Detecting aseismic strain transients from seismicity data

A. L. Llenos[1,2,3] and J. J. McGuire[4]

Received 8 March 2010; revised 25 February 2011; accepted 22 March 2011; published 17 June 2011.

[1] Aseismic deformation transients such as fluid flow, magma migration, and slow slip can trigger changes in seismicity rate. We present a method that can detect these seismicity rate variations and utilize these anomalies to constrain the underlying variations in stressing rate. Because ordinary aftershock sequences often obscure changes in the background seismicity caused by aseismic processes, we combine the stochastic Epidemic Type Aftershock Sequence model that describes aftershock sequences well and the physically based rate- and state-dependent friction seismicity model into a single seismicity rate model that models both aftershock activity and changes in background seismicity rate. We implement this model into a data assimilation algorithm that inverts seismicity catalogs to estimate space-time variations in stressing rate. We evaluate the method using a synthetic catalog, and then apply it to a catalog of $M \geq 1.5$ events that occurred in the Salton Trough from 1990 to 2009. We validate our stressing rate estimates by comparing them to estimates from a geodetically derived slip model for a large creep event on the Obsidian Buttes fault. The results demonstrate that our approach can identify large aseismic deformation transients in a multidecade long earthquake catalog and roughly constrain the absolute magnitude of the stressing rate transients. Our method can therefore provide a way to detect aseismic transients in regions where geodetic resolution in space or time is poor.

Citation: Llenos, A. L., and J. J. McGuire (2011), Detecting aseismic strain transients from seismicity data, *J. Geophys. Res.*, *116*, B06305, doi:10.1029/2010JB007537.

1. Introduction

[2] Transient aseismic processes such as fluid flow, magmatic intrusions, or slow slip events can alter the stress state and trigger seismicity in a variety of tectonic environments. For example, earthquake swarms associated with aseismic deformation have been observed in subduction zones such as Japan and New Zealand [*Ozawa et al.*, 2007; *Delahaye et al.*, 2009], continental strike-slip faults [*Lohman and McGuire*, 2007], and volcanic regions [*Dieterich et al.*, 2000; *Toda et al.*, 2002; *Segall et al.*, 2006; *Wolfe et al.*, 2007; *Montgomery-Brown et al.*, 2009]. However, there is not a straightforward relationship between the magnitude of the aseismic transient and the amount of seismicity triggered. In the Salton Trough of California, a M_w 5.7 creep event triggered a swarm of ~1000 $1 \leq M \leq 5.1$ earthquakes [*Lohman and McGuire*, 2007], whereas offshore of central Honshu, Japan, recurrent $M_w \sim 6.5$ slow earthquakes on the plate interface typically trigger swarms of ~10s of $2 \leq M \leq 5$ earthquakes [*Ozawa et al.*, 2007]. Moreover, a M_w 4.7 creep event detected on the Superstition Hills fault in southern California did not trigger any earthquakes [*Wei et al.*, 2009]. There is evidence, however, that the seismicity rate varies in proportion to the stressing rate increase caused by the aseismic transient [*Toda et al.*, 2002; *Segall et al.*, 2006; *Lohman and McGuire*, 2007], although this is likely complicated by a dependence on ambient stress conditions related to the timing within a fault's seismic cycle. Thus, seismicity rate variations observed in earthquake catalogs can potentially indicate when and where aseismic processes are occurring on some faults, which would ordinarily require high-quality geodetic data to resolve. Moreover, the variations in earthquake triggering observed during aseismic deformation transients suggest an opportunity to learn about the physics of earthquake triggering. They may also help constrain the extent to which potentially observable geophysical processes, such as slow slip, trigger seismicity that could then be forecastable in a probabilistic sense.

[3] The rate- and state-dependent friction model [*Dieterich*, 1994; *Dieterich et al.*, 2000] provides one approach for mapping seismicity rate variations to underlying stressing rate variations. This model has been successfully used to estimate stress changes caused by a dike intrusion [*Dieterich et al.*, 2000] and slow slip events [*Segall et al.*, 2006; *Montgomery-Brown et al.*, 2009] on Kilauea volcano as well as a large intrusion in the Izu islands [*Toda et al.*, 2002]. However, aftershock sequences in a catalog can obscure the background seismicity

[1]Massachusetts Institute of Technology/Woods Hole Oceanographic Institution Joint Program, Cambridge, Massachusetts, USA.
[2]Department of Geophysics, Stanford University, Stanford, California, USA.
[3]Now at U.S. Geological Survey, Menlo Park, California, USA.
[4]Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA.

0148-0227/11/2010JB007537

rate changes due to aseismic processes, particularly in regions such as southern California that are characterized by high aftershock productivity [*Helmstetter*, 2003; *Helmstetter and Sornette*, 2003a]. In such cases, the stressing rate changes estimated by the rate-state inversion will be a combination of those due to the underlying aseismic process and those due to the static stress changes resulting from the seismicity itself. Therefore, the rate-state inversion should be used with caution.

[4] The stochastic Epidemic Type Aftershock Sequence (ETAS) model [*Ogata*, 1988] provides an alternative approach to estimate the time dependence of underlying driving mechanisms. This model represents earthquake occurrence as a point process, characterized by a background rate and a set of parameters associated with Omori's law that can be optimized to fit a particular catalog. The ETAS model effectively detects when an external process is triggering anomalous seismicity [*Hainzl and Ogata*, 2005; *Ogata*, 2004, 2005; *McGuire et al.*, 2005]. However, ETAS lacks a procedure for estimating smooth variations in the background earthquake rate and a quantitative way to relate these directly to stressing rate variations.

[5] The rate-state model and the ETAS model have complementary strengths which, when combined, can provide an effective tool for detecting seismicity rate anomalies and relating them to underlying stress fields. In a previous study, we combined these two models into a single seismicity rate model that explains both aftershock activity as well as variations in background seismicity occurring in a region [*Llenos et al.*, 2009]. In this model, the observed seismicity rate R in a catalog is approximately a linear combination of an aseismically triggered component (reflecting seismicity triggered by long-term tectonic loading and aseismic processes) and an earthquake-earthquake triggered component (reflecting aftershock sequences). Applying the rate-state model to both aseismic and coseismic stress changes can be complicated, due to the different spatial scales involved. Coseismic stress changes typically occur at scales of one to two rupture lengths [*Rubin and Gillard*, 2000], which is on the order of a kilometer for a M4 earthquake. It would be extremely difficult to resolve variations at such scales given current earthquake detection thresholds. Moreover, the rate-state model has been shown to be highly sensitive to stress and slip heterogeneities at small spatial scales [*Helmstetter and Shaw*, 2006; *Marsan*, 2006; *Hainzl et al.*, 2009], and so using the rate-state model to accurately map the coseismic stress variations to seismicity rate variations would require very high spatial resolution.

[6] However, the rate-state model has had considerable success at identifying the sign and magnitude of the average stress change in a volume that results from nearby, large aseismic stressing transients. *Dieterich et al.* [2000] applied the rate-state model to a dike intrusion in Kilauea to relate seismicity rate variations observed in the catalog to the stress changes caused by the opening dike. The stress changes estimated from the seismicity using the rate-state model compared favorably with stress changes obtained from independent boundary element methods, demonstrating that the model can be successfully applied to detect aseismic stressing rate variations acting on large spatial scales.

[7] Therefore, rather than using the rate-state model for the coseismically triggered component of seismicity rate, we instead use the ETAS model to estimate it, because ETAS models the spatial fall-off of aftershocks with a continuous function of space. We then subtract the ETAS estimate of triggered seismicity from the observed (binned) catalog rate to estimate temporal changes in background rate, and then relate the residual seismicity rate directly to an aseismic stressing rate with the rate-state model. Thus, our approach essentially consists of combining the ETAS model to estimate coseismically triggered seismicity rate variations with the rate-state model to estimate aseismically triggered background rate variations.

[8] In this paper, we implement this combined model in a data assimilation algorithm that uses seismicity catalogs to estimate stressing rate variations. Data assimilation algorithms provide a way to combine a dynamic model with observed data in order to estimate the time dependence of the underlying (and often unobservable) state variables in the model. Therefore, we incorporate the rate-state equations into a state-space model and use an extended Kalman filter to estimate spatial and temporal variations in stressing rate from the observed seismicity rates in a catalog.

[9] We evaluate this method with a synthetic test, and apply it to a catalog of events from the Salton Trough in southern California in which a geodetically detected aseismic transient occurred. To validate our approach, we compare the estimated stressing rate peak during this aseismic slip event to that obtained from a slip inversion of geodetic data [*Lohman and McGuire*, 2007]. The results suggest that our seismicity inversion method provides an accurate way to detect and locate transient deformation strictly from earthquake catalogs and may provide a means to constrain the absolute magnitude of stressing rate variations. Our approach therefore may enable the detection of these transients in regions where geodetic resolution is poor, such as offshore in subduction zones, and in time periods prior to the availability of high-quality geodetic data.

2. Method

[10] The seismicity rate R observed in a catalog is in general a function of the stressing rate acting on the population of faults in a region [*Dieterich*, 1994]. We assume that the seismicity in a catalog is primarily triggered by three mechanisms: earthquake-earthquake interactions (i.e., as aftershocks), transient changes to the stress field from aseismic processes such as fluid flow or slow slip, and long-term tectonic loading. Therefore, $R(t,x) = f(\dot{S}) = f(R_A, R_C, R_T)$, where R_A reflects seismicity triggered by aseismic processes, R_C represents a coseismically triggered component (i.e., earthquakes triggered by other earthquakes), and R_T is triggered by long-term tectonic loading. We assume that R_T is small relative to R_A if an aseismic deformation/fluid-flow transient is occurring and therefore combine it with R_A, so that $R(t, x) \approx f(R_A, R_C)$.

[11] In previous studies, the ETAS model and the rate-state model have each been used to investigate both R_A and R_C. Recent studies however have found that aftershock rates predicted by the rate-state model may be complicated by factors such as the influence of small-scale stress heterogeneities and a dependence on stressing rate [e.g., *Helmstetter and Shaw*, 2006; *Marsan*, 2006; *Hainzl et al.*, 2009; *Cocco et al.*, 2010]. Therefore, we will model the coseismically triggered R_C using the ETAS model. Aseismic stressing rate variations of interest, however, typically occur on larger spatial and temporal scales, therefore we believe that the rate-state model may still be applied to relate the aseismically

Table 1. Mathematical Notation

Symbol	Description
	ETAS Model Notation
d, q, η	Aftershock spatial distribution parameters
K	Aftershock productivity
p, c	Omori decay parameters
α	Magnitude scaling parameter
μ	Background seismicity rate
M, M_c	Event magnitude, cutoff magnitude
	Rate-State Model Notation
A	Fault constitutive parameter
r	Reference seismicity rate
R, R_A, R_C, R_T	Observed, aseismically, coseismically and long-term tectonically triggered seismicity rates
σ	Effective normal stress
$\dot{S}_r$	Reference stressing rate
$\dot{S}, \dot{S}_A, \dot{S}_p$	Stressing rate, aseismic stressing rate, plate tectonic stressing rate
γ	Seismicity state variable
	Filter Notation
$\mathbf{d}$	Data (observation) vector
$\mathbf{g}$	Kalman gain
$\mathbf{h}$, $\mathbf{H}$	Measurement (model) matrix and Jacobian
L	Likelihood
N, N_d, N_e	Number of spatial cells, observations, epochs
$\mathbf{Q}$	Process covariance matrix
$\mathbf{R}$	Data covariance matrix
$\mathbf{t}$, $\mathbf{T}$	State transition model and Jacobian
$\mathbf{x}$	State variables
δ, ω	Measurement, process noise
τ, ε	Random walk scale parameters for $\dot{S}_A$, $\ln(\gamma)$
ξ	Data covariance scalar multiplier
$\mathbf{\Sigma}$	Error covariance matrix
$\boldsymbol{\nu}$	Innovation
$\mathbf{V}$	Variance of innovation

triggered R_A to an aseismic stressing rate. Table 1 summarizes the mathematical symbols used in the following sections.

2.1. ETAS Model

[12] The space-time ETAS model is a point process model that represents the seismicity rate $R(t,x,y)$ as a summation of the aftershock sequences produced by each event prior to time t plus a time-independent background seismicity rate μ [*Ogata*, 1998; *Zhuang et al.*, 2005; *Ogata and Zhuang*, 2006]:

$$R(t,x,y) = \mu(x,y) + \sum_{i:t_i<t} \kappa(M_i) \cdot \psi(t-t_i) \cdot f(x-x_i, y-y_i; M_i) \tag{1}$$

where

$$\kappa(M) = Ke^{\alpha(M-M_c)} \tag{2}$$

$$\psi(t) = \frac{p-1}{c}\left(1+\frac{t}{c}\right)^{-p} \tag{3}$$

$$f(x,y;M) = \frac{q-1}{\pi d e^{\eta(M-M_c)}}\left(1+\frac{x^2+y^2}{d e^{\eta(M-M_c)}}\right)^{-q} \tag{4}$$

The function $\kappa(M)$ is the expected number of events triggered by an earthquake of magnitude M and involves the aftershock productivity K and scaling parameter α. The function $\psi(t)$ is the probability density function form of the modified Omori law and is specified by the Omori decay parameters p and c [*Omori*, 1894; *Utsu*, 1961]. The function $f(x,y;M)$ describes the spatial distribution of the events triggered by an earthquake of magnitude M, specified by the parameters d, q, and scaling parameter η. The ETAS parameters μ, K, c, p, α, d, q, and η are generally obtained using maximum likelihood estimation from the observed occurrence times t_i and magnitudes M_i of earthquakes in a catalog with a magnitude of completeness M_c [*Ogata and Zhuang*, 2006].

2.2. Rate- and State-Dependent Friction Model

[13] In the rate- and state-dependent friction model, the seismicity rate R is a function of the stressing rate $\dot{S}$ acting on a population of faults governed by rate- and state-dependent friction [*Dieterich*, 1994]:

$$R = \frac{r}{\gamma \dot{S}_r} \tag{5}$$

$$d\gamma = \frac{dt}{A\sigma}\left(1-\gamma\dot{S}\right) \tag{6}$$

where r is a steady state reference seismicity rate associated with a reference stressing rate $\dot{S}_r$, $\dot{S}$ is the modified Coulomb stressing rate, γ is a state variable, and A is a fault constitutive parameter. We assume the normal stress σ and therefore the frictional parameter $A\sigma$ remain constant.

2.3. Combined ETAS/Rate-State Model

[14] *Llenos et al.* [2009] examined how the ETAS and rate-state model parameters changed during periods of high stressing rate. In the ETAS model, the increase in stressing rate due to aseismic transients primarily increases the background seismicity rate (i.e., ETAS parameter μ), while other aftershock parameters, in particular the aftershock productivity K, remain relatively unaffected. This contrasts with the rate-state model, where aftershock productivity, while theoretically independent of stressing rate, can appear to increase proportionally with it.

[15] Aftershock productivity is typically defined as the cumulative number of events following a main shock. In the rate-state model, this number can be obtained by integrating the seismicity rate predicted following a stress step after the steady state seismicity rate has been subtracted. When an increase in stressing rate occurs, the seismicity rate evolves toward a new steady state rate (equations (5)–(6)). Simulations indicate that if this new steady state rate is known and used to calculate the productivity of a stress step that occurs during a period of high stressing rate, then the productivity does not change compared to the case where the same stress step occurs during the background stressing rate. This demonstrates that aftershock productivity is independent of stressing rate in the rate-state model at steady state.

[16] However, in real applications, a new steady state seismicity rate will most likely not be resolvable, if it is even achieved. Transients such as slow slip events and magma intrusions can last on the order of 1–2 weeks, during which

time the stressing rate is often varying continuously, resulting in a similarly varying seismicity rate. Moreover, the evolution of the state variable γ, which governs the relationship between stressing and seismicity rate and therefore the time it would take to achieve a new steady state (equation (6)), depends on the value of the frictional parameter $A\sigma$. For larger values of this parameter (~0.1 MPa), a new steady state seismicity rate may never be achieved during the transient. Therefore, the long-term background seismicity rate may be the only resolvable steady state rate. When this rate is used rather than the larger transient rate to estimate the productivity by subtracting the background events, this can cause an apparent increase in the productivity (i.e., cumulative number of events) that is actually due to the increase in the background seismicity rate during the transient. This is consistent with the observation that when the ETAS model is fit to swarms that occurred during transients, the background seismicity rate μ increased by orders of magnitude while the productivity and other aftershock parameters remained relatively unaffected [*Llenos et al.*, 2009].

[17] The total seismicity rate R can then be approximately modeled as a linear combination of the aseismically triggered component R_A and the coseismically triggered component R_C:

$$R(t) = R_A(t) + R_C(t) \approx R_A(t) + \sum_{t_i \le t} \frac{Ke^{\alpha(M_i - M_c)}}{(t - t_i + c)^p} \tag{7}$$

where R_C is represented by the ETAS model, and R_A is essentially a time-dependent version of μ that will reflect the variations in seismicity rate triggered by transients. We first approximate R_A by subtracting out an ETAS-predicted rate from the observed rate R in a catalog. We then directly relate our estimate of R_A to an aseismic stressing rate $\dot{S}_A$ through the rate-state model equations. We utilize the space-time version of the ETAS model [*Ogata and Zhuang*, 2006], so that the model becomes:

$$\begin{aligned} R_A(t,x,y) &= R(t,x,y) - R_C(t,x,y) = R(t,x,y) \\ &\quad - \sum_{i:t_i<t} Ke^{\alpha(M_i - M_c)} \cdot \frac{p-1}{c}\left(1 + \frac{t - t_i}{c}\right)^{-p} \\ &\quad \cdot \frac{q-1}{\pi d e^{\eta(M_i - M_c)}}\left(1 + \frac{(x-x_i)^2 + (y-y_i)^2}{d e^{\eta(M_i - M_c)}}\right)^{-q} = \frac{r}{\dot{S}_r \gamma} \end{aligned} \tag{8}$$

$$d\gamma = \frac{dt}{A\sigma}\left[1 - \gamma\left(\dot{S}_A + \dot{S}_p\right)\right] \tag{9}$$

where $\dot{S}_p$ is the background plate tectonic stressing rate. This combined model of seismicity rate now reduces the impact of aftershock sequences on the estimation of R_A by using ETAS to model R_C and provides a way to quantitatively relate seismicity rate variations to stressing rate variations.

2.4. Extended Kalman Filter Algorithm

[18] We incorporate the aseismically triggered rate R_A into a state-space model, which describes a system using a state vector that evolves over time and can be related to observable data. These systems can be solved with data assimilation algorithms, which are a way to combine a dynamic model (e.g., the rate-state model) with noisy observed data (e.g., seismicity rates) to produce temporal estimates of the underlying (and often unobservable) state variables that describe the system (e.g., stressing rate). Commonly used in fields such as hydrology, meteorology, and engineering, data assimilation also provides a way to combine maximum likelihood estimation of time-independent model parameters with the estimation of time-dependent variables.

[19] Our state-space model, detailed in Appendix A, consists of a nonlinear measurement equation that relates the observations $\mathbf{d}_k$ to the state vector $\mathbf{x}_k$ at each time step k:

$$\mathbf{d}_k = \mathbf{h}_k(\mathbf{x}_k) + \omega_k, \;\; \omega_k \sim N(0, \mathbf{R}_k) \tag{10}$$

and a nonlinear state transition equation that describes how the state vector evolves over time:

$$\mathbf{x}_{k+1} = \mathbf{t}(\mathbf{x}_k) + \delta_{k+1}, \;\; \delta_{k+1} \sim N(0, \mathbf{Q}_{k+1}) \tag{11}$$

The data vector $\mathbf{d}_k$ consists of the aseismically triggered rate R_A estimated in each spatial box and relates to the state variables through the measurement matrix $\mathbf{h}_k$, which incorporates the rate-state model equations.

[20] We assume the measurement and process noises ω_k and δ_{k+1} are Gaussian and described by the data and process covariance matrices $\mathbf{R}_k$ and $\mathbf{Q}_{k+1}$ respectively. Measurement noise accounts for uncertainties in the seismicity rate observations, and incorporates the uncertainties in the ETAS parameters used to estimate R_A. Process noise accounts for uncertainties in the dynamic system (i.e., the rate-state model) used to model the state variables over time. Our state variables $\mathbf{x}_k$ consist of the long-term tectonic loading rate $\dot{S}_p$, the aseismic stressing rate $\dot{S}_A$, and the logarithm of the rate-state seismicity state variable γ to ensure positivity. The variables $\dot{S}_A$ and $\ln(\gamma)$ are modeled as random walk processes with scale parameters of τ and ε respectively to allow for arbitrary stressing rate histories. We divide the region into N spatial boxes and estimate the state variables in each box.

[21] The system is solved using an extended Kalman filter algorithm [e.g., *Gelb*, 1974; *Anderson and Moore*, 2005]. Kalman filters estimate the state of a system over time, assuming the system can be described by linear models and Gaussian noise. This makes it relatively straightforward to predict the mean and covariance of the state variables at the next time step, then update or correct that prediction with the measurement observed at that time step. Kalman filters are widely applied in many branches of engineering and science, including geodesy [e.g., *Segall and Matthews*, 1997; *McGuire and Segall*, 2003]. Because of the nonlinearity in both our model and state transition matrices (see Appendix A), we utilize an extended Kalman filter, which linearizes the model at each time step.

[22] We first make a priori estimates of the state vector $\mathbf{x}_{0|0}$ and covariance $\mathbf{\Sigma}_{0|0}$ and update them with the first observations to obtain $\mathbf{x}_{1|1}$ and $\mathbf{\Sigma}_{1|1}$. The subscripts $\mathbf{x}_{i|j}$ indicate the estimate of $\mathbf{x}$ at time step i given data through time step j.

We then use the filter prediction equations to estimate the state vector and covariance at the next time step $k + 1$:

$$\mathbf{x}_{k+1|k} = \mathbf{t}\left(\mathbf{x}_{k|k}\right) + \delta_{k+1}, \quad \delta_{k+1} \sim N(0, \mathbf{Q}_{k+1}) \tag{12}$$

$$\mathbf{\Sigma}_{k+1|k} = \mathbf{T}_k \mathbf{\Sigma}_{k|k} \mathbf{T}_k^T + \mathbf{Q}_{k+1} \tag{13}$$

Here $\mathbf{T}$ is the Jacobian of the nonlinear state transition model $\mathbf{t}$, which incorporates the rate-state evolution equation (equation (6)) to describe how the state variables evolve from one time step to the next (see Appendix A).

[23] The predictions are then compared with the observations at the current time step and updated using the filter update equations:

$$\mathbf{x}_{k+1|k+1} = \mathbf{x}_{k+1|k} + \mathbf{g}_{k+1}\left[\mathbf{d}_{k+1} - \mathbf{h}\left(\mathbf{x}_{k+1|k}\right)\right] \tag{14}$$

$$\mathbf{\Sigma}_{k+1|k+1} = \left[\mathbf{I} - \mathbf{g}_{k+1}\mathbf{H}_{k+1}\right]\mathbf{\Sigma}_{k+1|k} \tag{15}$$

where the Kalman gain $\mathbf{g}$ is computed by:

$$\mathbf{g}_{k+1} = \mathbf{\Sigma}_{k+1|k}\mathbf{H}_{k+1}^T\left[\mathbf{H}_{k+1}\mathbf{\Sigma}_{k+1|k}\mathbf{H}_{k+1}^T + \mathbf{R}_{k+1}\right]^{-1} \tag{16}$$

and $\mathbf{H}$ is the Jacobian of the nonlinear measurement matrix $\mathbf{h}$, which incorporates the rate-state model (equation (5)) to relate the observed seismicity rates to the underlying state variables (see Appendix A).

[24] These prediction and update steps iterate forward through the entire data set from $k = 1$ to N_e, where N_e is the total number of time steps, resulting in estimates of the state variables $\mathbf{x}_{k|k}$ at each time step k given the data up to and including k. We can run the filter backward in time in a process known as smoothing to obtain a back-smoothed estimate of the state variables (i.e., $\mathbf{x}_{k|Ne}$, or estimates of $\mathbf{x}$ at time step k given the entire data set).

2.5. Likelihood Calculations

[25] There are a total of 11 time-independent parameters in our model, seven of which are the space-time ETAS parameters (K, c, p, α, d, q, and η), two of which are rate-state parameters ($A\sigma$ and r), and two of which are filter hyperparameters (τ and ε). We assume that all of the time-independent parameters are also constant in space over the entire study area. We first optimize the model over the set of ETAS parameters given a history of occurrence H_t using a point process likelihood function [*Ogata*, 1998; *Daley and Vere-Jones*, 2002]:

$$\ln L(K, c, p, \alpha, d, q, \eta | H_t) = \sum_{i=1}^{n} \ln R(t_i, x_i, y_i | H_t) - \int_0^T \iiint_A R(t, x, y | H_t) dA dt \tag{17}$$

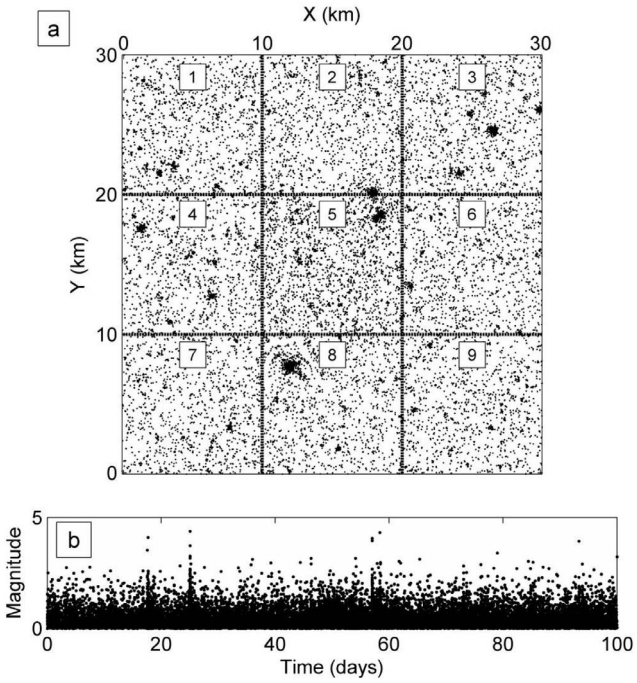

Figure 1. (a) Synthetic catalog locations and boxes used in analysis. Parameters used to generate the catalog are K = 4.5e-4 events/day/deg^2, α = 0.8, p = 1.2, c = 0.0001 days, d = 0.001 deg^2, q = 3, η = 0.3, r = 10 events/day, and $A\sigma$ = 0.001 MPa. (b) Synthetic catalog earthquake magnitudes and occurrence times. Aftershock sequences dominate both the map and the magnitude-time history, obscuring the anomalous seismicity triggered by the transient.

While this likelihood computation works well to estimate the ETAS parameters, it is relatively insensitive to changes in the non-ETAS parameters. Based on synthetic tests, it is preferable to optimize the set of non-ETAS parameters by computing a likelihood based on prediction error decomposition [*Harvey*, 1989; *Segall and Matthews*, 1997]. For Gaussian data, the likelihood can be expressed as:

$$\log L(r, A\sigma, \tau, \varepsilon) = -\frac{1}{2}(N_d - N_d \log N_d) - \frac{1}{2}\sum_{k=1}^{N_e} \log|\mathbf{V}_k| - \frac{1}{2} N_d \log\left[\sum_{k=1}^{N_e} \boldsymbol{\nu}_k^T \mathbf{V}_k^{-1} \boldsymbol{\nu}_k\right] \tag{18}$$

where N_d is the total number of observations, $\boldsymbol{\nu}_k$ is the innovation at each time step k:

$$\boldsymbol{\nu}_k = \mathbf{d}_k - \hat{\mathbf{d}}_{k|k-1} = \mathbf{d}_k - \mathbf{h}\left(\mathbf{x}_{k|k-1}\right) \tag{19}$$

and $\mathbf{V}_k$ is the variance of the k^{th} innovation:

$$\mathbf{V}_k = \mathbf{H}_k \mathbf{\Sigma}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \tag{20}$$

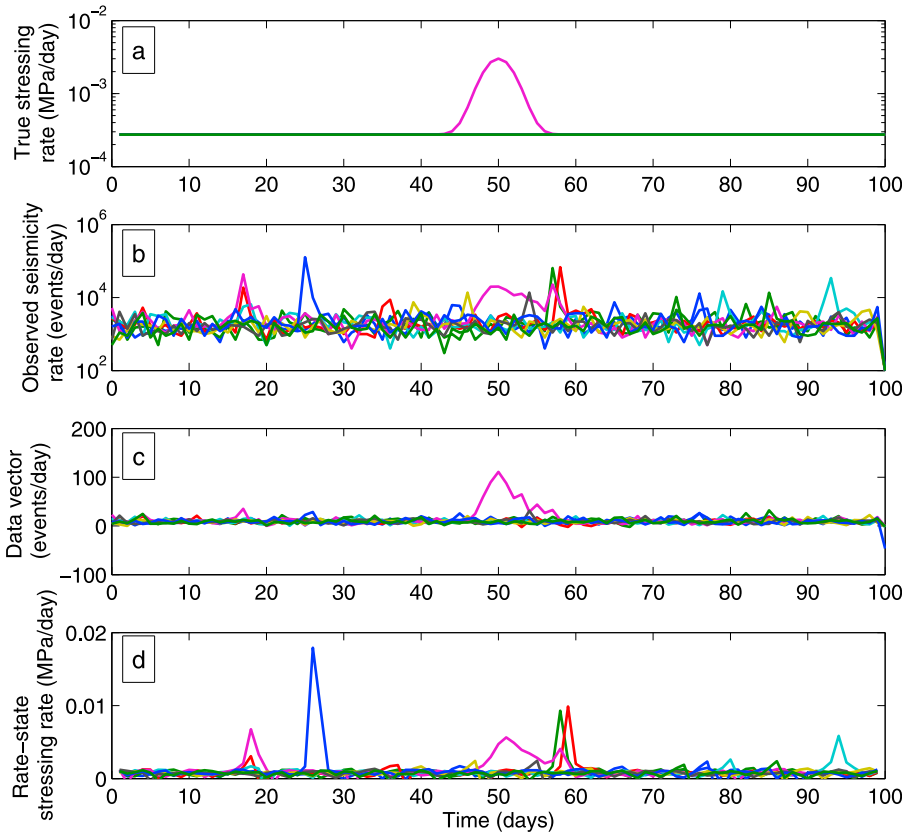

Figure 2. (a) True stressing rate $\dot{S}$ used in each spatial box for the synthetic test. A transient rate is imposed in box 5 (purple), involving an increase in $\dot{S}$ by a factor of 10 over the background rate of 0.1 MPa/yr imposed in the other boxes (green). (b) Observed seismicity rates in each spatial box (colors) obtained by binning occurrence times in time windows of 1 day. The anomalous seismicity rate due to the transient in box 5 (purple) is overshadowed by jumps in rate due to aftershock sequences, therefore the stressing rate estimate from a straightforward rate-state inversion will primarily reflect changes due to these larger peaks. (c) Data vectors (i.e., the aseismically triggered seismicity rate R_A) in each box obtained by subtracting out the ETAS-predicted rate. The anomalous seismicity in box 5 (purple) due to the transient is now much more visible. (d) Stressing rate in each box estimated using the rate-state inversion of *Dieterich et al.* [2000]. The transient in box 5 is overshadowed by the larger stressing rate changes due to the aftershock sequences in the catalog.

This likelihood function also employs a maximum likelihood estimate (MLE) for the data covariance scalar multiplier ξ^2, which is the value to which the data covariance matrix **R** is assumed to be known (see Appendix A):

$$\hat{\xi}^2 = \frac{1}{N_d}\sum \boldsymbol{\nu}_k^T \mathbf{V}_k^{-1} \boldsymbol{\nu}_k \tag{21}$$

With these equations, we then grid search over multiple non-ETAS parameter sets to determine their maximum likelihood estimate, with initial grid spacing of a factor of 10 for each parameter.

[26] A fundamental assumption we have made in this initial implementation is that the data and errors are Gaussian distributed, which is likely incorrect in that earthquake statistics typically involve non-Gaussian distributions. Future implementations could incorporate non-Gaussian error distributions by utilizing particle filter methods [e.g., *Fukuda and Johnson*, 2008; *Werner*, 2008; *Werner et al.*, 2011].

2.6 Summary

[27] Our complete algorithm can be summarized as follows:

[28] 1. Determine magnitude of completeness for the catalog.

[29] 2. Optimize the ETAS parameters for the entire catalog using a space-time ETAS algorithm [*Zhuang et al.*, 2005; *Ogata and Zhuang*, 2006].

[30] 3. Subtract the ETAS-predicted rate from the observed (binned) seismicity rate in each of N spatial boxes to calculate R_A, which is the data vector for the filter.

[31] 4. For a given set of non-ETAS parameters, run the extended Kalman filter to obtain the time history of the state variables.

[32] 5. Compute the Gaussian likelihood associated with the set of non-ETAS parameters.

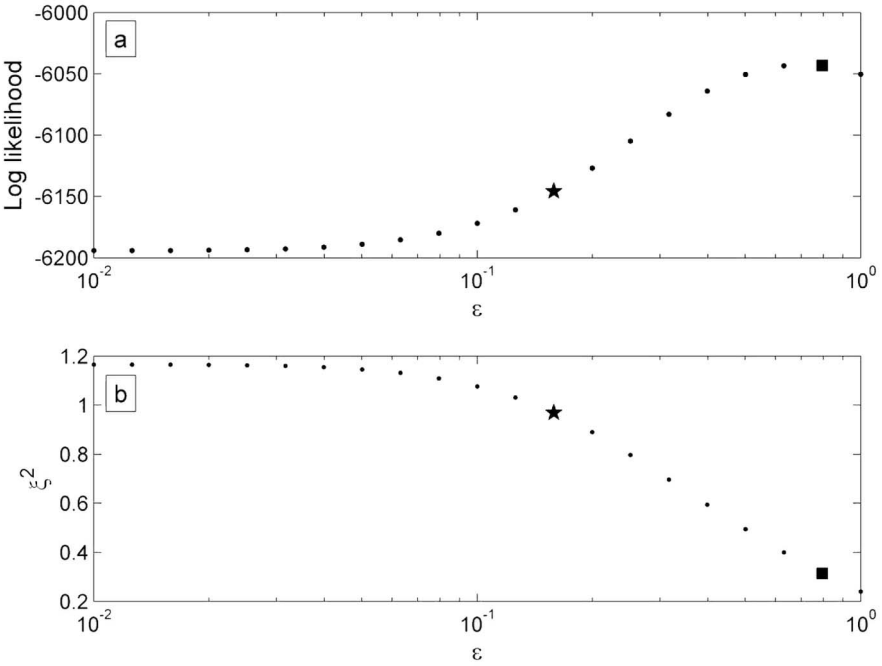

Figure 3. (a) Log likelihood versus smoothing parameter ε. (b) Data covariance multiplier $\hat{\xi}^2$ versus ε. Ideally $\hat{\xi}^2$ should be close to 1, because the data covariance $\mathbf{R}_k$ is estimated a priori. The MLE of ε is indicated by a square. Our preferred value for ε, for which $\hat{\xi}^2 \sim 1$, is indicated by a star.

[33] 6. Repeat steps 4 and 5 as a grid search for different parameter sets to obtain maximum likelihood estimates of the non-ETAS parameters.

3. Synthetic Test

[34] To test the validity of our inversion algorithm, we generate a synthetic catalog of events using known model parameters and a stressing rate history that we will attempt to recover. We subdivide a 30 km by 30 km region into 9 spatial boxes (Figure 1a) and impose a factor of 10 increase in stressing rate above a background rate of 0.1 MPa/yr in the center box (box 5). The stressing rate in the other boxes remains constant at the background rate (Figure 2a). Given these stressing rate histories, we use the rate-state equations (equations (5)–(6)) to calculate associated background seismicity rate histories, which are used in an ETAS simulator [*Felzer et al.*, 2002; *Helmstetter and Sornette*, 2002, 2003b] to generate a synthetic catalog with a magnitude of completeness $M_C = 0$. Figure 1 shows the catalog event locations, magnitudes and times of occurrence. Occurrence times were binned in time windows of 1 day to obtain seismicity rates. Aftershocks of the largest events dominate the catalog, obscuring the seismicity rate variations in both space and time due to the transient (Figure 2b). The anomalous seismicity triggered by the transient only becomes readily apparent following the removal of the ETAS-predicted rate (Figure 2c). If the *Dieterich et al.* [2000] rate-state inversion was applied to this catalog prior to the removal of the ETAS-predicted rate, the resulting stressing rate estimates would be dominated by stress changes due to the aftershocks rather than the aseismic transient of interest (Figure 2d).

3.1. Parameter Estimation

[35] We first estimate the best fitting ETAS parameters using the space-time algorithm of *Ogata and Zhuang* [2006]. The parameters used in the simulator were K = 4.5e-4 events/day/deg^2, $\alpha = 0.8$, $p = 1.2$, $c = 0.0001$ days, $d = 0.001$ deg^2, $q = 3$, and $\eta = 0.3$. The maximum likelihood estimates of these parameters from the estimation algorithm were $K = 0.15$ events/day/deg^2, $\alpha = 1.8$, $p = 1.2$, $c = 0.0001$ days, d = 0.5e-8 deg^2, $q = 1.35$, and $\eta = 0.58$. The simulator used a different form for the cluster size factor $\kappa(M)$ (equation (2)) to generate the synthetic catalog [*Felzer et al.*, 2002; *Helmstetter and Sornette*, 2003b] than the form utilized in the estimation algorithm [*Ogata and Zhuang*, 2006], resulting in differences between the true and estimated values of the parameters K and α. The simulator also uses a slightly different form of the spatial probability density function (equation (4)) than that used by the estimation algorithm, causing differences in the estimates of the spatial scaling parameters d, q, and η. However, both the simulation and estimation algorithms employed the same temporal function based on the modified Omori law (equation (3)), and the estimation algorithm recovers the true values for p and c even though the assumption of constant background seismicity rate is violated. Thus, the space-time ETAS parameters are well resolved by the estimation algorithm, despite the background rate being time-dependent.

[36] We subtract the ETAS-predicted seismicity rate from the binned seismicity rate to obtain the data vector (equations (8) and (10)), which is the input for the extended Kalman filter. We ran the filter with a number of different values for the non-ETAS parameters (r, $A\sigma$, τ, and ε) and calculated the Gaussian

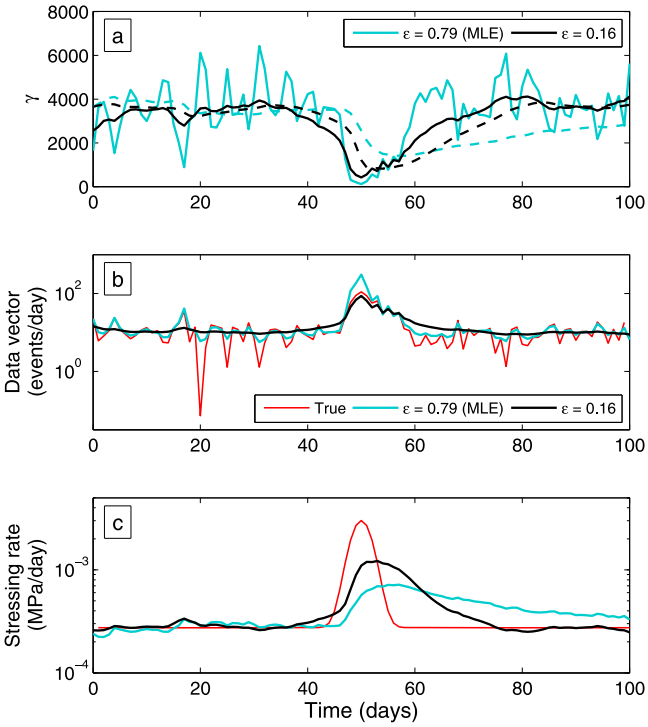

Figure 4. Comparison of back-smoothed filter results in box 5 for different values of ε. (a) Estimates for the rate-state seismicity state variable γ for the MLE of $\varepsilon = 0.79$ (teal), and our preferred value of $\varepsilon = 0.16$ (black). Also shown are the estimates for γ directly derived from the filter (solid lines) and estimates for γ integrated from the filter estimate of $\dot{S}$ using the rate-state model (equations (5)–(6)) (dashed lines). Ideally the two estimates at a given value of ε should agree, which is better achieved when $\varepsilon = 0.16$. (b) Fit to the observed data vector in box 5 (red) by each value of ε. For the MLE (teal), the filter clearly overfits the data. The $\varepsilon = 0.16$ case (black) better fits the largest peak. (c) Forward filter estimates for $\dot{S}$ in box 5 with $\varepsilon = 0.79$ (teal) and $\varepsilon = 0.16$ (black), compared to the true $\dot{S}$ (red). The $\varepsilon = 0.16$ estimate comes closer to recovering the true stressing rate.

likelihood for each filter run (equation (18)). Again, the fundamental assumption made here is that the data and errors are Gaussian distributed (see section 2.5). The parameter r could only be constrained to within an order of magnitude of the true value, and we observed a trade-off in likelihood between the parameters $A\sigma$ and τ, suggesting there is little constraint in $A\sigma$. Moreover, an order of magnitude change in the value of $A\sigma$ (e.g., from 0.001 MPa to 0.01 MPa) resulted in a change of the estimated peak stressing rate by a factor of 2, suggesting that our method is not sensitive to this parameter. Therefore, we fixed $A\sigma$ and r to be their true values (0.001 MPa and 10 events/day) and grid searched over τ and ε, the scaling parameters for the random walk processes used to model $\dot{S}_A$ and $\ln(\gamma)$ respectively.

[37] We employ two further constraints to select appropriate values for τ and ε. The first constraint comes from enforcing consistency between the filter's estimates of the state variables $\dot{S}_A$ and $\ln(\gamma)$. Because we model these variables as stochastic processes with noise specified by covariance matrix $\mathbf{Q}_k$ (equations (11) and (A8)), they do not have to strictly obey equations (5)–(6). Thus, the filter estimate of total stressing rate $\dot{S}$(the sum of the estimates of $\dot{S}_A$ and $\dot{S}_p$) can be integrated using the rate-state equations (equations (5)–(6)) to produce an estimate of γ that is not identical to the filter estimate for γ. Ideally these two estimates of γ should closely agree with one another. This places a constraint on the value of τ (i.e., the temporal smoothing of $\dot{S}_A$), since it must be large enough to allow enough variation in $\dot{S}$ for it to integrate to γ. However, if τ is too large (i.e., $\dot{S}_A$ too rough), the error covariance matrix $\mathbf{\Sigma}$ becomes nonpositive semidefinite, and the prediction step (equation (13)) can no longer be computed.

[38] A second constraint is employed to help choose a value for ε. Figure 3 shows the Gaussian likelihood and maximum likelihood estimate of the data covariance multiplier $\hat{\xi}^2$ for various values of ε, calculated using equations (18) and (21). The maximum likelihood occurs when $\varepsilon = 0.79$ (Figure 3a); however, at this value of ε, $\hat{\xi}^2$ is significantly less than 1 (Figure 3b). Because we estimate $\mathbf{R}_k$ a priori based on the variance of the data vector, $\hat{\xi}^2$ should be ~1. Values other than 1 will overfit or under fit the data. Therefore, we choose ε such that $\hat{\xi}^2 \sim 1$. Moreover, at higher values of ε, the filter estimate for $\dot{S}$ fails to integrate to match the filter estimate for γ as well as at smaller values of ε, because the stressing rate estimate cannot vary enough to match the variance in γ (Figure 4a). Additionally, higher values of ε ultimately overfit the data (Figure 4b). While the filter fits the smaller peaks in the data vector well, it overpredicts the largest peak. Smaller values of ε smooth out the smaller peaks in seismicity rate but end up fitting the largest peak well. Therefore, we identify a preferred value for ε by determining the MLE and then running the filter for increasingly smaller values of ε until $\hat{\xi}^2 \sim 1$. Our preferred value for ε is thus 0.16, for which $\hat{\xi}^2 \sim 1$, rather than the MLE of 0.79. The stressing rate estimate results show that the estimate for $\dot{S}$ with $\varepsilon = 0.16$ actually comes closer to recovering the true $\dot{S}$ than the MLE value of ε (Figure 4c).

[39] To summarize, of the 11 model parameters used in our algorithm, the seven space-time ETAS parameters are well-resolved by maximizing a point process likelihood function [*Ogata and Zhuang*, 2006]. The two random walk scaling parameters τ and ε can be optimized using a Gaussian likelihood function subject to constraints. However, neither likelihood function well constrains the rate-state parameters r and $A\sigma$ and thus these parameters should be fixed a priori.

3.2. Synthetic Test Results

[40] Figure 5 shows the resulting filter estimates of $\dot{S}$ for each box using $\varepsilon = 0.16$. Both the forward- and back-smoothed filter estimates clearly identify the transient in box 5. While not quite as large as the true stressing rate, the filter-estimated peak stressing rate is still within a factor of 2–3 of the true rate. The forward filter estimate delays the onset of the transient compared to the true onset, although the back-smoothed estimate mitigates this effect somewhat. Clearly, the filter detects when and in which region the transient is occurring (Figure 5). This is an encouraging result, because this test involved an increase in stressing rate of only a factor of 10 over the tectonic rate.

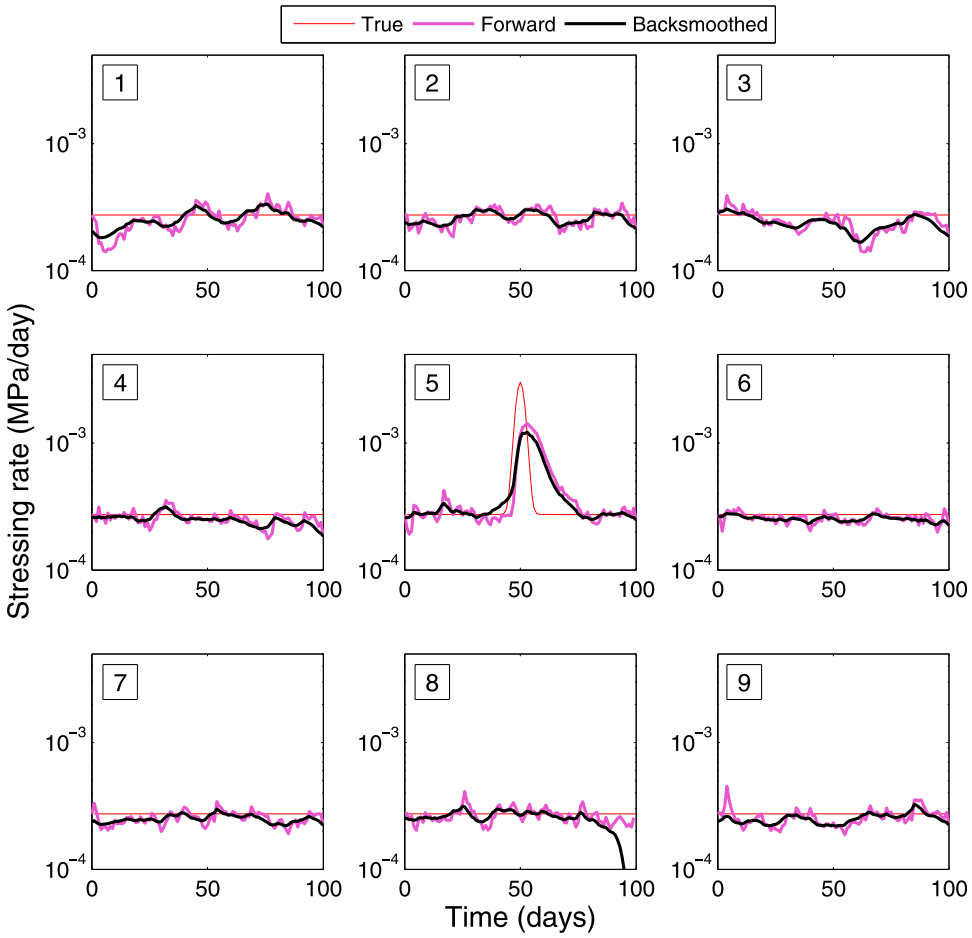

Figure 5. Filter estimates of stressing rate $\dot{S}$ in each box. The red line indicates true stressing rate, the purple line indicates the forward estimate, and the black line indicates the back-smoothed estimate. Both the forward- and back-smoothed results clearly identify in which box the transient is located and produce estimates of the peak stressing rate of the transient that are within a factor of 2–3 of the actual peak.

The stressing rate jumps that occur during slow earthquakes and magma intrusions are likely to be several orders of magnitude larger [*Thatcher*, 2001; *Toda et al.*, 2002; *Lohman and McGuire*, 2007].

4. Data Analysis: Salton Trough

[41] We next apply our algorithm to real earthquake data from the Salton Trough in southern California. In the Salton Trough, a transition occurs from a divergent plate boundary setting in the Gulf of California to the south, to the San Andreas strike-slip fault system to the north. The region is characterized by high heat flow [*Kisslinger and Jones*, 1991], which potentially acts to subdue aftershock activity [*Ben-Zion and Lyakhovsky*, 2006; *Yang and Ben-Zion*, 2009; *Enescu et al.*, 2009]. A high rate of earthquake swarm activity has been observed [e.g., *Richter*, 1958; *Brune and Allen*, 1967; *Hill et al.*, 1975; *Johnson and Hadley*, 1976; *Lohman and McGuire*, 2007; *Roland and McGuire*, 2009], possibly driven by magmatic intrusion [*Hill*, 1977] or aseismic fault creep [*Lohman and McGuire*, 2007; *Roland and McGuire*, 2009]. Geodetic measurements have also detected a number of aseismic transients in this region, including afterslip following the 1987 M6.6 Superstition Hills earthquake [*Williams and Magistrale*, 1989], creep events on the Superstition Hills fault [*Wei et al.*, 2009], and aseismic creep on the Obsidian Buttes fault [*Lohman and McGuire*, 2007].

4.1. Data Binning

[42] We analyze a catalog of M $\geq$ 1.5 earthquakes that occurred in the Salton Trough from February 1990 to August 2009. We choose a magnitude cutoff of 1.5 based on frequency magnitude plots of the entire data set, as well as in each spatial bin and in time windows following the largest events to ensure that the cutoff magnitude does not change over the space-time windows under consideration. Because of the partial derivatives involved in setting up the state-space model equations (e.g., equations (A5) and (A7)), we need to be able to resolve the background seismicity rate in each space-time window. Therefore we divide the region up into 4 spatial boxes (Figure 6) and bin the occurrence times into time windows of 20 days to obtain seismicity rates in

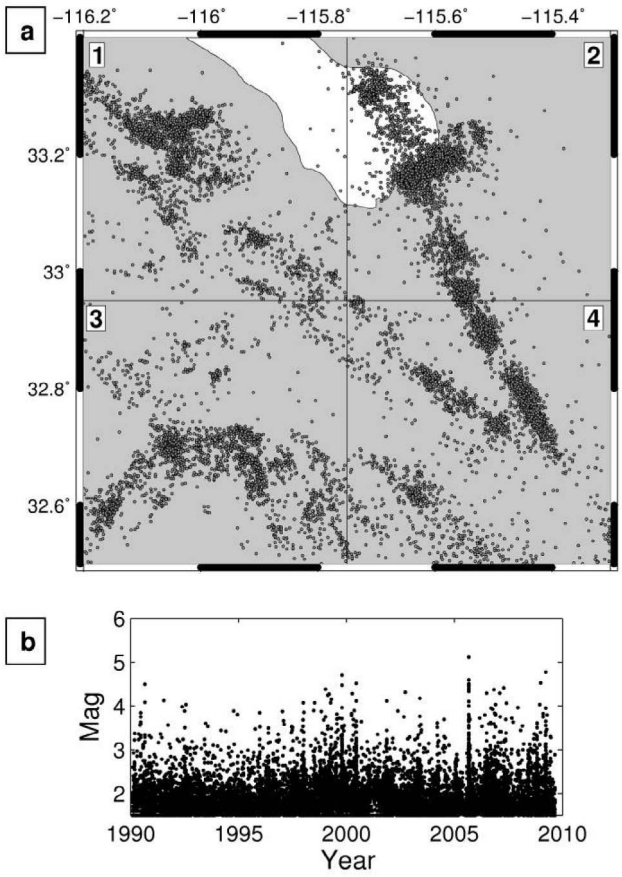

Figure 6. (a) Map of the Salton Trough region in California showing M ≥ 1.5 seismicity occurring from February 1990 to August 2009, obtained from the Southern California Earthquake Data Center. For our analysis, the region is divided into the 4 boxes indicated. (b) Magnitude-time history of the Salton Trough catalog.

each box. This space-time window allows us to obtain enough earthquakes in each bin to resolve the background rate. Any spatial binning scheme that obeys this criterion, such as a fault-based algorithm similar to those used in stress inversions [e.g., *Hardebeck and Hauksson*, 1999], would work and can be handled using the numerical integration technique described by *Ogata* [1998].

4.2. Parameter Estimation

[43] We first fit the space-time ETAS model to the 2005 M5.1 Obsidian Buttes earthquake to account for as much of the aftershock behavior as possible. From the space-time ETAS estimation algorithm [*Ogata and Zhuang*, 2006], the MLE for the ETAS parameters are $K = 0.53$ events/day/deg^2, $\alpha = 0.92$, $p = 1.3$, $c = 0.01$ days, $d = 4.8\text{e-}5$ deg^2, $q = 2.63$, and $\eta = 0.23$. However, the data vector formed using these parameters resulted in a number of negative seismicity rate values primarily due to the estimate of p, which can lead to instabilities in the filter due to the assumption of a Gaussian error distribution. Therefore, we instead use the ordinary time ETAS parameters ($K = 0.61$ events/day/deg^2, $\alpha = 0.88$, $p = 1.1$, and $c = 0.001$ days) fit to this catalog [*Llenos et al.*, 2009].

[44] We subtract the ETAS-predicted rate from the observed seismicity rate to form the data vector and estimate the data covariance $\mathbf{R}_k$. Again because of the tradeoff between the parameters τ and $A\sigma$ and our lack of sensitivity to $A\sigma$, we fix $A\sigma$ to 1 MPa. As the synthetic test demonstrates (see section 3.1), uncertainties in the parameter $A\sigma$ ultimately translates into uncertainties in the stressing rate estimate. Assuming that $A = 0.01$ from laboratory observations [*Dieterich*, 1994], this value of $A\sigma$ is consistent with faults that fail under hydrostatic conditions at a depth of ~4 km, the depth at which the Obsidian Buttes swarm occurred [*Chester and Higgs*, 1992; *Blanpied et al.*, 1998; *Lohman and McGuire*, 2007]. We then grid search over values of τ and ε. Using the constraint that the integral of the filter estimate of stressing rate must be consistent with the filter estimate of γ, we choose a value of $\tau = 2.5\text{e-}4$ and maximize the likelihood over values of ε (Figure 7). The maximum likelihood estimate occurs when $\varepsilon = 0.5$ (Figure 7a) but again, as in the synthetic test, at this value the data covariance multiplier $\hat{\xi}^2$ is significantly less than 1 (Figure 7b), and the filter overfits the data. Our preferred value is $\varepsilon = 0.04$, for which $\hat{\xi}^2 \sim 1$, resulting in a better fit for the largest peaks in the data.

4.3. Results

[45] Figure 8 shows the stressing rate estimates for each box using $\varepsilon = 0.04$, illustrating the filter's ability to detect when and in which box the largest transient in the region occurs. The largest anomaly occurs in box 2 and is associated with a geodetically observed shallow aseismic creep event on the Obsidian Buttes fault that triggered an earthquake swarm in 2005 [*Lohman and McGuire*, 2007]. The peak forward estimate of stressing rate is 0.042 ± 0.004 MPa/day and the back-smoothed estimate is 0.022 ± 0.006 MPa/day, roughly two orders of magnitude above tectonic loading.

[46] The second largest signal also occurred in box 2 and corresponds to the Bombay Beach earthquake swarm that occurred in March 2009. The swarm consisted of ~100s of events, the largest of which was a M4.8 that occurred three days after the swarm initiated. We also identify small anomalies in boxes 2 and 4 in May 2003 that may be related to an earthquake swarm that occurred in the Imperial fault zone (located near the boundary of the two boxes) [*Roland and McGuire*, 2009]. We can also associate smaller anomalies in box 2 with earthquake swarms that occurred in the Brawley seismic zone in 1996, 1998, and 2008 [Southern California Earthquake Center, http://www.data.scec.org/monthly/index.php]. While we cannot rule out the possibility of fluid flow triggering these smaller swarms, other swarms in the Salton Trough exhibit migration rates of 0.1–1 km/h which correspond to typical rupture propagation velocities of aseismic creep events [*Roland and McGuire*, 2009, and references therein], rather than the rates of fractions of kilometers per day associated with fluid flow [e.g., *Hainzl and Ogata*, 2005].

5. Discussion

[47] Our results highlight the need for a time-dependent background seismicity rate to account for variations in seis-

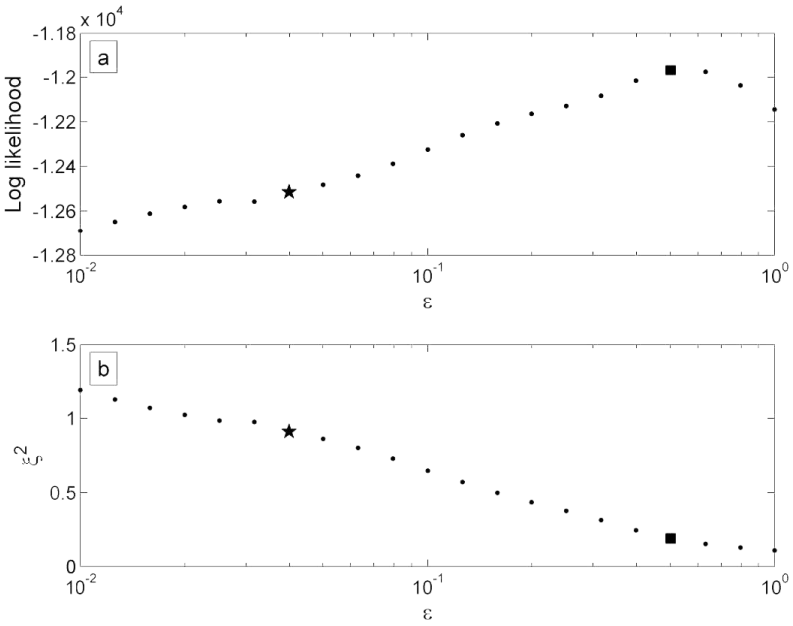

Figure 7. (a) Log likelihood versus ε for the Salton Trough data set. (b) Data covariance multiplier $\hat{\xi}^2$ versus ε. Again the MLE of ε (square) differs from our preferred value for ε (star), for which $\hat{\xi}^2 \sim 1$. As in the synthetic test, the MLE value of ε (0.5) overfits the data, while the preferred value of ε (0.04) fits the largest peaks better and results in closer agreement between the direct and integrated filter estimates of γ.

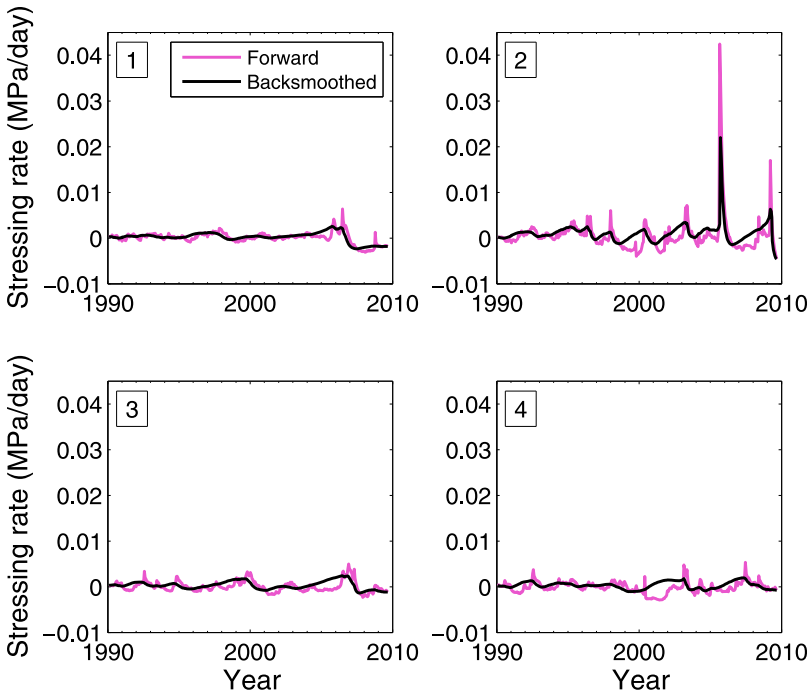

Figure 8. Filter estimates of $\dot{S}$ in each spatial box for the Salton Trough. The purple line indicates the forward estimate, the black line indicates the back-smoothed estimate. The largest signal, in box 2, corresponds to a geodetically observed aseismic transient in the Obsidian Buttes in 2005 [*Lohman and McGuire*, 2007]. The next largest signal, also in box 2, relates to the 2009 Bombay Beach earthquake swarm. Other smaller anomalies may be related to an earthquake swarm on the Imperial Fault in 2003 [*Roland and McGuire*, 2009] in boxes 2 and 4, earthquake swarms in the Brawley seismic zone in 1996, 1998, and 2008 in box 2, and the 2009 Bombay Beach swarm in box 2.

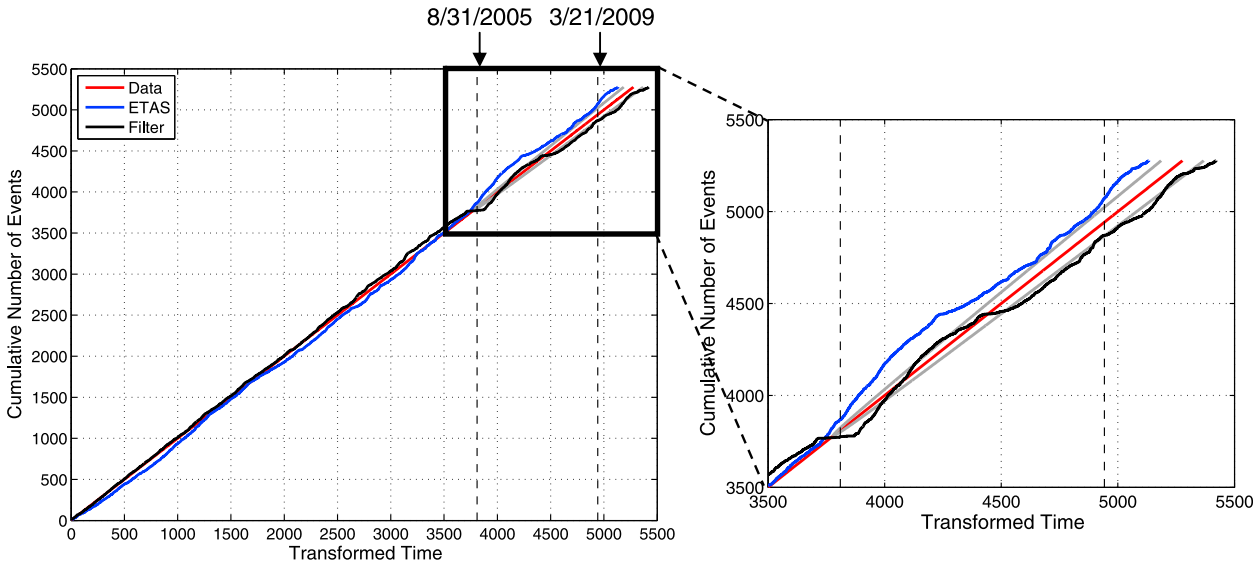

Figure 9. Cumulative number of events versus transformed time (i.e., predicted cumulative number of events). The red line is a one-to-one line indicating a perfect fit to the observed data. ETAS transformed times are calculated with seismicity rates estimated from the space-time ETAS model optimized to just prior to the 2005 Obsidian Buttes earthquake swarm (event 3779) and extrapolated for the remainder of the catalog (blue line). Transformed times are also calculated using the filter estimate of seismicity rate (black line). The significant deviation of the blue line from the data (red line; 2σ bounds shown by gray line) shows that the ETAS model (with a time-independent background rate) underpredicts the amount of seismicity, particularly during the 2005 swarm. The filter estimate (with a time-dependent background rate) provides a better fit to the observed cumulative number of events.

micity rate due to aseismic processes, as other studies have suggested [e.g., *Hainzl and Ogata*, 2005; *Lombardi et al.*, 2006; *Lombardi and Marzocchi*, 2007; *Lombardi et al.*, 2010]. Figure 9 compares the observed cumulative number of events in the Salton Trough catalog with the number of events predicted from optimizing the space-time ETAS model to the part of the catalog that occurred prior to the 2005 Obsidian Buttes swarm, and the number of events predicted from the filter estimate of seismicity rate. We transform the occurrence times t_i of the events in the catalog with the theoretical cumulative function $\tau_i = \int_0^{t_i} \lambda(s)ds$, where λ is the predicted seismicity rate from either ETAS or the filter [*Ogata*, 1988, 2005]. A plot of the cumulative number of events versus transformed time should be linear if the seismicity in the catalog is well described by a particular model. The 2σ error bars of the extrapolation can be calculated using $\sigma = [\tau - \Lambda(0,T) + \{\tau - \Lambda(0,T)\}^2/\Lambda(0,T)]^{1/2}$, based on the fact that the cumulative curves of the transformed times after $\Lambda(0,T)$ (where $\Lambda(0,T)$ is the transformed time of the last event that occurred during the time period $[0, T]$ over which the ETAS model was optimized) should behave as a standard Brownian process [*Ogata*, 2005]. Positive (or negative) deviations from this linear trend indicate that the model underpredicts (or overpredicts) the amount of seismicity. Figure 9 shows that with the space-time ETAS model, a significant positive deviation from this trend occurs near the beginning of the Obsidian Buttes swarm, suggesting that anomalous seismicity is occurring that the ETAS model alone cannot explain. The filter prediction however matches the observed seismicity well. Therefore, the time-dependent background seismicity rate produced by our filter algorithm can account for the seismicity rate anomalies that appear with respect to the space-time ETAS model, which utilizes a time-independent background seismicity rate.

[48] To validate the estimates of $\dot{S}$ obtained from these seismicity rate variations, we compare our peak stressing rate estimate for the 2005 Obsidian Buttes aseismic transient to an estimate based on a slip model of the deformation inverted from Interferometric Synthetic Aperture Radar (InSAR) data [*Lohman and McGuire*, 2007]. The seismicity triggered by this transient occurred primarily in the depth range of 4–6 km. We calculate the Coulomb stress change [*King et al.*, 1994; *Lin and Stein*, 2004; *Toda et al.*, 2005] at this depth range due to the aseismic slip on the shallow part of the fault and obtain an average total Coulomb stress change of 0.6 MPa (Figure 10). Based on GPS line length change data, the transient lasted ~1–10 days [*Lohman and McGuire*, 2007]. Given this range of durations, the average stressing rate during the transient then becomes ~0.06–0.6 MPa/day. For a duration of 5 days (which appears to best describe the GPS data), the average stressing rate is 0.12 MPa/day.

[49] We compare this average stressing rate to our peak stressing rate estimates for the transient. We obtained a peak stressing rate of ~0.04 MPa/day from the forward filter estimate and a peak stressing rate of ~0.02 MPa/day from the back-smoothed estimate (Figure 8). Thus, the results from inverting the seismicity catalog are within a factor of 5

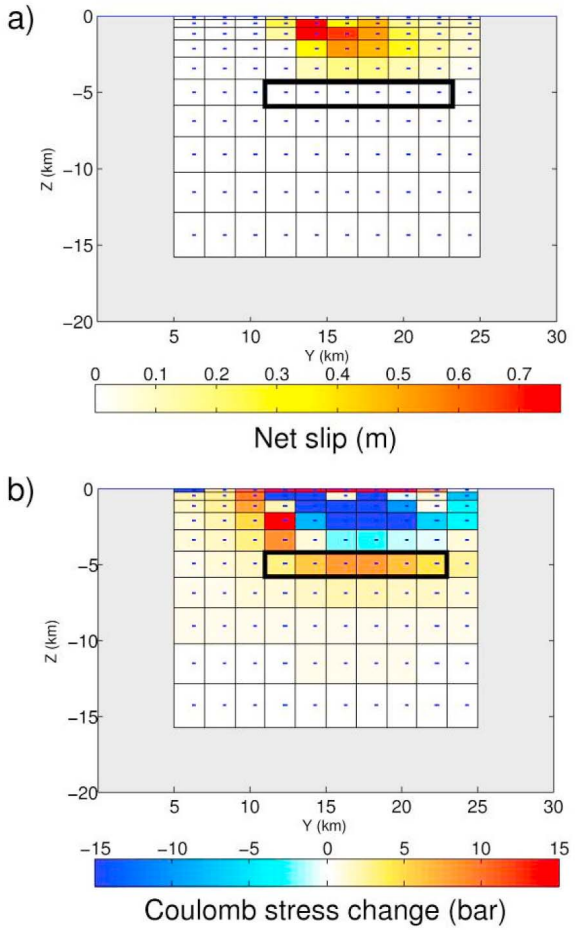

Figure 10. (a) Slip model of shallow aseismic creep on the Obsidian Buttes fault inverted from InSAR data [*Lohman and McGuire*, 2007]. Black box indicates depth at which microseismicity triggered by the transient occurred. (b) Calculated Coulomb stress change on the fault due to the shallow aseismic slip. The Coulomb stress change averaged over the outlined boxes is used to calculate an average stressing rate for the transient (~0.06–0.6 MPa/day for transient durations of 1–10 days).

of the average stressing rate estimated from the geodetic data. Moreover, if we take the duration of the transient to match the time step in our filter (20 days) then the estimates agree extremely well (0.03 MPa/day from the stress calculation versus 0.02–0.04 MPa/day from the filter). Given that stressing rate increases are likely to be many orders of magnitude over background plate tectonic rates [*Thatcher*, 2001; *Toda et al.*, 2002; *Lohman and McGuire*, 2007], the Salton Trough example demonstrates the feasibility of utilizing our approach to both detect and constrain the magnitude of stressing rate transients.

[50] The second largest anomaly in the filter stressing rate estimate (Figure 8) is related to the 2009 Bombay Beach earthquake swarm, which began on 21 March 2009 and lasted ~1 week. The largest event was a M4.8 that occurred on 24 March 2009, three days after the swarm began. The swarm occurred on the northernmost part of a series of ladder faults offshore of Bombay Beach (Figure 11a). Geodetic data are limited because of the fault's offshore location, but the nearest GPS station (DHLG) observed an offset at the time of the swarm. We fit daily GPS solutions from the routine Plate Boundary Observatory analyses and found a 1.0 mm offset in the east component and a −0.8 mm offset in the north component, relative to station P504 (Figures 11b and 11c). To determine if the signal can be explained by the earthquake swarm or if it requires an aseismic deformation event, we constructed a simple conservative forward model of the ground deformation due to the swarm. Summing the seismic moment released during the swarm, we calculated an average focal mechanism and placed it on the northernmost ladder fault. We estimated a rupture length and width from the moment release using empirical scaling relations [*Wells and Coppersmith*, 1994] and assumed a shear modulus of 25 GPa. From this forward model, we obtained offsets of −0.1 mm in the east component and −0.2 mm in the north component. Our modeled displacements are factors of 10 (east component) and 4 (north component) smaller than the observed, which allows for the possibility that aseismic deformation occurred. However, the GPS data are not conclusive, and deformation was not observed in the laser strainmeter data at Durmid Hills (D. Agnew, personal communication, 2010).

[51] In the rate-state model, stressing rate estimates depend to an extent on the value of $A\sigma$, because this parameter controls both the instantaneous change in seismicity rate following a stress change as well as the evolution of the state variable γ [*Catalli et al.*, 2008; *Llenos et al.*, 2009; *Cocco et al.*, 2010]. However, as both the synthetic test and data analysis demonstrate, our method is relatively insensitive to variations in this parameter. For the synthetic test, varying $A\sigma$ by a factor of 10 led to a change in the peak stressing rate estimate of a factor of 2. For the Salton Trough, varying this parameter by a factor of 10 led to a change in the peak stressing rate estimate of a factor of ~4. Therefore, our method can still be used to constrain relative changes in stressing rate on an order-of-magnitude scale.

[52] Because of our lack of sensitivity to the actual value of $A\sigma$, we utilized a physically motivated value in our Salton Trough analysis. This approach was successfully applied to detect stress changes due to a dike intrusion at Kilauea [*Dieterich et al.*, 2000]. Using a value of $A\sigma$ consistent with hydrostatic fault conditions at the depth at which the triggered seismicity occurred, *Dieterich et al.* [2000] obtained estimates of stress changes from the seismicity data that agreed with the stress changes calculated from geodetically constrained boundary element models within an order of magnitude. Similarly, we do not attempt to constrain the true value of $A\sigma$ and instead choose a value consistent with local conditions, because we are primarily concerned with detecting order-of-magnitude changes in stressing rate. Our results and subsequent validation with a geodetically derived model of deformation suggest that our method is successfully able to do so.

[53] Last, it is possible that the peaks in seismicity rate observed in box 2 are artifacts due to undetected seismicity. The ETAS estimate of background seismicity may be overestimated, particularly following large events, due to the

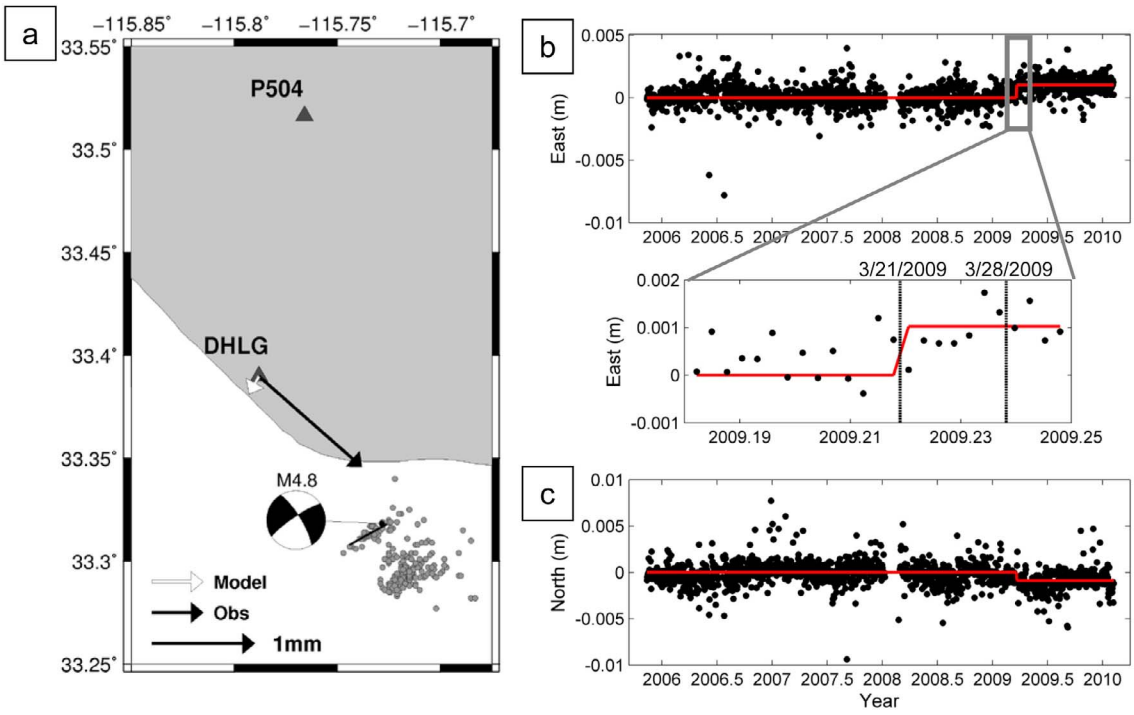

Figure 11. (a) Map of the Bombay Beach region showing M ≥ 1.5 seismicity during the swarm (gray dots). Vectors indicate the observed (black) and modeled (white) displacement at station DHLG relative to station P504. Displacements were modeled assuming the total moment release of the swarm occurred on the northernmost ladder fault. The focal mechanism of the M4.8 event is also shown. (b) Eastward component of displacement at DHLG relative to P504, observed from daily GPS solutions from PBO analyses (black dots). Linear trend (red line) obtained by fitting the time periods before and after the swarm. An offset of ~1mm occurs around the time of the swarm. (c) North component of displacement at DHLG relative to P504 (black dots), with linear trend (red line) fit to before and after the swarm. An offset of 0.8mm to the south occurs around the time of the swarm.

effect of seismicity below the magnitude cutoff triggering larger events above the cutoff [*Sornette and Werner*, 2005]. However, in this case the two largest spikes that we find correspond to transient signals that were also observed on GPS (i.e., the 2005 Obsidian Buttes transient and the 2009 Bombay Beach transient). Moreover, geodetically observed transients in other regions such as Kilauea and Boso have been shown to trigger significant changes in the ETAS-estimated background seismicity rate [*Llenos et al.*, 2009]. Finally, *Lombardi et al.* [2010] recently used simulated catalogs to demonstrate that a bias from undetected seismicity could not explain the changes in background seismicity rate observed during the 1997 Umbria-Marche earthquake sequence, which they attribute to fluid flow. Therefore, while it is possible that undetected events may cause apparent spikes in background seismicity rate, it is likely that the largest signals we detect are real variations, particularly since they correlate with GPS-detected transients.

6. Conclusion

[54] We have developed a technique to detect aseismic transients in time and space from earthquake catalog data by combining the ETAS and rate-state models of seismicity rate into a single data assimilation algorithm to invert catalogs for stressing rate variations. We applied it to a synthetic test and successfully detected an aseismic transient that involved a factor of 10 increase in stressing rate. We then applied it to a catalog from the Salton Trough in California, and successfully detected the onset and constrained the absolute magnitude of the largest aseismic transient in a 20 year catalog to within a factor of five of the stressing rate estimated with geodetic data. We also detected an anomaly related to the 2009 Bombay Beach swarm occurring around the same time as an offset observed at a nearby GPS station, suggesting that aseismic deformation may have occurred.

[55] Overall, the synthetic test and Salton Trough results suggest that our algorithm is a feasible way to detect aseismic stressing rate transients strictly from seismicity catalog data. This method may ultimately enable aseismic transient detection in regions lacking good geodetic data resolution, such as the (offshore) updip part of subduction zone faults, and in time periods prior to the widespread availability of geodetic data. Additionally, a seismicity based approach may be more sensitive to small (M4–5) and/or shallow slow-slip transients that are not detected by even dense geodetic networks such as the Plate Boundary Observatory [*Wei et al.*, 2009]. The results suggest that our seismicity inversion method provides an accurate way to detect and locate transient deformation strictly from seis-

micity catalogs and can constrain the absolute magnitude of stressing rate variations.

Appendix A: State-Space Model Specification

[56] This appendix details the state-space system used in our algorithm to invert for stressing rate variations. Our state-space model consists of an observation equation:

$$\mathbf{d}_k = \mathbf{h}_k(\mathbf{x}_k) + \omega_k, \quad \omega_k \sim N(0, \mathbf{R}_k) \tag{A1}$$

and a state transition equation:

$$\mathbf{x}_{k+1} = \mathbf{t}(\mathbf{x}_k) + \delta_{k+1}, \quad \delta_{k+1} \sim N(0, \mathbf{Q}_{k+1}) \tag{A2}$$

The observations $\mathbf{d}$ are the aseismically triggered seismicity rates R_A in each spatial bin (i.e., the observed seismicity rates R with the ETAS-predicted Omori sequences removed), and the state variables $\mathbf{x}_k$ consist of long-term tectonic loading rate $\dot{S}_p$, aseismic stressing rate $\dot{S}_A$, and the logarithm of the rate-state seismicity state variable γ in each of N spatial boxes. Thus the full state vector is:

$$\mathbf{x}_k = \left[\dot{S}^1_{p_k}, \dot{S}^1_{A_k}, \ln\gamma^1_k, \dot{S}^2_{p_k}, \dot{S}^2_{A_k}, \ln\gamma^2_k, ..., \dot{S}^N_{p_k}, \dot{S}^N_{A_k}, \ln\gamma^N_k\right] \tag{A3}$$

The $N \times 3N$ measurement matrix $\mathbf{h}$, relating the observed seismicity rates to the underlying state variables, is nonlinear, and formed with the rate-state model (equation (5)):

$$\mathbf{h}(\mathbf{x}_k) = \begin{bmatrix} \begin{bmatrix} 0 & 0 & \dfrac{r}{\gamma^1_k \dot{S}_r} \end{bmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{bmatrix} 0 & 0 & \dfrac{r}{\gamma^N_k \dot{S}_r} \end{bmatrix} \end{bmatrix} \tag{A4}$$

It is then linearized to form the matrix $\mathbf{H}$ used in the update step of the extended Kalman filter (equations (14)–(15)) by taking partial derivatives around $\mathbf{x}_k$:

$$\mathbf{H}_k = \left.\frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_k} = \begin{bmatrix} \begin{bmatrix} 0 & 0 & \dfrac{-r}{\gamma^1_k \dot{S}_r} \end{bmatrix} & & 0 \\ & \ddots & \\ 0 & & \end{bmatrix} \tag{A5}$$

The elements of the $3N \times 3N$ state transition matrix $\mathbf{t}$ are formed from the rate-state evolution equation for γ (equation (6)):

$$\mathbf{t}(\mathbf{x}_k) = \begin{bmatrix} \begin{bmatrix} \dot{S}^1_{p_k} & 0 & 0 \\ 0 & \dot{S}^1_{A_k} & 0 \\ 0 & 0 & \ln\left(\gamma^1_k + \dfrac{dt}{A\sigma}\left(1 - \gamma^1_k\left(\dot{S}^1_{p_k} + \dot{S}^1_{A_k}\right)\right)\right) \end{bmatrix} & 0 & \cdots \\ 0 & \ddots & 0 \\ \vdots & 0 & \begin{bmatrix} \dot{S}^N_{p_k} & 0 & 0 \\ 0 & \dot{S}^N_{A_k} & 0 \\ 0 & 0 & \ln\left(\gamma^N_k + \dfrac{dt}{A\sigma}\left(1 - \gamma^N_k\left(\dot{S}^N_{p_k} + \dot{S}^N_{A_k}\right)\right)\right) \end{bmatrix} \end{bmatrix} \tag{A6}$$

It is then linearized to form the matrix $\mathbf{T}$ used in the prediction step of the extended Kalman filter (equation (13)) by taking partial derivatives with respect to the state variables at time k in each spatial cell:

$$\mathbf{T}_k = \left.\frac{\partial \mathbf{t}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_k} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \dfrac{-\dfrac{dt}{A\sigma}\gamma^1_k}{\dfrac{dt}{A\sigma} + \gamma^1_k\left(1 - \dfrac{dt}{A\sigma}\left(\dot{S}^1_{p_k} + \dot{S}^1_{A_k}\right)\right)} & \dfrac{-\dfrac{dt}{A\sigma}\gamma^1_k}{\dfrac{dt}{A\sigma} + \gamma^1_k\left(1 - \dfrac{dt}{A\sigma}\left(\dot{S}^1_{p_k} + \dot{S}^1_{A_k}\right)\right)} & \dfrac{\gamma^1_k\left(1 - \dfrac{dt}{A\sigma}\left(\dot{S}^1_{p_k} + \dot{S}^1_{A_k}\right)\right)}{\dfrac{dt}{A\sigma} + \gamma^1_k\left(1 - \dfrac{dt}{A\sigma}\left(\dot{S}^1_{p_k} + \dot{S}^1_{A_k}\right)\right)} \end{bmatrix} & 0 \\ 0 & \ddots \end{bmatrix} \tag{A7}$$

We assume the measurement and process noises ω_k and δ_{k+1} (equation (A1)) are Gaussian and described by the data and process covariance matrices $\mathbf{R}_k$ and $\mathbf{Q}_{k+1}$ respectively. The data covariance matrix $\mathbf{R}_k$ is assumed to be known only up to a scalar multiplier ξ^2, such that $\mathbf{R}_k = \xi^2\Sigma_k$, where Σ_k is a matrix that contains any known correlations between the spatial cells and components [*Segall and Matthews*, 1997]. In our case, we assume $\xi^2 \sim 1$ and estimate Σ_k a priori based on the variance of the data vector during stable time periods.

[57] The elements of the $3N \times 3N$ process covariance matrix $\mathbf{Q}_{k+1}$ are determined by the properties of the random walks used to model the state variables $\dot{S}_A$ and $\ln(\gamma)$:

$$\mathbf{Q}_{k+1} = \begin{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \tau^2\Delta t & 0 \\ 0 & 0 & \varepsilon^2\Delta t \end{bmatrix} & & 0 \\ & \ddots & \\ 0 & & \end{bmatrix} \tag{A8}$$

where τ and ε are the scale parameters for the random walks that model $\dot{S}_A$ and $\ln(\gamma)$ respectively.

[58] **Acknowledgments.** We thank Y. Ogata and J. Zhuang for help with the ETAS modeling, R. Lohman for providing the Obsidian Buttes slip model, and J. Lin for help with the Coulomb calculations. Comments from and discussions with R. Reves-Sohn, B. Hager, J. Lin, M. Cocco, B. Enescu, and an anonymous reviewer helped improve the manuscript. We are also grateful for an enlightening discussion about aftershock productivity with an anonymous reviewer and Associate Editor David Marsan. The earthquake catalog was obtained from the Southern California Earthquake Data Center. Map figures were prepared using the Generic Mapping Tools software freely distributed by *Wessel and Smith* [1998]. This work was supported by NSF EAR grant 0738641 and USGS NEHRP grant G10AP00004.

References

Anderson, B. D. O., and J. B. Moore (2005), *Optimal Filtering*, Dover, Mineola, N. Y.

Ben-Zion, Y., and V. Lyakhovsky (2006), Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology, *Geophys. J. Int.*, *165*, 197–210, doi:10.1111/j.1365-246X.2006.02878.x.

Blanpied, M. L., C. J. Marone, D. A. Lockner, J. D. Byerlee, and D. P. King (1998), Quantitative measure of the variation in fault rheology due to fluid–rock interactions, *J. Geophys. Res.*, *103*, 9691–9712, doi:10.1029/98JB00162.

Brune, J. N., and C. R. Allen (1967), A low-stress-drop, low-magnitude earthquake with surface faulting: The Imperial, California, earthquake of March 4, 1966, *Bull. Seismol. Soc. Am.*, *57*, 501–514.

Catalli, F., M. Cocco, R. Console, and L. Chiaraluce (2008), Modeling seismicity rate changes during the 1997 Umbria-Marche sequence (central Italy) through a rate- and state-dependent model, *J. Geophys. Res.*, *113*, B11301, doi:10.1029/2007JB005356.

Chester, F. M., and N. G. Higgs (1992), Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions, *J. Geophys. Res.*, *97*, 1859–1870, doi:10.1029/91JB02349.

Cocco, M., S. Hainzl, F. Catalli, B. Enescu, A. M. Lombardi, and J. Woessner (2010), Sensitivity study of forecasted aftershock seismicity based on Coulomb stress calculation and rate- and state-dependent frictional response, *J. Geophys. Res.*, *115*, B05307, doi:10.1029/2009JB006838.

Daley, D. J., and D. Vere-Jones (2002), *An Introduction to the Theory of Point Processes*, vol. 1, 2nd ed., Springer, New York.

Delahaye, E. J., J. Townend, M. E. Reyners, and G. Rogers (2009), Microseismicity but no tremor accompanying slow slip in the Hikurangi subduction zone, New Zealand, *Earth Planet. Sci. Lett.*, *277*, 21–28, doi:10.1016/j.epsl.2008.09.038.

Dieterich, J. (1994), A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, *99*, 2601–2618, doi:10.1029/93JB02581.

Dieterich, J., V. Cayol, and P. Okubo (2000), The use of earthquake rate changes as a stress meter at Kilauea volcano, *Nature*, *408*, 457–460, doi:10.1038/35044054.

Enescu, B., S. Hainzl, and Y. Ben-Zion (2009), Correlations of seismicity patterns in Southern California with surface heat flow data, *Bull. Seismol. Soc. Am.*, *99*, 3114–3123, doi:10.1785/0120080038.

Felzer, K. R., T. W. Becker, R. E. Abercrombie, G. Ekström, and J. R. Rice (2002), Triggering of the 1999 M_W 7.1 Hector Mine earthquake by aftershocks of the 1992 M_W 7.3 Landers earthquake, *J. Geophys. Res.*, *107*(B9), 2190, doi:10.1029/2001JB000911.

Fukuda, J., and K. M. Johnson (2008), A fully Bayesian inversion for spatial distribution of fault slip with objective smoothing, *Bull. Seismol. Soc. Am.*, *98*, 1128–1146, doi:10.1785/0120070194.

Gelb, A. (1974), *Applied Optimal Estimation*, MIT Press, Cambridge, Mass.

Hainzl, S., and Y. Ogata (2005), Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.*, *110*, B05S07, doi:10.1029/2004JB003247.

Hainzl, S., B. Enescu, M. Cocco, J. Woessner, F. Catalli, R. Wang, and F. Roth (2009), Aftershock modeling based on uncertain stress calculations, *J. Geophys. Res.*, *114*, B05309, doi:10.1029/2008JB006011.

Hardebeck, J. L., and E. Hauksson (1999), Role of fluids in faulting inferred from stress field signatures, *Science*, *285*, 236–239, doi:10.1126/science.285.5425.236.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge Univ. Press, New York.

Helmstetter, A. (2003), Is earthquake triggering driven by small earthquakes?, *Phys. Rev. Lett.*, *91*, 058501, doi:10.1103/PhysRevLett.91.058501.

Helmstetter, A., and B. E. Shaw (2006), Relation between stress heterogeneity and aftershock rate in the rate-and-state model, *J. Geophys. Res.*, *111*, B07304, doi:10.1029/2005JB004077.

Helmstetter, A., and D. Sornette (2002), Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, *J. Geophys. Res.*, *107*(B10), 2237, doi:10.1029/2001JB001580.

Helmstetter, A., and D. Sornette (2003a), Båth's law derived from the Gutenberg-Richter law and from aftershock properties, *Geophys. Res. Lett.*, *30*(20), 2069, doi:10.1029/2003GL018186.

Helmstetter, A., and D. Sornette (2003b), Importance of direct and indirect triggered seismicity in the ETAS model of seismicity, *Geophys. Res. Lett.*, *30*(11), 1576, doi:10.1029/2003GL017670.

Hill, D. P. (1977), A model for earthquake swarms, *J. Geophys. Res.*, *82*, 1347–1352, doi:10.1029/JB082i008p01347.

Hill, D. P., P. Mowinckel, and L. G. Peake (1975), Earthquakes, active faults and geothermal areas in the Imperial Valley, California, *Science*, *188*, 1306–1308, doi:10.1126/science.188.4195.1306.

Johnson, C. E., and D. M. Hadley (1976), Tectonic implications of the Brawley earthquake swarm, Imperial Valley, California, January 1975, *Bull. Seismol. Soc. Am.*, *66*, 1133–1144.

King, G. C. P., R. S. Stein, and J. Lin (1994), Static stress changes and the triggering of earthquakes, *Bull. Seismol. Soc. Am.*, *84*, 935–953.

Kisslinger, C., and L. M. Jones (1991), Properties of aftershock sequences in southern California, *J. Geophys. Res.*, *96*, 11,947–11,958, doi:10.1029/91JB01200.

Lin, J., and R. S. Stein (2004), Stress triggering in thrust and subduction earthquakes, and stress interaction between the southern San Andreas and nearby thrust and strike-slip faults, *J. Geophys. Res.*, *109*, B02303, doi:10.1029/2003JB002607.

Llenos, A. L., J. J. McGuire, and Y. Ogata (2009), Modeling seismic swarms triggered by aseismic transients, *Earth Planet. Sci. Lett.*, *281*, 59–69, doi:10.1016/j.epsl.2009.02.011.

Lohman, R. B., and J. J. McGuire (2007), Earthquake swarms driven by aseismic creep in the Salton Trough, California, *J. Geophys. Res.*, *112*, B04405, doi:10.1029/2006JB004596.

Lombardi, A. M., and W. Marzocchi (2007), Evidence of clustering and nonstationarity in the time distribution of large worldwide earthquakes, *J. Geophys. Res.*, *112*, B02303, doi:10.1029/2006JB004568.

Lombardi, A. M., W. Marzocchi, and J. Selva (2006), Exploring the evolution of a volcanic seismic swarm: The case of the 2000 Izu Islands swarm, *Geophys. Res. Lett.*, *33*, L07310, doi:10.1029/2005GL025157.

Lombardi, A. M., M. Cocco, and W. Marzocchi (2010), On the increase of background seismicity rate during the 1997–1998 Umbria-Marche, central Italy, sequence: Apparent variation or fluid-driven triggering?, *Bull. Seismol. Soc. Am.*, *100*, 1138–1152, doi:10.1785/0120090077.

Marsan, D. (2006), Can coseismic stress variability suppress seismicity shadows? Insights from a rate-and-state friction model, *J. Geophys. Res.*, *111*, B06305, doi:10.1029/2005JB004060.

McGuire, J. J., and P. Segall (2003), Imaging of aseismic fault slip transients recorded by dense geodetic networks, *Geophys. J. Int.*, *155*, 778–788, doi:10.1111/j.1365-246X.2003.02022.x.

McGuire, J. J., M. S. Boettcher, and T. H. Jordan (2005), Foreshock sequences and short-term earthquake predictability on East Pacific Rise transform faults, *Nature*, *434*, 457–461, doi:10.1038/nature03377.

Montgomery-Brown, E. K., P. Segall, and A. Miklius (2009), Kilauea slow slip events: Identification, source inversions, and relation to seismicity, *J. Geophys. Res.*, *114*, B00A03, doi:10.1029/2008JB006074.

Ogata, Y. (1988), Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, *83*, 9–27, doi:10.2307/2288914.

Ogata, Y. (1998), Space-time point process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, *50*, 379–402, doi:10.1023/A:1003403601725.

Ogata, Y. (2004), Space-time model for regional seismicity and detection of crustal stress changes, *J. Geophys. Res.*, *109*, B03308, doi:10.1029/2003JB002621.

Ogata, Y. (2005), Detection of anomalous seismicity as stress change sensor, *J. Geophys. Res.*, *110*, B05S06, doi:10.1029/2004JB003245.

Ogata, Y., and J. Zhuang (2006), Space-time ETAS models and an improved extension, *Tectonophysics*, *413*, 13–23, doi:10.1016/j.tecto.2005.10.016.

Omori, F. (1894), On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo*, *7*, 111–200.

Ozawa, S., H. Suito, and M. Tobita (2007), Occurrence of quasi-periodic slow-slip off the east coast of the Boso peninsula, central Japan, *Earth Planets Space*, *59*, 1241–1245.

Richter, C. F. (1958), *Elementary Seismology*, W. H. Freeman, San Francisco, Calif.

Roland, E., and J. J. McGuire (2009), Earthquake swarms on transform faults, *Geophys. J. Int.*, *178*, 1677–1690, doi:10.1111/j.1365-246X.2009.04214.x.

Rubin, A. M., and D. Gillard (2000), Aftershock asymmetry/rupture directivity among central San Andreas fault microearthquakes, *J. Geophys. Res.*, *105*, 19,095–19,109.

Segall, P., and M. Matthews (1997), Time dependent inversion of geodetic data, *J. Geophys. Res.*, *102*, 22,391–22,409, doi:10.1029/97JB01795.

Segall, P., E. K. Desmarais, D. Shelly, A. Miklius, and P. Cervelli (2006), Earthquakes triggered by silent slip events on Kilauea volcano, Hawaii, *Nature*, *442*, 71–74, doi:10.1038/nature04938.

Sornette, D., and M. J. Werner (2005), Apparent clustering and apparent background earthquakes biased by undetected seismicity, *J. Geophys. Res.*, *110*, B09303, doi:10.1029/2005JB003621.

Thatcher, W. (2001), Silent slip on the Cascadia subduction interface, *Science*, *292*, 1495–1496, doi:10.1126/science.1061770.

Toda, S., R. S. Stein, and T. Sagiya (2002), Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity, *Nature*, *419*, 58–61, doi:10.1038/nature00997.

Toda, S., R. S. Stein, K. Richards-Dinger, and S. Bozkurt (2005), Forecasting the evolution of seismicity in southern California: Animations built on earthquake stress transfer, *J. Geophys. Res.*, *110*, B05S16, doi:10.1029/2004JB003415.

Utsu, T. (1961), A statistical study on the occurrence of aftershocks, *Geophys. Mag.*, *30*, 521–605.

Wei, M., D. Sandwell, and Y. Fialko (2009), A silent M_w 4.7 slip event of October 2006 on the Superstition Hills fault, southern California, *J. Geophys. Res.*, *114*, B07402, doi:10.1029/2008JB006135.

Wells, D. L., and K. J. Coppersmith (1994), New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.*, *84*, 974–1002.

Werner, M. J. (2008), On the fluctuations of seismicity and uncertainties in earthquake catalogs: Implications and methods for hypothesis testing, Ph.D. thesis, 308 pp., Univ. of Calif., Los Angeles.

Werner, M. J., K. Ide, and D. Sornette (2011), Earthquake forecasting based on data assimilation: Sequential Monte Carlo methods for renewal point processes, *Nonlinear Processes Geophys.*, *18*, 49–70, doi:10.5194/npg-18-49-2011.

Wessel, P., and W. H. F. Smith (1998), New, improved version of generic mapping tools released, *Eos Trans. AGU*, *79*(47), 579, doi:10.1029/98EO00426.

Williams, P. L., and H. W. Magistrale (1989), Slip along the Superstition Hills fault associated with the 24 November 1987 Superstition Hills, California, earthquake, *Bull. Seismol. Soc. Am.*, *79*, 390–410.

Wolfe, C. J., B. A. Brooks, J. H. Foster, and P. G. Okubo (2007), Microearthquake streaks and seismicity triggered by slow earthquakes on the mobile south flank of Kilauea Volcano, Hawai'I, *Geophys. Res. Lett.*, *34*, L23306, doi:10.1029/2007GL031625.

Yang, W., and Y. Ben-Zion (2009), Observational analysis of correlations between aftershock productivities and regional conditions in the context of a damage rheology model, *Geophys. J. Int.*, *177*, 481–490, doi:10.1111/j.1365-246X.2009.04145.x.

Zhuang, J., C.-P. Chang, Y. Ogata, and Y.-I. Chen (2005), A study on the background and clustering seismicity in the Taiwan region by using point process models, *J. Geophys. Res.*, *110*, B05S18, doi:10.1029/2004JB003157.

A. L. Llenos, U.S. Geological Survey, 345 Middlefield Rd. #977, Menlo Park, CA 94025, USA. (allenos@usgs.gov)

J. J. McGuire, Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA.