

Dynamics of Seismicity Patterns in Systems of Earthquake Faults

John B. Rundle
Department of Physics
and
Colorado Center for Chaos & Complexity
and
Cooperative Institute for Research in Environmental Sciences
University of Colorado
Boulder, CO 80309

W. Klein
Department of Physics
and
Center for Computational Science
Boston University
Boston, MA 02215

Kristy Tiampo and Susanna Gross
Colorado Center for Chaos & Complexity
Cooperative Institute for Research in Environmental Sciences
University of Colorado
Boulder, CO 80309

ABSTRACT

Systems of earthquake faults frequently show space-time behavior that is difficult to interpret. We describe a new technique based upon a Karhunen-Loeve expansion that allows the observed space-time seismicity patterns to be understood as eigenstates of a suitably constructed correlation operator. The evolution of seismicity can then be viewed in terms of a phase dynamical system giving rise to a "Pattern Dynamics" that can be obtained directly from observed data. Since the method does not use information on stress, displacements, or other fault-related variables, it necessarily represents an incomplete description of the dynamics. In principle, these methods can be used to construct forecast algorithms. We then apply our methods to seismicity from both numerical simulations and data obtained from southern California. Finally, we perform statistical tests to measure the quality of the forecasts computed using simulation data.

KEYWORDS: Threshold Systems, Stochastic Resonance, Earthquakes, Friction, Nonequilibrium Systems, Driven Dissipative Systems, Earthquake Forecasting, Nonlinear Prediction

1. Introduction

Earthquake faults can be considered to be driven nonlinear threshold systems comprised of interacting spatial networks of sliding surfaces that are subjected to persistent driving stresses. To first order, a fault segment slips when the shear stress $\sigma(\mathbf{x},t)$ on a cell at position \mathbf{x} and time t rises to an effective threshold level at which frictional resistance is overcome. The result is an increase in the slip $s(\mathbf{x},t)$ of the segment, as well as a decrease in the stress supported by the segment to a lower, residual value. Failure thresholds, residual stresses, and slip magnitudes may be modified by the presence of spatial heterogeneity in physical properties. Elastic interactions between segments may be excitatory (positive) in the sense that failure of interacting neighbors brings a segment closer to sliding, or inhibiting (negative) in the opposite case. Observations, together with numerical simulations of these systems (Rundle and Klein, 1995; Fisher et al., 1997; Scholz, 1990), reveal spatial and temporal patterns $\Psi(\mathbf{x},t)$ of seismic activity that are richly complex, and difficult to understand from any deterministic point of view (Nijhout, 1997). Understanding such patterns will increase our understanding of the physics of the fault systems themselves, since these space-time patterns are clearly emergent processes that reflect the structures, dynamics, and properties of the underlying high dimensional nonlinear system.

In earthquake fault systems, there exists the further difficulty that the deterministic dynamics, together with values of the stress $\sigma(\mathbf{x},t)$ and slip $s(\mathbf{x},t)$ state variables, are often unknown or effectively unobservable. Only very limited information is available at the earth's surface from drilling, GPS, InSAR, or other technologies. However, the associated pattern variables $\Psi(\mathbf{x},t)$, the seismic events themselves, *are* observable. Quantitatively, the seismicity $\Psi(\mathbf{x},t)$ can be represented as a set of time series at all positions \mathbf{x} , with $\Psi(\mathbf{x},t)$ defined by $\Psi(\mathbf{x},t) = 1$ if an earthquake occurs between $\{t, t+\Delta t\}$; $\Psi(\mathbf{x},t) = 0$ otherwise. The basic problem is illustrated in figure 1. The slip $s(\mathbf{x},t)$ at position \mathbf{x} and time t evolves to a value $s(\mathbf{x},t + \Delta t)$ under the action of the deterministic dynamics D_t :

$$D_t \{ s(\mathbf{x},t) \} \rightarrow s(\mathbf{x},t + \Delta t) \quad (1)$$

The stress $\sigma(\mathbf{x},t)$ also evolves to $\sigma(\mathbf{x},t+\Delta t)$. The values of the slip and stress state variables are effectively hidden from view, below the horizontal dashed line in figure 1.

However, the time, location, and magnitude of seismic events on fault systems are

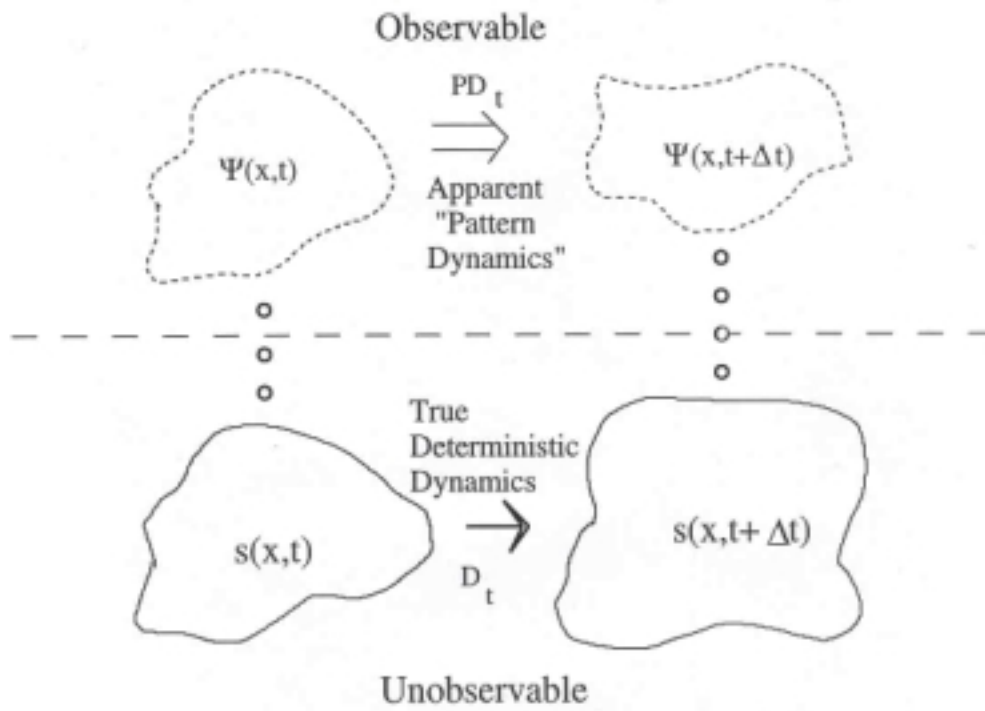


Figure 1. Schematic diagram of threshold systems, including unobservable state variables $s(\mathbf{x},t)$ and observable pattern variables $\Psi(\mathbf{x},t)$.

readily observable, and are tabulated routinely (Hill et al., 1990; Sieh et al., 1989; Bakun and McEvilly, 1984). The seismicity gives rise to observable space-time patterns $\Psi(\mathbf{x},t)$, which should not be confused with the variables in which the underlying dynamics is formulated, $s(\mathbf{x},t)$ and $\sigma(\mathbf{x},t)$. These latter variables also display patterns and correlations, but these are not the patterns of interest to us here since they are effectively unobservable. Since both $\Psi(\mathbf{x},t)$ and $\Psi(\mathbf{x},t + \Delta t)$ are observable, it might be possible in principle to learn about D_t , $s(\mathbf{x},t)$ and $\sigma(\mathbf{x},t)$ by studying the evolution of seismicity patterns. Unfortunately, there exists no evidence at present that $\Psi(\mathbf{x},t)$ is uniquely related to $s(\mathbf{x},t)$ and $\sigma(\mathbf{x},t)$. In other words, there may be a large number of slip states $\sigma(\mathbf{x},t)$ that produce a similar seismicity pattern $\Psi(\mathbf{x},t)$, or a large number of patterns $\Psi(\mathbf{x},t)$ associated with similar states $s(\mathbf{x},t)$. This weak association between $s(\mathbf{x},t)$ and $\Psi(\mathbf{x},t)$ is represented schematically in figure 1 by a vertical dotted line.

In attempting to forecast the future evolution of slip on fault systems, a common approach is to attempt to 1) use observations of seismicity $\Psi(\mathbf{x},t)$ to infer $s(\mathbf{x},t)$; then 2) use $s(\mathbf{x},t)$ together with an assumed model for the dynamics D_t to calculate $s(\mathbf{x},t + \Delta t)$; and finally 3) use $s(\mathbf{x},t + \Delta t)$ to infer $\Psi(\mathbf{x},t + \Delta t)$, which represents the observable seismic pattern of interest. However, this method involves a long chain of inference and assumption, particular links of which may not be well-justified. As we will show, earthquakes can be considered to be an example of a phase dynamical system (Mori and Kuramoto, 1998), in which evolution of the observables can be associated with rotation of a state or pattern vector within a high dimensional correlation space. In this new approach, we propose the construction of an approximate phase dynamics based on Karhunen-Loeve (KL) methods that associates a future seismicity pattern state $\Psi(\mathbf{x},t + \Delta t)$ with an earlier pattern state $\Psi(\mathbf{x},t)$ by means of an assumed "Pattern Dynamics" operator PD_t :

$$PD_t \{ \Psi(\mathbf{x},t) \} \Rightarrow \Psi(\mathbf{x},t + \Delta t) \quad (2)$$

Moreover, we propose to construct PD_t without any knowledge of the true deterministic dynamics D_t , using only knowledge of the seismicity pattern states preceding $\Psi(\mathbf{x},t)$.

The pattern dynamics operator that we construct operates on a space of functions $\Psi_R(\mathbf{x},t)$ that we call "reconstructed pattern states". The $\Psi_R(\mathbf{x},t)$ are continuous-valued in time, and represent approximations, in the sense described below, to the pattern states $\Psi(\mathbf{x},t)$. In fact, the $\Psi_R(\mathbf{x},t)$ will be seen to represent probability amplitudes from which probability density functions for seismicity may be calculated. This identification also follows from the fact that $\Psi_R(\mathbf{x},t)$ evidently arises from phase dynamics (Mori and Kuramoto, 1998), which are known to

have deep similarities to quantum systems. In contrast to the underlying deterministic dynamics D_t , which are most probably strongly nonlinear, the operator PD_t is assumed to be linear over small time intervals Δt . A convenient analogy for the relationship between PD_t and D_t is the correspondence between an equilibrium dynamical system, such as an Ising model that is governed by a nonlinear Langevin equation, and the corresponding functional Fokker-Planck equation (Haken, 1983; Gardiner, 1985). Evolution of dynamical variables such as the spin density are governed by a strongly nonlinear Langevin equation, but the probability density functions evolve according to a linear functional Fokker-Planck equation.

The assumption of a linear dynamics governing the $\Psi_R(\mathbf{x}, t)$ follows from the statistical time-stationary properties of the underlying dynamics, which in turn justifies the use of Karhunen-Loeve methods. We have shown in previous work (Rundle et al., 1995; Klein et al., 1997; Ferguson et al., 1999) that mean field threshold systems such as earthquake faults, characterized by long range interactions between sites \mathbf{x} and \mathbf{x}' , can be treated as equilibrium systems and have dynamics that are statistically stationary over long time intervals. These long intervals are eventually punctuated by rare, large events that serve to reorder the entire system. As the range of interaction increases, these equilibrium-like time intervals increase. In the limit of infinite range interactions where mean field conditions hold and the system is ergodic, Poincare's theorem (Huang, 1987) implies that dynamical variables can be written as sums and integrals of complex exponentials. In this paper, we construct a pattern dynamics for these mean field earthquake fault systems.

Our approach is similar in some respects to previous methods that use correlation methods (Nijhout et al., 1997; Fukunaga, 1990), but also has important differences. In most of these methods, a formalism is constructed that in effect uses space-time patterns to extrapolate future system behavior. However, our approach differs in that we retrieve a complete set of space-time patterns represented by the eigenvectors $\phi_n(\mathbf{x}_i)$ and eigenfrequencies ω_n of an equal time, rate-rate correlation function. The $\phi_n(\mathbf{x}_i)$ provide information about spatial correlations of patterns, whereas the ω_n provide information regarding how often each spatial correlation (eigenvector) is represented in the data. If the process is statistically stationary, as we have established for the mean field threshold systems we consider, these $\phi_n(\mathbf{x}_i)$ and ω_n *will themselves be independent of time*. Hence $\phi_n(\mathbf{x}_i)$ and ω_n can justifiably be determined from an equal time correlation operator. From these eigenvectors and eigenfrequencies, as well as initial-value data on the most recent event activity at each spatial location, we obtain the reconstructed pattern state $\Psi_R(\mathbf{x}, t)$, which in turn is used to compute a probability density function $P(\mathbf{x}, t)$.

Forecasting future evolution of other nonlinear systems, such as climate, weather, or El Nino-Southern Oscillations, often involve the computation of an unequal time correlation function (UTC) over short time intervals (Preisendorfer, 1988; Penland, 1989; Garcia and Penland, 1991; Penland and Sardeshmukh, 1995; Penland and Magorian, 1993; Broomhead and King, 1986; Vautard and Ghil, 1989). However, in these systems, which are not "threshold systems" in the sense we consider here, the state variables $s(\mathbf{x},t)$ and $\sigma(\mathbf{x},t)$ that define the true deterministic dynamics *can* be directly observed. These include wind speeds, temperatures, and pressures, and so forth. The dynamics, as represented by the Navier-Stokes equation, are essentially known. It would not be useful to construct a UTC operator for the state variables $s(\mathbf{x},t)$ and $\sigma(\mathbf{x},t)$, since these are unobservable. In the case of earthquake mainshocks, computing a UTC operator in the pattern variables $\Psi(\mathbf{x},t)$ is difficult because of the apparently random nature of the long intervals of quiescence that separate mainshocks. Long periods of activity are punctuated by short bursts of activity, making the definition of one or more time steps for the UTC problematic.

Our first attempts at forecasting the patterns seen in the earthquake simulations described below used just such a UTC technique, which are based on a linear operator acting on a state at time t to predict a state or probability at time $t + \Delta t$. As a result, the UTC forecast method would be inherently unable to distinguish between a precursory quiescence that precedes a large earthquake mainshock and the interseismic quiescent interval that follows it. Yet the probability of the earthquake would clearly be greater before the event than following. Moreover, there is often a wide spectrum of time scales among the pattern states, a property which is extremely difficult for a method using a one-time-step UTC state transition matrix to adequately capture. For these reasons, our attempts at forecasting simulated earthquakes using an unequal time correlation operator were not satisfactory, and we were thus motivated to develop the new method described in the following.

2. Examples: Dynamics and "Pattern Dynamics"

As illustrations to which our methods will be applied below, we show two examples of pattern activity in earthquake fault systems. In the first example, natural seismicity from southern California were obtained from the Caltech catalogue. Seismicity data between longitudes -115° and -121° , and latitudes 36° and 39° were considered. Only A and B quality events were accepted and all blast events were removed. Epicentral locations for seismic events were then binned into squares of 0.1° latitude and 0.1° longitude on a side. Figure 2 is a plot of the relative activity at all sites in southern California at which at least one earthquake occurred

Relative Southern California Seismicity, 1932-1998

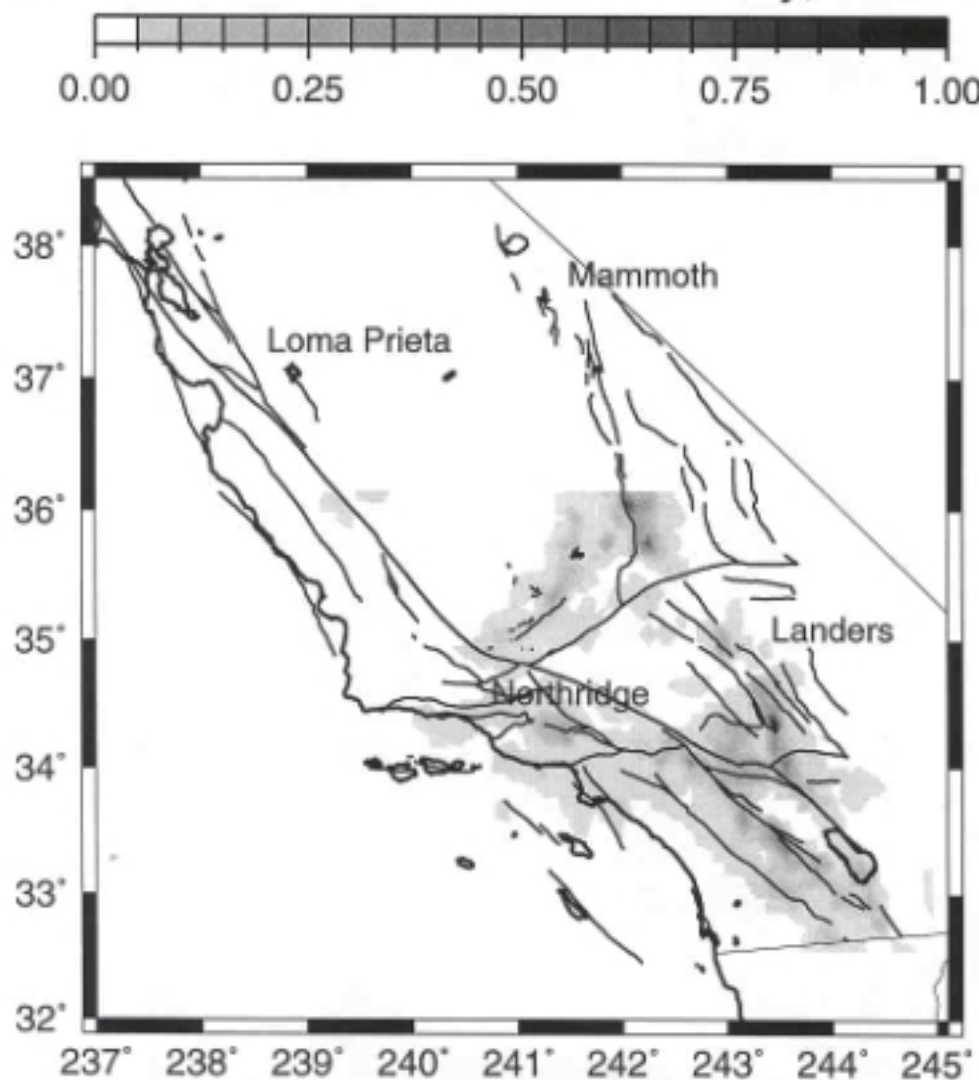


Figure 2. Relative seismicity in southern California from 1932-1998.

during the time interval 1932-1998. These are the data which define $\Psi(\mathbf{x},t)$. Obviously, no information is available about slips or stresses, thresholds, residual stresses, or friction constitutive laws on the various faults in the region.

Figure 3 illustrates space-time patterns of activity produced by a realistic earthquake fault model (Rundle, 1988). In this model, 80 fault segments are endowed with static and kinetic coefficients of friction, and are driven by a persistently increasing shear stress that, over long time intervals, produces slip events whose time-averaged rate of slip is equal to that observed in nature. In the model, there are 6 major faults: the San Andreas (SA), Imperial (I), Cerro Prieto (CP), San Jacinto (SJ), Elsinore (E), and Garlock (G). The top diagram is a plot with positions on all 6 faults concatenated together along the horizontal axis, and with event times (lines) on the vertical axis for a simulation run. The middle diagram is a plot of friction against position, and the bottom is a map view of the faults. For this simulation, the coarse-grained time interval $\Delta t = 1$ year. Details are given in appendix A.

Referring to the top diagram in figure 3, each horizontal line represents an earthquake that occurred at the time indicated on the vertical axis. These data correspond to the observable patterns $\Psi(\mathbf{x},t)$ described above. At each position \mathbf{x}_i along the earthquake faults, an activity time series can be constructed for which $\Psi(\mathbf{x}_i,t) = 1$ if an earthquake occurs at time t , and $\Psi(\mathbf{x}_i,t) = 0$ otherwise. Thus we have 80 time series, one for each fault segment, and each with 2000 time entries (2000 years in the simulation). These data are meant to be a somewhat realistic representation of the actual earthquake fault system. Like the real faults, the function $\Psi(\mathbf{x},t)$ represents the observable data, and encodes all space-time patterns that may exist. In contrast, there are many unobservable parameters in the simulations. As described in the appendix, each fault segment has coefficients of friction, a slip state variable, two stress variables (shear and normal stress), and so forth. In the real earth, there is no realistic possibility of measuring values for these variables.

3. Basic Method

Viewed over the long time spans of hundreds to thousands of years characteristic of the earthquake cycle, it has long been observed that earthquake mainshocks are recurrent at quasiperiodic intervals (Scholz, 1990). For example, in the Nankai Trough near southwest Japan, the average interval between mainshocks is 180 ± 67 years. For the San Andreas fault at Pallet Creek in the Big Bend region of southern California, the interval is 131 ± 10 years. Elsewhere along the San Andreas, deviations from the mean are more significant (Scholz (1990)).

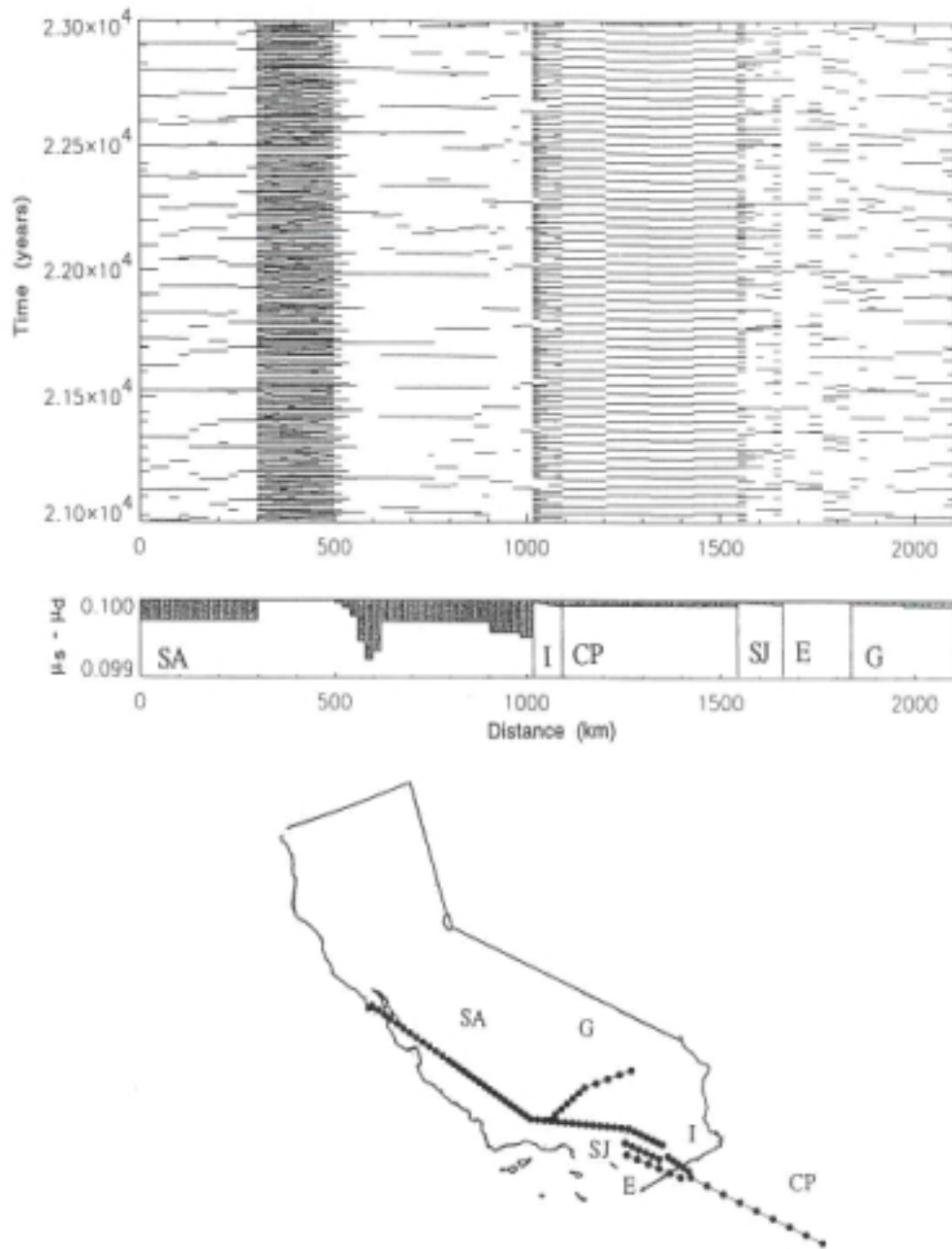


Figure 3. (Top). Time-distance plot of simulated seismic activity. $N=80$ fault segments (abbreviations in text) are plotted end to end, with north to left in all cases. Horizontal line over a given segment represents slip of that segment at time indicated. (Middle). Plot of difference in friction coefficients $\mu_s - \mu_d$ as a function of distance along faults. (Bottom). Map view of fault segments shown superposed on a map of California. A segment is represented by the interval between dots.

At the famous Parkfield, California site, as well as along the Alaska-Aleutian trench, and along the central and South American subduction zones, average recurrence intervals are also well defined. However, temporal and spatial clustering is also evident in these data, as has been discussed elsewhere (Scholz, 1990; Kagan and Jackson, 1991; Gross and Kisslinger, 1994; Rundle et al., 1996; Sornette and Knopoff, 1997) . Analysis in refs. Gross and Kisslinger (1994) and Rundle et. al. (1996) suggests that neither the recurrent nature of the mainshocks, nor the observed spatial-temporal clustering of activity, is compatible with a pure Poisson probability function.

Nevertheless, if we were to attempt the construction of a probability density function $P(\mathbf{x}_i, t)$ to forecast mainshocks on a fault segment \mathbf{x}_i , a logical approach would be to measure the average frequency of events on \mathbf{x}_i by calculating:

$$\omega_i = \pi v_i = \pi \left\{ \frac{\text{Number of Events on } i^{\text{th}} \text{ Segment}}{\text{Time Interval}} \right\} \quad (3)$$

and then write an expression such as:

$$P(\mathbf{x}_i, t) = Z_i^{-1} \{ \cos \omega_i (t - t_0) \}^2 \quad (4)$$

The constant t_0 would be determined as the time of the latest slip event, and Z_i is a normalizing factor. This and other approaches have been examined in the literature (Nijhout et al., 1997; Gray et al., 1997; Shatz et al., 1997; Constantine-Paton et al., 1990; Abeles, 1991; Gross and Rundle, 1998; Crutchfield and Hanson, 1993; Eneva and Ben-Zion, 1997a,b; Pepke and Carlson, 1997; Pepke et al., 1994; Kossobokov and Carlson, 1995; Minster and Williams, 1992; Molchan and Kagan, 1992; Sammis et al., 1996), but success has so far been mixed.

We have therefore been motivated to search for an alternative method in which $P(\mathbf{x}_i, t)$ is constructed by taking account of the variability in the seismicity data, as well as space-time correlations that may exist with activity at other segments \mathbf{x}_j . We have therefore developed a method based on a discrete Karhunen-Loeve expansion (Holmes et al., 1996; Funkunaga, 1990; Preisendorfer, 1988) that allows computation of the "normal modes" of the space-time seismicity pattern data from an equal-time correlation operator, yielding a set of eigenvectors $\phi_n(\mathbf{x})$ and a corresponding set of eigenfrequencies ω_n , $n = 1, \dots, N$. If the process is statistically stationary, as has been established for the mean field systems we consider, these $\phi_n(\mathbf{x})$ and ω_n *will themselves be independent of time*. Hence $\phi_n(\mathbf{x})$ and ω_n can justifiably be determined from an

equal time correlation operator. It is important to stress here that, due to the long range nature of the stress transfer Green's function describing real faults, earthquake fault systems can be considered to be mean field systems.

In the case of figure 3, we have $N = 80$ such eigenvectors and eigenfrequencies. From these, we compute the complex-valued, reconstructed seismicity pattern state $\Psi_R(\mathbf{x}_i, t)$:

$$\Psi_R(\mathbf{x}_i, t) = \sum_{n=1, N} \alpha_n e^{-i \omega_n t} \phi_n(\mathbf{x}_i) \quad (5)$$

The constants α_n are determined by fitting $\Psi_R(\mathbf{x}_i, t)$ to the time t_{oi} of the most recent slip event on segment \mathbf{x}_i , $\Psi_R(\mathbf{x}_i, t_{oi}) = 1$ (see equation 21 below). Expression (5) can in principle describe phenomena in a system characterized by a large number of greatly diverse time scales $\tau_i = 2\pi / \omega_i$. Furthermore, equation (5) implies that time evolution of the pattern state $\Psi_R(\mathbf{x}_i, t)$ occurs via rotation in a complex, N -dimensional space spanned by the eigenfunctions $\phi_n(\mathbf{x}_i)$. In this example, seismicity is thus an example of a phase dynamical system (Mori and Kuramoto, 1998).

From $\Psi_R(\mathbf{x}_i, t)$, we take the real, observable part:

$$\Psi_{\text{obs}}(\mathbf{x}_i, t) = \frac{1}{\sqrt{2}} \{ \Psi_R(\mathbf{x}_i, t) + \Psi_R^*(\mathbf{x}_i, t) \} \quad (6)$$

$\Psi_R(\mathbf{x}_i, t)$ is written in terms of a sum of eigenvectors $\phi_n(\mathbf{x})$, which arise from an equal-time correlation operator. A correlation function is related to a probability density function. In this sense, $\Psi_R(\mathbf{x}_i, t)$ thus represents the "square root" of a probability, therefore a probability amplitude:

$$P(\mathbf{x}_i, t) = | \Psi_{\text{obs}}(\mathbf{x}_i, t) |^2 = \{ \Psi_{\text{obs}}(\mathbf{x}_i, t) \}^2 \quad (7)$$

Thus the *probability amplitudes* associated with the "Pattern Dynamics" evolve linearly over the time interval $\{t, t+\Delta t\}$, even though the *underlying dynamics* do not.

As an example, for a model with N independent, non-interacting segments, we have

$$\phi_n(\mathbf{x}_i) = \delta_{ni} \quad (8)$$

where δ_{ni} is the Kronecker symbol. Using (5) - (8) for this case, we find an expression for $P(\mathbf{x}_i, t)$ that is identical to (4). The more interesting question arises when all of the segments in the fault system interact. In that case, it is not clear from the above discussion how the eigenvectors and eigenfrequencies can be computed. We address this topic in the following section.

4. Correlation Operators

When segments of an earthquake fault system or other nonlinear system interact, it is likely that space-time correlations and patterns will develop (see, e.g., Nijhout et al., 1997; Ball, 1999). Correlations that exist in the slip $s(\mathbf{x}, t)$ and stress $\sigma(\mathbf{x}, t)$ state variables lead to correlations in the seismicity pattern variables $\Psi(\mathbf{x}, t)$. To describe the correlations in seismicity, one can define any of a large number of equal time correlation operators. We consider two, a *static* correlation operator, and a *rate* correlation operator. The *static* correlation operator $C(\mathbf{x}_i, \mathbf{x}_j)$ between discrete segments centered at \mathbf{x}_i and \mathbf{x}_j is:

$$C(\mathbf{x}_i, \mathbf{x}_j) \equiv \frac{1}{T} \int_0^T dt \ z(\mathbf{x}_i, t) \ z(\mathbf{x}_j, t) \quad (9)$$

where $z(\mathbf{x}_i, t)$ is the mean-zero, univariant time series obtained from $\Psi(\mathbf{x}_i, t)$. The *rate* correlation operator $K(\mathbf{x}_i, \mathbf{x}_j)$ is:

$$K(\mathbf{x}_i, \mathbf{x}_j) \equiv - \frac{1}{T} \int_0^T dt \ \frac{\partial z(\mathbf{x}_i, t)}{\partial t} \ \frac{\partial z(\mathbf{x}_j, t)}{\partial t} \quad (10)$$

Both of the operators (9) and (10) are symmetric, rank-N operators that can be diagonalized using standard methods. The eigenvalues of $C(\mathbf{x}_i, \mathbf{x}_j)$ represent relative weights (probabilities) p_n (Holmes et al., 1996; Fukunaga, 1990; Preisendorfer, 1988), with corresponding eigenvectors $\chi_n(\mathbf{x}_i)$. The eigenvalues of $K(\mathbf{x}_i, \mathbf{x}_j)$ are the squares of inverse times $(1/\tau_n)^2$, therefore related to frequencies $\omega_n = 2\pi/\tau_n$, with corresponding eigenvectors $\phi_n(\mathbf{x}_i)$. The latter are just the eigenfrequencies and eigenvectors that we use to define the reconstructed pattern state $\Psi_R(\mathbf{x}_i, t)$ in (5). In Appendix B, we discuss the physical significance of the eigenvalues of $K(\mathbf{x}_i, \mathbf{x}_j)$.

Other correlation operators can be defined as well. For example, one can define a time series $y(\mathbf{x}_i, t)$ whose entries are the slip on segment \mathbf{x}_i in the time interval $\{t, t + \Delta t\}$, and whose eigenvalues have the units of slip squared. Or one can define a time series $w(\mathbf{x}_i, t)$ whose entries are the number of events, and whose eigenvalues are the squares of relative frequencies. Or one can define a covariance matrix if $z(\mathbf{x}_i, t)$ is not univariant, and so forth. Any of these possible operators can be examined using the methods we describe, some of which we are in the process of studying. Each have different physical interpretations and yield different information about correlations in the patterns and the processes in the system. The eigenvectors of each contain information related to the spatial correlations between segments.

Since $\Psi(\mathbf{x}, t)$ and $z(\mathbf{x}_i, t)$ represent time series whose elements are impulse functions, it is operationally difficult to compute the time derivatives in (10). Computing $K(\mathbf{x}_i, \mathbf{x}_j)$ is therefore difficult, whereas computing $C(\mathbf{x}_i, \mathbf{x}_j)$ is easy. As a result, we propose an alternative approach based on Fourier transforms:

$$z(\mathbf{x}_i, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\varpi_i e^{i \varpi_i t} \hat{z}(\mathbf{x}_i, \varpi_i) \quad (11)$$

Then:

$$C(\mathbf{x}_i, \mathbf{x}_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\varpi_i d\varpi_j P_*(\mathbf{x}_i, \mathbf{x}_j, \varpi_i, \varpi_j) \quad (12)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\varpi_i d\varpi_j \varpi_i \varpi_j P_*(\mathbf{x}_i, \mathbf{x}_j, \varpi_i, \varpi_j) \quad (13)$$

$P_*(\mathbf{x}_i, \mathbf{x}_j, \varpi_i, \varpi_j)$ is the joint probability density of the two variables ϖ_i, ϖ_j :

$$P_*(\mathbf{x}_i, \mathbf{x}_j, \varpi_i, \varpi_j) = \hat{z}(\mathbf{x}_i, \varpi_i) \hat{z}(\mathbf{x}_j, \varpi_j) \left\{ \frac{1}{2\pi T} \int_0^T dt e^{i(\varpi_i + \varpi_j)t} \right\} \quad (14)$$

The covariance $\langle \varpi_i \varpi_j \rangle$ is then:

$$\langle \varpi_i \varpi_j \rangle = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\varpi_i d\varpi_j \varpi_i \varpi_j P_*(\mathbf{x}_i, \mathbf{x}_j, \varpi_i, \varpi_j)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\varpi_i d\varpi_j P_*(\mathbf{x}_i, \mathbf{x}_j, \varpi_i, \varpi_j)} \quad (15)$$

If we assume that the joint spectra of the time series are simultaneously sharply peaked about the average frequencies $\langle \varpi_i \rangle$ and $\langle \varpi_j \rangle$:

$$\left| \frac{\langle (\varpi_i - \langle \varpi_i \rangle) (\varpi_j - \langle \varpi_j \rangle) \rangle}{\langle \varpi_i \rangle \langle \varpi_j \rangle} \right| \ll 1 \quad (16)$$

we obtain the result:

$$K(\mathbf{x}_i, \mathbf{x}_j) \approx C(\mathbf{x}_i, \mathbf{x}_j) \langle \varpi_i \rangle \langle \varpi_j \rangle \quad (17)$$

The frequencies $\langle \varpi_i \rangle$ are observables, they are simply:

$$\langle \varpi_i \rangle = \pi v_i \quad (18)$$

from (3).

Once $K(\mathbf{x}_i, \mathbf{x}_j)$ is found from (17), it can be diagonalized to obtain the normal modes of the patterns, the eigenfrequencies ω_n and the eigenvectors $\phi_n(\mathbf{x}_i)$. These eigenvectors provide information about specific patterns of spatial correlations between sites, and the eigenfrequencies provide information about how often these recur.

5. Forecast Methodology

Our forecast methodology makes use of equations (5)-(7). Given an observed pattern $\Psi(\mathbf{x}, t)$, we wish to compute $P(\mathbf{x}_i, t + \Delta t)$, the probability density for activity at the next time step. From (6) we have:

$$\Psi_{\text{obs}}(\mathbf{x}_i, t) = \frac{1}{\sqrt{2}} \{ \Psi_R(\mathbf{x}_i, t) + \Psi_R^*(\mathbf{x}_i, t) \}$$

$$= \sum_{n=1,N} \beta_n \cos \{ \omega_n t \} \phi_n(\mathbf{x}_i) \quad (19)$$

Our approach is to use the observed pattern data $\Psi(\mathbf{x},t)$ prior to $t+\Delta t$ to find $\phi_n(\mathbf{x}_i)$ and ω_n from the operator $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)$. Then if the β_n are known, (6) can then be used to compute $P(\mathbf{x}_i, t+\Delta t)$:

$$P(\mathbf{x}_i, t+\Delta t) = | \Psi_{\text{obs}}(\mathbf{x}_i, t+\Delta t) |^2 = \{ \Psi_{\text{obs}}(\mathbf{x}_i, t+\Delta t) \}^2 \quad (20)$$

To fix the N values β_n in (19), we find the most recent time t_{mi} prior to time t at which the segment or cell at \mathbf{x}_i "fires". We then require that:

$$\Psi_{\text{obs}}(\mathbf{x}_i, t_{oi}) = (-1)^m \quad (21)$$

where the first earthquake at site \mathbf{x}_i in the sequence has $m=0$, the next $m=1$, and so forth (both values ± 1 correspond to an event). Condition (21) then leads to a set of $N \times N$ independent equations that can be solved by standard methods. This process is then repeated at time $t + \Delta t$ to calculate $P(\mathbf{x}_i, t+2\Delta t)$, the forecast for the next time $t + 2\Delta t$. One should always take care to use the most recent data to calculate the β_n . As discussed above, we assume that the forecast probability amplitudes evolve linearly over the time from t to $t + \Delta t$.

6. Spatial Correlation Patterns in Southern California Seismicity

As a first application of these methods, we demonstrate that the observed data shown in figure 2 from southern California may be decomposed into pattern basis states $\chi_n(\mathbf{x}_i)$ having relative probabilities p_n , by computing the static correlation operator $\mathbf{C}(\mathbf{x}_i, \mathbf{x}_j)$. For this example, the coarse-grained time interval was .001 years, or approximately 8-hour units, yielding a total of 66,500 time steps over the interval 1932-present (see figure 2). Epicentral locations for seismic events were binned into squares of 0.1° latitude and 0.1° longitude on a side, yielding $N = 1,329$ boxes (sites) \mathbf{x}_i in which at least one event occurred. At each site \mathbf{x}_i and coarse grained time step t_q , $\Psi(\mathbf{x}_i, t_q) = 1$ if at least one event occurred, $\Psi(\mathbf{x}_i, t_q) = 0$ otherwise. The correlation operator $\mathbf{C}(\mathbf{x}_i, \mathbf{x}_j)$ was then computed according to (12), then diagonalized to obtain (real) eigenvectors $\chi_n(\mathbf{x}_i)$ and eigenvalues (probability weights) p_n . The results for the first two eigenvectors $\chi_1(\mathbf{x}_i)$ and $\chi_2(\mathbf{x}_i)$ are shown in figure 4. Values of $\chi_n(\mathbf{x}_i)$ are scaled so that $1 \geq \chi_n(\mathbf{x}_i) \geq -1$. $\chi_n(\mathbf{x}_i) = 1$ is represented by dark blue, $\chi_n(\mathbf{x}_i) = -1$ by dark red, with intermediate values represented by lighter colors.

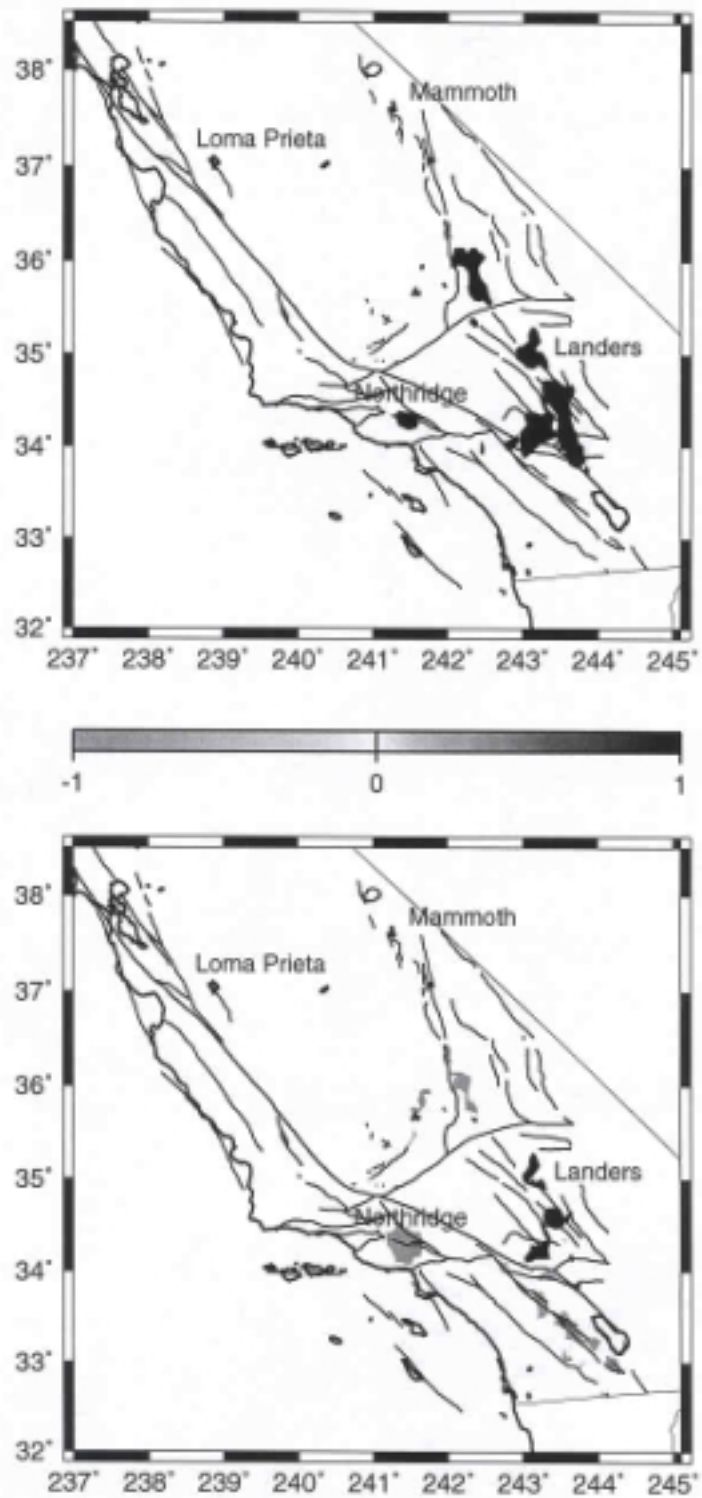


Figure 4. Eigenvectors for southern California seismicity. (Top) Eigenvector $\chi_1(\mathbf{x}_i)$. (Bottom) Eigenvector $\chi_2(\mathbf{x}_i)$.

The physical interpretations of p_n and $\chi_n(\mathbf{x}_i)$ are motivated by the fact that the eigenvalues p_n of a real, symmetric matrix like $\mathbf{C}(\mathbf{x}_i, \mathbf{x}_j)$ are all positive. Therefore, all of the information about correlation between sites must be carried by the eigenvectors $\chi_n(\mathbf{x}_i)$. Since correlation operators are also probabilities, the p_n must represent the probability, or fraction of the variance, that is associated with $\chi_n(\mathbf{x}_i)$. Also, two sites that have the same value, say $\chi_n(\mathbf{x}_i) = \chi_n(\mathbf{x}_j) = 1$, are either always active together, or always quiescent together. The same is true if $\chi_n(\mathbf{x}_i) = \chi_n(\mathbf{x}_j) = -1$ since $\chi_n(\mathbf{x}_i)$ may be multiplied by an overall constant without changing the physical content. On the other hand, if $\chi_n(\mathbf{x}_i) = -\chi_n(\mathbf{x}_j) = 1$, the physical interpretation is that site \mathbf{x}_i tends to be active when \mathbf{x}_j is not, and vice versa. Finally, if $\chi_n(\mathbf{x}_i) \neq 0$ and $\chi_n(\mathbf{x}_j) = 0$, activity at the two sites is uncorrelated. Thus one can use the information encoded in each $\chi_n(\mathbf{x}_i)$ to compare two or more sites for their tendencies to be simultaneously active in a given correlation mode.

Turning to figure 4 (a), we see that eigenvector $\chi_1(\mathbf{x}_i)$ reflects the relative intensity of seismicity in the region, all active sites being red to varying degrees. On the other hand, eigenvector $\chi_2(\mathbf{x}_i)$ in figure 4(b) is more interesting: the sites near the 1992 Landers-Big Bear sequence are all red, while the vast majority of sites in the remainder of southern California are all blue. This eigenvector can be interpreted in either of two ways. The first is that the Landers-Big Bear area is active only when the remainder of southern California is inactive. The second is the "Mogi donut" interpretation -- that the quiescent Landers area is surrounded by active regions. We note that the values of p_n corresponding to figure 4 are $p_1 = .11$, $p_2 = .045$, implying that $\chi_1(\mathbf{x}_i)$ is present in 11% of the southern California seismicity, while $\chi_2(\mathbf{x}_i)$ is present in 4.5% of the seismicity. Also, it should be pointed out that the value of this approach is precisely the reason that all possible spatial patterns are represented in the eigenvectors, which therefore serve as "pattern basis states" for the seismicity data. Many further results will be discussed in Tiampo et al. (1999).

Finally, we note that it is not possible to straightforwardly use this particular seismicity data set to make forecasts using the methods outlined in the previous section. In order to adequately define ω_n and $\phi_n(\mathbf{x}_i)$ from $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)$, at least several cycles of seismic activity are needed at each site, thus several thousand years of data would be necessary to define the $\Psi(\mathbf{x}_i, t_q)$ for the spatial sites shown. Instead, we illustrate below the forecast technique using the relatively realistic southern California simulation described in section 3, for which arbitrarily long synthetic seismicity catalogs can be obtained. This suggests that in order to eventually implement

a forecast technique for a seismically active region such as southern California, one should use the combination of all available instrumental and historic seismicity data, together with a the good simulation data to define as well as possible the ω_n and $\phi_n(\mathbf{x}_i)$.

7. Forecasting Synthetic Seismicity in Southern California

As an illustration of the forecast method discussed in section 5, we use the relatively realistic model of an earthquake fault system described in the appendix. Selected results from tens of thousands of model years of simulations are shown in figure 3, which is a time-distance plot of all the events occurring in 2000 years of simulation data, with a reference map view of the model fault system. Figure 5 shows the events at year 21664, and in figure 6, the events at year 22554. The data shown in figure 3 (Top) were used to construct $\mathbf{C}(\mathbf{x}_i, \mathbf{x}_j)$ and $\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)$, from which were obtained the eigenvectors $\phi_n(\mathbf{x})$ and eigenfrequencies ω_n . The first three spatial eigenvectors for the 6000 year time period preceding the data shown in figure 3 are shown in figures 7-9. It is interesting that between distances ~ 500 km to ~ 1000 km, the eigenvector in figure 7 has 0 spatial nodes, the eigenvector in Figure 8 has 1 spatial node, and the eigenvector in figure 9 has 2 spatial nodes.

The probability $P(\mathbf{x}_i, t)$ for fault segment \mathbf{x}_i to slip at time t was computed using the eigenstates and eigenperiods obtained from a simulation data set over the 6000 years prior to the data to be predicted ("training data set"). The results are shown in figure 10, which shows a 300 year subset of simulation data taken from that shown in figure 3. Figure 10 also shows contours of probability $P(\mathbf{x}_i, t)$ superposed on the previous simulation data. It can be seen that in many cases, there is reasonable agreement between probabilities and the time and location of events, although the "false alarm" rate is high.

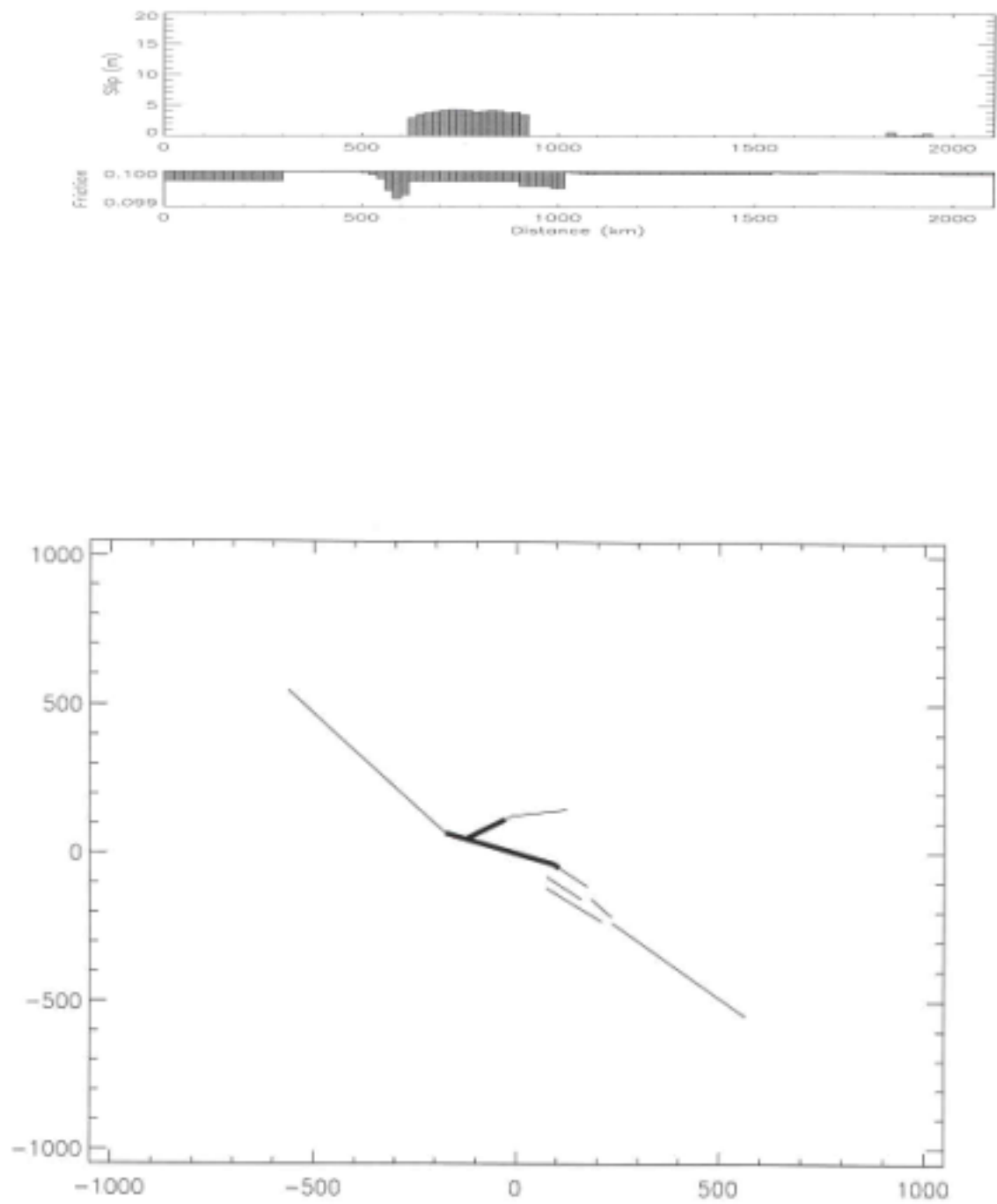


Figure 5. Events at year (time step) 21664. (a). Slip plotted against distance. (b) Map view of events.

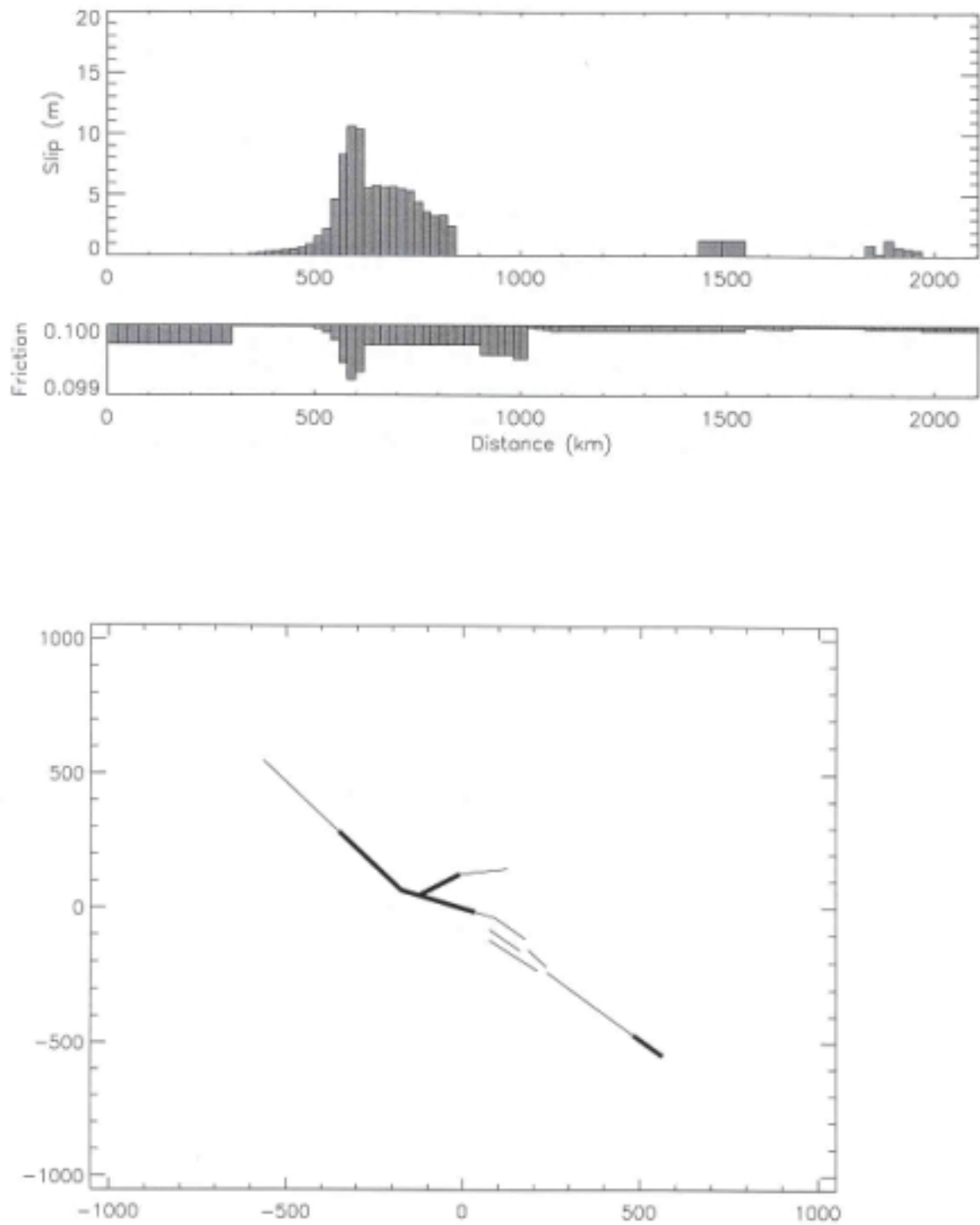


Figure 6. Events at year (time step) 22554. (a). Slip plotted against distance. (b) Map view of events.

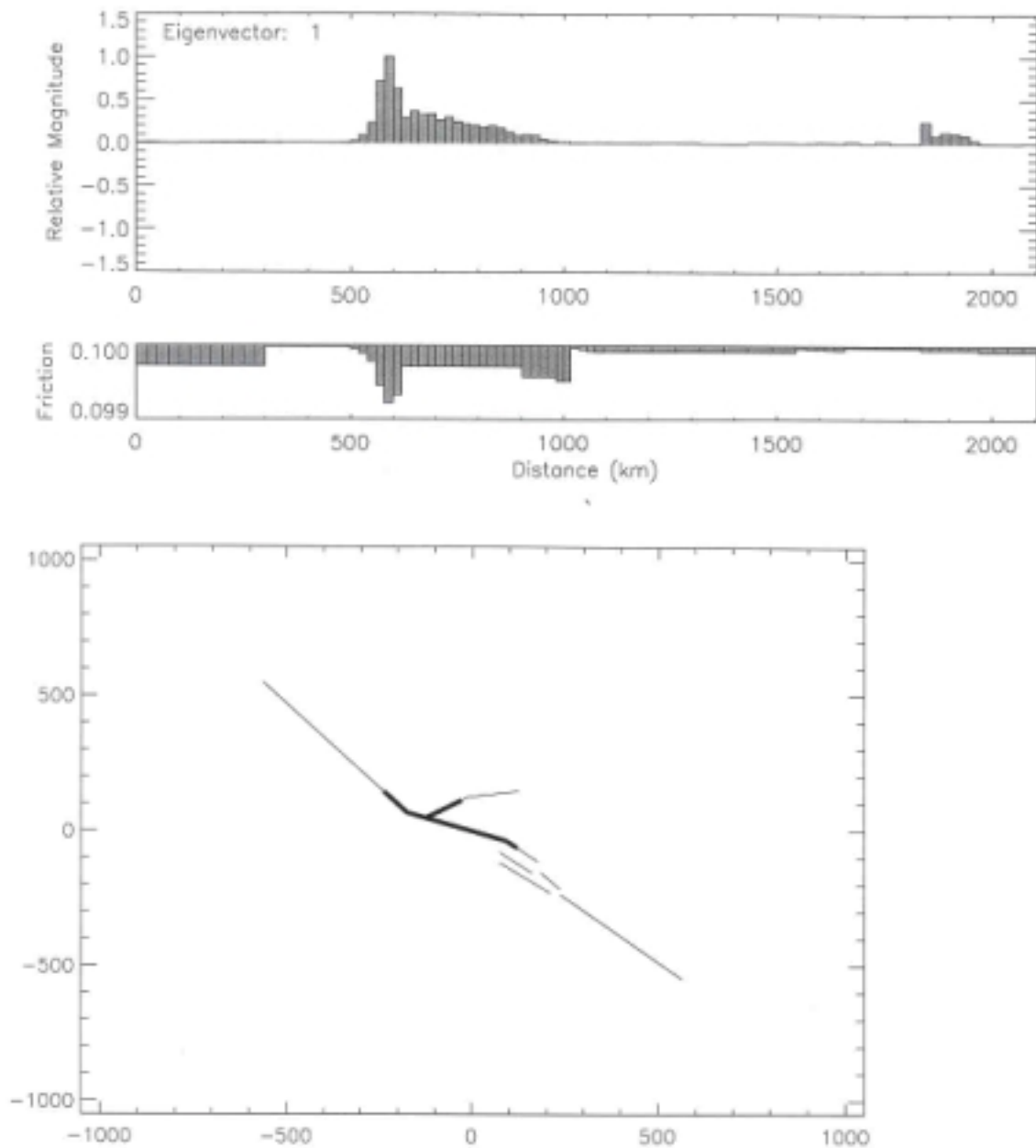


Figure 7 Eigenvector ϕ_1 . (a) Eigenvector plotted against distance. For display purposes, ϕ_1 are scaled so that $\text{Max}\{|\phi_1|\} = 1$. (b) Map view. Values at all segments \mathbf{x}_i at which $\phi_1(\mathbf{x}_i) > .05$ are shown with a bold line. Values at all segments \mathbf{x}_i at which $\phi_1(\mathbf{x}_i) < -.05$ are shown with a dotted line. All other segments shown with a thin line. Period = 804 years.

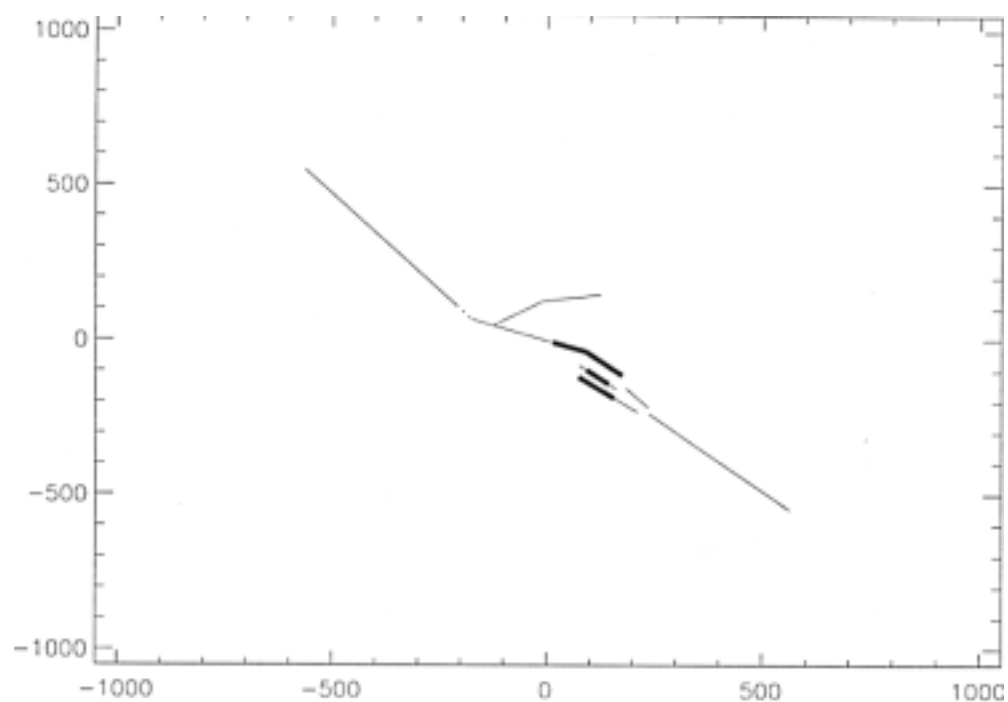
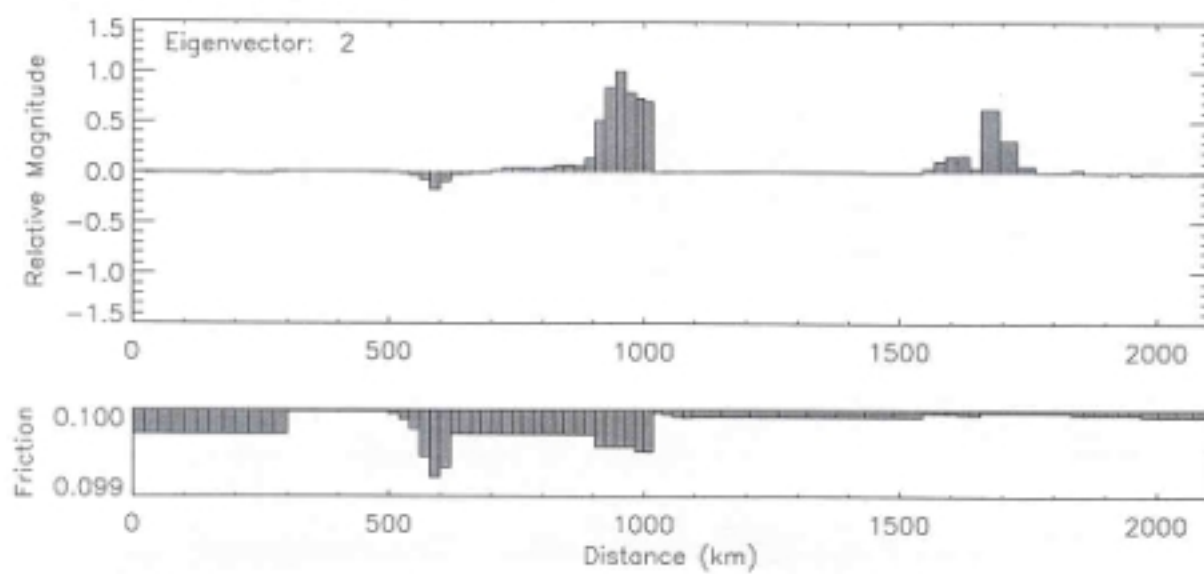


Figure 8. Eigenvector _2. As in figure 5a,b. Period = 680 years.

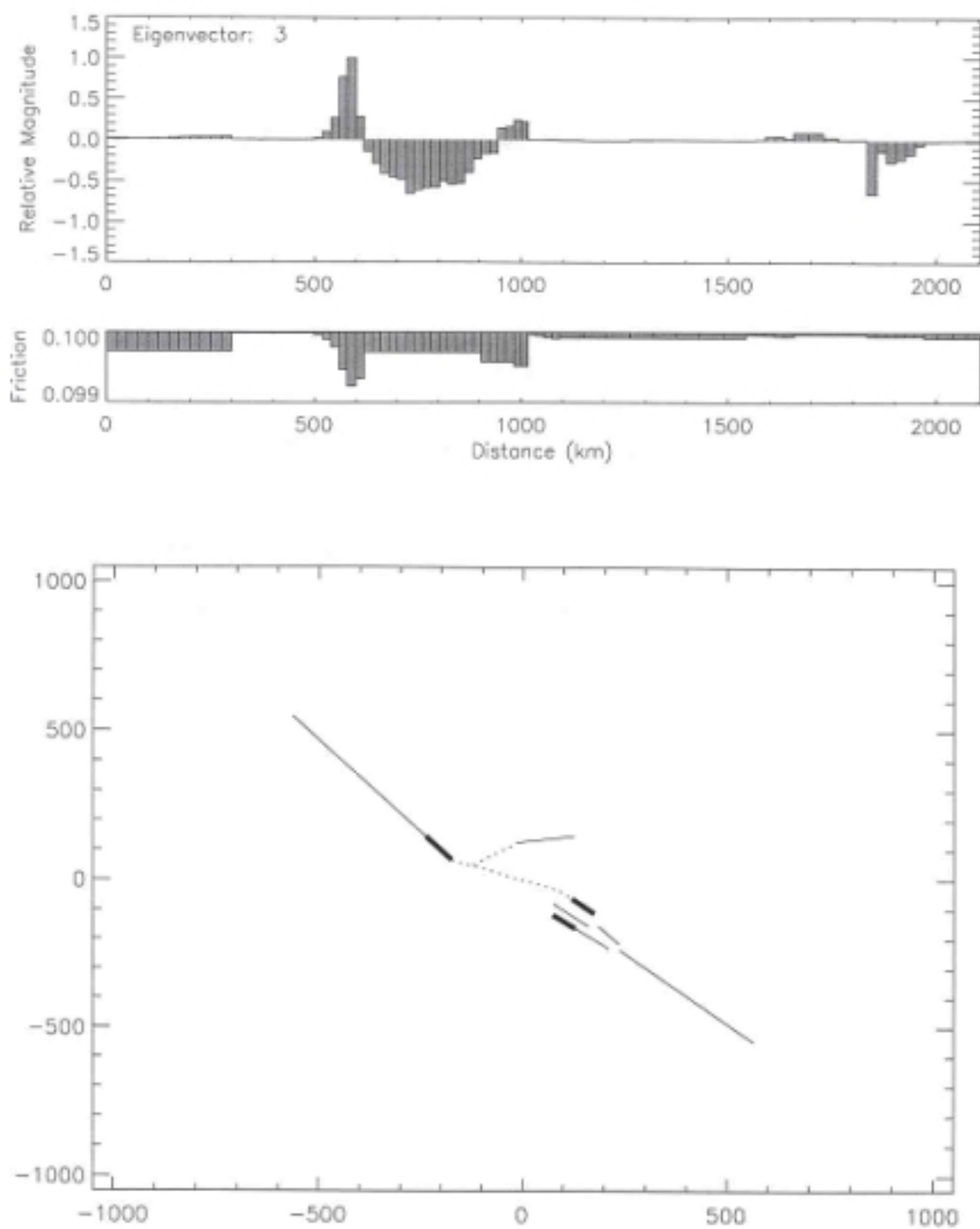


Figure 9. Eigenvector 3 . As in figure 5 a,b. Period = 597 years.

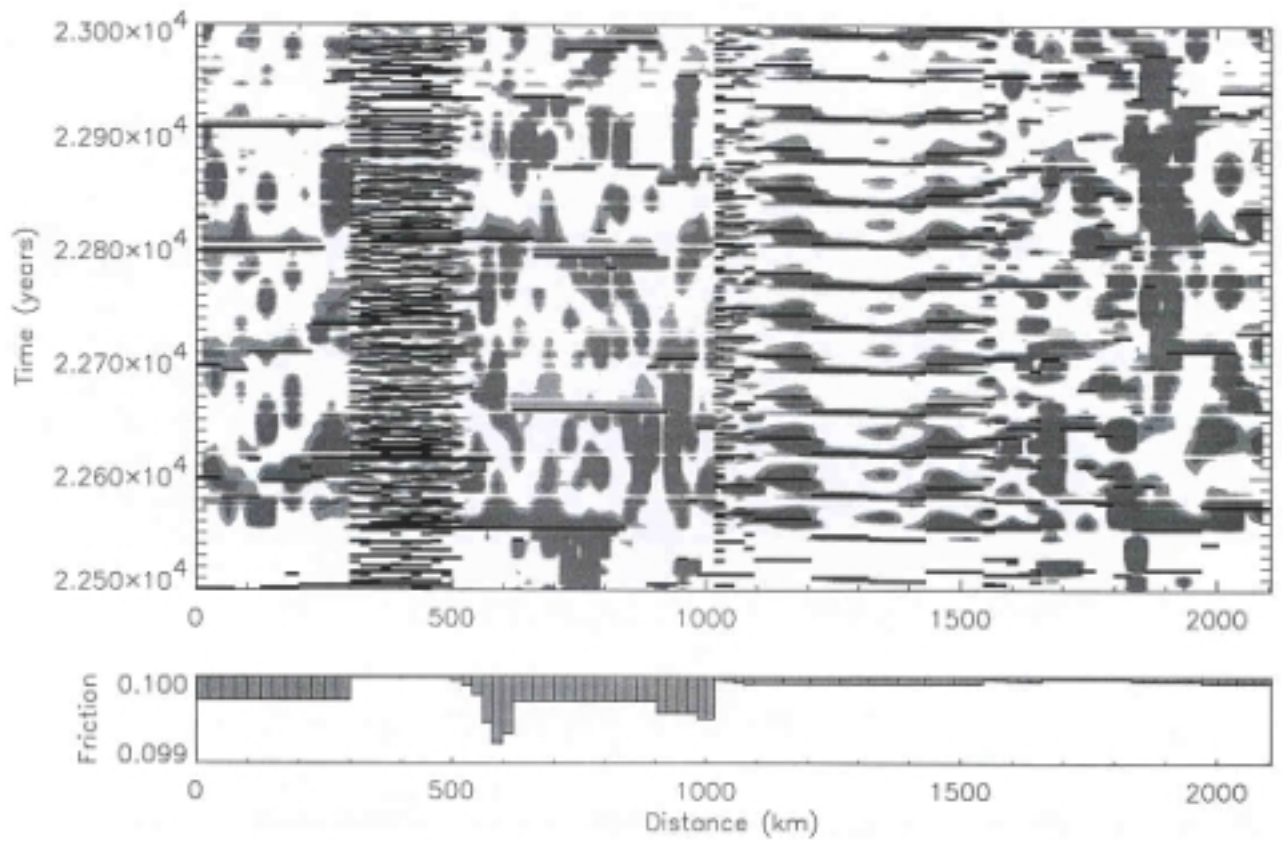


Figure 10. Superposition of events (solid horizontal lines) from year 22,500 to year 23,000 top with probabilities $P(x_i, t)$ calculated according to (20). Light shaded regions are for $P(x_i, t) > 1/3$, dark shaded regions are for $P(x_i, t) > 1/2$. Method used was basic method described in section 5.

8. "Randomized Model" and Uncertainty Principle

Success in forecasting future events using the pattern dynamics method depends on the capability to accurately retrieve both the eigenvectors \mathbf{x}_i , and the eigenfrequencies ω_i . For purposes of comparison, it is useful to have an alternative model, that we call the "Randomized Model", that illuminates trade-offs in different approaches. This new model defines the eigenfrequencies at each site much more precisely, but at the expense of losing important information about the spatial correlation patterns.

The impulsive time series defining the correlation operators consist of signals that are sharply localized in time. However, time series having signals localized in time have a relatively

flat spectrum over a broad range of frequencies. The uncertainty principle for Fourier transforms [28] defined using equation (23) indicates that:

$$\{\langle \omega \rangle, \langle t \rangle\} \geq \frac{1}{2}, \quad (22)$$

This result suggests a simple, but illustrative strategy to improve the frequency resolution of the method described above: that we add a random time increment, drawn from a uniform distribution of times about the original event time, to each event time of the $y_i(t)$. In this way, we degrade the temporal resolution without changing the average frequency ω_i or the mean event time. We should then expect that resolution of normal mode frequencies will improve. However, as remarked above, important information will be lost about the spatial correlations of the various sites.

To test this idea, we defined a set of new random time series $y'_i(t)$, in which the time of each slip event was altered by the addition of a random number uniformly distributed on the interval $[-1/2\omega_i, 1/2\omega_i]$. Because the randomized times are uniformly distributed around the original event times, a memory of the original event times is retained. We then repeated the entire procedure and constructed new forecast probabilities $P(x_i, t)$. The eigenvectors corresponding to the three largest eigenvalues are shown in figures 11, 12, and 13. It can be seen from these figures that better frequency resolution is obtained at the cost of less knowledge about spatial correlations between sites x_i and x_j .

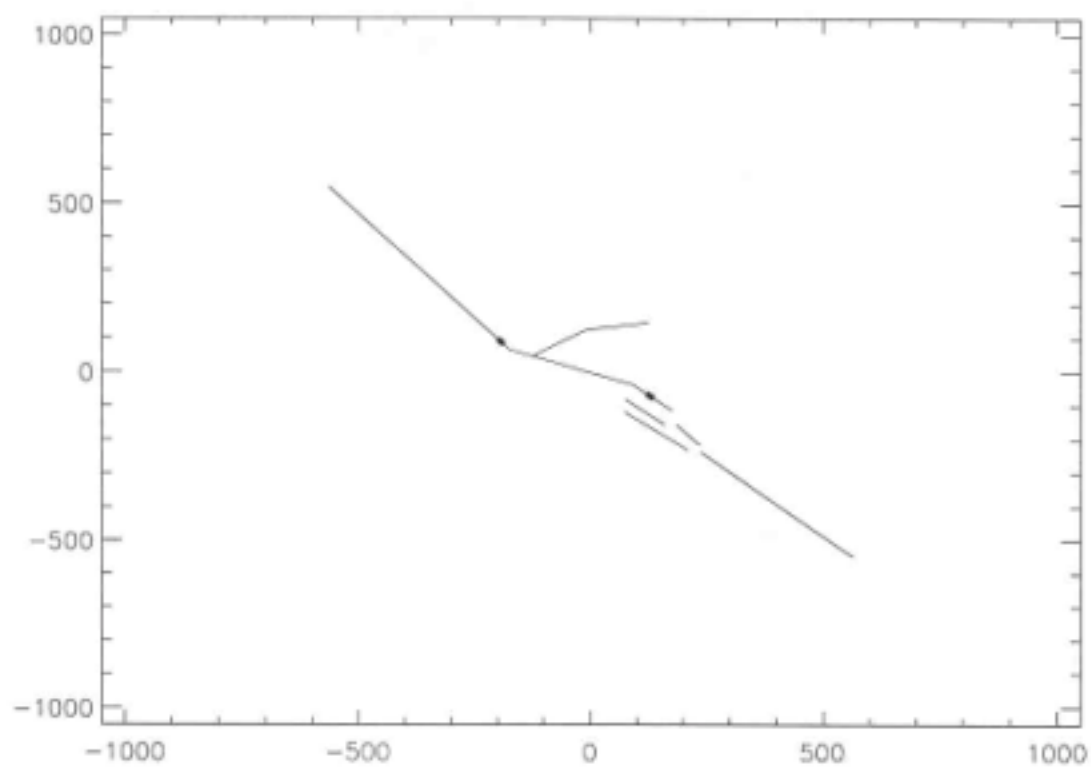
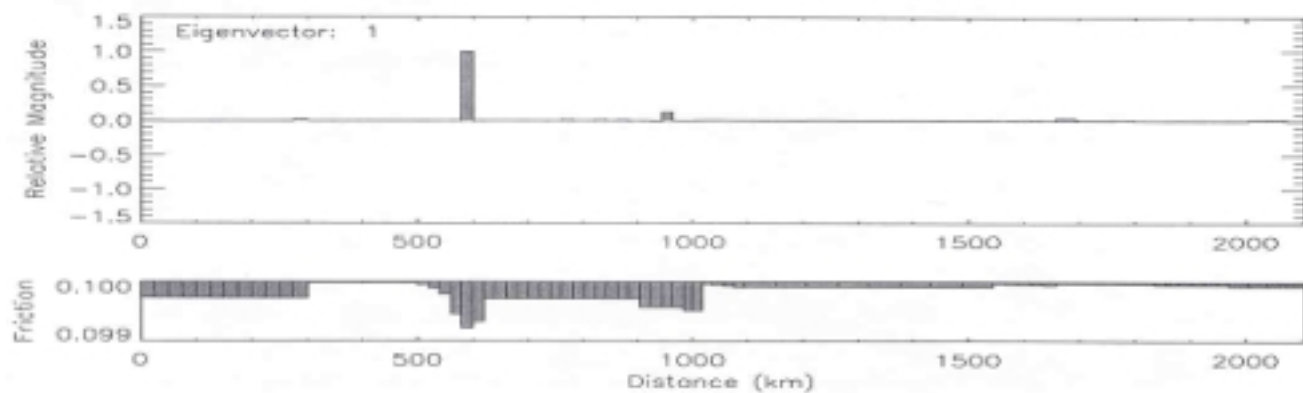


Figure 11. Eigenvector ϕ_1 for randomized model. As in figure 5 a,b. Period = 580 years.

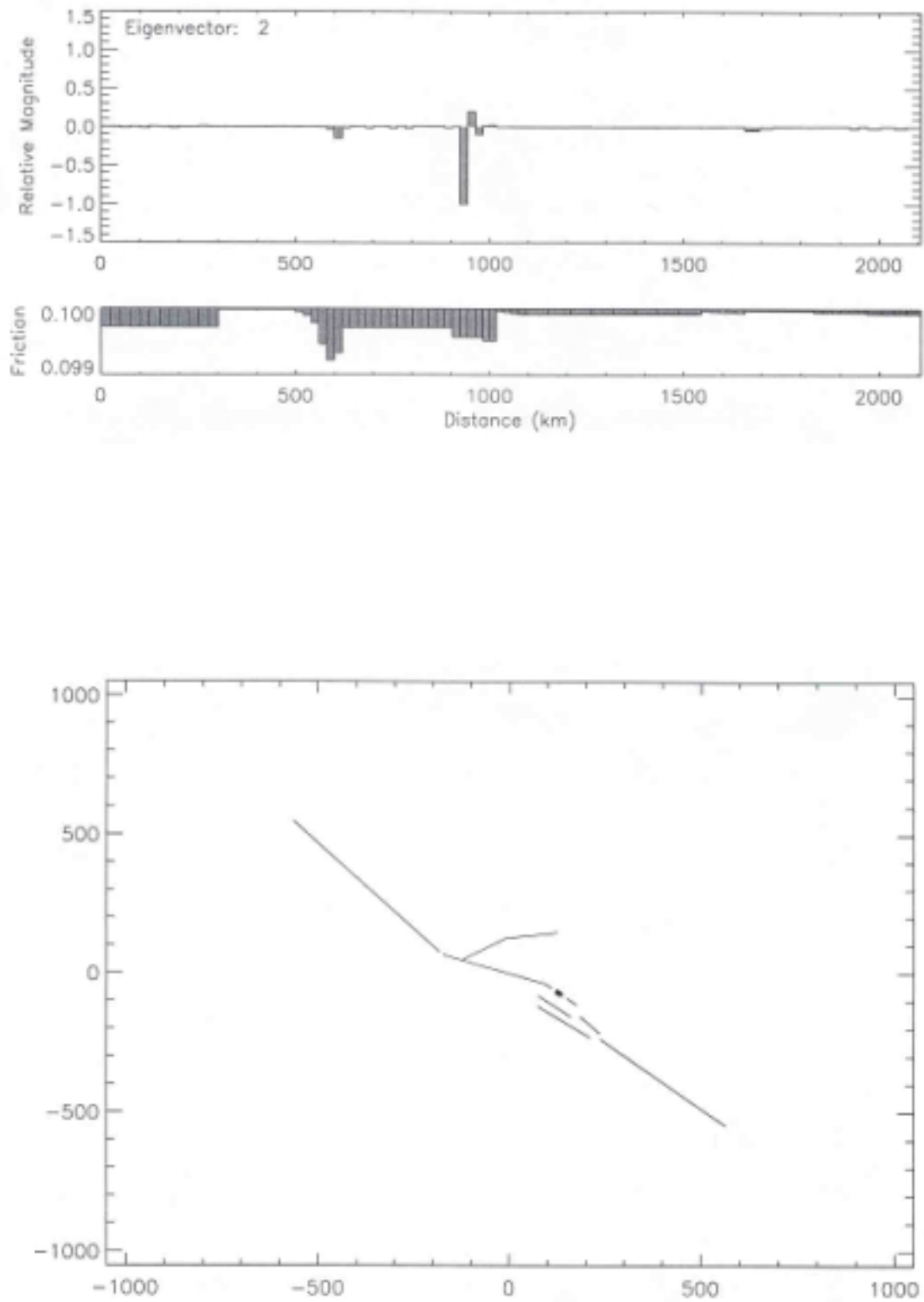


Figure 12. Eigenvector ϕ_2 for randomized model. As in figure 5 a,b. Period = 476 years.

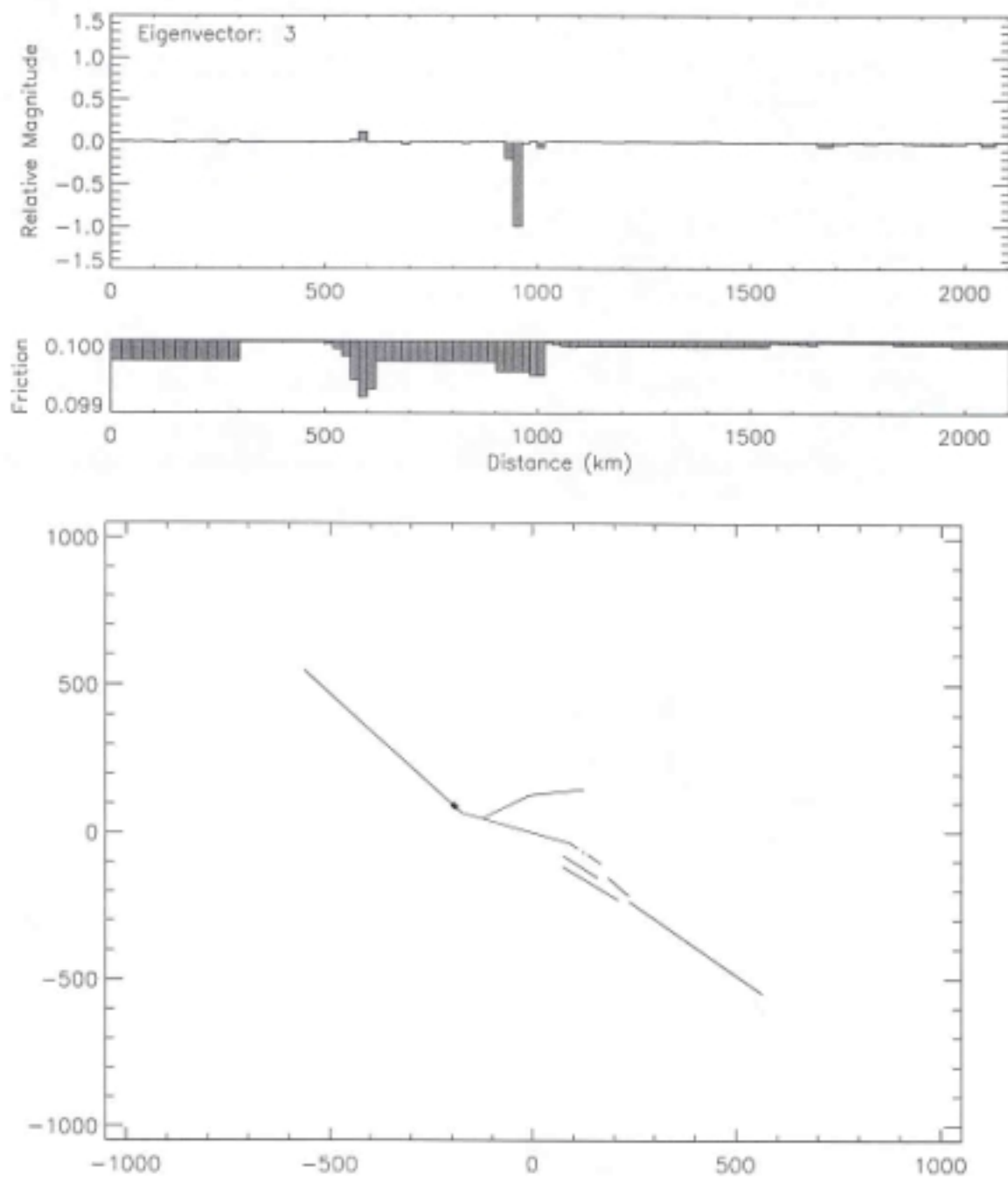


Figure 13. Eigenvector ϕ_3 for randomized model. As in figure 5 a,b. Period = 473 years.

Results for the new forecast probabilities are shown in figure 14 for the same simulation data as was used in figure 10. In several locations, for example at distances ranging from 0 to 300 km, and from 500 to 1000 km, forecasts in figure 14 are superior to those in figure 10. However, the Likelihood ratio test described below indicates that the original method with non-randomized times is more accurate overall.

9. Statistical Test of Forecast Probabilities

The method we describe can be tested by statistical means to assess the quality of the forecast. The predictive skill of a candidate method can be compared, for example, to a standard Poisson model by computing a Likelihood ratio. The same method has been used to evaluate other proposed forecast methods, including time-to-failure models, for real earthquakes (Sammis et al., 1996; Gross and Rundle, 1998). The Likelihood test is based on the idea that a cost is incurred if a forecast method produces a false alarm, i.e. an event during a time interval when no event occurs. Using this test, we find that the Pattern Dynamics (PD) method associates significantly more probability measure with the actual simulation events than does a competing, Poisson probability model.

The Likelihood ratio method is described in detail elsewhere (Gross and Rundle, 1998), and only a brief summary is given here. Consider a given normalized probability density $P(\mathbf{x}, t)$. If the i^{th} of N fault segments slips Q_i times, so that there are N_T total events on all segments, the Likelihood function is:

$$L = \prod_{i=1}^{N_T} \prod_{q=1}^{Q_i} P(\mathbf{x}_i, t_q) \quad (23)$$

We wish to compare the quality of forecasts given by two probability distributions, for example the PD method $P_{PD}(\mathbf{x}_i t)$ and a reference time-independent Poisson model $P_P(\mathbf{x})$. The corresponding Likelihoods are designated as L_{PD} and L_P . To compare the two probability models, we define a composite Likelihood ratio that involves a mixing ratio r :

$$L(r) = \prod_{i=1}^{N_T} \prod_{q=1}^{Q_i} P(\mathbf{x}_i, t_{q,r}) \quad (24)$$

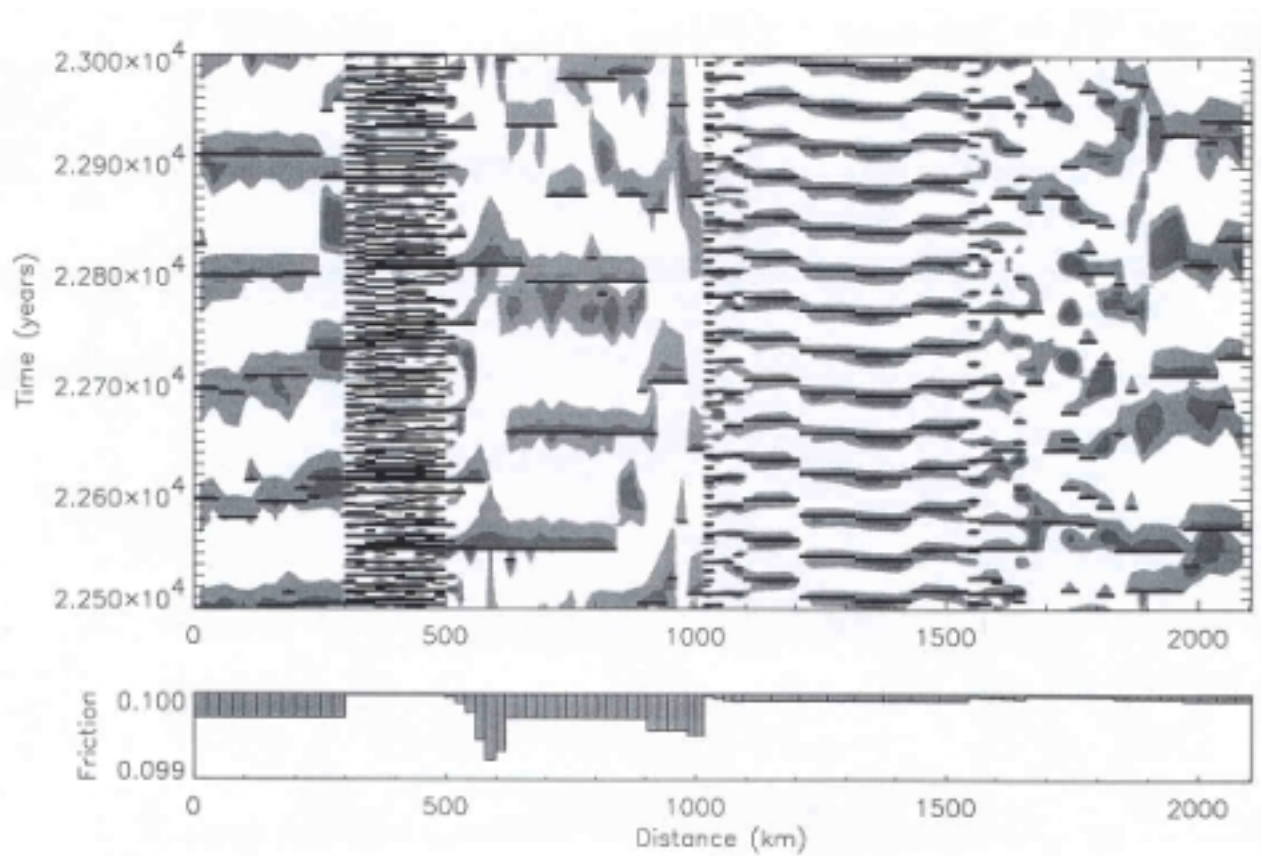


Figure 14. Superposition of events (solid horizontal lines) from year 22,500 to year 23,000 top with probabilities $P(\mathbf{x}_i, t)$ calculated according to (20) using "improvements" described in section 7. Light shaded regions are for $P(\mathbf{x}_i, t) > 1/3$, dark shaded regions are for $P(\mathbf{x}_i, t) > 1/2$.

where:

$$P(\mathbf{x}_i, t_q, r) = r P_{PD}(\mathbf{x}_i, t_q) + (1 - r) P_P(\mathbf{x}_i) \quad (25)$$

Equations (23)-(25) assumes that the earthquakes are all statistically independent, a standard assumption that is clearly not correct in the case of our simulations. However, statistical independence is an assumption routinely used in earthquake forecasting, and the Poisson model represents one of the current most widely used method for calculating earthquake hazard (Scholz, 1990; Gross and Rundle, 1998).

Significance of Results: In figure 15, we plot the log-likelihood ratio $\log_{10}\{L_R(r)\}$, which is defined by:

$$\begin{aligned} \log_{10}\{L_R(r)\} &= \log_{10}\left\{\frac{L(r)}{L_P}\right\} \\ &= \log_{10}\{L(r)\} - \log_{10}\{L_P\} \end{aligned} \quad (26)$$

In figure 15, we show a plot of $\log_{10}\{L_R(r)\}$ against the mixing ratio r , corresponding to the forecasts and simulation data in figures 10 and 14. In both cases:

$$L(r) \gg L_P \quad ; \quad r > 0 \quad (27)$$

implying that $L_{PD} \gg L_P$. These results show that the PD model is substantially better (~ 20 in figure 8, ~ 5 in figure 15 for large r) than the Poisson model at predicting the occurrence times and locations of the synthetic events. The best probability model for both randomized and nonrandomized forecasts are those with the largest r . Cases with $r = 1$ cannot be evaluated: if an event were not predicted, $P_{PD}(\mathbf{x}_i, t_q) = 0$ there. Since L_{PD} is a product of all the probabilities $P_{PD}(\mathbf{x}_i, t_q)$ at the times and locations of the events, L_{PD} would be zero, and $\log_{10}\{L_{PD}\}$ would be undefined.

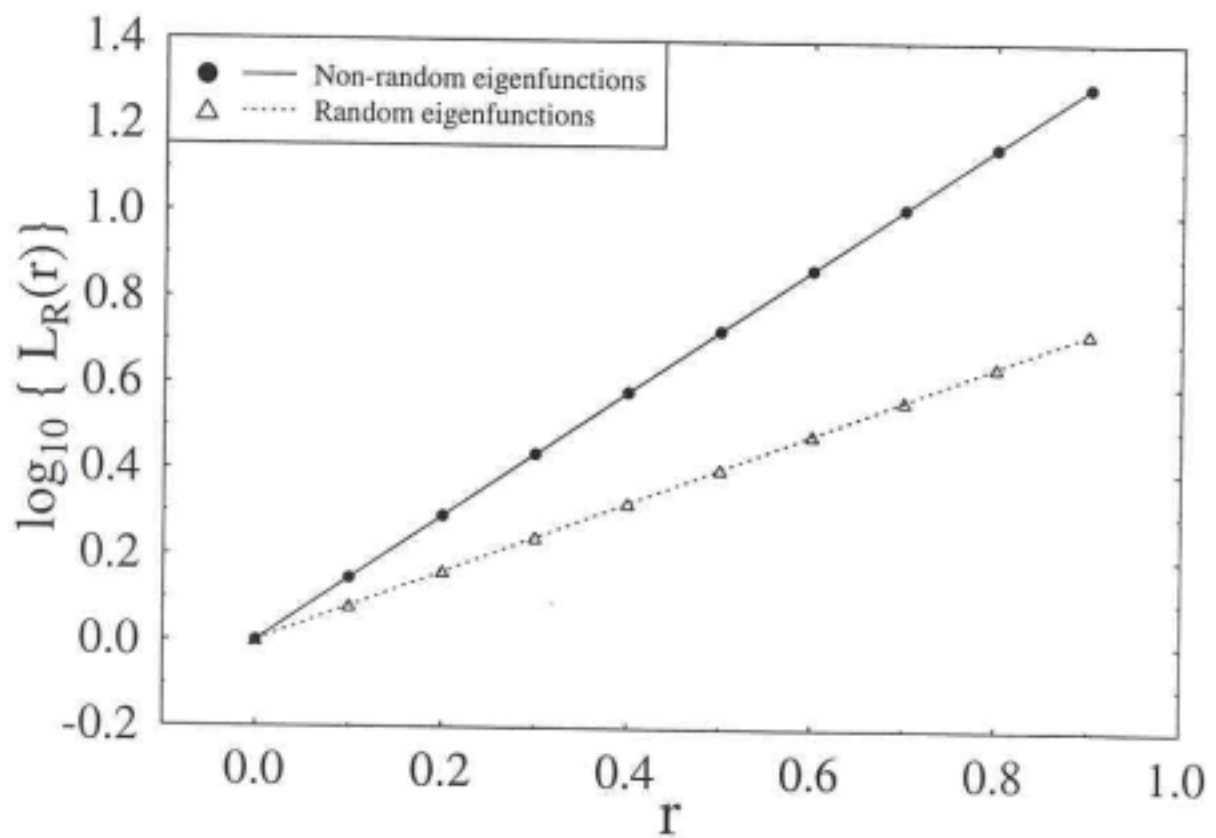


Figure 15. Plot of log-likelihood ratio $\log_{10} \{ L_R(r) \}$ against mixing ratio r corresponding to event forecasts in figures 8 (solid, circles) and 12 (dashed, triangles).

From examination of figures 10 and 14, it can be seen that there exist a number of false positive space-time forecasts in using the probability model $P_{PD}(\mathbf{x}_i, t_m)$, as well as some "failures to predict". However, the result (27) as shown in figure 15 means that $P_{PD}(\mathbf{x}, t)$ still puts far more of its probability measure at the times of the slip events than does $P_p(\mathbf{x})$.

10. Extensions to Observed Seismicity in Southern California

Application of the forecast methods described above to observed seismicity in a region such as southern California will be more difficult, since one does not have available accurate seismicity data spanning many earthquake cycles over thousands of years. For example, in southern California, instrumental records began in 1932 with only a handful of seismographs. This network has been augmented and improved almost continually ever since. However, the methods described in the preceding sections suggest similar approaches that may prove fruitful. A variety of methods based upon learning patterns and using these to forecast future events have been tried, recent examples being those of Keilis-Borok et al. (1988) and Press and Allen (1995).

We can explicitly make use of the idea that earthquakes can be considered as an example of phase dynamics, as in equation (5). Thus evolution of pattern states is via rotations in N -dimensional complex correlation space. This implies that the time average of such pattern states is zero. If we assume that the dynamics is represented by an incoherent superposition of such pattern states over a sufficiently large region, the spatial mean of pattern states will vanish also. Defining $R(\mathbf{x}_i, t)$ as the instantaneous rate of seismicity (number of events of any size) at time t and location \mathbf{x}_i , and $\langle R(t) \rangle$ as the spatial average of $R(\mathbf{x}_i, t)$:

$$\Psi_{obs}(\mathbf{x}_i, t) \approx R(\mathbf{x}_i, t) - \langle R(t) \rangle \quad (28)$$

We also need to account for the fact that the seismicity $R(\mathbf{x}_i, t)$ in a real region such as southern California is an extremely noisy function (e.g., Scholz, 1990), and little information has been obtained about systematic behavior of the amplitude of this function in previous studies. However, over longer time intervals we can expect $S(\mathbf{x}_i, t_0, t)$ to be more stable (figure 2). Thus it is usually practical to deal only with time-averaged quantities. We therefore define a time-averaged seismicity function $S(\mathbf{x}_i, t_0, t)$:

$$S(\mathbf{x}_i, t_0, t) = \frac{1}{t - t_0} \int_{t_0}^t dt' \Psi_{\text{obs}}(\mathbf{x}_i, t') \quad (29)$$

$S(\mathbf{x}, t_0, t)$ thus represents the seismic activity throughout the southern California region over the time interval (t_0, t) . $S(\mathbf{x}, t_0, t)$ is also presumably the real part of a complex-valued seismicity phase function $\Sigma(\mathbf{x}, t_0, t)$, the imaginary part of which is not defined from the data at hand. The function $S(\mathbf{x}_i, 1932, 1998)$ has the same spatial dependence as the temporal average of seismicity given in figure 2, but has zero spatial mean.

Now $S(\mathbf{x}_i, t_0, t)$ can be considered to be an N -dimensional vector $\mathbf{S}(t_0, t)$ that has an inner product and a well-defined L^2 norm, $\|\mathbf{S}(t_0, t)\|$ (e.g., Bachman and Narici, 1966):

$$\|\mathbf{S}(t_0, t)\| = \left\{ \int d\mathbf{x} |\mathbf{S}(t_0, t)|^2 \right\}^{1/2} = \left\{ \int d\mathbf{x} \{ S(\mathbf{x}_i, t_0, t) \}^2 \right\}^{1/2} \quad (30)$$

From these considerations, we can then define $\hat{\mathbf{S}}(t_0, t)$, which can be understood to be the unit vector in N -dimensional correlation space pointing in the direction of $\mathbf{S}(\mathbf{x}, t_0, t)$, with tail fixed at the origin and head lying on the N dimensional unit sphere:

$$\hat{\mathbf{S}}(t_0, t) = \frac{\mathbf{S}(t_0, t)}{\|\mathbf{S}(t_0, t)\|} \quad (31)$$

On the surface, equation (31) may seem to be only a minor alteration of (29), but this is not at all the case. As expressed by the activity function $\mathbf{S}(t_0, t)$, seismicity at \mathbf{x}_i over the interval (t_0, t) depends only upon local conditions near \mathbf{x}_i . As expressed by $\hat{\mathbf{S}}(t_0, t)$ however, seismicity now depends upon conditions everywhere in the system, because the norm $\|\mathbf{S}(t_0, t)\|$ is a non-local function. For a system in which the correlation length is small, the difference between (29) and (31) should be unimportant. However, for a strongly correlated system, the difference may be profound.

In a phase dynamical system, the phase functions $\hat{\mathbf{S}}(t_0, t)$ represent probability amplitudes, so we define the probability associated with the state $\hat{\mathbf{S}}(t_0, t)$ to be:

$$P(\mathbf{x}, t_0, t) = |S(\mathbf{x}, t_0, t)|^2 = |\hat{\mathbf{S}}(t_0, t)|^2 \quad (32)$$

In view of the definition (31) of $\hat{\mathbf{S}}_{(t_0,t)}$, we have the condition that the total probability mass at any time over all sites \mathbf{x} is constant:

$$\int d\mathbf{x} \ P(\mathbf{x}, t_0, t) = 1 \quad (33)$$

The physically meaningful changes in $\hat{\mathbf{S}}_{(t_0,t)}$ occur only via rotations about the origin. Over the time interval $(t_1, t_2 = t_1 + \Delta t)$, such a rotation $\Delta\hat{\mathbf{S}}_{(t_1,t_2)}$ can be denoted by:

$$\Delta\hat{\mathbf{S}}_{(t_1,t_2)} = \hat{\mathbf{S}}_{(t_0,t_2)} - \hat{\mathbf{S}}_{(t_0,t_1)} \quad (34)$$

Furthermore, it can be seen that, as $t_2 \rightarrow t_1$:

$$|\Delta\hat{\mathbf{S}}_{(t_1,t_2)}|^2 \propto (\Delta t)^2 \quad |\Psi_{\text{obs}}(\mathbf{x}_i, t_1)|^2 \quad (35)$$

In view of (7), we see that $|\frac{1}{2} \Delta\hat{\mathbf{S}}|^2$ defines a change in probability $\Delta P(\mathbf{x}, t_1, t_2)$ over the time interval (t_1, t_2) . Moreover:

$$|\frac{1}{2} \Delta\hat{\mathbf{S}}_{(t,t)}|^2 = 0 \quad (36)$$

which implies that the change in total probability mass over all sites \mathbf{x} is zero over any time interval $\Delta t = 0$. Because probability is conserved, we must have the condition:

$$\int d\mathbf{x} \ \Delta P(\mathbf{x}, t_1, t_2) = 0 \quad (37)$$

for all times (t_1, t_2) . These conditions can all be satisfied if we take:

$$\Delta P(\mathbf{x}, t_1, t_2) = |\frac{1}{2} \Delta\hat{\mathbf{S}}_{(t_1,t_2)}|^2 - \mu_P(t_1, t_2) \quad (38)$$

where:

$$\mu_P(t_1, t_2) = \frac{1}{A} \int d\mathbf{x} \quad |\frac{1}{2} \Delta\hat{\mathbf{S}}_{(t_1,t_2)}|^2 \quad (39)$$

and A is the integral over all sites \mathbf{x} .

The physical meaning of these equations is that an increase in ΔP above the background level measured by μ_P should be interpreted as a tendency to form a spatially correlated region of

seismic activity, and that such regions evidently must be present for larger earthquakes to occur. In work to be reported elsewhere (Tiampo et al., 1999) we examine these methods using data from the southern California seismic network over the time interval 1932-present. We find that the probability function $\Delta P(\mathbf{x}, t_1, t_2)$ evidently can be used to detect the formation and existence of regions correlated in space and time that are subsequently the sites of major earthquakes.

11. Final Comments

A variety of methods have been proposed for understanding the space-time patterns in driven nonlinear systems including Principal Component Analysis (Penland and Magorian, 1993), Principal Oscillation Pattern (Preisendorfer, 1988; Penland, 1989; Garcia and Penland, 1991) analysis, and Singular Spectrum Analysis (Fukunaga, 1990; Broomhead and King, 1986; Vautard and Ghil, 1989). However, in these systems, the variables that define the underlying dynamics are directly amenable to observation, in contrast to most driven threshold systems such as neural systems, earthquakes, disordered magnetic systems and so forth. In the latter, only the associated "derivative variables" such as the correlated-event activity (e.g., earthquakes, action potentials, depinning avalanches), can be directly observed. Thus the standard methods cannot be applied to threshold systems without serious qualifications. Other methods that are being examined elsewhere are described in references (Eneva and Ben-Zion, 1997a,b; Pepke and Carlson, 1994; Pepke et al., 1994; Kossobokov and Carlson, 1995; Minster and Williams, 1992; Molchan and Kagan, 1992; Sammis et al., 1996; Ouchi, 1993), but these have so far not yielded adequate results for threshold systems.

Understanding the patterns of activity in earthquake fault systems must proceed via the construction of representative simulations, and the development of analysis methods for the patterns that result. Once the analysis methods are in place, they should be applied to forecasting and understanding data from natural systems. In particular, it is often the case that the largest events are the most infrequent, but the most important to understand, since they usually act to slave the entire system. For earthquake models, these correspond to the largest and most destructive events. Detailed pattern analysis may allow the patterns of smaller, more frequent events leading up to the largest events to be detected and recognized. If this can be demonstrated, reliable forecasting of the largest events may be possible.

Acknowledgements: The authors would like to acknowledge helpful conversations with C. Penland and J. Hopfield. Research by JBR has been supported by NASA grant NAG5-5168 (simulations), and by USDOE grant DE-FG03-95ER14499 (theory), to Univ. of Colorado.

Research by WK has been supported by USDOE grant DE-FG02-95ER14498 to Boston University. Research by KT has been supported by NGT5-30025 (NASA).

12. Appendix A: Earthquake Fault System Model

The dynamical model of the earthquake fault system was developed in Rundle (1988). In contrast to many current models in the literature describing a single planar fault, the model in Rundle (1988) includes all of the major faults in southern California, albeit at a relatively crude scale (figure 3, bottom). In this Cellular Automaton-type model, each of $N = 80$ fault segments is driven at its own particular long term rate V of a few centimeters/year.

Segment i slips when the shear stress $\sigma_{sh,i}(t) = \mu_s \sigma_{norm,i}(t)$, where μ_s is a preassigned coefficient of "static" friction, and $\sigma_{norm,i}(t)$ is the normal stress on the segment. At failure, slip occurs that is sufficient to reduce $\sigma_{sh,i}(t)$ to the value $\sigma_{sh,i}(t) = \mu_d \sigma_{norm,i}(t)$, where μ_d is a coefficient of "dynamic" friction. The difference in frictional coefficients $\mu_s - \mu_d$ is tuned so that realistic event sequences on the middle ("big bend") and southern parts of the model San Andreas (SA) fault occur at realistic intervals of ~ 150 years. Some tuning was also carried out to produce "realistic" sequences on the Imperial (I), the San Jacinto (SJ), the Elsinore (E) and Garlock (G) faults, in contrast to frictional properties on the Cerro Prieto (CP) and Northern San Andreas fault segments, which were not tuned. Time steps of 1 year are used. Interactions between segments allow stress transfer so that one slipping segment can induce others to slip as well in the same event. Realistic directionally-dependent elastic and viscoelastic interactions are included by means of the stress Green's functions. In figure 3 (Bottom) we show a map view of the 80 segments of the model, each of which is a vertical rectangular slip surface. These segments extend to a depth of 18 km in an elastic plate of thickness 30 km. The plate overlies, and is in welded contact with, a Maxwell viscoelastic half space in which shear stress can relax via viscoelastic flow. The reader is referred to Rundle (1988) for other details.

13. Appendix B. Eigenvalues of $K(x_i, x_j)$

In this appendix we give an argument as to why the eigenvalues of the rate correlation matrix $K(x_i, x_j)$ should equal the square of the frequencies of the eigenmodes. The main assumptions we make for this illustration is that each fault segment that can fail has only one frequency, and that the time averaging period can be approximated by infinity.

Consider first the simple case where each fault segment has a different frequency ω_i in the Fourier domain. Referring to equation (11), we can represent each time series as:

$$z(\mathbf{x}_i, t) = \text{Re} \{ z(\mathbf{x}_i, 0) e^{i\omega_i t} \} \quad (\text{B1})$$

In this case, with normalization, the static correlation matrix is the identity matrix. The rate correlation matrix is also diagonal with elements of ω_i^2 . Consequently in this simple case the eigenvectors of the static correlation matrix and the rate correlation matrix are the same, and the eigenvalues of the rate matrix are simply the frequencies.

In the slightly more complicated case where the frequencies can be the same the matrices are no longer diagonal. However, it is simple to see that any two rows have the following property; either they are identical or they are disjoint. By disjoint we mean that the non-zero elements in one row, say row one, are zero in any other row that is not identical to row one. Since the numbering of the fault segments is arbitrary we can renumber so that all of the rows that are identical are grouped together. This will put the matrices in Jordan-normal form. That is the matrices, both static and rate correlation are now made up of blocks of non-zero elements arranged along the diagonal. Each block has the same non-zero element which is simply the frequency squared of some element in the rate correlation matrix and one in the static matrix. Clearly, the eigenvalues are the rate matrix are the square of the frequencies times the eigenvalues of the static matrix.

In the case of real data, it is certainly true that the time averaging period is not infinite and the number of frequencies for an individual fault element is not one. However, there is evidence that we discussed in the body of the paper that supports the point of view that the assumptions we made are good approximations. It is clear that there will be modifications that will need to be made to incorporate additional modes and finite averaging times if we wish to forecast over long times and that the less periodic the individual fault elements are, the more important the corrections will be.

14. References

- Abeles, M. , *Corticonics, Neural Circuits of the Cerebral Cortex*, Cambridge University Press, Cambridge, UK, 1991.
- Bachman, G. and L. Narici, *Functional Analysis*, Academic Press, New York, 1966.
- Bakun, W. , and T.V. McEvilly, Recurrence models and Parkfield, California, earthquakes, *J. Geophys. Res.*, 89, 3051-3058, 1984.
- Ball, P. , *The Self-Made Tapestry, Pattern Formation in Nature*, Oxford University Press, Oxford, UK, 1999.
- Broomhead, D.S. and G.P. King, Extracting qualitative dynamics from experimental data, *Physica D*, 20, 217-236, 1986.
- Constantine-Paton, M. , H.T. Cline. and E. Debski, Patterned activity, synaptic convergence, and the NMDA receptor in developing visual pathways, *Ann. Rev. Neurosci.*, 13, 129-154, 1990.
- Crutchfield, J.P. and J.E. Hanson, Turbulent pattern bases for cellular automata, *Physica D*, 69, 279-301, 1993.
- Eneva, M. and Y. Ben-Zion, Application of pattern recognition techniques to earthquake catalogs generated by model of segmented fault systems in three-dimensional elastic solids, *J. Geophys. Res.*, 102, 24513-24528, 1997a.
- Eneva, M. and Y. Ben-Zion, Techniques and parameters to analyze seismicity patterns associated with large earthquakes, *J. Geophys. Res.*, 102, 17785-17795, 1997b.
- Ferguson, C.D. , W. Klein and J.B. Rundle, Phys. Rev. E., Spinodals, scaling and ergodicity in a model of an earthquake fault with long-range stress transfer, *Phys. Rev E*, 60, 1359-1373, 1999.
- Fisher, D.S., K. Dahmen, S. Ramanathan and Y. Ben-Zion, Statistics of earthquakes in simple models of heterogeneous faults, *Phys. Rev. Lett.*, 78, 4885-4888, 1997.
- Fukunaga, K. , *Introduction to Statistical Pattern Recognition*, Academic Press, NY, 1990. (See especially chapter 9)
- Garcia, A., and C. Penland, Fluctuating hydrodynamics and principal oscillation pattern analysis, *J. Stat. Phys.*, 64, 1121-1132, 1991.
- Gardiner, C.W. , *Handbook of Stochastic Methods*, Springer-Verlag, Berlin, 1985.
- Gray, C.M. , p. 93, in H.F. Nijhout, L. Nadel and D.L. Stein, eds., *Pattern Formation in the Physical and Biological Sciences*, Lecture Notes V, Santa Fe Inst., Addison Wesley, Reading, MA, 1997.
- Gross, S.J. and C. Kisslinger, Tests of models of aftershock rate decay, *Bull. Seism. Soc. Am.*, 84, 1571-1579, 1994.
- Gross, S. and J.B. Rundle, A systematic test of time-to-failure analysis, *Geophys. J. Int.*, 133, 57-64, 1998.

- Haken, H., *Synergetics, An Introduction*, Springer-Verlag, Berlin, 1983.
- Hill, D., J.P. Eaton and L.M. Jones, in R.E. Wallace, ed., *The San Andreas Fault System, California*, US Geol. Surv. Prof. Paper 1515, 1990).
- Holmes, P.J., J.L. Lumley and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, UK, 1996.
- Huang, K., *Statistical Mechanics*, John Wiley, New York, 1987.
- Kagan, Y.Y. and D.D. Jackson, Long term earthquake clustering, *Geophys. J. Int.*, *104*, 117-133, 1991.
- Keilis-Borok, V.I., L. Knopoff, I.M. Rotwain and C.R. Allen, Intermediate-term prediction of occurrence times of strong earthquakes, *Nature*, *371*, 1988.
- Klein, W., J.B. Rundle and C. Ferguson, Scaling and nucleation in models of earthquake faults, *Phys. Rev. Lett.*, *78*, 3793-3796, 1997.
- Kossobokov, V. and J.M. Carlson, Active zone size versus activity: a study of different seismicity patterns in the context of the prediction algorithm M8, *J. Geophys. Res.*, *100*, 6431-6441, 1995.
- Minster, J.-B. and N. Williams, M8 intermediate-term earthquake prediction algorithm; performance update for $M > 7.5$, 1985-1996, *Trans. Am. Geophys. Un. EOS*, *73(43)*, Fall Mtg. Suppl., 456-456, 1996.
- Molchan, G.M. and Y.Y. Kagan, Earthquake prediction and its optimization, *J. Geophys. Res.*, *97*, 4823-4838, 1992.
- Mori, H and Y. Kuramoto, *Dissipative Structures and Chaos*, Springer-Verlag, Berlin, 1998.
- Nijhout, H.F., L. Nadel and D.L. Stein, *Pattern Formation in the Physical and Biological Sciences*, Lecture Notes V, Santa Fe Inst., Addison Wesley, Reading, MA, 1997.
- Ouchi, T., Population dynamics of earthquakes and mathematical modeling, *Pure Appl. Geoph.*, *140*, 15-28, 1993.
- Penland, C., Random Forcing and Forecasting Using Principal Oscillation Pattern Analysis, *Monthly Weather Rev.*, *117*, 2165-2185, 1989.
- Penland, C. and P.D. Sardeshmukh, The Optimal Growth of Tropical Sea Surface Temperature Anomalies, *J. Climate*, *8*, 1999-2024, 1995.
- Penland, C. and T. Magorian, Prediction of Nino 3 Sea Surface Temperatures Using Linear Inverse Modeling, *J. Climate*, *6*, 1067-1075, 1993.
- Pepke, S. L., J.M. Carlson and B.E. Shaw, Prediction of large events on a dynamical model of a fault, *J. Geophys. Res.*, *99*, 6769-6788, 1994.
- Pepke, S.L. and J.M. Carlson, Predictability of self-organizing systems, *Phys. Rev. E*, *50*, 236-242, 1994.

- Preisendorfer, R.W. , Principal Component Analysis in Meteorology and Oceanography, ed. C.D. Mobley, *Develop. Atm. Sci.* 17, Elsevier, Amsterdam, 1988.
- Press, F and C.R. Allen, Patterns of seismic release in the southern California region, *J. Geophys. Res.*, 100, 6421-6430, 1995.
- Rundle, J.B., A physical model for earthquakes; 2, Application to Southern California, *Jour. Geophys. Res.*, 93, 6255-6274, 1988.
- Rundle, J.B. , W. Klein, S. Gross, and D.L. Turcotte, Boltzmann fluctuations in numerical simulations of nonequilibrium lattice threshold systems, *Phys. Rev. Lett.*, 75, 1658-1661, 1995.
- Rundle, J.B. and W. Klein, New ideas about the physics of earthquakes, *Rev. Geophys. Space Phys. Suppl.*, July, 283, 1995.
- Rundle, J.B., W. Klein and S.J. Gross, Rupture characteristics, recurrence, and predictability in a slider-block model for earthquakes, pp. 167-203 in *Reduction and Predictability of Natural Hazards*, Santa Fe. Inst. ser. Sci. Complexity, XXV, ed. J.B. Rundle, D. Turcotte and W. Klein, Addison-Wesley, Reading, MA., 1996.
- Sammis, C.G. , D. Sornette and H. Saleur, pp. 143-156 in *Reduction and Predictability of Natural Hazards*, Santa Fe. Inst. ser. Sci. Complexity, XXV, ed. J.B. Rundle, D. Turcotte and W. Klein, Addison-Wesley, Reading, MA., 1996.
- Scholz, C. H. , *The Mechanics of Earthquakes and Faulting*, Cambridge, UK, 1990.
- Shatz, C.J. , p. 299, in H.F. Nijhout, L. Nadel and D.L. Stein, eds., *Pattern Formation in the Physical and Biological Sciences*, Lecture Notes V, Santa Fe Inst., Addison Wesley, Reading, MA, 1997.
- Sieh, K. , M. Stuiver and D. Brillinger, A more precise chronology of earthquakes produced by the San Andreas Fault in Southern California, *J. Geophys. Res.*, 94, 603-623, 1989.
- Sornette, D. and L. Knopoff, The paradox of the expected time until the next earthquake, *Bull. Seism. Soc. Am.*, 87, 789-798, 1997.
- Tiampo, K., J.B. Rundle, S. McGinnis, W. Klein, and S.J. Gross, Observation of systematic variations in non-local space-time seismicity patterns in southern California prior to major earthquakes, submitted to *Science*, 1999.
- Vautard, R. and M. Ghil, Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, *Physica D*, 35, 395-424, 1989.