Pure appl. geophys. 159 (2002) 2429–2467
0033–4553/02/102429–39 $ 1.50 + 0.20/0

Pure and Applied Geophysics

Pattern Dynamics and Forecast Methods in Seismically Active Regions

KRISTY F. TIAMPO,[1] JOHN B. RUNDLE,[2]
SETH A. MCGINNIS,[1] and WILLIAM KLEIN[3]

Abstract— Large, extended fault systems such as those in California demonstrate complex space-time seismicity patterns, which include repetitive events, precursory activity and quiescence, and aftershock sequences. Although the characteristics of these patterns can be qualitatively described, a systematic quantitative analysis remains elusive. Our research suggests that a new pattern dynamics methodology can be used to define a unique, finite set of seismicity patterns for a given fault system. In addition, while a long-sought goal of earthquake research has been the reliable forecasting of these events, very little progress has been made in developing a successful, consistent methodology. In this report, we document the discovery of systematic space-time variations in seismicity from southern California using a new technique. Here we present examples of this analysis technique on data obtained *prior* to events in seismically active areas that show coherent regions associated with the future occurrence of major earthquakes in the same areas. These results strongly support the hypothesis that seismic activity is highly correlated across many space and time scales within large volumes of the earth's crust.

Key words: Fault system dynamics, pattern dynamics, mathematical methods in geophysics, seismicity.

Introduction

While the historic earthquake record is not complete, it has long been recognized that earthquake mainshocks occur at quasi-periodic intervals and that, for some parts of the world, average recurrence intervals are well defined (KANAMORI, 1981). In addition, both temporal and spatial clustering is evident in the data, with the result that neither the recurrent nature of the mainshocks, nor the observed phenomena such as foreshocks, aftershocks, seismic gaps, or mainshock triggering, is compatible with a Poisson probability function (JONES and HAUKSSON, 1997; KAGAN and

[1] CIRES, University of Colorado, Boulder, CO U.S.A.
E-mails: kristy@fractal.colorado.edu; sethmc@turcotte.colorado.edu
[2] Department of Physics, Colorado Center for Chaos and Complexity, CIRES, University of Colorado, Boulder, CO, 80309, U.S.A, and Distinguished Visiting Scientist, Jet Propulsion Laboratory, Pasadena, CA, 91125, U.S.A. E-mail: rundle@cires.colorado.edu
[3] Department of Physics, Boston University, Boston, MA USA and Center for Nonlinear Science, Los Alamos National Laboratory, Los Alamos, NM U.S.A. E-mail: klein@buphyc.bu.edu

JACKSON, 1992; SAVAGE, 1993; DIETERICH, 1994; GRANT and SIEH, 1994; RUNDLE and KLEIN, 1995; TURCOTTE, 1997; MAIN, 1999b). Much of the recent geophysical research associated with earthquakes themselves has centered on investigating these spatial and temporal patterns in local and regional seismicity data (KANAMORI, 1981). Notable examples include characteristic earthquakes (SWAN *et al.*, 1980; ELLSWORTH and COLE, 1997), repeating earthquakes (BAKUN *et al.*, 1986; MARONE *et al.*, 1995), seismic gaps (HABERMAN, 1981; HOUSE *et al.*, 1981; KAGAN, 1981; KAGAN and JACKSON, 1992; WYSS and WIEMER, 1999), well-defined recurrence intervals (BAKUN and MCEVILLY, 1984; LYZENGA *et al.*, 1991; SAVAGE, 1993), Mogi donuts (MOGI, 1969; MOGI, 1977), temporal clustering (FROHLICH, 1987; PRESS and ALLEN, 1995; DODGE *et al.*, 1996; ENEVA and BEN-ZION, 1997; JONES and HAUKSSON, 1997; RUNDLE *et al.*, 1997; HUANG *et al.*, 1998), 'slow' earthquakes (LINDE *et al.*, 1996; MCGUIRE *et al.*, 1996; KERR, 1998), precursory quiescence (YAMASHITA and KNOPOFF, 1989; WYSS *et al.*, 1996; KATO *et al.*, 1997; WYSS *et al.*, 2000), aftershock sequences (GROSS and KISSLINGER, 1994; NANJO *et al.*, 1998), earthquake triggering over large distances (STEIN *et al.*, 1992; HILL *et al.*, 1993,1995; KING *et al.*, 1994; DENG and SYKES, 1996; GOMBERG, 1996; STARK and DAVIS, 1996; POLLITZ and SACKS, 1997; STEIN, 1999), scaling relations (RUNDLE, 1989; PACHECO *et al.*, 1992; ROMANOWICZ and RUNDLE, 1993; RUNDLE, 1993; SALEUR *et al.*, 1995; RUNDLE *et al.*, 1999), and time-to-failure analyses (BUFE and VARNES, 1993; BOWMAN *et al.*, 1998; GROSS and RUNDLE, 1998; BREHM and BRAILE, 1999; JAUMÉ and SYKES, 1999; MAIN, 1999a). Although much of this work represents important attempts to describe these characteristic patterns using empirical probability density functions, none of these observations or methodologies systematically identifies all possible seismicity patterns. The quantification of all possible space-time patterns would seem to be a necessary first step in the process of identifying which patterns are precursory to large events, leading to the possible development of new approaches in forecast methodology. Yet, as can be seen in Figure 1, a plot of relative southern California seismicity during the time period 1932–1991, the identification and quantification of these patterns is no easy matter.

Recent large earthquakes include the M$\sim$7.4 event that struck Izmit, Turkey in August of 1999, the M$\sim$7.6 Taiwan earthquake that occurred in September of 1999, and the M$\sim$7.1 Hector Mine, California earthquake of October 1999. Many similar examples have been documented over the course of time (RICHTER, 1958; SCHOLZ, 1990), yet despite the fact that the largest of these events span distances exceeding 500 km, no reliable precursors have been detected with any repeatability (KANAMORI, 1981; GELLER *et al.*, 1997). It is difficult for most scientists to understand why events of this magnitude are not preceded by at least some causal process. While various patterns of seismic activity centered on the source region have been proposed, as noted above, these efforts to identify the premonitory signals have focused predominantly on local regions near the earthquake source. As a result, these techniques often require intensive and expensive monitoring efforts and have been largely unsuccessful

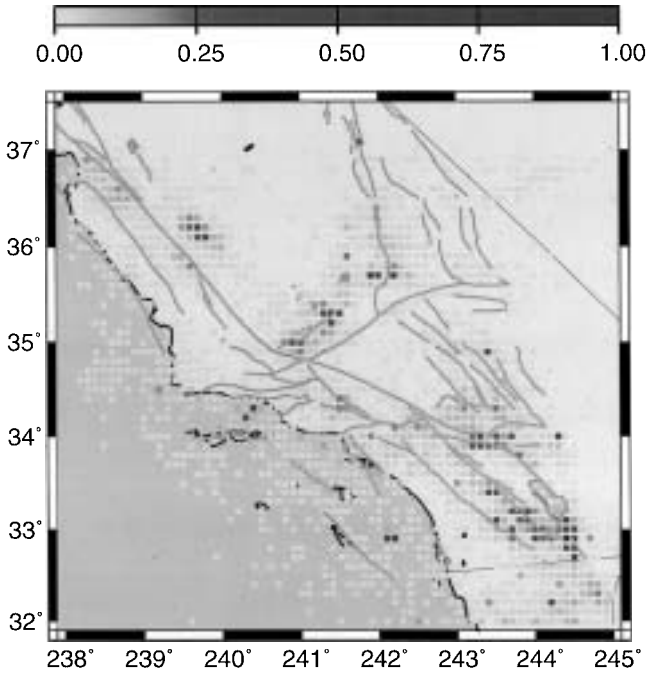

Figure 1
Seismicity for southern California, 1932–1991, normalized to the maximum number of events for the period.

(KANAMORI, 1981). Since these hypothesized patterns are localized on the eventual source region, the fact that one must know or suspect where the event will occur before they can be applied is a major drawback to their implementation.

In this report, we discuss a new pattern dynamics methodology that can be used to define a unique, finite set of seismicity patterns for a given fault system. Similar in nature to the empirical orthogonal functions historically employed in the analysis of atmospheric and oceanographic phenomena (PREISENDORFER, 1988), this method derives the eigenvalues and eigenstates from the diagonalization of a correlation matrix using a Karhunen-Loeve expansion (FUKUNAGA, 1970). This pattern dynamics technique has been successfully applied to the study of numerically modeled seismicity for fault networks similar in character and extent to those found in California (RUNDLE *et al.*, 2000a). We implement this same methodology in order to analyze historical seismicity in California and derive space-time eigenvalue patterns for the San Andreas fault system. The significant eigenstates for this relatively short period of time can be directly correlated with the known California faults and associated events (TIAMPO *et al.*, 2000).

In this work we present a method for identifying these areas of increased probability of an event. The success of the Karhunen-Loeve decompositions

discussed above, coupled with recent observational evidence, suggests that earthquake faults are characterized by strongly correlated space-time dynamics (BUFE and VARNES, 1993; PRESS and ALLEN, 1995; BOWMAN *et al.*, 1998; TIAMPO *et al.*, 2000). We have extended these results and observations of their application to numerical simulations of earthquake fault systems, to the development of a method for identifying areas of increased probability of an event, $\Delta\mathbf{P}$. Realistic numerical simulations of earthquakes also suggest that space-time pattern structures are non-local in character, a consequence of strong correlations in the underlying dynamics (RUNDLE, 1988; RUNDLE *et al.*, 2000a). The procedure described below is based upon the idea that seismic activity corresponds geometrically to the rotation of a pattern state vector in the high-dimensional correlation space spanned by the eigenvectors of an equal-time correlation operator (RUNDLE *et al.*, 2000b; TIAMPO *et al.*, 2000).

Background and Theory

Earthquake fault systems are now thought to be an example of a complex nonlinear system (BAK *et al.*, 1987; RUNDLE and KLEIN, 1995). Interactions among a spatial network of fault segments are mediated by means of a potential that allows stresses to be redistributed to other segments following slip on any particular segment. For faults embedded in a linear elastic host, this potential is a stress Green's function whose exact form can be calculated from the equations of linear elasticity, once the current geometry of the fault system is specified. A persistent driving force, arising from plate tectonic motions, increases stress on the fault segments. Once the stresses reach a threshold characterizing the limit of stability of the fault, a sudden slip event results. The slipping segment can also trigger slip at other locations on the fault surface whose stress levels are near the failure threshold as the event begins. In this manner, earthquakes occur that result from the interactions and nonlinear nature of the stress thresholds.

The Karhunen-Loeve method, a linear decomposition technique in which a dynamical system is decomposed into a complete set of orthonormal subspaces, has been applied to a number of other complex nonlinear systems over the last fifty years, including the ocean-atmosphere interface, turbulence, meteorology, biometrics, statistics, and even solid earth geophysics (HOTELLING, 1993; FUKUNAGA, 1970; AUBREY and EMERY, 1983; PREISENDORFER, 1988; SAVAGE, 1988; PENLAND, 1989; VAUTARD and GHIL, 1989; GARCIA and PENLAND, 1991; PENLAND and MAGORIAN, 1993; PENLAND and SARDESHMUKH, 1995; HOLMES *et al.*, 1996; MOGHADDAM *et al.*, 1998). The notable success of this method in analyzing the ocean-atmosphere interface and such features as the El Niño Southern Oscillation (ENSO), a nonlinear system whose underlying physics is governed by the Navier-Stokes equation, suggested its application to the analysis of the earthquake fault system (NORTH, 1984; PREISEN-

DORFER, 1988; PENLAND and MAGORIAN, 1993; PENLAND and SARDESHMUKH, 1995). Building on these methods for analyzing nonlinear threshold systems, space-time seismicity patterns can be identified in both numerical simulations using realistic earthquake models for southern California (BUFE and VARNES, 1993; BOWMAN *et al.*, 1998; GROSS and RUNDLE, 1998; BREHM and BRAILE, 1999; JAUMÉ and SYKES, 1999; RUNDLE *et al.*, 2000a) and actual historic seismicity records (TIAMPO *et al.*, 1999, 2000). In this paper we apply this Karuhunen-Loeve expansion (KLE) technique (FUKUNAGA, 1970; HOLMES *et al.*, 1996) to the analysis of observed seismicity data from southern California in order to identify basis patterns for all possible space-time seismicity configurations. These basis states represent a complete, orthonormal set of eigenvectors and associated eigenvalues, obtained from the diagonalization of the correlation operators computed for the regional historic seismicity data, and, as such, can be used to reconstitute the data for various subset time periods of the entire data set.

Variables in many dynamical systems can be characterized by a phase function that involves both amplitude and phase angle (MORI and KURAMOTO, 1998). Our simulations have suggested that seismicity can be described by pure phase dynamics (MORI and KURAMOTO, 1998; RUNDLE *et al.*, 2000a, 2000b), in which the important changes in seismicity are associated primarily with rotations of the vector phase function in a high-dimensional correlation space (FUKUNAGA, 1970; HOLMES *et al.*, 1998; RUNDLE *et al.*, 2000a). Changes in the amplitude of the phase function are unimportant, or not relevant. Examples of pure phase dynamical systems in the classical world include weak turbulence in fluids and reaction-diffusion systems (MORI and KURAMOTO, 1998). Another nonclassical example is a quantum system in which the wave function is the phase function. By mapping our problem into the mathematics of quantum mechanics, we are treating the underlying stress-strain dynamics as the hidden variables and the seismicity patterns as the wave functions (RUNDLE and KLEIN, 1995; RUNDLE *et al.*, 2000a, 2000b).

Observations and numerical simulations suggest that space-time patterns of seismic activity directly reflect the existence of space-time correlations in the underlying stress and strain fields (RUNDLE *et al.*, 2000a, 2000b). A spatially coherent, uniformly high-level of stress on a fault is a necessary precondition for the occurrence of a large earthquake. Recently, several groups have found that spatial coherence in the stress field is reflected in a similar coherence in the seismic activity (RUNDLE, 1988; BUFE and VARNES, 1993; KNOPOFF *et al.*, 1996; MAIN, 1999a; SALEUR *et al.*, 1996; BREHM and BRAILE, 1998; RUNDLE *et al.*, 2000a, 2000b). It should therefore be possible to compute the increase in probability of observing such an anomalous correlation, $\Delta\mathbf{P}$, directly from the observed seismicity data. Using the fact that seismicity is an example of phase dynamics, it follows that $\Delta\mathbf{P}$ can be calculated from the square of the phase function for the associated pattern state vector (FUKUNAGA, 1970; HOLMES *et al.*, 1998; RUNDLE *et al.*, 2000a). To emphasize the connection to phase dynamics, we call the function $\Delta\mathbf{P}$ the Phase Dynamical Probability Change (PDPC).

Finally, we note that the space-time correlations or patterns that lead to a uniformly high stress field on the fault represent emergent space-time structures, which evidently form and evolve over time intervals of years preceding the mainshock. Longer time intervals and larger correlated areas should be associated with larger mainshocks.

Data

The primary seismicity data set for southern California employed in this analysis is the entire Caltech catalog from 1932 through December of 1999, obtained from the Southern California Earthquake Center (SCEC) database, with all blast events specifically removed from the catalog [http://www.scecdc.scec.org.]. Relevant data include location, in latitude and longitude, and the time the event occurred. Seismic events between $-115°$ and $-122°$ longitude and 32° and 37° latitude were selected, and all quality events were acquired. Separate analyses were performed for the entire data set, consisting of all events of magnitude greater than or equal to 0.0, and on another data set in which only those events of magnitude greater than or equal to 3.0 were included in the binning process described below. The time periods evaluated spanned 1932 to 1991, and 1932 through August of 1999. In both cases, the seismicity was binned into squares of 0.1° latitude and 0.1° longitude to a side, and a time series constructed for each location square, boxes of approximately 11 km to a side. Each time step is given an initial value of 1.0 if one or more events occur in that time period, or a value of 0.0 otherwise. Subsequently, the mean for each time series is removed from the data.

For the time period 1932 through August of 1999, the seismicity was analyzed using the entire data set, including the entire areal extent and events of all quality. The time interval for this decomposition was increased to one day, so that the total number of time steps is approximately 24,333. In addition, all locations from the entire database, and all quality events, were included, even those where no event occurred for the more than 67 years. The number of location time series affected by the seismicity, p, therefore is 3621. For the time period 1932 to 1991, the time interval for the analysis was again one day, so that the total number of time steps is approximately 21,535. Again, all locations, and all quality events were included, even those where no event occurred for the entire 59 years. The total number of location time series remains constant at 3621.

Method

Karhunen-Loeve Decomposition

Pattern evolution and prediction in nonlinear systems is complicated by nonlinear mode coupling and noise, however understanding such patterns, which are the

surface expression of the underlying dynamics, is critical to understanding and perhaps characterizing the physics which control the system. Karhunen-Loeve expansion (KLE) methods can be used to define a unique, complete pattern basis set for a given dynamical system (FUKUNAGA, 1970; NORTH, 1984; PENLAND, 1989; HOLMES *et al*., 1996). For driven threshold systems, an adaptation of these KLE methods can be employed to characterize both the space-time patterns of threshold transitions, i.e., "firings," as well as the underlying, usually unobservable Markov variables that define the dynamics (FUKUNAGA, 1970; HOLMES *et al*., 1996; RUNDLE *et al*., 2000). In either case, the patterns are defined by the eigenstates and eigenvalues of one of an appropriately constructed family of correlation operators.

Earthquake fault systems are examples of driven nonlinear threshold systems, comprised of interacting spatial networks of statistically identical, nonlinear units that are subjected to a persistent driving force (SCHOLZ, 1990; RUNDLE and KLEIN, 1995; FISHER *et al*., 1997; RUNDLE *et al*., 1997; FERGUSON *et al*., 1999). Numerous examples of such systems exist, including neural networks (HERTZ *et al*., 1990; HERZ and HOPFIELD, 1995), sandpiles (BAK *et al*., 1987), and superconductors (FISHER *et al*., 1997), of which earthquakes are but another example. Such systems re composed of cells which fire, or fail, when the driving force causes the force or potential, $\sigma(\mathbf{x},t)$, on a cell at location $\mathbf{x}$ and time t, to reach a predefined threshold value $\sigma^{\mathbf{F}}$. The behavior of these systems is determined by parameters such as threshold values, residual stresses, quenched disorder and noise (BAK *et al*., 1987; RUNDLE and KLEIN, 1995; FISHER *et al*., 1997). Complex spatial and temporal firing patterns result that are difficult to analyze deterministically (OUCHI, 1993; NIJHOUT, 1997). In the case of an earthquake fault system, the driving force is tectonic plate motion, and the internal potential is the stress on each fault cell or patch. The firing, or failure of each patch results in an increase in the internal state variable $s(\mathbf{x},t)$ and a decrease in the cell potential to some residual value $\sigma^{\mathbf{R}}$. The interactions between the cells, or fault patches, may be excitatory, bringing another closer to failure, or inhibitory, in which the failure of one cell can move neighboring cells further from failure. The spatial and temporal firing patterns, $\psi(\mathbf{x},t)$, of these driven threshold systems are complex and often difficult to understand and interpret from a deterministic perspective, as these patterns are emergent processes that develop from the obscure underlying structures, parameters, and dynamics of a multidimensional nonlinear system (OUCHI, 1993; NIJHOUT, 1997).

Analysis of a number of these driven threshold systems often is complicated by the fact that the underlying dynamics and the state variables which control the physics of the system, $s(\mathbf{x},t)$, are unknown and difficult to observe. The earthquake fault system is no exception. While it is not only probable, but essential, that space-time patterns and correlations exist in the variables and interactions which control earthquake dynamics, $s(\mathbf{x},t)$ and $\sigma(\mathbf{x},t)$, from which the observable surface patterns and correlations, $\psi(\mathbf{x},t)$ arise, those true dynamical patterns are difficult or impossible to observe within the earth (SCHOLZ, 1990; TURCOTTE, 1997). The

schematic shown in Figure 2 illustrates the physical problem. As the state variable $s(\mathbf{x},t)$ at a particular location $\mathbf{x}$ evolves in time under the deterministic dynamics $\mathbf{D_t}$ to a value $s(\mathbf{x},t + \Delta t)$, the force or potential $\sigma(\mathbf{x},t)$, also evolves to $\sigma(\mathbf{x},t + \Delta t)$. While the values of $s(\mathbf{x},t)$ and $\sigma(\mathbf{x},t)$, along with the specifics of $\mathbf{D_t}$, are hidden from view below the dashed line of Figure 2, the firing patterns of $\psi(\mathbf{x},t)$ are observable. In the earth, there is no means at present to measure the stress and strain at every point in an earthquake fault system, or the constitutive parameters, which characterize the heterogeneous medium and its dynamics. However, the seismicity, which is the surface expression of its firing activity, can be located in both space and time with considerable accuracy (BAKUN and MCEVILLY, 1984; SIEH *et al.*, 1989; HILL *et al.*, 1990). For example, the firing activity, $\psi(\mathbf{x},t)$, can be represented as a set of time series at all positions $\mathbf{x}$, where $\psi(\mathbf{x},t) = 1$ if an event occurs in the time interval between t and $t + \Delta t$, and $\psi(\mathbf{x},t) = 0$ otherwise.

Observations and numerical simulations suggest that space-time patterns of seismic activity directly reflect the existence of space-time correlations in the underlying stress and strain fields (PRESS and ALLEN, 1995). A spatially coherent, uniformly high level of stress on a fault is a necessary precondition for the occurrence of a large earthquake. Recently, several groups have found that spatial coherence in

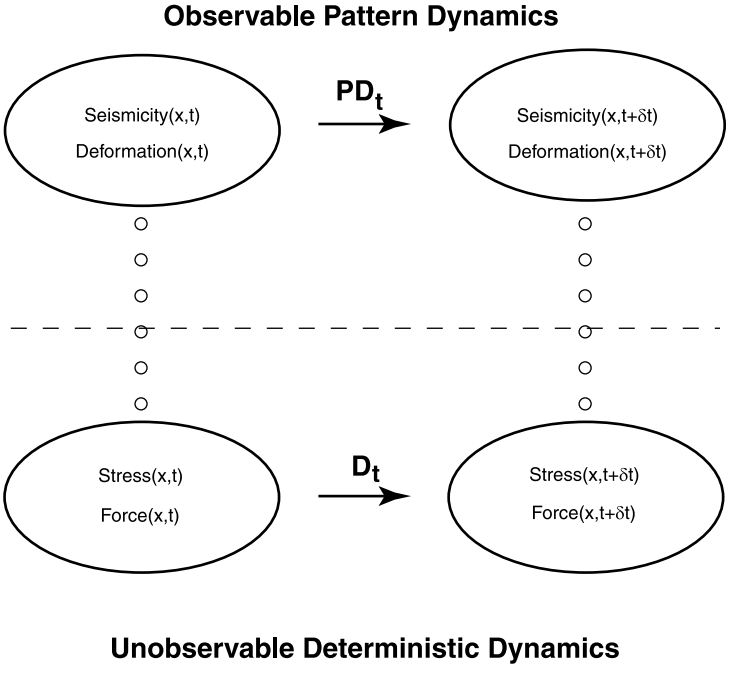

Figure 2
Schematic diagram of threshold systems.

the stress field is reflected in a similar coherence in the seismic activity (BUFE and VARNES, 1993; BOWMAN *et al.*, 1998; RUNDLE *et al.*, 2000a).

RUNDLE *et al.* (2000) extended the standard KLE methods to include the construction of pattern states that can be used to forecast events in time, in much the same manner as EOF analysis is used to predict El Niño events in meteorology (PREISENDORFER, 1988; PENLAND; GARCIA and PENLAND, 1991). This procedure involves constructing a correlation operator, $C(x_i,x_j)$, for the sites that contain the spatial relationship of slip events over time. $C(x_i,x_j)$ is decomposed into the orthonormal spatial eigenmodes for the nonlinear threshold system, e_j, and their associated time series, $a_j(t)$.

The Karhunen-Loeve expansion is obtained from the p time series that record the deformation history at particular locations in space. Each time series, $y(x_s,t_i) = y_i^s$, $s = 1,\ldots, p$, consists of n time steps, $i = 1,\ldots, n$. The goal is to construct a time series for each of numerous locations that records, for a given short period of time, whether an earthquake occurred at that location (value = 1) or did not occur (value = 0). If, for example, the time interval was decimated into units of days, the result would be a time series of 365 time steps for every year of data, with either a zero or a one at each time step. These time series are incorporated into a matrix, **T**, consisting of time series of the same measurement for p different locations, i.e.,

$$\mathbf{T} = [\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_p] = \begin{bmatrix} y_1^1 & y_1^2 & \cdots & y_1^p \\ y_2^1 & y_2^2 & \cdots & y_2^p \\ \vdots & \vdots & \ddots & \vdots \\ y_n^1 & y_n^2 & \cdots & y_n^p \end{bmatrix}.$$

T is therefore an $n \times p$, matrix of real values (FUKUNAGA, 1970). The covariance matrix, $S(x_i,x_j)$, for these events is formed by multiplying **T** by $\mathbf{T}^{\mathbf{T}}$, where S is a $p \times p$ real, symmetric matrix. The covariance matrix, $S(x_i,x_j)$, is converted to a correlation operator, $C(x_i,x_j)$, by dividing each element of $S(x_i,x_j)$, by the variance of each time series, $y(x_i,t)$ and $y(x_j,t)$, as follows:

$$\sigma_p = \sqrt{\frac{1}{n}\sum_{k=1}^{n} (y_k^p)^2},$$

and

$$\mathbf{C} = \begin{bmatrix} \frac{s_{11}}{\sigma_1\sigma_1} & \frac{s_{12}}{\sigma_1\sigma_2} & \cdots & \frac{s_{1p}}{\sigma_1\sigma_p} \\ \frac{s_{21}}{\sigma_2\sigma_1} & \frac{s_{22}}{\sigma_2\sigma_2} & \cdots & \frac{s_{2p}}{\sigma_2\sigma_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{s_{p1}}{\sigma_p\sigma_1} & \frac{s_{p2}}{\sigma_p\sigma_2} & \cdots & \frac{s_{pp}}{\sigma_p\sigma_p} \end{bmatrix}.$$

This equal-time correlation operator, $C(x_i,x_j)$, is decomposed into its eigenvalues and eigenvectors in two parts. The first employs the trireduction technique to reduce

the matrix **C** to a symmetric tridiagonal matrix, using a Householder reduction. The second part employs a ql algorithm to find the eigenvalues, λ_j^2, and eigenvectors, e_j, of the tridiagonal matrix (PRESS *et al.*, 1992).

These eigenstates thus represent the orthonormal basis vectors arranged in order of decreasing correlation, and reflect the relative importance of the various modes over the time interval of interest. Dividing each corresponding eigenvalue, λ_j^2, by the sum of the eigenvalues, yields that percent of the correlation accounted for by that particular mode. The associated orthonormal time series can be reconstructed by projecting the initial data set onto these basis vectors (PREISENDORFER, 1988; HOLMES *et al.*, 1996). The time dependent expansion coefficients, $a_j(t)$, which represent temporal eigenvectors, are reconstructed by multiplying the original data matrix by the eigenvectors, i.e.,

$$a_j(t_i) = \vec{e}^T \cdot T = \sum_{s=1}^{p} e_j y_i^s,$$

where $j,s = 1, \ldots, p$ and $i = 1, \ldots, n$. This eigenstate decomposition technique produces the orthonormal spatial eigenmodes for this nonlinear threshold system, e_j, and the associated principal component time series, $a_j(t)$. These principal component time series represent the signal associated with each particular eigenmode over time. For purposes of clarity, the spatial eigenvectors are designated "KLE modes" and the associated time series "Principal Component (PC)" vectors.

Phase Dynamical Probability Change

Our method is based on the idea that the time evolution of seismicity can be described by pure phase dynamics (MORI and KURAMOTO, 1998; RUNDLE *et al.*, 2000a, 2000b). We therefore construct a real-valued seismic phase function $\hat{\mathbf{S}}(x_i, t_0, t)$. For our analysis, the phase function $\hat{\mathbf{S}}(x_i, t_0, t)$ characterizes the seismic activity in southern California between 32° and 37° latitude, and −115° to −122° longitude. It should be noted that this southern California catalog *has not been declustered*, because the space-time clustering carries the information concerning the space-time correlations that we seek to identify. Since seismicity in active regions is a noisy function (KANAMORI, 1981), we work with temporal averages of seismic activity. The geographic area is partitioned into N = 3621 square regions approximately 11 km on a side, centered on a point x_i. Within each box, a time series is defined using the Caltech seismic catalog obtained from the online SCEC database. For southern California, the instrumental data begins in 1932 and extends to the present.

We define the activity rate $\psi_{\text{obs}}(x_i, t)$ as the number of earthquakes per unit time, of any size, within the box centered at x_i at time t. The geographic region that $\hat{\mathbf{S}}(x_i, t_0, t)$ represents is taken large enough so that seismic activity can be considered an incoherent superposition of phase functions. To construct the phase function $\hat{\mathbf{S}}(x_i, t_0, t)$, we define the time-averaged seismicity function $\mathbf{S}(x_i, t_0, t)$ over the interval $(t - t_0)$:

$$\mathbf{S}(x_i, t_0, t) = \frac{1}{(t - t_0)} \int_{t_0}^{t} \psi_{\text{obs}}(x_i, t) dt.$$

Since there are N numbers $\mathbf{S}(x_i, t_0, t)$, and if we assume t_0 to be a fixed time, we can then consider $\mathbf{S}(x_i, t_0, t)$ to be the ith component of a general, time-dependent vector evolving in an N-dimensional space. In previous work, we showed that this N-dimensional *correlation space* is spanned by the eigenvectors of an $N \times N$ correlation matrix (RUNDLE *et al.*, 2000a, 2000b).

Denoting spatial averages over the N boxes by $\langle\rangle$, the phase function $\hat{\mathbf{S}}(x_i, t_0, t)$ is then defined to be the mean-zero, unit-norm function obtained from $\mathbf{S}(x_i, t_0, t)$:

$$\hat{\mathbf{S}}(x_i, t_0, t) = \frac{\mathbf{S}(x_i, t_0, t) - \langle \mathbf{S}(x_i, t_0, t) \rangle}{\|\mathbf{S}(x_i, t_0, t)\|},$$

where $\|\mathbf{S}(x_i, t_0, t)\|$ is the norm over all spatial boxes. Note that

$$\|\mathbf{S}(x_i, t_0, t)\| = \{\langle (\mathbf{S}(x_i, t_0, t) - \langle \mathbf{S}(x_i, t_0, t) \rangle)^2 \rangle\}^{1/2}.$$

Based upon the assumption of pure phase dynamics (RUNDLE *et al.*, 2000a, 2000b), the important changes in seismicity will be given by $\Delta\hat{\mathbf{S}}(x_i, t_1, t_2) = \hat{\mathbf{S}}(x_i, t_0, t_2) - \hat{\mathbf{S}}(x_i, t_0, t_1)$, which is a pure rotation of the N-dimensional unit vector $\hat{\mathbf{S}}(x_i, t_0, t)$ in time.

In order to both remove the last free parameter in the system, the choice of base year, as well as to reduce the random noise component, we sum over all base years, t_0, from 1932 through t_2, such that

$$\Delta\hat{\mathbf{S}}(x_i, t_1, t_2) = \frac{\int_{1932}^{t_2} [\Delta\hat{\mathbf{S}}(x_i, t, t_2) - \Delta\hat{\mathbf{S}}(x_i, t, t_1)]\, dt}{\int_{1932}^{t_2} dt}.$$

In phase dynamical systems, probabilities are related to the square of the associated vector phase function (MORI and KURAMOTO, 1998; RUNDLE *et al.*, 2000b). Thus we find

$$\Delta\mathbf{P}(x_i, t_1, t_2) = \{\Delta\hat{\mathbf{S}}(x_i, t_1, t_2)\}^2 - \mu_P.$$

In other words, μ_P is the spatial mean of $\{\Delta\hat{\mathbf{S}}(x_i, t_1, t_2)\}^2$, and $\Delta\mathbf{P}(x_i, t_1, t_2)$ is the square of the value of $\Delta\hat{\mathbf{S}}(x_i, t_1, t_2)$ at each x_i minus that spatial mean. Note that the integral of $\Delta\mathbf{P}(x_i, t_1, t_2)$ over all N sites vanishes for all time intervals $[t_1,t_2]$, as it should, to conserve probability. Finally, note again that there are *no remaining free parameters in this method*. For any given catalog, up to the present, it is possible to compute the relative change in probability of an event over any given time period.

Results

The decomposition of nonlinear systems into their orthonormal eigenfunctions has been used successfully in the atmospheric sciences for many years (PREISENDORFER, 1988; PENLAND and MAGORIAN, 1993). The Karhunen-Loeve approach, the theoretical basis for EOF techniques, represents these space-time patterns as a set of eigenvectors, or normal modes, of an equal-time correlation function, their associated time series, and N total eigenfrequencies, where N is the total number of locations. The eigenvectors provide information pertaining to the spatial correlations of the patterns; the time series characterize each eigenvectors temporal pattern; the eigenfrequencies provide information regarding how often they occur in the data. In a number of applications, a complex, linear correlation operator is constructed in order to extrapolate future system behavior such as the El Niño southern oscillation (PENLAND and MAGORIAN, 1993). Here we apply this decomposition method to historical seismicity data in southern California in order to identify basis patterns for all possible space-time seismicity configurations. These basis states are a complete, orthonormal set of eigenvectors and associated eigenvalues that are obtained from the diagonalization of the correlation operators computed for this regional historic seismicity data.

Time Period 1932 through August, 1999

In our first analysis example, the time period starting in 1932 and continuing through the end August of 1999 was analyzed using the entire data set, including the entire areal extent and events of all quality. Figure 3a is a plot of the first 25 normalized eigenvalues, while Figure 3b is the first 1000 normalized eigenvalues, plotted on a log-log scale. Figure 4 shows the first two modes for southern California, for this data set. The absolute maximum value in each plot is normalized to one, where squares are positive and diamonds are negative, i.e., squares and diamonds are anticorrelated. The correct interpretation is that while a square location is "on," a diamond location is "off," and *vice versa*. The first mode is effectively a background hazard map, where small events are mutually correlated throughout southern California, while the second mode is the Landers event. The accompanying PC time series are shown in Figures 4a and 4c. The influence of spatial and temporal variations due to the density and completeness of network coverage is visible in the PC time series. For example, note the distinctive wave associated with the Landers sequence and its large numbers of aftershocks, punctuated by the occurrence of the Northridge event.

Figure 4d displays the second KLE mode. Here the region surrounding the 1992 Landers event is "on" (squares) whereas the rest of the southern San Andreas fault system is "off" (diamonds). The Coalinga earthquake is visible in this mode, and an apparent correlation between Landers and a set of events in eastern Nevada is revealed. As can be seen in Figures 5 and 6, a number of the KLE eigenpatterns are

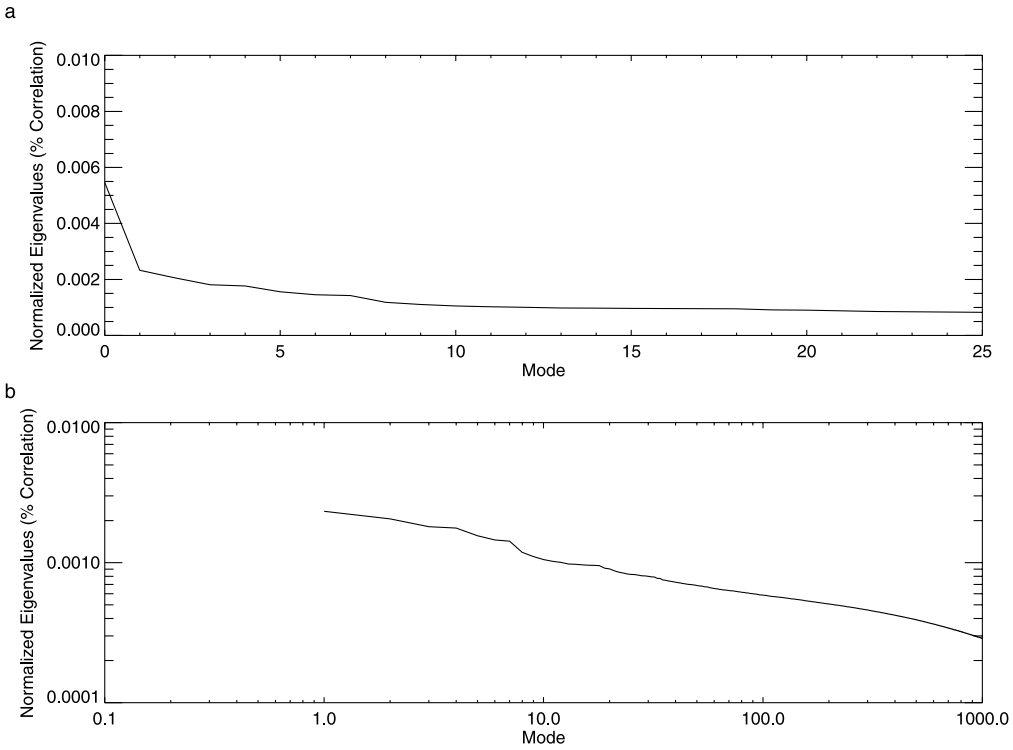

Figure 3
a) The first 25 normalized eigenvalues, for the time period 1932–1999. (b) The first 1000 normalized eigenvalues, plotted on a log-log scale.

lower order harmonics of the second mode. Interestingly, Figure 6b, KLE7, illustrates the correlations between the North Palm Springs event and other major southern California earthquakes, although with minimal correlation with the Landers sequence.

Typically, the higher order KLE modes display signal on shorter spatial and temporal scales than the initial, lower modes. This is illustrated in Figure 6. Figure 6c, KLE8, is a smaller scale harmonic of Figure 6a, KLE6, while KLE9, Figure 6d, shows the Landers sequence, essentially isolated, with the Joshua Tree earthquake anticorrelated with the Landers event to the north and Big Bear to the northwest.

Time Period: 1932 through December, 1991

One of the interesting questions which arises in studying the results above is how exactly the dominance of the Landers sequence, and its associated instrumentation, affects the eigenpatterns. This encouraged the removal of that event from the data set by cutting off the time series before its occurrence, at the end of 1991. For this time

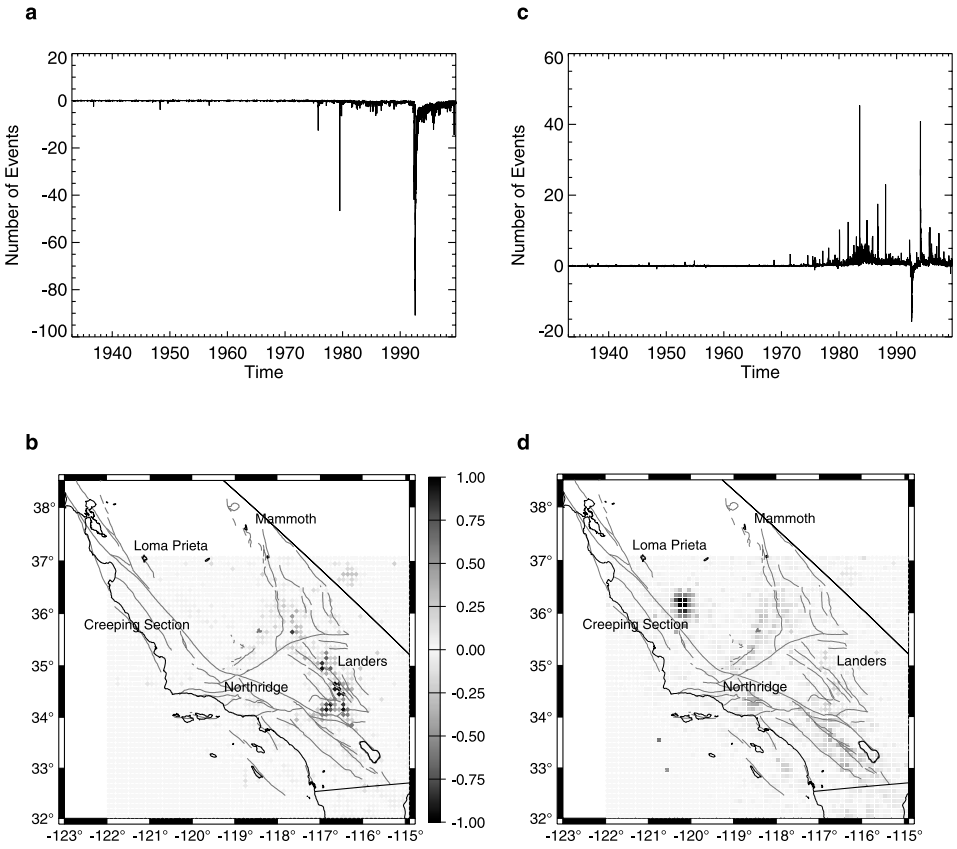

Figure 4
First two KLE modes for southern California seismicity, 1932–1999. Diamonds are negative, square boxes are positive. a) PC time series for first KLE mode; b) first KLE mode, normalized to maximum; c) PC time series for second KLE mode; and d) second KLE mode, also normalized to the maximum.

period, 1932 to 1991, the time series interval was again one day, so that the total number of time steps is approximately 21,535. Again, all locations from the entire database, and all quality events, were included in the decomposition, consequently the number of location time series is 3621. Figure 7a shows the first 25 normalized eigenvalues, while Figure 7b is the first 1000 normalized eigenvalues, plotted on a log-log scale. The eigenvalue plot is now smoother, without the large drop after the first mode that was a function of the large wave in the first PC time series generated by the Landers aftershocks (see Fig. 4b).

KLE modes one and two are shown in Figure 8, where KLE1 is the mode associated with background seismicity. As expected, there is no large signal for Landers visible in Figure 8b. In addition, many of what were the lower modes in the previous analysis have moved up in the eigenvalue ranking, replacing the large

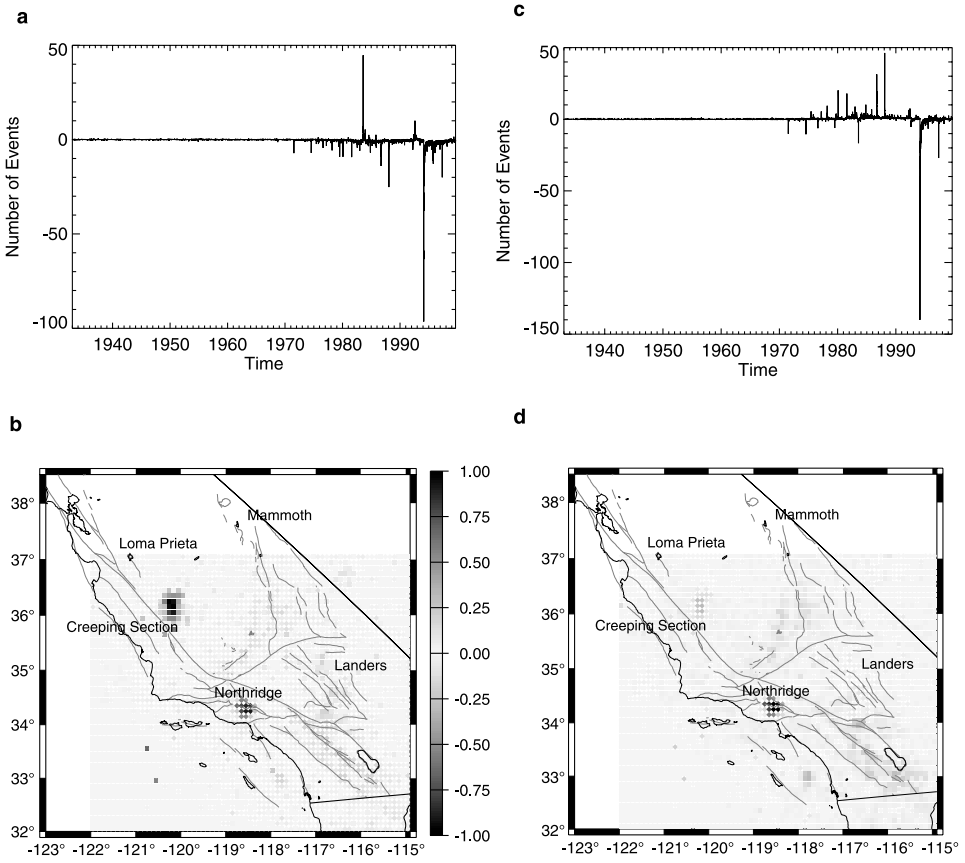

Figure 5
KLE modes three and four for southern California seismicity, 1932–1999. a) PC time series for third KLE mode; b) third KLE mode, normalized to maximum; c) PC time series for fourth KLE mode; and d) fourth KLE mode, also normalized to the maximum.

number of Landers harmonics. KLE2 (Fig. 8d) is the 1983 Coalinga earthquake, anticorrelated with the 1986 Oceanside and North Palm Spring events. The 1971 San Fernando event is a single diamond, correlated with the Coalinga earthquake.

Figure 9 again demonstrates both the addition of lower order harmonics and increasingly smaller spatial and temporal scales with increasing mode number. The fourth KLE mode shown in Figure 9a is the 1987 Superstition Hills and Elsinore Ranch events correlated with the Whittier Narrows earthquake and anticorrelated with the North Palm Springs and Oceanside events. Figure 9b, KLE mode five, also shows the North Palm Springs and Oceanside events correlated with each other. In KLE7, Point Mugu appears correlated with the 1971 San Fernando event. The Tejon Ranch earthquake of 1988, in addition to the 1979 Homestead Valley earthquake, south of Landers, is also correlated with these events, while Imperial Valley is

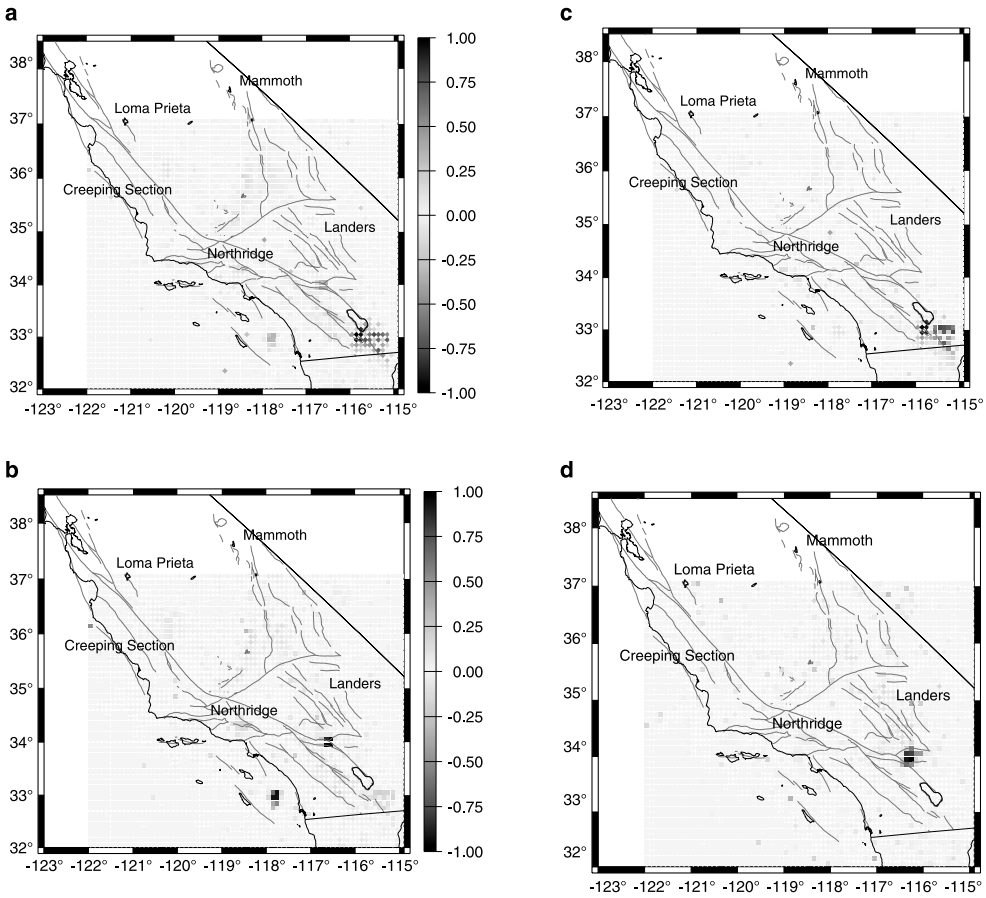

Figure 6
KLE modes six through nine for southern California seismicity, 1932–1999, each normalized to its maximum. a) Sixth KLE mode; b) seventh KLE mode; c) eighth KLE mode; and d) ninth KLE mode.

anticorrelated with these events. KLE8 again presents the 1979 Homestead Valley event, now anticorrelated with Point Mugu. Note the arcuate structure, which cuts across the faults at the location of the 1992 Landers earthquake. This is a feature of the local seismicity that has only been recognized in recent years with the occurrence of the 1992 Landers sequence and the 1999 Hector Mine earthquake, but which was clearly visible in this decomposition as early as 1991 (Fig. 9d). It should be noted again that *no seismicity data after December, 1992, are included in this analysis.*

Analysis of this same data set, 1932 through 1991, but with a magnitude cut of 3.0 applied to the data, yields the KLE modes shown in Figures 10 and 11. The application of a magnitude cut has allowed events from earlier in the data set to achieve a greater prominence. Interestingly, while the background mode has dropped out, the first mode is not Coalinga, the second mode in the entire data set (see Fig. 8b

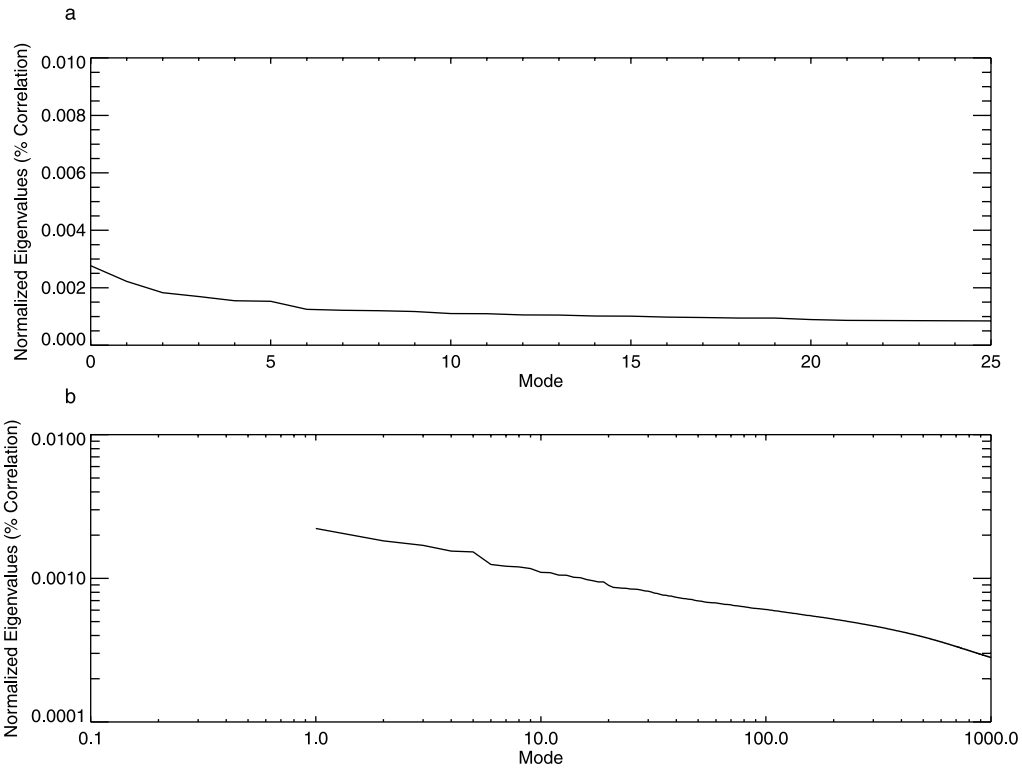

Figure 7
a) The first 25 normalized eigenvalues, for the time period 1932–1991. b) The first 1000 normalized eigenvalues, plotted on a log-log scale.

above), but the 1971 San Fernando event. The 1983 Coalinga earthquake is now the third KLE mode, behind the 1969 Avila Beach earthquake (see Fig. 10). The correlation between the 1979 Imperial Valley and 1987 Superstition Hills sequence is the fourth mode. The 1952 Kern County event has appeared as the seventh mode (Fig. 11a), while the eighth mode is also new to the decomposition – the 1968 Borrego Mountain earthquake (Fig. 11b). Figure 11 also shows KLE modes ten and eleven, now harmonics of the earlier modes, on smaller spatial and temporal scales. One interesting feature of KLE mode 11 is the anticorrelation between the 1983 Coalinga event and the historic location of the Parkfield earthquake to the west, which has not taken place in recent years as expected. This mode supports speculation that the Coalinga event delayed its occurrence.

Precursory Modes

The presence of both large- and small-scale correlations in the data, evident in the KLE decompositions shown above, prompted a study of the change in these

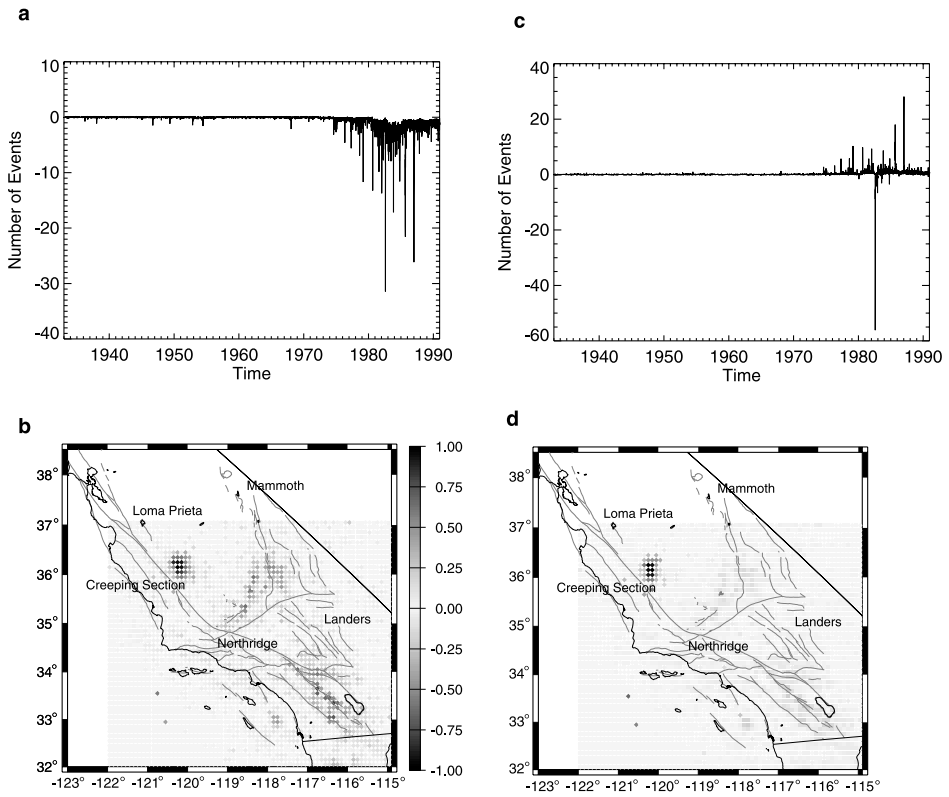

Figure 8
First two KLE modes for southern California seismicity, 1932–1991. a) PC time series for first KLE mode; b) first KLE mode, normalized to maximum; c) PC time series for second KLE mode; and d) second KLE mode, also normalized to the maximum.

modes for each year, in an attempt to identify modes which consistently appear over some identifiable time period prior to an event. While a complex rate correlation operator, $K(x_i,x_j)$, can be used to compute the probability of future events on a fault patch model producing events over time periods of thousands of years (RUNDLE *et al.*, 2000a), its application to historic seismicity data is limited. Neither the long-time periods nor the abundance of moderate to large events produced by numerical simultions are available in the actual data, nor is the same accuracy in time and space possible. Consequently, the following method was developed.

If the seismicity in a given year, **S** is known, and the eigenmodes, or eigenvectors e_i, are calculated using all seismicity data ($i = 1, \ldots N, N = 3621$ sites), then the eigenvectors are a complete, orthonormal set of basis functions, and any seismicity over that space can be decomposed into those eigenvectors.

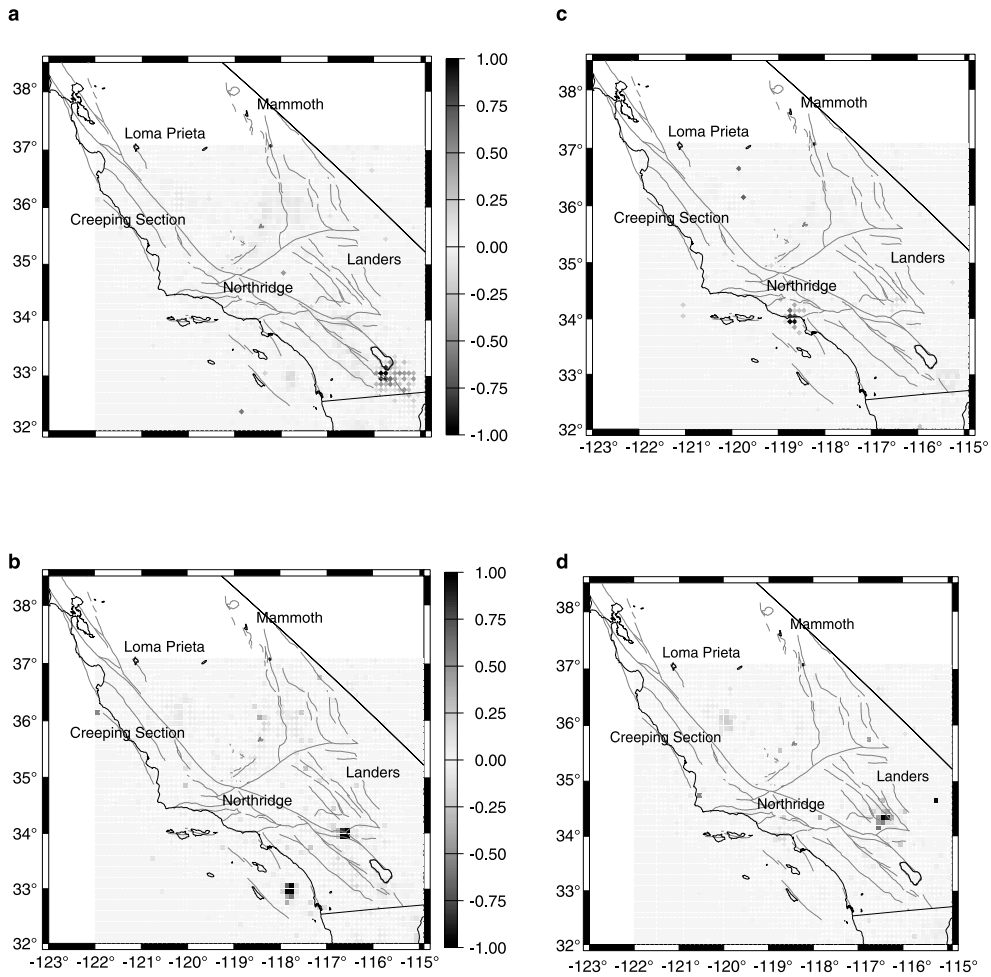

Figure 9
a) Fourth KLE mode; b) fifth KLE mode; c) seventh KLE mode; and d) eighth KLE mode for southern California seismicity, 1932–1991, each normalized to its maximum.

$$\mathbf{S} = \sum_{i=1}^{N} \alpha_i e_{ji},$$

where α_i are the eigenvalues for that particular year. The eigenvalues, α_i, are then computed from

$$\alpha_i = \sum_{i=1}^{N} e_{ji} S_i.$$

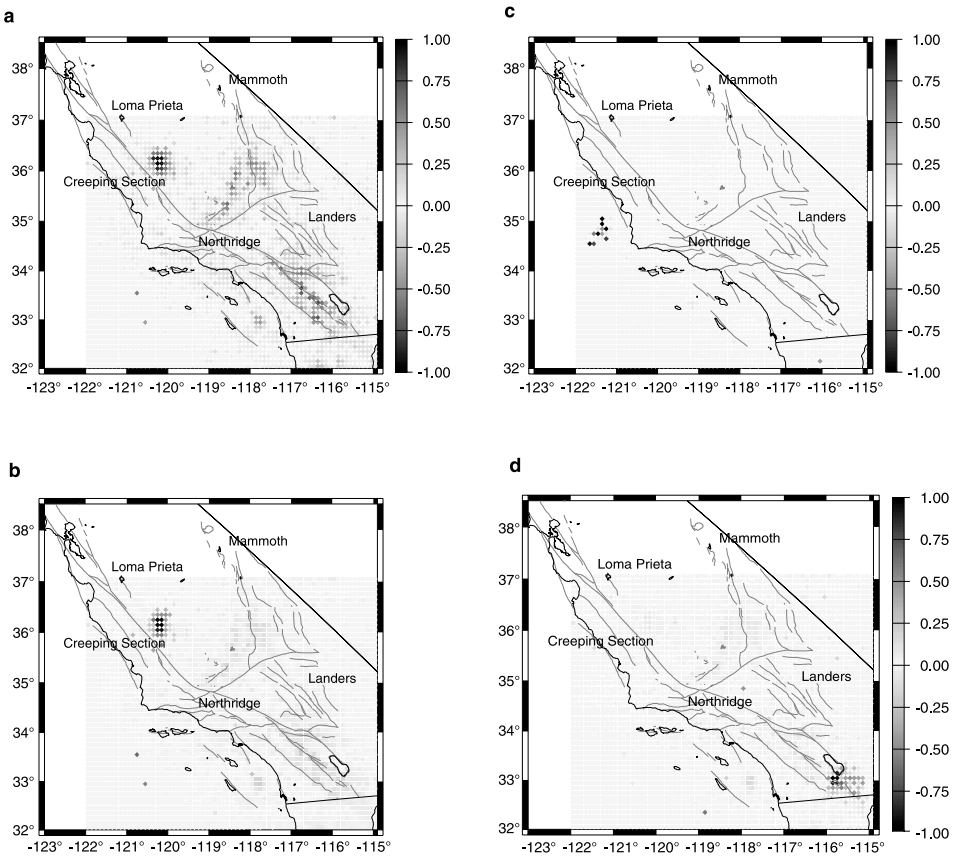

Figure 10
a) First KLE mode; b) second KLE mode; c) third KLE mode; and d) fourth KLE mode for southern California seismicity, magnitudes ≥ 3.0, 1932-1991, each normalized to its maximum.

Computing the α_i for any given year, given the KLE decompositions above, is a relatively simple process. The data set used was that described above, for the time period 1932 through 1991, with a magnitude cutoff of 3.0. The resulting $(\alpha_i)^2$, i.e., the power spectrum of the eigenmodes, for each year prior to the 1992 Landers sequence, are plotted in Figure 12. Note again, that no data after December 1991 are included in this analysis. Events that occur in the data set, for example the 1979 Imperial Valley event or the 1983 Coalinga earthquake, have signal in the corresponding eigenmodes, and would be expected to produce signal prior to those events. In addition, several of the higher order eigenmodes, on smaller spatial and temporal scales, appear to change substantially over the several years prior to 1992. Interestingly, a number of the modes that increase are those that include signal for the 1992 Landers earthquake sequence. For example, in Figure 13 is shown KLE modes 46, 49, 58, and 62, which, while noisy, contain significant correlations for the Joshua Tree, Landers and Big Bear events.

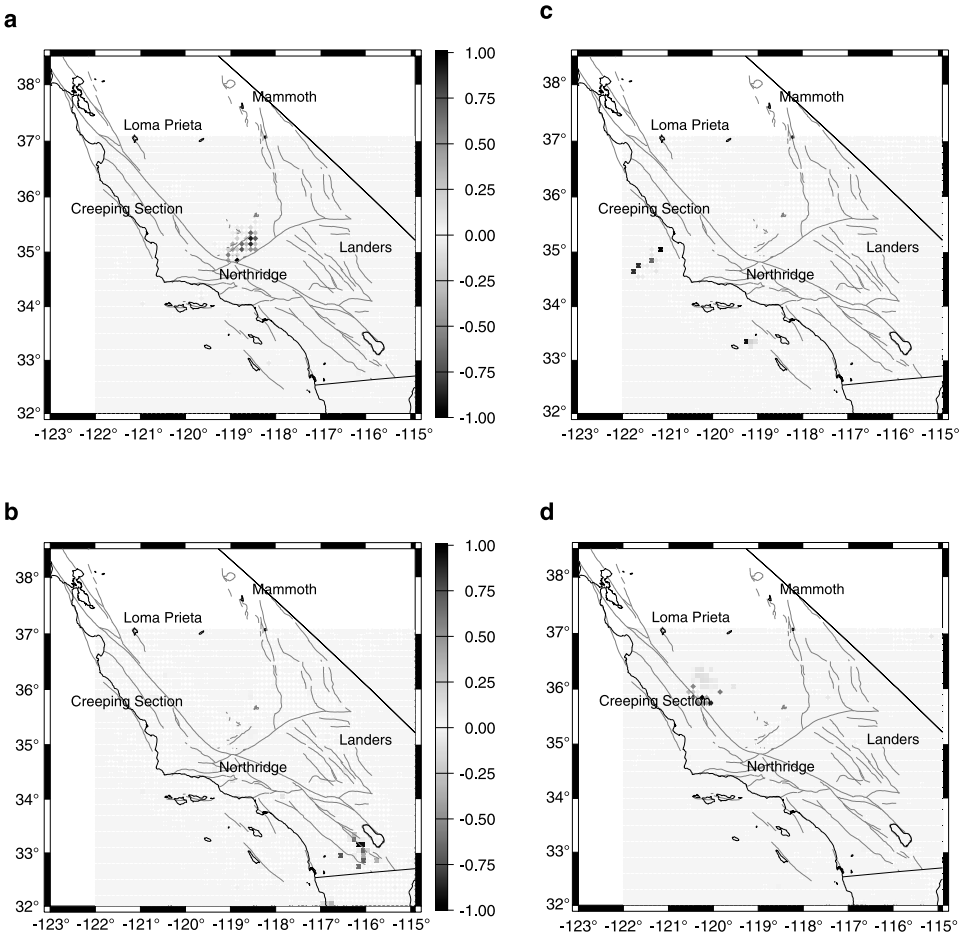

Figure 11
a) Seventh KLE mode; b) eighth KLE mode; c) tenth KLE mode; and d) eleventh KLE mode for southern California seismicity, magnitudes ≥ 3.0, 1932–1991, each normalized to its maximum.

Figure 14 depicts the $(\alpha_i)^2$, plotted for the seismicity over each year, summed sequentially from 1987 through 1991. Systematic changes now appear in the $(\alpha_i)^2$. These show the growth in the smaller scale Landers modes over time, as the fault system becomes increasingly correlated prior to the 1992 event. Again, it is important to remember that no data after December 31, 1991, is included in the analysis of either these eigenmodes or the associated $(\alpha_i)^2$.

Note that a multidimensional vector can be formed from the seismic activity by assuming that each of the locations on the grid of southern California is one of the dimensions, and that as the amount of seismicity at each location changes in time, the vector undergoes a rotation about some mean. If the seismicity

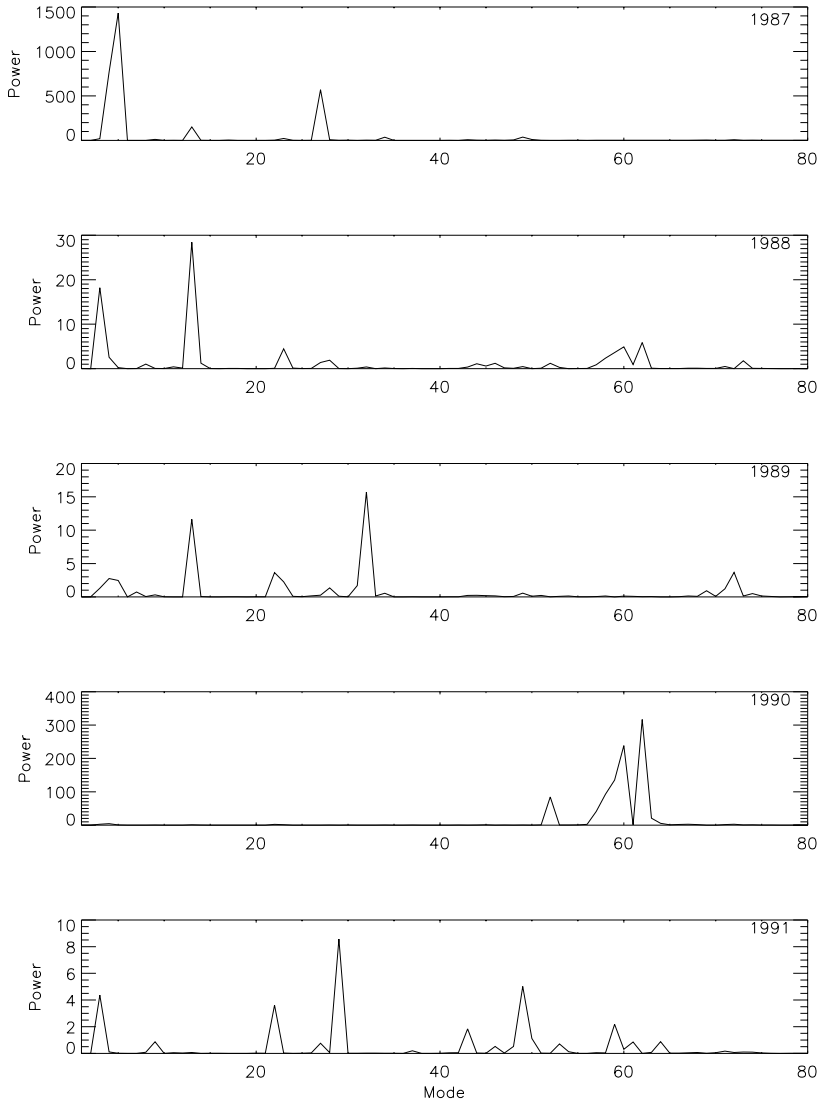

Figure 12
Eigenvalue power, $(\alpha_i)^2$, for the individual years 1987–1991, decomposed using eigenmodes derived using southern California seismicity, magnitudes ≥ 3.0, 1932 through 1991.

patterns, as described by the correlations in the $(\alpha_i)^2$, are undergoing a systematic change due to the growth of precursory modes, such as seen in Figures 12 and 14, then the vector is no longer experiencing a random walk about some mean. It undergoes a persistent rotation away from that mean that is quantifiable and can be converted into a probability of current and future events.

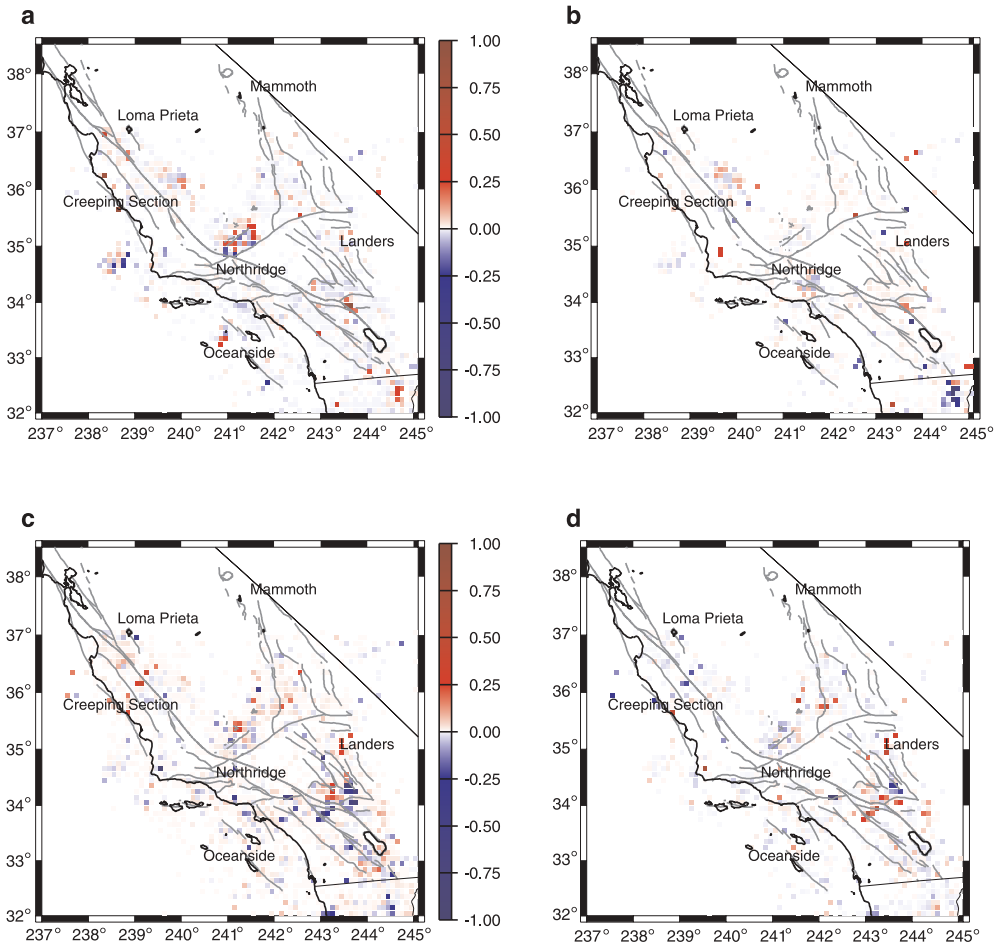

Figure 13
a) KLE mode 46; b) KLE mode 49; c) KLE mode 58; and d) KLE mode 62 for southern California seismicity, magnitudes ≥ 3.0, 1932–1991. The color scale is linear, blue to white to red, where blue is negative.

Probability of an Earthquake

Our simulations have suggested that the correlations in the seismicity represented by the KLE modes above can be described by phase dynamics. Phase dynamics is a method used in various branches of physics to describe the behavior of important parameters of the physical system (FUKUNAGA, 1970; MORI and KURAMOTO, 1998). Variables in many dynamical systems can be characterized by using this phase dynamical technique, represented as a phase function that involves both amplitude and phase angle. Changes in the amplitude of the phase function are unimportant, or not relevant. Examples of pure phase dynamical systems in the classical world

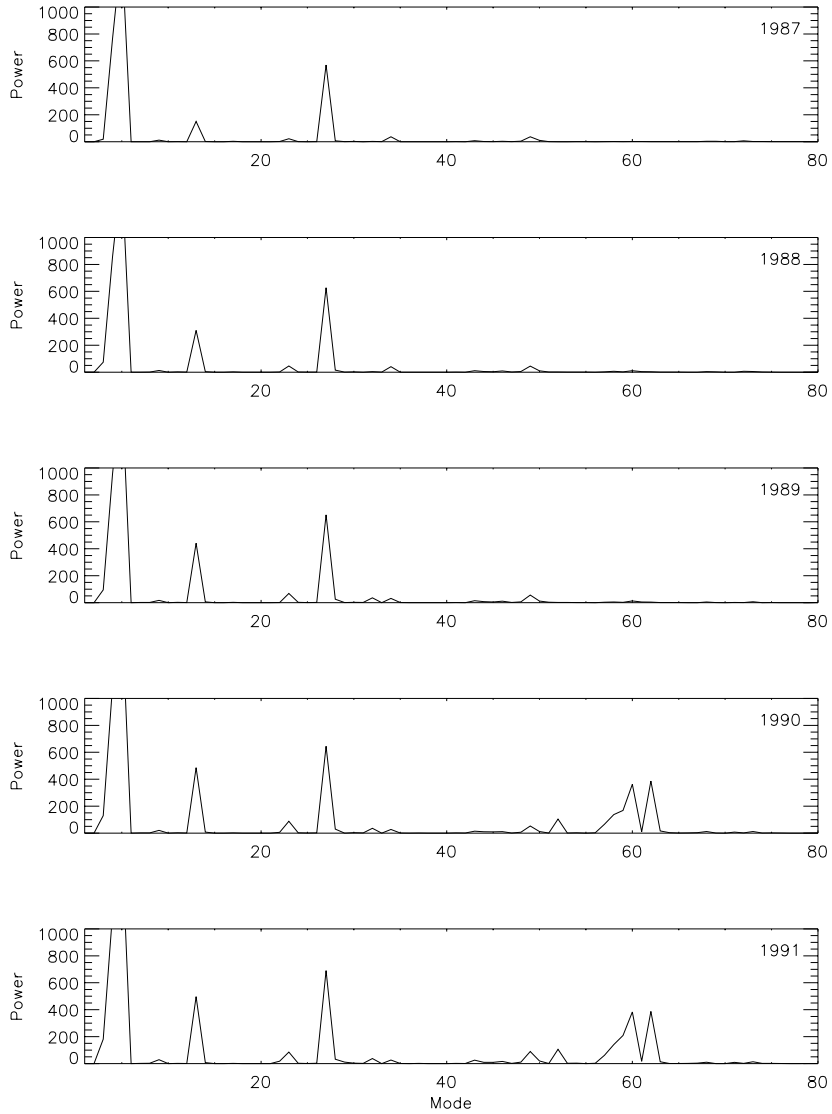

Figure 14
Eigenvalue power, $(\alpha_i)^2$, for the individual years *summed* from 1987 through 1991, decomposed using eigenmodes derived using southern California seismicity, magnitudes ≥ 3.0, 1932–1991.

include weak turbulence in fluids and reaction-diffusion systems. Another nonclassical example is a quantum system in which the wave function is the phase function.

In the case of earthquake fault systems, the important changes in seismicity are associated primarily with rotations of the vector phase function in a high-dimensional correlation space. Using the assumption that seismicity is an example

of pure phase dynamics, it follows that $\Delta\mathbf{P}$ can be calculated from the square of the phase function for the associated pattern state vector (RUNDLE *et al.*, 2000a). It should therefore be possible to compute the increase in probability of observing an anomalous correlation, $\Delta\mathbf{P}$, directly from the observed seismicity data and its rotation. To emphasize the connection to phase dynamics, we call the function $\Delta\mathbf{P}$ the Phase Dynamical Probability Change (PDPC) (TIAMPO *et al.*, 2000).

The technique described here is not a model, rather it is a new method for processing seismicity data to reveal underlying space-time structure. The purpose of the remainder of this paper is to exhaustively test the method, without regard to its formulation, and to apply it to the forecast of future earthquakes.

The seismicity data employed in our analysis are the same as those used above. Using only the subset of these data at locations $\mathbf{x}$ in southern California and covering the period from January 1, 1932 through December 31, 1991, we compute the PDPC function $\Delta\mathbf{P} = \Delta\mathbf{P}(x_i, 1932, 1991)$ for detecting anomalous spatial correlations at sites in southern California *prior* to January 1, 1992. Note that we use only events having magnitude $M \geq 3$, to ensure completeness of the catalog. The hypothesis to be tested is that these anomalous correlated regions are associated with large mainshocks that occurred *after* January 1, 1992. Figure 15a showed the relative seismic activity for this period, $\mathbf{S}(x_i, 1932, 1991)$, superimposed on a map of southern California. The intensity scale is logarithmic, where the scale value indicates the exponent, to the base ten, of the grayscale intensity value. It is clear that $\mathbf{S}(x_i, 1932, 1991)$ is an unremarkable function. For example, there is little evidence of any phenomena precursory to the M$\sim$7.3 Landers, California event that occurred on June 28, 1992.

Figure 15b exhibits the PDPC anomalies in southern California for the time period 1978 to 1991. Note that no data after December of 1991 were used in this analysis. The triangles denote events of $M > 5$ which go off during this time period, while the open circles are events which occur after 1991. Note the frequent occurrence of large earthquakes at the locations of increased relative probability.

Figure 15b shows plots of all $\Delta\mathbf{P} > 0$, using only existing seismicity data acquired *prior* to January 1, 1992; six months before the occurrence of the Landers earthquake sequence. The increase in $\Delta\mathbf{P}$ above the background level, as measured by μ_P, the spatial mean of $\Delta\mathbf{P}$, should be interpreted as the formulation of a spatially coherent region of either anomalous activation or quiescence, associated with an increased chance of a major earthquake. The scale is again logarithmic, scaled to the largest value of $\Delta\mathbf{P}$ on any of Figure 15b. $\Delta\mathbf{P}$ indicates only a relative, not an absolute change in probability. The inverted triangles represent events that occurred during the time period covered by the plot, indicating locations associated with large events that are present in the data used to construct $\Delta\mathbf{P}$. At a minimum, the method should identify areas of increased $\Delta\mathbf{P}$ associated with the triangles, and it is clearly successful. Of most interest are regions of increased $\Delta\mathbf{P}$ with no nearby triangles. According to our hypothesis, these locations may represent sites for future large earthquakes.

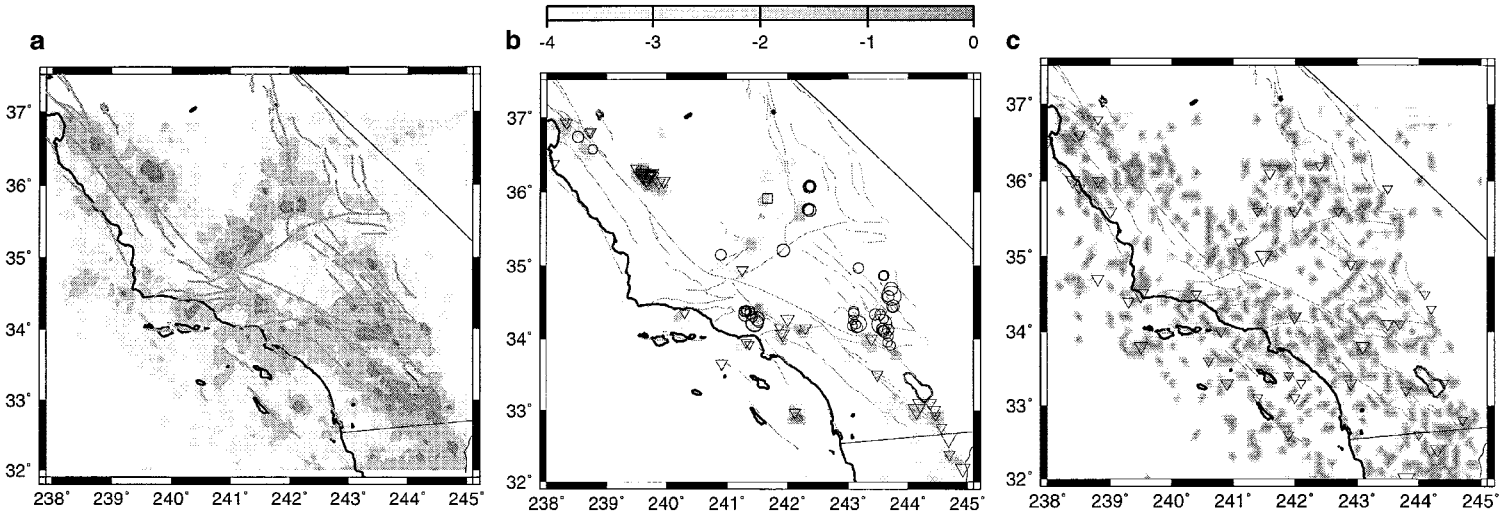

Figure 15

a) Seismicity function, **S**(1932, 1991), M ≥ 3.0, on a logarithmic scale, normalized to the maximum number of events, for southern California; b) maps of $\Delta\mathbf{P}$ for the time interval 1991–1978 for actual data; and c) randomized catalog. In constructing these maps, no data are used from time periods after December 31, 1991. The coding is logarithmic, in which the scale value indicates the exponent, to the base ten, of the gray intensity value. Inverted triangles represent events that occurred during the indicated time periods, with three sizes corresponding to magnitudes M of: $5 < M < 6$, $6 \leq M < 7$, $7 \leq M$. Open circles represent events that occur after January 1, 1992. Three increasing circle sizes again correspond to the same magnitude ranges as for the inverted triangles.

We then superimpose on this map the locations of mainshocks larger than 5.0 that occurred between January 1, 1992 and November 1, 1999, that is, the ~8 years *following* the time interval from which we computed the probabilities. Black circles denote these locations. One can observe an obvious correspondence between regions of increased probability and the location of the subsequent mainshocks, tending to support the results first observed in our simulations. Included are circles representing the 1992 Landers sequence, the 1994 Northridge earthquake, and the recent Hector Mine event.

Figure 15b shows that the recent large earthquakes occurring *after* January 1, 1992 are associated with areas of $\Delta \mathbf{P} > 0$ that formed *prior* to January 1, 1992. In particular, a bright area has developed close to the epicenter of the Joshua Tree event, latitude 33.95°, longitude 243.7°, which occurred in April of 1992. The anomalous area north of that location corresponds to the Landers earthquake sequence of June 1992. The epicenter of the recent October 16, 1999, M~7.0 Hector Mine also occurred on the northernmost end of that anomaly.

There is clearly variability, particularly for smaller events, depending on the choice of time interval. Larger events tend to be associated with larger anomalous regions that form earlier and persist longer after the mainshock. Since earthquake fault dynamics are now believed to be associated with critical phenomena (RUNDLE and KLEIN, 1995; GELLER *et al.*, 1997; KLEIN *et al.*, 1997) we hypothesize that there may be a scaling relation between the area A of the correlated region and the time interval τ prior to the mainshock at which the anomalous correlation begins to form, such that $\tau \propto A^{\eta}$, where η is a critical exponent near 1. Since the linear size of our location grid boxes is approximately 11 km, one should not expect events significantly smaller than M~6, whose characteristic linear source dimension is 10 km, to be well resolved by our procedure. Yet even the smaller circles associated with M~5 – 6 events often seem to occur in proximity to these anomalous areas. One such example is the China Lake event that occurred at approximately latitude 36.75°, longitude 242.25°. Here we also should note that a number of the anomalies correspond to events with magnitude less than 5, but that display persistent moderated seismicity, such as the Durrwood Meadows swarm, highlighted with a black square, at 35.5° latitude, 241.75° longitude. This set of events, which included an earthquake of magnitude 4.9, occurred from late 1983 through mid-1984.

These results suggest that the anomalies are correlated with total seismic moment release. The large, intense anomaly associated with the Coalinga event in 1983, for example, was followed by numerous aftershocks of magnitude 5.0 or greater. While there are earthquakes of $M > 5$ which occur without the formation of an area of increased probability ("false negatives"), this may be the result of a lower total seismic moment release. In the future, we will investigate whether this technique can be extended to estimate the potential magnitude and time of occurrence of forecast events from both the sizes of the candidate source regions and the temporal persistence and duration.

Likelihood Tests

To test the hypothesis that the formation of correlated regions of anomalous activity can be identified by our method, and are related to future large events, we carried out thousands of likelihood ratio tests (BEVINGTON and ROBINSON, 1992; GROSS and RUNDLE, 1998) on the method using values of $\Delta\mathbf{P}$ obtained from random seismicity catalogs that were used as null hypotheses. In the likelihood ratio test, a probability density function (PDF) is required. For our PDF we used a Gaussian distribution whose width is that of our original location grid, approximately 11 km, and whose peak value is given by $\Delta\mathbf{P} + \mu_P$, because in a likelihood ratio test the value of probability at all sites must be positive. The log likelihood is then calculated for the circles, which identify the locations of future events, as shown in Figure 15b, and provides a measure of how well the colored regions predict the actual events as quantified by the locations of the circles.

Each random catalog was constructed from the instrumental catalog by using the same total number of events, but assigning occurrence times from a uniform probability distribution over the years 1932–1991, and distributing them uniformly over the original locations. This procedure produces a Poisson distribution of events in space with an exponential distribution of interevent times. Randomizing the catalog in this way destroys whatever coherent space-time structure may exist in the data, *thus effectively declustering the catalog*. While random "Poisson clustering" may remain, there will be no "Omori clustering" left in these catalogs.

Likelihood ratio tests were then performed on thousands of such random catalogs, comparing calculations of $\Delta\mathbf{P}$ from the actual catalog with calculations of $\Delta\mathbf{P}$ from each random catalog. Success of the prediction is scored by how well the circles are forecasted by the colored areas. The better the prediction, the larger the likelihood value. Figure 16 shows the distribution of likelihood values for five hundred of the random catalogs, for the same time period as shown in Figure 15b. Superimposed on this plot as a dashed line is the likelihood ratio value for Figure 15b, computed from the actual catalog. One can see that the value of likelihood computed from the actual catalog is always substantially better than likelihoods computed from the random catalogs.

For purposes of illustration, we compute $\Delta\mathbf{P}$ for one random catalog whose likelihood value is close to the mean of the distribution in Figure 16, as shown in Figure 15c. One can see that there are many more anomalous areas in Figure 15c than in Figure 15b; these regions are far more broadly distributed in space, and contain many more dark areas. The specific likelihood value for Figure 15b equals −311.665, whereas that for Figure 15c equals −363.344. These values correspond to a likelihood ratio $e^{51.68} \approx 10^{22.44}$, indicating that the locations of increased probability obtained from the actual instrumental catalog are far more likely to be associated with the locations of the circles than those obtained from the

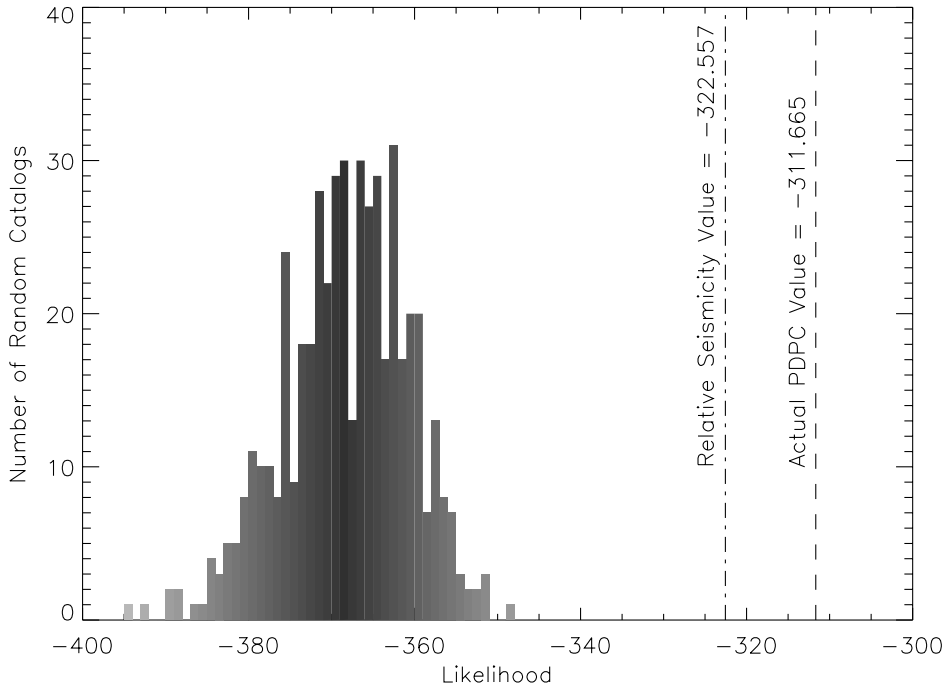

Figure 16
Histogram of likelihood values for one hundred random catalogs of southern California seismicity. In addition, the likelihood value for the actual catalog, as shown in Figure 15b, is superimposed on this plot as a dashed line, and the likelihood value for the relative southern California seismicity as shown in Figure 15a is plotted as the dash-dotted line.

random catalog. The physical reason for this large ratio is that the likelihood test invokes a penalty for predictions that are not sufficiently near the circles (''false positives''), and there are many more such locations in Figure 15c than in Figure 15b.

In addition, we computed the likelihood ratio using Figure 15a as the null hypothesis that may be taken to correspond to a map of relative hazard based upon recent instrumental seismic intensity. The likelihood value for this map is −322.557, equating to a likelihood ratio of $e^{11.22} \approx 10^{4.87}$. When compared to the computations for Figure 15b, this suggests that our method is considerably more successful at forecasting the locations of future major earthquakes than the current practice of relying on hazard maps to predict broad areas of increased seismic potential. From these statistical tests, we conclude that there are coherent, anomalous space-time correlation structures in the instrumental catalog that our method identifies, and that these correlations are effectively destroyed by the common practice of declustering seismicity catalogs.

Discussion

We emphasize that while our method may identify higher risk areas, there is no certainty at this time that every location of increased $\Delta\mathbf{P}$ will be located near the site of a future large earthquake. There are numerous examples in Figure 15b where a colored area appears without the occurrence of a major earthquake between 1992 and 1999. One example appears near 34° latitude, 244.25° longitude. Further attempts at optimization of the method must focus on better spatial location of events and minimizing the numbers of both false positives and false negatives.

As discussed below, we have seen that regions with $\Delta\mathbf{P} > 0$ may correspond either to anomalous seismic activity or anomalous quiescence. In some locations, a region that had been the site of a recent major earthquake may evidently indicate values of $\Delta\mathbf{P} > 0$ that are linked with a future large event that will occur at a neighboring, although somewhat disjointed, location. The positive value of $\Delta\mathbf{P}$ may appear only at the location of the past event, rather than at the neighboring location of the future event. In these cases, we have found that the neighboring aftershock zone of the past event is participating in the anomalous activity or quiescence that defines the future event. In addition, it is possible that anomalous activation or quiescence may appear in events having magnitudes less than the uniform cutoff we use, $M = 3$. While we have performed this analysis with other magnitude cuts and have found the results to be relatively insensitive between values of 2.5 and 4.0, this may not apply to other tectonic settings, based on the quality of the data available. The appropriateness of various magnitude cutoffs should therefore be tested in future work.

An example of these effects is the anomalous area at the location of the 1971 San Fernando earthquake, shown in Figure 15b, that is evidently associated with the coming 1994 Northridge event. This area is present *not* because it represents aftershock activity from the 1971 San Fernando event (anomalous seismic activity), rather it represents an area of relative anomalous quiescence. To further illustrate the point that the San Fernando region is anomalously quiescent prior to Northridge, we plot the change in the phase function $\Delta\hat{\mathbf{S}}(x_i, 1978, 1991)$, as defined in the method section above, spanning the years 1978–1991. In Figure 17, a positive value of $\Delta\hat{\mathbf{S}}$ represents anomalous seismic activity, and a negative value of $\Delta\hat{\mathbf{S}}$ represents anomalous quiescence. The intense anomaly near Northridge-San Fernando seen in the plot of $\Delta\mathbf{P}$ in Figure 15b is seen to be negative in Figure 17, indicating that the positive value of $\Delta\mathbf{P}$ arises from anomalous quiescence in the San Fernando aftershock zone. In addition, it will be seen below that the construction of $\Delta\mathbf{P}(x_i, 1978, 1991)$ actually subtracts the effect of any changes in seismicity prior to 1978, therefore San Fernando aftershock activity between 1971–1978 does not contribute to the calculation. Figure 17 also indicates that the 1983 Coalinga earthquake displays seismic activation, as would be expected since it occurred during this time period, while the Landers sequence is a mix of anomalous quiescence and activation. These results support the conclusion that the PDPC function does not simply identify areas associated with past

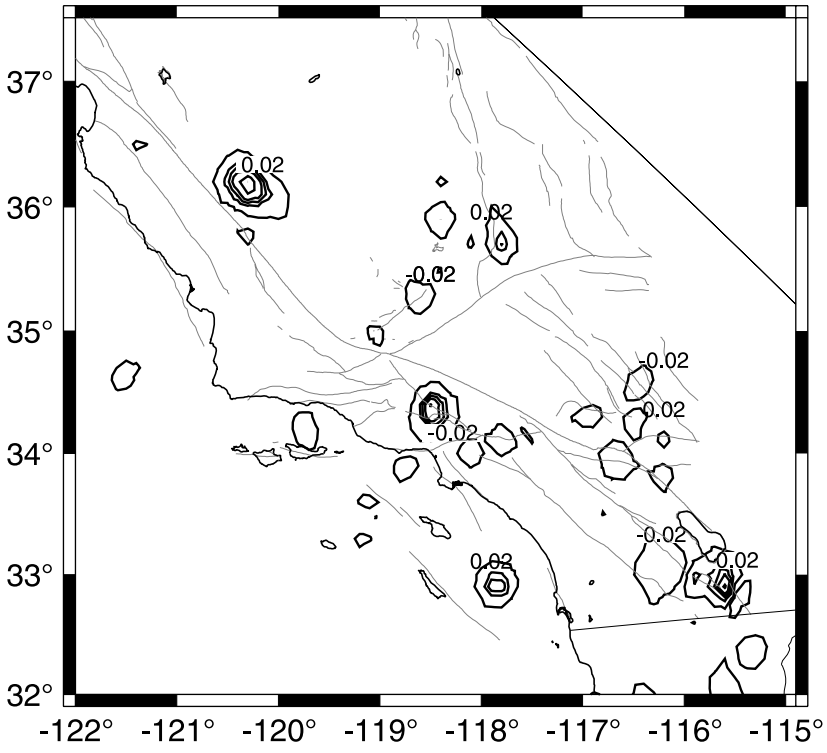

Figure 17
Contour plot of $\Delta\hat{\mathbf{S}}(x_i, t_1, t_2)$, normalized to the maximum absolute value, for the time period 1991–1978.

events and their aftershock sequences, rather it quantifies the underlying stress coherence and correlations associated with the regional seismicity.

Finally, we calculate the PDPC for all of California over three additional time periods. Figure 18a shows the increased probability for the time period 1968 through 1978. Again, the intensity scale is as shown in Figure 15, and the inverted triangles represent those events that occur during the time period, while the open circles are those earthquakes that occur throughout the next ten years. Despite the sparser networks in place in California at that time, which affects the spatial completeness of the catalog, the 1979 Imperial Valley earthquake, $M = 6.4$, occurs near a bright set of anomalies, which includes anomalies for the 1987 Superstition Hills and Elmore Ranch events. Also clearly visible is the upcoming 1983 Coalinga event at 36.25° latitude, 239.75° longitude, a previously unanticipated hidden thrust event.

Figure 18b shows the PDPC encompassing 1978 through 1988. Clearly visible is another relatively unexpected event at the time, the 1989 Loma Prieta earthquake. The epicenter of the mainshock occurs within 10 kms of the nearest anomaly, which extends down through the aftershock zone. Again, the Durrwood Meadows swarm,

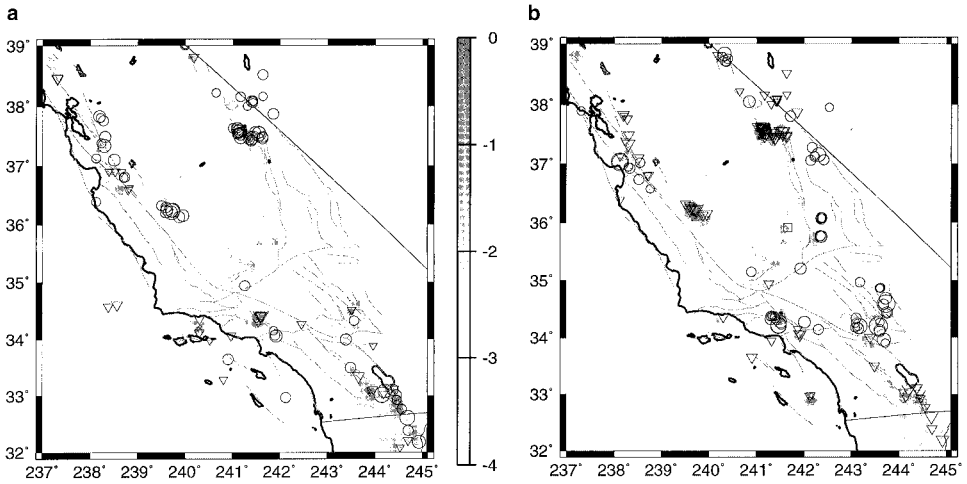

Figure 18
Maps of $\Delta\mathbf{P}$ for a) for the time period 1978–1968, and for b) the time period 1988–1978. In constructing these maps, no data is used for time periods after the final year. The color scale is the same as in Figure 15. Inverted triangles represent events that occurred during the indicated time periods, as before. Open circles represent events that occur during the ten years after the referenced time period. Three increasing sizes again correspond to the same magnitude ranges as Figure 15.

at 35.5° latitude, 241.75° longitude, including one earthquake of magnitude 4.9, is plotted as a square box.

Finally, Figure 19 shows a forecast for roughly the next ten years, using a white to yellow to red logarithmic intensity scale. The PDPC analysis is performed for the time period 1989 through 1999. The persistent anomaly located at the northeast end of the White Wolf Fault, 35.25° latitude, 241.5° longitude, corresponds to an area of recurrent seismicity throughout the 1980s and 1990s, which may represent the birth of a newly forming structure, a blind strike-slip fault connecting the Kern County and Walker Pass, California, earthquakes (BAWDEN *et al.*, 1999). Or, the orange zones at 33° latitude, 244° longitude, may identify a silent earthquake that occurred in the mid-1990s at that location (Paul Vincent, personal communication). Note that the northwesternmost anomaly, at 38.75° latitude, 237.2° longitude, identifies the location of a series of events, $M = 4$, followed by a number of smaller events, which occurred in January of 2000.

Figure 19 identifies certain areas that may be destined for activity in the next 5–10 years. For example, the yellow area just to the north of the 1988 Loma Prieta event, along the southern Hayward fault, is one potential area of concern. However, in agreement with deformation work by BURGMANN *et al.* (2000), the northern section of the Hayward fault manifests a lesser potential for an independent event. In southern California, the area just to the north of the 1992 Landers sequence, as well as the site of the 1983 Coalinga earthquake, display a

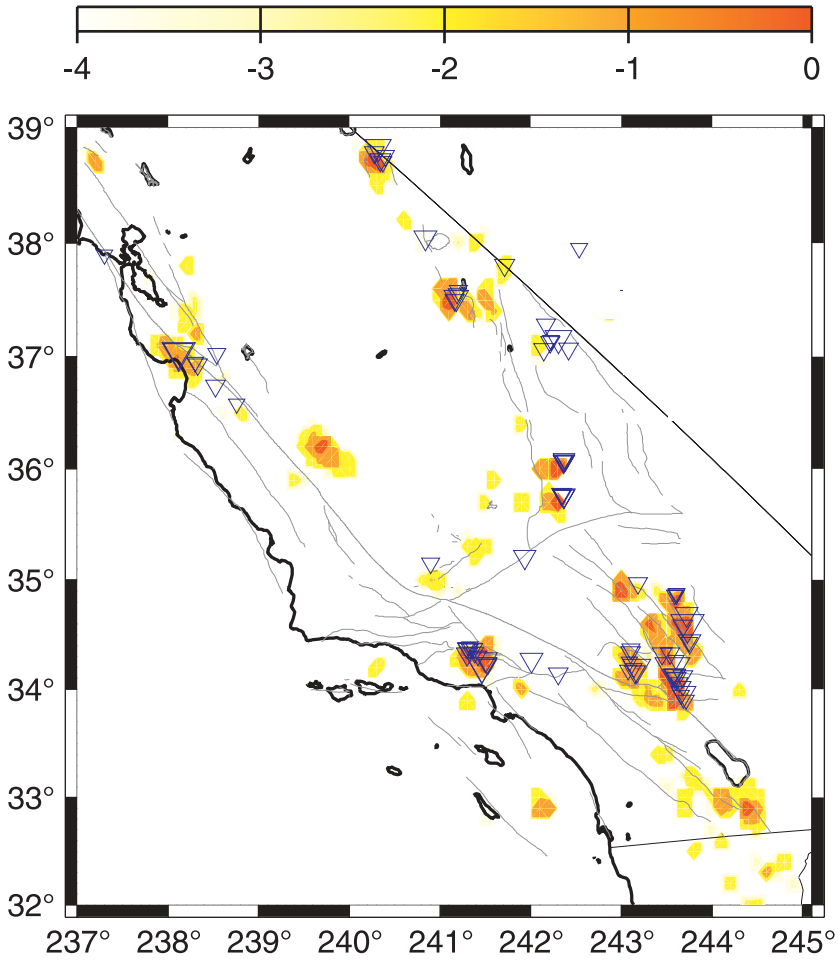

Figure 19

Map of $\Delta \mathbf{P}$ for the time period 1999–1989. In constructing this map, no data are used for the year 2000. The color scale is still logarithmic, as in Figure 15. Inverted triangles represent events that occurred during the indicated time periods, as before. Open circles represent events that occur during the ten years after the referenced time period. Three increasing sizes again correspond to the same magnitude ranges as Figure 15.

large PDPC anomaly. Finally, several areas of relative probability increase are visible along the southern San Andreas and San Jacinto faults. We must point out that, for reasons already detailed, absolute probabilities cannot be attached to these anomalies at this time. In addition, these anomalies represent areas of varying seismic activity – it is possible that some or all of these locations may represent areas of either recurring moderate events of magnitude less than 5.0, or locations of silent earthquakes occurring between 1989 and 1999. Finally, the occurrence of one

or more of these events could modify the underlying dynamics of the system, requiring the revision of subsequent forecasts.

Conclusions

This pattern dynamics approach that we have applied to historical seismicity data in southern California reveals a wealth of interesting spatial patterns. In particular, it provides a new methodology for classifying all of the possible seismicity patterns that can exist in terms of mutually orthogonal eigenstates. In fact, a number of the descriptive patterns cited earlier can be readily identified among the eigenstates depicted in Figures 4 through 12. For example, Figure 4d can be interpreted either as seismic activation within the Landers epicentral region, or else as quiescence near Landers coincident with seismic activity surrounding the region, i.e., a ''Mogi donut.''

Our results argue strongly for the development of realistic numerical simulations of fault systems such as those in southern California (RUNDLE, 1988; TIAMPO *et al.*, 1999; RUNDLE *et al.*, 2000a). Because the historic data set is incomplete worldwide, construction of such numerical simulations is necessary to more accurately define the most significant eigenpatterns, which can be applied to understanding the nature of the observed seismicity. For example, while a magnitude cutoff of 3.0 may be appropriate for southern California, it may not be applicable for other tectonic settings. Simulations can help to better define the sensitivity of the analysis to parameters such as this.

Moreover, computer simulations will be of critical importance for relating the observable pattern basis set for the seismicity data to the pattern basis set for stress, strains, and displacements through time. While seismicity is readily observable by standard methods, stress and strain within the earth are not. However, stress and strain are the primary dynamical variables, and are also the Markov variables in which the underlying nonlinear dynamics are most likely formulated. It will be most important to relate a readily observable, seismicity pattern basis set to the actual, unobservable dynamical pattern basis set, so that mode-shaping techniques can be applied to the underlying dynamics (FUKUNAGA, 1970; HOLMES *et al.*, 1996). In this manner it may be possible to characterize the spatially coarse-grained features of local and regional stress levels, coefficients of friction, failure and residual stress levels, and fault interactions. Finally, RUNDLE *et al.* (2000a), demonstrated, using simulations, that such methods can in principle forecast future events with accuracies considerably better than a standard Poisson process. With the development of a quantitative, readily reproducible technique for characterizing all possible seismicity patterns, these methods may allow us to test the hypothesis that large damaging earthquakes on a given subset of faults are preceded by one of a small set of characteristic precursory seismicity patterns. If this hypothesis is true, then it may well be possible to develop a quantitative method to forecast large, infrequent events

using the patterns of seismic activation and quiescence associated with smaller, more frequent events on local and regional fault systems.

Finally, our results suggest that systematic variations in seismicity prior to recent southern California earthquakes can be observed. Our method employs data from existing seismic monitoring networks as well as a theoretical understanding obtained from numerical computer simulations to identify anomalous activity or quiescence in seismicity. These space-time patterns in the seismic activity directly reflect the existence of correlated structure in the underlying stress and strain fields, a necessary precondition for the occurrence of large earthquakes. While at this time we offer this forecast for scientific evaluation only, depending on the nature of future seismicity in the region, as well as ongoing modifications and extensions of the theory and technique, this procedure may prove a useful tool for forecasting seismic activity.

Acknowledgements

This work has been supported by CIRES and NASA student fellowships to KFT and SM; by US DOE grant DE-FG03-95ER14499 to JBR. Research by WK was supported by US DOE/OBES grant DE-FG02-95ER14498 and W-7405-ENG-6 at LANL. WK would like to acknowledge the hospitality and support of CNLS at LANL.

References

Aubrey, D. G. and Emery, K. O. (1983), *Eigenanalysis of Recent United States Sea Levels*, Continental Shelf Res. *2*, 21–33.

Bak, P. and Tang, C. (1989), *Earthquakes as Self-organized Critical Phenomena*, J. Geophys. Res. *94*, 15,635–15,637.

Bawden, G. W., Michael, A. J., and Kellogg, L. H. (1999), *Birth of a Fault: Connecting the Kern County and Walker Pass, California, Earthquakes*, Geology *27*, 601–601.

Bakun, W. H., King, G. C. P., and Cockerham, R. S. (1986), *Seismic slip, aseismic slip, and the mechanics of repeating earthquakes on the Calaveras fault, California*. In *Earthquake Source Mechanics*, AGU Monograph, AGU, Washington, D.C., 195–208.

Bakun, W. H. and McEvilly, T. V. (1984), *Recurrence Models and Parkfield, California, Earthquakes*, J. Geophys. Res. *89*, 3051–3058.

Bowman, D. D., Ouillon, G., Sammis, C. G., Sornette, A., and Sornette, D. (1998), *An Observational Test of the Critical Earthquake Concept*, J. Geophys. Res. *103*, 24,359–24,372.

Brehm, D. J. and Braile, L. W. (1999), *Intermediate-term Earthquake Prediction Using the Modified Time-to-failure Method in Southern California*, BSSA *89*, 275–293.

Bufe, C. G. and Varnes, D. J. (1993), *Predictive Modeling of the Seismic Cycle of the Greater San Francisco Bay Region*, J. Geophys. Res. *98*, 9871–9883.

Burgmann, R. *et al.* (2000), *Earthquake Potential along the Northern Hayward Fault, California*, Science *289*, 1178–1182.

Deng, J. S. and Sykes, L. R. (1996), *Triggering of 1812 Santa Barbara Earthquake by a Great San Andreas Shock: Implications for Future Hazards in Southern California*, Geophys. Res. Lett. *23*, 1155–1158.

Dieterich, J. H. (1994), *A Constitutive Law for Rate of Earthquake Production and an Application to Earthquake Clustering*, J. Geophys. Res. *99*, 2601–2618.

DODGE, D. A., BEROZA, G. C., and ELLSWORTH, W. L. (1996), *Detailed Observations of California Foreshock Sequences: Implications for the Earthquake Initiation Process*, J. Geophys. Res. *101*, 22,371–22,392.

ELLSWORTH, W. I. and COLE, A. T. (1997), *A Test of the Characteristic Earthquake Hypothesis for the San Andreas Fault in Central California*, Seis. Res. Lett. *68*, 298.

ELLSWORTH, W. I., COLE, A. T., and DIETZ, L. (1998), *Repeating Earthquakes and the Long-term Evolution of Seismicity on the San Andreas Fault near Bear Valley*, *California*, Seis. Res. Lett. *69*, 144.

ENEVA, M. and BEN-ZION, Y. (1997), *Techniques and Parameters to Analyze Seismicity Patterns Associated with Large Earthquakes*, J. Geophys. Res. *102*, 17,785–17,795.

EVISON, F. F. (1977), *Fluctuations of Seismicity before Major Earthquakes*, Nature *266*, 957–972.

FERGUSON, C., KLEIN, W., and RUNDLE, J. B. (1999), *Spinodals, Scaling and Ergodicity in a Model of an Earthquake Fault with Long-range Stress Transfer*, Phys. Rev. E. *60*, 1359–1373.

FISHER, D. S., DAHMEN, K., RAMANATHAN, S., and BEN-ZION, Y. (1997), *Statistics of Earthquakes in Simple Models of Heterogeneous Faults*, Phys. Rev. Lett. *78*, 4885–4888.

FROHLICH, C. (1987), *Aftershocks and Temporal Clustering of Deep Earthquakes*, J. Geophys. Res. *92*, 13,944.

FUKUNAGA, K., *Introduction to Statistical Pattern Recognition* (Academic Press, N.Y., 1970).

GARCIA, A. and PENLAND, C. (1991), *Fluctuating Hydrodynamics and Principal Oscillation Pattern Analysis*, J. Stat. Phys. *64*, 1121–1132.

GELLER, R. J., JACKSON, D. D., KAGAN, Y. Y., and MULARGIA, F. (1997), *Enhanced: Earthquakes Cannot Be Predicted*, Science *278*, 488–490.

GOMBERG, J. (1996), *Stress/Strain Changes and Triggered Seismicity Following the M_W 7.3 Landers, California, Earthquake*, J. Geophys. Res. *101*, 751–764.

GRANT, L. and SIEH, K. (1994), *Paleoseismic Evidence of Clustured Earthquakes on the San Andreas Fault in the Carrizo Plain*, J. Geophys. Res. *99*, 6819–6842.

GRAY, C. M. (1997), *Synchronous oscillations in neuronal systems: Mechanisms and functions*. In *Pattern Formation in the Physical and Biological Sciences*, Lecture Notes V, SFI, Addison-Wesley, Reading, MA, 93–134.

GROSS, S. J. and KISSLINGER, C. (1994), *Tests of Models of Aftershock Rate Decay*, BSSA *84*, 1571–1579.

GROSS, S. and RUNDLE, J. B. (1998), *A Systematic Test of Time-to-failure Analysis*, Geophys. J. Int. *133*, 57–64.

HABERMAN, R. E. (1981), *Precursory seismicity patterns: Stalking the mature seismic gap*. In *Earthquake Prediction: An International Review*, AGU Monograph, AGU, Washington, D.C., 29–42.

HERTZ, J., KORGH, A., and PALMER, R. G. (1990), *Introduction to the Theory of Neural Computation*, Lecture Notes I, SFI, Addison-Wesley, Reading, MA.

HERZ, A. V. and HOPFIELD, J. J. (1995), *Earthquake Cycles and Neural Reverberations: Collective Oscillations in Systems with Pulse-coupled Threshold Elements*, Phys. Rev. Lett. *75*, 1222–1225.

HILL, D., EATON, J. P., and JONES, L. M. (1990), *Seismicity, 1980–86. The San Andreas Fault System, California*, USGS Prof. Paper 1515, U.S. GPO, Washington, D.C., 115–152.

HILL, D. P., JOHNSTON, M. J. S., LANGBEIN, J. O., and BILHAM, R. (1995), *Response of Long Valley caldera to the $M_W = 7.3$ Landers, California, Earthquake*, J. Geophys. Res. *100*, 12,985–13,005.

HILL, D. P., REASENBERG, P. A., MICHAEL, A., ARABAZ, W. J., BEROZA, G., BRUMBAUGH, D., BRUNE, J. N., CASTRO, R., DAVIS, S., DEPOLO, D., ELLSWORTH, W. L., GOMBERG, J., HARMSEN, S., HOUSE, L., JACKSON, S. M., JOHNSTON, M. J. S., JONES, L., KELLER, R., MALONE, S., MUNGUIA, L., NAVA, S., PECHMANN, J. C., SANFORD, A., SIMPSON, R. W., SMITH, R. B., STARK, M., STICKNEY, M., VIDAL, A., WALTER, S., WONG, V., and ZOLLWEG, J. (1993), *Seismicity Remotely Triggered by the Magnitude 7.3 Landers, California, Earthquake*, Science *260*, 1617–1623.

HOLMES, P., LUMLEY, J. L., and BERKOOZ, G., *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge University Press, Cambridge, U.K., 1996).

HOTELLING, H. (1933), *Analysis of a Complex of Statistical Variables into Principal Components*, J. Educ. Psych. *24*, 417–520.

HOUSE, L. S., SYKES, L. R., DAVIES, J. N., and JACOB, K. H. (1981), *Identification of a possible seismic gap near Unalaska Island, eastern Aleutians, Alaska*. In *Earthquake Prediction: An International Review*, AGU Monograph, AGU, Washington, D.C., 81–92.

HUANG, Y., SALEUR, H., SAMMIS, C., and SORNETTE, D. (1998), *Precursors, Aftershocks, Criticality and Self-organized Criticality*, Europhys. Lett. *41*, 43–49.

JAUMÉ, S. C. and SYKES, L. R. (1999), *Evolving Towards a Critical Point: A Review of Accelerating Seismic Moment/Energy Release prior to Large and Great Earthquakes*, Pure Appl. Geophys. *155*, 279–306.

JONES, L. M. and HAUKSSON, E. (1997), *The Seismic Cycle in Southern California: Precursor or Response?* Geophys. Res. Lett. *24*(4), 469–472.

KAGAN, Y. Y. and JACKSON, D. D. (1992), *Seismic Gap Hypothesis, Ten Years After*, J. Geophys. Res. *96*, 21,419–21,431.

KANAMORI, H. (1981), *The nature of seismicity patterns before large earthquakes*. In *Earthquake Prediction: An International Review*, AGU Monograph, AGU, Washington, D.C., 1–19.

KATO, N., OHTAKE, M., and HIRASAWA, T. (1997), *Possible Mechanism of Precursory Seismic Quiescence: Regional Stress Relaxation due to Preseismic Sliding*. Pure Appl. Geophys. *150*, 249–267.

KERR, R. A. (1998), *A Slow Start for Earthquakes*, Science *279*, 985.

KING, G. C. P., STEIN, R. S., and LIN, J. (1994), *Static Stress Changes and the Triggering of Earthquakes*, BSSA *84*, 935–953.

KLEIN, W., RUNDLE, J. B., and FERGUSON, C. (1997), *Scaling and Nucleation in Models of Earthquake Faults*, Phys. Rev. Lett. *78*, 3793–3796.

LINDE, A. T., GLADWIN, M. T., JOHNSTON, M. J. S., GWYTHER, R. L., and BILHAM, R. G. (1996), *A Slow Earthquake Sequence on the San Andreas Fault*, Nature *383*, 65–68.

LYZENGA, G. A., RAEFSKY, A., and MULLIGAN, S. G. (1991), *Models of Recurrent Strike-Slip Earthquake Cycles and the State of Crustal Stress*, J. Geophys. Res. *96*, 21,623–21,640.

MAIN, I. G. (1999a), *Applicability of Time-to-failure Analysis to Accelerated Strain Release before Earthquakes and Volcanic Eruptions*, Geophys. J. Int. *139*, F1–F6.

MAIN, I. G. (1999b), *Is the Reliable Prediction of Individual Earthquakes a Realistic Scientific Goal?* Nature website debate, http://helix.nature.com/debates/earthquake/equake_frameset.html.

MARONE, C., VIDALE, J. E., and ELLSWORTH, W. L. (1995), *Fault Healing Inferred from Time-dependent Variations in Source Properties of Repeating Earthquakes*, Geophys. Res. Lett. *22*, 3095–3098.

MCGUIRE, J. J., IMHLE, P. F., and JORDAN, T. H. (1996), *Time-domain Observations of a Slow Precursor to the 1994 Romanche Transform Earthquake*, Science *274*, 82–85.

MOGHADDAM, B., WAHID, W., and PENTLAND, A. (1998), *Beyond Eigenfaces: Probabilistic Matching for face recognition*. In *Third IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, 1–6.

MOGI, K. (1969), *Some Features of Recent Seismic Activity in and near Japan, (2) Activity before and after Great Earthquakes*, Bull. Earthquake Res. Inst., Tokyo Univ. *47*, 395–417.

MOGI, K. (1977), *Seismic activity and earthquake predictions*, Proc. Symp. on Earthquake Prediction, Seis. Soc. Japan, 203–214.

MOGI, K. (1979), *Two Kinds of Seismic Gaps*, Pure Appl. Geophys. *117*, 1172–1186.

MORI, H. and KURAMOTO, Y., *Dissipative Structures and Chaos* (Springer-Verlag, Berlin, 1998).

NANJO, K., NAGAHAMA, H., and SATOMURA, M. (1998), *Rates of Aftershock Decay and the Fractal Structure of Active Fault Systems*, Tectonophysics *287*, 173–186.

NIJHOUT, H. F. (1997), *Pattern formation in biological systems*. In *Pattern Formation in the Physical and Biological Sciences*, Lecture Notes V, SFI, Addison Wesley, Reading, MA, 269–298.

NORTH, G. R. (1984), *Empirical Orthogonal Functions and Normal Modes,* J. Atm. Sci. *41*(5), 879–887.

OUCHI, T. (1993), *Population Dynamics of Earthquakes and Mathematical Modeling*, Pure Appl. Geophys. *140*, 15–28.

PACHECO, J. F., SCHOLZ, C. H., and SYKES, L. R. (1992), *Changes in Frequency-size Relationship from Small to Large Earthquakes*, Nature *355*, 71–73.

PENLAND, C. (1989), *Random Forcing and Forecasting Using Principal Oscillation Pattern Analysis,* Monthly Weather Rev. *117*, 2165–2185.

PENLAND, C. and MAGORIAN, T. (1993), *Prediction of Niño 3 Sea Surface Temperatures Using Linear Inverse Modeling*, J. Climate *6*, 1067–1076.

PENLAND, C. and SARDESHMUKH, P. D. (1995), *The Optimal Growth of Tropical Sea Surface Temperature Anomalies*, J. Climate *8*, 1999–2024.

POLLITZ, F. F. and SACKS, I. S. (1997), *The 1995 Kobe, Japan, Earthquake: A Long-delayed Aftershock of the Offshore 1944 Tonankai and 1946 Nankaido Earthquakes*, BSSA *87*, 1–10.

PREISENDORFER, R. W., *Principle Component Analysis in Meteorology and Oceanography* (Elsevier, Amsterdam. 1988).

PRESS, F. and ALLEN, C. (1995), *Patterns of Seismic Release in the Southern California Region*, J. Geophys. Res. *100*, 6421–6430.

PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A., and VETTERING, W. T., *Numerical Recipes in C*, 2nd edit. (Cambridge University Press, Cambridge, U.K., 1992).

RICHTER, C. F., *Elementary Seismology*, (Freeman, San Francisco, 1958).

ROMANOWICZ, B. and RUNDLE, J. B. (1993), *On Scaling Relations for Large Earthquakes*, BSSA *83*, 1294–1297.

RUNDLE, J. B. (1988), *A Physical Model for Earthquakes: 2. Applications to Southern California*, J. Geophys. Res. *93*, 6255–6274.

RUNDLE, J. B. (1989), *Derivation of the Complete Gutenberg-Richter Magnitude-frequency Relation Using the Principle of Scale Invariance*, J. Geophys. Res. *94*, 12,337–12,342.

RUNDLE, J. B. (1993), *Magnitude-frequency Relation for Earthquakes Using a Statistical Mechanical Approach*, J. Geophys. Res. *98*, 21,943–21,950.

RUNDLE, J. B., GROSS, S., KLEIN, W., FERGUSON, C., and TURCOTTE, D. L. (1997), *The Statistical Mechanics of Earthquakes*, Tectonophysics *277*, 147–164.

RUNDLE, J. B. and KLEIN, W. (1995), *New Ideas about the Physics of Earthquakes*, Rev. Geophys. Space Phys. Suppl. (July) *283*, 283–286.

RUNDLE, J. B., KLEIN, W., and GROSS, S. (1999), *Physical Basis for Statistical Patterns in Complex Earthquake Populations: Models, Predictions and Tests*, Pure Appl. Geophys. *155*, 575–607.

RUNDLE, J. B., KLEIN, W., TIAMPO, K., and GROSS, S. (2000a), *Linear Pattern Dynamics in Nonlinear Threshold Systems*, Phys. Rev. E. *61*, 2418–2432.

RUNDLE, J. B., KLEIN, W., GROSS, S., and TIAMPO, K. F. (2000b), *Dynamics of seismicity patterns in systems of earthquake faults*. In *Geocomplexity and the Physics of Earthquakes*, AGU Monograph, Washington, D.C., 127–146.

SALEUR, H., SAMMIS, C. G., and SORNETTE, D. (1995), *Discrete Scale Invariance, Complex Fractal Dimensions, and Log-periodic Fluctuations in Seismicity*, J. Geophys. Res. *100*, 17,661–17,677.

SAVAGE, J. C. (1988), *Principal Component Analysis of Geodetically Measured Deformation in Long Valley Caldera, Eastern California, 1983–1987*, J. Geophys. Res. *93*, 13,297–13,305.

SAVAGE, J. C. (1993), *The Parkfield Prediction Fallacy*, BSSA *83*, 1–6.

SCHOLZ, C. H., *The Mechanics of Earthquakes and Faulting* (Cambridge University Press, Cambridge, U.K., 1990).

SCHWARTZ, D. P. and COPPERSMITH, K. J. (1984), *Fault Behavior and Characteristic Earthquakes – Examples from the Wasatch and San Andreas Fault Zones*, J. Geophys. Res. *89*, 5681–5698.

SIEH, K., STUIVER, M., and BRILLINGER, D. (1989), *A More Precise Chronology of Earthquakes Produced by the San Andreas Fault in Southern California*, J. Geophys. Res. *94*, 603–623.

STARK, M. A. and DAVIS, S. D. (1996), *Remotely Triggered Microearthquakes at The Geysers Geothermal Field, California*, Geophys. Res. Lett. *23*, (9), 945–948.

STEIN, R. S. (1999), *The Role of Stress Transfer in Earthquake Occurrence*, Nature *402*, 605–609.

STEIN, R. S., KING, G. C. P., and LIN, J. (1992), *Change in Failure Stress on the Southern San Andreas Fault System Caused by the 1992 Magnitude = 7.4 Landers Earthquake*, Science *258*, 1328–1331.

SWAN, F. H., SCHWARTZ, D. P., and CLUFF, L. S. (1980), *Recurrence of Moderate to Large Magnitude Earthquakes Produced by Surface Faulting on the Wasatch Fault Zone, Utah*, BSSA *70*, 1431–1462.

TIAMPO, K. F., RUNDLE, J. B., KLEIN, W., and GROSS, S. J. (1999), *Systematic Evolution of Nonlocal Space-time Earthquake Patterns in Southern California*, EOS Trans. AGU *80*, 1013.

TIAMPO, K. F., RUNDLE, J. B., KLEIN, W., MCGINNIS, S., and GROSS, S. J., *Observation of Systematic Variations in Non-local Seismicity Patterns from Southern California*. In *Geocomplexity and the Physics of Earthquakes*, AGU Monograph, Washington, D.C. (2000).

TURCOTTE, D.L., *Fractals and Chaos in Geology and Geophysics,* 2nd edition (Cambridge University Press, Cambridge, U.K. 1997).

VAUTARD, R. and GHIL, M. (1989), *Singular Spectrum Analysis in Nonlinear Dynamics, with Applications to Paleodynamic Time Series*, Physica D *35*, 395–424.

WYSS, M. and HABERMAN, R. E. (1988), *Precursory Quiescence before the August 1982 Stone Canyon, San Andreas Fault, Earthquakes*, Pure Appl. Geophys. *126*, 333–356.
WYSS, M., SCHORLEMMER, D., and WIEMER, S. (2000), *Mapping Asperities by Minima of Local Recurrence Time: San Jacinto-Elsinore Fault Zones*, J. Geophys. Res. *105*, 7829–7844.
WYSS, M., SHIMAZIKI, K., and URABE, T. (1996), *Quantitative Mapping of a Precursory Seismic Quiescence to the Izu-Oshima 1990 (M 6.5) Earthquake, Japan*, Geophys. J. Int. *127*, 735–743.
WYSS, M. and WIEMER, S. (1999), *How Can One Test the Seismic Gap Hypothesis? The Case of Repeated Ruptures in the Aleutians*, Pure Appl. Geophys. *155*, 259–278.
YAMASHITA, T. and KNOPOFF, L. (1989), *A Model of Foreshock Occurrence*, Geophys. J. *96*, 389–399.

(Received February 20, 2001, revised June 11, 2001, accepted June 15, 2001)