A simulation-based approach to forecasting the next great San Francisco earthquake

John B. Rundle [1,2*], Paul B. Rundle [1], Andrea Donnellan [2], Donald L. Turcotte [3], Robert Shcherbakov [1], Peggy Li [4], Bruce D. Malamud [5], Lisa B. Grant [6], Geoffrey C. Fox [7], Dennis McLeod [8], Gleb Yakovlev [1], Jay Parker [4], William Klein [9], and Kristy F. Tiampo [10]

[1] Center for Computational Science and Engineering, University of California, Davis, CA 95616, USA

[2] Earth & Space Sciences Division, Jet Propulsion Laboratory, Pasadena, CA 91125, USA

[3] Department of Geology, University of California, Davis, CA 95616, USA

[4] Exploration Systems Autonomy Section, Jet Propulsion Laboratory, Pasadena, CA 91125, USA

[5] Department of Geography, Kings College London, WC2R 2LS, UK

[6] Department of Environmental Health, Science & Policy, University of California, Irvine, CA 92697, USA

[7] Departments of Computer Science and Physics, School of Informatics, Indiana University, Bloomington, IN 47405, USA

[8] Department of Computer Science, University of Southern California, Los Angeles, CA 90089, USA

[9] Department of Physics, Boston University, Boston, MA 02215, USA

[10] Department of Earth Sciences, Biological & Geological Sciences, University of Western Ontario, London, ON N6A 5B7, Canada

[*] To whom correspondence should be addressed; E-mail: rundle@cse.ucdavis.edu.

In 1906 the great San Francisco earthquake and fire destroyed much of the city. As we approach the 100 year anniversary of this event, a critical concern is the hazard posed by another such earthquake. In this paper we examine the assumptions presently used to compute the probability of occurrence of these earthquakes. We also present the results of a numerical simulation of interacting faults on the San Andreas system. Called *Virtual California,* this simulation can be used to compute the times, locations and magnitudes of simulated earthquakes on the San Andreas fault in the vicinity of San Francisco. Of particular importance are new results for the statistical distribution of interval times between great earthquakes, results that are difficult or impossible to obtain from a purely field-based approach.

The great San Francisco earthquake (18 April 1906) and subsequent fires killed more than 3,000 persons, and destroyed much of the city leaving 225,000 out of 400,000 inhabitants homeless. The 1906 earthquake occurred on a 470 km segment of the San Andreas fault that runs from the San Juan Bautista north to Cape Mendocino

(Fig. 1) and is estimated to have had a moment magnitude $m \approx 7.9$ (1). Observations of surface displacements across the fault were in the range $2.0 - 5.0$ m (2). As we approach the hundredth anniversary of the great San Francisco earthquake, a timely question is the extent of the hazard posed by another such event, and how this hazard may be estimated.

The San Andreas fault is the major boundary between the Pacific and North American plates, which move past each other at an average rate of 49 mm yr^{-1} (3), implying that to accumulate $2.0 - 5.0$ m of displacement, $40 - 100$ years are needed. One of the simplest hypotheses for the recurrence of great earthquakes in the San Francisco area is that they will occur at approximately these $40 - 100$ year time intervals. This would indicate that the next earthquake may be imminent. However, there are two problems with this simple "periodic" hypothesis. The first is that it is now recognized that only a fraction of the relative displacement between the plates occurs on the San Andreas fault proper. The remaining displacement occurs on other faults in the San Andreas system, which in northern California is primarily in the east San Francisco Bay region, on the Hayward and Calaveras faults (see Fig. 1). Hall et al. (4) concluded that the mean displacement rate on just the northern part of the San Andreas Fault is closer to 24 mm yr^{-1}. With the periodic hypothesis this would imply recurrence intervals of 80 to 200 years.

The second and more serious problem with the periodic hypothesis involves the existence of complex interactions between the San Andreas Fault and other adjacent faults. It is now recognized (5-7) that these interactions lead to chaotic and complex non-periodic behavior so that exact predictions of the future evolution of the system are

not possible. Only probabilistic hazard forecasts can be made. For the past fifteen years a purely statistical approach has been used by the Working Group on California Earthquake Probabilities (WGCEP) (8-11) to make risk assessments for northern California. Their statistical approach is a complex, collaborative process that uses observational data describing earthquake slips, lengths, creep rates and other information on regional faults as inputs to a San Francisco Bay Regional fault model. Using their forecast algorithm, the WGCEP (11) found that the conditional probability for earthquakes having $M \geq 6.7$ during the 30 year period 2002–2031 is 18.2%.

As described in the WGCEP (11) report, the critical assumption in computing the hazard probability is the choice of probability distribution, or *renewal model*. In their study, distributions the WGCEP (11) utilized include the Brownian passage time (BPT); the log normal; the Poisson; and the empirical model (a variation on the Poisson model). The means and standard deviations of the distributions for event times on the fault segments were constrained by geological and seismological observations.

Virtual California

In this paper, we present the results of a topologically realistic numerical simulation of earthquake occurrence on the San Andreas fault in the vicinity of San Francisco. This simulation, called *Virtual California,* includes fault system physics such as the complex elastic interactions between faults in the system, as well as friction laws developed with insights from laboratory experiments and field data. Simulation-based approaches to forecasting and prediction of natural phenomena have been used with considerable success for weather and climate. When carried out on a global scale these simulations are referred to as *General Circulation Models* (12,13).

Turbulent phenomena are represented by parameterizations of the dynamics, and the equations are typically solved over spatial grids having length scales of tens to hundreds of kilometers. Although even simple forms of the fluid dynamics equations are known to display chaotic behavior (5), general circulation models have repeatedly shown their value. In many cases ensemble forecasts are carried out, which use simulations computed using multiple models to test the robustness of the forecasts.

The *Virtual California* simulation, originally developed by Rundle (14), includes stress accumulation and release, as well as stress interactions between the San Andreas and other adjacent faults. The model is based on a set of mapped faults with estimated slip rates, prescribed long term rates of fault slip, parameterizations of friction laws based on laboratory experiments and historic earthquake occurrence, and elastic interactions. An updated version of *Virtual California* (15-17) is used in this paper. The faults in the model are those that have been active in recent geologic history. Earthquake activity data and slip rates on these model faults are obtained from geologic databases of earthquake activity on the northern San Andreas fault. A similar type of simulation has been developed by Ward and Goes (18) and Ward (19). A consequence of the size of the fault segments used in this version of *Virtual California* is that the simulations do not generate earthquakes having magnitudes less than about $m \approx 5.8$.

Virtual California is a backslip model - the loading of each fault segment occurs due to the accumulation of a slip deficit at the prescribed slip rate of the segment. The vertical rectangular fault segments interact elastically, the interaction coefficients are computed by means of boundary element methods (20). Segment slip and earthquake initiation is controlled by a friction law that has its basis in laboratory-derived physics

(17,). Onset of initial instability is controlled by a static coefficient of friction. Segment sliding, once begun, continues until a residual stress is reached, plus or minus a random overshoot or undershoot of typically 10%. Onset of instability is also possible by means of a stress-rate dependent effect, in that segment sliding can initiate if stress on a segment increases faster than a prescribed value due to failure of a nearby segment. Finally, the friction law used in *Virtual California* also includes a term that promotes a small amount of stable segment sliding as stress increases. This latter term has been shown to promote stress-field smoothing along neighboring segments, offsetting the stress-roughening effects of increasing fault complexity, and allowing larger earthquakes to occur. To prescribe the friction coefficients we use historical earthquakes having moment magnitudes $m \geq 5.0$ in California during the last ~ 200 years (17).

The topology of *Virtual California* is shown in Fig. 1 superimposed on a LandSat image. The 650 fault segments are represented by red, blue, and yellow lines. The combined blue and yellow lines represent the San Andreas fault, stretching from the Salton trough in the south to Cape Mendocino in the north. The yellow line represents the "San Francisco section" of the San Andreas fault, about 250 km in length, which is the section of the fault whose rupture would be strongly felt in San Francisco and is considered in this paper.

Our goal is to forecast waiting times until the next great earthquake on the yellow section of the fault for two minimum magnitudes: (i) $m_{SF} = 7.0$ and (ii) $m_{SF} = 7.3$. It should be emphasized that the magnitude m_{SF} is based only on the slip of the earthquake over the San Francisco section of the San Andreas fault. An earthquake

with a particular value of m_{SF} might actually have a "total" magnitude $m > m_{SF}$ if the rupture includes segments outside the San Francisco section, so that the magnitude m is based on the extent of the entire rupture. We also note that Virtual California does produce earthquakes on the northern San Andreas fault having magnitudes $m \sim 7.8$ to 7.9, but these seem to occur at intervals of ~ 700 years or longer. We have generally found that the complex fault interactions in this region tend to inhibit the occurrence of great earthquakes on the northern San Andreas fault, representing interesting physics that needs further study.

Using standard seismological relationships (21), we estimate that an earthquake having $m_{SF} = 7.0$ with an average slip of 4 m and a depth of 15 km would rupture a 20 km length of fault. With similar conditions, an earthquake having $m_{SF} = 7.3$ would rupture a 66 km length of fault. Earthquakes like these would produce a considerable damage, destruction, and injury in San Francisco.

Using *Virtual California*, we advance our model in 1 year increments, and simulate 40,000 years of earthquakes on the entire San Andreas Fault system. We note that although the average slip on the fault segments and the average recurrence intervals are tuned to match the observed averages, the variability in the simulations is a result of the fault interactions. Slip events in the simulations display highly complex behavior, with no obvious regularities or predictability.

In Fig. 2, we show examples of the distribution of earthquakes on the "San Francisco Section" of the San Andreas fault for a 3000 year period. The left panel shows the slip in each earthquake as a function of distance along the fault from Fort

Ross (FR) in the north to San Juan Bautista (SJB) in the south. The center panel shows the moment magnitudes associated with each of the events at left. The frequency-magnitude distribution of earthquakes for a 40,000 year simulation of events on this section of fault, including the 3000 year interval illustrated, are shown at right.

One output of our simulations is the distribution of surface displacements caused by each model earthquake. Synthetic aperture radar interferometry (InSAR) is routinely used to obtain the coseismic displacements that occur after earthquakes (22). The displacements associated with two sets of our model earthquakes are illustrated in Fig. 3 as interferometric patterns. Each complete interferometric fringe color cycle corresponds to a displacement along the line-of-sight to the hypothetical spacecraft of 56 mm.

Earthquake Risk

A quantitative output of our simulations is the statistical distribution of time intervals between successive great earthquakes on a given fault or group of faults. For the northern section of the San Andreas fault near San Francisco, this distribution is required if the risk of future earthquakes on the fault is to be specified. We associate the properties of this distribution directly with the elastic interactions between faults, which are an essential feature of our model. Current estimates of risk are based on the observed statistics of fault intervals. However, Savage (23) has argued convincingly that actual sequences of earthquakes on specified faults are not long enough to establish the statistics of interval times with the required reliability.

In the WGCEP (11) report, several probability distributions are proposed to

describe the statistics of failure of a single geological fault segment. The two most frequently used are the Brownian passage time (BPT) (24) and the log normal (LN) (25). The parameters of these distributions, the mean μ and standard deviation σ for each geological segment are found from field data by procedures described in WGCEP (11). Once the statistical distribution for a single geological segment is defined, the distribution appropriate for rupture of multiple geological segments is computed by combining the single-segment probabilities. A critical assumption made by the WGCEP (9) in this procedure is that geological fault segments should be statistically independent and uncorrelated. This assumption is not satisfied by *Virtual California* earthquakes, since the elastic interactions allow stress changes on one fault to alter the stress level and failure time of other faults. Our numerical simulations using *Virtual California* allow us to measure both single fault segment statistics, as well as multiple segment statistics for arbitrary combinations of faults, fault segments, or fault systems. By the nature of the simulation procedure, the multiple segment probabilities include the effects of fault interactions. We illustrate this approach by using numerical simulations to obtain interval statistics for a suite of synthetic earthquakes on the San Francisco section of the San Andreas fault over 40,000 years. We then compare these statistics to curves computed using proposed *a priori* probability distributions.

We consider earthquakes on the section of the northern San Andreas fault shown in yellow in Fig. 1. Over the 40,000 year simulation, we obtained 395 simulated $m_{SF} \geq 7.0$ events having an average recurrence interval of 101 years, and 159 $m_{SF} \geq 7.3$ events having an average recurrence interval of 249 years. From the simulations, we measured the distributions of inter-event time intervals *t* between great earthquakes on the San Francisco segment. The time t is defined as the time interval

between two successive great earthquakes.

A second important distribution that we will consider is the distribution of waiting times Δt until the next great earthquake, given that the time elapsed since the most recent great earthquake is t_0. If we take the time of the last great earthquake to be 1906 and the present to be 2005, we find for San Francisco $t_0 = 2005 - 1906 = 99$ years. The waiting time Δt is measured forward from the present, thus $t = t_0 + \Delta t$. We will express our results in terms of the cumulative conditional probability $P(t, t_0)$ that an earthquake will occur in the waiting time $\Delta t = t - t_0$ if the elapsed time since the last great earthquake is t_0 (26).

A probability distribution that was not considered by the WGCEP (9) is the Weibull distribution, which has been used widely in Japan (27-30). Here the fraction of the waiting times $P(t, 0)$ that are less than t can be expressed as

$$P(t) = 1 - \exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right], \qquad (1)$$

where β and τ are fitting parameters. Sieh et al. (31) fit this distribution to the interval times of great earthquakes on the southern San Andreas fault obtained from paleoseismic studies with $\tau = 166.1 \pm 44.5$ years and $\beta = 1.5 \pm 0.8$. In its extension to the cumulative conditional probability the Weibull distribution is given by (32)

$$P(t,t_0) = 1 - \exp\left[\left(\frac{t_0}{\tau}\right)^{\beta} - \left(\frac{t}{\tau}\right)^{\beta}\right]. \qquad (2)$$

Equation (2) specifies the cumulative conditional probability that an earthquake will have occurred at a time t after the last earthquake if the earthquake has not occurred by a time t_0 after the last earthquake.

We first consider the type of statistical forecast described in the WGCEP (9) report. In Fig. 4a, the solid blue line is the cumulative probability $P(t,0)$ that a simulated great $m_{SF} \geq 7.0$ earthquake will occur on the San Andreas Fault near San Francisco, at the time *t* following the last such great earthquake. For comparison, we plot three other cumulative probability distributions having the same mean μ = 101 years and standard deviation σ = 61 years as the simulation data.. The solid black line is the best-fitting Weibull distribution; the dashed line is the BPT distribution; and the dotted line is the LN distribution. For the Weibull distribution, these values of mean and standard deviation correspond to $\beta = 1.67$ and $\tau = 114$ years.

Fig. 4b shows the same type of conditional probability as computed by the WGCEP (9), obtained from the simulation data in Fig 4a.. The solid blue line is then the simulation-based conditional probability $P(t_o \leq t < t_o + 30 \text{ yrs} \mid t \geq t_o)$ that a magnitude $m_{SF} \geq 7.0$ event will occur in the next 30 years, given that it has not occurred during the time t_o since the last such event. For comparison, the solid black line is the corresponding conditional probability for the Weibull distribution; the dashed line is for the BPT; and the dotted line is for the LN.

From the results shown in Fig. 4, it can be seen that the Weibull distribution describes the simulation data substantially better than either the BPT or LN distributions. At least in *Virtual California*, we can conclude that among these three statistical distributions, the Weibull distribution is the preferred distribution to describe the failure of a group of fault segments interacting by means of elastic stress transfer.

Examples of cumulative conditional distributions of interval times are given in Fig. 5, for $m_{SF} \geq 7.0$ in Fig. 5a and for $m_{SF} \geq 7.3$ in Fig. 5b. The left-most curves in each figure, $P(t,0)$, are the curves that pass through $t = 0$. The left-most curve in Fig 5a is the same as the distribution of interval times given in Fig 4a. Also included are the Weibull distributions from Eq. (1) that best fit the data. For $m_{SF} \geq 7.0$ in Fig. 5a our best fit for the leftmost curve requires $\beta = 1.67$ and $\tau = 114$ years as described above, and for the left most curve Fig. 5b, our best fit requires $\beta = 2.17$ and $\tau = 289$ years.

We next determine the cumulative conditional probabilities that an earthquake will occur at a time t after the last earthquake if it has not occurred at a time t_0. We therefore remove interval times that are less than or equal to t_0 and plot the cumulative distribution of the remaining interval times. The resulting distributions $P(t,t_0)$ are given in Fig. 5a for $m_{SF} > 7.0$ with $t_0 = 25$, 50, 75, 100, 125, and 150 years and in Fig. 5b for $m_{SF} > 7.3$ with $t_0 = 50$, 100, 150, 200, 250, and 300 years. With the fitting parameters β and τ used to fit Eq. (1) to the cumulative distributions of waiting times $P(t)$, we again compare the predictions of the Weibull distribution for $P(t,t_0)$ from Eq. (2), the smooth curves, with data from our simulations in Fig. 5, the irregular curves.

The data given in Fig. 5 can also be used to determine the waiting times to the next great earthquake $\Delta t = t - t_0$ as a function of the time since the last great earthquake occurred t_0. This dependence is given in Fig. 6a for $m_{SF} \geq 7.0$ and in Fig. 6b for $m_{SF} \geq 7.3$. The green stars in Fig. 6 are the median waiting times Δt, $P(t_0 + \Delta t, t_0) = 0.5$, to the next great earthquake as a function of the time t_0 since the last great earthquake. These stars are the intersections of the dashed red lines with $P(t, t_0) = 0.5$ with the cumulative distributions in Fig. 5. Also given as circles in Fig. 6 are the waiting times for $P(t, t_0) = 0.25$ (lower limit of the yellow band) and for $P(t, t_0) = 0.75$ (upper limit of the yellow band). The dashed red lines are the forecasts of risk based on the Weibull distributions from Eq. (2).

Immediately after a great earthquake, e.g., in 1906, we have $t_0 = 0$ years. At that time, Figs. 5a and 6a indicate that there was a 50% chance of having an earthquake $m_{SF} \geq 7.0$ in the next $t = 90$ years, i.e., in 1996. Also at that time ($t_0 = 0$ years), there was a 50% chance of having an earthquake with $m_{SF} \geq 7.3$ in the next $t = 249$ years, as shown in Figs. 5b and 6b.

In 2006 it will have been 100 years since the last great earthquake occurred in 1906. The cumulative conditional distributions corresponding to this case have $t_0 = 100$ years. We see from Figs. 5a and 6a that there is a 50% chance of having a great earthquake ($m_{SF} \geq 7.0$) in the next $\Delta t = 45$ years ($t = 145$ years). This is the red star in Fig. 6a. It can also be seen that there is a 25% chance for such an earthquake in the next

$\Delta t = 20$ years ($t = 120$ years), and a 75% chance of having such an earthquake in the next $\Delta t = 80$ years ($t = 180$ years). During each year in this period, to a good approximation, there is a 1% chance of having such an earthquake. These figures are consistent with the information in Fig. 4b, which indicates a 30% chance of an $m_{SF} \geq 7.0$ earthquake during the period 2006-2036.

Similarly, Figs. 5b and 6b indicate that there is a 75% chance of having a great earthquake with $m_{SF} \geq 7.3$ in the next $\Delta t = 250$ years, a 50% chance in the next $\Delta t = 180$ years (the red star in Fig. 5b), and a 25% chance in the next $\Delta t = 75$ years. To a good approximation, there is a 0.3% chance of having such an earthquake during each year in this period.

We see from Figs. 3-6 that the Weibull distribution that fits the distribution of interval times also does an excellent job of also fitting the conditional probabilities and the waiting times. In both simulations and in our Weibull fit, the median waiting times systematically decrease with increases in the time since the last great earthquake. This is not the case for other distributions that provide a good fit to interval times (9). Our results therefore support the use of Weibull distributions to carry out probabilistic hazard analyses of earthquake occurrences.

Discussion

There are major differences between the simulation-based forecasts given in this paper, and the statistical forecasts given by the WGCEP (11). In our approach, it is not necessary to prescribe a probability distribution of inter-event times. The distribution of event intervals is obtained directly from simulations, which include the

physics of fault interactions and frictional physics. Since both methods use the same database for mean fault slip on fault segments, they give approximately equal mean inter-event times. The major difference between the two methods lies in the way in which inter-event times and probabilities for joint failure of multiple segments are computed. In our simulation approach, these times and probabilities come from the modeling of fault interactions through the inclusion of basic dynamical processes in a topologically realistic model. In the WGCEP (11) statistical approach, times and probabilities are embedded in the choice of an applicable probability distribution function, as well as choices associated with a variety of other statistical weighting factors describing joint probabilities for multi-segment events.

It should be remarked that Fig. 2 indicates that there is a difference between measurements of "earthquake recurrence of a certain magnitude" on a fault, and "earthquake recurrence at a site" on a fault. Specifically, the latter is the quantity that is measured by paleoseismologists, who would observe very different statistics on the earthquakes shown in Fig. 2 (left) if they made observations at the locations of 50 km, 100 km and 150 km from Fort Ross. The former includes any earthquakes that rupture any set of segments on the given section of fault.

A measure of the interval times is the coefficient of variation c_v of the distribution of values. The coefficient of variation is the ratio of the standard deviation to the mean $c_v \equiv \frac{\sigma}{\mu}$. For periodic earthquakes, we have $\sigma = c_v = 0$; for random (Poisson) distribution of interval times, we have $\sigma = \mu$ and $c_v = 1$. For our simulations of great earthquakes on the San Francisco section of the San Andreas fault,

we found that $c_v = 0.6$ for earthquakes having $m_{SF} \geq 7.0$, and $c_v = 0.48$ for earthquakes with $m_{SF} \geq 7.3$. These figures apply to any earthquakes on the fault between Fort Ross and San Juan Bautista, rather than at a point on the fault. As mentioned previously, Ward and Goes (18) also simulated earthquakes on the San Andreas fault system. Although the statistics of the simulated earthquakes produced by their *Standard Physical Earth Model* (SPEM) are similar to those produced by *Virtual California* , there are important differences between the two simulation codes. Whereas *Virtual California* involves rectangular fault segments in an elastic half space, SPEM is a plain strain computation in an elastic plate of thickness H. The friction laws used in the two simulations are also entirely different. Ward and Goes (18) obtained the statistical properties of earthquake interval times for the San Francisco section of the San Andreas fault, and found $c_v = 1.16$ for earthquakes having $m_{SF} \geq 7.0$, and $c_v = 0.54$ for earthquakes with $m_{SF} \geq 7.5$. The two simulations therefore predict different variability in the occurrence time intervals of the largest events.

It is also of interest to compare the simulation results with the available statistical distributions of interval times for the San Andreas fault. Paleoseismic studies of $m_{SF} = 7+$ earthquakes on the southern San Andreas fault at Pallett Creek by Sieh et al (31) indicate seven intervals with $\mu = 155$ years and $\sigma = 109$ years, hence $c_v = 0.70$. This figure, which is for a single site, should be compared with those above, which were obtained for a ~240 km long section of fault. A second example is the sequence of $m_{SF} = 6+$ earthquakes that have occurred on the Parkfield section of the San Andreas fault between 1857 and 2004. The seven intervals give $\mu = 24.5$ years and $\sigma = 9.25$ years, and $c_v = 0.38$. Again, this is at a single site on a fault.

The inter-event statistics we have obtained in our simulations include the essential features of the fault interactions and are influenced also by the quality and quantity of the data available to constrain the parameters for each fault segment in the model. If different slip rates are prescribed, the mean and median waiting times will change accordingly. However, the statistical distribution of waiting times is likely to remain unchanged.

The statistical distribution of waiting times has long been a subject of controversy. The Weibull distribution utilized here is one of a number of distributions previously proposed (32). If $\beta = 1$, the Weibull distribution of waiting times reduces to a Poisson distribution. Waiting times are independent of the time since the last earthquake (no memory). In the limit $\beta \to \infty$ the intervals are constant and earthquakes are periodic. We find $\beta = 1.67$ for $m_{SF} \geq 7.0$ and $\beta = 2.17$ for $m_{SF} \geq 7.3$. The larger earthquakes are more periodic and less random. The validity of the Weibull distribution places important constraints on future probabilistic earthquake hazard analyses.

In this paper we have examined the statistics of great earthquake occurrence on the northern San Andreas fault in the San Francisco bay region using numerical simulations. For previous estimates of hazard, only purely statistical estimates have been made. Our approach is analogous to the simulations used to forecast the weather. An example of the type of statement that can be made about the seismic hazard is: "There exists a 5% chance of an earthquake with magnitude $m \geq 7.0$ occurring on the San Andreas fault near San Francisco prior to 2009 and 55% chance by 2054". The

practical use of statements like this for hazard estimation using numerical simulations must be validated by more computations and observations.

Acknowledgements. This work has been supported by grant DE-FG02-04ER15568 from the US Department of Energy, Office of Basic Energy Sciences to the University of California, Davis (JBR, PBR), grant ATM 0327558 from the National Science Foundation (DLT, RS), and supported by grants from the Computational Technologies Program of NASA's Earth-Sun System Technology Office to the Jet Propulsion Laboratory, the University of California, Davis, and the University of Indiana (JBR, PBR, AD).

References Cited

1. Engdahl, ER and Villasenor, A. (2002) *International Handbook of Earthquake and Engineering Seismology*, W. H. K. Lee, H. Kanamori, P. C. Jennings, C. Kisslinger, eds. (Academic Press, Amsterdam) pp. 665–690.

2. Thatcher, W, (1975) *J. Geophys. Res.* **80**, 4862-4872.

3. Demets, C, Gordon, RG, Argus, DF and Stein, S (1994) *Geophys. Res. Lett.* **21**, 2191-2194.

4. Hall, NT, RH Wright, Clahan, KB (1999) *J. Geophys. Res.* **104**, 23,215-23,236.

5. Lorenz, EN, (1963) *J. Atmos. Sci.* **20**, 130-141.

6. Rundle, J.B. (1988) *J. Geophys. Res.*, **93**, 6237 - 6254.

7. Turcotte, DL (1997) *Fractals and Chaos in Geology and Geophysics*, 2nd ed. (Cambridge University Press, Cambridge).

8. Working Group on California Earthquake Probabilities (1988) USGS Open File Report 88-398.

9. Working Group on California Earthquake Probabilities (1990) USGS Circular 1053.

10. Working Group on California Earthquake Probabilities (1995) *Bull. Seism. Soc. Am.*, 85, 379-439.

11. Working Group on California Earthquake Probabilities (2003) USGS Open File Report 03-214.

12. Kodera, K, Matthes, K, Shibata, K, Langematz, U and Kuroda, Y (2003) *Geophys. Res. Lett.* **30**, Art. No. 1315.

13. Covey C, Achutarao KM, Gleckler, PJ, Phillips, TJ, Taylor, KE and Wehner, MF (2004) *Global Planet. Change*, **41**, 1-14

14. Rundle, JB (1988) *J. Geophys. Res.* **93**, 6255-6271.

15. Rundle, JB, Rundle, PB, Klein, W, Martins, JSS, Tiampo, KF, Donnellan, A, and Kellogg, LH (2002) *Pure Appl. Geophys.* **159**, 2357-2381.

16. Rundle, PB, Rundle, JB, Tiampo, KF, Martins, JSS, McGinnis, S, Klein, W (2001) Nonlinear network dynamics on earthquake fault systems, *Phys. Rev. Lett.* **8714**, Art. No. 148501.

17. Rundle, JB, Rundle, PB, Donnella,n A, and Fox, G (2004) *Earth Planets Space* **56**, 761-771.

18. Ward, SN and Goes, SDB (1993) *Geophys. Res. Lett.*, **20**, 2131-2134.

19. Ward, SN (2000) *Bull. Seismol. Soc. Amer.* **90**, 370-386.

20. Boundary Elements XXIV. (2002) 24th International Conference on Boundary

21. Kanamori, H and Anderson, DL (1975) *Bull. Seism. Soc. Am.* **65**, 1073-1096.

22. Massonnet D, Rossi, M, Carmona, C, Adragna, F, Peltzer, G, and Feigl, K (1993) *Nature* **364**, 138-142 (1993).

23. Savage, JC (1994) *Bull. Seism. Soc. Am.*, **84**, 219-221.

24. Matthews, MV, Ellsworth, WL, Reasenberg, PA (2002) *Bull. Seism. Soc. Am.*, **92**, 2233-2250.

25. Nishenko, SP and Buland, R (1987) *Bull. Seism. Soc. Am.*, **77**, 1382-1399.

26. Wesnousky, SG, Scholz, CH, Shimazaki, K and Matsuda, T (1984) , *Bull. Seismol. Soc. Amer.* **74**, 687-708.

27. Hagiwara, Y (1974) *Tectonophysics*, **23**, 313-318.

28. Rikitake, T (1976) *Tectonophysics*, **35**, 335-362

29. Rikitake, T, (1982) *Earthquake Forecasting and Warning* (D. Reidel, Dordrecht).

30. Utsu, T (1984) *Bull. Earthquake Res. Inst.*, **59**, 53-66.

31. Sieh, K, Stuiver, M and Brillinger, D. (1989) *J. Geophys. Res.* **94**, 603-623.

32. Sornette, D and Knopoff, L (1997) *Bull. Seismol. Soc. Amer.* **87**, 789-798.

Figure Captions.

Fig. 1. Faults segments making up Virtual California. The model has 650 fault segments, each approximately 10 km in length along strike and a 15 km depth. The yellow and blue segments make up the San Andreas fault. In this paper we consider earthquakes only on the yellow “San Francisco” segment of the San Andreas fault.

Fig. 2. Illustration of simulated earthquakes on the "San Francisco section" of the San Andreas fault. In the left panel, the slip is given as a function of the distance along the fault for each earthquake over a 3000 year period. The center panel shows the corresponding moment magnitude of each of the simulated earthquakes. The right panel gives the cumulative frequency-magnitude distribution during the entire 40,000 year simulation. The dashed line has a slope (b-value) of 1.

Fig. 3. Interferometric patterns of the coseismic deformations associated with two sets of model earthquakes. Each complete interferometric fringe color cycle corresponds to a displacement of 56 mm.

Fig 4. (a) The solid blue line is the simulation-based cumulative probability e that a great $m_{SF} \geq 7.0$ earthquake will occur on the San Andreas Fault near San Francisco at a time t years after the last great earthquake, just after the last great earthquake with $m_{SF} \geq 7.0$. For comparison, we plot three cumulative probability distributions having the same mean μ = 101 years and standard deviation σ = 61 years as the simulation data.. The solid black line is the best-fitting Weibull distribution; the dashed line is the Brownian passage time (BPT) distribution; and the dotted line is the log normal (LN) distribution. For the Weibull distribution, these values of mean and standard deviation

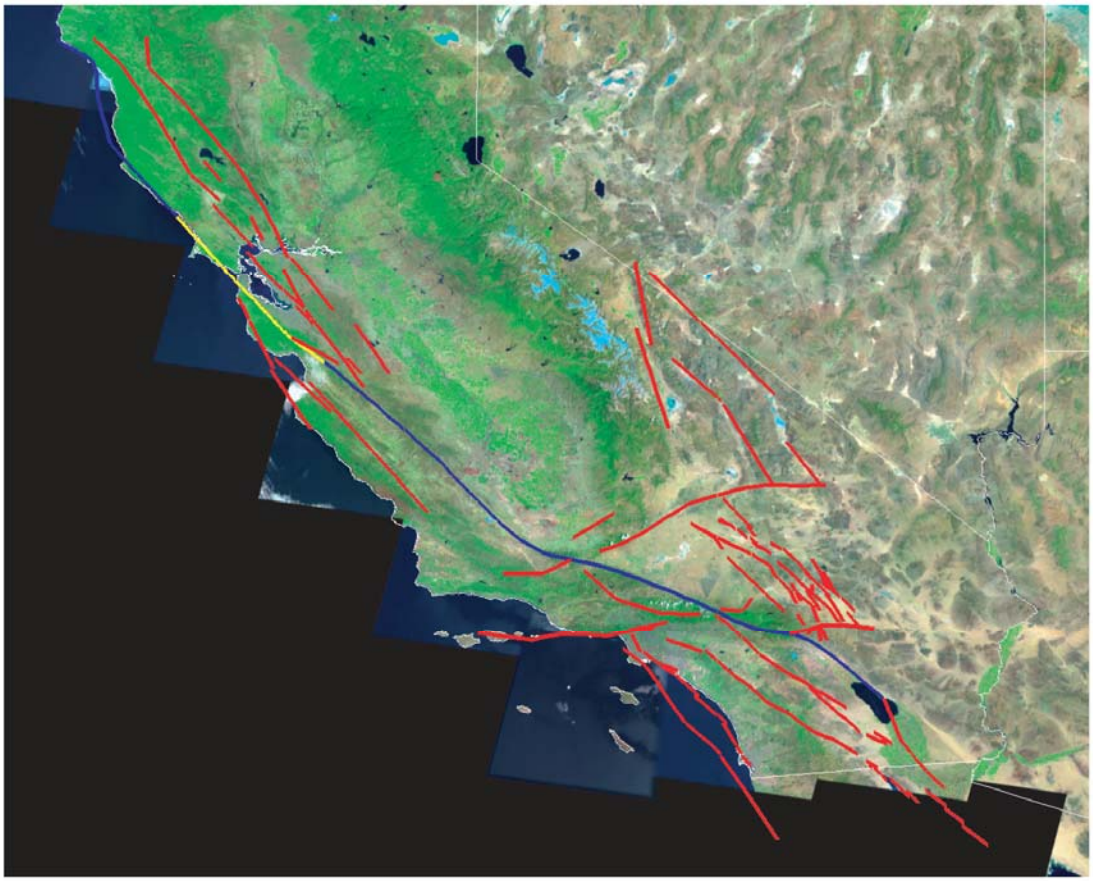
correspond to $\beta = 1.67$ and $\tau = 114$ years.

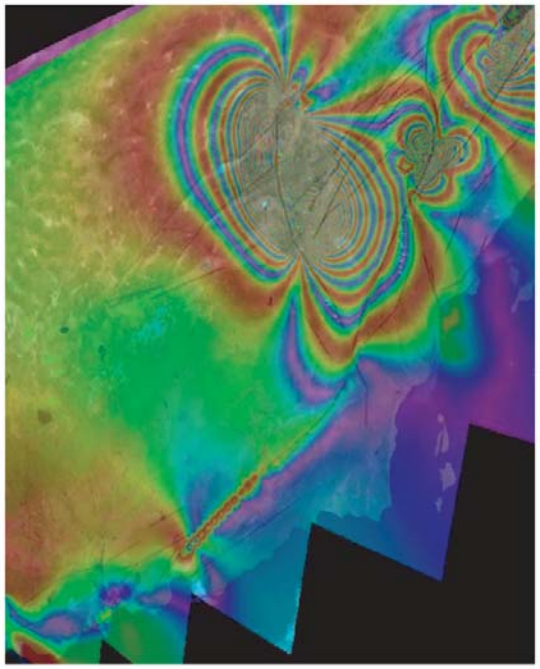
(b) The solid blue line is the conditional probability $P(t_o \le t < t_o + 30 \mid t \ge t_o)$ that a magnitude $m_{SF} \ge 7.0$ event will occur in the next 30 years, given that it has not occurred by a time t_o since the last such event. The black solid line is the corresponding conditional probability for the Weibull distribution; the dashed line is for the BPT; and the dotted line is for the LN.

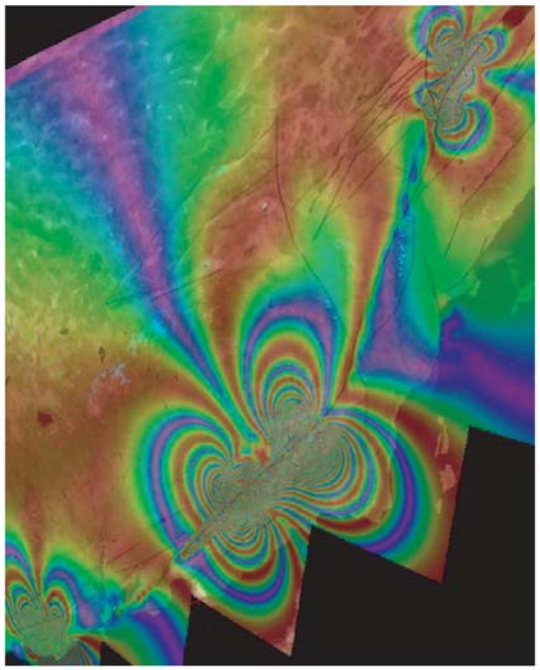
Fig. 5. (a) The conditional cumulative probability $P(t,t_0)$ that a great $m_{SF} \ge 7.0$ earthquake will occur on the San Andreas Fault near San Francisco at a time t years after the last great earthquake, if the last great earthquake occurred t_0 years ago in the past. Results are given for $t_0 = 0$, 25, 50, 75, 100, 125, and 150 years. Also included are the fits to the data of the Weibull distribution. First, the best fit of Eq. (1) to the complete distribution of interval times ($t_0 = 0$) is obtained taking $\beta = 1.67$ and $\tau = 114$ years as in Fig. 4a. These values are then substituted into Eq. (2) taking $t_0 = 25$, 50, ..., 150 years. The Weibull fits are shown as colored curves. (b) Results for $m_{SF} \ge 7.3$. In this case we take $t_0 = 0$, 50, 100, 150, 200, and 250 years. The best fit of Eq. (1) to the complete distribution of interval times ($t_0 = 0$) requires $\beta = 2.17$ and $\tau = 289$ years.

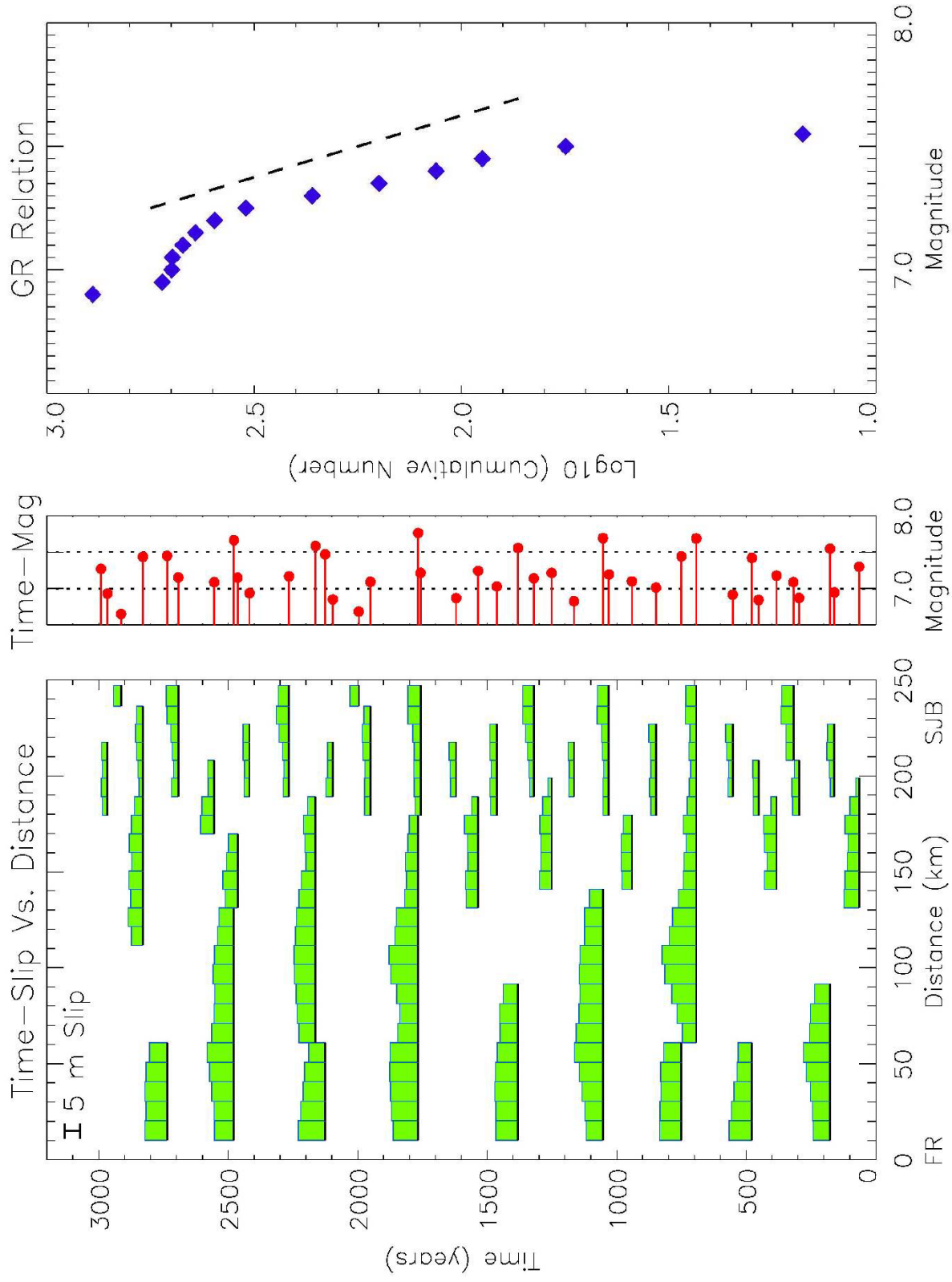
Fig. 6. The green stars (corresponding to the 50% probability of the distributions in Fig. 5) and the green solid line give the median waiting times until the next great earthquake as a function of the time t_0 since the last great earthquake. The red star is the median waiting time (50% probability) from today. The yellow band represents

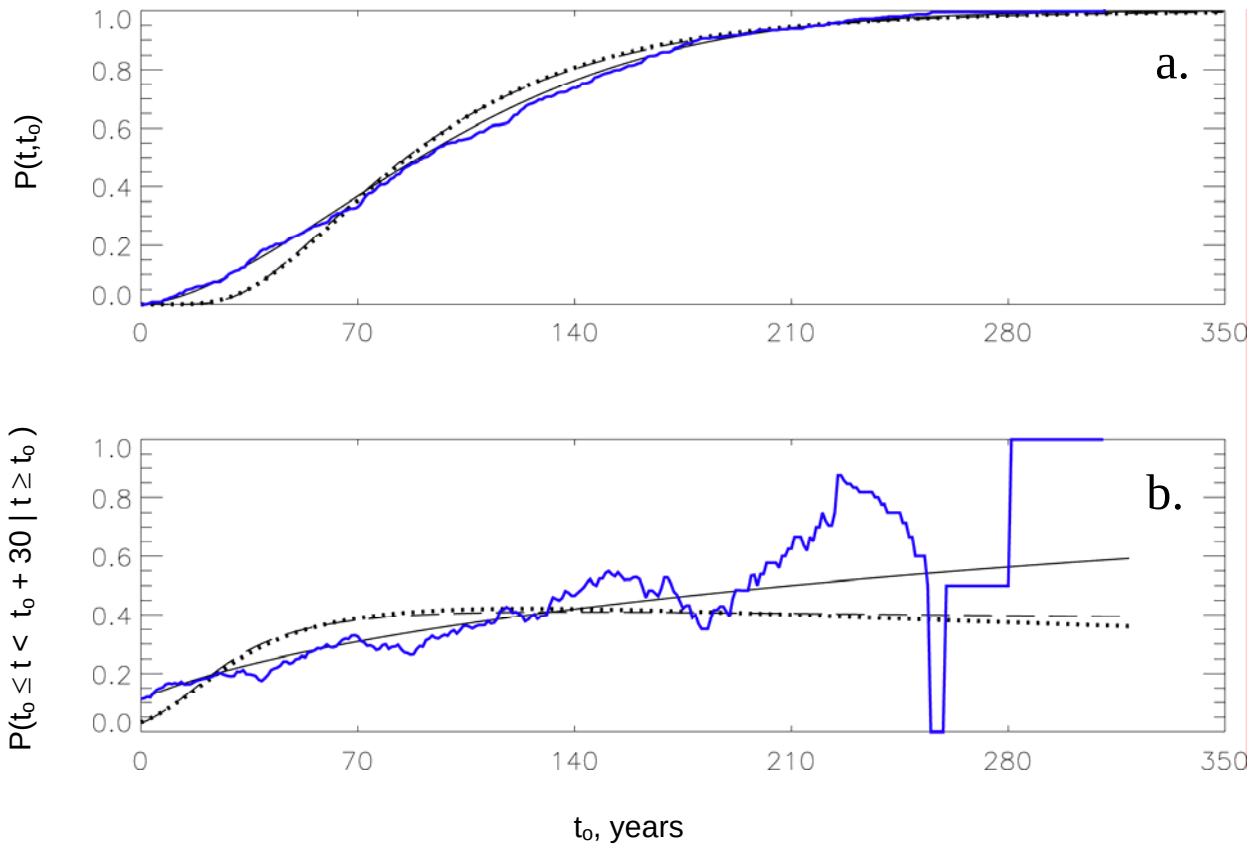
waiting times with 25% probability (lower edge of yellow band) to 75% probability (upper edge of yellow band). The dashed red lines are the forecast using the Weibull distribution in Eq. (2) (a) for $m_{SF} \geq 7.0$ and (b) for $m_{SF} \geq 7.3$.

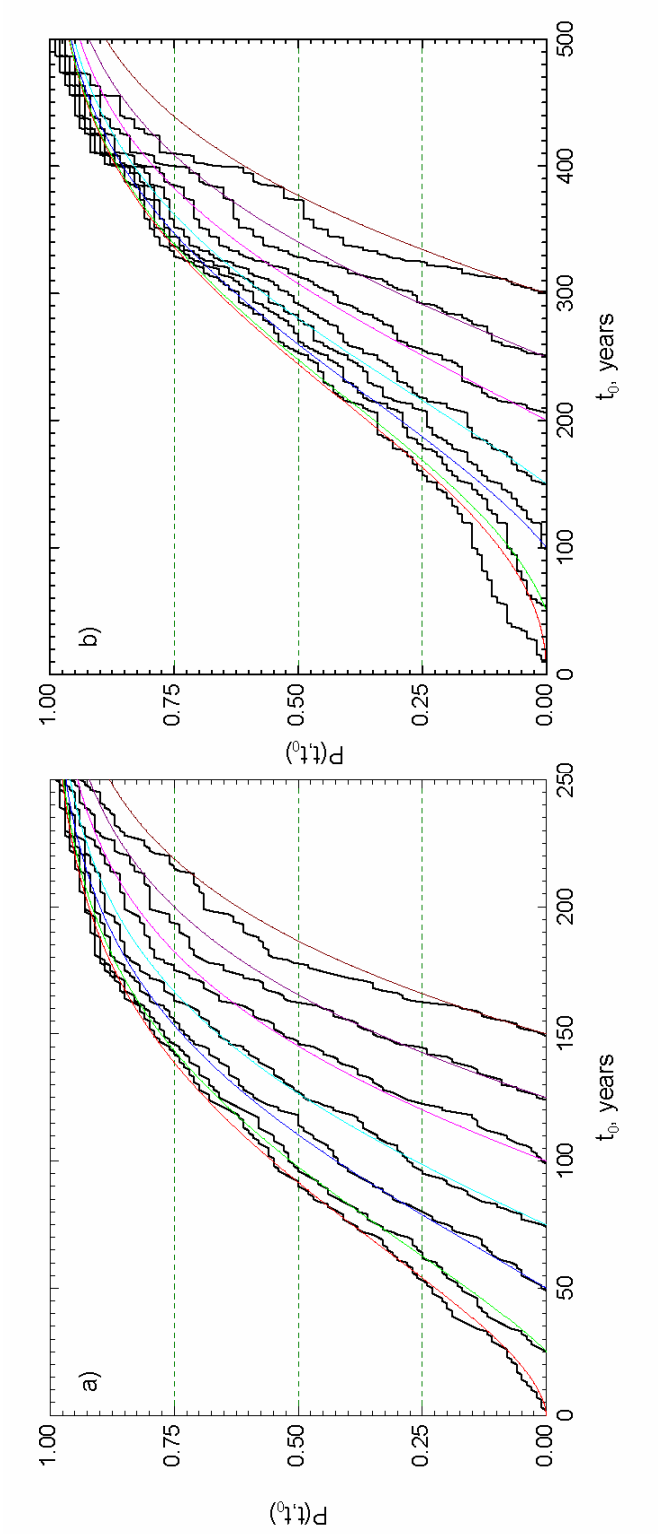


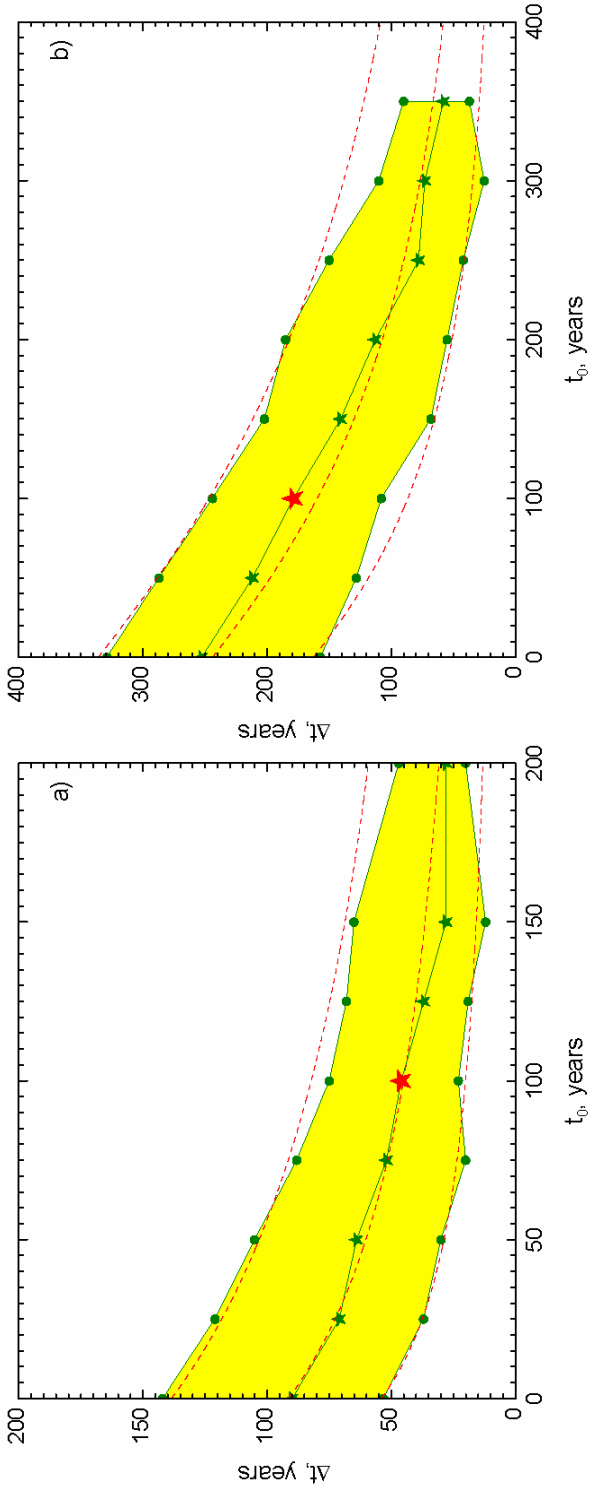





Time–Slip Vs. Distance
5 m Slip
Time (years)
Distance (km)
FR
SJB
Time–Mag
Magnitude
GR Relation
Log10 (Cumulative Number)
Magnitude


a.
b.
P(t,t$_o$)
P(t$_o$ ≤ t < t$_o$ + 30 | t ≥ t$_o$)
1.0
0.8
0.6
0.4
0.2
0.0
0
70
140
210
280
350
t$_o$, years


a)
b)
P(t,t0)
t0, years
0.00
0.25
0.50
0.75
1.00
0
50
100
150
200
250
300
400
500


a)
b)
Δt, years
t_0, years
0
50
100
150
200
300
400