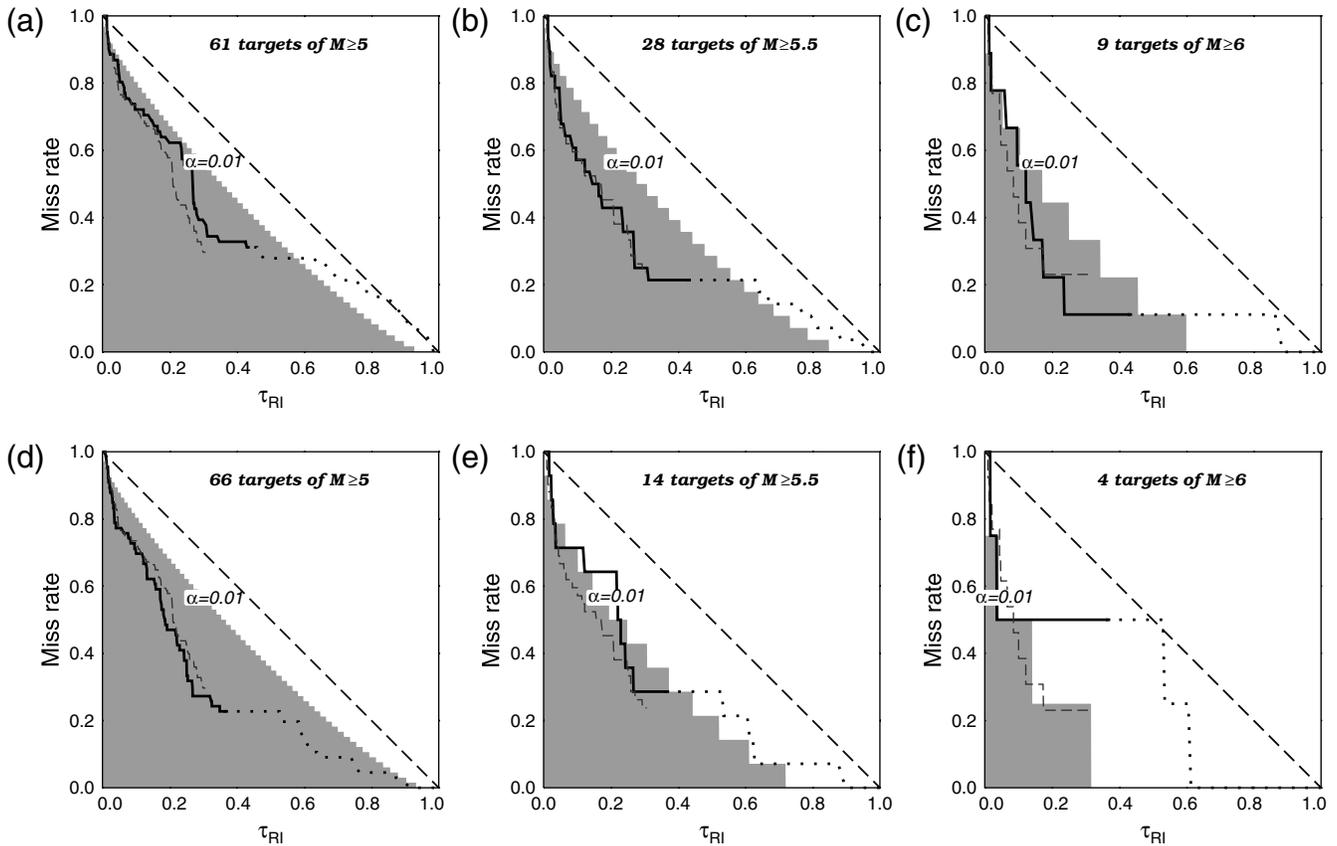


Figure 6. Retrospective evaluation of the EAST model in California for two time periods and three M_{target} values: (a and d) $M_{\text{target}} = 5$, (b and e) $M_{\text{target}} = 5.5$, and (c and f) $M_{\text{target}} = 6$. Using a Molchan diagram, we compare the prediction of the EAST model to the prediction of the RI reference model. Arbitrarily, we choose two time periods with approximately the same number of $M \geq 5.0$ earthquake from January 1984 to June 1992 (top diagrams) and from July 1992 to December 2008 (bottom diagrams). For comparison, a thin dashed line shows the result obtained in Figure 3 from January 1984 to December 2008.

do not exhibit any substantial changes when compared to the main retrospective result that covers the entire period (Fig. 3). We conclude that the EAST model is stable in time despite the permanent evolution of seismic activity and the associated variation of alarm distributions.

We believe that the EAST model can be applied worldwide in a large diversity of active tectonic settings. To verify that the EAST model can be generalized to different seismic zones, we use the same set of parameters to perform a retrospective analysis of three subregions in California (dotted lines in Fig. 2). Namely, we consider southern, central, and northeast California. Figure 7 shows the corresponding Molchan diagrams for three M_{target} values, $M_{\text{target}} \in \{5, 5.5, 6\}$. We see that the EAST model has a good performance in two of the three subregions. In northeast California, despite the small number of targets, the result is even better than for entire California (Fig. 3). About 50% of $M \geq 5.0$ earthquakes, 60% of $M \geq 5.5$ earthquakes, and both $M \geq 6.0$ earthquakes are forecast with τ_{RI} value as small as 4%. Nevertheless, results are worse in central California because we cannot distinguish between the miss rate of the EAST model and the miss rate of the RI reference model at a significance level of 1%. The main reason is that the epicenters of the Loma Prieta

(17 October 1989, M_w 6.9) and Parkfield (28 September 2004, M_w 6.0) events are in the space–time regions where the E_a value cannot be defined. However, high E_a values are observed nearby. Actually, the results in central California become as good as in other subregions if we decrease the minimum aftershock magnitude threshold from $M_{\text{min}}^A = 1.8$ to $M_{\text{min}}^A = 1.7$ and take the centroid locations instead of the epicenter locations for the Loma Prieta and Parkfield events.

Short-term and long-term estimations of $\langle t_g \rangle$ values range from 5×10^{-2} to 5×10^{-1} day in more than 99% of the cases. These values correspond to (1) E_a values close to unity, with extreme values in the range of $[10^{-2}, 10^2]$, and (2) c values that range from 10^{-4} to 10^{-2} day, taking $t_{\text{start}} = 10^{-7}$ day and $t_{\text{stop}} = 10^{-1}$ day in equation (9). These observations are in good agreement with reported c values in California for the same aftershock magnitude ranges (Narteau *et al.*, 2002; Shcherbakov *et al.*, 2004). Figure 8 shows EAST forecast maps before the three largest earthquakes in southern California during the period of the retrospective test: Landers (28 June 1992, M_w 7.3), Northridge (17 January 1994, M_w 6.7), Hector Mine (16 October 1999, M_w 7.1). These maps also give an idea of the temporal

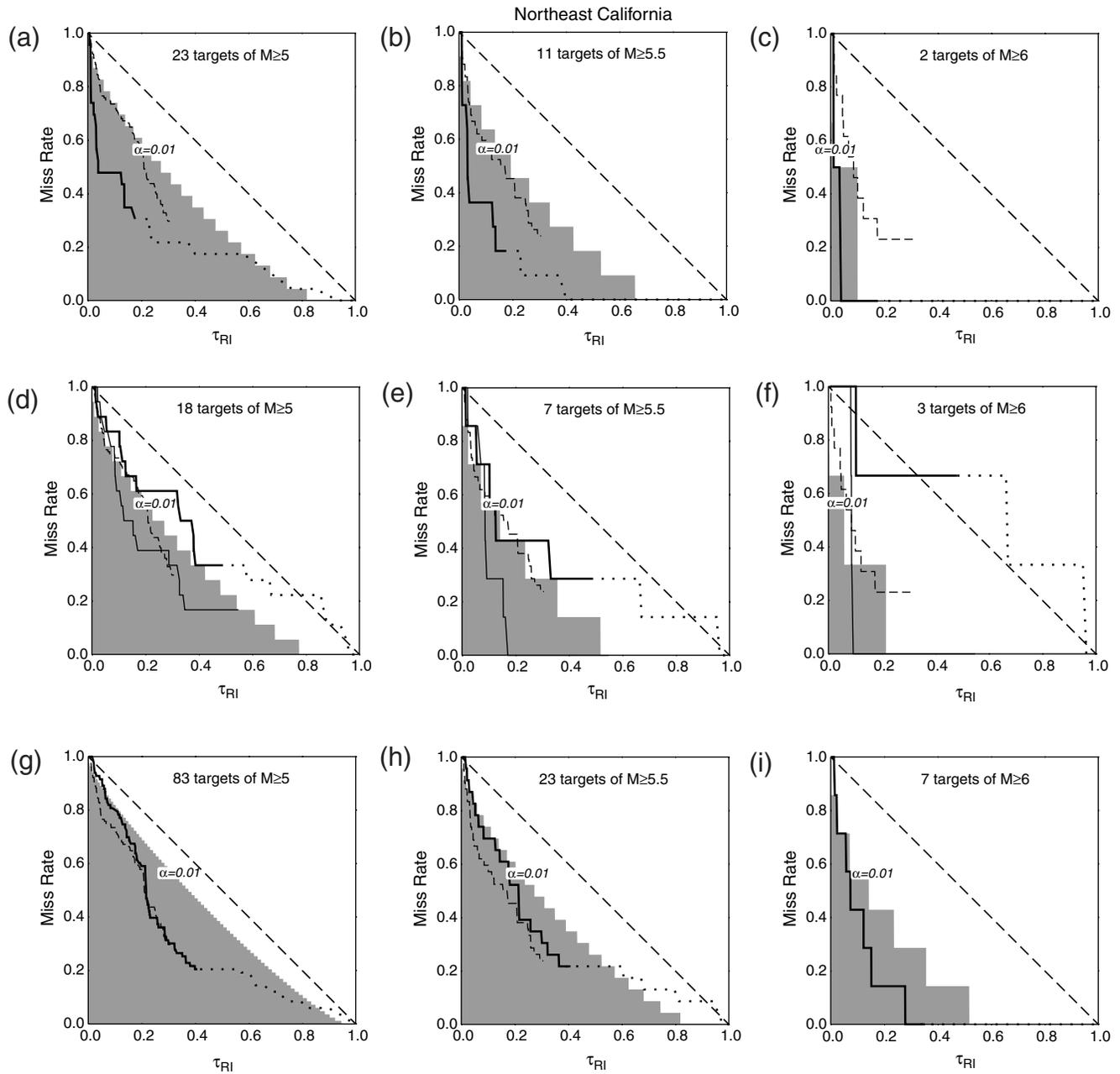


Figure 7. Retrospective evaluation of the EAST model in (top) northeast, (center) central, and (bottom) southern California from 1984 to 2008 for three M_{target} values: (left) $M_{\text{target}} = 5$, (center) $M_{\text{target}} = 5.5$, and (right) $M_{\text{target}} = 6$. Using a Molchan diagram, we compare the prediction of the EAST model to the prediction of the RI reference model. For comparison, thin dashed lines show the result obtained in Figure 3 for California. For central California, thin solid lines show the Molchan trajectory using $M_{\text{min}}^A = 1.7$ and centroid locations instead of the epicenter locations for Loma Prieta and Parkfield (see text).

variation of the forecasts from our model. For comparison, Figure 8d shows the forecast of the RI model that we used as a reference model in the Molchan diagrams. We use the same smoothing procedure for all these maps, but the RI map looks more smoothed than the EAST maps. The main reason is the absence of E_a values in about 90% of the space–time regions $c(x, y, t)$ due to the restrictive aftershock selection process (see the [Mainshock and Aftershock Selection](#) section). The RI reference model has nonzero values in many of these cells because of the long time period over which it has been

calculated; this is in contrast to the EAST maps, which are more fragmented and with sharp peaks.

The EAST Model in the CSEP Testing Center for California

We submitted the EAST model to the CSEP testing center for California in April 2009. To entirely satisfy the CSEP requirements, we use $M_{\text{target}} = 3.95$ and expand the prediction to the entire state of California (dashed lines in Fig. 2).

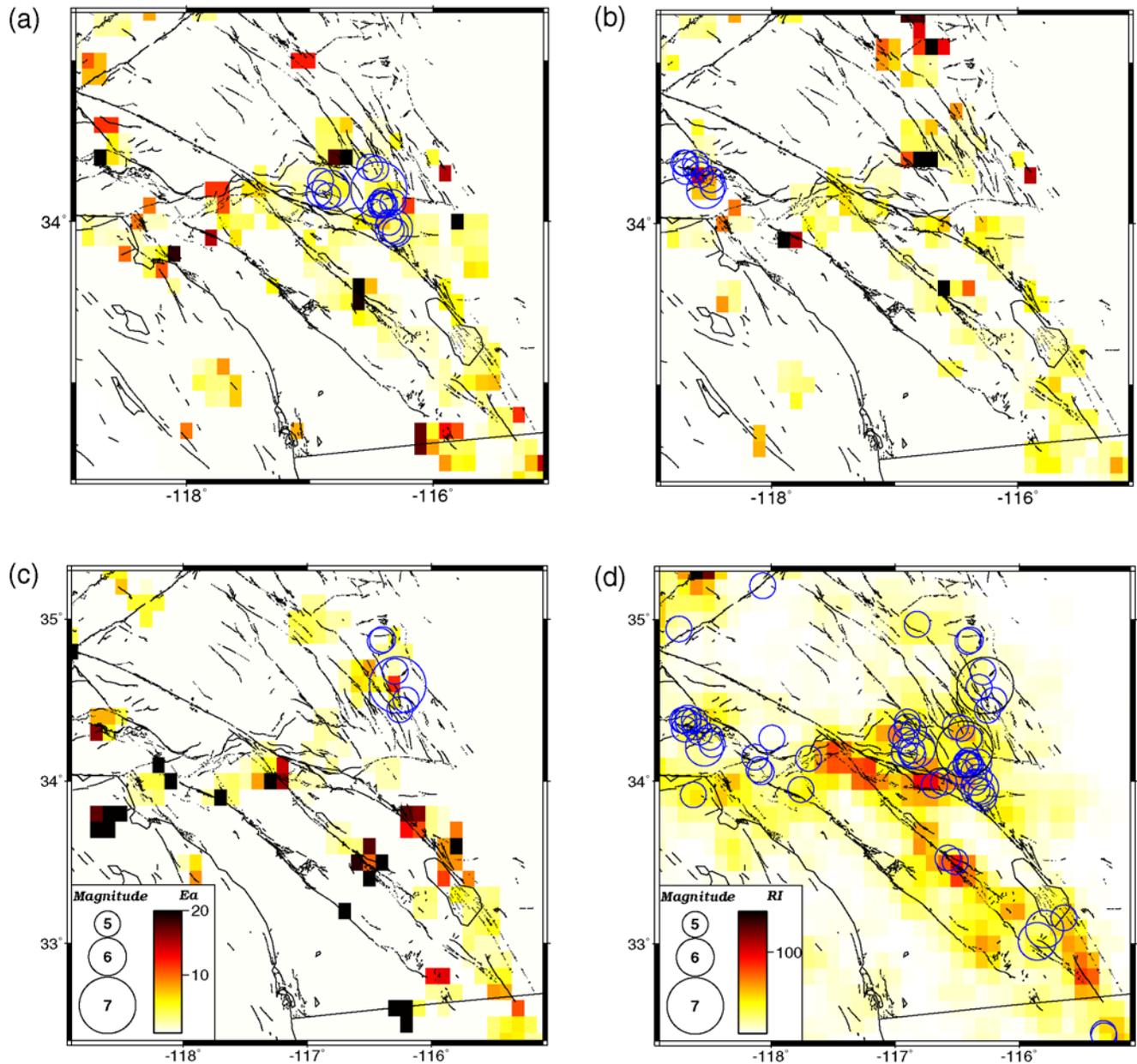


Figure 8. Examples of EAST forecast maps and RI map. EAST forecast maps show the distribution of E_a values before the three largest events in southern California from 1984 to 2008: (a) Landers forecast map from 1 April to 30 June 1992. (b) Northridge forecast map from 1 January to 31 March 1994. (c) Hector-Mine forecast map from 1 October to 31 December 1999. (a) and (b) have the same color scale as (c). Circles show epicenters of $M > 5$ earthquakes that occurred during the corresponding forecasts. For the RI map (d), shading represents a number of earthquakes. For each cell $c(x, y, t)$, we count $M \geq 3.0$ earthquakes in circles of diameter $D = 25$ km from 1960 to 1984. The color version of this figure is available only in the electronic edition.

Note that we expect an increasing number of misses outside the region of the retrospective test (solid lines in Fig. 2) because of an increase of the completeness threshold of the network, especially in offshore areas.

The prospective test of the EAST forecast model for California started officially on 1 July 2009. To date, three three-month forecast periods have ended; and, based on these nine months of seismicity, we can perform a comparison between the predictive power of the EAST model and the predictive power of the RI reference model. To produce

the RI maps, we use the same smoothing procedure as previously described here but count the number of $M \geq 3.0$ events in the period from 1960 to 2008. Nevertheless, in the retrospective analysis, we have observed that target earthquakes often occur in cells of high E_a value but of moderate RI frequency. Hence, we currently use a modified E_a value

$$E_{a1}(x, y, t) = \frac{E_a(x, y, t)}{\lambda_g(x, y)}, \quad (12)$$

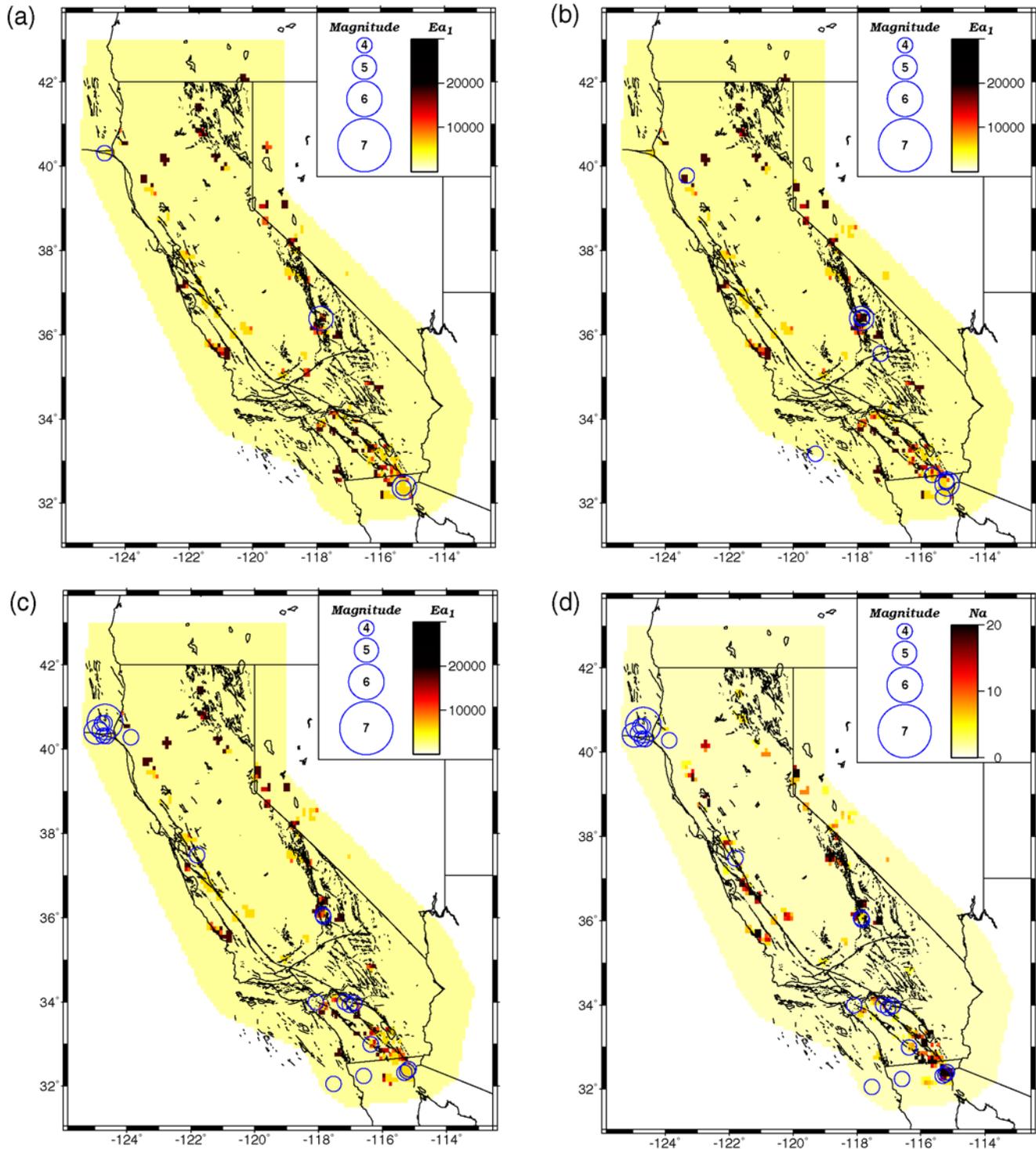


Figure 9. EAST forecast maps for the periods from (a) 1 July–30 September 2009, (b) 1 October–31 December 2009, and (c) 1 January–30 March 2010. For this last period, the E_{a1} value is calculated from the number of aftershocks shown in (d). Circles show epicenters of $M \geq M_{\text{target}}$ earthquakes that occurred during the corresponding forecasts. The color version of this figure is available only in the electronic edition.

where $\lambda_g(x, y)$ is the smoothed frequency of $M \geq 4$ earthquakes since 1960, calculated using a two-dimensional Gauss filter with standard deviation of 10 km. All the other model parameters have the same values as in the third column of Table 1.

During the first nine months of the test, 44 target earthquakes have occurred in the region under consideration, 21 of them in cells $c(x, y, t)$ with relatively high E_{a1} value (Fig. 9). Figure 10a shows that the EAST model has better predictive power than the RI reference model at a significance level

of 1%. For example, approximately 40% of the target earthquakes are predicted with a τ_{RI} value that is less than 0.05. In addition, the current performance of the EAST model is better than that obtained retrospectively in the whole CSEP testing region. Not surprisingly, the predictive power of the EAST model is even higher within the reduced area of the retrospective test (Fig. 10b). These stimulating results are mostly due to the successful forecast of a swarm of 11 $M \geq M_{\text{target}}$ earthquakes in central California (Fig. 9b). However, 10 other events occurred in cells with high E_{a1} values in northern California and near the United States–Mexico border (Fig. 9b,c). Even if we exclude the 11 events of the swarm from the comparison, the EAST model still has better predictive power than the RI reference model. Accordingly, we suggest that our successful prospective results are not due only to the space–time clustering of aftershocks.

To test the influence of earthquake clustering, we take advantage of the CSEP experiment by comparing the prediction of the EAST model with the prediction of four time-dependent rate-based clustering models. Two are three-months models proposed by Rhoades (2007): Every Earthquake a Precursor According to Scale (EEPAS; 5 versions) and Parkfield Prediction Experiment (PPE; 2 versions). Two others are one-day models: epidemic-type aftershock sequence (ETAS; Ogata, 1998, prepared for the test in California by Zhuang and Liukis) and short-term earthquake probability (STEP; Gerstenberger *et al.*, 2005). The three-months models are tested in California for $M_{\text{target}} = 4.95$. Only six $M > 4.95$

earthquakes have occurred during the first nine months of the test. This number is obviously too small to obtain statistically significant results. Then, we have extrapolated the EEPAS and PPE forecasts to $M_{\text{target}} = 3.95$ and performed the joint analysis for both M_{target} values. The Molchan diagrams for all these tests are available as an electronic supplement to this paper (Figs. S2–S7). In all cases, the EAST model outperformed the time-dependent reference models. For $M_{\text{target}} = 3.95$, this result is obtained with a 1% significance level. Furthermore, the three-month EAST model shows much better performance than one-day ETAS and STEP models. This may be partially the result of a lower spatial resolution in those models. Nevertheless, the final evaluation of the EAST model, as well as a reliable estimation of its relative accuracy compared with time-independent and time-dependent models, obviously requires a longer-term analysis of the prospective test.

Discussion and Perspectives

The working hypothesis of the EAST forecast model is that the time delay before the onset of the aftershock decay rate is anticorrelated with the level of stress in the seismogenic crust. To characterize this time delay, we work at small spatial (< 10 km) and temporal (< 12 hr) scales to concentrate on the effect of coseismic stress perturbation and to reduce the influence of other factors that may control the distribution of aftershocks (e.g., afterslip [Peng and Zhao, 2009] or fluids [Mori *et al.*, 2008]). In addition, we introduce a new quantity $\langle t_g \rangle$, the geometric mean of the elapsed times

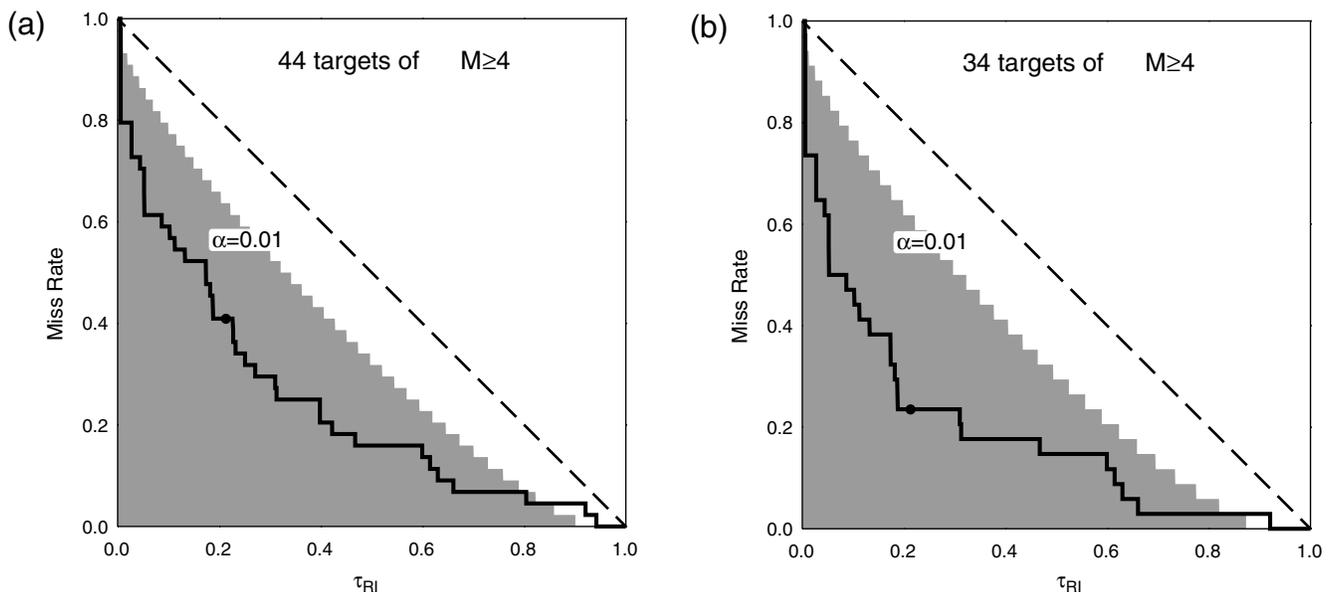


Figure 10. Evaluation of the EAST model in California from 1 July to 31 December 2009 for $M_{\text{target}} = 3.95$ in (a) the CSEP testing region and (b) the region used in the retrospective test (Fig. 2). Using a Molchan diagram, we compare the prediction of the EAST model to the prediction of the RI reference model. The solid line is the Molchan trajectory calculated from the highest to the lowest threshold value E_a^0 of the alarm function. The dotted line is the Molchan trajectory that incorporate zones where the E_a value cannot be defined (see text). The dashed diagonal line corresponds to an unskilled forecast model with respect to the reference model. The shaded area indicates the zone of the Molchan diagram in which the prediction of the EAST model is better than the prediction of the RI reference model at a level of significance $\alpha = 1\%$

from mainshocks to aftershocks. We could have continued on to estimate the c value of the Omori–Utsu law (Utsu, 1961) or the λ_b value of the limited power-law model (Narteau *et al.*, 2002). However, any model explaining the onset of the power-law aftershock decay rate is an idealization of a real behavior that is more complex. For this reason, we consider a nonparametric value such as the geometric mean $\langle t_g \rangle$ as a more meaningful quantity than the c value and the λ_b value to assess the time delay before the onset of the aftershock decay rate.

In both, retrospective and prospective tests, the EAST model shows better performance than the RI reference model. Nevertheless, we complement our analysis by using this reference model in space–time regions in which we cannot identify at least N_{\min} aftershocks. The results of the alarm-based model therefore may be biased by aftershock productivity, and we have to test that the E_a value is the dominant contribution to the model. With this objective in mind, we repeat the retrospective analysis using only the space–time regions $c(x, y, t)$ in which the E_a value is defined. Accordingly, we only consider $M > M_{\text{target}}$ earthquakes that occurred in these regions from t to $t + \delta t$. Thus, we keep 91 events out of 124 for $M_{\text{target}} = 5$, 31 out of 42 for $M_{\text{target}} = 5.5$, and 10 out of 13 for $M_{\text{target}} = 6$. Figure 11 also shows that in these cases the predictive power of the EAST model is better than the predictive power of the RI reference model at a 1% significance level. Consequently, there is no clear influence of the fragmented distribution of aftershocks on the performance of the forecast of the EAST model.

From the Molchan diagrams, we can systematically estimate the probability gain γ (Aki, 1981):

$$\gamma = \frac{1 - \nu}{\tau_{\text{RI}}}. \quad (13)$$

In a vast majority of cases, the comparisons between the EAST model and the RI reference model show γ values close to 10. In other forecast models, similar probability gain

values are mostly due to short-term forecasts of triggered events (Helmstetter *et al.*, 2006, Kagan and Jackson, 2000). Considering only mainshocks, the prediction power of these models is significantly lower (Kossobokov, 2006; Schorlemmer *et al.*, 2010). This may be also the case for the EAST model, especially for small M_{target} values for which an increasing number of target earthquakes may be aftershocks. For this reason, we now estimate the predictive power of the EAST model when all $M > M_{\text{target}}$ aftershocks are removed. To identify mainshocks and aftershocks, we use the declustering algorithm of Gardner and Knopoff (1974) and repeat the retrospective evaluation of the EAST model using only $M > M_{\text{target}}$ mainshocks as the target. Using E_{a1} values, Figure 12 shows the corresponding probability gain curves for three M_{target} values, $M_{\text{target}} \in \{4, 4.5, 5\}$, and a comparison with the probability gain curves obtained with $M > M_{\text{target}}$ aftershocks. In all cases, the EAST model has better predictive power than the RI reference model. For $M_{\text{target}} < 5$, the EAST model has a higher predictive power with $M > M_{\text{target}}$ aftershocks for a small threshold value of the alarm function (high rate of hits). For high threshold values of the alarm function (low rate of hits) the EAST model works better if the $M > M_{\text{target}}$ aftershocks are removed. For $M_{\text{target}} = 5$, the performance of the EAST model is always better without than with aftershocks. Combined with the fact that the number of $M > M_{\text{target}}$ aftershock decrease with an increasing M_{target} value, this result may also explain why the EAST model works better for higher M_{target} values (see Fig. 3).

This paper has two goals. First, it proposes a new alarm-based model that may improve seismic hazard assessment. Second, it tests whether the time delay before the onset of the power-law decay rate can be used to characterize the level of stress in the seismogenic crust. We think that the results presented here extend the set of evidence that allows us to answer positively to this question. In Figure 12, the

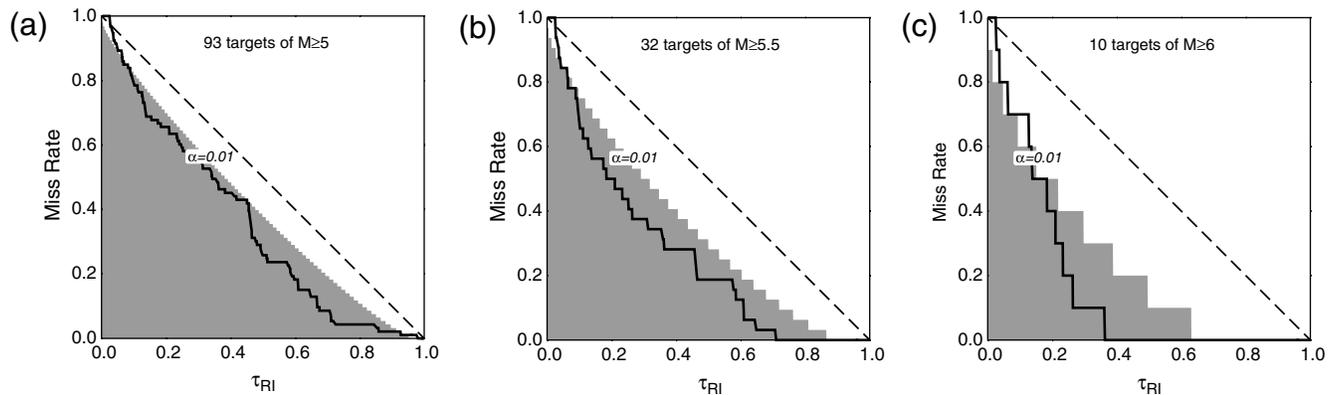


Figure 11. Retrospective evaluation of the EAST model in the space–time regions of California in which the E_a value can be defined. This evaluation covers a period from 1984 to 2008 for three M_{target} values: (a) $M_{\text{target}} = 5$, (b) $M_{\text{target}} = 5.5$, and (c) $M_{\text{target}} = 6$. Using a Molchan diagram, we compare the prediction of the EAST model to the prediction of the RI reference model. The solid line is the Molchan trajectory calculated from the highest to the lowest threshold value E_a^0 of the alarm function. The dashed diagonal line corresponds to an unskilled forecast model with respect to the reference model. The thin line limits the zone of the Molchan diagram in which the prediction of the EAST model is better than the prediction of the RI reference model at a level of significance $\alpha = 1\%$.

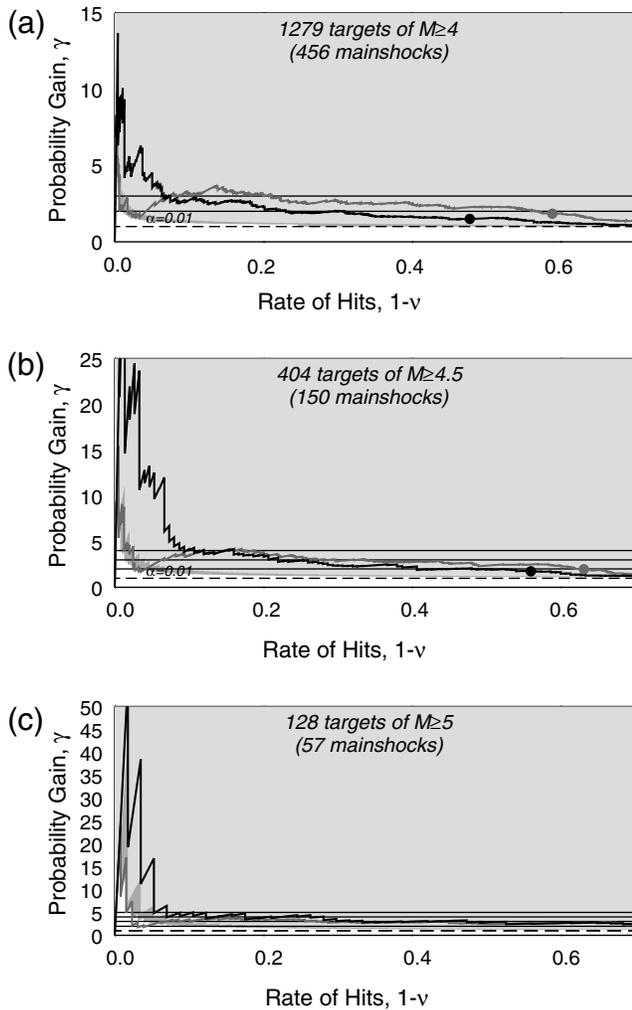


Figure 12. Probability gain plots for the retrospective evaluation of the EAST model: (black line) with and (gray line) without $M > M_{\text{target}}$ aftershocks. This evaluation covers a period from 1984 to 2008 for three M_{target} values: (a) $M_{\text{target}} = 4$, (b) $M_{\text{target}} = 4.5$, and (c) $M_{\text{target}} = 5$. The γ values are calculated from the RI reference model. Circles indicate the limit at which we start to incorporate space–time regions in which the E_a value is not defined. As in a Molchan diagram, the shaded areas show the $\alpha = 1\%$ quantiles of the miss-rate distribution.

difference between the probability gain curves with and without $M > M_{\text{target}}$ aftershocks also support this hypothesis. In fact, a vast majority of $M > M_{\text{target}}$ aftershocks are likely to result from a perturbation of the state of stress in zones where the level of stress may be significantly lower before the mainshock. At the three-month time step of the EAST model, it is impossible to capture such abrupt variations, and the model performs better when we remove aftershocks. A possible direction for further investigations is to refine the space–time mesh of the EAST model to forecast large aftershocks with the same accuracy.

Another perspective is to develop a frequency-based forecast model using retrospective results of the EAST model. In fact, for different ranges of E_a value, the long-term averaged expected frequency of events may be calculated from

the rate of hits per space–time cell $c(x, y, t)$. In this case, the probability of having an $M > M_{\text{target}}$ earthquake during the next time step is likely to decrease extremely rapidly with respect to the M_{target} value. As a result, for practical applications, it may be more useful to detect peak areas and peak periods in the probability map. Additional development of the EAST model in combination with other methods might further resolve this problem, too (Schorlemmer and Wiemer, 2005).

A definitive evaluation of the EAST model in the CSEP testing center requires long-term analysis. Using the Molchan diagrams of the retrospective evaluation, we can try to estimate this time for different M_{target} values. In practice, we take the Molchan trajectories obtained in California from 1984 to 2008 for $M_{\text{target}} = \{4, 5, 6\}$ (two of them are plotted in Fig. 3). Then, we recalculate the $\alpha = 1\%$ quantiles of the miss-rate distribution for an increasing number of events. As this number increases, the standard deviation of this binomial distribution is decreasing, and the zone in which the null-hypothesis cannot be rejected narrows and focuses around the diagonal $1 - \tau_{\text{RI}}$. When the Molchan trajectory is entirely below this critical level α , we note the corresponding number of events. We obtain values of 65 for $M_{\text{target}} = 4$, 40 for $M_{\text{target}} = 5$, and 8 for $M_{\text{target}} = 6$. Using the corresponding average seismic rate, we conclude that the EAST model need to be tested for 1 year for $M_{\text{target}} = 4$, 5 years for $M_{\text{target}} = 5$, and 10 years for $M_{\text{target}} = 6$. Taking into account that, in a forward test, results may be slightly worse, the duration of the test may become a little longer.

Conclusions

In a new alarm-based model, called the EAST model, the time delay before the onset of the aftershock decay rate is used to identify space–time regions with a higher level of stress and, consequently, a higher seismogenic potential. Retrospective analysis and the current evaluation in the CSEP testing center show some promising results that indicate the feasibility of the approach and that support our working hypothesis. Hence, we conclude that early aftershock decay rate may be a powerful way to estimate the evolution of the level of stress within the seismogenic crust and, more generally, a diagnostic tool for earthquake activity at both regional and local scales.

Data and Resources

The prospective test of the EAST model is carried out in the framework of the Collaboratory for the Study of Earthquake Predictability (CSEP; <http://www.cseptest.org/>). The Advanced National Seismic System (ANSS) earthquake catalog was searched using <http://quake.geo.berkeley.edu/cnss/catalog-search.html>. Most of the plots were made using the Generic Mapping Tools version 4.2.1 (www.soest.hawaii.edu/gmt/; Wessel and Smith, 1998).

Acknowledgments

This paper has been improved by the constructive comments of J. Zechar and two anonymous reviewers. The authors are also grateful to M. Liukis for invaluable help in installing the EAST model in CSEP and the evaluation of the prospective forecasts. This work has been partially supported by the French Ministry of Research (ANR-09-RISK-02-001/CASAVA).

References

- Aki, K. (1981). A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction: An International Review*, Maurice Ewing Series 4, D. Simpson and P. Richards (Editors), American Geophysical Union, Washington, D.C., 566–574.
- Atkinson, B. K. (Editor) (1987). *Fracture Mechanics of Rocks*, Academic Press, New York, 534 pp.
- Ben-Zion, Y., and V. Lyakhovskiy (2006). Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology, *Geophys. J. Int.* **165**, 197–210.
- Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.* **99**, 2601–2618.
- Enescu, B., J. Mori, M. Miyazawa, and Y. Kano (2009). Omori-Utsu law c -values associated with recent moderate earthquakes in Japan, *Bull. Seismol. Soc. Am.* **99**, 884–891.
- Gardner, J., and L. Knopoff (1974). Is the sequence of earthquakes in southern California with aftershocks removed Poissonian?, *Bull. Seismol. Soc. Am.* **5**, 1363–1367.
- Gerstenberger, M. C., S. Wiemer, L. M. Jones, and P. A. Reasenberg (2005). Real-time forecasts of tomorrow's earthquakes in California, *Nature* **435**, 328–331.
- 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**, 90–106.
- Jordan, T. H. (2006). Earthquake predictability, brick by brick, *Seismol. Res. Lett.* **77**, 3–6.
- Kagan, Y. Y. (2004). Short-term properties of earthquake catalogs and models of earthquake source, *Bull. Seismol. Soc. Am.* **94**, 1207–1228.
- Kagan, Y. Y., and D. D. Jackson (2000). Probabilistic forecasting of earthquakes, *Geophys. J. Int.* **143**, 438–454.
- Kossobokov, V. (2006). Testing earthquake prediction methods: The west Pacific short-term forecast of earthquakes with magnitude $m_w \geq 5.8$, *Tectonophysics* **413**, 25–31.
- Kossobokov, V., and P. Shebalin (2003). Earthquake prediction, chap. 4 in *Nonlinear Dynamics of the Lithosphere and Earthquake Prediction*, V. I. Keilis-Borok and A. A. Soloviev (Editors), 141–205, Springer-Verlag, New York, 141–207.
- Lolli, B., and P. Gasperini (2006). Comparing different models of aftershock rate decay: The role of catalog incompleteness in the first times after main shock, *Tectonophysics* **423**, 43–59.
- Molchan, G. (1990). Strategies in strong earthquake prediction, *Phys. Earth Planet. In.* **61**, 84–98.
- Molchan, G. (1991). Structure of optimal strategies in earthquake prediction, *Tectonophysics* **193**, 267–276.
- Molchan, G., and V. Keilis-Borok (2008). Earthquake prediction: Probabilistic aspect, *Geophys. J. Int.* **173**, 1012–1017.
- Mori, J., Y. Kano, and B. Enescu (2008). Comparison of early aftershock sequences for the 2004 Mid-Niigata and 2007 Noto Hanto earthquakes in central Japan, *Earth Planets Space* **60**, 151–154.
- Nanjo, K. Z., B. Enescu, R. Shcherbakov, D. L. Turcotte, T. Iwata, and Y. Ogata (2007). Decay of aftershock activity for Japanese earthquakes, *J. Geophys. Res.* **112**, no. B08309, 12 pp., doi [10.1029/2006JB004754](https://doi.org/10.1029/2006JB004754).
- Narteau, C., P. Shebalin, and M. Holschneider (2002). Temporal limits of the power law aftershock decay rate, *J. Geophys. Res.* **107**, no. B122359, 14 pp., doi [10.1029/2002JB001868](https://doi.org/10.1029/2002JB001868).
- Narteau, C., P. Shebalin, and M. Holschneider (2005). Onset of power law aftershock decay rates in southern California, *Geophys. Res. Lett.* **32**, no. B03306, 5 pp., doi [10.1029/2005GL023951](https://doi.org/10.1029/2005GL023951).
- Narteau, C., P. Shebalin, and M. Holschneider (2008). Loading rates in California inferred from aftershocks, *Nonlin. Proc. Geophys.* **15**, 245–263.
- Narteau, C., S. Byrdina, P. Shebalin, and D. Schorlemmer (2009). Common dependence on stress for the two fundamental laws of statistical seismology, *Nature* **462**, 642–645.
- Ogata, Y. (1998). Space-time point-process models for earthquake occurrences, *Ann. Inst. Statist. Math.* **50**, 379–402.
- Peng, Z., and P. Zhao (2009). Migration of early aftershocks following the 2004 Parkfield earthquake, *Nature Geosci.* **28**, 877–881.
- Peng, Z. G., J. E. Vidale, and H. Houston (2006). Anomalous early aftershock decay rate of the 2004 M_w 6.0 Parkfield, California, earthquake, *Geophys. Res. Lett.* **33**, L17307, doi [10.1029/2006GL026744](https://doi.org/10.1029/2006GL026744).
- Peng, Z. G., J. E. Vidale, M. Ishii, and A. Helmstetter (2007). Seismicity rate immediately before and after main shock rupture from high frequency waveforms in Japan, *J. Geophys. Res.* **112**, no. B03306, doi [10.1029/2006JB004386](https://doi.org/10.1029/2006JB004386).
- Rhoades, D. A. (2007). Application of the EEPAS model to forecasting earthquakes of moderate magnitude in southern California, *Seismol. Res. Lett.* **78**, 110–115.
- Scholz, C. (1968). Microfractures, aftershocks and seismicity, *Bull. Seismol. Soc. Am.* **58**, 1117–1130.
- Schorlemmer, D., and M. Gerstenberger (2007). RELM Testing Center, *Seismol. Res. Letts.* **78**, 30–36.
- Schorlemmer, D., and S. Wiemer (2005). Microseismicity data forecast rupture area, *Nature* **434**, 1086–1086.
- Schorlemmer, D., S. Wiemer, and M. Wyss (2005). Variations in earthquake-size distribution across different stress regimes, *Nature* **437**, 539–542.
- Schorlemmer, D., J. D. Zechar, M. Werner, D. D. Jackson, E. H. Field, T. H. Jordan (, and The RELM Working Group). First results of the regional earthquake likelihood models experiment, *Pure Appl. Geophys.* **167**, no. 8–9, 859–876.
- Shcherbakov, R., and D. L. Turcotte (2004). A damage mechanics model for aftershocks, *Pure Appl. Geophys.* **161**, 2379–2391.
- Shcherbakov, R., D. L. Turcotte, and J. B. Rundle (2004). A generalized Omori's law for earthquake aftershock decay, *Geophys. Res. Lett.* **31**, L11613, 5 pp., doi [10.1029/2004GL019808](https://doi.org/10.1029/2004GL019808).
- Shebalin, P. (2004). Aftershocks as indicators of the state of stress in a fault system, *Dokl. Earth Sci.* **398**, 978–982.
- Shebalin, P., V. Kellis-Borok, A. Gabriellov, I. Zaliapin, and D. Turcotte (2006). Short-term earthquake prediction by reverse analysis of lithosphere dynamics, *Tectonophysics* **413**, 63–75.
- Sibson, R. H. (1974). Frictional constraints on thrusts, wrench and normal faults, *Nature* **249**, 542–544.
- Utsu, T. (1961). A statistical study on the occurrence of aftershocks, *Geophys. Mag.* **30**, 521–605.
- Vidale, J. E., Z. Peng, and M. Ishii (2004). Anomalous aftershock decay rates in the first hundred seconds revealed from the Hi-net borehole data, *Eos Trans. AGU* **85**, no. 47, Fall Meet. Suppl., Abstract S23C-07.
- Wessel, P., and W. Smith (1998). New, improved version of Generic Mapping Tools released, *Eos Trans. AGU* **79**, no. 47, 579.
- Zechar, J. D., and T. H. Jordan (2008). Testing alarm-based earthquake predictions, *Geophys. J. Int.* **172**, 715–724.
- Zechar, J., and T. Jordan (2010). The area skill score statistic for evaluating earthquake predictability experiments, *Pure Appl. Geophys.* **167**, doi [10.1007/s00024-010-0086-0](https://doi.org/10.1007/s00024-010-0086-0).
- Zechar, J. D., T. H. Jordan, D. Schorlemmer, and M. Liukis (2007). Comparison of two earthquake predictability evaluation approaches, Molchan error trajectory and likelihood, *Seismol. Res. Lett.* **78**, 250.
- Zechar, J. D., D. Schorlemmer, M. Liukis, J. Yu, F. Euchner, P. J. Maechling, and T. H. Jordan (2010). The Collaboratory for the Study of Earth-

quake Predictability perspective on computational earthquake science,
Concurrency Comput. Pract. Ex. **22**, 1836–1847. doi [10.1002/
cpe.1519](https://doi.org/10.1002/cpe.1519).

International Institute of Earthquake Prediction Theory and Mathematical
Geophysics
Moscow, 84/32 Profsoznaya
Moscow 117997, Russia
(P.S.)

Laboratoire de Dynamique des Fluides Geologiques
Institut de Physique du Globe de Paris
75252 Paris Cedex 05, France
(C.N.)

Institutes of Applied and Industrial Mathematics
Universität Potsdam
POB 601553, 14115
Potsdam, Germany
(M.H.)

Department of Earth Sciences
University of Southern California
3651 Trousdale Parkway
Los Angeles, California 90089
(D.S.)

Manuscript received 1 May 2010