

A fault and seismicity based composite simulation in northern California

M. B. Yıkılmaz¹, E. M. Heien², D. L. Turcotte¹, J. B. Rundle^{1,3}, and L. H. Kellogg¹

¹Department of Geology, University of California, Davis One Shields Avenue, Davis, CA 95616-8605, USA

²Computational Infrastructure for Geodynamics, University of California, Davis One Shields Avenue, Davis, CA 95616-8605, USA

³Department of Physics, University of California, Davis One Shields Avenue, Davis, CA 95616-8605, USA

Received: 21 September 2011 – Revised: 23 November 2011 – Accepted: 24 November 2011 – Published: 14 December 2011

Abstract. We generate synthetic catalogs of seismicity in northern California using a composite simulation. The basis of the simulation is the fault based “Virtual California” (VC) earthquake simulator. Back-slip velocities and mean recurrence intervals are specified on model strike-slip faults. A catalog of characteristic earthquakes is generated for a period of 100 000 yr. These earthquakes are predominantly in the range $M = 6$ to $M = 8$, but do not follow Gutenberg-Richter (GR) scaling at lower magnitudes. In order to model seismicity on unmapped faults we introduce background seismicity which occurs randomly in time with GR scaling and is spatially associated with the VC model faults. These earthquakes fill in the GR scaling down to $M = 4$ (the smallest earthquakes modeled). The rate of background seismicity is constrained by the observed rate of occurrence of $M > 4$ earthquakes in northern California. These earthquakes are then used to drive the BASS (branching aftershock sequence) model of aftershock occurrence. The BASS model is the self-similar limit of the ETAS (epidemic type aftershock sequence) model. Families of aftershocks are generated following each Virtual California and background main shock. In the simulations the rate of occurrence of aftershocks is essentially equal to the rate of occurrence of main shocks in the magnitude range $4 < M < 7$. We generate frequency-magnitude and recurrence interval statistics both regionally and fault specific. We compare our modeled rates of seismicity and spatial variability with observations.

1 Introduction

Probabilistic seismic hazard analysis (PSHA) is of importance in assessing risk, allocating resources for risk mitigation, and establishing the basis for earthquake insurance premiums and establishing licensing criteria for nuclear power plants. There are two divergent approaches to PSHA, the first is fault based and the second is seismicity based. In the first approach, faults are identified, magnitude and recurrence statistics for earthquakes on each fault are specified, and the risk of future earthquakes on these faults is estimated. Of particular importance in nuclear power-plant licensing is the concept of capable faults. A capable fault is defined to be a fault that has been active in the last 35 000 yr (Machette, 1999). The primary problem with this approach is that many earthquakes occur on faults that have not been recognized. This is certainly true for smaller earthquakes ($M < 5$), but is also true for a fraction of the larger earthquakes ($5 < M < 8$).

In the second approach to PSHA, past observed seismicity is extrapolated to provide an estimate of future seismic risk. This depends on the validity of the Gutenberg-Richter (GR) frequency-magnitude scaling. For instance, the rate of occurrence of $M < 3$ earthquakes in a region is extrapolated to determine the expected rate of occurrence of larger earthquakes. A problem with this approach is that many recognized faults where earthquakes are expected to occur are seismically quiet. Examples are segments of the San Andreas Fault in California.

In this paper we propose a PSHA approach that combines the fault based and seismicity based approaches. In the fault based analysis a set of regional faults are specified along with their mean slip rates and recurrence times.



Correspondence to: M. B. Yıkılmaz
(mbyikilmaz@ucdavis.edu)

A simulation model is used to generate synthetic catalogs of earthquakes on these faults. These faults tend to generate large earthquakes. In order to complete the specification of main shock seismicity we generate a set of background earthquakes that satisfy Gutenberg–Richter frequency-magnitude statistics, these earthquakes occur randomly in time and are spatially associated with the mapped faults. We then take both the fault based and the background earthquakes as parents which generate families of aftershocks. As a specific example we consider northern California and generate a synthetic 100 000 yr catalog of all earthquakes with $M \geq 4$ and compare the results with observed seismicity. Our results model both the spatial and temporal components of seismicity.

Examples of a fault based PSHA are the sequence of studies carried out by the Working Group on California Earthquake Probabilities (WGCEP). These studies have been summarized by Field (2007a). Statistical forecasts of the probability of rupture on specified fault segments were made. Mapped faults were considered and geologic and geodetic data were used to specify slip rates on each fault. Using historical earthquakes and paleoseismicity, a characteristic earthquake magnitude was assigned to each fault, resulting in mean earthquake recurrence intervals. A statistical distribution of recurrence times was assumed with a coefficient of variation near 0.5. A typical product is the probability of having an earthquake with magnitude greater than 6.5 on specified faults in 30 yr.

Examples of seismicity based PSHA are the sequence of regional likelihood models (RELM) that have been tested for California (Field, 2007b). In 2004, forecasts were solicited to predict the locations of $M > 5$ earthquakes in California during the five year 1 January 2006 to 31 December 2010. Some 16 forecasts were submitted and these will be evaluated during 2011. These forecasts were primarily based on the locations of small earthquakes prior to the test period. The direct extrapolations of this seismicity based on GR scaling is referred to as a “relative intensity” (RI) forecast (Holliday et al., 2005; Shcherbakov et al., 2010).

A number of fault-based simulations of seismicity have been carried out. A limited simulation model for earthquakes on specified strike-slip faults in California called Virtual California (VC) was given by Rundle (1988) and was subsequently updated (Rundle et al., 2002, 2004, 2006). Vertical fault segments were embedded in an elastic half space. VC is a 3-D simulation code that includes static interactions of stress using dislocation theory. Stress accumulation is by means of “backslip”, a linearly increasing negative displacement across each fault segment is applied at a prescribed velocity. A model earthquake occurs when the resulting stress exceeds a prescribed static coefficient of friction. Related models have been proposed by Ward (1992, 1996, 2000), by Robinson and Benites (1995, 1996) and by Dietrich and Richards-Dinger (2010).

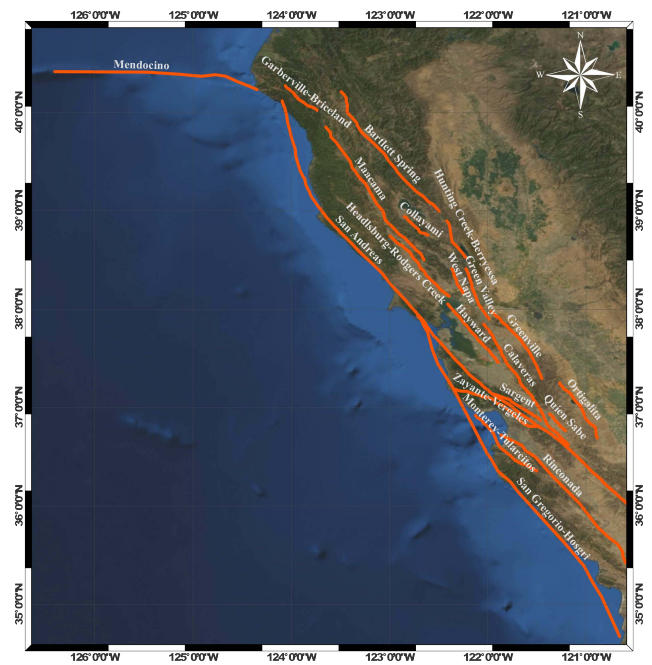


Fig. 1. Map of the strike-slip faults considered in our VC simulation of northern California seismicity. These are the major strike-slip faults that make up the San Andreas plate-boundary deformation in northern California.

In this paper our objective is to model seismicity with $M \geq 4$ in northern California and to compare the results with observations. The basis of this model is a fault based simulation using VC which includes major mapped strike-slip faults in the region. We select northern California for our initial simulation because the seismicity is dominated by strike-slip ruptures. To complete our composite model we will also include background seismicity and aftershocks. We will obtain a 100 000 yr catalog of earthquakes and will examine the statistics of this seismicity.

2 VC simulation

The primary basis for our composite simulation of the seismicity in northern California is fault based. We utilize the VC model to generate earthquakes on a specified set of strike-slip faults. These faults are illustrated in Fig. 1. The reasons for choosing northern California is that the zone of deformation between the Pacific plate and the Central Valley micro-plate is relatively narrow and is dominated by displacements on strike-slip faults. The faults have a depth of 12 km and we utilize 4 layers of $3 \text{ km} \times 3 \text{ km}$ elements in our cellular grid computation. There are a total of 3348 elements, 837 elements per layer (837×4). An element ID identifies each $3 \text{ km} \times 3 \text{ km}$ element; the numbering is from north to south. The association of element ID's with specific faults is shown in Fig. 2.

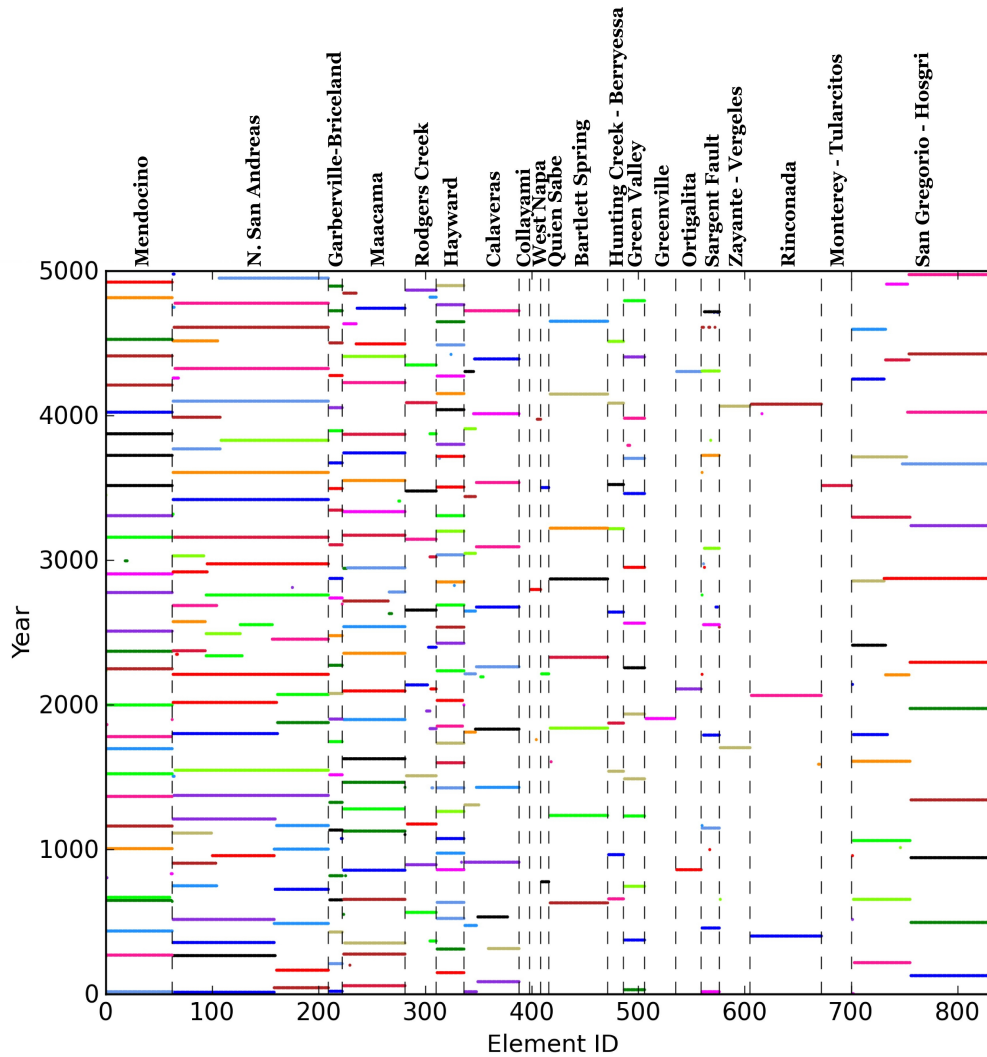


Fig. 2. A catalog of earthquakes that occurred during 5000 yr of our VC simulation of northern California seismicity. Elements on a fault are numbered from north to south and have a length of 3 km. Surface ruptures are shown as colored lines.

In order to constrain the behavior of the faults we prescribe the mean slip rate on each fault or fault segment and the mean recurrence interval between earthquakes on each fault or fault segment. The mean slip rates are based on geological and geodetic data and, when available, the mean recurrence times are based on paleoseismic data (California Department of Conservation, Division of Mines and Geology, 1996; Bryant and Lundberg, 2002; US Geological Survey, 2002; Wills, Weldon, and Bryant, 2008; Petersen, et al., 2008; Southern California Earthquake Data Center, 2010). When paleoseismic data are not available a “characteristic” earthquake magnitude is assumed and combined with the slip rate to obtain a recurrence interval. The assumed distributions of mean slip velocities and recurrence intervals are given in Fig. 3. Noise has been added to the recurrence intervals to simulate fault roughness. The mean interval times are used to calculate static coefficients of friction on the faults.

The VC simulation code imbeds the prescribed faults in an elastic half-space. A fault displacement on a fault segment produces stress changes on all other fault segments. These static interactions are obtained using dislocation theory. Stress accumulation on a fault is by means of backslip. A linearly increasing displacement is applied to the fault at the prescribed slip rate. A model earthquake occurs on the fault when the shear stress difference across the fault exceeds the prescribed static coefficient of friction on the fault. When a fault segment fails the stress on the fault is reduced to the dynamic coefficient of friction. The advantage of the backslip approach is that it produces cyclic variations of fault displacements. If displacements on faults accumulate, faults grow and die and geometric incompatibilities develop.

Alternative friction laws are available for fault behavior. These include (1) static-dynamic, (2) velocity/slip weakening, and (3) rate and state. In the simulations reported here

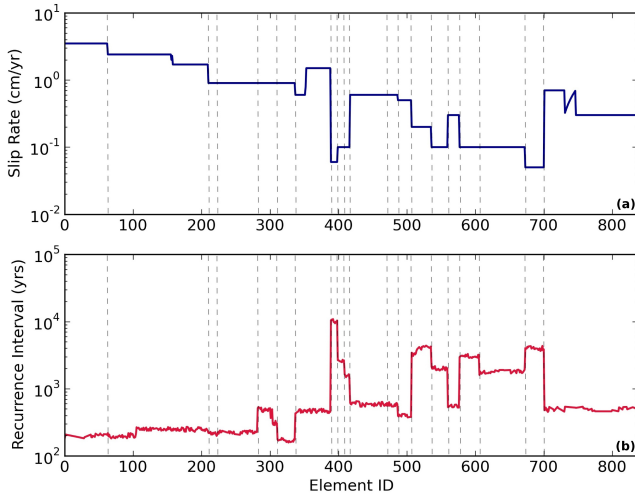


Fig. 3. (a) Prescribed slip rates on each element. (b) Prescribed recurrence intervals on each element.

the static-dynamic law is used. Because our model is quasi-static, an approximate model must be used to treat rupture propagation. Rupture propagation is enhanced by the stress singularity at a crack tip. Fully dynamic simulations utilize a detailed treatment of the stresses at the vicinity of the propagating fault rupture. With our relatively large, $3 \text{ km} \times 3 \text{ km}$ elements, the stress transfer to adjacent elements is unrealistically low. In order to model the rupture tip stress singularity we introduce a dynamic triggering factor F . Rupture is initiated on an element on a fault, the stress on another element is σ before any stress is transferred and the static failure stress on that element is σ_s . The dynamic triggering factor is defined by

$$F = \frac{\sigma_R - \sigma}{\sigma_S - \sigma} \quad (1)$$

where $\sigma \leq \sigma_R \leq \sigma_S$ is the reduced stress on other elements when a rupture has been initiated. When $F = 1$ we have $\sigma_R = \sigma_S$ and there is no enhanced failure. When $F = 0$ any increase in stress on the element due to stress transfer causes the element to rupture.

If the dynamic triggering factor is small, many short ruptures occur on the prescribed faults. In this limit it is not possible to generate large earthquakes such as the 1906 San Francisco earthquake on the northern San Andreas Fault. If the dynamic triggering factor is near unity all earthquakes rupture the entire fault on which they initiate.

In order to constrain the value of the dynamic triggering factor we utilize the record of past earthquakes obtained using paleoseismic studies. Unfortunately paleoseismic studies are not available to provide a synthesis of the earthquake history on the northern San Andreas Fault. However, Biasi and Weldon (2009) have used a series of paleoseismic studies on the southern San Andreas Fault. In their Fig. 15 they

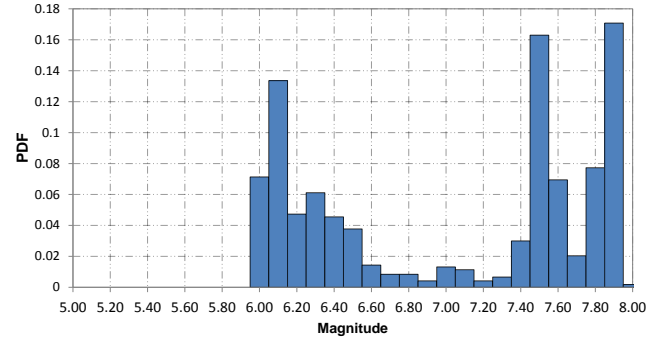


Fig. 4. Probability distribution function (PDF) of earthquake magnitudes on the northern San Andreas Fault during the 100 000 yr VC simulation. Bins are 0.1 magnitude units.

give two scenarios for the great earthquakes on the 550 km length of fault during the last 1600 yr. These scenarios give a mixture of long and intermediate length ruptures. In order that our simulations produce these types of failures we take $F = 0.5$. With $F = 0.3$ almost every earthquake ruptures the entire San Andreas Fault. With $F = 0.7$ short ruptures dominate. We also introduce noise in our calculations. The stress drop $\sigma_{st} - \sigma_d$ is randomly increased/decreased by a factor of 0.125 on each element. The influence of this noise is similar to the influence of the stress intensity factor. With no noise, long ruptures become dominant. With the random factor at 0.25 short ruptures tend to dominate. The noise is a proxy for fault roughness. Large noise simulates rough faults. Small noise simulates smooth faults and ruptures propagate freely. Details of parameter selection and other aspects of the VC simulation code can be found in Rundle et al. (2006).

We have carried out our simulation of seismic activity on the specified faults for 110 000 yr, the first 10 000 yr were discarded to remove the initial transient. A typical 5000 yr record of simulated seismicity is illustrated in Fig. 2. The behavior is dominated by characteristic (full fault) ruptures but there are examples of smaller ruptures, for example, on the northern San Andreas Fault. This behavior is quite similar to that given by Biasi and Weldon (2009) for the southern San Andreas Fault based on paleoseismic studies.

To illustrate the statistical nature of the simulated seismicity we consider two faults, the northern San Andreas (elements 63 to 210) and the Hayward (elements 311 to 337). The probability distribution function (PDF) of earthquake magnitudes on the San Andreas Fault for the full 100 000 yr simulation is given in Fig. 4. Two peaks stand out, the first in the magnitude range $7.4 < M < 7.9$ and the other in the magnitude range $6 < M < 6.5$. The associated variability of rupture lengths can be seen in Fig. 2. The PDF of earthquake magnitudes on the Hayward Fault for the 100 000 yr simulation is given in Fig. 5. This distribution is dominated by characteristic earthquakes with $M \approx 7.0$. The PDF of recurrence intervals of earthquakes on the northern San

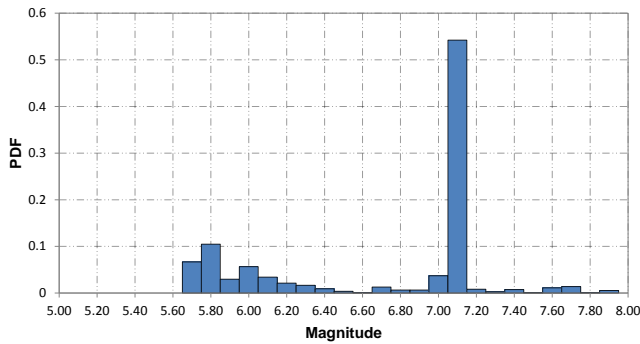


Fig. 5. Probability distribution function (PDF) of earthquake magnitudes on the Hayward fault during the 100 000 yr VC simulation. Bins are 0.1 magnitude units.

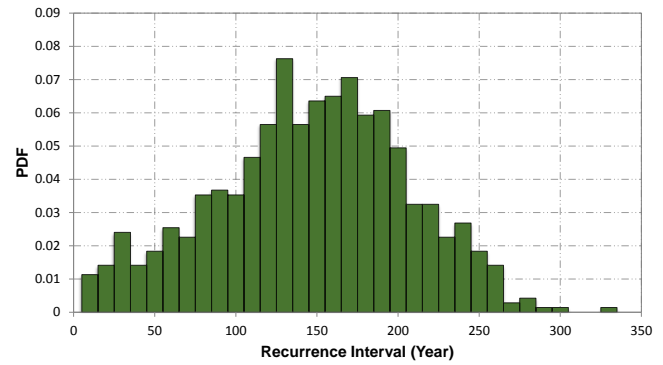


Fig. 7. Probability distribution function (PDF) of recurrence intervals for earthquakes with $M > 6.5$ on the Hayward Fault during the 100 000 yr simulation. The mean recurrence time is 141 yr with $C_v = 0.42$. The bin width is 10 yr.

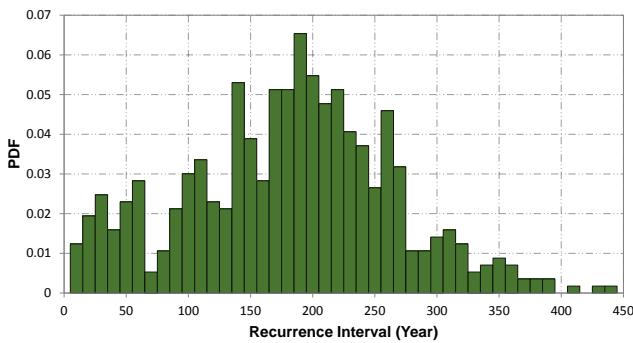


Fig. 6. Probability distribution function (PDF) of recurrence intervals for earthquakes with $M > 7.0$ on the northern San Andreas Fault during the 100 000 yr simulation. The mean recurrence time is 176 yr with $C_v = 0.48$. The bin width is 10 yr.

Andreas Fault with $M > 7.0$ is given in Fig. 6. The mean recurrence interval is 176 yr with coefficient of variation $C_v = 0.48$. The PDF of recurrence intervals of earthquakes on the Hayward fault with $M > 6.5$ is given in Fig. 7. The mean recurrence interval is 141 yr with coefficient of variation $C_v = 0.42$. Although there is little variability of magnitudes on the Hayward Fault, there is considerable variability of recurrence intervals. This is quite similar to the behavior of characteristic earthquakes on the Parkfield segment of the San Andreas Fault (Bakun et al., 2005). These earthquakes occurred in 1857, 1881, 1901, 1922, 1934, 1966 and 2004. The mean recurrence interval and coefficient of variation of these earthquakes are $\mu = 24.5$ yr and $C_v = 0.378$. It is well documented that the magnitude of the last four earthquakes in the sequence were very close to 6 and it is likely that the earlier ones also were. Recently, Lienkaemper et al. (2010) studied 12 historical earthquakes on the southern Hayward Fault that occurred in the past 1900 yr and found a mean recurrence interval of 161 ± 10 yr and a $C_v = 0.2$ which is considerably lower than our value ($C_v = 0.42$) and other values deduced from observations on paleoseismic results.

The cumulative frequency-magnitude statistics of all earthquakes during our 100 000 yr simulation are given in Fig. 8. The statistics are dominated by the large earthquakes in the $6 < M < 8$ range. However there is a strong deviation from GR statistics even over this restricted range. It should be noted that single element ruptures, $3 \text{ km} \times 3 \text{ km}$, could generate $M \approx 5$ earthquakes. We see very few small earthquakes in the range $5 < M < 6$ relative to GR statistics.

3 Background seismicity

The GR statistics of the VC earthquakes given in Fig. 8 indicate a relatively complete catalog of $M > 7$ events. However, the VC simulation does not generate the required GR statistics for smaller earthquakes. Our objective is to produce a complete catalog of earthquakes in northern California with magnitudes greater than four. We conclude that most main shocks in the magnitude range $4 < M < 7$ occur on faults not included in the VC simulation. In order to complete our catalog of main shocks in northern California we carry out a statistical simulation of background seismicity. To simulate these earthquakes we could use a version of the RI method described above. However, this would imply that recent seismicity (in the last 40 yr) is representative of the long term (100 000 yr) activity. Instead of making this assumption we utilize an alternative statistical model. We specify the background seismicity as follows:

(1) We assume that the earthquakes occur randomly in time (Wang et al., 2010). The time intervals t between background earthquakes for this Poisson process are selected randomly from an exponential. The cumulative distribution function (cdf) P_{ct} is given by

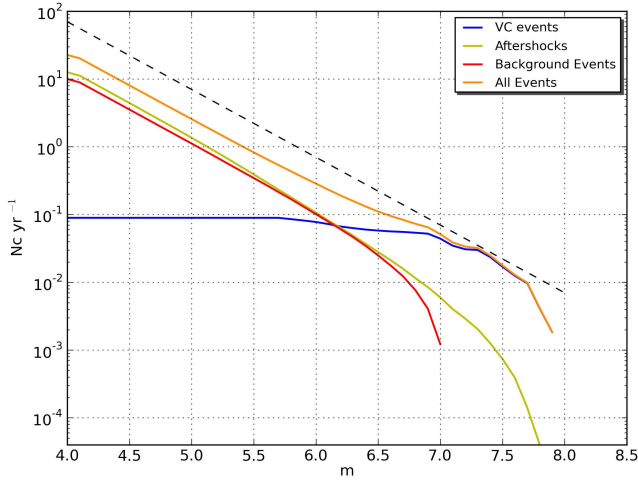


Fig. 8. Cumulative frequency-magnitude statistics for simulated earthquakes in northern CA with magnitudes greater than 4. The cumulative numbers of earthquakes per year N_c with magnitudes equal to or greater than m are given as a function of M . Included are: (1) Earthquakes on the VC faults. These earthquakes are predominantly in the magnitude range 7 to 8, the mean recurrence interval is about 10 yr. (2) Other background events, they have the prescribed GR magnitude statistics with a mean recurrence interval of about 0.1 yr. (3) Aftershocks generated by the BASS model. The mean recurrence interval is also about 0.1 yr. The distribution for all earthquakes is also shown. On average there are about 20 $M > 4$ earthquakes per year. The dashed line is the GR distribution, Eq. (2), with $b = 1.0$.

$$P_{ct} = 1 - \exp\left(\frac{-t}{\tau}\right) \quad (2)$$

where the time constant τ is the mean waiting time between earthquakes and is a measure of the seismic intensity. We select random numbers in the interval $0 < P_{ct} < 1$ and the corresponding time intervals t are obtained from Eq. (2). We consider earthquakes in the range $4 \leq M \leq 7$ and in order to match the recent rate of seismicity in northern California we take $\tau = 1$ month in Eq. (2).

(2) We require that this random seismicity satisfies Gutenberg-Richter (GR) frequency-magnitude statistics. The cumulative probability that an earthquake has a magnitude greater than M , P_{cm} is given by

$$P_{cm} = 10^{-b(M-M_{\min})} \quad (3)$$

where the b value is generally in the range $0.8 < b < 1.2$ and M_{\min} is the minimum magnitude considered. We again select random numbers in the interval $0 < P_{cm} < 1$ and the corresponding magnitudes are obtained from Eq. (3). We assume $b = 1$ and, since we are interested in generating seismicity with $M > 4$, we take $M_{\min} = 4$. Since we postulate that main shocks with $M > 7$ are well represented by the VC synthetic catalog, we discard selected random magnitudes with $M > 7$.

(3) We must also specify the location of each background earthquake. Clearly it is not appropriate to assume that these

earthquakes are randomly distributed over northern California. We hypothesize that the random earthquakes are clustered near the VC faults illustrated on Fig. 1. We pick a random point on a VC fault. The distance r from this point to the point where the background earthquake occurs is assumed to satisfy the cdf P_{cr} given by

$$P_{cr} = \frac{1}{\left(1 + \frac{r}{d_b}\right)^{q-1}} \quad (4)$$

where d_b is a characteristic distance to the clustered earthquakes. We again select random numbers in the interval $0 < P_{cr} < 1$ and the corresponding radial distances are obtained from Eq. (4). The justification for the use of Eq. (4) will be given when we consider aftershocks. The direction from the point on the VC fault is picked randomly. For our characteristic length we take $d_{bg} = 4$ km. This gives a mean distance from our selected point of $r = 25$ km. Based on the studies of aftershocks given by Felzer and Brodsky (2006) we take $q = 1.35$. The direction of the main shock from the point on the VC fault is picked randomly.

The frequency-magnitude statistics of the background main shocks are given in Fig. 8. Before discussing this result we will complete our composite simulation by including aftershocks.

4 Aftershocks

Regional seismicity is made up of foreshocks and aftershocks as well as main shocks. We have simulated main shocks using a combination of VC earthquakes as well as statistical background earthquakes. We now turn our attention to foreshock, main shock, aftershock simulations. The epidemic type aftershock sequence (ETAS) model has been widely used to treat aftershock statistics. The model was developed by Ogata and colleagues (Ogata 1998, 1999; Ogata and Zhuang, 2006) and by Helmstetter, Sornette and colleagues (Helmstetter and Sornette 2002, 2003a, b; Sornette and Werner, 2005). The ETAS model utilizes the concept of parent and daughter earthquakes. The first earthquake in a sequence is a parent earthquake and it generates a sequence of daughter earthquakes. Each of these daughter earthquakes becomes a parent earthquake that produces second-generation daughter earthquakes. Further generations are considered until the sequence dies out. One or more of the generated earthquakes may be larger than the original parent earthquake. In this case the original parent earthquake is a foreshock and the largest subsequent earthquake is the main shock. In the ETAS model the mean number of daughter earthquakes generated by a parent earthquake is given by a productivity relation. The magnitude of a daughter earthquake is determined randomly from a Gutenberg-Richter distribution, the time of occurrence is determined randomly using the modified form of Omori's law and the location is determined randomly from an Omori like relation.

In this paper we will use the branching aftershock sequence (BASS) model. This model was introduced by Turcotte et al. (2007) and Holliday et al. (2008a, b) and has also been used by Lennartz et al. (2008, 2011). The reason we choose the BASS model over the ETAS model is that the BASS model is self-similar. That is, it is not sensitive to the choice of M_{\min} . The BASS model is the self-similar limit of the ETAS model. Because of the self-similar property the statistics of $M > 4$ aftershocks are identical to $M > 3$ or $M > 1$, this is not the case for ETAS simulations. In the BASS model the productivity relation in ETAS is replaced by the modified form of Båth's law (Shcherbakov and Turcotte, 2004a). Each VC and background earthquake is used as a parent earthquake for a BASS simulation of aftershocks. We require that the frequency-magnitude distribution for each sequence of aftershocks satisfies a Gutenberg-Richter frequency-magnitude relation:

$$\log_{10}[N_d(\geq M_d)] = a - bM_d \quad (5)$$

where M_d is the magnitude of a daughter earthquake, $N_d(\geq M_d)$ is the number of daughter earthquakes with magnitudes greater than or equal to M_d and a and b are constants.

We further require that a modified form of Båth's law is applicable for each generation of aftershocks. Båth's law (Båth, 1965) states that the difference in magnitude ΔM between a mainshock with magnitude M_{ms} and its largest aftershock with magnitude M_{asmax} , $\Delta M = M_{\text{ms}} - M_{\text{asmax}}$, is approximately constant independent of the main-shock magnitude, typically $\Delta M \approx 1.2$. The modified form of Båth's law (Shcherbakov and Turcotte, 2004a) is based on an extrapolation of the GR scaling for aftershocks. The magnitude of the "largest" aftershock consistent with GR scaling for aftershocks is obtained by setting;

$$N_d(\geq (M_p - \Delta M^*)) = 1 \quad (6)$$

in Eq. (5). When this is done we obtain

$$a = b(M_p - \Delta M^*) \quad (7)$$

Using this result, the GR relation for aftershocks is given by

$$\log_{10}[N_d(\geq M_d)] = b(M_p - \Delta M^* - M_d) \quad (8)$$

In order to terminate the sequence of aftershocks, it is necessary to specify a minimum magnitude earthquake M_{\min} in the sequence. From Eq. (8), the total number of daughter earthquakes N_{dT} is given by

$$N_{dT} = N(\geq M_{\min}) = 10^{b(M_p - \Delta M^* - M_{\min})} \quad (9)$$

Following Turcotte et al. (2007) we take $b = 1$ and $\Delta M^* = 1.2$. Since we restrict our simulation to earthquakes with $M \geq 4$ we take $M_{\min} = 4$. From Eqs. (8) and (9) we obtain the cumulative distribution function P_{Cm} for the magnitudes of the daughter earthquakes

$$P_{Cm} = \frac{N_d(\geq M_d)}{N_{dT}} = 10^{-b(M_d - M_{\min})} \quad (10)$$

For each daughter earthquake a random value $0 < P_{Cm} < 1$ is generated, and the magnitude of the earthquake is determined from Eq. (9), using the parameter values given above. To stabilize our solution we neglect foreshocks and require $M_d \leq M_p$.

We require that the time delay t_d until each daughter earthquake after the parent earthquake satisfies a rewritten form of Omori's law (Shcherbakov et al., 2004b)

$$R(t_d) = \frac{dN_d}{dt} = \frac{1}{\tau(1 + \frac{t_d}{c})^p} \quad (11)$$

where $R(t_d)$ is the rate of aftershock occurrence and τ , c , and p are parameters. The number of daughter aftershocks that occur after a time t_d is given by

$$N_d(\geq t_d) = \int_{t_d}^{\infty} \frac{dN_d}{dt} dt' = \frac{c}{\tau(p-1)(1 + \frac{t_d}{c})^{p-1}} \quad (12)$$

The total number of daughter earthquakes N_{dT} is obtained by setting $t_d = 0$ in Eq. (11) with the result

$$N_{dT} = \int_0^{\infty} \frac{dN_d}{dt} dt = \frac{c}{\tau(p-1)} \quad (13)$$

From Eqs. (12) and (13) we obtain the cumulative distribution function P_{Ct} for the times of occurrence of the daughter earthquakes:

$$P_{Ct} = \frac{N_d(\geq t_d)}{N_{dT}} = \frac{1}{(1 + \frac{t_d}{c})^{p-1}} \quad (14)$$

For each daughter earthquake a random value $0 < P_{Ct} < 1$ is generated and the time of occurrence of the earthquake is determined from Eq. (14). Based on the study of California earthquakes by Shcherbakov et al. (2004b) we take $c = 0.1$ days and $p = 1.25$.

We utilize a spatial form Omori's law to specify the location of each daughter earthquake (Helmstetter and Sornett, 2003b; Felzer and Brodsky, 2006). The cumulative distribution function P_{Cr} for the radial distance r_d of each daughter earthquake from the parent earthquake is given by

$$P_{Cr} = \frac{N_d(\geq r_d)}{N_{dT}} = \frac{1}{\left(1 + \left[\frac{r}{d_a 10^{0.5 M_p}}\right]\right)^{q-1}} \quad (15)$$

where d_a and q are parameters. For each daughter earthquake a random value $0 < P_{Cr} < 1$ is generated and the radial distance r to the aftershock is determined from Eq. (14). The origin of the distance is randomly selected from the parent earthquake rupture zone for the first generation of aftershocks; the direction from their origin is also selected randomly. In our simulations we take $d_a = 40$ m and $q = 1.35$.

Each first-generation daughter earthquake becomes a parent for the second-generation daughter earthquake. The procedure given above is used to determine the magnitudes,

times of occurrences, and locations of the second-order daughter earthquakes relative to their parents. This procedure is continued until the sequence of aftershocks with magnitudes greater than M_{\min} dies out.

The frequency-magnitude statistics of the aftershocks during our 100 000 yr simulation are given in Fig. 8. It is interesting to note that the number of aftershocks of both VC and background main shocks are approximately equal to the number of background main shocks.

5 Composite seismicity

The frequency-magnitude statistics of our 100 000 yr simulation of northern California are also given in Fig. 8. Earthquakes in the magnitude range $7 < M < 8$ are dominated by VC events. Earthquakes in the magnitude range $4 < M < 6.5$ are dominated by background main shocks and aftershocks in near equal parts. It is not surprising that both background seismicity and aftershocks are well approximated by GR frequency-magnitude statistics since the magnitudes were selected from GR distributions. An important question is whether the VC earthquakes are consistent with the extrapolation of GR statistics for smaller events. It is seen in Fig. 8 that the smaller events occur about a factor of two less frequently than that predicted by an extrapolation of the VC data. The level of the VC earthquakes is prescribed by the slip rates and magnitudes on the VC faults. This level is near one $M \geq 7$ earthquake each 12 yr and one $M \geq 7.5$ earthquake each 50 yr. These values may be a bit high. They could be reduced if more fault segmentation was introduced by reducing the dynamic triggering effect. More, small and fewer large earthquakes would occur on the VC faults.

The number of smaller earthquakes could be increased by decreasing the value of τ in Eq. (1). We have taken $\tau = 1$ month (12 $M \geq 4$ earthquakes per year on average) based on seismicity in the last 50 yr. It is quite possible that this level is a factor of two less than the long term average. It is clear that we could adjust the parameters F and τ to obtain a near continuous GR distribution but is not clear which adjustment is justified.

As a measure of seismic activity in northern California we consider earthquakes with magnitudes $M > 4$. The numbers of these earthquakes per year for the period 1970 to 2009 are given in Fig. 9a. For comparison we give the number of $M > 4$ earthquakes for a typical 40 yr period in our composite simulation. These are given in Fig. 9b. It would appear desirable to make a quantitative comparison of these two data sets. The mean and standard deviations of the magnitude distributions could be given. However, these statistics would be misleading because of the large numbers of aftershocks associated with the larger VC and actual earthquakes. The mainshock-aftershock sequences are similar. However, the actual background seismicity appears to be more variable than the simulated seismicity. This difference may be

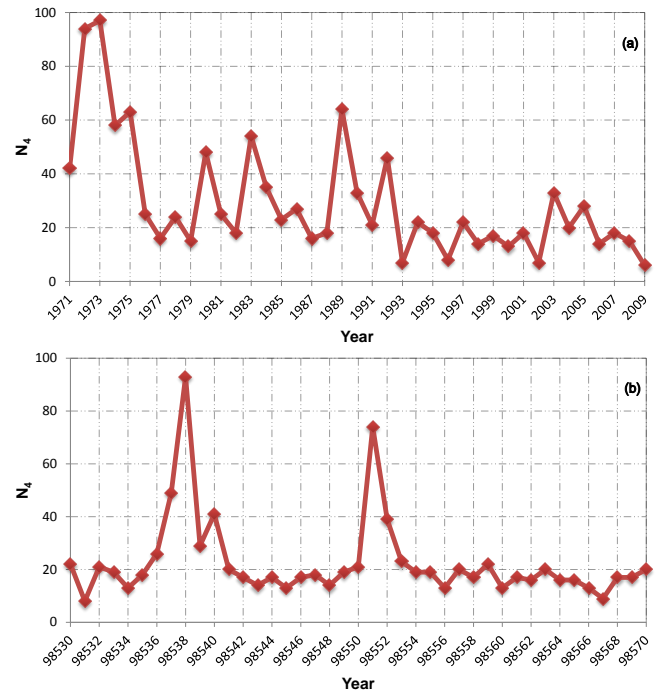


Fig. 9. (a) Numbers of $M \geq 4$ earthquakes per year N_4 in northern California for the period 1970 to 2009. (b) Numbers of $M \geq 4$ earthquakes per year N_4 for a typical 40 yr period obtained from our composite simulation.

due to long-term correlations in the background seismicity as discussed by Lennartz et al. (2011), which can lead to the clustering of large earthquake. An example would be the occurrence of swarms of earthquakes in the magnitude range $4 < M < 6$. In order to address this question more directly we consider the spatial distributions of actual and simulated earthquakes.

The locations of the northern California earthquakes with $M > 4$ for the period 1970 to 2009 are given in Fig. 10. No attempt has been made to separate aftershocks from main shocks. Comparing Fig. 10 with Fig. 1 there is a clear association of the seismic activity during this relatively short period with the major strike-slip faults considered in our VC simulation. In Fig. 11 the locations of our simulated seismicity in northern California during a 40 yr period are given. This period was selected to include a VC earthquake in order to illustrate the relative roles of VC, background, and aftershock seismicity. The $M = 7.79$ VC earthquake on the San Andreas Fault is shown as the solid black line. The background earthquakes during this period are shown in green and the aftershocks in red. It is interesting to compare the simulation seismicity given in Fig. 11 with the actual seismicity given in Fig. 10. In both cases there is a strong correlation of the seismicity with the VC strike-slip faults given in Fig. 1. However, there are also significant differences. The seismicity in Fig. 10 on and adjacent to the Mendocino fault is both

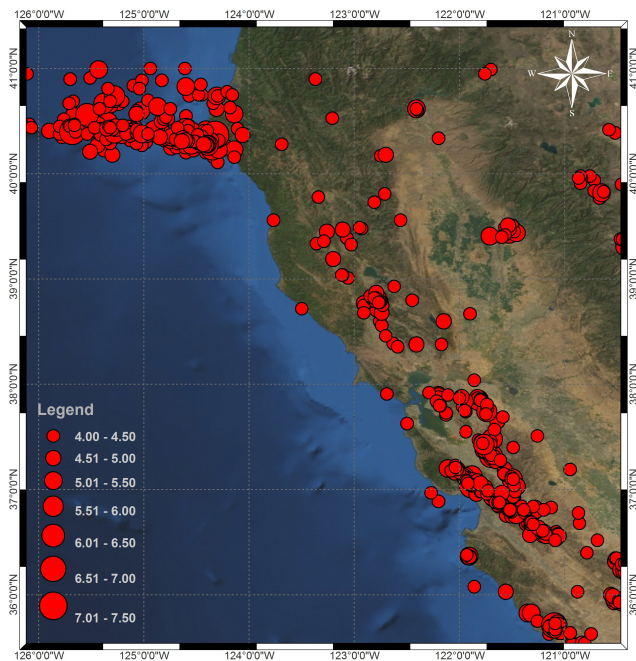


Fig. 10. Locations of the northern California earthquakes with $M \geq 4$ for the period 1970 to 2009.

very strong and diffuse. This seismicity is characterized by swarms of earthquakes in the range $4 \leq M \leq 6$. The temporal variability of this seismicity is responsible for the high fluctuation level seen in Fig. 9a. These swarms can be attributed to the behavior of this younger oceanic transform fault relative to the older continental strike-slip faults elsewhere in the simulation. One possible way to introduce this behavior into our simulations would be to prescribe a lower maximum magnitude earthquake in this region.

Another striking difference between the simulated and actual seismicity is the lack of earthquakes on the northern San Andreas Fault between Cape Mendocino and Redwood City (ruptured in 1906). To reproduce this behavior we would have to suppress our background activity on this fault. A justification for this is not obvious.

6 Discussion

In this paper we have presented a composite model for the simulation of regional seismicity. The model includes fault based seismicity using the VC earthquake simulator. Also included are random background seismicity and aftershocks. The model has been applied to northern California. Earthquakes on the major strike-slip faults are generated using the VC simulation; a typical catalog of seismicity has been given in Fig. 2. However, many earthquakes in northern California and elsewhere do not occur on mapped faults. We include these earthquakes using a random generator for background

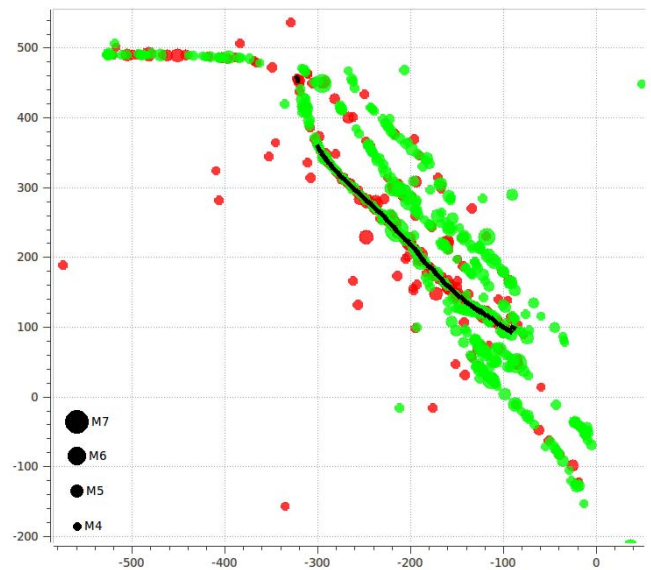


Fig. 11. Locations of our simulated northern California seismicity with $M \geq 4$ during a 40 yr period. The period selected included a $M = 7.79$ VC earthquake on the northern San Andreas Fault. The rupture during this simulated earthquake is shown by the solid black line. Aftershocks during the 40 yr are shown in red and background seismicity in green.

seismicity. We also generate aftershocks using a random simulation based on the BASS model. Using these three components we obtain simulated earthquake catalogs for northern California.

One output of our simulations is the recurrence statistics of earthquakes on a specified fault, or a specified point on a fault. A detailed study of these statistics using VC from some faults in California has been given by Yakovlev et al. (2006). It was argued that the Weibull distribution was preferred for recurrence statistics. Typical coefficients of variation were in the range $0.3 \leq C_v \leq 0.5$, generally in agreement with published values. Examples of these statistics based on the VC simulation in this paper are given in Fig. 6 for the northern San Andreas Fault and in Fig. 7 for the Hayward Fault. A statistical output unique to our composite model is the relative roles of main-shocks versus aftershocks. As illustrated in Fig. 8 about 50% of our seismicity with $M \geq 4$ can be attributed to aftershocks. This is certainly a reasonable value but it is difficult to compare it with observations. Observed seismicity has a strong temporal dependence because of the aftershock association with the largest earthquakes. Also, distinguishing aftershocks from main-shocks is extremely difficult.

A major question concerning the behavior of earthquakes on a fault is the distribution of earthquake magnitudes on the fault. Two alternative limiting cases are the Gutenberg-Richter (GR) and the characteristic earthquake (CE) hypotheses. It is accepted that small earthquakes on or near

a fault obey GR scaling. In the GR hypothesis, all earthquakes on or near a fault satisfy GR scaling. In the competing CE hypothesis, the large “characteristic” earthquakes lie above GR scaling. They occur more frequently than predicted by the GR distribution of small events in the region. Arguments favoring the CE hypothesis have been given by Wesnousky (1994), Hofmann (1996), Ishibe and Shimazaki (2009) and others. Arguments favoring the GR hypothesis have been made by Parsons and Geist (2009) and Naylor et al. (2009) among others.

A principal focus of this paper is to consider this question. As discussed above we can control rupture propagation on a fault by varying the dynamic triggering factor F . If F is small the required stress for a rupture to propagate is small. Taking $F = 0.1$ we find only characteristic earthquakes. Once a rupture is initiated it propagates across the entire fault. If F approaches unity, only limited rupture propagation occurs. Small earthquakes dominate and a near GR distribution of magnitudes is found. However, large earthquakes that rupture a large fraction of a fault do not occur. To some extent small F simulates a smooth fault and F near unity simulates a rough fault.

In the simulation given in Fig. 2 and discussed in some detail in this paper we have chosen $F = 0.5$. The reason for this choice is that the behavior of our simulated ruptures on the northern San Andreas Fault is similar to the paleoseismic synthesis of ruptures on the southern San Andreas Fault given by Biasi and Weldon (2009). It would be desirable to quantify this qualitative similarity. There are a number of difficulties in doing this. First, the data from the paleoseismic studies are quite limited so quantification is not really possible. Also, the details of the fault segmentation in our simulation are sensitive to the spatial distribution of fault strengths along the fault. Variability of fault strengths increases segmentation with the spectrum of variability influencing the frequency-magnitude statistics of the earthquakes generated.

Another question that can be addressed by simulation is the role of fault jumps. In a fault jump a rupture jumps from one fault to another. The Landers earthquake was an example of a rupture jumping so that the rupture spread over several faults that were thought to rupture separately (Johnson et al., 1994). In terms of northern California one question is whether a rupture could jump from the Hayward Fault to the Rodgers Creek Fault. In the simulations given here fault jumps do not occur. However, in a previous paper (Yıkılmaz et al., 2010) we showed that fault jumps can be included in VC simulations. A model fault must bridge the gap. If the mean recurrence interval on the bridging fault is large, it will act as an asperity. Our simulation results were in good agreement with the sequence of observed subduction zone earthquakes at the Nankai Trough, Japan.

A number of modifications of our VC simulation are possible:

- Replace backslip with an evolving driven plate model. This may be possible, but the accumulating displacements on faults cause serious problems for long term simulations. As the total displacement on a fault increases, so does the fault’s length (Marrett and Allmendinger, 1991). Faults must grow and die. To avoid overlaps and holes, a partial continuum rheology (damage, plastic or fluid) appears to be necessary.
- Replace the half-space elasticity with a plate over a viscous substrate. This can certainly be done and would introduce a relaxation process. However, continuous viscous relaxation is computationally intensive.
- Replace the static-dynamic friction with another friction law such as rate and state friction. Again this can be done but generally increases computational requirements.
- Our VC simulation consisted of 3348 3 km by 3 km elements. These computations were carried out on a desktop computer using a single CPU. We could reasonably extend the VC model to 1 km \times 1 km using high performance computers but this is about the current resolution limit.

Simulations of regional seismicity are clearly at an early stage of development. The goal is to create ensemble forecasts of earthquake activity similar to ensemble forecasts of weather. One approach is to search a simulation for a period of seismic activity similar to recently observed seismicity. The subsequent activity in the simulation would then be a forecast of future seismic activity.

Acknowledgements. This work was funded by NSF grant # 0753407. The contributions of J. B. Rundle have been partially supported by NASA grant # NNX08AF69G.

Edited by: L. Telesca

Reviewed by: Q. Wang and two other anonymous referees

References

- Bakun, W. H., Aagaard, B., Dost, B., Ellsworth, W. L., Hardebeck, J. L., Harris, R. A., Ji, C., Johnston, M. J. S., Langbein, J., Lienkaemper, J. J., Michael, A. J., Murray, J. R., Nadeau, R. M., Reasenber, P. A., Reichle, M. S., Roeloffs, E. A., Shakal, A., Simpson, R. W., and Waldhauser, F.: Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, *Nature*, 437, 969–974, 2005.
- Bath, M.: Lateral inhomogeneities of upper mantle, *Tectonophysics*, 2, 483–514, 1965.
- Biasi, G. P. and Weldon, R. J.: San Andreas fault rupture scenarios from multiple paleoseismic records: Stringing pearls, *Bull. Seismol. Soc. Am.*, 99, 471–498, 2009.
- Bryant, W. A. and Lundberg, M. M.: Quaternary Fault and Fold Database of the United States, US Geological Survey website:

- <http://earthquakes.usgs.gov/regional/qfaults>, last access: 16 October 2010, 2002.
- California Department of Conservation, Division of Mines and Geology: Probabilistic Seismic Hazard Assessment for the State of California, DMG Open-file Report 96-08, 1996.
- Dieterich, J. H. and Richards-Dinger, K. B.: Earthquake recurrence in simulated fault systems, *Pure Appl. Geophys.*, 167, 1087–1104, 2010.
- Felzer, K. R. and Brodsky, E. E.: Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature*, 441, 735–738, 2006.
- Field, E. H.: A summary of previous Working Groups on California Earthquake Probabilities. *Bull. Seismol. Soc. Am.*, 97, 1033–1053, 2007a.
- Field, E. H.: Overview of the working group for the development of regional earthquake likelihood models (RELM). *Seismol. Res. Lett.*, 78, 7–16, 2007b.
- Helmstetter, A. and Sornette, D.: Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, *J. Geophys. Res.-Solid Earth*, 107, 2237, doi:10.1029/2001JB0001580, 2002.
- Helmstetter, A. and Sornette, D.: Foreshocks explained by cascades of triggered seismicity, *J. Geophys. Res.-Solid Earth*, 108, 2457, doi:10.1029/2003JB002409, 2003a.
- Helmstetter, A. and Sornette, D.: Predictability in the epidemic-type aftershock sequence model of interacting triggered seismicity, *J. Geophys. Res.-Solid Earth*, 108, 2482, doi:10.1029/2003JB002485, 2003b.
- Hofmann, R. B.: Individual faults can't produce a Gutenberg-Richter earthquake recurrence, *Eng. Geol.*, 43, 5–9, 1996.
- Holliday, J. R., Nanjo, K. Z., Tiampo, K. F., Rundle, J. B., and Turcotte, D. L.: Earthquake forecasting and its verification, *Nonlin. Processes Geophys.*, 12, 965–977, doi:10.5194/npg-12-965-2005, 2005.
- Holliday, J. R., Turcotte, D. L., and Rundle, J. B.: Self-similar branching of aftershock sequences, *Physica A*, 387, 933–943, 2008a.
- Holliday, J. R., Turcotte, D. L., and Rundle, J. B.: A review of earthquake statistics: Fault and seismicity-based models, ETAS and BASS, *Pure Appl. Geophys.*, 165, 1003–1024, 2008b.
- Ishibe, T. and Shimazaki, K.: Seismicity in source regions of large interplate earthquakes around Japan and the characteristic earthquake model, *Earth Planet. Space*, 61, 1041–1052, 2009.
- Johnson, A. M., Fleming, R. W., and Cruikshank, K. M.: Shear zones formed along long, straight traces of fault zones during the 28 June 1992 Landers, California, earthquake, *Bull. Seismol. Soc. Am.*, 84, 499–510, 1994.
- Lennartz, S., Bunde, A., and Turcotte, D.: Missing data in aftershock sequences: Explaining the deviations from scaling laws, *Phys. Rev. E*, 78, 041115, doi:10.1103/PhysRevE.78.041115, 2008.
- Lennartz, S., Bunde, A., and Turcotte, D.: Modelling seismic catalogues by cascade models: Do we need long-term magnitude correlations?, *Geophys. J. Int.*, 184, 1214–1222, 2011.
- Lienkaemper, J. J., Williams, P. L., and Guilderson, T.: Evidence for a twelfth large earthquake on the southern Hayward Fault in the past 1900 years, *Bull. Seismol. Soc. Am.*, 100, 2024–2034, 2010.
- Machette, M. N.: Active, capable, and potentially active faults – a paleoseismic perspective, *J. Geodynam.*, 29, 387–392, 2000.
- Marrett, R. and Allmendinger, R. W.: Estimates of strain due to brittle faulting: Sampling of fault populations, *J. Struct. Geol.*, 13, 735–738, 1991.
- Naylor, M., Greenhough, J., McCloskey, J., Bell, A. F., and Main, I. G.: Statistical evaluation of characteristic earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions, *Geophys. Res. Lett.*, 36, L20303, doi:10.1029/2009GL040460, 2009.
- Ogata, Y.: Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, 50, 379–402, 1998.
- Ogata, Y.: Seismicity analysis through point-process modeling: A review, *Pure Appl. Geophys.*, 155, 471–507, 1999.
- Ogata, Y. and Zhuang, H. C.: Space-time ETAS models and an improved extension, *Tectonophysics*, 413, 13–23, 2006.
- Parsons, T. and Geist, E. L.: Is there a basis for preferring characteristic earthquakes over a Gutenberg-Richter Distribution in probabilistic earthquake forecasting?, *Bull. Seismol. Soc. Am.*, 99, 2012–2019, 2009.
- Petersen, M. D., Frankel, A. D., Harmsen, S. C., Mueller, C. S., Haller, K. M., Wheeler, R. L., Wesson, R. L., Zeng, Y., Boyd, O. S., Perkins, D. M., Luco, N., Field, E. H., Wills, C. J., and Rukstales, K. S.: Documentation for the 2008 Update of the United States National Seismic Hazard Maps, US Geological Survey Open-File Report 2008–1128, 61 pp., 2008.
- Robinson, R. and Benites, R.: Synthetic seismicity models of multiple interacting faults, *J. Geophys. Res.-Solid Earth*, 100, 18229–18238, 1995.
- Robinson, R. and Benites, R.: Synthetic seismicity models for the Wellington Region, New Zealand: Implications for the temporal distribution of large events, *J. Geophys. Res.-Solid Earth*, 101, 27833–27844, 1996.
- Rundle, J. B.: A physical model for earthquakes, 2. Application to southern-California, *J. Geophys. Res.-Solid Earth*, 93, 6255–6274, 1988.
- Rundle, J. B., Rundle, P. B., Donnellan, A., and Fox, G.: Gutenberg-Richter statistics in topologically realistic system-level earthquake stress-evolution simulations, *Earth Planet. Space*, 56, 761–771, 2004.
- Rundle, J. B., Rundle, P. B., Klein, W. d., Tiampo, K. F., Donnellan, A., and Kellogg, L. H.: GEM plate boundary simulations for the plate boundary observatory: A program for understanding the physics of earthquakes on complex fault networks via observations, theory and numerical simulation, *Pure Appl. Geophys.*, 159, 2357–2381, 2002.
- Rundle, P. B., Rundle, J. B., Tiampo, K. F., Donnellan, A., and Turcotte, D. L.: Virtual California: Fault model, frictional parameters, applications, *Pure Appl. Geophys.*, 163, 1819–1846, 2006.
- Shcherbakov, R. and Turcotte, D. L.: A modified form of Bath's law, *Bull. Seismol. Soc. Am.*, 94, 1968–1975, 2004a.
- Shcherbakov, R., Turcotte, D. L., and Rundle, J. B.: A generalized Omori's law for earthquake aftershock decay, *Geophys. Res. Lett.*, 31, L11613, doi:10.1029/2004GL019808, 2004b.
- Shcherbakov, R., Turcotte, D. L., Rundle, J. B., Tiampo, K. F., and Holliday, J. R.: Forecasting the locations of future large earthquakes: An analysis and verification, *Pure Appl. Geophys.*, 167, 743–749, 2010.
- Sornette, D. and Werner, M. J.: Constraints on the size of

- the smallest triggering earthquake from the epidemic-type aftershock sequence model, Bath's law, and observed aftershock sequences, *J. Geophys. Res.-Solid Earth*, 110, B08304, doi:10.1029/2004JB003535, 2005.
- Southern California Earthquake Data Center: available at: http://www.data.sceec.org/fault_index/alphadex.htm, last access: 16 October 2010, 2010.
- Turcotte, D. L., Holliday, J. R., and Rundle, J. B.: BASS, an alternative to ETAS, *Geophys. Res. Lett.*, 34, L12303, doi:10.1029/2007GL029696, 2007.
- US Geological Survey.: National Seismic Hazard Maps – Fault Parameters, available at: http://gldims.cr.usgs.gov/webapps/cfusion/Sites/C2002_Search/index.cfm, last access: 16 October 2010, 2002.
- Wang, Q., Jackson, D., and Zhuang, J.: Are spontaneous earthquakes stationary in California?, *J. Geophys. Res.-Solid Earth*, 115, B08310, doi:10.1029/2009JB007031, 2010.
- Ward, S. N.: An application of synthetic seismicity in earthquake statistics – the Middle America trench, *J. Geophys. Res.-Solid Earth*, 97, 6675–6682, 1992.
- Ward, S. N.: A synthetic seismicity model for southern California: Cycles, probabilities, and hazard, *J. Geophys. Res.-Solid Earth*, 101, 22393–22418, 1996.
- Ward, S. N.: San Francisco Bay Area earthquake simulations: A step toward a standard physical earthquake model, *Bull. Seismol. Soc. Am.*, 90, 370–386, 2000.
- Wesnousky, S. G.: The Gutenberg-Richter or characteristic earthquake distribution, which is it?, *Bull. Seismol. Soc. Am.*, 84, 1940–1959, 1994.
- Wills, C. J., Weldon, R., and Bryant, W. A.: California fault parameters for the National Seismic Hazard Maps and Working Group on California Earthquake Probabilities, Appendix A in The Uniform California Earthquake Rupture Forecast, version 2 (UCERF 2), US Geological Survey Open-File Report 2007-1437A, and California Geological Survey Special Report 203A, 48 pp., 2008.
- Yakovlev, G., Turcotte, D. L., Rundle, J. B., and Rundle, P. B.: Simulation-based distributions of earthquake recurrence times on the San Andreas fault system, *Bull. Seismol. Soc. Am.*, 96, 1995–2007, 2006.
- Yikilmaz, M. B., Turcotte, D. L., Yakovlev, G., Rundle, J. B., and Kellogg, L. H.: Virtual California earthquake simulations: simple models and their application to an observed sequence of earthquakes, *Geophys. J. Int.*, 180, 7340–742, 2010.