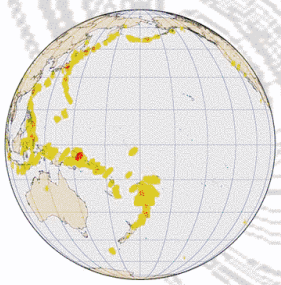
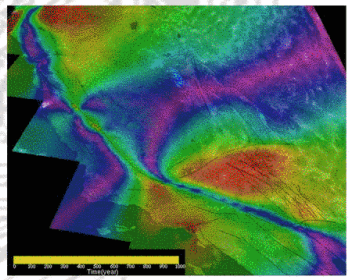
Relating Observable Earthquake Pattern Formation to Unobservable Earthquake Physics: Implications for Earthquake Forecasting

John B. Rundle
Center for Computational Science & Engineering
University of California, Davis



Global Hotspot Map



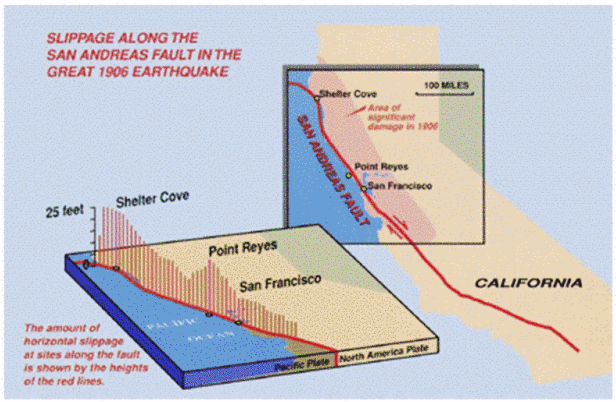
Virtual California Simulation
InSAR Surface Deformation

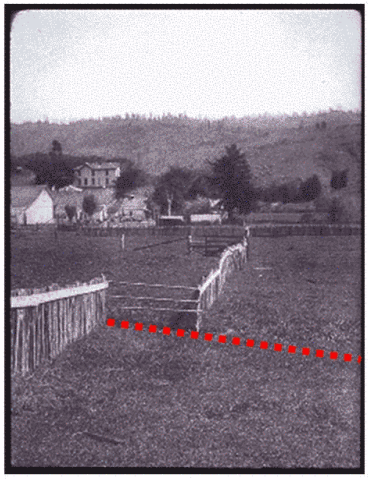
Jet Propulsion Laboratory
California Institute of Technology
QUAKESIM

Collaborators - All

Chien-chih Chen, Dept. of Earth Sciences, National Central University, Taipei, Taiwan
Andrea Donnellan, Division of Earth and Space Science, JPL
Geoffrey Fox, Dept. of Computer Science, University of Indiana
Robert Granat ,Data Understanding Systems Group, JPL
Lisa Grant, Dept. of Environmental Science, University of California, Irvine
James Holliday, Dept. of Physics and CSE, University of California, Davis
Peggy Li, Parallel Applications Technologies Group, JPL
Bill Klein, Dept. of Physics, Boston Univ.; and CNLS, LANL
Greg Lyzenga, Dept. of Physics, Harvey Mudd College
Jorge Martins, Universidad Fluminense, Niteroi, Brazil
Dennis McLeod, Dept. of Computer Science, USC
Gleb Morein, CSE, University of California, Davis
Kazuyoshi Nanjo, CSE, UC Davis and Tohoku University, Japan
Jay Parker, Satellite Geodesy and Geodetic Systems, JPL
Marlon Pierce, Community Grids Lab, University of Indiana
Paul Rundle, Department of Physics, Harvey Mudd College; and CSE, UC Davis
Robert Shcherbakov, CSE, University of California, Davis
Kristy Tiampo, Dept. of Geology, University of Western Ontario, London, Ontario
Terry Tullis, Dept. of Geology, Brown University
Don Turcotte, Dept. of Geology, University of California, Davis
Jordan Van Aalsburg, Dept. of Physics, University of California, Davis

Many Spatial Scales in Earthquakes are Strongly Coupled

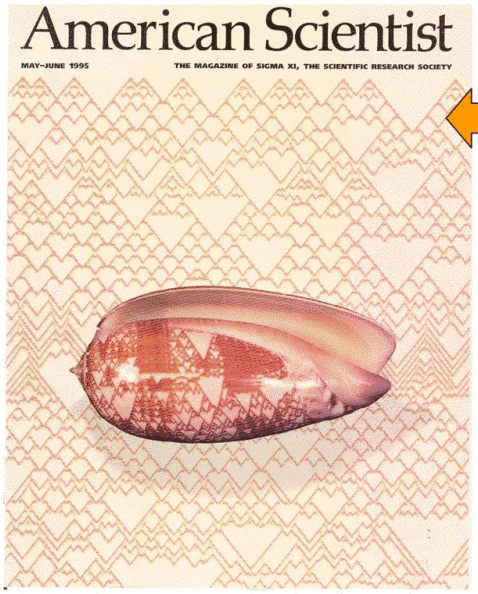




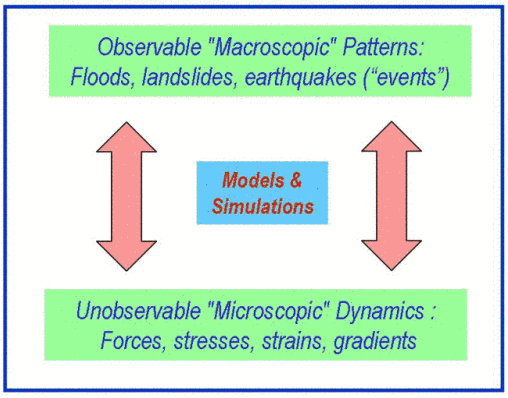
Wayne Thatcher, US Geological Survey

Three coupled spatial scales visible: 1. Fault length (~350 km); 2. Slip (~m); 3. Fault thickness (~cm)

Discovery in Science and Engineering Often Focuses on Understanding the Meaning of Space-Time Patterns



Example: The nonlinear dynamical systems that generate patterns and regularities in natural biosystems can be simulated in the computer as a means of understanding the important biological feedback loops.



Mean Field Slider Block Models

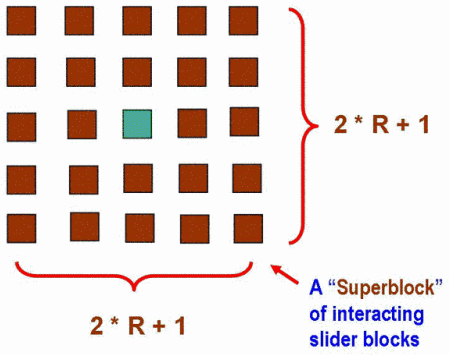
A simple model of an earthquake fault - CA Model (R Burridge and L Knopoff, BSSA, 1967; JBR & DD Jackson, BSSA, 1977)

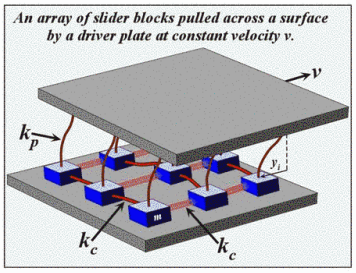
As in mean field Ising models, fluctuations in total energy diminish as the mean field regime is approached. We have used three methods to approach the mean field regime:

1. Set $K_C / K_L \rightarrow \infty$ in a nearest neighbor model
2. Let each block interact with q other blocks, keeping K_c = constant (flat interaction)
3. Use a dipole interaction K_c / r^3 between blocks

In all of these methods, the effective range of interaction $R \rightarrow \infty$. Relation between two flat q-models:

$$q_1 K_{C1} = q_2 K_{C2}$$





Flat interaction: In the figure above, the central block (green) interacts with

$$q = (2R + 1)^2 - 1 = 5^2 - 1 = 24$$

surrounding blocks (brown), each interaction having strength K_C.

Scaling in Mean Field Slider Block Models

Like mean field Ising models, mean field slider block models demonstrate scaling near a critical velocity $V = V_{SP}$.

At upper right, 18 million clusters in a 512 x 512 system produce a number-size relation:

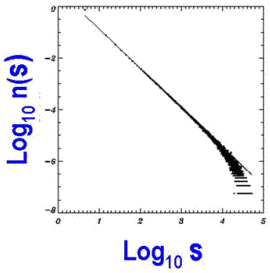
$$n(s) \sim \exp\{ - |V - V_{SP}| s^{\sigma} \} / s^{\tau - 1} \qquad (1)$$

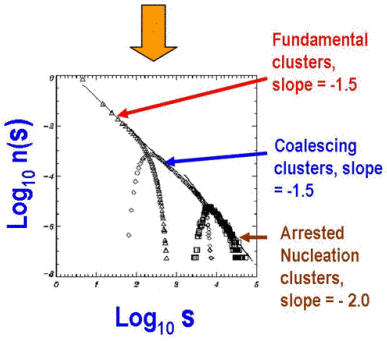
where:

$\sigma = 1$

$\tau = 2.5$

Careful analysis indicates that the scaling region is actually is a superposition of 3 separate scaling regimes (Anghel et al., Phys. Rev. E, in press, 2001), each of which obeys an equation like (1).

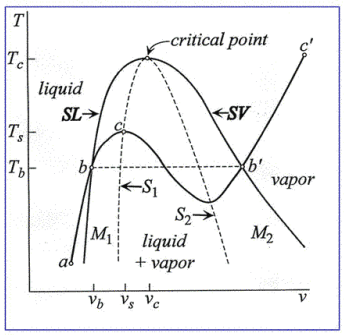


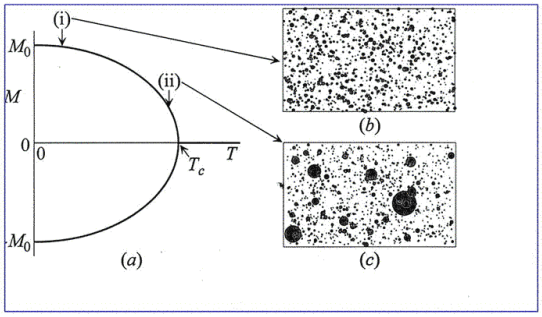


Universality and Critical Phenomena

Scaling and fractal distributions are now know to occur in all fields of science and have universal features. Chief among these features is proximity to a critical point of a phase diagram for the dynamical system, or proximity to a line of critical points such as a spinodal line.

At right is the phase diagram for a liquid-gas isotherms showing the critical point, the spinodal lines (dashed lines S_1 and S_2), coexistence curves (SL and SV), the Maxwell construction (b, b'), and a Van der Waals isotherm, or "loop" (a,b,c,b'c')





At left is the phase diagram for a magnetic system, showing the coexistence curves (solid line) and the critical point (Curie Temperature T_c)

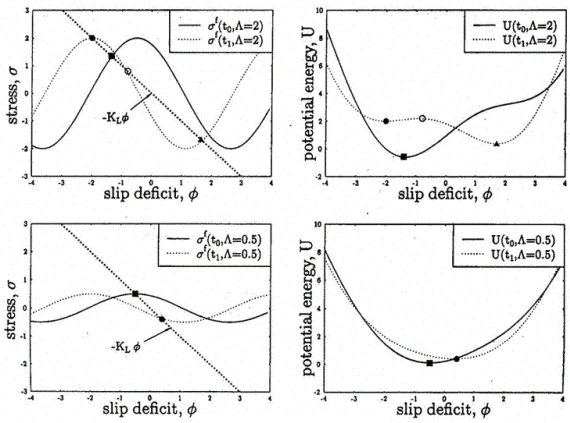
Traveling Density Wave Model of a Threshold System

JBR et al., Phys. Rev. Lett., 76, 4285, 1996.

Implies that earthquakes are an example of Phase Dynamics

$\Lambda>1$ Spiking

$\Lambda<1$ No Spiking



This model has the physics of a first order phase transition. It has a classical limit of stability, or spinodal. Here we show the mean field behavior of the model.

$$U[\phi] = \frac{1}{2} K_L \phi^2 - 2\gamma \cos\{\kappa(\phi + Vt)\}$$

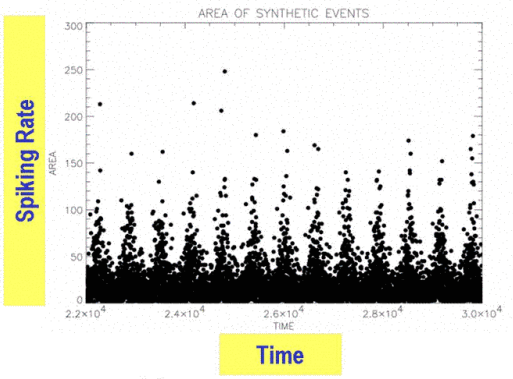
$$\frac{\partial \phi}{\partial t} = -\frac{\partial U}{\partial \phi} = -K_L \phi - 2\gamma\kappa \sin\{\kappa(\phi + Vt)\}$$

$$\Lambda \equiv \frac{2\gamma\kappa^2}{K_L}; \quad \phi \equiv s - Vt$$

Principle of Minimum Free Energy

Traveling Density Wave Model with Noise

Broad Dispersion in Noise Produces a Complex Spiking (Firing) Process



A more detailed TDW model with noise is needed to produce complex output. Here $T_{ij} = K_C / r_{ij}^3$ represents an exciting cell-cell interaction, and η is a noise. The average period τ of spiking is $\tau = \kappa / V$.

In this simulation, the noise η is uniformly distributed on the interval $(0,2\pi)$.

The lattice is 64 x 64 = 4096 sites.

$$\frac{\partial \phi_i}{\partial t} = - K_L \phi_i + \sum_{j \neq i} T_{ij} \phi_j + 2\gamma\kappa \sin\{\kappa(\phi_i + Vt + \eta)\}$$

Test of MF Slider Block Ergodicity via Energy Metric

CD Ferguson et al., Phys. Rev. E, 60, 1358 (1999)

An independent test of the idea that mean field slider block models display ergodicity and Boltzmann statistics is at right. There we plot the inverse of the Thirmulai-Mountain (Phys. Rev. A, 42, 4574, 1990) fluctuation metric $1 / \Omega(t)$ for a system with N slider blocks:

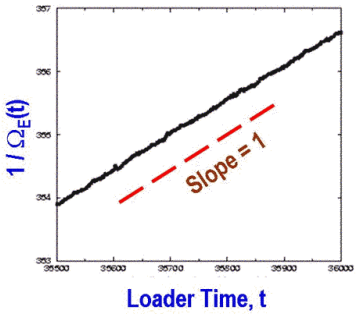
$$\Omega_E(t) \equiv (1/N) \sum_{i=1}^{N} [\, E_i(t) - E(t) \,]^2 \qquad (1)$$

For an ergodic system, Thirumalai and Mountain show that:

$$\Omega_E(t) \sim 1 / t \qquad (2)$$

The data at right indicate that (2) is observed in mean field slider block models.

Note: A similar metric $\Omega_F(t)$ can be computed for any fluctuating quantity F(t).

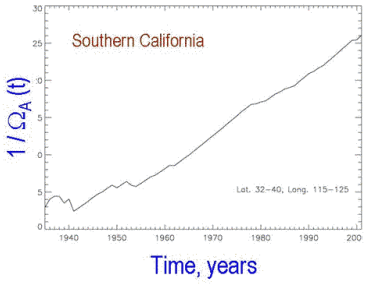


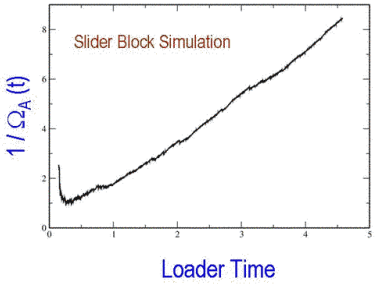
Other work has independently confirmed the existence of ergodicity and Boltzmann statistics in mean field slider block models (W Klein et al., *Phys. Rev. Lett.*, 78, 3793, 1997; G. Morein et al., *Geophys. J. Int.*, 131, 552, 1997; D. Fisher et al., *Phys. Rev. Lett.*, 78, 4885, 1997; CD Ferguson et al., *Phys. Rev. E.*, 60, 1359, 1999; IG Main et al., *J. Geophys. Res.*, 105, 6105, 2000).

Testing for Ergodicity in Natural Seismicity

KF Tiampo et al., PRL, Phys. Rev. Lett., 91, 238501(1-4) (2003)

We apply the TM fluctuation metric using earthquake activity as the fluctuating quantity to compute the metric $\Omega_A(t)$ (D. Thirmulai and R. Mountain, Phys. Rev. A, 42, 4574, 1990)



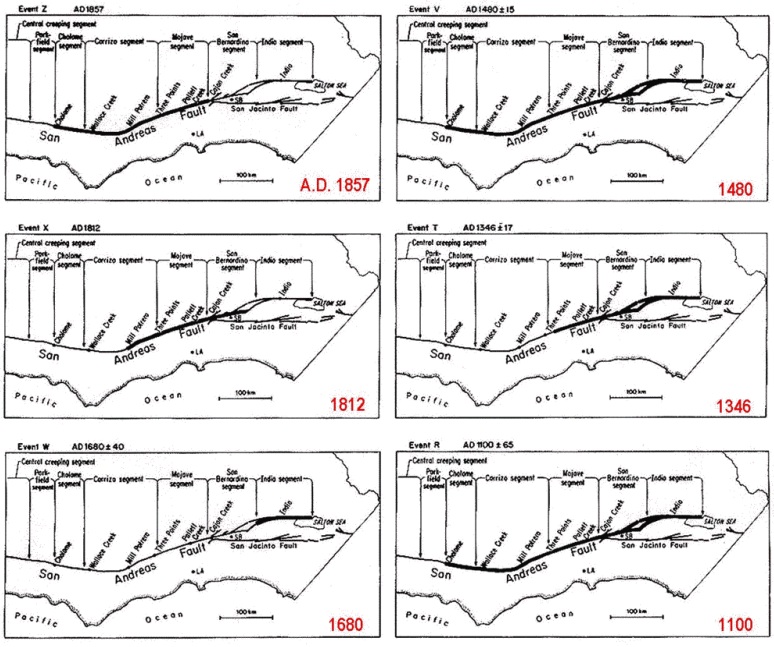


Work confirming the existence of ergodicity and Boltzmann statistics in mean field slider block models (W Klein et al., *Phys. Rev. Lett.*, 78, 3793, 1997; G. Morein et al., *Geophys. J. Int.*, 131, 552, 1997; D. Fisher et al., *Phys. Rev. Lett.*, 78, 4885, 1997; CD Ferguson et al., *Phys. Rev. E.*, 60, 1359, 1999; IG Main et al., *J. Geophys. Res.*, 105, 6105, 2000. D Egolf found similar behavior in mean field coupled map lattices - *D.A. Egolf, Science, 287, 101 , 2000*).

Space-Time Patterns of Historic Earthquakes

Large earthquakes on major faults occur irregularly in time

Major events on the San Andreas Fault



From K. Sieh et al., JGR, 94, 603 (1989)

Simulation based methods: Virtual California

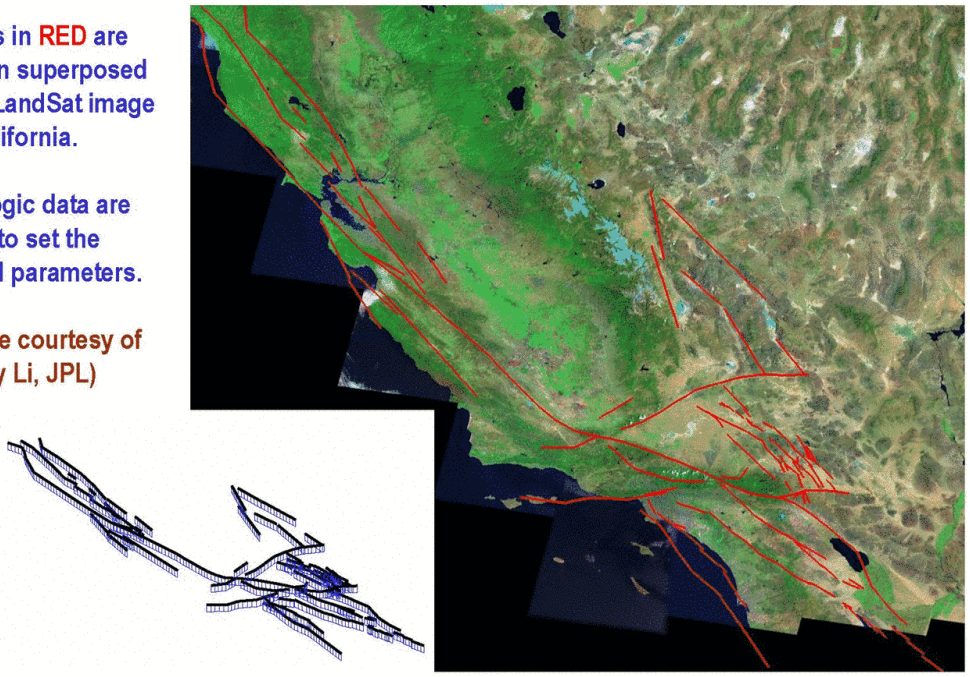
Virtual California 2001 Includes All the Major Active Strike Slip Faults in California

(PB Rundle et al., in press, 2005)

Faults in RED are shown superposed on a LandSat image of California.

Geologic data are used to set the model parameters.

(Image courtesy of Peggy Li, JPL)



The Virtual California Simulation:

Present Characteristics & Properties

Backslip model – Topology of fault system does not evolve. Stress accumulation occurs as a result of negative slip, or "backslip".

Linear interactions (stress transfer) -- Interactions are purely elastic, but linear viscoelastic or poroelastic interactions can be added

Arbitrarily complex 3D fault system topologies -- At the moment, all faults are vertical strike-slip faults. Boundary element mesh is ~ 10 km horizontal, 15 km vertical. Faults are embedded in an elastic half space, but layered media are possible as well.

Friction laws -- are based on laboratory experiments of Tullis-Karner-Marone, with additive stochastic noise. Method of solution for stochastic equations is therefore via Cellular Automaton methods. Friction laws based on general theoretical law obtained by Klein et al. (1997). Both TKM and Rate-and-State friction can be derived as special cases.

Remarks on Friction

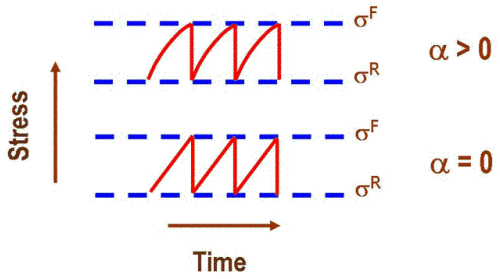
Thresholds and stress-dependent aseismic slip ("Stress Leakage") factor α. Sharp threshold consistent with recent results from Parkfield data. Aseismic slip consistent with laboratory data of Dieterich-Tullis-Karner-Marone

Baseline values for parameters α are determined for each fault segment. It can easily be shown that:

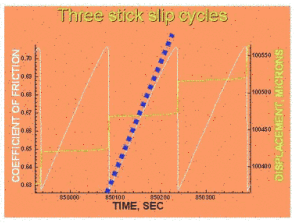
$$2\alpha = \frac{\text{Average stable aseismic slip}}{\text{Total slip}}$$

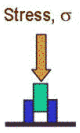
So α is an observable quantity.

We use data from Deng and Sykes (1997) to determine the average values of stable interseismic, aseismic slip for many faults in California.



α determines fraction of total slip that is stable aseismic slip.





Data from T Tullis, PNAS, 1996

Stochastic Friction: Mathematical Details

Friction experiments can be described by a stochastic "Leaky Threshold Equation", or "Hopfield Equation":

Friction Equation on each segment:

$$d\,\Delta\sigma/dt = KV_{Load} - \Delta\sigma\,\{\alpha + \delta(t - t_F) + \delta(\partial\sigma/\partial t - \eta)\} + \varepsilon(t)$$

"Rate of change of stress = Rate of stress supply – Rate of stress dissipation"

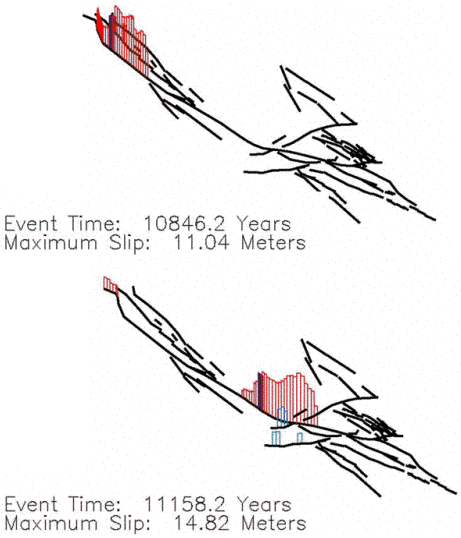
"Dynamic" stress triggering term

Also: $\Delta\sigma = \sigma - \sigma^R$

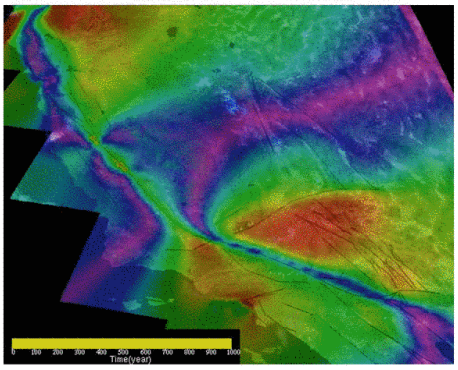
t_F represents all times for which: $\sigma(t_F) = \sigma^F$. **V_{Load} is the segment loading rate (long term fault slip rate), and K is the segment stiffness.**

The δ - function parametrizes the sudden slip event. The factor η characterizes the rate at which stress increases on fault segments. ε is a stochastic noise.

Examples from Virtual California Simulations

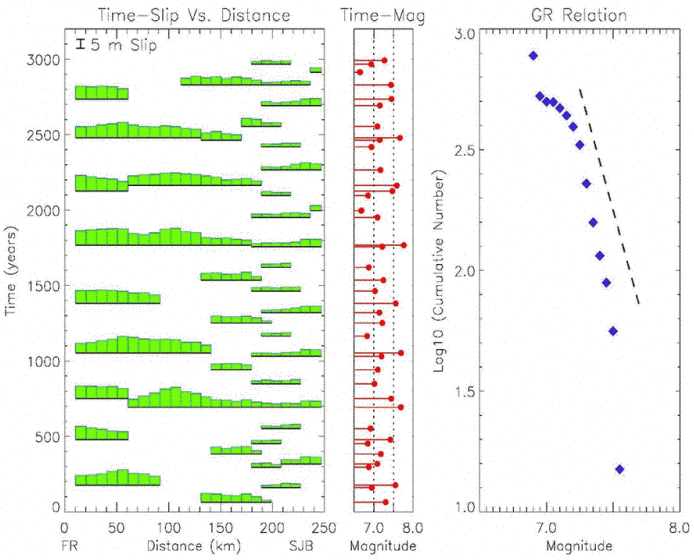


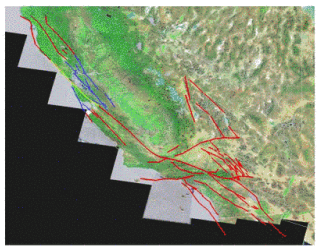
Two large events, dark bar is epicenter location. Red indicates right-lateral slip. Blue indicates left-lateral slip



Synthetic Aperature Radar Interferograms from a VC simulation. 1000 years of events, each frame is the change in surface deformation over 5 years, movie advances 1 year per frame.

Activity on the San Francisco Segment of the San Andreas Fault

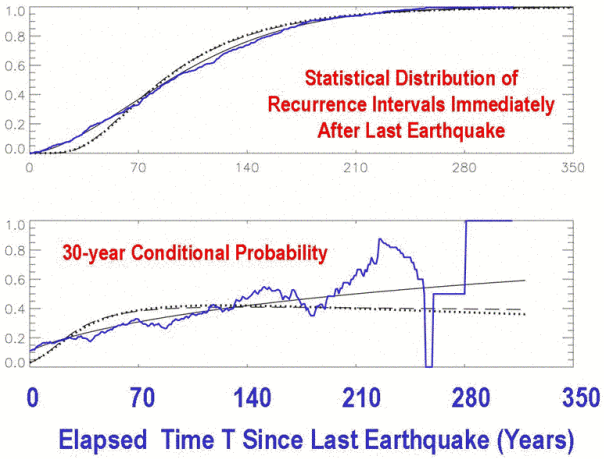


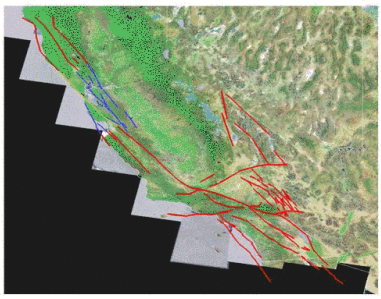


San Francisco Bay section of the San Andreas fault is shown by the yellow fault line.

Forecasting Simulated San Andreas Events

Statistics of M > 7.0 earthquakes on San Francisco Bay section of the San Andreas fault (yellow fault line).
Testing the method used by the Working Group on California Earthquake Probabilities





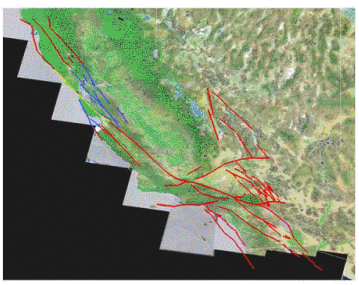
30-Year conditional probability that failure occurs during the 30 year time window, given that it has not occurred prior to the start of the time window.

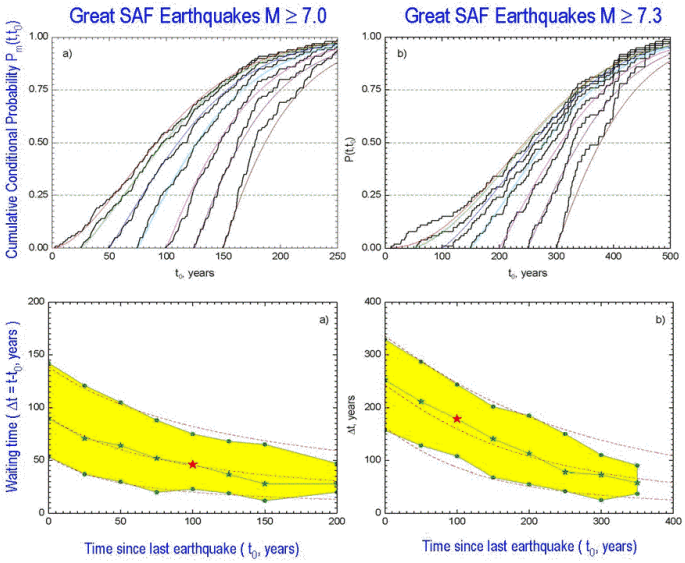
Dotted Line – LogNormal Distribution. Dashed line – Brownian Passage Time Distribution.
Solid Line – Weibull Distribution. Mean = 114 years, Standard Dev. = 62 years
Weibull Distribution is a stretched exponential or Avrami law.

Waiting Time Statistics

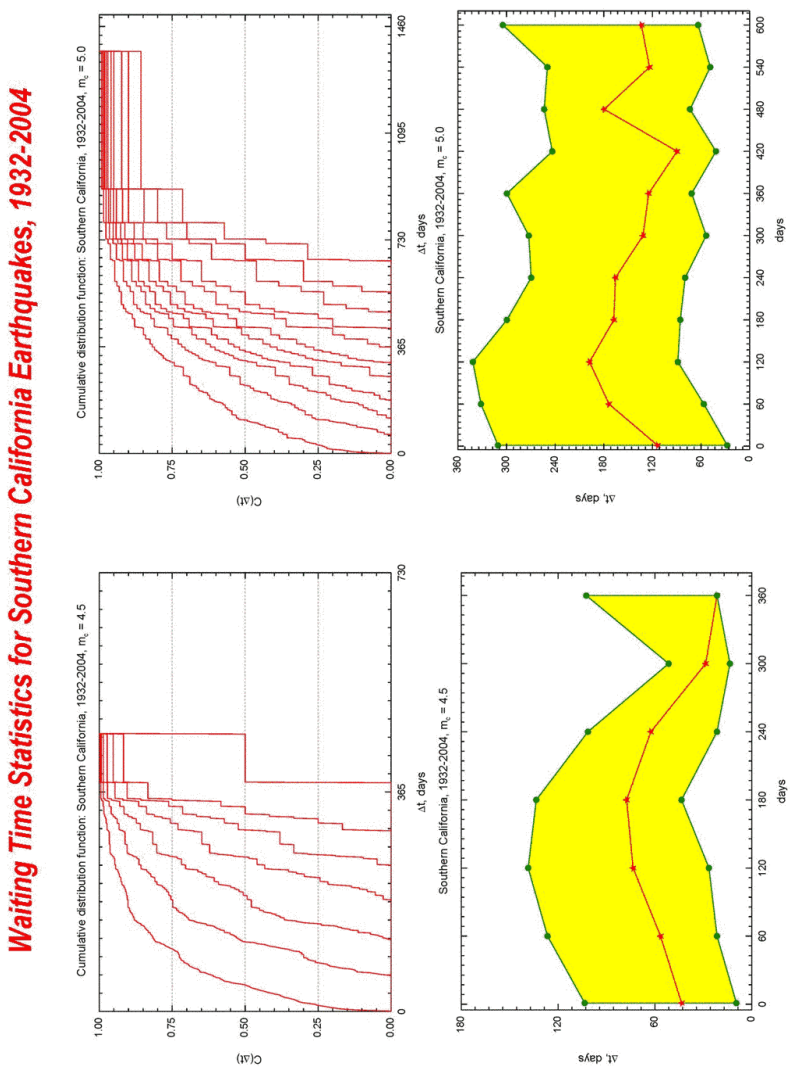
A Method to Forecast the Next Great San Francisco Earthquake on the Northern San Andreas Fault

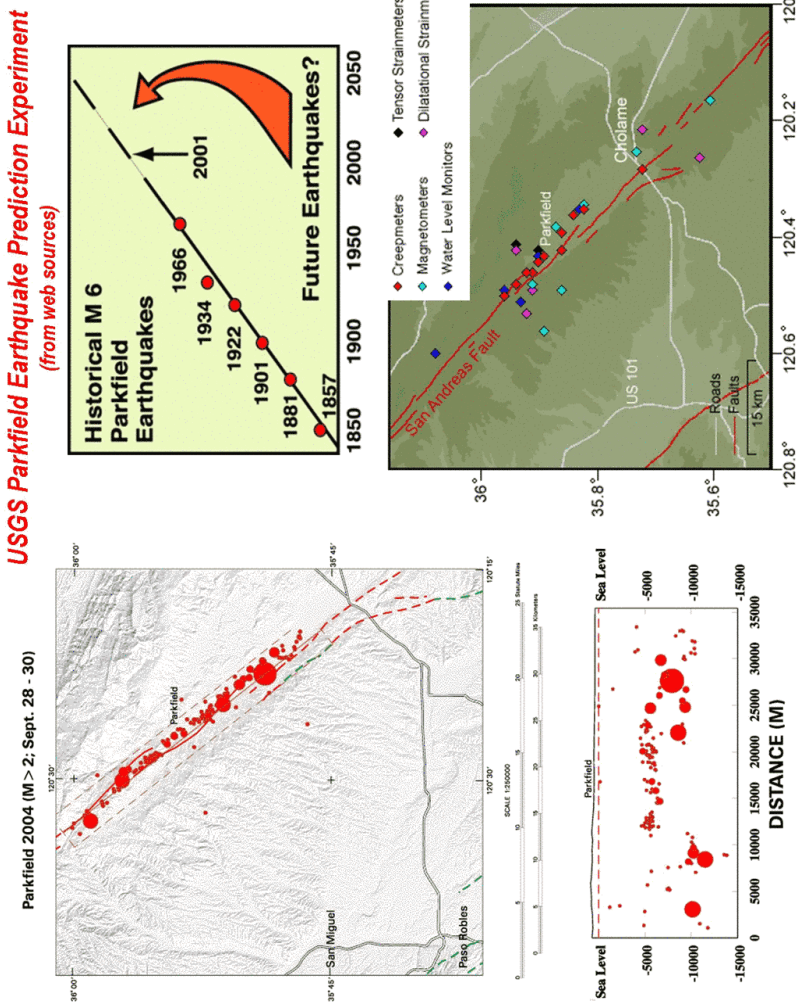
We compute (measure) $P_m(t,t_0)$, which is the conditional cumulative probability that an event with magnitude M > m will occur prior to a time t from the present, given that it has not occurred during the elapsed time t_0 since the last event. From this, we compute the median waiting time $\Delta t = t - t_0$ until the next event, and the 25% - 75% envelope (JBR et al, in press, PNAS, 2005)





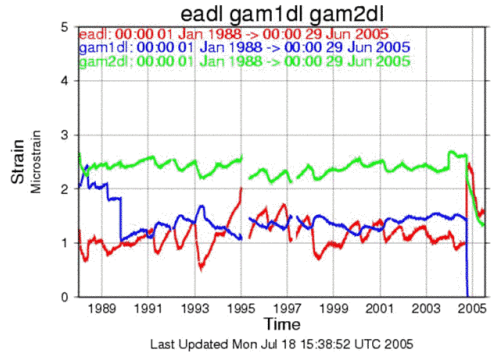
The yellow region is $.25 \le P_m(t,t_0) \le .75$, the middle 50%. The red star represents the value for today, 99 years after the great 1906 San Francisco earthquake. Lines represent best-fitting Weibull statistics (stretched exponential) to the distributions and waiting time data.

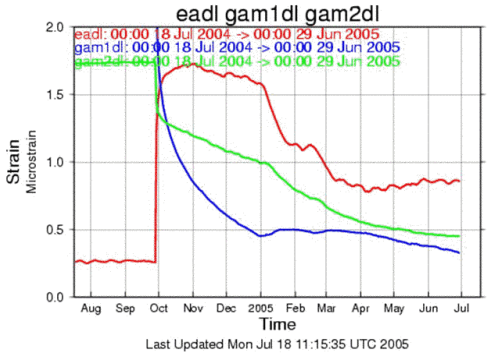

Waiting Time Statistics for Southern California Earthquakes, 1932-2004
Cumulative distribution function: Southern California, 1932-2004, m_c = 5.0
C(Δt)
Δt, days
Southern California, 1932-2004, m_c = 5.0
Δt, days
days
Cumulative distribution function: Southern California, 1932-2004, m_c = 4.5
C(Δt)
Δt, days
Southern California, 1932-2004, m_c = 4.5
Δt, days
days

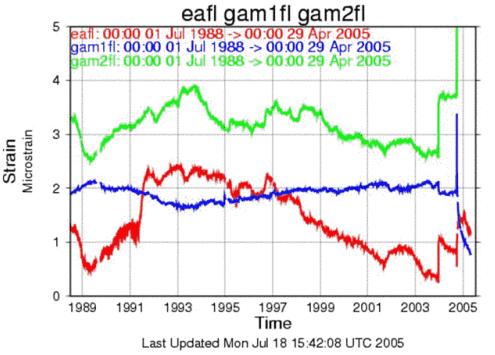

USGS Parkfield Earthquake Prediction Experiment
(from web sources)
Historical M 6 Parkfield Earthquakes
1857
1881
1901
1922
1934
1966
2001
Future Earthquakes?
1850 1900 1950 2000 2050
Creepmeters
Magnetometers
Water Level Monitors
Tensor Strainmeters
Dilatational Strainmeters
San Andreas Fault
Parkfield
Cholame
US 101
Roads
Faults
15 km
36°
35.8°
35.6°
120.8°
120.6°
120.4°
120.2°
120°
Parkfield 2004 (M > 2; Sept. 28 - 30)
Parkfield
San Miguel
Paso Robles
Sea Level
-5000
-10000
-15000
0 5000 10000 15000 20000 25000 30000 35000
DISTANCE (M)

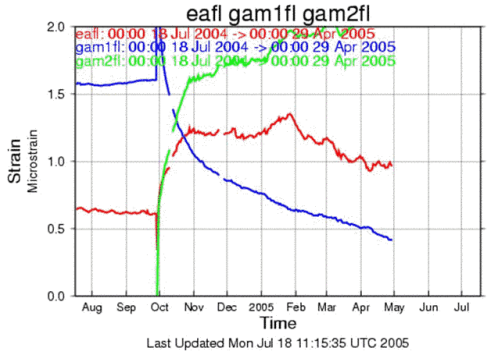
Tensor Strain Prior to Parkfield

No precursors observed, in contrast to predictions from some models



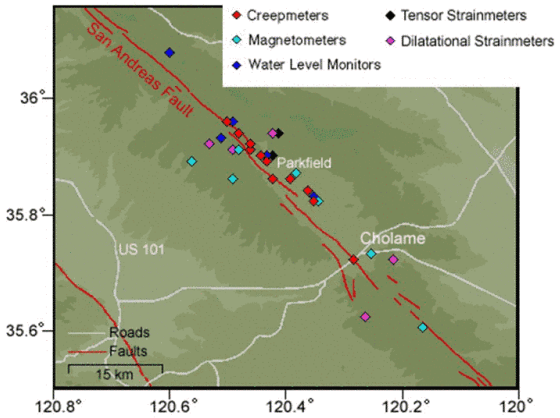


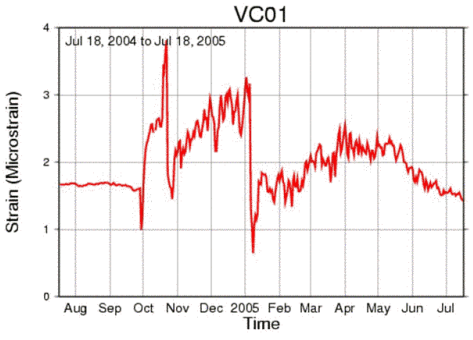


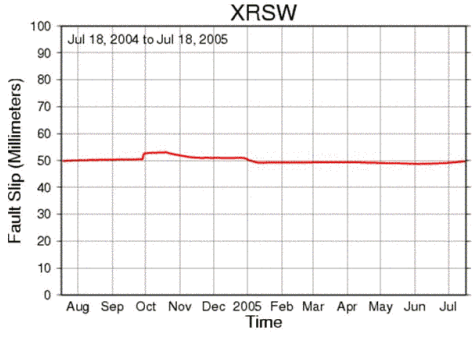


Parkfield Borehole Dilatometer and Fault Creep Data

No precursors observed, in contrast to predictions from some models

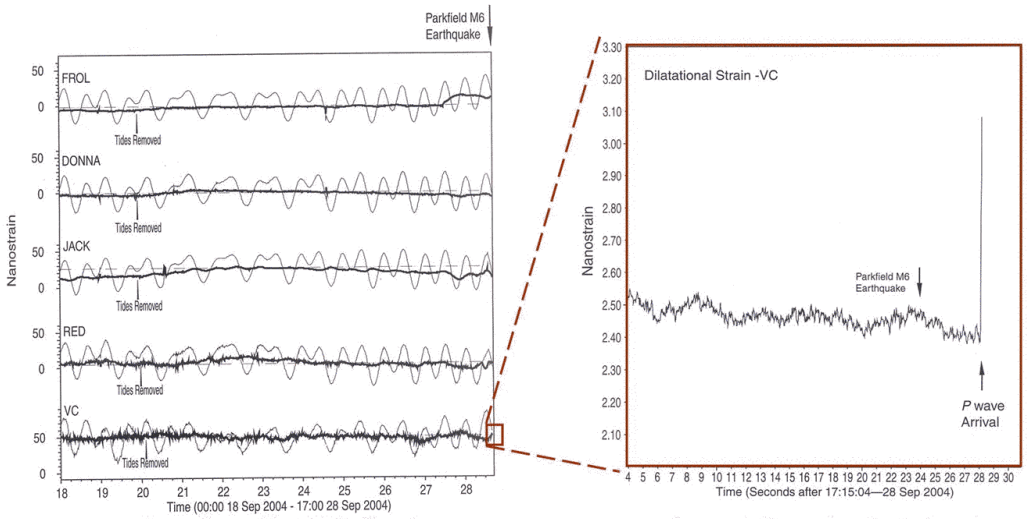






Parkfield Strain - High Resolution

No precursors observed at any scale of resolution, contrast to predictions from some models.



Days Preceding Main Shock

Seconds Preceding Main Shock

Borehole volumetric strainmeter data. Thin lines corrected for atmospheric pressure. Thick lines are data with Earth tides removed.

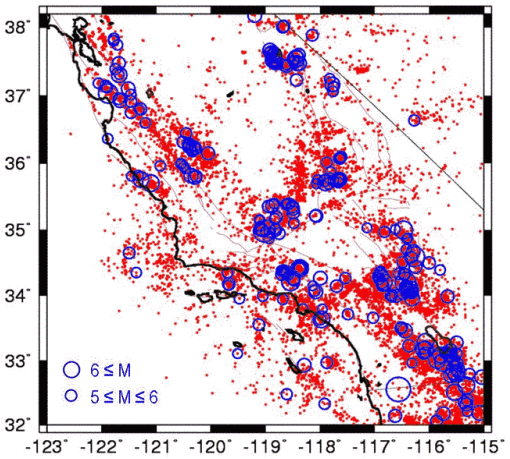
Data from Langbein et al., *Seis. Res. Lett.*, 76, 10-26 (2005)

Earthquake Fault Systems are Phase Dynamical Systems

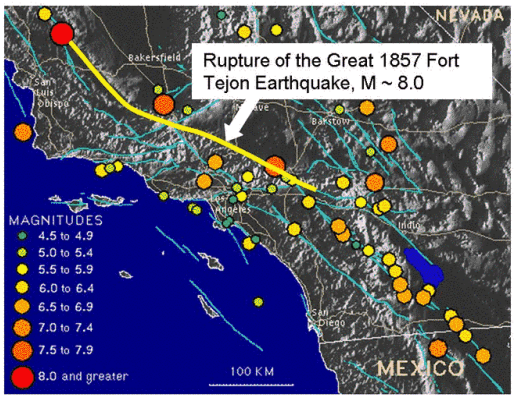
JBR et al., Phys. Rev E, 61, 2418-2431 (2000)

Historic Records (since ~ 1800 AD) and Seismographs (since 1932) in California Inform Us About Locations of Large California Earthquakes of the Past

Epicenters from Instrumental Seismicity
1932 - Dec 31, 1999; M ≥ 3.0



Epicenters from Historic Seismicity
Large Earthquakes, Last ~200 Years



http://www.data.scec.org/clickmap.html

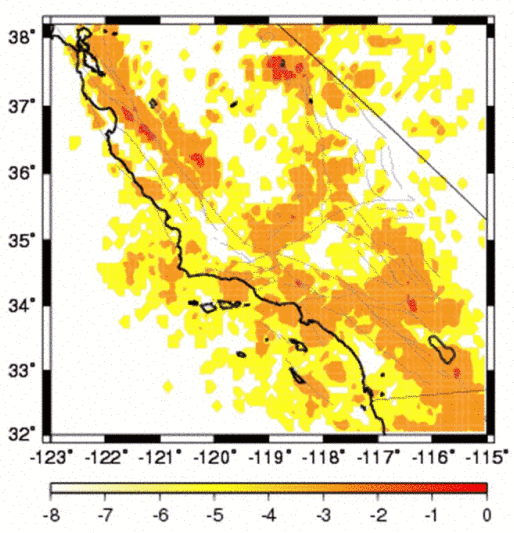
From Seismicity Data to Intensity Maps

Start with raw seismicity data

Define spatial coarse-grained grid at a size of .1° x .1°
(.1° ~ linear size of an M ~ 6 earthquake)

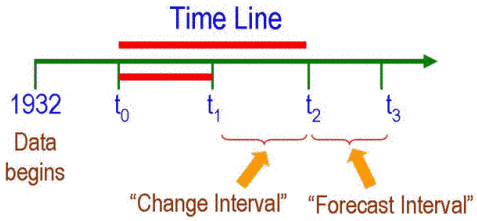
Compute and contour the seismic intensity $I(x,t_0,t)$ over the time interval (t_0,t). Units are
(Number / $Number_{MAX}$)

Figures Courtesy of James Holliday



Log_{10} (Normalized Intensity with $M \geq 3$)
[Intensity normalized to maximum value]

Computing a PI Earthquake Forecast



1. Compute intensity maps:

   $I_1 = I(x,t_0,t_1)$; $I_2 = I(x,t_0,t_2)$.

2. Normalize them so they have the same statistics with respect to area averages.

3. Define the average Change Map:

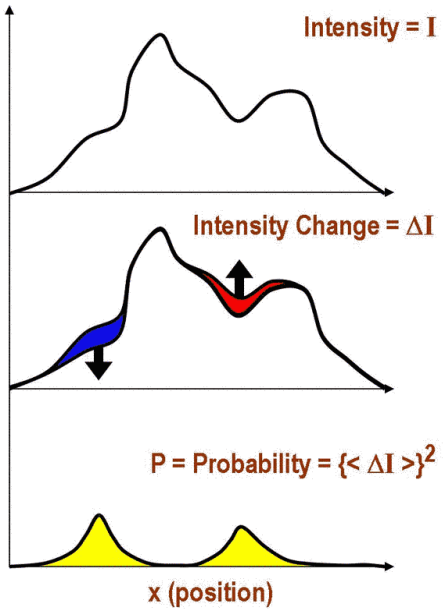
   $\langle \Delta I \rangle = \langle I_2 - I_1 \rangle$

   The average is over all change maps having the same change interval (t_1,t_2).

4. Define $P = \{\langle \Delta I \rangle\}^2$, and subtract the area mean to obtain ΔP

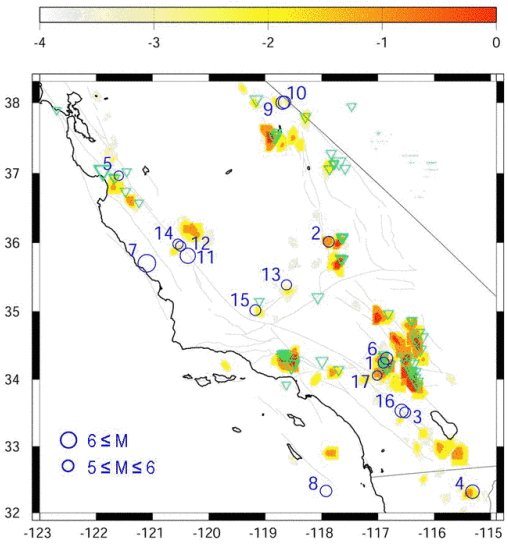
5. Color-contour $Log_{10}\{\Delta P\}$ on a map

Schematic: Spatial Cross Section Of Intensity Map



A Real Time Earthquake Forecast Experiment:
Forecasting Locations of Future Significant Earthquakes 2000 - 2010

(JB Rundle et al., PNAS, v99, Supl 1, 2514-2521, Feb 19, 2002; KF Tiampo et al., Europhys. Lett., 60, 481-487, 2002; JB Rundle et al.,Rev. Geophys. Space Phys., 41(4), DOI 10.1029/2003RG000135 ,2003. http://quakesim.jpl.nasa.gov)



Plot of Log_{10} (Seismic Potential)

Increase in Potential for significant earthquakes, ~ 2000 to 2010

Seventeen significant earthquakes (blue circles) have occurred in Central or Southern California. Margin of error of the anomalies is +/- 11 km; Data from S. CA. and N. CA catalogs:

After the work was completed

1. Big Bear I, M = 5.1, Feb 10, 2001
2. Coso, M = 5.1, July 17, 2001

After the paper was in press (September 1, 2001)

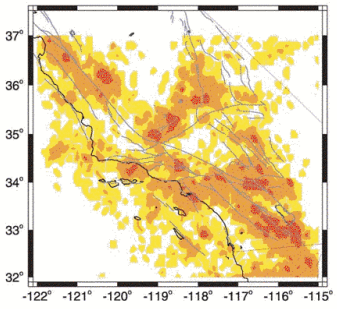
3. Anza I, M = 5.1, Oct 31, 2001

After the paper was published (February 19, 2002)

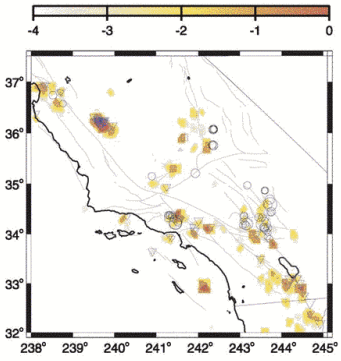
4. Baja, M = 5.7, Feb 22, 2002
5. Gilroy, M=4.9 - 5.1, May 13, 2002
6. Big Bear II, M=5.4, Feb 22, 2003
7. San Simeon, M = 6.5, Dec 22, 2003
8. San Clemente Island, M = 5.2, June 15, 2004
9. Bodie I, M=5.5, Sept. 18, 2004
10. Bodie II, M=5.4, Sept. 18, 2004
11. Parkfield I, M = 6.0, Sept. 28, 2004
12. Parkfield II, M = 5.2, Sept. 29, 2004
13. Arvin, M = 5.0, Sept. 29, 2004
14. Parkfield III, M = 5.0, Sept. 30, 2004
15. Wheeler Ridge, M = 5.2, April 16, 2005
16. Anza II, M = 5.2, June 12, 2005
17. Yucaipa, M = 4.9 - 5.2, June 16, 2005

Note: This original forecast was made using both the full Southern California catalog plus the full Northern California catalog. The S. Calif catalog was used south of lattitude 36°, and the N. Calif. catalog was used north of 36° . No corrections were applied for the different event statistics in the two catalogs. Green triangles mark locations of large earthquakes (M ≥ 5.0) between Jan 1, 1990 – Dec 31, 1999.

CL#03-2015



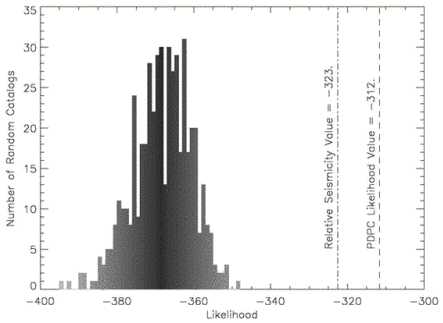
Original version of historic intensity map, 1932-2002



Original retrospective forecast for the period 1992-2002

First Tests of Pattern Informatics Method for Earthquake Forecasting Using Likelihood Ratios (ca. 2002)

Here we compared the forecast skill of the historic intensity map (significant earthquakes with M > 5 will occur where the most small earthquakes with M ~ 3 have historically occurred) with the forecast skill of the PI method (then called the PDPC method).



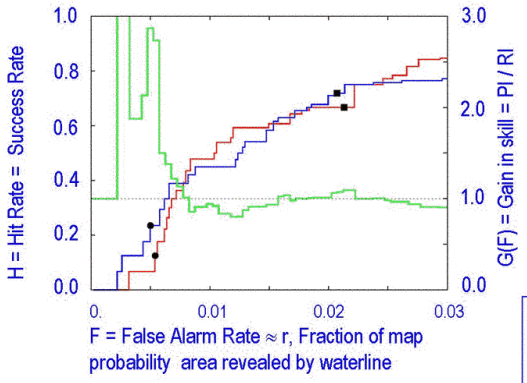
Original results of likelihood tests for two null hypotheses 1) 500 random seismicity catalogs (histogram) and 2) the historic intensity map shown at upper left; and 3) the Pattern Informatics map at lower left (originally called the "Phase Dynamics Probability Change" (PDPC) method.

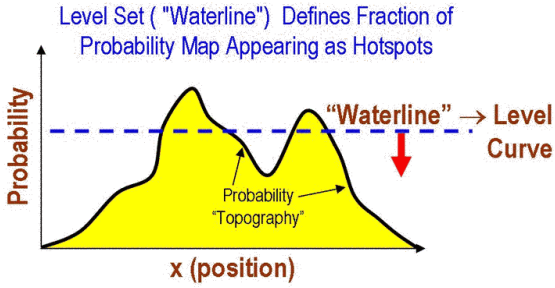
Note that the Likelihood test has an extremely serious defect - it multiplies probabilities rather than adding them - thus the likelihood value is extremely sensitive to the probability of the least probable event.

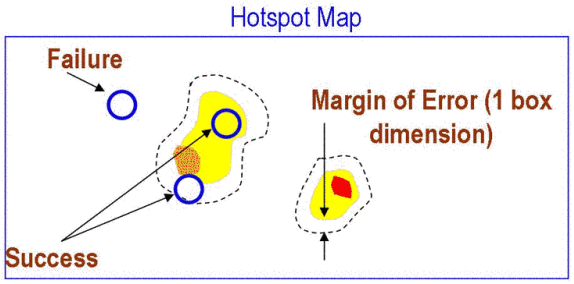
A Better Method of Testing

Evaluating Success Using Relative Operating Characteristic (ROC) Diagrams

Forecast Interval: 1992 – 2005 (about 60 "large" earthquakes M ≥ 5). Historic Intensity computed through end of 1991.
PI computed over change interval from 1978-1991.







"Success" is defined if epicenter of large event falls within the area enclosed by dashed line

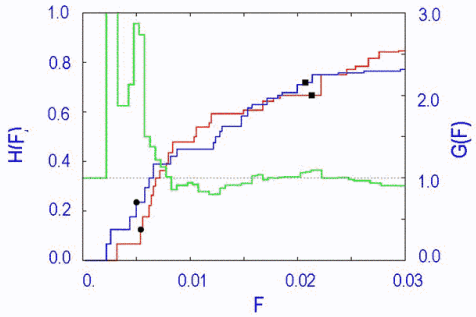
Detail on Contingency Tables, ROC Curves and "Hotspot Pixel" Maps

Pattern informatics (PI) forecast

Forecast	Observed		
	Yes	No	Total
Yes	(a) 4	(b) 25	29
No	(c) 13	(d) 4998	5011
Total	17	5023	5040

Relative intensity (RI) forecast

Forecast	Observed		
	Yes	No	Total
Yes	(a) 2	(b) 27	29
No	(c) 14	(d) 4997	5011
Total	16	5024	5040



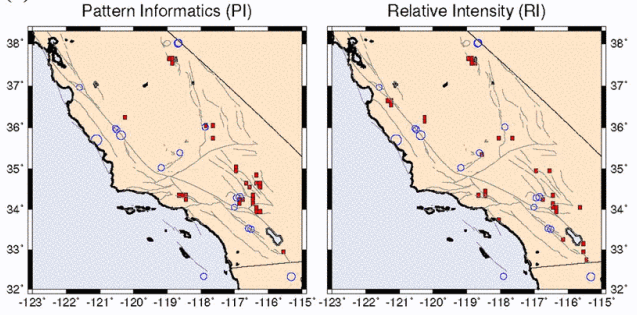
Pattern informatics (PI) forecast

Forecast	Observed		
	Yes	No	Total
Yes	(a) 23	(b) 104	127
No	(c) 9	(d) 4904	4913
Total	32	5008	5040

Relative intensity (RI) forecast

Forecast	Observed		
	Yes	No	Total
Yes	(a) 20	(b) 107	127
No	(c) 10	(d) 4903	4913
Total	30	5010	5040

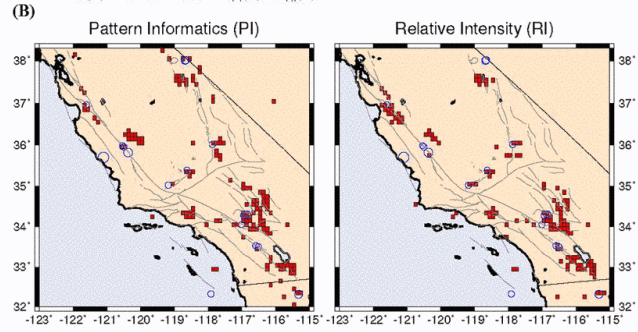
(A)



Contingency Tables:
F = False Alarm Rate = b /(b +d)
H(F) = Hit Rate = a /(a +c)
G(F) = Gain in Skill = $H_{PI}(F) / H_{RI}(F)$

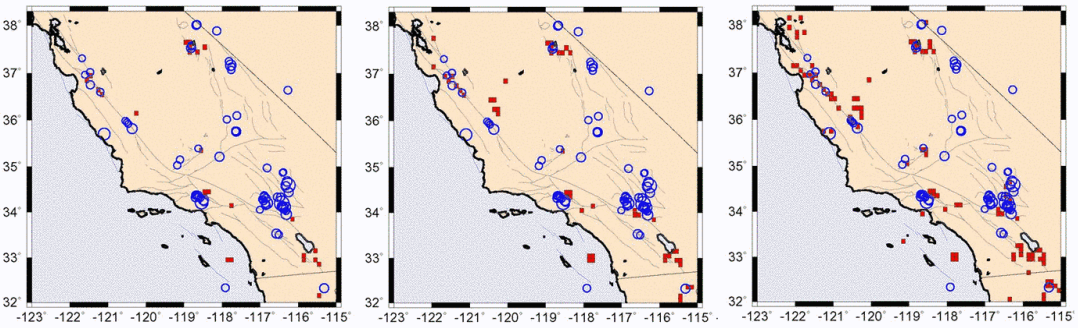
Note:
F ≈ r = fraction of map probability area revealed by level curve or waterline

(B)



Successive Level Curves on PI Pixel Maps

Change Interval: 1978-1991. Forecast Interval: 1992-2005



1.5% of active boxes colored | 3% of active boxes colored | 6% of active boxes colored

More Failures to Predict — *Fewer Failures to Predict*

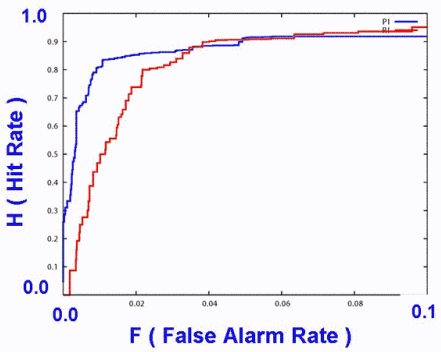
More Red Hotspot Pixels

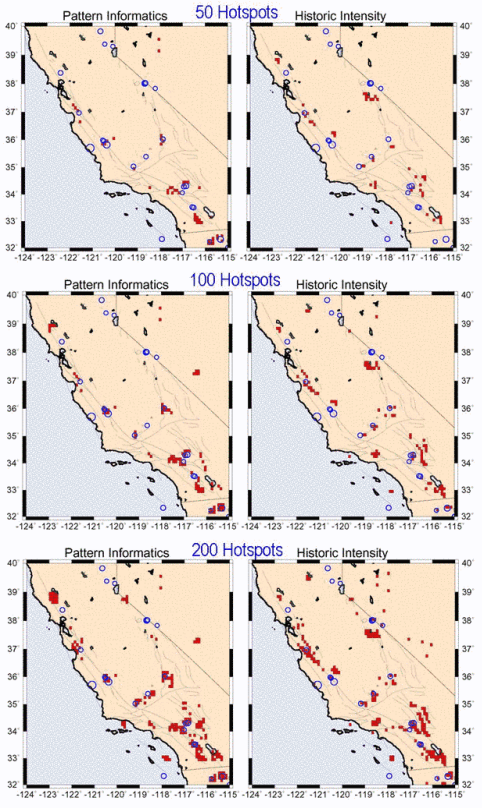
Fewer False Alarms — *More False Alarms*

New Modification: PI-10M Method Applied to California Earthquakes

We have developed a new modification of the original PI method that begins with a forecast based on the historic intensity map and improves upon it (CC Chen et al., 2005). At right are maps based on this modification corresponding to the forecast for 2000 - 2010.

Details: We use only the top 10% most active sites, and normalize all time series in the remaining boxes to have the same statistics. The new algorithm weights the change maps made using longer time series more heavily than change maps made using shorter time series. Here t_0 = 1950, t_1 = 1985, t_2 = 1999.. Here we use M > 2.8 events to forecast M > 4.8 earthquakes.





Summary:

Any information about the future is valuable.

Information about the locations of future events within a multiyear time window and above a "significant" magnitude threshold of M > ~ 5 is available.

Various types of multiyear forecasts are possible - some are based only on data mining (PI), others are model- and simulation-based (Virtual California).

Practical earthquake forecasting on multiyear time scales is a work in progress - it is not impossible, only very difficult.