Pure appl. geophys. 163 (2006) 1819–1846
0033–4553/06/091819–28
DOI 10.1007/s00024-006-0099-x

Pure and Applied Geophysics

Virtual California: Fault Model, Frictional Parameters, Applications

P. B. RUNDLE,[1] J. B. RUNDLE,[1] K. F. TIAMPO,[2] A. DONNELLAN,[3] and D. L. TURCOTTE[4]

Abstract—Virtual California is a topologically realistic simulation of the interacting earthquake faults in California. Inputs to the model arise from field data, and typically include realistic fault system topologies, realistic long-term slip rates, and realistic frictional parameters. Outputs from the simulations include synthetic earthquake sequences and space-time patterns together with associated surface deformation and strain patterns that are similar to those seen in nature. Here we describe details of the data assimilation procedure we use to construct the fault model and to assign frictional properties. In addition, by analyzing the statistical physics of the simulations, we can show that that the frictional failure physics, which includes a simple representation of a dynamic stress intensity factor, leads to self-organization of the statistical dynamics, and produces empirical statistical distributions (probability density functions: PDFs) that characterize the activity. One type of distribution that can be constructed from empirical measurements of simulation data are PDFs for recurrence intervals on selected faults. *Inputs* to simulation dynamics are based on the use of *time-averaged* event-frequency data, and *outputs* include PDFs representing measurements of *dynamical variability* arising from fault interactions and space-time correlations. As a first step for productively using model-based methods for earthquake forecasting, we propose that simulations be used to generate the PDFs for recurrence intervals instead of the usual practice of basing the PDFs on standard forms (Gaussian, Log-Normal, Pareto, Brownian Passage Time, and so forth). Subsequent development of simulation-based methods should include model enhancement, data assimilation and data mining methods, and analysis techniques based on statistical physics.

Key words: Earthquakes, simulations, fault Systems, forecasting, statistics.

1. Introduction—Virtual California

Virtual California is a model representing the structure and dynamics of the vertical, strike-slip earthquake fault system in California (RUNDLE, 1988; see RUNDLE *et al*., 2004 for a recent description). It is a type of model called a *backslip model* (see the Appendix for details), so-called because the loading in the model arises

[1] Center for Computational Science and Engineering and Departments of Physics and Engineering, University of California at Davis, Davis, CA, U.S.A.
E-mail: prundle@cse.ucdavis.edu; jbrundle@ucdavis.edu
[2] Department of Earth Sciences, University of Western Ontario, London, ON, Canada.
E-mail: ktiampo@uwo.ca
[3] Earth and Space Science Division, Jet Propulsion Laboratory, Pasadena, CA, U.S.A.
E-mail: andrea.donnellan@jpl.nasa.gov
[4] Department of Geology, University of California at Davis, Davis, CA, U.S.A.
E-mail: turcotte@geology.ucdavis.edu

from ''negative slip'' or ''backslip'' applied to each fault segment at its geologically observed long-term rate of offset, $V(\boldsymbol{x})$. We note that in this context, a ''fault segment'' is regarded simply as a ''degree of freedom'', rather than a spatially coherent entity with geological meaning. All of the rectangular fault segments are embedded in an elastic half space, and they interact with each other by means of quasistatic elastic interactions, whose stress Greem's functions are computed by means of a Boundary Element Method.

Frictional coefficients are assigned to each fault segment, along with other frictional parameters, by means of a *data assimilation* technique described in this paper. When the model is used to produce a simulation, the result is a history of slip on the fault segments in response to the driving forces. The interactions between the fault segments serve to organize the system so that, instead of a sequence of single segments breaking individually, multiple segments break simultaneously, producing large earthquakes. The dynamical evolution in the model is also produced by means of a stochastic, cellular automaton method, in which a random overshoot or undershoot component is added at the time of sliding of a segment. Once an earthquake history is computed, the surface deformation can be computed as well by using the appropriate kinematic Green's functions, which at the moment are also computed by means of Boundary Element Methods.

As an aside, we note that, since the Green's function code is an entirely separate code from the actual time-stepping, dynamical evolution simulation code, the stress and kinematic Green's functions could also be computed, for example, by a Finite Element Method. For that reason, this approach to earthquake simulation is quite general and in principle, any level of geometric or topological complexity is possible in the structure of the fault system. In addition, this generality will allow us to include any general type of faulting, including thrust faults, normal faults, and dipping strike-slip faults, in future models.

The usefulness of models such as Virtual California is that: 1) They allow the user to investigate all of the multitude of time and space scales in the problem that are inaccessible to direct human observation; 2) they allow the basic physics of the system to be investigated through the use of a ''numerical laboratory'' approach to earthquake fault system science; 3) they allow the user to have access not only to the observable surface deformation, but in addition to all the other physical variables that are generally impossible to observe, such as internal stresses and strains anywhere within the medium; and 4) they allow the user to develop a number of forecasting and prediction methodologies, such as statistical methods, ensemble methods, and model steering methods, based on data assimilation and parameter estimation and refinement techniques.

A similar type of simulation, the *Standard Physical Earth Model* (*SPEM*) has been developed and used by WARD (2000, 1996) and WARD AND GOES (1993). *SPEM* has two important differences from *Virtual California*. The first is that whereas *Virtual California* utilizes the full equations of three-dimensional elasticity, *SPEM* is

a plane-strain model. The second difference is in the form of the friction laws. The equations describing the laboratory-based friction model for *Virtual California* are detailed in the appendix here. In contrast, *SPEM* uses two scales: an inner and an outer scale, which control the conditions for rupture initiation and healing. Despite these important differences, it can be shown that many of the statistical results produced by the two simulations are very similar (RUNDLE *et al.*, 2006). This indicates that the self-organization of the system induced by the elastic interactions is probably the dominant feature of these types of simulations, and that details of the local frictional physics are probably less important in the system-level physics.

In the remainder of this paper, we describe details of the construction of the Virtual California fault and friction models at a greater level of detail than has been previously published. We describe details of the data assimilation method by which the frictional parameters are set for each fault segment. In an appendix, we also describe details of the physics of the Virtual California simulation, in which the friction and other parameters are introduced.

What is described here is the "Standard Model" for Virtual California. Minor variations of this model have been used in recent work (RUNDLE *et al.*, 2004) to improve computational performance and efficiency. Moreover, since several versions of the model have appeared in the literature over recent years (e.g., RUNDLE *et al.*, 2001, 2004), part of this paper is a chronology of successive versions of the model, and how they were constructed.

2. Compilation of the Fault Geometry

A fault segment in the Virtual California model is geometrically represented by a two-dimensional rectangular object embedded in an elastic half space. Only strike-slip faults with the potential to be involved in magnitude $M \geq 6$ earthquakes are included, which allows us to reasonably assume that each fault segment has a vertical dip and horizontal rake. Furthermore, although we collect depth measurements for each segment (in order to prepare for the inclusion of depth-dependent dynamics), for now we assume that all faults run from 0 to 20 km deep. Thus, the two endpoints of the surface trace, measured in latitude and longitude[1], define a fault segment.

Our earliest model (RUNDLE *et al.*, 2001), used in simulations through 1999, is based mostly on the data in Table 1 of DENG and SYKES (1997). As such, it is limited to southern California and a small area of northern Mexico, from roughly 36.2° to 31.8° N. The Mojave block and offshore regions are noticeably incomplete. Our model includes only strike-slip faults, so just the segments labeled RL (right-lateral strike-slip) and LL (left-lateral strike-slip) are used. In the DENG and SYKES (1997)

[1] North latitudes and west longitudes are given positive values.

Table 1

Segments and geologic rates of offset for the modified version of VC 2001 used in Figures 1–7

Fault or Fault System Name	Segment Nos.		Chart Distance (km)		Average Slip Rate (mm/yr)
	Begin	End	Begin	End	
Bartlett Springs	0	7	0.0	84.7	6
Calaveras	8	22	84.7	238.9	15 (8- > 17) 6 (18- > 22)
Collayomi	23	25	238.9	266.8	.6
Concord-Green Valley	26	31	266.8	322.2	6
Death Valley	32	55	322.2	569.6	5 (32- > 49) 4 (50- > 55)
Garberville-Briceland	56	59	569.6	609.2	9
Greenville	60	66	609.2	682.2	2
Hayward	67	77	682.2	793.3	9 (67- > 74) 3 (75- > 77)
Hunter Mtn.-Saline Val.	78	84	793.3	861.3	2.5
Hunting Creek-Berryessa	85	90	861.3	920.3	6
Lake Mountain	91	93	920.3	953.7	6
Maacama	94	111	953.7	1133.3	9
Monterey Bay-Tularcitos	112	119	1133.3	1213.6	.5
Ortigalita	120	126	1213.6	1280.1	1
Owens Valley	127	138	1280.1	1401.6	1.5
Palo Colorado-Sur	139	146	1401.6	1479.8	3
Panamint Valley	147	156	1479.8	1584.5	2.5
Quien Sabe	157	158	1584.5	1607.6	1
Rinconada	159	177	1607.6	1796.9	1
Rodgers Creek	178	183	1796.9	1858.9	9
Round Valley	184	189	1858.9	1914.3	6
San Gregorio	190	198	1914.3	2003.3	5
Sargent	199	203	2003.3	2056.0	3
West Napa	204	206	2056.0	2085.9	1
White Mountains	207	216	2085.9	2186.5	1
San Andreas North	217	263	2186.5	2653.6	24 (217- > 248) 17 (249- > 263)
San Andreas Creeping	264	273	2653.6	2751.3	34
San Andreas South	274	335	2751.3	3330.7	34 (274- > 298) 30 (299- > 312) 24 (313- > 321) 25 (322- > 335)
San Jacinto	336	364	3330.7	3622.1	12 (336- > 352) 14 (353- > 364)
Elsinore	365	388	3622.1	3857.5	3 (365- > 368) 5 (369- > 384) 4 (385- > 388)
Imperial Valley	389	406	3857.5	4020.0	30
Laguna Salada	407	416	4020.0	4118.5	4

Table 1

(Contd.)

Garlock	417	440	4118.5	4353.0	−5 (417- > 426)−7 (427- > 440)
Palos Verdes	441	447	4353.0	4428.6	3
Santa Cruz Island	448	452	4428.6	4481.9	-3
Brawley	453	457	4481.9	4533.8	25
Santa Monica	458	468	4533.8	4653.3	-3
Cleghorn	469	470	4653.3	4676.4	-3
Tunnel Ridge	471	472	4676.4	4695.6	-1.3
Helendale	473	481	4695.6	4781.7	.8
Lenwood-Lockhart	482	499	4781.7	4955.2	.8
Pipes Canyon	500	501	4955.2	4970.8	.7
Gravel Hills-Harper	502	509	4970.8	5051.2	.9
Blackwater	510	516	5051.2	5113.0	2
Camp Rock-Emerson	517	527	5113.0	5227.2	1 (517- > 524) .6 (525- > 527)
Homestead Valley	528	530	5227.2	5254.4	.6
Johnson Valley	531	536	5254.4	5320.4	.6
Calico-Hidalgo	537	549	5320.4	5455.5	1 (537) 1.7 (538) 2.6 (539- > 545) .6 (546- > 549)
Pisgah-Bullion	550	562	5455.5	5571.2	1
Mesquite Lake	563	564	5571.2	5592.2	1
Pinto Mountain	565	573	5592.2	5676.0	-1
Morongo Valley	574	574	5676.0	5690.6	-.5
Burnt Mountain	575	576	5690.6	5707.6	.6
Eureka Peak	577	578	5707.6	5725.8	.6
Hollywood-Raymond	579	582	5725.8	5763.7	-1 (579- > 580) -.5 (581- > 582)
Inglewood-Rose Cyn	583	604	5763.7	5979.2	1 (583- > 590) 1.5 (591- > 604)
Coronado Bank	605	623	5979.2	6179.5	3
San Gabriel	624	637	6179.5	6310.8	3 (624- > 628) 2 (630- > 633) 1 (634- > 637)
Big Pine	638	644	6310.8	6379.5	-4
White Wolf	645	649	6379.5	6427.6	-5

data tabulation, the surface traces are defined by a midpoint (latitude/longitude), a length (km), and a strike angle (degrees clockwise from North), so a simple algorithm is used to find the segment endpoints.

Assume a spherical Earth with radius 6371 km. Define midpoint longitude θ, midpoint latitude ϕ, strike angle ω, and length L. Take $\omega' = 90°-\omega$. The endpoint latitudes are

$$\phi \pm \frac{L}{2} \sin \omega' \frac{180°}{\pi(6371\text{km})}. \quad (1)$$

The endpoint longitudes are

$$\theta \pm \frac{L}{2}\cos\omega' \frac{180^\circ}{\pi(6371\text{km})\cos\frac{\phi+\phi_C}{2}}, \tag{2}$$

where ϕ_C is the latitude of the corresponding endpoint. The $\cos\phi + \phi_C/2$ term is a compromise to avoid integration.

DENG and SYKES (1997) do not contain the Landers fault, consequently three Landers segments were added, based on WALD and HEATON (1994). Measurements are taken directly from their map of three fault segments labeled Figure 1.

In 2000, the model geometry was again revised to include northern California faults with the added segments ranging from approximately 35.5° to 41.3° N. The great majority of fault segments are taken from a table of fault parameters used in (http://eqhazmaps.usgs.gov/html/faults.html) USGS, 1996 Seismic Hazard maps, compiled by BARNHARD and HANSON (1996). Faults included are those in California that are, for the most part, north of latitude 35.5°N. Any faults without a dip angle

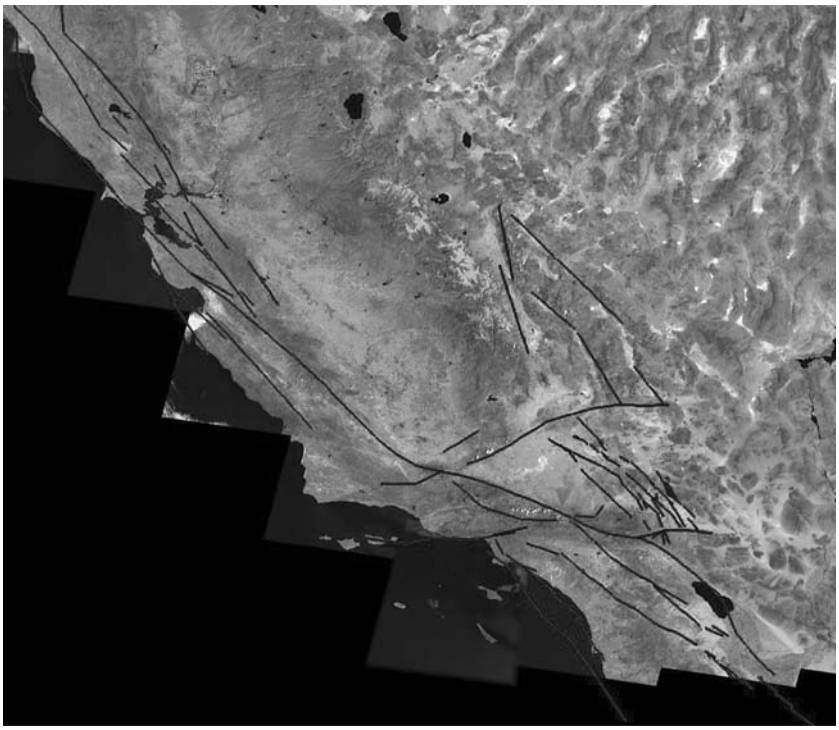

Figure 1
VC 2001 superposed on map of California. (Map of California courtesy of About.com.)

of 90° are excluded, based on the assumption that they are not strike-slip[2]. The endpoints of the faults were entered as they appeared.

Data for the northern San Andreas fault segments were assembled from several sources. First, the northern endpoint of the DENG and SYKES (1997) segment SACreeping1 is connected to the southern endpoint of the BARNHARD and HANSON (1996) segment San Andreas Fault, 1906. Both endpoints of San Andreas Fault, 1906 are used, however, because of this model segment's enormous span, intermediate endpoints are added to better reflect the natural trace. The southernmost such endpoints belong to the BARNHARD and HANSON (1996) segments 'San Andreas Fault, Santa Cruz Mtn.' and San Andreas Fault, Peninsula Segment. Four additional intermediate endpoints are added between 39.1° and 40.1°N' based on measurements taken from the map in Figure 2A of WGCEP (1999). Finally, north of the San Andreas Fault, 1906 segment, there is an intermediate endpoint near the Mendocino Triple Junction measured from Figure 1 of WARD (2000), and a final endpoint is taken from the westernmost endpoint of BARNHARD and HANSON (1996) Mendocino Fault Zone.

The final major revision to the Virtual California fault model came in 2001. The focus was the full realization of southern California, compensating for the omissions of DENG and SYKES (1997). The principal changes are that the Mojave block is fleshed out and offshore faults are added. The majority of the fault trace information for this revision is taken directly from the clickable fault maps on the Southern California Earthquake Data Center website (http://www.data.scec.org/faults/faultmap.html). The digitized maps were enlarged and endpoints were specified so as to reasonably approximate the fault trace shapes. These endpoints were measured and, where possible, corrected against the source material for BARNHARD and HANSON (1996), Appendix A of USGS Open-File Report 96-706 by PETERSEN *et al*., (1996). Some smaller faults, as well as continuations of offshore faults beyond the scope of the SCEDC maps, are taken directly from PETERSEN *et al*. (1996). Non-strike-slip faults are excluded, based on information on the SCEDC website and the first column of PETERSEN *et al*. (1996). In addition, some strike-slip faults are excluded if they are inactive during modern times, so that they have not ruptured during the Holocene epoch. Length and geometrical significance are also considered, especially if a fault seems likely to transfer stress from one fault subsystem to another. The reconstruction of the Mojave block also necessitated the removal of the WALD and HEATON (1994) Landers faults and the DENG and SYKES (1997) Pisgah fault.

Once the fault segment endpoints in a version of the Virtual California model are compiled, they are converted from latitude and longitude to distances (in km) north and east of an arbitrary reference point—31°N, 121°W. The resulting plot is an equal-area sinusoidal projection based on a central meridian of 121°W. As relative

[2] The table does not explicitly specify fault type. It should also be noted that the table lists no dip angles between 90° and 75°, making 90° the logical cutoff point.

distances in such a plot grow more distorted farther from the central meridian, future models should move the central meridian eastward to minimize distortion. Currently, linear distortion does not exceed 2% for any portion of the model, and segment strike angles are distorted by no more than 10°, and generally much less.

As the Virtual California simulation code is designed to work with fault segments of approximately 10 km in length, the length of a M $\approx$ 6 earthquake, the final step is to split the segments into appropriate pieces. The following algorithm is used. The length L of a given fault with endpoints ($[x_1, y_1]$, $[x_2, y_2]$) is computed by $L = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$. L is then divided by 10 and rounded to the nearest integer. This new value N represents the number of smaller segments into which the initial fault will be split. The endpoints of the ith smaller segment, $1 \leq i \leq N$, are then

$$\left(\left[x_1 + \frac{i-1}{N}(x_2 - x_1)y_1 + \frac{i-1}{N}(y_2 - y_1)\right], \left[x_1 + \frac{i}{N}(x_2 - x_1)y_1 + \frac{i}{N}(y_2 - y_1)\right]\right).$$

In the event that the fault-specific depth parameters are used in future models, the appropriate sources are hereby stated: DENG and SYKES (1997) faults, BARNHARD and HANSON (1996) faults, and PETERSEN *et al.* (1996) faults are all associated with depth values in the tables from which the endpoints are taken. The WALD and HEATON (1994) Landers faults are assigned depths of 15 km based upon that paper's Figure 5 and comments from the text. Depth values for faults taken from the SCEDC clickable fault maps are based on corresponding or similar neighboring faults in PETERSEN *et al.* (1996) tables. For depth values along the northern San Andreas, the respective sources (traveling from south end to north end) are: DENG and SYKES (1997), SACreeping1; BARNHARD and HANSON (1996), San Andreas Fault, Santa Cruz Mtn.; BARNHARD and HANSON (1996), San Andreas Fault, Peninsula Segment; WGCEP (1999), Table 2, SAF—North Coast South; WGCEP (1999), Table 2, SAF–North Coast North; BARNHARD and HANSON (1996), Mendocino Fault Zone.

3. Activity and Friction Parameters

The dynamics of the Virtual California model are largely governed by the frictional parameters associated with individual faults. Each model segment is

Table 2

Mean (λ), standard deviation (θ), and Coefficient of Variation (CoV) for inter-event time interval statistics measured from the simulation data in Figures 4 and 5

Fault (Magnitude Range)	λ (years)	θ (years)	*CoV*
Northern SAF ($M \geq 5.8$)	47	28	.59
Southern SAF ($M \geq 5.8$)	57	36	.63
Northern SAF ($M \geq 7.5$)	181	100	.56
Southern SAF ($M \geq 7.5$)	221	116	.52

associated with three values: slip rate, aseismic fraction, and mean recurrence interval. The fault slip rate, in cm/year, is the long-term mean rate and includes both aseismic and coseismic slip. Right-lateral slip is considered positive and left-lateral slip negative. As before, we use only vertical strike-slip faults, allowing us to assume horizontal rake in all cases. The aseismic fraction is the fraction of the fault's total slip that is attributable to creep, and can theoretically take on any value from 0 to 1. We set a minimum aseismic fraction of 0.1 to enhance the verisimilitude of our results (the reasons are complex; see RUNDLE *et al.* (2001) for details). The mean recurrence interval is the mean time, in years, between large seismic events on the fault. We bound the mean recurrence interval between 1 and 5000 years for reasons of model functionality. Presently, we are interested only in the most active fault segments, therefore we do not consider segments that slip on average more infrequently than once every 5000 years. While slip rates, including aseismic and coseismic components, are relatively easy to find in literature, mean recurrence intervals are much harder to quantify; no fault is free to rupture in isolation. It was necessary to develop a special method to compute tractable recurrence intervals, which we discuss below.

With the exception of the northern San Andreas and Mojave block systems, assignment of total slip rate (combined aseismic and coseismic) values to faults is straightforward. For any fault in the model, we note that we enforce a minimum long-term slip rate value of 0.1 mm/yr. Slip rates for DENG and SYKES (1997) faults are taken from Table 1 of DENG and SYKES (1997); slip rates for the WALD and HEATON (1994) Landers faults are assumed to be 0.3 mm/yr; and slip rates for BARNHARD and HANSON (1996) faults are taken from BARNHARD and HANSON (1996).

For the northern San Andreas, the respective sources for slip rate values (traveling from south end to north end) are: DENG and SYKES (1997), SACreeping1; WGCEP (1999), Table 2, SAF – Santa Cruz Mtns; BARNHARD and HANSON (1996), San Andreas Fault, Peninsula Segment; WGCEP (1999), Table 2, SAF – North Coast South; WGCEP (1999), Table 2, SAF – North Coast North; BARNHARD and HANSON (1996), Mendocino Fault Zone. For the Mojave block segments based on the SCEDC clickable fault maps, slip rate values are taken from corresponding entries in the accompanying Alphabetical Fault Index whenever possible. For uncertain or missing slip values, PETERSEN *et al.* (1996) and Table 1 of PETERSEN and WESNOUSKY (1994) were consulted, and the final slip rate value is something of a consensus.

The aseismic slip factors α for fault segments (RUNDLE *et al.*, 2001) are less stringently based on the field data. Early in the construction of the model, it became clear that this parameter has a disproportionately significant impact on the simulated seismicity patterns. RUNDLE *et al.* (2001) found that the effect of this parameter, which is modeled after laboratory results of TULLIS (1996) and KARNER and MARONE (2000), generally acts to smooth the stress field on a fault, thereby leading to large events. Very similar effects are seen in the rate-and-state friction models of DIETERICH (1979). Although much of the available data states that there is no

significant aseismic slip on many faults, reflecting this condition in simulations yields seismicity patterns with very few large events. Mandating a token amount of creep for each fault gives rise to larger, coordinated event patterns that better reflect observed California seismicity.

The very first versions of the Virtual California model assign an aseismic slip fraction of 0.2 to every fault segment (i.e., 20% of a fault's slip rate is creep). In subsequent versions, field data are used with no minimum slip fraction. Initially, a minimum aseismic slip fraction of 0.2 was then introduced, and later lowered to 0.1. From the 2000 version of the model onward, all faults have aseismic slip values based on field data, when available, with a minimum value of 0.1. Fractions for DENG and SYKES (1997) faults are calculated from their endnotes to Table 1. As the DENG and SYKES (1997) model San Andreas stops at the edge of the creeping section, the bordering segment of the northern San Andreas is designated the creeping section and assigned an aseismic slip fraction of 1.0. The San Andreas segment just north of this section is given a slip fraction identical to that of the DENG and SYKES (1997) segment SACreeping1 to affect a symmetrical tapering of the aseismic slip rate on either side of the creeping section. The northwestern Lenwood fault is given creep because of a non-quantitative note in the Alphabetical Fault Index on the SCEDC website, but the 0.25 fraction is based solely on a comparison of total slip rates from PETERSEN and WESNOUSKY (1994) the SCEDC website, and PETERSEN *et al.* (1996).

The final parameter to set is the mean recurrence interval for large events. The principal difficulty is that reliable and useful seismicity records for California extend back in time no more than 200 years, a shorter period than most of the major California faults' average recurrence intervals. Also, geological records of ruptures are often quite imprecise, even when they exist. All of this makes finding consistent, useful recurrence intervals in the literature a difficult proposition and necessitates a different approach. It is assumed that, as both the coefficient of friction and the normal force at the fault plane are reasonably constant for a strike-slip fault, a segment would be predisposed to have ruptures of a particular magnitude; i.e., a "characteristic earthquake" (SCHWARTZ and COPPERSMITH, 1984). By making reasonable generalizations about properties of earthquakes, the seismic moment of an earthquake is correlated with the coseismic slip it induces. Then, by determining the seismic moment of a segment's characteristic earthquake, a characteristic slip per event is calculated. Dividing this quantity by the segment's coseismic slip rate yields the average interval between ruptures. (For creeping segments, a recurrence interval of 1 year is used.)

The relationship of seismic moment to coseismic slip in each event was developed first. We have the equations $m_o = \mu\langle s\rangle A$ and $\langle s\rangle = f\Delta\sigma\sqrt{A}/\mu$, where m_o is seismic moment, μ is shear modulus, $\Delta\sigma$ is stress drop, $\langle s\rangle$ is average slip, A is fault rupture area, and f is a dimensionless fault shape factor (KANAMORI and ANDERSON, 1975). Combining these, we get

$$\langle s \rangle = \frac{(\Delta\sigma)^{2/3} f^{2/3} m_o^{1/3}}{\mu}. \tag{3}$$

We use reasonably assumed values of $\mu = 3 \times 10^{10}$ Pa, $\Delta\sigma = 5 \times 10^6$ Pa, and $f = 1$ to yield a moment-slip relation that gives realistic values. Numerically, it is quite comparable to the strike-slip moment magnitude regression in Figure 11b of WELLS and COPPERSMITH (1994).

The next step is to determine each segment's characteristic event moment. The first method tried was to simply examine the historic earthquake catalog and attempt to associate a major ($M \geq 6$) historic event with each segment (prior to the segments' being split)[3]. The seismic moment of this event becomes the characteristic major event moment for the fault segment. Since the 1999 fault model was current at the time this method was tried, the catalog used is Table 2 of DENG and SYKES (1997). Known foreshocks and aftershocks are excluded (as they are unlikely to represent a characteristic event), as is any event whose mechanism dip and rake indicated that it is a thrust or reverse fault. Finally, a single event from the catalog is assigned to each segment, based on the known circumstances of the catalog events, or, as a last resort, based on proximity.

This method has serious problems. First, not every segment has an easily associated major historic earthquake, or even one in close proximity relative to other faults. Since every segment needs an event, a number of the event assignments are dubious. A more chronologically extensive catalog would certainly help. Second, this method quite often gives rise to situations in which neighboring, ostensibly continuous fault segments are associated with earthquakes of vastly different moments, and thus possess markedly different frictional properties. It creates a pronounced artificial frictional discontinuity at the shared endpoint, which can impede the realistic propagation of slip during an earthquake.

The problems created by the initial attempts to set characteristic major event moments led to a change in approach. We decided to assume that a fault's characteristic earthquake would be similar to the historic earthquakes that occurred in its vicinity; i.e., that regions, as opposed to individual faults, have characteristic events[4]. Then, a fault segment's characteristic major event moment may become a manner of average of the moments of the historic major events in close proximity. It is most logical to use a distance-weighted average so that faults are most closely

[3] The cutoff for characteristic events is set at $M = 6$ because each of the approximately 10-km-long fault segments must fail as a unit. A 10-km rupture length produces an event on the order of $M = 6$.

[4] The definition of *region* is naturally somewhat nebulous. As will become clear, our regions of similar characteristic event moment are essentially defined by the spatial density and moment value diversity of events in the historic earthquake catalogs. There are large regions of similar characteristic event moment where there are few recorded events or where there are many recorded events of similar magnitude. Regions of similar characteristic event moment are much smaller where there are many recorded events of differing magnitude in close spatial proximity.

associated with the nearest possible events. The current method to determine faults' characteristic major event moment is based on this reasoning.

The details of this method are as follows. First, a catalog of all major historic strike-slip events ($M \geq \sim 6$) in the region of the fault system model, excluding foreshocks and aftershocks, is compiled. Each event is characterized by seismic moment and hypocentral latitude and longitude. (The Virtual California model assumes that fault frictional attributes do not change with time, thus the dates of the earthquakes are unnecessary.) Since most earthquake catalogs list events according to magnitude, magnitude values are instead compiled and plugged into a standard moment-magnitude relation

$$m_o = 10^{1.5M+9.0}. \tag{4}$$

Then, the characteristic major event moment m_χ of the j-th fault segment is a heavily distance-weighted average of the moments of all faults in the catalog:

$$m_{\chi j} = \frac{\sum_i m_{oi} r_{ij}^{-3}}{\sum_i r_{ij}^{-3}}, \tag{5}$$

where $m_{\chi j}$ is the characteristic major event moment of the j-th fault segment, m_{oi} is the seismic moment of the i-th catalog event, and $r_{ij} = |\boldsymbol{x}_i - \boldsymbol{x}_j|$ is the distance between the j-th fault segment and the hypocenter of the i-th catalog event.

A few notes about the distance-weighted average method for determining characteristic major event moments: The distance-weighting exponent of –3 is used for two reasons. First, it is large enough to ensure that a fault's characteristic major event magnitude is only significantly influenced by the events closest to it—distant faults contribute to the average in a numerically insignificant way, and a single quake in close proximity dominates the average. Second and more importantly, it mimics the attenuation observed for changes in static stress transfer, so that faults are affected by distant earthquakes in a numerically consistent way. To address the issue of discontinuities between neighboring segments, the distance method is applied after the fault segments are split into ˜10-km segments to create the smoothest possible transitions of frictional properties along fault traces at the system level. Note that all characteristic major event moment values fall between the largest and smallest event moments in the historic catalog. One may think of this method as defining a function, continuous everywhere except at the catalog hypocenters, mapping location in the latitude-longitude domain to a characteristic major event moment—a function completely independent of the locations of fault segments.

In compiling the historic earthquake catalogs, source material was added for each incarnation of Virtual California. The 1999 version of the model uses the catalog in Table 2 of DENG and SYKES (1997), which covers southern California $M \geq 6$ events from 1812–1994. Moment magnitude and/or local magnitude are used wherever possible. The 2000 version adds two more sources. One is Table 1

of JAUMÉ and SYKES (1996), which lists 'moderate' earthquakes in the San Francisco Bay region from 1850–1993. All events with $M < 5.8$ are excluded, and all earthquakes in this region are assumed to be strike-slip events. At some point, the magnitude of the 1906 San Francisco event was changed from 7.8 to 8.2, perhaps to better match the $M = 8.3$ value found in THATCHER (1975), but this change now appears to conflict with prevailing opinion. The other source added in the 2000 model is the UC Berkeley-based CNSS earthquake catalog, via the online search interface (http://quake.geo.berkeley.edu, since changed to the ANSS earthquake catalog). We use all earthquakes from 1900–2000 with $M \geq 5.8$ located between 35° and 42.5°N latitude and 127° and 116° W longitude. There is no easy way to distinguish strike-slip events in this catalog, therefore no further exclusions are made; this obviously presents a problem. The 2001 model adds only the CNSS earthquake catalog events from 1900–2001 with $M \geq 5.8$ located between 31.5° and 35° N latitude and 122° and 114.25°W longitude. In the final version of the model, all events with $M < 6.0$ are excluded.

During initial testing of the 2001 Virtual California model, a problem with the treatment of the largest earthquakes in the catalog was noticed. For earthquakes whose rupture areas fall along more than about 100 km of a fault, there are segments with significant slip far from the epicenter. In slipping, these segments reveal accumulated slip deficits large enough to indicate that their own characteristic events would be similarly large. However, because earthquakes are treated as point sources in our distance-weighted average method, the 'influence' of these large events is diluted for these segments' characteristic major event moments. As a result, their mean recurrence intervals are artificially low. To correct for this, large historic events, defined as events whose length along strike is significantly larger than the depth, are now represented as a set of smaller concatenated events spaced along the rupture length. The magnitude of a concatenated event is based upon the slip observed at its location during the historic large event. Thus the local observed slip is inserted into the moment-slip correlation (Eq. 3) and the moment-magnitude relation (Eq. 4) to yield a magnitude value. In this way, the distance-weighted average method reproduces the observed slip patterns.

The three large earthquakes singled out for splitting into smaller events are the 1812 Wrightwood event, the 1857 Fort Tejon event, and the 1906 San Francisco event. The Wrightwood and Fort Tejon earthquakes are based on the rupture models in Tables 3a and 3b of DENG and SYKES (1997). A smaller event is placed at the location of each listed rupture segment, and each segment is associated with its listed displacement value. In the case of the Fort Tejon event, the surface slip (SS) values are used. The San Francisco event is based on the plot of surface slip vs. distance in Figure 2 of THATCHER (1975), along with the accompanying rupture area map. Seventeen evenly spaced artificial events are created along the rupture between 39.00°N, 123.69°W and 36.82°N, 121.51°W, and the corresponding local slip values are taken from the plot (with an imposed minimum slip of 1 m to enhance slip continuity).

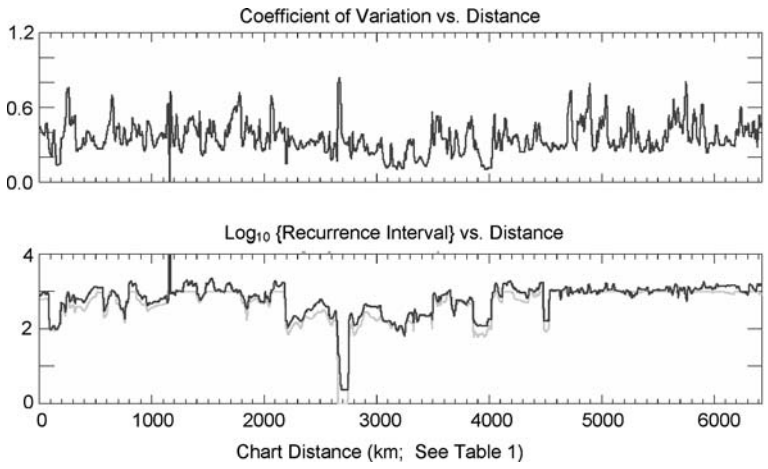

Figure 2
(a) Plot of coefficient of variation along faults. (b) Plot of average recurrence intervals along faults. Light lines: Recurrence intervals as computed from observed data. Dark lines: Recurrence intervals as measured from simulations.

4. Applications

We present computations below for a slightly modified version of the 2001 model. In these calculations, we have used a subset of the 650 of the 680 fault segments defined in the full VC 2001 model. Most of the neglected segments are at the extreme northern end of the northern San Andreas fault, where it takes a sharp westward turn north of Cape Mendocino. Scant information is available relative to the long-term slip rate on that section of fault, and its presence seemed to produce unwanted irregularities in the dynamics of the remaining segments. About five other fault segments in southern California on the Manix fault and three segments on the Verdugo fault were found to be interfering with the dynamics of the model for reasons that are presently obscure. This problem is the subject of current investigations. This somewhat reduced model is shown in Figure 1, superposed on a map of California. The fault segments are listed in Table 1 (see also RUNDLE *et al.*, 2004).

We have also assumed a uniform depth of 15 km for all faults in the model, which is in contrast to the ''standard depth'' of 20 km that was assumed in the construction of the model as described above. The depth value of 15 km was adopted because of a desire to first investigate models with uniform depth, and the value of 15 km is more reflective of the depth of seismicity in California's seismically active regions. In addition, although the data assimilation procedure described above provides our best initial estimate of the friction parameters on the faults, we find that simulated earthquakes often have a somewhat longer time interval between them on a given

segment than was computed from the data assimilation procedure (see Fig. 2). This difference is a result of the fault interactions, and their effect on the dynamics. Thus we apply an overall multiplicative factor to the friction values obtained by our data assimilation procedure in order to optimally match the average time intervals between simulated large earthquakes on the San Andreas fault to the intervals observed in nature.

We now discuss an application of our simulations to the problem of earthquake forecasting. The report from the Working Group on California Earthquake Probabilities, WG02 (2003), represents the latest version of a method that has been developed for using observed geological and geophysical data, together with broad assumptions about the appropriate statistical distribution functions that describe the stochastic nature of the process. In the WG02 (2003) method, the models that are used to compute a forecast from the data are statistical in nature, that is, the physics included is typically encoded in 1) the form of the statistical distribution, and 2) the parameters of the distribution, usually the mean and variance. WG02 then uses the statistical distribution for all the fault segments in a region, with means and variances set using the geological and geophysical data, to compute the conditional probability that a large ($M \geq 6.7$) event will occur in the San Francisco Bay region during the period 2002–2031, given that the last such event occurred a time T years ago.

In our approach, we use our Virtual California simulations to generate (''measure'') the statistical distribution needed to compute the conditional probability. In addition, we focus on sections of the northern and southern San Andreas fault. The basic idea is to use essentially the same methodology to evaluate the probability for earthquakes on these faults, but to replace the assumptions made about the form of the statistical distribution with measurements of the statistics of earthquake occurrence on the faults of interest that are obtained from our simulations, which include fault interactions.

The primary statistical distributions investigated by the WG02 are the Log-normal distribution (e.g., EVANS *et al.*, 1993), and the Brownian first passage, or Brownian Passage Time (BPT) distribution (RANGARAJAN and DING, 2000; MATTHEWS *et al.*, 2002). The Log-normal distribution is known to describe a number of physical situations where aging, fatigue, and failure processes occur. Then the probability density function (PDF) for failure between time t and $t + \delta t$ is given by:

$$P_{LN}(t) \;=\; \frac{1}{\omega\, t\sqrt{2\pi}}\, \exp\left\{\frac{-[\log(t/\rho)]^2}{2\omega^2}\right\}. \tag{6}$$

Here the parameters ρ and ω are related to the mean λ and standard deviation θ of the population statistics of waiting times t by the usual relations:

$$\rho \;\equiv\; \frac{\lambda}{\sqrt{\left(\frac{\theta}{\lambda}\right)^2 \;+\; 1}}, \tag{7}$$

$$\omega \equiv \sqrt{\log\left\{\left(\frac{\theta}{\lambda}\right)^2 + 1\right\}}. \tag{8}$$

Likewise, the Brownian Passage Time PDF for failure between time t and $t+\Delta t$ is given by (RANGARAJAN and DING, 2000; MATTHEWS *et al.*, 2002):

$$P_{BPT}(t) = \sqrt{\frac{\lambda^3}{2\pi\theta^2 t^3}} \exp\left\{-\frac{\lambda(t-\lambda)^2}{2\theta^2 t}\right\}. \tag{9}$$

We also discuss the Weibull PDF (TURCOTTE, 1997), which is the type of statistics often used in Japan (RIKITAKE, 1982). This can be written in the form (EVANS *et al.*, 1993):

$$P_W(t) = \beta\left[\frac{t^{\beta-1}}{\tau^\beta}\right] \exp\left\{-\left(\frac{t}{\tau}\right)^\beta\right\}. \tag{10}$$

The mean λ and standard deviation θ of $P_W(t)$ are given in terms of the parameters τ, β by:

$$\lambda = \tau\,\Gamma\left[\frac{\beta+1}{\beta}\right], \tag{11}$$

$$\theta = \tau^2\left\{\Gamma\left[\frac{\beta+2}{\beta}\right] - \left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^2\right\}. \tag{12}$$

In the WG02 method, the values for the standard deviation of earthquake recurrence time θ_i, and the mean recurrence time interval λ_i are inferred from a variety of observations on the selected Bay Area fault segments. A problem that is sometimes encountered is that for some segments, observations for only the last earthquake involving that segment exist; thus, the values of θ_i and λ_i are estimated by indirect means. Once the PDFs $P(t)$ for the segments are established, the Cumulative Distribution Function $C(t)$, also called simply the "distribution function", is then used to compute the conditional probability (also called the "discrete hazard function") for failure of the segment during the interval $(T, T+\Delta T)$, given that it has been a time T since the last such failure (WG02, 2003):

$$H_{\Delta T}(t,T) \;=\; P(T \le t \le T+\Delta T \,|\, t \ge T) \;=\; \frac{C(T+\Delta T) - C(T)}{1-C(T)}. \tag{13}$$

The WG02 (2003) computes the $\Delta T = 30$ year conditional probabilities (Eq. 13) for the interval 2002–2031, for the occurrence of an earthquake with magnitude $M \ge 6.7$, occurring on the active earthquake faults in the San Francisco Bay region.

Here we compute the entire Discrete Hazard Function $H_{\Delta T}(t, T)$ for the occurrence of an earthquake with magnitude $M \ge 7.5$, for two intervals, 10 years and 30 years, as a function of the time since the last such event anywhere on the fault.

Instead of assuming the PDF $P(t)$ and its associated distribution $C(t)$, we measure these quantities from simulations. Our method uses basically the same type of observational data as does the WG02, but instead of using these data to estimate the statistical parameters θ_i and λ_i for each fault segment, the data are assimilated into physical values for the Virtual California fault segments as described above. We then use the simulations to directly measure $P(t)$ and $C(t)$ for the fault segments of interest, and finally compute $H_{\Delta T}(t,T)$ as in Eq. (8).

The results are shown in Figures 2–7, which are based on 40,000 years of simulated earthquakes on the fault system shown in Figure 1. Figure 2a shows the Coefficient of Variation (CoV) on each fault segment, plotted as a function of "chart distance" along the fault system (see Table 1). Recall that the CoV is defined as the ratio of standard deviation θ to the mean λ of a probability density:

$$CoV \equiv \frac{\theta}{\lambda}. \tag{14}$$

5. Computational Results

Figure 3 shows the sections of the San Andreas fault that are considered for our computation of conditional probabilities. Figures 4–7 compare the output of Virtual California simulations to results obtained using WG02 (2003)-type methods employing the Log-normal distribution (dotted lines in the figures), the BPT distribution

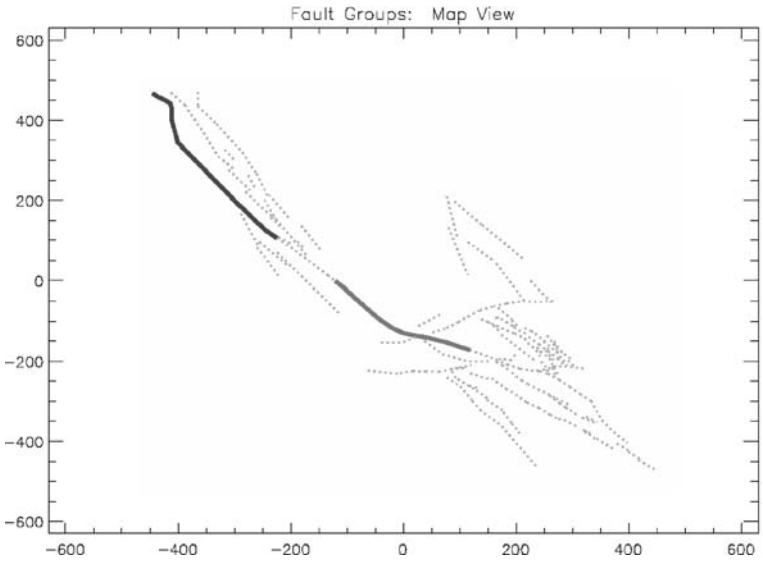

Figure 3
Map of VC 2001 fault model with northern San Andreas fault (NSAF) segments (dark gray) and southern San Andreas fault (SSAF) segments (light gray) indicated. Axes are in km.

(dashed lines in the figures), and the Weibull distribution (solid lines in the figures). The irregular curve is obtained from the simulation data. The Log-normal, BPT and Weibull distributions are chosen to have the same mean and standard deviations as the simulation results (see Table 2). In Figure 4, we show four histograms for inter-event

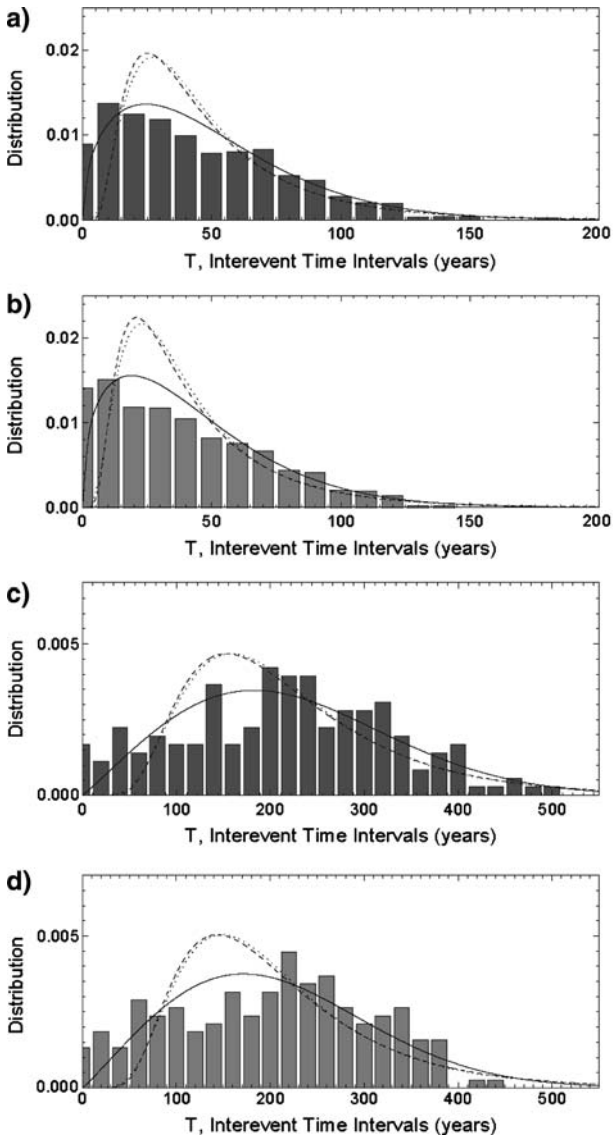

Figure 4

Histograms for interevent time intervals for earthquake data from Virtual California simulations. (a) NSAF events having $M \geq 7.0$. (b) SSAF events having $M \geq 7.0$. (c) NSAF events having $M \geq 7.5$. (d) SSAF events having $M \geq 7.5$.

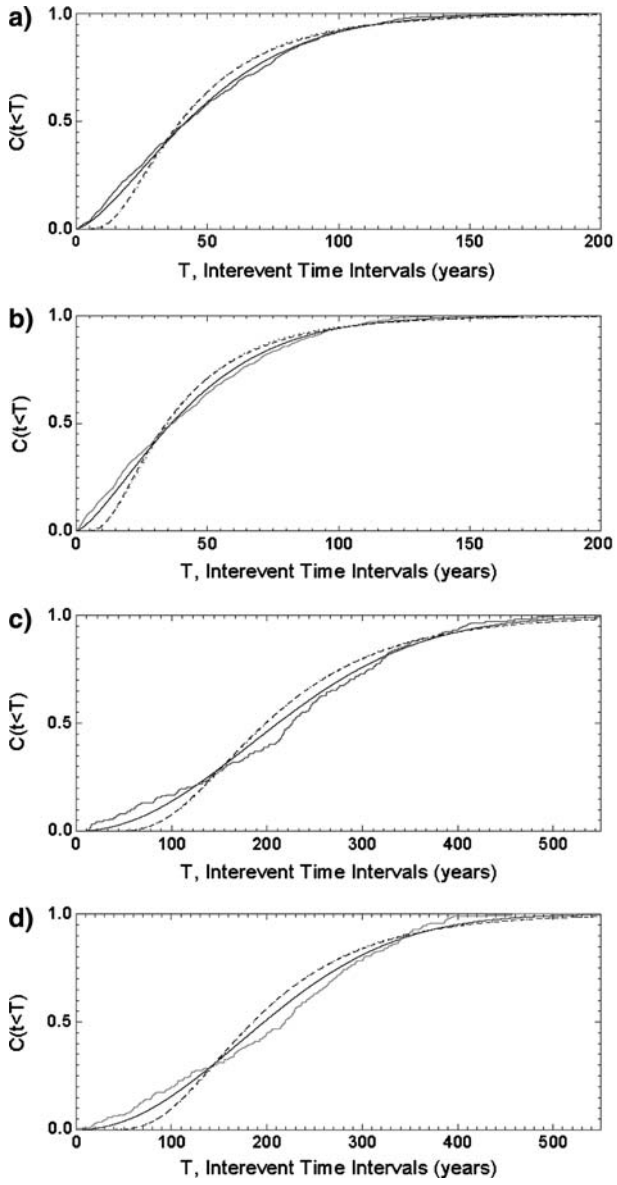

Figure 5
Cumulative Distribution Functions for interevent time intervals for earthquake data from Virtual California simulations. (a) NSAF events having $M \geq 7.0$. (b) SSAF events having $M \geq 7.0$. (c) NSAF events having $M \geq 7.5$. (d) SSAF events having $M \geq 7.5$.

time intervals for a) events having $M \geq 7.0$ on the northern San Andreas fault; b) events having $M \geq 7.0$ on the southern San Andreas fault; c) events having $M \geq 7.5$ on the northern San Andreas fault; and d) events having $M \geq 7.5$ on the southern San Andreas

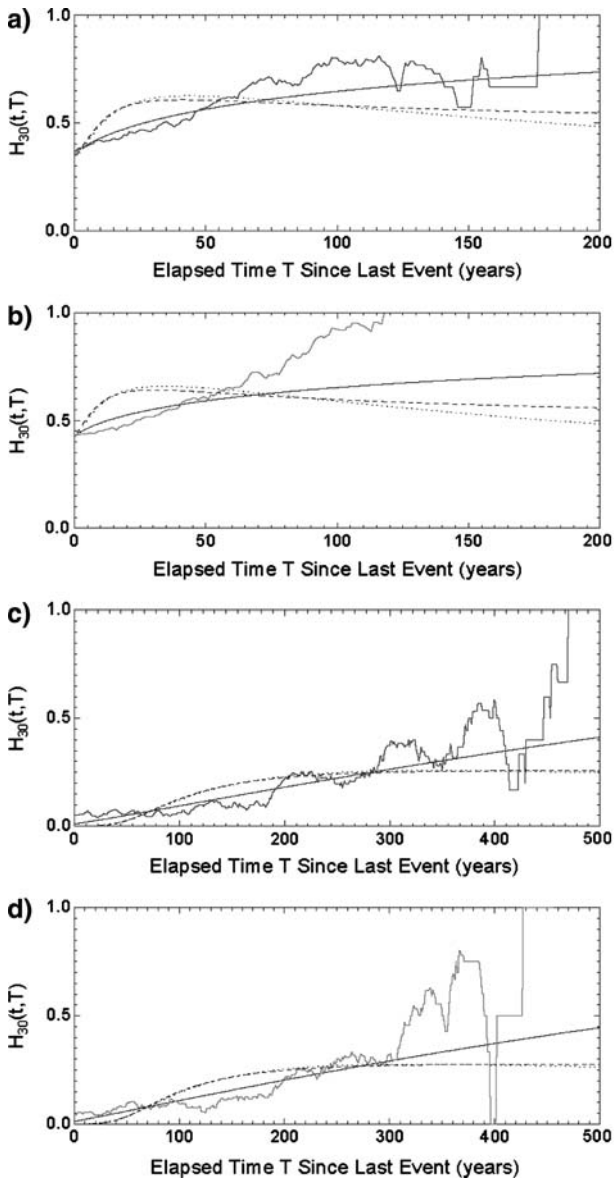

Figure 6
30-year Discrete Hazard Functions (Conditional Probabilities) for earthquake data from Virtual California simulations. (a) NSAF events having $M \geq 7.0$. (b) SSAF events having $M \geq 7.0$. (c) NSAF events having $M \geq 7.5$. (d) SSAF events having $M \geq 7.5$.

fault. In all plots, histograms with the same normalization and the same statistical parameters are shown for the Log-normal (solid) and BPT statistics. When integrated and normalized to 1, the histograms can be used as PDFs to compute $P(t)$. In Figures

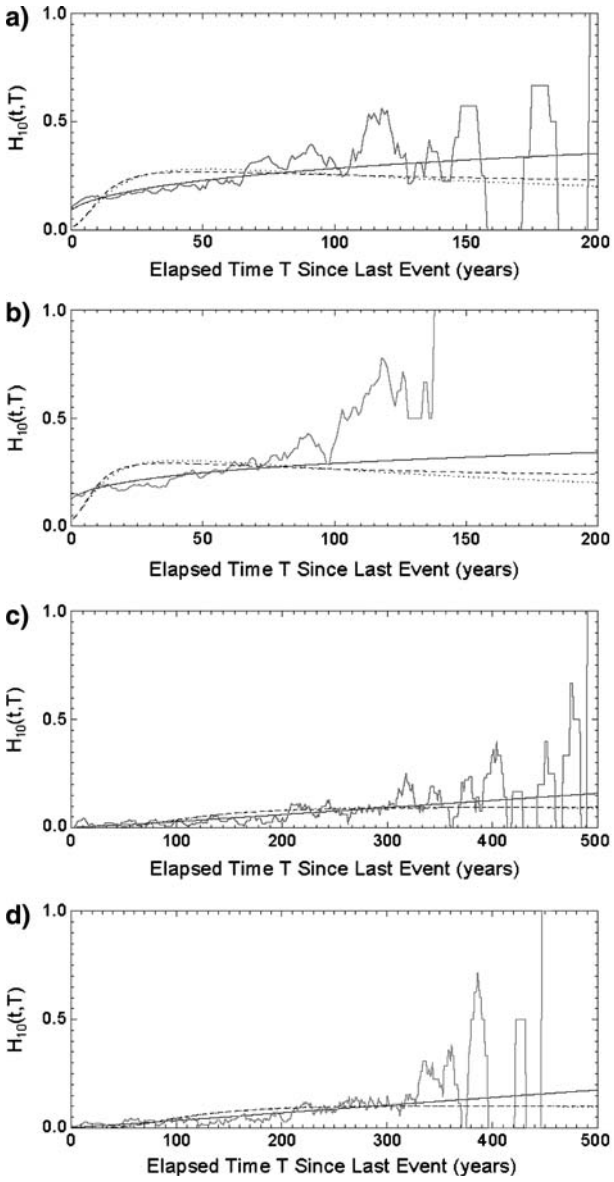

Figure 7
10-year Discrete Hazard Functions (Conditional Probabilities) for earthquake data from Virtual California simulations. (a) NSAF events having $M \geq 7.0$. (b) SSAF events having $M \geq 7.0$. (c) NSAF events having $M \geq 7.5$. (d) SSAF events having $M \geq 7.5$.

4a and 4b, it can be seen that the Log-normal and the BPT curves are characteristically too high at the small-interval end, and too low at the high-interval end. However, the Weibull distribution provides a reasonable, average representation of the simulation

data. Figures 5(a–d) show CDFs, or $C(t)$ curves corresponding to the histograms of Figure 4. Differences between the measured data and the theoretical curves for the Log-normal and BPT CDFs are not as apparent to the eye as they are for the histograms; a well-known result that arises due to the fact that the CDFs represent the integration of the histogram or PDF data. However, the Weibull distribution is clearly different and again represents a better fit to the simulation data.

Figures 6(a–d) plot the 30-year values for $H_{30}(t,T)$. Figure 6a is for all northern San Andreas events having $M \geq 7.0$; Figure 6b is for all southern San Andreas events having $M \geq 7.0$; Figure 6c is for all northern San Andreas events having $M \geq 7.5$; and Figure 6d is for all southern San Andreas events having $M \geq 7.5$. It can be seen that the $H_{30}(t,T)$ curves for the simulation data with $M \geq 7.0$ (Figs. 6a,b) are not well-represented by the $H_{30}(t,T)$ curves obtained from either the Log-normal or BPT statistics, but are reasonably well represented by the Weibull statistics. Moreover, the $H_{30}(t,T)$ curves eventually rise to the asymptotic value $H_{30}(t,T) \rightarrow 1$ as t increases, due to the fact that all faults in the simulation eventually fail. By contrast, the $H_{30}(t,T)$ curves for the Log-normal and BPT distributions do not tend asymptotically to $H_{30}(t,T) \rightarrow 1$, which is a well-understood feature of these statistics (WG02 2003). The Weibull distribution produces $H_{30}(t,T)$ curves that do eventually tend asymptotically to 1.

Finally, Figures 7(a–d) plots the 10-year values for $H_{10}(t,T)$. Figure 7a is for all northern San Andreas events having $M \geq 7.0$; Figure 7b is for all southern San Andreas events having $M \geq 7.0$. Figure 7c is for all northern San Andreas events having $M \geq 7.5$; Figure 7d is for all southern San Andreas events having $M \geq 7.5$. Again, the $H_{10}(t,T)$ curves eventually rise to the asymptotic value $H_{10}(t,T) \rightarrow 1$ as t increases, due to the fact that all faults in the simulation eventually fail. The $H_{10}(t,T)$ curves for the Log-normal and BPT distributions also do not tend asymptotically to $H_{10}(t,T) \rightarrow 1$, but the Weibull distributions do. Although the Log-normal and BPT curves approximately represent the simulation data for about 130 years' inter-event time, the $H_{10}(t,T)$ curves thereafter rise abruptly to $H_{10}(t,T) \rightarrow 1$ as all fault segments eventually fail.

6. Conclusions

We conclude that topologically realistic simulations such as Virtual California can be easily used to develop forecast statistics for earthquakes. An advantage of simulation data such as that shown here is that interactions among the faults in the system are fully included—they do not have to be assumed or added *a posteriori*. Moreover, all of the faults in the system eventually fail, meaning that there is no paradox that failure becomes increasingly unlikely as the time since the last event recedes, as is the case with renewal models such as the Log-normal distribution.

As an example of how these forecast figures are used, we can consider the Loma Prieta earthquake of 1989 ($T = 15$ years). Note that the southern San Andreas fault depicted in Figure 3 does not include the Parkfield segment. Assuming that the Loma

Prieta earthquake was on the San Andreas fault, Figures 6a and 7a show that the 30-year conditional probability for failure $H_{30}(t,15)$ on the northern San Andreas fault, given the most recent earthquake larger than $M \geq 5.8$, is now approximately 50%. The corresponding 10-year conditional probability for failure $H_{10}(t,15)$ is approximately 18%.

There are two important qualifications on these computations, which are: 1) these hazard probabilities represent lower bounds since we have not included all possible observed (and unobserved) faults in the model; and 2) we have not controlled for the time of earthquake intervals with respect to occurrence time within the earthquake cycle of the greatest events.

The validity of our data assimilation procedure also depends importantly on the characteristics (location, magnitude) of the large historical events upon which it is based. To the extent that these basic parameters are unknown, or that events were not observed or adequately characterized, our procedures for assigning friction values will be in error. We fully expect that knowledge of these parameters will improve as better data are obtained.

Acknowledgements

This work has been supported by a grant from US Department of Energy, Office of Basic Energy Sciences to the University of California, Davis DE-FG03-95ER14499 (JBR); with additional funding from the National Aeronautics and Space Administration under grants to the Jet Propulsion Laboratory and the University of California, Davis (JBR; PBR, GM, DT; AD, PL). This research was also supported in part by the Southern California Earthquake Center (PBR). SCEC is funded by NSF Cooperative Agreement EAR-0106924 and by USGS Cooperative Agreement 02HQAG0008. SCEC contribution number for this paper is 777.

Appendix

Physics of the Virtual California Model

The *Virtual California* model (RUNDLE *et al.*, 2001, 2004, 2005) is a stochastic, cellular automata instantiation of an earthquake *backslip* model, in that loading of each fault segment occurs via the accumulation of slip deficit $\phi(\boldsymbol{x},t) = s(\boldsymbol{x},t) - Vt$, where $s(\boldsymbol{x},t)$ is slip, V is long-term slip rate, and t is time. At the present time, faults used in the model are exclusively vertical strike-slip faults; the most active faults in California, and upon which most of the seismic moment release is localized. Thrust earthquakes, such as the 1994 Northridge and 1971 San Fernando faults, are certainly damaging, but they occur infrequently and are therefore regarded as

perturbations on the primary strike slip fault structures. The Virtual California model also has the following additional characteristics.

1. *Surfaces of discontinuity* (faults) across which slip is discontinuous at the time of an earthquake, and which are subject to frictional resistance. Here we restrict the model to only topologically complex systems of vertically dipping faults mirroring the complexity found on the natural fault networks of southern California.
2. *Stochastic dynamics*. In these models, we are interested in the space-time patterns and correlations that emerge from the underlying stress-strain dynamics. These correlations evolve over many hundreds or thousands of years, time scales much longer than the time scales associated either with rupture or elastic wave periods. Most of the elastic and frictional parameters for faults and earth materials, although known in the laboratory, will likely remain poorly defined in nature. For this reason it makes little sense to attempt a deterministic solution to the equations of motion. Instead, we use a Cellular Automaton (CA) approach, in which the dynamics is parameterized by random variables chosen from well defined probability distributions.
3. *Linear elastic stress transfer* or interactions between fault surfaces. Again, although most of the significant parameters associated with rupture, such as friction coefficients and friction law constants and functions can be defined and measured in the laboratory, current experience indicates they will likely always be poorly known for faults in nature. We therefore use quasistatic stress interaction (Green's function) tensors $T_{ij}^{kl}(\boldsymbol{x}-\boldsymbol{x}')$, which we will write henceforth schematically as $T(\boldsymbol{x}-\boldsymbol{x}')$.
4. *Persistent increase of stresses* on the fault surfaces arising from plate tectonic forcing parameterized via the backslip method. This method has the advantage of matching the long-term rate of offset V in model faults with the geologically known long-term slip rate on faults in nature. Stress increase occurs via the following physics. The stress tensor $\sigma_{ij}(\boldsymbol{x},t)$ is related to the slip $s_l(\boldsymbol{x},t)$ by:

$$\sigma_{ij}(\boldsymbol{x},t) = \int d\boldsymbol{x}_k \; T_{ij}^{kl}(\boldsymbol{x}-\boldsymbol{x}') \; s_l(\boldsymbol{x}',t). \tag{A1}$$

   Now if $\boldsymbol{x} = \boldsymbol{x}'$, a positive slip $s_l(\boldsymbol{x},t) > 0$ results in a *decrease* in stress, $\Delta\sigma_{ij}(\boldsymbol{x},t) < 0$. Therefore, if we write the equation:

$$\sigma_{ij}(\boldsymbol{x},t) = \int d\boldsymbol{x}_k \; T_{ij}^{kl}(\boldsymbol{x}-\boldsymbol{x}') \; \{ s_l(\boldsymbol{x}',t) \; - \; V_l(\boldsymbol{x}')t \}, \tag{A2}$$

   where $V_l(\boldsymbol{x}) = \langle s_l(\boldsymbol{x},t) \rangle$ is the average long-term rate of slip at $\boldsymbol{x}'$, then the second term - $V_l(\boldsymbol{x})t$ leads to an *increase* in the stress, $\Delta\sigma_{ij}(\boldsymbol{x},t) > 0$. Therefore the second term is the stress accumulation term.
5. Parameters for friction laws and fault topology that are determined by assimilating seismic, paleoseismic, geodetic, and other geophysical data from events occurring over the last ~200 years in California (RUNDLE *et al.*, 2001, 2002).

6. Frictional resistance laws that range from the simplest Amontons-Coulomb stick-slip friction, to heuristic laws such as slip- or stress-rate dependent weakening laws based on recent laboratory friction (TULLIS, 1996) and fracture experiments. These laws are related to rate-and-state and leaky threshold laws (RUNDLE *et al.*, 2002).

In general, several of the friction laws described above can be written in the following representative, equivalent forms on an element of fault surface:

$$\begin{aligned} \frac{\partial \sigma}{\partial t} &= K_L V - f(\sigma V), \\ K_L \frac{\partial s}{\partial t} &= f(\sigma, V). \end{aligned} \tag{A3}$$

Here $s(\boldsymbol{x},t)$ is slip at position $\boldsymbol{x}$ and time t, $\sigma(\boldsymbol{x},t)$ is shear stress, K_L is the self-interaction or "stress drop stiffness" and $f[\sigma,V]$ is the *stress dissipation function* (RUNDLE *et al.*, 2002). For example, the "Amontons" or Coulomb friction law, having a sharp failure threshold, can be written in the form (2) using a Dirac delta function:

$$\frac{\partial s}{\partial t} = \frac{\Delta \sigma}{K_L} \delta(t - t_F), \tag{A4}$$

where the stress drop $\Delta\sigma = \sigma - \sigma^R(V)$ and $\sigma^R(V)$ is the velocity-dependent residual stress. For laboratory experiments, K_L is the{machine + sample}stiffness, and for simulations, represents the stiffness of a coarse-grained element of the fault of scale size L. $\delta(\,)$ is the Dirac delta, and t_F is any time at which $\sigma(\boldsymbol{x},t_F) = \sigma^F(V)$. Both σ^F and σ^R can also be parameterized as functions of the normal stress χ by means of coefficients of static μ_S and ("effective") kinetic μ_K coefficients of friction, $\sigma^F = \mu_S \chi$, $\sigma^R = \mu_K \chi$.

In recent work (RUNDLE *et al.*, 2002), we have introduced another parameter α, which allows for stable stress-dependent aseismic sliding. The process described by α is seen in laboratory friction experiments (TULLIS, 1996), and is expressed by a generalization of equation (A4):

$$\frac{\partial s}{\partial t} = \frac{\Delta \sigma}{K_L} \{\alpha + \delta(t - t_F)\}. \tag{A5}$$

We found that the parameter α, which can be fixed either through laboratory experiments or through field observations (TULLIS, 1996; DENG and SYKES, 1997), acts to smooth the stress field a fault when $\alpha > 0$, and to roughen the fault stress field when $\alpha < 0$.

In the model results that we describe here, we further generalize (A5) to include an additional term which depends on rate of stress increase:

$$\frac{\partial s}{\partial t} = \frac{\Delta \sigma}{K_T} \left\{ \alpha + \delta(t - t_F) + \beta \delta\left(\frac{\partial \sigma}{\partial t} - \eta \right) \right\}, \tag{A6}$$

where β is a constant with appropriate units (stress/time^2) and η is a critical stress rate. Here K_T represents the total spring constant associated with a fault segment. The last term can be considered to be parameterization of effects associated with a dynamic stress intensity factor. It is known that stress rate effects are important in the process of dynamic fracture, such as might be expected during an earthquake. For example, the stress intensity factor K_{I} for mode I tensile fracture is thought to be of the form:

$$K_{ID} = K_{ID}\left(\frac{\partial\sigma}{\partial t}\ ,\ T\right), \tag{A7}$$

where T is temperature. More specifically, for a crack propagating at velocity v, it has been proposed that the time-dependent dynamic stress intensity factor $K_D(t)$ is of the general form:

$$K_D(t) = k(v)\ K_D(0) = k(v)\ K_S, \tag{A8}$$

where K_S is the static stress intensity factor. While not of the exact form of either equation (A7) or (A8), equation (A6) is an expression of the idea that the onset of earthquake sliding depends on the stress rate through a critical threshold value η.

In the simulations described below, we implement equation the physical process described by equation (A6) in our Virtual California CA simulations as follows. We define the Coulomb Failure Function $CFF(\boldsymbol{x},t)$:

$$CFF(\boldsymbol{x},t)\ =\ \sigma(\boldsymbol{x},t)\ -\ \mu_S\ \chi(\boldsymbol{x},t), \tag{A9}$$

According to the first term in the equation, stable slip can occur with amplitude proportional to α for nonzero $\Delta\sigma$. In addition, according to the second term, unstable failure of a fault occurs of when $CFF(\boldsymbol{x},t) = 0$. To implement a failure mechanism in a simple way that demonstrates physics similar to the third term, we allow unstable slip of amplitude:

$$\frac{\Delta\sigma}{K_T}\ =\ \frac{\sigma(\boldsymbol{x},t)\ -\ (\mu_S\ -\ \mu_K)\,\chi(\boldsymbol{x},t)}{K_T}, \tag{A10}$$

when the condition:

$$-\frac{\partial}{\partial t}\ \mathrm{Log}\,\{\,CFF(\boldsymbol{x},t)\,\}\ >\ \eta \tag{A11}$$

where typically $0 < \eta < 1$, or in discrete terms:

$$\frac{CFF(\boldsymbol{x},t)\ -\ CFF(\boldsymbol{x},t+\Delta t)}{CFF(\boldsymbol{x},t)}\ >\ \eta. \tag{A12}$$

In equation (A12), we interpret Δt as being the time since the beginning of the earthquake at time t. Implicitly, it is assumed in (A6), (A11) and (A12) that:

$$\eta \gg \left.\frac{\partial \sigma(\boldsymbol{x},t)}{\partial t}\right|_{\text{Interseismic}} = -\int d\boldsymbol{x}_k \; T_{ij}^{kl}(\boldsymbol{x}-\boldsymbol{x}')\, V_l(\boldsymbol{x}'), \tag{A13}$$

i.e., that the η-value for stress-rate triggering is much larger than the stress rate characterizing interseismic stress accumulation.

References

Barnhard, T. and Hanson, S., compilers (2000–2001 accessed), *Fault parameters used in USGS 1996 Seismic Hazard maps*, http://eqhazmaps.usgs.gov/html/faults.html.

Council of the National Seismic System, in association with the Northern California Earthquake Data Center (2000–2001 accessed), *CNSS earthquake catalog*, http://quake.geo.berkeley.edu/cnss/catalog-search.html.

Deng, J. S. and Sykes, L. R. (1997), *Evolution of the stress field in southern California and triggering of moderate-size earthquakes: A 200-year perspective*, J. Geophys. Res. *102*(B5), 9859–9886.

Dieterich, J. S. (1979), *Modeling of rock friction. 1. Experimental results and constitutive equations*, J. Geophys. Res. *84*, 2161–2168.

Evans, M., Hastings, N., and Peacock, B., *Statistical Distributions* (Wiley Interscience, New York. 1993).

Jaumé, S. C. and Sykes, L. R. (1996), *Evolution of moderate seismicity in the San Francisco Bay region, 1850 to 1993: Seismicity changes related to the occurrence of large and great earthquakes*, J. Geophys. Res. *101*(B1), 765–789.

Kanamori, H. and Anderson, D. L. (1975), *Theoretical basis of some empirical relations in seismology*, Bull. Seismol. Soc. Am. *65*, 1073–1095.

Karner, S. L. and Marone, C., *Effects of loading rate and normal stress on stress drop and stick-slip recurrence interval*, pp. 187–198 in (Rundle J. B., Turcotte, D. L., and Klein, W., eds.) *GeoComplexity and the Physics of Earthquakes*, Geophysical Monograph 120 (American Geophysical Union, Washington, D.C., 2000).

Matthews, M. V., Ellsworth, W. L., and Reasenberg, P. A. (2002), *A Brownian model for recurrent earthquakes*, Bull. Seismol. Soc. Am. *92*, 2233–2250.

Petersen, M. D. and Wesnousky, S. G. (1994), *Fault slip rates and earthquake histories for active faults in Southern California*, Bull. Seismol. Soc. Am. *84*(5), 1608–1649.

Petersen, M. D., Bryant, W. A., Cramer, C. H., Cao, T., Reichle, M., Frankel, A. D., Lienkaemper, J. J., McCrory, P. A., and Schwartz, D. P. (1996), *Probabilistic seismic hazard assessment for the state of California*, USGS Open-File Report 96–706, U. S. Govt. Printing Office.

Rangarajan, G. and Ding, M. (2000), *First passage time problem for biased continuous-time random walks*, Fractals *8*, 139–145.

Rikitake, T., *Earthquake Forecasting and Warning* (Center for Acad. Publ. Japan, D. Reidel, Hingham, MA, USA. 1982).

Rundle, J. B. (1988), *A physical model for earthquakes, 2, Application to southern California*, J. Geophys. Res. *93*, 6255–6274.

Rundle, P. B., Rundle, J. B., Tiampo, K. F., Sá Martins, J. S., McGinnis, S., and Klein, W. (2001), *Nonlinear network dynamics on earthquake fault systems*, Phys. Rev. Lett. *87*(14), Art. No. 148501 (1–4).

Rundle, J. B., Tiampo, K. F., Sá Martins., and Klein, W. (2002), *Self-organization in leaky threshold systems: The influence of near mean field dynamics and its implications for earthquakes, neurobiology and forecasting,* Proc. Nat. Acad. Sci. USA, *99*, Supplement 1, 2514–2521.

Rundle, J. B., Rundle, P. B., Donnellan, A., and Fox, G. C. (2004), *Gutenberg-Richter statistics in topologically realistic system-level earthquake stress-evolution simulations*, Earth Planets Space *56*, 761–771.

Rundle, J. B., Rundle, P. B., Donnellan, A., Turcotte, D.L., Shcherbakov, R. Li, P., Malamud, B. D., Grant, L. B., Fox, G. C., McLeod, D., Yakovlev, G., Parker, J., Klein, W., and Tiampo, K. F.

(2006, in press), A simulation-based approach to forecasting the next great San Francisco earthquake, Proc. Nat. Acad. Sci.

SCHWARTZ, D. P. and COPPERSMITH, K. J. (1984), Fault Behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones, J. Geophys. Res. *89*, 5681–5698.

SOUTHERN CALIFORNIA EARTHQUAKE DATA CENTER (2001 accessed), *Clickable Fault Map of Southern California*, http://www.data.scec.org/faults/faultmap.html.

THATCHER, W. (1975), *Strain accumulation and release mechanism of 1906 San Francisco Earthquake*, J. Geophys. Res. *80*(35), 4862–4872.

TULLIS, T. E. (1996), *Rock friction and its implications for earthquake prediction examined via models of Parkfield earthquakes*, Proc. Nat. Acad. Sci. USA *93*, 3803–3810.

TURCOTTE, D.L., *Fractals and Chaos in Geology and Geophysics*, (Cambridge University Press, Cambridge, UK. 1997).

WALD, D. J. and HEATON, T. H. (1994), *Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake*, Bull. Seismol. Soc. Am. *84*(3).

WARD, S.N. (1996), *A synthetic seismicity model for southern California: Cycles, probabilities and hazard*, J. Geophys. Res. *101*, 22393–22418.

WARD, S. N. (2000), *San Francisco Bay Area earthquake simulations: A step toward a standard physical earthquake model*, Bull. Seismol. Soc. Am. *90*(2), 370–386.

WARD, S.N. and GOES, S.D.B. (1993), *How regularly do earthquakes recur?-A synthetic seismicity model for the San Andreas fault*, Geophys. Res. Lett. *20*, 2131–2134.

WELLS, D. L. and COPPERSMITH, K. J. (1994), *New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement*, Bull. Seismol. Soc. Am. *84*(4), 974–1002.

WORKING GROUP ON CALIFORNIA EARTHQUAKE PROBABILITIES (1999), *Earthquake Probabilities in the San Francisco Bay Region: 2000 to 2030 – A Summary of Findings*, USGS Open-File Report 99–517, U. S. Govt. Printing Office.

(Received December 30, 2004, revised August 10, 2005, accepted September 5, 2005)
Published Online First: August 31, 2006