DEPARTMENT OF THE INTERIOR

U.S. GEOLOGICAL SURVEY

A linguistic model of earthquake frequencies applied to the seismic history of California

by

Herbert R. Shaw[1]

Open-File Report 87-296

This report is preliminary and has not been reviewed for conformity with U.S. Geological Survey editorial standards.

[1] U.S. Geological Survey, Menlo Park, CA 94025

UNITED STATES DEPARTMENT OF THE INTERIOR
DONALD PAUL HODEL, Secretary
GEOLOGICAL SURVEY
Dallas L. Peck, Director

For additional information write to:

U.S. Geological Survey
IGP Branch MS 910
345 Middlefield Road
Menlo Park, CA 94025

Copies of this report can be purchased from:
Open-File Report Section
U.S. Geological Survey
Box 25425, Federal Center
Denver, CO 80225
Call (303) 236-7376

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

ABSTRACT

Distributions of historical earthquakes in California and environs have attributes analogous to the statistical properties of language. The availability of complex networks of faults at many length scales (generically called "dendrites") represents the skeletal structure of seismic information sources which play a role like that of lexicographic trees in language and neural dendrites in the central nervous system. A general linguistic model is explored by comparing rank-frequency and size-frequency data of English with the historic seismic catalog in California for earthquakes equal to or greater than M = 5 (referring generally to surface wave magnitudes). Frequency distributions are compared both for explicit word sizes and magnitudes and for mean values of averaged data. The latter have overlapping distributions at comparable frequencies without rescaling. Incremental data are rescaled. An idealized length parameter is defined by $L^* = Base^{y/q}$, where y = M refers to seismic magnitude and y = m refers to word "magnitude" expressed in letters per word. The scaling exponent 1/q is an adjustable renormalization parameter for intercomparisons of different frequency-magnitude sets. The product qc, where c is the coefficient in moment-magnitude correlations, performs the same scaling function for idealized length-moment correlations. Cumulative and(or) incremental plots of log (NL^*_i) vs. log L^*_i, where N is the number frequency at a given value L^*_i, are used to compare frequency distributions and to define aggregate and(or) partial fractal dimensions of the data sets. Shannon's information measures of uncertainty (entropy) and redundancy are also calculated for idealized frequency distributions and for the historic seismic data, and the relations between information and fractal measures are illustrated. Redundancies for the most complete seismic data sets are roughly 25 percent, which compares with about 50 percent for the English language (the same calculations are made for an idealized bond rupture model, providing an alternative method of interpretation). It is concluded that such measures of seismic information are much less complete than language repertoires. With sufficiently comprehensive data, however, seismic redundancies may approach those of language. If so, seismic records could be read with certainties approaching those demonstrated in language reading tests, and(or) in the construction of crossword puzzles. One reason for this suggestion is the fact that the fractal structures of word samples are similar to those of seismically defined fractals when they are comparably rescaled (e.g., the exponents q = 1 for seismic data and q = 3 for word data result in approximate fractal correspondences). Each of these sets is multifractal with subtrends corresponding to partial fractals ranging between about 0 and 3. Multifractal plots are illustrated for the seismic data as a whole and for each of eight subregions in California. There is sufficient indication of self-similarity among the grid-like patterns of partial fractal subsets to provide criteria of potential events based on incompleteness. These results for information measures and multifractal structures also have conceptual parallels with renormalization methods applied to critical phenomena in fluids, the droplet theory of the Ising model of magnetization, and structures of fluid turbulence. In all these cases partial fractal sets reflect structural domains having many scales of length, where different sets dominate under different sets of conditions (e.g., temperature in critical models, generalized deformation states in the cases of turbulent flow and earthquakes). Multifractal seismic distributions are expressed in units of fault dimensions and are compared with fractal fault dimensions obtained previously. Correspondences are used to propose that there is a general fractal coordination of the earthquake process at many scales of length and of

deformation states throughout the western margin of the continent (the distribution is loosely analogous to the coordinations among small and large domains in the process of magnetization near the Curie point). Support for this viewpoint is provided by the peculiar location of California relative to Plate Tectonic features, and also by comparisons of local and global seismic moment rates. An arbitrary map of block domains is constructed and used to generate a space-time plot of the seismic history of California. When the determination of multifractal structures and domain maps are extended to lower magnitudes and larger map scales, it is thought that models combining linguistic, fractal singularity, and mean-field renormalization techniques will offer new insights to earthquake forecasting.

INTRODUCTION

The central hypothesis developed by Shaw et al. (1981) during a study of faulting at the continental scale is that the earthquake process occurs within a volumetrically distributed pattern of fault branching networks that are directly analogous to areally distributed stream drainage networks (stream drainage also has a depth distribution, but the processes are generally confined to a near-surface thickness; exceptions occur where underground streams interact with more complicated flows characteristic of aquifers). A secondary hypothesis is that the total energy of deformation is transmitted through this volumetric network in a manner analogous to hydrologic flows (a percolation process). That is, there are paths of volumetrically diffuse to highly focused trajectories of energy flow over many different scales of length. From the standpoint of length scales, as well as general considerations of energy partitioning, there are parallels with concepts of structural domains in critical phenomena, droplet theories of magnetization, and vortex structures in fluid turbulence (Wilson, 1979). Only the latter will receive further attention in later discussions. "Energy" is meant to refer to all processes contributing to seismic instabilities, including such things as pore fluid flow, mineral reactions, and the rheology of crustal deformations (another hypothesis advanced by Shaw, 1980, is that magma transport in the Earth is a phenomenon that plays a central role, both directly and indirectly, in processes that lead to seismic instabilities; I comment on this relationship in the concluding discussion).

If we can identify characteristic source regimes and characteristic fault patterns (generically abbreviated "dendrites" in the following), then we may also be able to map patterns of energy channeling according to frequency analyses of convergent domains. The purpose of this paper is to illustrate an approach using some statistical techniques borrowed from linguistics. This is possible because there appear to be a number of statistical properties of dendritic structures of the above types that resemble statistical structures known to exist in language. The paper makes both some general and specific comparisons with the earthquake process, derives relationships between concepts of information and fractal geometry, and applies concepts of multifractal singularity spectra to relations between fault-length frequencies, b-values of recurrence curves, and c-values of moment-magnitude relations. I illustrate the conclusions using the historic seismicity of California and environs for earthquake magnitudes of 5 and larger.

THE EVOLUTION OF REDUNDANCY

The linguistic model is shown in Figure 1 as a guide to subsequent illustrations of patterns of earthquake occurrences. Metaphorically, the

Figure 1 near here

linguistic development of a particular story line is assumed to be analogous to a system of evolving slip events that become focused within the branching hierarchy of a system of earthquake faults. The analogy is between the statistical measures of generalized story lines and generalized fault lines along which earthquakes are expressed. Seismic events are therefore hypothetical analogs of some set of linguistic units of measurement represented by letters, words, and(or) syntactical combinations of words. Thematic ramifications of limited subject matter are analogous to fault systems. The statistics of letter and word occurrences constitute the data

base for event frequencies in the linguistic case; these are the analogs of seismic catalogs.

In the generalized case (linguistic or seismic) probabilities become enhanced during the building of a story from a structurally diffuse and highly uncertain reservoir of words to form a relatively redundant array. This happens during ordinary discourse, or in the process of writing about a particular theme. Something like this also happens in the processes that produce earthquake recurrences within certain segments of fault zones. That is, probabilities are greatly enhanced in those localities relative to other zones within the same system, or within other fault systems. Figure 1 symbolizes the idea of overlapping subject matter and intersecting story lines. Specific subject matter is either potential (never actually written

Figure 1 near here

before) or represents some aspect of previous themes that recurs because it has acquired new interest; these alternatives are analogous, respectively, to the formation of new fault breaks and to reactivation of existing faults.

Because the logic of Figure 1 is central to the idea that there is widespread coordination among earthquake events (globally, but specifically in California), I attempt to explain with other parallels and analogies why I think it is relevant. Documentation given later is expressed in more conventional terms, but the notion that there is a fabric of seismic information flow in space and time persists as the linguistic theme of this paper (e.g., Figure 25 is essentially a space-time plot of information flow involving the entire western margin of North America encompassing California).

The coordinate system of Figure 1 is vague because it is impossible to adequately portray linguistic variability on a two-dimensional map. The general idea, however, can be illustrated in terms of an imaginary language in which each statement is made up of a pair of words. Temporarily ignoring the dendritic lines, consider each circle to be a source represented by a person who may become a protagonist in a potential three-way conversation. I say protagonist because a priori any one of the three potential sources might take the lead after a conversation begins, or the lead might switch from one to another source. Also I say potential because the conversation may or may not happen depending on factors involving such things as proximity, common interests, and so on. Also, imagine that there is a coordinate grid of some sort within each circle, such as a matrix, that identifies the probability that a particular word-pair "sentence" may be uttered by that protagonist. Words exist outside the circles, but the probabilities that they will be used is effectively zero; that is, each circle identifies the working vocabulary and repertoire of statements (articulation factor) of the speaker. The circles are not necessarily the same size, reflecting differences in sizes of vocabulary and abilities of articulation. Thus, every entry in the matrix of available word pairs within each circle identifies a possible statement.

In order to consider how an interaction might evolve, initially consider the location of each circle and its probability structure to be independent of any other circle. The simplest example is that of three people who are initially either not within audible range of each other, or who have not caught each others attention (e.g., at a cocktail party). As proximity increases, however, the circles of influence begin to overlap in one or more of the above respects. While the statements continue to be independent, it is now possible that within areas of overlap between pairs of circles, or all three circles, a coincident statement might occur. At that stage of evolution

two alternative futures are available: (a) Coincident statements may be noted in passing but may not be pursued (e.g., everyone has had the experience of saying or thinking something related to what someone else has just said, but then electing not to respond with thoughts suggested by the incident). (b) A coincidence stimulates a response that triggers another response that leads to a longer series of more or less coincidental statements (in the example this is represented by greater redundancies for word pairs activated within each circle than would occur for less specific language samples).

The areas of independent coincidences between two sources are indicated by the hachured areas in Figure 1, and a three-way coincidence by the crosshatched area. However, once a responsive mode occurs, such as (b) above, the subsequent interactions are no longer independent, although the general ranges of probabilities still exist. At this point an interplay of all three circles begins, so that the definition of a conversation subsequently applies. One way to describe the resulting influence on linguistic structure is to imagine that the probability matrix becomes rearranged so that certain word pairs have enhanced probabilities in the areas of overlap. This is equivalent to saying that the statements emanating from one source influence those of one or both other sources which in turn reciprocate, and so on. That is, a system of feedback ensues (note that written literature could also be described as a conversation, but one where there are many potential sources of feedback; also cause-effect responses are less direct and are sometimes buried deeply in the thematic structure).

The evolution of such rearrangements, and the consequent systems of feedback, are schematically indicated by the converging dendritic paths in Figure 1. Branches that tap two circles symbolize mutual reinforcement of statements that progressively dominate the three-way conversation, as indicated by the heavy arrows. It is assumed in the simplified case that sources take turns, but in any actual situation the sources can operate simultaneously. As in ordinary conversation, then, the statistical structure of thematic content sorts itself out via the processes of feedback.

To be complete, at this point there should be a set of symbols indicating the reciprocal effects of the three-way interaction on the subsequent probabilities in each circle, but that is harder to illustrate even in a schematic way. In other words, Figure 1 shows a stage of evolution where a conversation (or written dialogue, in the loose sense, as in the scientific literature) is just beginning. Linguistic trends, or story lines, would then propagate in specialized directions depending on the strengths of the feedback circuits among the sources. If the dialogue persists indefinitely a special pattern of word probabilities may become entrenched, meaning that a characteristic dendrite of feedback interactions has become redundantly incised (this is analogous to the development of a long-lived pattern of stream erosion, or of earthquake faults).

The above type of evolving interaction is reminiscent of schemes used to describe other complex systems. An example sometimes cited is Young's nineteenth century theory of color vision; see description by Erickson (1984). In that case, each circle would represent a pattern of receptors (color receptors being the systems of cones in the retina of the eye rather than persons or systems of faults) each of which is sensitive to the full spectrum of colors but has a different sensitivity function for each color.

The analogous receptors of the regions in Figure 1 represent sensitivity distributions to stimuli; the responses of each receptor are similar in kind but differ in terms of their maximum sensitivity to different types of stimuli (e.g., the response functions of different persons to words heard in conversation, or the responses of different faults to a source of

deformation). If the functions were exactly the same in every case there would be no variety in the responses of different receptors to different stimuli; there would be total uncertainty as to the association of patterns of reception and the patterns of their source functions. A conversation would be more like a chant in unison, perception of colors in vision would always be mixed the same way hence could not be discriminated, and all fault systems of the same size and geometry would have the same frequency distribution of earthquakes.

In the case of vision, Young hypothesized that only three types of cones are required to adequately codify the full spectrum of color perception. Neural feedback then governs subsequent reactions to the codified color patterns that are perceived, an aspect beyond the scope of Young's optical theory (see remarks by Erickson, 1984, concerning applications to mechanisms of perception in general). A seismic analog of color perception might be constructed by assuming that there were a few basic types of fault in each region, each having a different but similar response function; in combination, the response functions uniquely codify the complete repertoire of source interactions. Patterns of seismic radiation might be viewed as analogous to the visualization of color stimuli, the latter being the spatially and temporally diverse deformation phenomena impinging on a system of faults.

This model of communication, where a system of response functions operates in such a way that a more-or-less unique mapping of patterns of source functions is codified by a simple system of receptors (the neural network, or a network of potential sites for seismic slip events), also illustrates a sometimes contradictory aspect of information flow. In the theory of vision this ambiguity is easy to understand. For example, if cones were identical in response (individually or in sets that give the same kind of output everywhere on the retina), then their construction could be described as perfectly redundant. As a consequence, however, one cone or set of cones would see exactly the same color pattern as the entire retina, hence different color fields would not be identified in a mappable way, except in terms of intensity contrasts. Therefore, from the standpoint of color perception the response functions would be totally uncertain in that every part of the retina would be receiving and passing along the same indiscriminate pattern; there would be no correlation between a neural coordinate system and a source coordinate system.

A system of seismic receptors within a fault dendrite representing available loci of slip events has analogous attributes. For example, consider an idealized subduction-type fault system that is characterized completely by a linear array of segments each of which experiences earthquakes of the same total moment over a long enough time. If each segment responds in exactly the same way to a source of deformation that is uniformly applied, and if there are many such segments, then the sequence of activations is totally uncertain. If each segment responds in a way that spans the same frequency spectrum but at different characteristic magnitudes, then there is a potential basis for discriminating patterns of seismic events. In the same sense as in vision, a high degree of mechanistic redundancy can imply high discriminatory uncertainty, and a certain degree of mechanistic uncertainty (representing a diversity of response functions) can lead to high redundancy of spectral discrimination (i.e., the same combinations of complex stimuli produce a codifiable mapping of responses among what would otherwise be an indistinguishable array).

Descriptive ambiguities between uncertainty and redundancy can be avoided to some extent by defining them quantitatively in terms of specific numerical measures of frequency in space and(or) in time. This is attempted in later

illustrations of information contents of frequency distributions. Often, however, the data are inadequate to make the explicit distinctions that were just described in terms of idealized models.

These considerations suggest a more general theory of neural response systems. One compatible approach is illustrated by Edelman (1978). There, the analogous feedback shown in Figure 1 has been termed "phasic reentrant signaling". Phasic reentrant signaling refers to the fact that phases of stimulation (signals from and to the sources of the present model) become recognized in a manner such as those described above and are recorded by recognizer neurons. These recognizers then act as sources of signals for other neural responses, meaning that information is both received and reentered as new signals by the same arrays of neurons. The analogy with the earthquake process is represented by the fact that the responses of fault systems are influenced by changes that lead up to each seismic event, and each event reciprocally influences probabilities of future occurrences in the same and other regions of earthquake potential (that is, there is feedback among foreshock and aftershock effects of both local and regional scale that can influence the character of behavior in terms of greater or lesser degrees of sensory recognition, hence long-term predictability).

Another analogy that demonstrates the roles of uncertainty and redundancy is provided by the way in which neural dendrites in the skin permit the localization of a pinprick. Sperry (1974) describes the effect with a diagram that also resembles Figure 1. One system of nerve fibers is distributed so broadly that only the general region of pain can be recognized, but systems of interwoven fibers (the neural networks) permit very precise mapping of location on the brain. This kind of circuitry is similar to those aspects of feedback in Figure 1 by which information becomes localized and reentered to influence patterns of new responses (the part not shown).

It is such pinpointing effects that are envisioned to influence the development of a specific type of linguistic response and the pinpointing of a particular sequence of seismic events within a large system of potential earthquake faults (the development of a seismic theme). These cases, however, are the inverse of the pinprick in that the analogous fault dendrites are also an intrinsic part of the energy sources that provide the stimulus for an event (something similar is illustrated by the phenomenon of a virtual sensation, a reversal of circuitry in which a locus of pain is induced internally by real or latently imprinted causes and is projected to a specific locus in the body; latently imprinted refers to such things as a sensation of pain in an area of the body where there are no actual nerve endings, either naturally or by reason of removal of portions of the anatomy).

Returning to the linguistic context of Figure 1, the three areas of initially independent discourse might be viewed as revolving around a generalized subject, for example the physics of the earth. As the independent discourses grow and overlap, at some stage a common intersection occurs in a form that becomes focused on a particular geophysical context (that is, it can propagate indefinitely in the direction of a particular subset of word frequencies; "direction" in this sense is a highly generalized idea in that the actual path of word combinations represents a multidimensional and highly nonlinear process). An example suggested by the present subject is discussion of the physics of the earthquake process. After the idea was first introduced, a specialized linguistic theme evolved that has a probability structure characteristic of the subject. For example, a special repertoire could have started with the word "earthquake" which was followed by the array of usages that attend discussions of earthquakes. In this context the word earthquake and many other related words occur with probabilities that are

greatly enhanced over what they would be in ordinary conversation or in unspecialized writing. An exception, which makes the same point, is that in certain areas of the world the word earthquake, or its equivalent, is a common household word because of a particularly high frequency of earthquakes. I describe the dynamical complexity of this kind of evolving process by the phrase "redundancy of action", because it is generally impossible to separate out and describe all of the acting forces and their specific response functions.

MEASURES OF REDUNDANCY

Qualitatively, Figure 1 symbolizes a map of disperse earthquake epicenters among a potential population of epicenters for which some characteristic pattern is just beginning to emerge. To give numerical perspective, imagine a matrix of all possible English words arranged as in a distance chart. In a large dictionary there are more than 10^5 different words. Thus there would be over a hundred thousand columns and rows in the matrix. But even if such a matrix were constructed it would identify only those syntactical units limited to word pairs, as was assumed initially above. A story line corresponds to a route through such a chart, a route selected from all possible word combinations by some feedback process. To describe such a process, however, the data would have to be plotted in a very large hyperspace (i.e., one with 10^5 degrees of freedom rather than two). This is far too many variables to trace or to visualize any specific geometric path through the data. The analogous size of a seismic "vocabulary" is poorly documented compared with language, but the number of degrees of freedom must also be large.

The situation is clarified with a simpler example. Ignoring spaces and punctuation there are $26^2 = 676$ possible pairs of English letters. All these pairs can be shown in a two dimensional array of squares each containing the value of probability that a given pair of letters will occur in ordinary English (Konheim, 1981, Table 2.3.4). It is well known that the probabilities are highly nonuniform and reveal a certain measure of informational uncertainty that can be calculated. Language, however, produces ordered sequences of letter pairs. These appear as linear trends, localized subsets, or as tortuous linear paths through the square array; patterns of pair-probabilities look something like a crossword puzzle (e.g., Konheim, 1981, Table 2.3.4). Some pairs (e.g., t followed by h) are encountered frequently, others are rarer, and some may not be observed at all.

A model of seismic frequencies based on statistical concepts applicable to language might be useful in developing generalized methods of long-term earthquake forecasting because we have an innate experience of how to read such patterns even when we do not understand exactly how they evolved. Present methods of seismic forecasting rely heavily on detailed knowledge of recurrence times in specific localities, or return times for some expected level of ground motion in a given vicinity; see, respectively, Lindh (1983) and Algermissen and others (1982). A linguistic type of forecast would not replace these methods of prediction, rather it would ideally provide a framework from which experienced forecasters could "read" a pattern of probabilities in the same sense that experienced readers can evaluate missing letters and(or) words in a sample of written text. For example, it has been found by means of various tests that the English language is about fifty percent redundant (see numerical definitions below). This is based on several types of evidence, one of which is the fact that under controlled conditions readers can assess the content of messages in which about half of the symbols are missing; see Shannon and Weaver (1949), Miller and Chomsky (1963), Pierce

(1980). Estimates of redundancy based on reading tests are consistent with estimates of redundancy based on numerical measures of information.

Shannon (see Shannon and Weaver, 1949, p. 48 ff.) often referred to the measure of information as "the uncertainty" of a given suite of words or symbols, but he also coined the synonymous term "informational entropy". The word entropy is now commonly used to indicate information measured in terms of the negative of the mean logarithmic probability of a set of probabilities. The term uncertainty, however, is descriptively apt for the problems of earthquake forcasting and will be used here unless a more specific comparison is being drawn with entropy concepts employed elsewhere. The equivalence between formulations of informational and thermodynamic entropy are discussed by Jaynes (1957), Tribus and McIrvine (1971), and by Pierce (1980). Applications of a "maximum entropy concept" to the statistics of earthquake frequencies have been given by Berrill and Davis (1980) and Main and Burton (1984).

Redundancy is a numerical measure explicitly derived from the numerical uncertainty of Shannon, although other meanings exist (e.g., Edelman, 1978). A simple demonstration of the relation between numerical uncertainty and redundancy is given by the ratio U_{obs}/U^*, where U_{obs} is the calculated uncertainty of a set of observed probabilities as given below (the observations may refer to counts within sets of letters, sets of words, etc.), and U^* is the calculated uncertainty of the same number of symbols of that type assumed to have equal probabilities. An example using the 676 letter-pairs of the English language makes the definition clearer. The observed uncertainty is calculated from the equation:

$$U_{obs} = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad (1)$$

where n (= 676 in the case of letter pairs) is the number of different symbols in the set, and p_i refers to the observed count of one of the 676 possible pairs of letters in English divided by the total count of different pairs in the sample that represents the context of interest (e.g., from the probability matrix in Konheim, 1981, Table 2.3.4). The equal-probability reference state is then given by:

$$U^* = -n(1/n)\log_2(1/n) = -\log_2(1/n) = \log_2 n \qquad (2)$$

Thus, in the equal-probability limit the reference uncertainty is given by the base 2 logarithm of the number of symbols in the set.

The terms "maximum entropy" or "minimum redundancy" are often encountered in descriptions of statistical models. Because these limits are relative to a particular system of symbols, the terms may be misleading. In any event it is necessary to refer such limits to a defined reference state. The equal-probability limit provides such a reference because it represents the inability to discriminate between possible outcomes among a specifically defined set of symbols. As usual the sums of probabilities of the symbols in a chosen set are normalized to unit value (dimensional constants required to distinguish different informational and physical entropy states are ignored).

Shannon called the ratio U_{obs}/U^* the relative entropy (relative uncertainty here). Designating this ratio by the notation $F = U_{obs}/U^*$, Shannon then defined the difference (1 - F) as the redundancy, called R here; see Shannon and Weaver (1949, p 56); Khinchin (1957); Gatlin (1972). These definitions provide a simple numerical scale that ranges between zero and one in either F or R. For F = 1, R = 0, the uncertainty is maximized for that set of symbols (i.e., the most uncertain state is given by a situation where every

symbol is equally probable; an analogy would be where each magnitude and(or) location is represented with equal liklihood in an earthquake catalog). Zero redundancy equivalently means there are no special recurrences of particular symbols or events in space and(or) in time. The inverse, F = 0, R = 1 can occur only if one symbol has unit probability and the rest zero probability (e.g., the monotonous recurrence of a particular letter or word, or the precise recurrence of a specific earthquake magnitude, especially at a given location, in the absence of any other events). In this case the pattern is said to be 100 percent redundant (R will be expressed as a percentage in numerical examples). Subsequently, when I use the term redundancy it is with these numerical definitions in mind.

Because these measures of information are sensitive to how symbols are defined and the sizes of the populations of given types of symbols, it is not immediately obvious how these concepts should apply to fault frequencies and seismic events. Therefore, in order to show that the linguistic approach has more than intuitive appeal it is necessary to show that sets of seismic symbols have properties in common with language. Later the statistics of the English language and the historic catalog of California earthquakes are directly compared.

Language codes, as mentioned above, have statistical structures in the neighborhood of 50 percent redundancy. In one sense this might be thought of as a fifty-fifty "mix" of both highly uncertain and highly redundant word combinations. However, this is not an accurate picture of the pervasive character of the statistical structure shown by experimental tests. Campbell (1982, p. 70 ff.) refers to experiments in which texts can be completed when more than half of the symbols are missing if the sample is large enough and the subject is reasonably familiar to the reader (if such an experiment were to apply to technical writing in a vigorously expanding field of research such as seismology or molecular biology, however, the reader would have to be up-to-date on every aspect of the changing terminology).

For the sake of argument, if it were found that the informational uncertainty of the earthquake process, neglecting dimensional constants, were the same as that of language, then it could be expected that with sufficiently large samples experienced forecasters could read the seismic record with the same confidence as readers of incomplete language texts. That is, if a seismic catalog could be shown to be about fifty percent complete with respect to stipulated bounds of thematic content (e.g., as expressed by structures and feedback processes contextually like Figure 1), then the rest of the expected events could be filled in. Like the text readers, the forecasters would be able to say what events are missing, or will occur next, in the catalog based on its overall statistical structure even if it is not possible to give a precise genetic description of each event and(or) there mechanical interactions. This conjecture has rather serious caveats, such as: (1) no compendium of seismic data exists with the degree of etymological definition found in a language dictionary, and (2) the statistical structure symbolized in Figure 1 must not change so rapidly that criteria of completeness are invalidated as fast as they are documented (e.g., a dictionary would soon be useless if the words were continuously being replaced with new ones).

Weighed against the first caveate is the fact that Main and Burton (1984) have shown that information theory itself provides some tests for the degrees of departure of given suites of seismic data from completeness (this approach is explored later in additional respects); Shaw and Gartner (1986) have also shown that it is possible in principle to learn how a seismic pattern corresponds to the time-dependent properties of fault dendrites. The second caveate is difficult to assess because it depends on the first one. The

following discussion of Figures 1 through 3, however, suggests that incised patterns of earthquake faults are not expected to suddenly change, although the internal patterns by which they are activated may be variable over both short and long time scales (see Shaw and others, 1981). That is, the vocabulary of fault motions available for seismic conversations is probably not as changeable as, say, the terminology of molecular genetics (e.g., compare current texts with those a few years old). But the ways in which a stable repertoire of earthquake faults is sampled by seismic events may be entirely unknown.

The linguistic approach to seismicity has some parallels with concepts of ramifying structures that describe geometric features of the faulting process (see Shaw and Gartner, 1986). Therefore, in the next two sections I consider some aspects of seismic prediction based on vocabularies represented by fault patterns.

A SCHEMATIC REDUNDANCY MODEL FOR CALIFORNIA SEISMICITY

An implication of the linguistic description of a process analogous to Figure 1 is the ability to state a conditional probability that a specific word will be uttered by one of the speakers given the history of an established conversation that has higher than average redundancy. One way to test this idea would be to analyze long samples of conversation in the same way as is done in reading tests mentioned above. Another might be to do a predictive test in which observers of the conversation are asked to anticipate such things as the next word spoken, who the next speaker will be, and so on. Carrying this approach to an extreme, it would imply such things as anticipation of all the words of the next sentence by one of the speakers.

Qualitative examples of such predictions are familiar. We have all observed conversations in which it is sometimes possible to make a good guess at who will say what next. Trivial examples of high conditional probabilities are given by responses to salutations. The repertoire of these responses, however, is very restricted compared to more general conversations. Obviously, such abilities are limited to special cases when the range of predictions extends beyond a few words (for discussions of ranges of statistical influence in written language see the previously cited references on information theory). If longer range predictions are possible at all they are based on great familiarity with the context in which an anticipated statement will be made (e.g., certain outcomes of stereotyped storybooks, movies, and so on, can be anticipated well in advance, the classic being "And they lived happily ever after.").

The idealized goal of earthquake prediction resembles the foregoing examples. Theoretically, we would like to be able to predict the locations and times of earthquake occurrences into the indefinite future. A prediction of what fault system is likely to be active next is analogous to specifying the next speaker, and the prediction of a specific event in that system is analogous to anticipating the next word the speaker will say. The greater goal of specifying sequences of events for a given fault system is analogous to anticipating an entire sentence structure by one of the speakers. Such analogies identify not only the complexity of the undertaking (analogous, say, to developing an ability to predict exactly what every head of state in the world is expected to say next), but they also suggest that the degree to which predictions will be possible is contingent on a familiarity with earthquake statistics comparable to that of language.

Lacking an ability to predict whole sentence structures, it is not an impossible goal to predict missing letters and words in an established and

ongoing dialogue. The corresponding analogy is the possibility that given a knowledge of those fault systems that interact it may be possible to make improved guesses of next events in a given system, and possibly even which system is most likely to make the next seismic statement. At any rate, a demonstration of the ability to evaluate the range of influences throughout a network of interacting fault systems is prerequisite to anything more sophisticated. A corollary of this implication is the idea that any physical hypotheses used to make long-term predictions of large events within any one of several interacting fault systems are suspect unless they can also make predictions of the types just mentioned. Rather than attempting the task of evaluating conditional probabilities using incomplete data, the approach taken here is to pose a model whereby patterns of seismic interactions among systems of faults can be studied as a prerequisite to further discussion of statistical structures, linguistic or otherwise.

A generalized model of earthquake fault dendrites is shown in Figure 2 following the logic of Figure 1, except that the pattern is now viewed as a

Figure 2 near here

map scaled roughly to the area of the state of California. The circles of Figure 1 are left open at the perimeters, symbolizing the idea that California is an area surrounded by Plate Tectonic activity in which the boundaries of motions are of a larger scale, areally and in depth, that is not sharply delineated. The symbols E, CN, and BJ stand, respectively, for the seismically active areas off the northern California coast near Eureka, eastern California to central and northwestern Nevada, and Baja California. These areas have distinct relationships to Plate Tectonic features (e.g., Suppe and others, 1975; Herd, 1978), but for the present they are based only on the observation that they seem to represent frequency-magnitude patterns of earthquake activity that resemble each other more than they do interior regions of California.

One reason to keep the discussion general is that attempts to correlate specific fault sets with specific earthquake models often prove unreliable if there is not a great deal of control on mappable faults that correlate directly with the characteristics of seismic events (examples of reasonable control are localities where there is a great deal of information on the distributions and histories of seismic gaps). In a general way, therefore, region E also corresponds to the Northern California Offshore Zone of Ryall and others (1966), and CN encompasses much of their Ventura-Winnemucca Zone. The trends E-BJ and CN-BJ represent a convergence between the Central and Southern California Zones and the Ventura-Winnemucca Zone (see Insets in Figure 3 below). The areal distribution of faulting and earthquake activity discussed by Wallace (1984) categorically represents the CN portion of the patterns of Figures 2 and 3 here. From the plate tectonic perspective, these three zones are essentially coincident with the unstable triple junctions shown by Suppe and others (1975, Figure 5) for the western U.S.

In addition to the open-endedness of the areas in Figure 2 they have unspecified distributions in depth. This aspect is also beyond the scope of discussion here, but it is implied that these domains effectively constitute "cornerstones" of the seismic distribution in which there is an extended three-dimensional interaction with processes in the Earth's mantle. As discussed in Figure 1, the dendritic patterns schematically indicate a hypothetical stage of converging deformation that leads to the development of redundant structures in the common areas of overlap. Two-way and three-way conversations now represent the activation of characteristic fault sets (the

incised story lines of the generalized history represented by the dendrites). The central common area can be viewed as the core region of present day fault patterns in California.

Obviously, the evolution of actual fault patterns did not develop from an original homogeneous distribution as would be implied by a literal parallel with Figure 1. The history of deformation and faulting discussed by Shaw and others (1981) is very long compared with activity considered here (the situation is analogous to ancient language themes that are structurally revised by renewed discussion).

Taking into account the orienting effects of older structures, Figure 3 shows the inferred pattern of the three-way overlap that would exist if it were assumed that present day activity followed the logic I have outlined.

Figure 3 near here

Rather than having circular or parabolically open shapes, however, the common area takes the form of an intersection of folded or bifurcating U-shaped zones. Although this is a schematic accommodation of Quaternary fault distributions, it also reflects a convergence of mantle sources of deformational energy with the Earth's crust. However, no implication is made for specific types of dynamic coupling (dynamical concepts are not considered here, except as are implied by the statistical relationships between earthquake frequencies, magnitudes, and moments). The insets in Figure 3 show how the generalized pattern relates to event distributions and contours of "tectonic flux" determined by Ryall and others (1966) from earthquake data; see Wallace (1984) for descriptive examples of dynamic coupling in western U.S. suggested by the geologic history.

The rest of the paper offers a preliminary evaluation of the linguistic model outlined in Figures 1 through 3 in four respects: (1) The relation between fault dendrites and concepts of steady states of seismic activity is explored. (2) Similarities and differences between the statistics of language and the statistics of earthquakes are illustrated. (3) Evidence required to evaluate the degree to which an area such as California has elements of large scale coordination is examined in terms of the evolution of seismicity within smaller spatial domains over time. (4) An approach to forecasting is outlined based on the concept of interacting domains.

FAULT DENDRITES AND PATTERNS OF STEADY OR TRANSIENT EVENT FREQUENCIES

A schematic overview of relations between temporal frequencies and event magnitudes for fault dendrites is shown in Figure 4 (the language analogy would be represented by frequencies of syntactical units plotted against a measure of size, such as letters per word, words per phrase, etc.; however, exactly corresponding lexicographic dendrites are not easy to portray because of the combinatorial degrees of freedom). The heavy solid line in Figure 4a

Figure 4 near here

symbolizes the steady state, meaning that counts per unit time remain invariant over a characteristic length of time, and there would be a characteristic and constant value of b in the equation:

$$\log f = a - bM \qquad (3)$$

for M at specified increments, or for f as a cumulative frequency of events

equal to or larger than M (the numerical range of frequencies will be normalized in various ways in later discussion; unless otherwise stated I use incremental distributions). The heavy dotted lines show how such a curve might be shifted parallel to itself according to global fluctuations in the spatially averaged deformation field. By contrast, the dashed lines indicate nonproportional transient variations relative to the time scales of the global variations. As a rough guide to the time scales, the orders of magnitudes for transient swings might be from decades to centuries, and for more global variations from thousands to millions of years (this is an artificial discrimination that defines steady or transient only in terms of time scales possibly relevant to hazards analyses).

Bifurcation styles are indicated in Figure 4b; Figures 4c and 4d show how subsets of trends might be distributed for the same or dissimilar branching ratios. If the latter represent steady states for each set they are also subject to the transient effects mentioned. Figures 4e and 4f show how individual fault sets might be activated in different styles, depending on local properties. Figure 4e represents a case where branches behave differently in terms of percentage activations, some sets reacting in groups over some part of their lengths (steeper trends) and some reacting over progressively longer sections of an individual branch (flatter trends). Figure 4f represents the latter mode on the assumption that there is a single fault branch activated in segments (a bamboo-like effect); the trend in such a case tends to be flat because the segmentation limits the potential number of active segments of given length. The same distribution is possible for a poplar-like dendrite of Figure 4b, but it is not similarly limited in potential bifurcation ratios (i.e., the cumulative length is not restricted as in the single segmentation model).

One basis for recognizing time-dependence comes from estimated ages of fault activation on the continent scale. Within the thirty different faulting regions outlined on the map of Howard and others (1978), the distributions of numbers of fault lengths measured by Shaw and others (1981) have remained broadly similar over different age categories ranging from the Holocene to at least the Miocene (see Wallace, 1984, for durations of activity on specific systems of faults in the Great Basin province during the late Quaternary). This means that variations roughly average out over time spans in the range 10^4 to 10^7 years. This happens by compensating shifts between systematically higher and lower levels of activity, either uniformly or in a teeter-totter fashion. By uniformly it is meant that all frequencies of rupture events are increased and decreased together over spans of time shorter than those represented by the data set as a whole. By teeter-totter fashion it is meant that transient increases in the frequencies of smaller rupture events are compensated by energetically equivalent decreases in the frequencies of larger rupture events, and vice versa (Figure 4a). Trenching studies by Sieh (1981, 1984), Sieh and Jahns (1984), and by Weldon and Sieh (1985) provide evidence of variations along the San Andreas system over the latest 2000 years; variations at the 100-year scale are represented by patterns of seismicity illustrated later.

For purposes of discussing departures from concepts of average long-term steady states, a "short" time scale means less than 10,000 years. It is within this temporal range that the schematic variations of Figure 4 are assumed to apply, and in which the tripartite distributions in Figures 2 and 3 are assumed to be characteristic of statistical structures. Within this time-frame, the regions designated E, CN, and BJ are viewed as relatively stable and persistent "cornerstones" that are analogous to the protagonists of the conversational modes of Figure 1. By contrast, when these same regions

are viewed as structural elements of Plate Tectonic motions they are considered to be unstable features subject to changing kinematic styles over longer times; see discussion of plate reconstructions in Suppe and others (1975, Figure 5). The possible existence of such dualities over different scales of time and space are important to keep in mind during later interpretations of statistical redundancies.

The concept of time-dependence favored here reflects changes in the number-length ratios within branching patterns of discrete length hierarchies as a function of time. The total length of fault that is active in the steady state is determined by overall bifurcation ratios and characteristic branch lengths that take part in the earthquake process. Geometrically, if bifurcation ratios are large, then there is a large increase in total length of faults with decreasing branch length within a particular length hierarchy. The converse is true when ratios are small. The former situation seems to describe most fault regions in the contiguous U.S. (Shaw and others, 1981). Faulting in some regions, however, can have nearly a constant total length per length order of both long and short faults. That is, the bifurcation ratio is such that the sum of fault lengths within a particular length order is about the same as those within each of every other length order; the product N_i x L_i per fault order (taking L_i to be a characteristic length per branch for that order consisting of N_i branches) is then constant for all L_i.

High bifurcation ratios tend to represent the bushier branching patterns, and low ratios the poplar-like patterns of Figure 4b. This is not necessarily true if the latter represents a densely braided structure, as occurs in many places along the main trunk of a major fault system such as the San Andreas; see Wallace (1973). If more length were represented in the longest fault branch than exists in the total of all shorter branches, then there would be an approach to the model of a segmented single main trunk (bamboo-like structure) in which segment lengths could be either constant or variable. The data of Shaw, and others (1981) suggests that the bamboo-like limit never actually occurs in faulting despite discussions of faults as individual linear entities. An alternative is a branching structure topologically collapsed into a surface trace with a vanishing overall width compared to the map scale at which the length is expressed. In such cases it would be difficult without intensive trenching studies to distinguish between activity on subsidiary branches of a ramifying system of related faulting and activity between segments of a single fault trace.

I prefer to view all faults as dendritic in structure with a distinctive bifurcation pattern wherein it is also possible for any branch to rupture over subintervals of a statistically averaged length (see Shaw and Gartner, 1986). This approach is not in conflict with results for several areas where it is assumed that each event is associated with movement along the total length of a specific fault (the so-called "maximum moment" or "characteristic earthquake" model) if one assumes that a particular strand or set of strands is associated with each event. An analogous discussion referred to the concept of a continuous single fault surface is given by Aki (1984) in terms of concepts of asperities and barriers. The San Andreas system represents families of structures of the above kinds, depending on model emphasis; the descriptions differ from concepts of characteristic earthquakes only if the San Andreas is viewed as a single segmented fault trace. Such a viewpoint is not supported by the mapping of excavated intervals in the trenching studies of Sieh and coworkers cited above.

The above distinctions may be semantic or real depending on how they influence concepts of recurrence times. That is, the recurrence time might refer to: (a) the activation of the total length of each one among a set of unique faults, (b) the activation of characteristic subintervals of a single long fault, or (c) the time between events of comparable magnitude in a fault dendrite, In case (c), there is not necessarily a one-to-one correspondence between recurrence time and geographic locality. Such distinctions may explain contrasts between estimates based on either seismic or geologic evidence of faulting; see Evernden and others (1978), Wallace (1984), Aki (1984). Alternatives are illustrated by comparing extreme versions of these three types of systems.

In type (a) imagine that there is an infinite number of discrete faults in which there are faults of just the right length to go with any conceivable earthquake in a seismic catalog. In this case the interpretation of recurrence time depends on evidence of discreteness of both the event magnitudes and the numbers of faults with charateristics suitable to those events. For example, if there were precisely 30 characteristic event magnitudes distributed at exactly 0.1 magnitude units over three decades of magnitude, and there were precisely 30 and only 30 unique faults of exactly the right lengths to produce each event, then the time between like events might reasonably be associated with the recurrence of a characteristic event on the same characteristic fault (i.e., in exactly the same geograhic location). But what if the number of faults of similar length in broadly the same geograhic vicinity was unlimited? Then the recurrence time for each characteristic event would be effectively infinite in terms of any exact location within the domain. That is, there are so many possible fault strands available that each one need be activated only once during any duration significant to hazards forecasts. The same fault trace, whether expressed as the trajectory on the ground surface or as a planar distribution in depth, is never activated exactly the same way twice. Therefore, the time between events of like magnitude is just a property of the sampling rate of a distribution of independent fault surfaces.

Regularity between the intervals at which the same magnitude occurs in the type (a) interpretation would be a property of the availability of the right size fault for local rates of deformation of the system as a whole. Prediction, then, would depend on an understanding of two general factors: (1) the geometric distribution of all discrete faults, and (2) the pattern of deformation rates that impinge on that distribution (I say pattern because, in general, deformation may be homogeneous or highly heterogeneous, and it may vary slowly to rapidly in time in any of its spatial modes). If one combines these factors in a global sense, they are referring to the rheology of the earthquake process. In that respect the type (a) model could also be expanded to allow for generation of new faults and(or) dimensional rearrangements of existing ones.

The type (b) interpretation suggests that there is a distinctly limited relation between the length of a single fault and the earthquake magnitude associated with that length. Therefore, if every event implies a finite segment length and there is a finite number of magnitude intervals, then each magnitude can be associated with a particular segment of the fault. If the segments are truly characteristic of particular nonmigrating patches of the fault, hence particular geographic locations, then recurrence times give an explicit representation of the hazard potential of each vicinity. Prediction then depends on evidence of regularity of deformation patterns and of how well

different characteristic patches of the fault can be described. The description must include evaluation of whether or not the segments are spatially invariant or may migrate along the main fault trace, the latter implying uncertainty in exactly where a specific recurrence time applies. If there is regularity in both the timing and the location of characteristic activation events, then the earthquake process is deterministically predictable (this does not require harmonic regularity, because time-predictable models also permit specific predictions for irregular timing if rate-of-recovery data are available from geodetic measurements).

The concept of recurrence times applied to fault dendrites in the type (c) interpretation combine elements of both the type (a) and (b) interpretations. If the dendritic pattern is unlimited, type (c) is like type (a) except that instead of an infinite number of discrete faults there is an infinite number of discrete dendrites (a forest effect). Whatever uncertainty exists in (a) is compounded. On the other hand, if a dendrite can be associated not only with a characteristic hierarchy of branch-lengths in a specifically ordered array, but also with a particular geographic domain, then the type (c) situation is analogous to type (b) relative to the "basin area" within which the dendrite is expressed. Therefore, deterministic prediction would also be possible if there is a fixed relation between branch lengths, bifurcation ratios per order, and a delimited geographic area associated with each branch order. This is the type of model explored by Shaw and others (1981) and by Shaw and Gartner (1984, 1986). Its apparent complexity relative to (b) is because there is presently an insufficient data base to test the time-dependent characteristics of dendritic systems of faulting. However, numerous studies by Wallace (see Wallace, 1984) suggest that characterization of fault systems according to principles of hierarchical ordering may be achievable.

Shaw and Gartner (1984) suggest that there may be a quantitative parallel between the type (c) interpretation and hydrologic principles of dendritic ordering where there are characteristic proportionalities between numbers, lengths, areas, and energy discharge (stream flow or energy of deformation). The simplest form of this approach would be analogous to having a time-predictable model for a characteristic dendrite in which the response of connected sequences of the smaller branches determines the magnitude-time relation for the next larger branch, and the integrated slip of all branches determines the time-magnitude potential of the main trunk fault.

The manner in which subsets of any of the above types of hierarchies are activated determines our experience of the characteristic areal and time variations of both large and small earthquakes. The difficulties of prediction already mentioned in the linguistic context are echoed by ambiguities that apparently exist among the types of faulting interpretations just described. That is, there are many examples where one or another type of model appears to be conditionally successful as a predictive tool (some data indicate high redundancy and others high uncertainty in the sense of numerical information). By this token, then, apparently the earthquake process contains aspects of all these viewpoints concerning recurrence times. This suggests that complexities of the sorts seen in the statistics of language, where there is a general dynamical process capable of producing simultaneously ordered and disordered sets of responses, are at least qualitatively paralleled by the earthquake process. Type (c) models geometrically resemble the notion of lexicographic trees, hence it might be expected that earthquake fault dendrites are capable of self-induced structural ordering. If this parallel can be strengthened, then there is a strong incentive to document geometric, kinematic, and linguistic properties of fault dendrites as an adjunct to

concepts of forecasting.

With these geometric properties in mind, an imaginary model for the qualitative description of time-dependence in (c) is given by the motion of trees of different characteristic structures during waxing and waning wind storms. Very leafy trees may absorb or deflect energy in the motions of the outer branches until the wind velocity reaches a magnitude sufficient to break limbs and(or) uproot the tree. By contrast, an isolated palm tree takes the brunt of the force on a single long trunk. Other shapes and bifurcation ratios, such as those characteristic of pine trees, range between these limits. During the waxing phase in any of these examples the strain is first felt in the smallest structures and progressively propagates throughout the dendritic hierarchy if the wind is sustained long enough. For example, the strain may ultimately influence even the dendritic root hierarchy, but there the analogous energy flow would be reversed (the alternatives mimic some of the distinctions between energy flow in river tributary-distributary systems, and in strains related to foreshock and aftershock sequences; the local sources of energy transport have different dendritic distributions, though both "ends" of the system are responding to the same total energy transfer).

As the wind continues to rise, the smaller structures appear to become relatively static as the motion is taken up by the larger branches. This static appearance represents the state where the local strain rate on the small branches drops because they have reached the limit of local displacements, unless they are stripped off. These conditions are analogous to local sequences of foreshock microfracturing and characteristic earthquakes. As the wind velocity wanes the reverse sequence is seen, but the temporal and spatial modes differ considerably; relaxation slip events appear to diffuse outward from the central focus of strain accumulation (the tree limbs, largest activated faults, etc.). This scenario is a dendritic analogy of relations between asperities, barriers, and foreshock-aftershock phenomena described by Aki (1984).

If the dendritic scenario were used to describe earthquake distributions in large fault dendrites, the following sequence of observations might be expected: (1) Initially there is a quasi-steady state where events with small to intermediate magnitudes are distributed over the smaller outer branches of a dendritic fault array within a characteristic geographic area, in contrast with a series of foreshock patches restricted to a specific fault plane; compare Kanamori (1981). (2) If the energy input continues, this distribution becomes suppressed in favor of larger and more widely spaced events enveloped by the spatial distribution of preceding smaller events, possibly culminating in a single largest event; the sequence parallels the pattern of information flow of Figure 1 within a system such as Figure 3). (3) As the deformation wave temporarily subsides, the reverse sequence results in another rise in the rates of small-scale events generally emanating from the loci of the strongest events and dispersed by progressively smaller branches of the dendrite; compare Kanamori (1986, Figure 12). (4) All events decrease in frequency, or the waning phase merges with the onset of a new deformation pulse so that the frequencies of the smaller events remain relatively steady for some time, or possibly even continue to increase (see later discussion of the 1952 Kern County event in Figures 18 and 19, and Table 2).

The frequency-magnitude distribution for this history corresponds in a general way to a fault number-length hierarchy that also has a characteristic pattern of b-values (see Shaw and Gartner, 1986, Figure 6); step-like and(or) oscillating recurrence curves are expected from such interpretations (e.g., curves plotted by Shaw and Gartner, 1986, Figures 7 and 8; Ryall and others, 1966, Figure 7, shown here in Figure 5 below). Since the average slopes of

recurrence curves depend on the way the gross available moment rate is distributed among regional patterns of slip rates or fault length activation rates, the steady state tends to alternate between limits that average to a characteristic b-value (see Figure 4).

In Shaw and Gartner (1986) the available moment for earthquakes is given by the total fault length activated during a specified time, hence the frequency-magnitude distribution for a fault-set of fixed bifurcation number is simply partitioned according to the logarithm of activated length. This automatically gives b = 1 for the condition of constant activation rate and constant bifurcation state. For steady state dendritic models of variable bifurcation states, however, b-value trends depend on how the total moment is partitioned among proportional length orders, hence the range and average of the net distribution of b-values can depart from unit slope depending on the sequence by which the dendrite as a whole is activated (compare Shaw and Gartner, 1986, Figure 8 with Figure 4 here).

This schematic description explains in a general way the transient teeter-totter effects in Figure 4a. That is, the net effect is a combination of partial and complete activations of branches in the fault dendrite. Consequently, aspects may represent the effect of either the progressive activation of portions of individual faults or may reflect complete activation of characteristic branches of all faults in the dendrite (the portion of a branch activated may be less than the total length and still be characteristic, depending on the failure conditions that affect all branches systemwide). In the terminology of Wesnousky and others (1983) these are the "b-value" and "maximum moment" types of trends, respectively. The latter type is the characteristic earthquake model of Coppersmith and Schwartz (1984) generalized to a fault dendrite with characteristic branching ratios as illustrated by Shaw and Gartner (1986). In the terminology of earthquake families discussed by Aki (1984), the former family is of the "barrier-type", and the latter of the "asperity-type" if the scaling is referred to patches restricted to a single fault surface rather than to a distribution of surfaces within a volume.

The above ideas are complicated by the feedback effects implied in Figures 1 through 3, because rather than dealing with the analogy of a single tree, the history of motions is more like that of a forest. In either case an important feature is the potential reciprocity between activity on the smaller and larger structures in relation to the propagation of a general wave of deformation throughout the aggregate system. Many features of this kinematic distribution represent a generalization of sequences described elsewhere in terms of seismic gaps, seismic cycles, and "Mogi doughnuts"; see Evernden and others (1978), Ellsworth and others (1981), Mogi (1985), Hill and others (1985), and Shaw and Gartner (1986). I use the general model to guide an inspection of earthquake frequencies in California and to examine additional relationships between the statistics of lexicograhic and fault dendrites.

PRELIMINARY REMARKS ON FREQUENCY-MAGNITUDE RELATIONS IN CALIFORNIA

The purpose of this section is to summarize the seismic data that will be used to draw more quantitative relationships between statistical descriptions of language and earthquakes. Figure 5 reviews the earthquake history of California in the form of a frequency-magnitude plot showing the pre-1900, post-1900, and 1800-1982 distributions of events equal to or larger than M = 5

as given by Real and others (1978), Toppozada and others (1979a,b), and Sherburne and others (1985). Toppozada and others (1986) give a more

Figure 5 near here

up-to-date but geographically restricted tabulation of events above M = 6. The latest event listed in their table is April 24, 1984, a 6.2 event SE of San Jose, CA. Figure 5 in Toppozada and others (1986) shows a map of historic earthquake damage in California that mimics the form of Figure 3 of this paper.

(Note: Events in California since October, 1982 are not included in the data base, but the later events do not significantly change the patterns shown; events that occurred during 1986 while this report was being written qualitatively fit the pattern described in later illustrations of areal distributions of activity, especially the Chalfant Valley foreshock-aftershock sequence beginning in July, 1986 east of Long Valley, California having maximum magnitudes around M = 6. Magnitudes given here are as reported in the literature cited without evaluation, unless otherwise stated. The moment-magnitude scale of Hanks and Boore, 1984, shifts the higher magnitudes to lower values at the same values of moment because it refers to local magnitudes, M_L. Their calibration offers a uniform basis for evaluation of effects of instrument saturation. The unqualified use of magnitude scales, however, can lead to confusion because the differences in local, surface-wave, and moment-based magnitudes are greater than a magnitude unit for the larger events in California: the evaluation of M_L for the 1906 San Francisco event according to Jennings and Kanamori, 1979, gives the range as 6 3/4 to 7 compared with the usually accepted surface wave magnitude, M_S, of 8 1/4. The value M = 8.2 is adopted here as the maximum event in the data base. Boore, 1984, reduces the values and range even more, as shown by the box plotted in Figure 2 of Hanks and Boore, 1984. Uncertainties related to magnitude scales are discussed by Kanamori, 1983a, and are not addressed here except with regard to the geometric implications of relations between moment and magnitude as reflected in the variability of the coefficient c in the equation $\log_{10}M_o = d + cM$. The range of variation in coefficient c is indicated by the slope in Figure 2 of Hanks and Boore, 1984. Presumably such variations occur for any scale based on instrument magnitudes that are subject to saturation effects. In the absence of explicit evaluation it is assumed that magnitudes cited here are roughly equivalent to M_S.).

The trends of log f vs. M in Figure 5 illustrate aspects of the frequency variations discussed above. This is shown by the episodic character and the flat portions at the highest magnitudes. Apparently this behavior reflects the occurrence of three great earthquakes within a half century in California (1857, 1872, 1906). The average slope for magnitudes less than about 7 is slightly below b = 1. If this represents an approach to the steady state it is of the oscillatory character described in the previous section.

The interpretation of these trends as effects of transient swings of fault activation statistics contrasts with the maximum entropy type of steady state model illustrated by Main and Burton (1984). There frequencies are assumed to vary in a regular progression modified by the effect of instrument saturation causing a roll-off toward low frequencies at the highest magnitudes. These saturation effects, and limiting geometric effects suggested by von Seggern (1980) and Caputo (1982) are significant aspects of any interpretation (compare Shaw and others, 1986, Figure 17c). Here, however, they appear to be overridden by oscillatory shifts among characteristic smaller and larger fault-length sets where the frequencies of the largest events are increased

rather than decreased relative to a uniform Gutenberg-Richter type of trend. This is shown more clearly later in terms of relations between frequency distributions, fractal geometry, and information measures.

Interpretations of the meaning of an increase or decrease in frequency obviously are model-dependent in the absence of both a standard earthquake catalog for the steady state and a standard geometry of failure that permits a one-to-one correlation with measurements of lengths, areas, and volumes involved in the deformation that contributes to each event. This is a variation on the theme of uncertainties related to the teeter-totter effects in Figure 4. The high frequencies of the larger events in Figure 5, relative to extrapolation of the trends of the smaller events, represents an increase only if the general trend is characteristic of a distribution that describes a quasisteady or standard sort of behavior of constant b-value. If the feedback effects involving foreshock-aftershock sequences are typical (as they must be in any sufficiently general picture of the earthquake process) then the concept of a standard state characterized by monotonic b-values may be illusory. If so, the description of frequency shifts is ambiguous in the absence of a more explicit geometric context.

An instructive example of this ambiguity is shown by the contrast between the geometric interpretation of fault rupture due to von Seggern (1980) and the geometric interpretation of activated fault dendrites of Shaw and Gartner (1986). The conclusions of these two studies appear to agree concerning variations of b-values related to the fractal geometry of fault failure. However, they differ on the question of temporal and geographic correlations. The von Seggern (1980) interpretation refers to the topography of a continuous fault surface of complex geometry and stress distributions (the fractal geometry of this case is analogous to that of rough surfaces, or to a variably contorted suture-like interface). It was shown by him that between major events on such a surface there is an expectable decrease in b-values; i.e., foreshock sequences in several examples are shown to have lower b-values than aftershock sequences.

Cycles of variations in b-values discussed by Shaw and Gartner (1986, Figures 8 and 18), invoke uncertainties of the model types discussed above. The interpretation favored there is that a drop in frequencies of small to intermediate events distributed widely over a regional fault dendrite compensates a frequency increase on geographically more restricted and longer strands of a main trunk fault system relative to longer-term averages (a hydrologic analogy is given by flood stages along different reaches of a major river dendrite). Event distributions derived from variations in fractal dimensions can be artifacts of the choice of model, as is also the case with asperity-barrier concepts, and the differences can be geographically significant. This difference is related to the temporal-spatial uncertainty in the interpretation of recurrence times already discussed.

An important conclusion by von Seggern (1980) is that monitoring of the time-dependence of b-values in the fractal context should have great predictive significance; similar conclusions are drawn by I.G. Main (written commun., 13 June, 1986) and by Main and Burton (in press). Von Seggern (1980, p. 639) stated that an interval of low b-values should be considered as a hazardous state. In California, however, there are intervals of time for differing geographic domains where b-values (for M greater than 5) appear to remain low during repeated large events. It is also observed that a rise in frequencies of smaller and presumably independent events (i.e., independent of the frequency increase related to aftershock sequences discussed by von Seggern, 1980) on a regional scale followed by a decrease may be associated with occurrences of large earthquakes. As shown later, such transients can

wax and wane over large areas with timespans apparently shorter than a decade (the scales of influence are indicated by regions defined in Figure 14).

Evidently the time-dependence of b-values has to be monitored over broad geographic domains throughout the area of Figure 3 at yearly (perhaps even monthly) time scales if premonitory patterns are to be adequately documented. Other criteria of coordination are needed to ameliorate such stringent monitoring requirements and to distinguish between concepts of predictable seismic cycles and random transients of the teeter-totter types.

The central point in these comparisons is that the geographic regions envisioned here are large compared to those implied by any single-fault type of interpretation of foreshock-aftershock sequences. Although the contrast may be semantic from a fractal viewpoint, there is a major distinction that should be of concern in a region such as the San Francisco Bay Area where there are several fault strands capable of M = 7 events or larger (an alternative way of expressing the contrast in descriptive geometry might be to say that the "far-field" stress distribution considered secondary in a fault-surface interpretation is a "near-field" stress distribution in a fractal dendrite model; compare Shaw and Gartner, 1986, Figure 19).

Another distinction between the present approach and previous ones is that the concept of complex geometries involves arrays of fractal dimensions rather than unique values characteristic of specific events. That is, there are many possible routes in fractal space from one seismic state to another, and any particular outcome represents a multifractal confluence of geometric distributions of energy sources. To illustrate, there is a parallel between the variable fractal paths constructed by Shaw and Gartner (1986, Figure 18) and the multifractal interpretations of trajectories in turbulent flow described by Halsey and others (1986). The concept of singularity spectra derived in the latter study is used later to describe seismic frequency distributions in terms of multifractal sets.

COMPARISON OF LINGUISTIC AND SEISMIC FREQUENCY-MAGNITUDE RELATIONS

Concepts of geometric reference states can be examined in a more general way on the basis of the linguistic approach previously outlined. That is, if there are no specifically unique geometries, perhaps it is possible that there is a characteristic statistical pattern within which there is imbedded more than one set of fractally self-similar distributions. In order to examine this question I compare some of the properties of English word statistics with earthquake statistics. This approach might be alternatively expressed as a search for characteristic departures from a maximum entropy type of model.

Frequencies of English words, expressed as logarithms of probabilities based on samples of two different total sizes, are plotted vs. the logarithm of word rank in Figure 6. Rank refers to the sequence of different words

Figure 6 near here

arranged in order of decreasing frequency. If all words occurred with the same frequency the curve would be horizontal and the order of words would be arbitrary (i.e., every word would be equally probable). If the slope were nearly vertical one word would be most probable and every other word would fall into some hierarchy of rapidly decreasing probabilities (the vertical limit is a one-word vocabulary of unit probability). The earthquake catalog of Figure 5 (1800-1982 data set for magnitude increments of 0.1) is plotted on the same graph for comparison, where frequencies have been normalized to unit probability. The seismic trend parallels the hyperbolic trend for English

words over the first decade of rank and then falls off to lower relative frequencies (the hyberbolic trend is usually called Zipf's law; see Shannon and Weaver, 1949; Miller and Chomsky, 1963; Mandelbrot, 1965). In general, such a roll-off can be an artifact of several kinds of truncation. The most obvious is the effect of a limited sample regardless of magnitude range. This is the case where the rank is truncated and the rank-frequency range is restricted. In addition, however, even in long records there appears to be a roll-off at the larger magnitudes that results from effects such as instrumental saturation or an approach to a limiting moment (see discussions by Berrill and Davis (1980), Main and Burton (1984, in press), and Shaw and Gartner (1986, Figures 17 and 18).

If a particular truncation hypothesis were to be tested in a manner comparable to the statistics of language, it would be necessary to have a data set resolved to a rank of at least ten thousand intervals, or ten thousand distinguishably different events by whatever criteria. By comparison the historical record in California, as shown, is of rank 33 for intervals of 0.1 units of magnitude, so ten thousand "words" defined in terms of the same range of seismic magnitudes would imply a resolution in M of 0.001.

If it were assumed that the roll-off seen in Figure 6 were statistically characteristic, it could be said that a steady-state seismic catalog was made up of a language-like part combined with a highly redundant tail. Similar effects are seen in small and(or) highly specialized samples of English. In extreme cases the first part of the curve may be nearly flat (equiprobable range) with an abrupt truncation; a square corner or stepwise distribution of this sort could be thought of as a hypothetical limit where there are mixed intervals of total uncertainty and total redundancy. Mandelbrot (1965) derived a more general relation than Zipf's law that has a flat portion at low rank and approaches a linear limit of nearly hyperbolic form at high rank. It is found here, however, that in large samples the simple hyperbolic trend is closely approximated, and much of the shape effect is related to either sample size, or, what may amount to the same thing, unusually stereotyped redundancies. These ideas are explored further below.

The same data are shown in Figure 7 in terms of the logarithm of probability plotted against word "magnitude", using the term generically to indicate the measure of word size as well as seismic event size. At first

Figure 7 near here

glance there would appear to be little resemblance between the language curve and the seismic trend. It is possible, however, to dimensionally renormalize the data in ways that bring out some features that are common to both data sets. In order to demonstrate the logic, it is assumed that the magnitude of a word (letters per word) is analogous to a seismic magnitude in the sense that it is a function of a length scale. A proper length scale is not known *a priori*, so I arbitrarily choose proportionalities that later can be compared with correlations among fault length, seismic moment, and magnitude. For example, empirical correlations suggest that M is proportional to activated fault length raised to a power somewhat larger than one (depending on the definition of length relative to moment and magnitude, the exponent might range from less than 1 to 3: see Mark and Bonilla, 1977; von Seggern, 1980; Shaw and Gartner, 1984, 1986). In such cases the meaning of length depends on the nature of source geometry, hence it depends on the same kinds of dimensional considerations we are attempting to deduce from fault patterns and frequency-magnitude relations. I do not resolve this circularity here, but quantitative patterns are found that illustrate constraints on self-consistent

geometries based on fractal concepts (compare von Seggern, 1980).

For simplicity I assume that both M and m (where m corresponds to letters per word) are directly proportional to the same logarithmic length scale. Therefore, this length is given by the log base raised to either M or m. For example, using base 2 to emphasize this convention, the relation is expressed by the following proportionalities:

$$L^*_s = 2^M \quad \text{(seismic)} \tag{4}$$

and

$$L^*_a = 2^m \quad \text{(English)} \tag{5}$$

It is simpler to use base 10 so that $\log_{10}L^*$ is numerically identical to either m or M, but the choice of a different base emphasizes the artificial convention, and base 2 is commonly used in information theory. This defines the hypothetical length L^* as a size measured in bits given by m or M. As shown later, renormalization can be accomplished by multiplying m or M by a coefficient required to give the same length-frequency relation.

The definition of L^* in the seismic case is tied to a particular magnitude relation (and logarithmic base) which can be redefined if a more rational basis is found. That is, the definition can be adjusted for relations that represent actual geometric length-moment-magnitude scaling, meaning that the definition of L^*_s (where the subscript refers to a seismically defined length scale) could be standardized by choosing a given relationship as standard. Thus L^* is a kind of fictive yardstick the meaning of which depends on calibration. I hope to show that a concept of multifractal sets may be useful in establishing an approach to such calibrations.

This demonstration is approached in several steps designed to establish the meaning of the fractal context in terms of the data already given. Subsequently I also demonstrate how these measures relate to uncertainty and redundancy so that there is a common thread of reasoning that permits some intercomparison of frequency-magnitude data with equivalent representations expressed in terms of complex geometric sets and Shannon's measures of information. I begin with frequency-length data for words expressed in a form that permits the definition of fractal dimensions based on the length scale L^*_a (the subscript refers to the alphabetic "length"). This is done in a manner directly parallel to Mandelbrot's (1982) derivations for dissection sets and coastline lengths. The respective total length values are obtained from products of frequencies multiplied by L^*_a, either incrementally or cumulatively. These values based on the curve for English words in Figure 7 are shown in Figure 8. The incremental basis considers the geometric object

Figure 8 near here

to be discontinuous, or dust-like, hence it can be "less-than-linear" (the topologic dimension, D_T, in Figure 8 is between 0 and 1). By contrast the cumulative curve expresses sets of line lengths distributed in a plane, so the corresponding range of D_T is between 1 and 2.

The latter range can be envisioned in a manner parallel to measurements of a coastline or suture line using calipers of varying aperture as the variable yardstick. Starting with the largest linear segment expressed by $N_iL^*_a$, length accumulates at a rate depending on the lengths of shorter segments $N_jL^*_a$ as L^*_a is decreased. This procedure parallels the measurements of cumulative fault segment lengths versus frequency class as illustrated by Shaw and Gartner (1986, Figure 5 and Figures 13 through 16). If a formally constructed set of line segments, such as a Cantor set (see Figure 11), is

viewed as a geometric array distributed in a plane, the same type of measurements can be made. In that case the dimension is also variable, as found in Figure 8, even though the set theoretic definition gives a constant fractal dimension (0.6309 in the case of the standard Cantor set of Figure 11). Other linear arrays, such as some fault dendrites measured in this way, are fractally self-similar over a characteristic range of length scales, as demonstrated by Shaw and others (1981) and Shaw and Gartner (1986) in terms of the slopes in logarithmic plots of dendrite number-length sets (the value of slope, S, is approximately equal to -D; e.g., Shaw and Gartner, 1986, Figure 5).

The essential nature of fractal sets is shown by numerical examples where the fractal dimension has integer values that are numerically the same as in the case of Euclidean forms. In contrast with Euclidean geometry, a fractal object of integer dimension can range from a dust-like complex array to a solid-like complex array. That is, it may look nothing like the idealized picture of a line, planar surface, or three-dimensional object. The fractal dimension does not necessarily suggest what the appearance of a macroscopic object may be like. A fractal object with D = 1 may look little like a line (although it could); it might be a cloud-like dust of volumetrically disperse, areally disperse, or linearly disperse trends. By the same token, a D = 2 object might, on the one hand, look more like a line than a plane, or on the other hand look more like a Euclidean three-dimensional object than a plane (examples are the diverse spatial arrangements of fault sets, which may scatter over large volume domains or cluster in restricted depth ranges with highly elongated trends).

A classic example of a linear form that approaches D = 2 is an ammonite's sutures (the surface trace looks like an intricate line but is of fractal dimension closer to two). The convolutions of the cortical "surface" of the brain offer an analogous example of a planar form that approaches D = 3 (the cortex layer nearly fills a topologic volume); see numerous other examples in Mandelbrot (1982). The concept of the degree-of-filling of a specified topologic domain offers the simplest verbalization of a fractal object, if the domain can be specified.

In the above terms the fractal structure of English, referred to the reference scale L^*_a, apparently is of composite nature and can't be identified with a particular topologic domain (linguistic topology relates to the degrees of freedom discussed in Figure 1). On the incremental basis, defined by values of log NL^*_a vs. log L^*_a in Figure 8, the shorter words are of point-like character (D near zero), whereas the longer words (m greater than 8) display fractal self-similarity with values closer to unity (D is roughly 0.9). I emphasize that this is relative to the definition of L^*_a, so there is still the uncertainty concerning the proper length scale; a different fractal set based on a redefined L^*_a will be shown. By the same token, these definitions do not necessarily resemble the discussion of lexicographic trees in Mandelbrot (1982), although the incremental set here is closely related to those ideas. If a relationship between the fractal structure of language can be related to the physical topology of a phenomenon such as earthquakes, the correspondence also has interesting implications for linguistics.

Miller and others (1958) in attempting to model the statistics of word lengths as a Markov process found two distinctly different ranges of behavior, as though there was a different alphabet for short and long words (the short-word model was fitted by a 4-letter alphabet, an observation that has interesting parallels with the genetic code; their research was prior to the Watson-Crick discoveries). They explained this contrast in terms of different

statistical structures for what they called "function" words (articles, prepositions, pronouns, numbers, conjunctions, auxiliary verbs, and certain irregular forms) vis a vis "content" words (nouns, verbs, adjectives and adverbs). The former compose most of the shorter words, hence they correspond roughly to the point-like fractal set here. The content words, on the other hand, are longer and more varied, hence they are more dendritically distributed in the fractal sense.

The great variety of big words in English would seem to be a major distinction between language and earthquake statistics. That is, long words individually tend to have low probabilites (exceptions being repetitive technical words such as "earthquake", "probability", etc.), but there are more different types of long words than short ones. If earthquakes were distributed that way there would be a great many different kinds of large magnitude events associated with relatively few types of small events repeated much more frequently (i.e., there would be meaningful recurrence times for small events but effectively infinite recurrence times for the same large event, as discussed earlier). Linguistically, the frequency of usage of specific small words is higher than that of large words in order of ranking but there are relatively few different types of small words (e.g., "the" is usually the most frequent word in samples I have examined, large or small). Frequencies per length class fall off rapidly at word lengths shorter than 4 letters (compare Figures 6 and 7). In my counts based on the Hanley (1951) catalog, words less that three letters in length make up about five percent of the sample while words longer than ten letters constitute almost ten percent. Words that are five through ten letters in length represent 75 percent of the sample.

The short-word, long-word distinction can be rationalized to some extent using the idea of multifractal sets as a demonstration of the composite character of word structures (earthquakes or language). The linguistic short-word limit is a form of truncation analogous to roll-off at the largest earthquake magnitudes (i.e., there is a physical limit at one letter which is reflected in the information potential and structural roles of two- and three-letter words). In this respect earthquakes and words are somewhat the inverse of each other. There is an irreducibility in word size (barring analysis of letter formation, speech sound patterns, etc.) analogous to the limits imposed by instrument response and(or) extendability of fault rupture controlling the source moment for a "maximum" earthquake. Inversely, there is no theoretical limit on the maximum potential size of words nor on the minimum size of seismic events (at least down to the atomic scale of intracrystalline slip).

Explanations for ultimate limits in either case relate to questions of dynamics as well as geometry. For example, there is no theoretical reason why we couldn't invent an alphabet with unlimited numbers of one-letter symbols. It is simply a matter of fact that the evolution of language reflects a dynamics wherein there is a balance between unlimited expansion or contraction of alphabetic degrees of freedom. In the context of word lengths, this tendency was described by Zipf (1949 p. 22) as a balance between "forces of unification and diversification", and by Mandelbrot (1961) as a kind of random walk in which there is an equilibrium between probabilities of lengthening a word and shortening it (compare Miller and Chomsky, 1963; Mandelbrot, 1982). The limits on earthquake sizes are analogously constrained, the maximum being limited by such things as crustal thiknesses and fault lengths relative to the velocities of sound transmission through rock (barring such things as continuous mega screw dislocations, supersonic shock waves, etc., that theoretically could feed on the total strain energy of the Earth). At the

small end, detection of the acoustic energy emitted during deformations is limited only by the sensitivity of recording devices for motions at the atomic scale. Analogously to the construction of alphabets and words, there is a certain probability for lengthening faults and raising potential event magnitudes by coalescence of ruptures, and there is a countering probability for shortening them by healing, dissection of longer faults, and so on.

It can be expected that the magnitudes of words and earthquakes would resemble each other more if the tendency of word lengths to be unlimited could be renormalized. Renormalization does not address the lower seismic limit because earthquakes smaller than M = 5 are not included in the data base. A clue to renormalization is suggested by the fact that the information of large words tends to be amplified over that of root words (i.e., letters per word are increased in additive groups, but the information conveyed is greater than the sum of the parts; if this weren't true there would be no "selective pressure" to construct long words in the first place). It is well known (particularly in this age dominated by communications media) that the ability to express complex ideas is expanded more than in direct proportion to the increase in the size of a person's vocabulary.

A literary example often cited for informational compression related to an expanded vocabulary is the writing of James Joyce; "compression" refers to the idea that information per word increases with vocabulary size, so the "information density" increases. This compression, however, apparently reflects linguistic structures at a more subtle level than is shown by overall word frequencies. For example, in Figures 6 and 7 Joyce's vocabulary, at least as it is expressed in Ulysses, is not different from more mundane samples of English. At the same time it is obviously unusual with respect to diverse usages and the occurrences of peculiar words unlikely to be found elsewhere (there are 29,899 different words in a total of 260,430 words listed alphabetically in Hanley, 1951). Many constructions characteristic of Joyce's style are beyond the range of word lengths included in my counts (e.g., words such as "mangongwheeltracktrolleyglarejuggernaut" are not included). However, his tendency to use words longer than ten letters is probably well represented (e.g. words such as the 14-letter "lizardlettered" and the 16-letter "philoprogenetive").

The data in Figure 7 from Miller and others (1958, Figure 2B) represents 5537 different words taken from three different styles of written English: (a) the King James version of the Bible, (b) William James Talks to Teachers, and (c) the Atlantic Monthly (department called "Atlantic Reports" for April and May, 1957). This is only 19 percent of the diversity shown by Joyce in a single work. If there are significant differences in the statistical structures of such seemingly diverse styles of writing, additional techniques are needed beyond those shown by trends of frequencies, rankings, and word lengths that are shared in common in plots such as Figures 6 and 7. It is proposed here that concepts of multifractal sets is such a technique, although its application to language is not pursued beyond the comparison with seismic frequencies (see Figure 9 below).

The notion that the written expression of complex thought patterns is compressed (or the "amount" of meaning per word expanded) by the use of large words seems to contradict another common notion that the most powerful expressions of scientific ideas are those described using the smallest possible words. If both are true, then they clearly reflect another (or the same?) optimization principle, because it is difficult to imagine that Einstein could have done a better job describing the theory of relativity using the 850-word vocabulary of Basic English (see remarks on sizes of vocabularies and relative redundancies in Shannon and Weaver, 1949, p. 56).

Someone might assert that the most powerful scientific ideas could be expressed mathematically without using any words at all. This just changes the type of symbolism, however, so analogous properties of relative uncertainty and redundancy (and the implicit, if nebulous, autocatalytic optimization principles) would exist in a "language-free" codification of mathematical principles.

Making arbitrary use of the concept of compression, I define a new length scale for language by use of the relation $L^*_a' = 2^{m/3}$; the exponent 1/3 is chosen arbitrarily to bring the range of word lengths somewhere near that of seismic lengths. The result is shown in Figure 9 compared with the seismic data of Figure 5 expressed in terms of the length scale $L^*_s = 2^M$. On the

Figure 9 near here

revised L^*_a' basis there is a much closer resemblance between the fractal structure of words and earthquakes. This resemblance is shown by the similar sets of slope trends. Note that crosstrends occur at fairly regular intervals, suggesting a tendancy toward geometric progressions. The trends of negative slope correspond to fractal dimensions (D) between 2 and 3, while the crosstrends, or jumps, of positive slopes correspond to values of D between 0 and 1. Slopes greater than +1 do not correspond to any standard topologic limits in that they imply negative D, while slopes less than -2 imply D greater than 3 (the latter are possible regardless of logarithmic base if the geometry refers to dynamic variables in an n-dimensional phase space). The grid of light lines shows reference sets of fractal loci corresponding to D = 0, D = 1, and D = 3 (relative to the base 2 definition of length). One line for D = 2 is shown because some subsets of the steeper cumulative seismic trend are closer to that value than to D = 3. Notably, the use of cumulative distributions smooths the data so that the fractal arrays are obscured (a similar observation was made by Shaw and Gartner, 1986, concerning the significance of irregularities in recurrence curves based on incremental rather than cumulative frequency data).

If the incremental distributions fell along these trends with relatively regular spacings it could be concluded that the linguistic structure of words and earthquakes was made up of a nearly periodic array, or fabric, of point-like to volume-filling sets. Seismic or linguistic style could then be represented in terms of the varieties of weaves and their intricacies, much in the sense that one judges handwoven rugs in these terms. Literary criticism seems to use a qualitatively similar concept in drawing comparisons between styles of authors. It could be conjectured that the weave for an author such as James Joyce would be relatively intricate. Another possibility, however, is that there are nonperiodic components that produce a more continuous spectrum of fractal sets. The former are analogous to what are termed mode-locked quasiperiodic structures in patterns of fluid convection; the latter are called singularity spectra. Physical systems in general may alternate between such styles of complex behavior, and the range of regimes may also include more strictly periodic and more chaotic patterns that do not display the same richness of structural organization. The analogy is explored farther in later illustrations (see Figures 12 and 19).

It would be desirable to test these ideas against other samples of language and earthquakes in order to ascertain whether or not a universally valid classification is possible. Rather than attempting this here, however, I tentatively accept the existence of such fractal sets in order to explore relationships between them and the information expressed by such patterns. Whatever the proper scaling of lengths may turn out to be, and whether or not

it is unique, exponents in the terms $2^{m/3}$ and 2^{M} can be generalized by defining a variable exponent, q, giving functions of the form $L^* = 2^{M/q}$. An expectable range of q might be from 1/3 to 3, depending on whether the length scale is being stretched or contracted.

In the fractal context, q is not necessarily limited to rational numbers. If there are parallels with structures of fluid turbulence, however, mode-locking tends to occur in the vicinities of rational values of dynamic forcing parameters; these might be revealed as fractional multiples of a standard value, q_0 (conceivably there could be hybrid sets with more than one exponential multiplier, $q_{0,i}$). Because q is a power of L^* it can be thought of as a fractal dimensional coefficient; that is, it is a parameter required to normalize the magnitude scale relative to a common set of fractal dimensions.

By the same token, we can arbitrarily raise all length scales to any other constant dimensional coefficient. Therefore we can also define a length scale by the relation $L^{*qc} = 2^{cM}$, where c is a coefficient of proportionality between the logarithm of seismic moment and magnitude. On this basis, L^{*qc} is proportional to seismic moment, M_0, while L^{*q} is proportional to magnitude (note that an empirical calibration of these exponents requires conversion to the same logarithmic base; e.g., for c as the coefficient in a base 10 correlation of moment and magnitude, the corresponding exponents require adjustment by the factor $\log_{10}2$).

These definitions are used in the next section to illustrate contrasting relations of fractal dimensions defined in different ways to frequencies, lengths, moments, and magnitudes. The relation between fractal dimensions and information measures is considered subsequently in sections concerning questions of the completeness of seismic data for subregions of California. Two fractional exponents of a fundamental length scale appear to be needed to bring the frequency-magnitude distributions simultaneously into alignment with other data sets representing actual length measurements of faults and earthquake sources as well as with dimensional interpretations of moments derived from the analysis of earthquake spectra using seismometer records (the b-value of frequency-magnitude relations can be directly converted to a fractal dimension, hence relations between q, c, b, and D are implied). Successful correlations would bring the present approach simultaneously into alignment with geologic measurements, observational seismology, and with spectral scaling relations such as those studied by Chouet and Aki (1978).

Before going ahead with a study of the parameters q and c in relation to M_0, b, and D for given frequencies and magnitudes, another comparison between language and earthquake statistics is illustrated. This is the case of average word sizes and event magnitudes at constant frequencies shown in Figure 10. When word lengths using the same data base are averaged at

Figure 10 near here

constant frequency, the range of their average size, $\underline{m}$, is similar to the range of average earthquake magnitudes at constant frequency, $\underline{M}$, based on trends compiled by Algermissen and Perkins (1976) for the contiguous U.S. In this case a compression of word length (previously accomplished by redefining L^*_a) is caused by the fact that individual words and seismic events of greatly different sizes can have the same frequency of occurrence. The corresponding ranges of fractal dimensions are shown in Figure 11 for the language data in comparison with the range based on incremental word lengths (the trends and ranges of D for the averaged earthquake data would be similar to those of the averaged word data). Because L^* is defined in the same way

Figure 11 near here (note: where $\underline{m}$ and $\underline{M}$ appear, place bar above the letter rather than below; in this paragraph only)

for both data sets, base 10 is used in Figure 11, so the numerical values of L^* are the same as either the average word length or the average magnitude.

Figures 10 and 11 illustrate the gross statistical resemblance of frequency distributions for language and earthquakes. They are more instructive, however, in revealing how the averaging of frequency data, either in terms of average magnitudes or in terms of cumulative plots, can blur the infrastructure expressed in the form of fractal sets. The relationships discussed in the next section refer to the more explicit, but more restricted, fractal sets illustrated in Figure 9 using incremental earthquake data in California.

SINGULARITY SPECTRA FOR LANGUAGE AND EARTHQUAKES

The number of possible combinations of dimensional exponents such as q and c defined above, and the small size of the seismic catalog used to draw fractal inferences in Figures 8 through 11, make it difficult to capture the full range of these relations in a simple diagram. However, a more formal concept of singularity spectra developed by Halsey and others (1986) may help explain the nature of multifractal sets. The idea of describing patterns of singularities generated by a physical process in terms of fractal sets was stimulated by attractor theory (see Shaw, in press) combined with experimental research on fluid convection and turbulence that illustrated attractor-like patterns of motion. This approach has grown from the work of many international researchers, converging on methods described in Halsey and others (1986).

In order to cast the seismic problem in a similar format, I use an idealized relation between fractal dimensions and the b-values and c-values of frequency-magnitude, and moment-magnitude relations as discussed by Aki (1981); to do the same thing with the linguistic problem would require introduction of a generating function analogous to moment, such as the "cost" of a word. Shaw and Gartner (1986, Figure 18a) relate the idealized fractal dimension $D = 3b/c$ discussed by Aki to measured fractal dimensions of fault dendrites. If a given suite of seismic observations can be expressed in terms of a fractal length scale of the sort already discussed together with both the seismically idealized and a fault-related set of fractal dimensions, then it should be possible to relate L^*_s to coordinates that have one-to-one correspondences with measurements of seismic moment and with geographic length scales.

To keep the different bases straight, I refer to the Aki (1981) type of fractal as D_s relative to length L_s, and the fault-related fractal as D_f relative to length L_f. Therefore, an important eventual goal is to find the transformation that will bring the sets (D^*_s, L^*_s), (D_s, L_s), and (D_f, L_f) into alignment. The fractal dimensions are formally derived from the variations of a specified length-set (in the present context related to each other by adjustable exponents) as follows: (1) $D^*_s = D$ based on graphs of L^*_s for specified M and q, (2) $D_s = D$ based on L_s intrinsic to the relation $D = 3b/c$, and (3) $D_f = D$ based on graphical plots using fault length L_f. Present information does not permit unambiguous correlations for the different dimensional cases, but if one should be found it would facilitate the mapping of seismic parameters onto appropriate geographic reference frames. The result would be a step toward global predictions of

event frequencies on fault structures within reference domains of the sorts shown in Figures 3 and 14.

The question of how D^*_s may relate to D_s can be illustrated using the constructions shown in Figure 12. These in turn can be related to D_f and

Figure 12 near here

L_f based on fault measurements as discussed by Shaw and Gartner (1986, Figure 18a) for compatible values of b and c. But first consider some of the options for relations defined by the family of functions $L^{*qc} = Base^{cM}$. Taking logarithms, and using the moment vs. magnitude equation $\log M_o = d + cM$, relations between L^*_s, magnitude and moment are defined as follows:

$$\log L^{*qc} = cM \qquad (6a)$$

$$\log L^{*qc} = \log M_o - d \qquad (6b)$$

The range of c-values are found from plots of $\log M_o$ vs. M on a specified basis, as discussed by Hanks and Boore (1984, Figure 2). I assume here that both q and c are bounded by the limits 0 and 3, and that average values of c are typically near 1.5 for intermediate magnitudes (see Hanks and Kanamori, 1979). In Figure 9 the value q = 1 was used for earthquakes and the value q = 3 for renormalization of L^*_a for English to give comparable ranges of fractal dimensions. If it were assumed that the topologic limit should be the same as that used in Aki's (1981) definition of $D_s = 3b/c$, then M_o is proportional to L_s^3. Also, if L_s and L^*_s were assumed to be the same, then we would have qc = 3 in Eq. 6b (i.e., M_o proportional to L^{*3}_s). However, if qc is constant, then q decreases as c increases, and vice versa, implying that there are continuously varying proportionalities among q, c, L^*_s, M, and M_o.

This may be valid as a general tendency even though it contradicts the idea that there is a fixed relation between M_o and L_s. If c exceeds 1.5 for large events, then $\log L^*_s$ approaches M (i.e., q = 1 for limit qc = 3 at c = 3). Inversely, for small events (c = 1, or less) $\log L^*_s$ decreases with M/q; for example, at c = 1, q = 3 if qc = 3, and in that case L^*_s is compressed by the same factor as the L^*_a scale in Figure 9. This fits with observations that magnitude is roughly proportional to length for long strike slip faults, while magnitudes may be larger for small faults than would be expected from length correlations for long faults (see Mark and Bonilla, 1977).

Another special case that illustrates a different dynamical limit is given by choosing qc = 1 as the constant value. Then, we have the proportionalities q = 1/c and $\log L^*_s = cM$. For this length scale M_o is a function of the first power rather than the cube of L^*_s. Such a correlation would be consistent with the definition of moment as the product of a force and the distance from an axis of rotation. To be dynamically valid, earthquakes would have to represent the action of simple torques about stationary axes having point-like distributions (D_f close to zero).

These and other specific comparisons show that no constant value of the product, qc, will simultaneously satisfy correlations of $\log L^*_s$ vs. M and(or) M_o as well as correlations of M_o vs. L_s or L_f, and(or) M vs. $\log L_f$. Special cases would have to be invented for each of the different geometric correlations and(or) interpretations of M and(or) M_o that exist in the literature. Some combinations of q and c that span a range of possible

empirical correlations are shown in Table 1 for the typical value c = 1.5.

Table 1 near here

In Figure 12, D_s represents values of b and c that are assumed to be characteristic of some range of frequency-magnitude data. If local b-values were estimated for each range, and the appropriate c-values were assumed to be known, then sets of D_s-values could be plotted and compared with those inferred from plots like Figure 9. Such plots are not shown because the derivation of D_s, outlined below, presupposes a fixed relation between D_s and L_s. Forcing the sets of D^*_s-values to correspond to D_s by manipulating q and c is contradictory in that D^*_s is not normalized to any particular topologic limit. Some idea of how the values of L^*_s given by Eqs. 6 and Table 1 are related to Figure 12a is indicated by the q and c scales marked on the abscissa.

The main purpose of Figure 12 is to show that within the constraint of a fixed topologic relation between L_s and M_o used to define D_s, typical paths shown by the zigzag lines and arrows in Figure 12b are multifractal, and that for special paths such as A, B, and C there are continuous spectra of singularities analogous to those found in other systems of complex fractal geometries (in this respect, note that both q and c are variable dimensional exponents hence are of the same character as fractals; the term partial fractals may be appropriate in a manner parallel to the definition of the f-alpha functions used to characterize multifractal variations in Halsey and others, 1986). These paths are shown as geometric idealizations rather than as realistic examples; more realistic paths would start near b = 1 at low magnitudes. As indicated by the different paths in Figure 12a connected by the dotted lines with arrows, and by the zigzag path in Figure 12b, singular arrays of D_s-, b-, and c-values are typical of observed frequency-magnitude distributions. In this respect conclusions based on D_s-L_s relations are qualitiatively the same as those based on L^*_s derived more directly from frequency-magnitude data.

Empirically M_o is often expressed by a function that is roughly proportional to L_f^2; see Shaw and Gartner (1986) and Wesnousky and others (1983). Since such relations span uncertain ranges of fault lengths, magnitudes, and c-values, the empirical proportions are imprecise and reflect a mixed hierarchy of length scales possibly interrelated by fractional exponents. In an approximate sense, the discussion by Shaw and Gartner (1986, Figure 17) showed that values of D_f for measured fault dendrites are about one dimensional unit smaller than fractal dimensions based on $D_s = 3b/c$. Because the former are relative to the limit $D_T = 2$ (map plane), while the latter are relative to $D_T = 3$, the corresponding fractal length scales are roughly compatible (see later discussion of Figure 19). Thus, in some cases it is found that L^*_s and L_f are consistent, and in others that L_s and L_f are consistent. Therefore, L^*_s and L_s should also be consistent over similar parameter ranges.

Consideration of how the relation $D_s = 3b/c$ was derived is useful here because it identifies the connection between L_s and a specific fault segmentation model described by Aki (1981). For example, he states (p. 571) that the slope in a plot of log N vs. log L is 3b/c, where "b is the slope in a log frequency-magnitude relation" (usually, b is written as a positive coefficient, so the slope would be -b as defined by Eq. 3). This definition is consistent with graphical derivations of D-values using Mandelbrot's formula D = (1 - Slope) in plots of log NL vs. log L, as was done in Figures 8 and 9. This is easily verified by inserting numerical values of b and c in the preceding log N vs. log L relation; see graphical resume in Figure 13.

Figure 13 near here

But Aki also states that D_s represents a segmentation model given by the formula D = log N / log (1/r), where N is the number of parts and r is the ratio of subsegment lengths to a given fault length scale (see Mandelbrot, 1977, p. 43). Accordingly D_s = 3b/c = log N / log (1/r). Since the quantitiy 3b/c evidently also represents the ratio -log N / log L for the previous relation normalized to N = 1 for L = 1, then the consistent segmentation is 1/r = 1/L, or the total fault length is made up of N contiguous subsegments of length r. This just means that the whole length is the sum of the partial lengths, in contrast with dissection sets where N is decreased relative to 1/r by cutting out some fraction of the parts (a distribution that attenuates in total length as the number of parts increases).

If no parts are removed, or numerically discounted, we would expect that D_s = 1 is a minimum because the continuity of an original assumed fault is not broken (the definition of a connected fractal). Thus if b can decrease below 0.5, c would have to decrease below 1.5 in this segmentation model (e.g., if D_s = 1 is the limit, then c varies as 3b). If so, low b-values would correlate with small event magnitudes according to a c-value correlation such as that of Hanks and Boore (1984). Since this does not seem to be generally true, it is likely that the segmentation model used to define the relation D_s = 3b/c is not appropriate for earthquakes on systems of small faults.

If the above segmentation model were generalized to dendritic systems it would imply hierarchies of subsegment lengths proportioned according to characteristic branch lengths for each length order. In this respect the derivation resembles the model used by Shaw and Gartner (1986) to illustrate the relation between fault dendrite models and D_f expressed relative to values of b and c. To this extent fractal dimensions for connected fractals based on either L_s or L_f, respectively, are reasonably consistent. That is, dendrite-type and barrier-type models can be statistically similar with regard to number-length hierarchies. At the same time, however, they can differ importantly as mentioned earlier with regard to implications concerning how the fractal sets are distributed geographically.

The extent to which L^*_s and L_s agree or disagree would appear to depend on the fractal correspondences between length scales for moment distributions. The dotted oscillatory path in Figure 12a reflects the alternations of b- and c-values in Figure 12b discussed by Shaw and Gartner (1986) in terms of correlations between seismic data and faulting data in Japan (in general, c-values increase with increasing moment). This is the same sort of oscillating effect found in Figures 8 and 9, where the maximum range alternates between about D = 0 and D = 3. Apparently this can only happen for D_s in Figure 12a at very low or very high c-values (D_s jumps back and forth between the abscissa and Path C; hypothetically these variations are assumed to extend beyond the limit D_s = 1 discussed above). Therefore the most oscillatory patterns would seem to be typical of either high or low moment systems. Note that when c approaches the limit 3, log L^*_s becomes proportional to M, because q = 1 in the limit qc = 3.

The fractal range 0 to 3 can also be decribed as an oscillation between point-like and volume-filling sets, the extreme being for the "ultimate event" at the top of Figure 12b. Physically this would imply cycles of rapid exchange of strain energy from within dense dendritic arrays of areally distributed active faults to an expression that is point-like in character

(point-like in this case is meant in the same context as a hypocentral locus of a spatial domain). Generally speaking, such patterns imply that some sets of events in segmented and(or) dendritic fault hierarchies will be more predictable (high fractal D) than others (low fractal D). This observation relates to the fractal interpretation of von Seggern (1980) and to the remark by Aki (1981, p. 571) that fractal patterns will require a very complete areal scrutiny to be detectable by direct monitoring techniques. Aki also suggested that this may be one reason why methods of intensive monitoring used in China are relatively successful. An analogous conclusion was drawn by Shaw and Gartner (1986) with respect to the need to monitor the smaller branches of fault dendrites in order to detect trends in seismic energy distributions interpreted along the lines of Figures 1 through 3 here.

The difficulties of making correlations among different types of fractal sets would be mitigated if better correspondences between L^*_s, L_s, and L_f could be established. Partial agreement of subsets has been found above, but it is difficult to state any precise functional relations because of the model-dependent and topology-dependent aspects of the different sets. This is unavoidable at present because there are few data permitting a description of fault dimensions with depth in the crust. At present L^*_s and D^*_s would appear to be the best contextual representations of topologically three-dimensional seismic distributions that can potentially be made precise by calibration. Correlations with paleoseismic fault length and(or) fault slip fractal sets are in progress to extend these relations to times characteristic of seismic cycles (see Shaw and Gartner, 1986, Figure 8).

An example of a specific correlation that is simultaneously consistent with definitions of L^*_s, L_s, and L_f is given by Row (3) in Table 1. The predicted relations for $q = 4/3$, $c = 3/2$ can be tested against empirical correlations from Shaw and Gartner (1986) that are typical of others found in the seismological literature:

$$M = 1.235 + 1.243 \log L_f \qquad \text{(meters)} \qquad (7a)$$

$$\log M_o = 23.50 + 1.94 \log L_f \qquad \text{(kilometers)} \qquad (7b)$$

From Table 1, Row 3 the equivalent coefficients are 1.333 compared with 1.243 in Eq. 7a, and 2 compared with 1.94 in Eq. 7b. However, because the correlation refers to $c = 1.5$, it describes only one section through the multifractal space of Figure 12. Therefore it appears to satisfy several aspects of the average behavior of the earthquake process for intermediate magnitudes. I say average because we still do not have one-to-one comparisons of fault data and moment data that correlate directly with the subsets of multifractal trends in Figure 9. Even so, the implication is that detailed correspondences among D^*_s, D_s, and D_f may exist when magnitude and moment are related to L^*_s and when q and c are varied compatibly relative to median values such as $q = 4/3$, $c = 3/2$.

AN INTERPRETATION OF FRACTAL LINGUISTIC CYCLES AND OSCILLATIONS

In their discussion of the statistical heterogeneity of language, Miller and others (1958, p. 384) state that: "Clearly, the function words are the heavily overworked glue that holds our sentences together". The function words correspond here to one of the point-like fractal sets ($D = 0$) in the oscillatory multifractal fabric. By the same token it might be said that the fractal sets with $D = 0$ in Figure 9 are the glue that holds the patterns of seismic distributions together. To the extent that a correlation between

L^*_s and fault length exists, as discussed in the prceding section, the same conclusion can be drawn for those faults that represent the point-like sets in a dendritic hierarchy of faults within a given geographic domain. In that case the point-like sets symbolize both the glue (in terms of earthquake frequencies) and the skeletal framework that gives fault dendrites a characteristic seismic pattern of expression. If so, the point-like sets should manifest the relationships that exist between geography, fault distributions, and seismicity. In this regard we have come almost full circle to the models outlined in Figures 1 through 4.

In the remainder of the paper this implication is explored for the distributions of earthquakes within smaller regions and over shorter intervals of time. The corresponding uncertainty-redundancy relations of information measures are examined, and the incompleteness of the data is discussed from the multifractal viewpoint.

From the previous analysis it would seem to be possible to identify the relations between fractal dimensions and Shannon's uncertainty measure, because both can be derived from the same sets of data. Because the uncertainty (or entropy) is formally defined only in terms of probabiltiy distributions, however, it is less "informative" than is the concept of multifractal sets. Alternatively, the fractal variables can be viewed as measures of other sorts of entropy distributions in a manner analogous to the configurational entropies of molecular structures (see treatment of partition functions in Halsey and others, 1986).

Another way of stating the situation is to observe that if a given probability distribution can be expected to reflect some hierarchical ranking, then the informational uncertainty will give a good representation of that distribution. This is the basis for maximum entropy methods applied to frequency-magnitude analysis by Berrill and Davis (1980) and Main and Burton (1984). As Jaynes (1957) put it, the maximum entropy assumption is logically to be preferred in an interpretation of a probability distribution in the absence of reasons to think otherwise (i.e., lacking knowlege of any other distribution). The fractal sets discussed here constitute reasons to think otherwise when it comes to detailed interpretations. That is, the two approaches only reflect differences in the degree to which the data can be shown to be complete or incomplete in an appropriate geometric context. Therefore, in principle geometric fault models (e.g., Main and Burton, 1984, Figure 1) could be used to derive partition functions and singularity spectra to compare with data sets such as those of Figures 9 and 12.

The concept of a statistical glue has parallels with other phenomena. An example is given by the concept of gluons in particle physics, where the gluon is represented by photons "as a kind of glue" that holds protons and electrons together in the atomic structure (Pagels, 1982, p. 253). The role of the point-like sets here is analogous. This can be seen fairly directly in the distributions of event frequencies where the point-like intervals of Figure 9 keep resetting the trends of falling frequencies. These oscillations are related to what was earlier described as the transient teeter-totter effect relative to steady state trends of constant b-value. Now it is seen that the entire fabric of variations is transient (obviously the events themselves are) within a general trend that depends on how the fractal sets are sampled. This in turn relates to how the L^*_s fabric relates to fractal dimensions of mappable dendrites.

Thus, as in the atomic structure, there is an active exchange between the fractal structures of an energy space and the fractal structures of a tangible material space. Geographic predictability depends on how well the former can be mapped onto the latter. (Note: In view of the parallels between language

and the genetic code, it can also be predicted that this sort of statistical relation between highly uncertain and highly redundant sets will be found in the intron-exon structure of genes in biologic systems; see Gilbert, 1985. So the possibility demonstrated here that such codes can be described in terms of singularity spectra suggests that genes also represent a multifractal fabric in which the interweaving of intron and exon segments provides a structural framework that is glued together by point-like functional sets the sampling of which is of high uncertainty. In this light, the idea that ordered regions are the only ones important in the functioning of the genome is untenable. The remark applies equally to biology and seismology; in the latter case the "genome" is represented by the integral activity of a system of interacting fault dendrites analogous to Figure 3.).

To clarify, I tentatively assume that D = 0 corresponds to a maximum relative uncertainty, F, and D = 3 corresponds to a maximum redundancy, R, in the sense of Shannon's measures defined earlier. The oscillations in D, and the teeter-totter effect, therefore represent extremal variations between uncertainty and redundancy. Qualitatively, an area such as the Parkfield vicinity of the San Andreas fault is analogous to generally high D and redundancy (self-similarity suggests, however, that there will also be local structures with oscillating fractal dimensions within that set), while the so-called locked portions (vicinities of the 1857 and 1906 events) have relatively low D and are sampled with high uncertainty. This refers to the current epoch where the fractal patterns based on historic seismicity may approximately correspond to fault behavior.

A similar remark applies to seismic gaps in general. In the dendritic stream analogy, these can be likened to major limb or trunk portions of an intermittent flow system wherein the smaller dendrites record greater continuity of distributed events that only occasionally contribute to the major flow channels. Prediction of the incidences of flow along particular reaches of the trunk system similarly depends on how well the smaller tributaries can be monitored. By the same token. however, the seismic gaps and point-like distributions at all scales indicate the framework within which more localized monitoring should be deployed.

An example of this kind that has become conspicuous in recent studies is the White Mountain seismic gap described in Hill and others (1985). Here there is a pattern of distributed events that recur regularly combined with uncertain occurrences of large events on individual long faults. Other examples are discussed in Evernden and others (1978), Aki (1981), Kanamori (1981, 1986, Figure 12), and Anderson and others (1986).

Unfortunately there is at present an unavoidable ambiguity in discussing redundancies in space and time. This was pointed out earlier in the examples of patterns of neural receptors in contrast with patterns of perception. If seismic gaps are geographically fixed, the spatial description of characteristic earthquakes representing the activation of specific fault intervals is redundant (no uncertainty of location). Activations in time, however, may be highly uncertain. Analogously, the locations of articles, prepositions, etc. in sentence structures are highly determined, but their temporal predictability is uncertain (i.e., both types of roles correspond to the point-set "gluons" above). The idea that there may be an optimal redundancy of events in both space and time depends on the existence of coordinations among multifractal sets. This permits a "reading" of space-time relations among D-values that define an overall redundancy pattern in the same manner as happens in language.

In the following sections I explore how imperfectly the seismic distributions in California can be read in this context in the absence of

direct correlations with fault patterns. The results indicate that more intensive efforts should be made to map multifractal seismic patterns onto fractal fault patterns down to the scale of microearthquakes.

FREQUENCY-MAGNITUDE DATA FOR SUBREGIONS OF CALIFORNIA

An index map of California and its immediate surroundings is shown in Figure 14. It is subdivided into arbitrary subregions for the purpose of

Figure 14 near here

compiling subsets of earthquake frequencies. The only basis for this classification is that tabulated earthquakes are broadly grouped in a manner that fell within these geographic localities (ambiguities noted in tabulating the data were numerous events near the SF-C boundary, and many events near the CN-SS-OV boundaries in the vicinity of Mammoth Lakes, CA). The triangles are the corners of arbitrary linkages that generally straddle the geographic areas. These linkages are later used to illustrate migrations of historic earthquake activity according to a kinematic "bond-breaking" model. The arbitrary areas and benchmarks referred to in subsequent illustrations are identified by name and number in Figure 14.

Tectonically defined seismic zones are intentionally avoided in this paper in order to emphasize the statistical aspects of linguistic concepts. This does not reflect a belief that geologic insight is out-of-place. On the contrary, it is expected that any patterns that may be revealed in such an arbitrary context can be made far more explicit using detailed fault maps at expanded scales.

The earthquake history summarized in Figure 5 is broken down into eight subregions in Figure 15 (some of the 14 different subregions of Figure 14 have been combined). The rationale is to portray historic patterns representing

Figure 15 near here

all areas of concern from a hazards viewpoint at the same time as showing the general relations to areas where attention, both scientific and popular, has been focused on special events. Thus, the subregions E+SH, CN+TC+NS, and BJ+IP represent the three Plate-Tectonic "cornerstones" of the seismic distribution drawn schematically in Figures 2 and 3. These three areas resemble each other in average slopes of log f *vs.* M (post-1900 data sets). More interesting, however, is the observation that oscillations in these curves have suggestions of synchronization. For example there are frequency maxima near magnitudes 5.5 and 6 that seem to correlate over time spans as short as decades (see Figure 16 below).

At first sight it may be surprising that the most widely separated of the subregions should have the greatest behavioral resemblances. It is inferred, however, that these three areas behave in a coordinated manner because they reflect similar classes of fractal oscillations of the kinds illustrated in Figures 9 and 12. That is, they may reflect the average tectonic behavior of the continental margin more consistently than do interior regions of California. For example, the culturally important subregions SF (vicinity of the 1906 San Francisco earthquake) and C (vicinity of the 1857 central California earthquake) have some resemblance to these patterns when the entire interval 1800-1982 is considered, but the resemblance weakens at shorter time intervals. In the linguistic context of Figures 1 through 3, the cornerstone regions represent communication between plate tectonic stimuli and the

statistical responses of receptor regions subject to their influence (note that this implies a coordinating role rather than a causative one; the latter resides in the overall plate motions expressed in the interplay of all subregions of Figure 14).

These inferences are examined more closely in Figure 16, where the same data are plotted for intervals of 30 years. The three "cornerstone regions"

Figure 16 near here

are again consistently similar to each other during the 1926-1955, and 1956-1982 intervals. These intervals reflect the advent of instrumented seismic networks, whereas earlier data are too fragmentary to make such comparisons. It is interesting that in Figure 15 the post-1900 behavior of the LA subregion has been perhaps the most steady in the fractal context, although oscillations exist. This might fit with the highly diversified patterns of faulting in that vicinity and with the high fractal dimension of fault dendrites inferrred from length frequencies documented by Shaw and others (1981) and Shaw and Gartner (1986, Figure 5). It might be inferred that there is good temporal predictability of small to intermediate events and a low probability of large earthquakes in the immediate vicinity of Los Angeles (this is supported by the high internal consistency of fractal sets in Figure 19 below). However, high population densities extend into neighboring subregions capable of great earthquakes.

A different interpretation might be that LA represents a relatively long-lived deviation from a cyclic pattern of oscillating frequencies that has begun to reveal itself in the post-1926 behavior of other subregions. That is, following the rapid-fire occurrence of California's three greatest historically recorded earthquakes over a span of less than fifty years (1857, 1872, 1906), all subregions eventually began to reexperience more frequent events of small to intermediate magnitudes. In the fractal context, the extremes of oscillations became less severe and the overall seismic fractal dimension of the mid-20th century is higher than it was during the 1857-1906 period.

Descriptively, fault activations during the most recent several decades have been areally disperse (dendritically more widely distributed) as though the entire state experienced a pervasive strain relaxation. The same seismic gaps persist and give definition to earthquake distributions, but we have been in intermediate and relatively damped modes of oscillating paths such as those illustrated by Figure 12. If true, we should watch for a statewide decrease in the higher frequencies reflected in a return to more extreme oscillations in fractal paths as premonitory indications of conditions attending those of great earthquakes. Analysis of multifractal patterns for events smaller than M = 5 and in smaller subregions may be an effective way to develop such a long-range forecast. Patterns appear to change with sufficient rapidity on the scale of 30-year intervals that it is not possible to infer exactly where we are in such a cycle based on the statistics of the larger events.

The mid-20th century relaxation effect can be seen in the statewide patterns of short-term frequencies shown in Figure 17. Here the data for all

Figure 17 near here

subregions are combined but are shown at 5-year intervals. Examination of regional and temporal variations together suggest that, with exceptions mentioned below, the log f vs. M relation tends to flatten and elongate during times when seismic gaps are activated, and to steepen and contract at other

times. This is a manifestation of several effects already discussed in different contexts (teeter-totter swings about a steady state, oscillations in fractal paths, episodic reddening of frequency spectra, and concepts of seismic cycles; compare von Seggern, 1980; Main and Burton, 1984, in press; Aki, 1984; Shaw and Gartner, 1986).

An exception to this tendency is shown during the 1951-1955 period when the frequencies of the smaller events remained very high during the decade preceding and following the M = 7.7 Kern County earthquake. This pattern suggests that there may be at least two routes to the occurrences of large earthquakes in California: one is manifested by a period of waxing frequencies of small to intermediate magnitudes, the other by a waning period of small high-frequency events (a crude visual analogy is provided by the occurrences of breaking water waves during different regimes of the wave cycle). This effect can be seen better in the distribution of earthquake counts and moments at five-year intervals broken down according to subregion. These data are given in Table 2 and also are shown in the form of histograms in Figure 18.

Table 2 and Figure 18 near here

The 1951-1955 period is instructive with regard to local and statewide patterns. The cluster of 24 events in BF shown in Table 2A during this period represents a unique sequence both temporally and spatially (an idealized foreshock-aftershock sequence associated with the 1952 Kern County event). But notice also the cluster of 33 events in the CN subregion. Some of these occurred during the 1954 Fallon, NV earthquake sequence (two of the three large events shown for CN were part of this sequence), but many of the other events during 1954 were widely separated geographically. Thus, there was a major difference between the 1952 and 1954 flurries of counts in Table 2A in that the 1952 events were strongly localized whereas the 1954 events occurred throughout the state.

One interpretation of these effects (the frequency-magnitude variations in time and the clustering in space) is physically analogous to the charging of a leaky reservoir or capacitor. A general shift in earthquake counts and moments from northern to southern California with time is recorded in Table 2 accompanied by increasing frequencies of small to intermediate magnitudes and local flurries clustered around the occurrences of large events. That is, these events occurred during a period of pervasively increasing seismicity. By contrast, events such as the 1857, 1872, and 1906 earthquakes appear to reflect the discharge of stored seismic potential so that regional seismicity was suppressed (i.e., somehow this type of event is able to discharge the seismic moment within a very large crustal volume that otherwise would be expressed more uniformly by high frequencies of smaller events over large areas of the state).

The variations of b-values are distinctly different in the above modes; apparently, a change in statewide patterns toward either higher or lower b-value trends may be premonitory to a large earthquake. Distinctions, however, may be aided by the development of better evidence concerning the completeness of seismic catalogs in the fractal context. From the linguistic point-of-view illustrated in Figures 1 through 3, if the 1857 and 1906 events along the San Andreas are considered to represent a dominant theme, then the waxing seismic mode with local clustering of events associated with other fault sets is analogous to subthemes or topical asides during a conversation focused on the dominant theme (e.g., the making of a major point during a conversation is often accompanied by a pause during which the "impact soaks in"; by contrast, statements that are not quite as profound can occur without interrupting an increasingly animated exchange).

MULTIFRACTAL ARRAYS AND COMPLETENESS OF SEISMIC DATA

An improvement on the significance of b-value variations might be made in terms of plots of fractal dimensions such as that shown in Figure 9. It was shown above that q = 4/3 represents a scaling exponent that gives approximate correspondences among D^*_s, D_s, and D_f (fractal dimensions based respectively on seismically defined lengths from magnitudes, seismically defined D from b- and c-values, and geologically defined fault lengths); a fourth fractal set is designated D^*_f, for L^*_s recalibrated to L^*_f, a set dimensionally scaled to L_f (see Figure 19B below). For q = 4/3 as the scale factor, plots of $\log_{10}L^*_s$ vs. $\log_{10}(NL^*_s)$ for eight composite surbregions of Figure 15 are shown in Figure 19. Each of the diagrams has two

Figure 19 near here

or more dominant fractal trends: one is the zero-set common to all plots regardless of logarithmic base. Others represent subtrends up to a maximum that depends on the logarithmic base. In Figure 19B the maximum characteristic fractal dimension of the set is roughly $D^*_f = 3$. Formally this would correspond to D^*_s-values in Figure 19A of about $D^*_s = 10$ because of the factor $\log_{10}2$ (see Inset in Figure 19B that gives conversions at constant values of b). The fact that there are fractal sets in Figure 19A at approximate integer values of D^*_s, including 3, suggests that the arrays are crudely invariant under exponential transformations. Said differently, the fractal subtrends appear to represent a wide range of geometric progressions, possibly including values exceeding the topologic limit $D_T = 3$ (the latter correspond to a phase space with more than three degrees of freedom; Shaw and Gartner (1986) discuss time as a fourth dimension).

The elongations of the overall arrays in Figure 19 represent a sort of fractally averaged b-value that can be estimated from Figure 19B (Inset). The fact that this trend is poorly defined, except where it is dominated by large events, emphasizes the fragmentary character of the distributions. Sugregion BF+MJ in Figure 19B is instructive in this regard. The mean trend of small events, many of which were aftershocks of the large event (the 1952 Kern Co. event, M = 7.7), corresponds roughly to an average near b = 1 (note contrast with the zero-set corresponding to b = 0).

If one imagines that domains exist within which there are deformation states that correspond to occupancies on a lattice or grid of multifractal states, then the trends shown in Figure 19 express the completeness or incompleteness of seismic expression of these states. If the full range of possibilities were available in the sense of self-similar structures near the critical conditions for transition from periodic to chaotic behavior (e.g., in fluids this is represented by the transition from mode-locked quasiperiodic to chaotic turbulence) then a singularity spectrum analogous to Figure 12a would be expected (see Halsey and others, 1986). The presence of characteristic fractal sets, therefore, suggests that at the space-time scales of these data the patterns in Figure 19 may represent incomplete sets and(or) quasiperiodic mode-locking. If the latter, the locking may be transient and (or) a function of spatial resolution at M = 5 and above. The same plots could be made for shorter time scales but the data are too few; it is possible that extending these plots to low magnitudes may reveal continuous singularity spectra within more specifically delineated spatial domains.

Assuming that the patterns in Figure 19 represent partial filling of a fractal grid suggested by the respective trends, then these plots offer

additional insight concerning potential events that are missing from the sets. For example, in Figure 19A, SF and C are incomplete with respect to higher frequencies of small to intermediate magnitude events, whereas BJ+IP is relatively complete at the lower magnitudes and incomplete at the highest potential magnitude (i.e., an event near 7.5 or larger would be needed to fill in the lower right corner of the grid). Similarly, LA is "missing" an event at about M = 7. The diagram for LA is also instructive because the pre-1900 and post-1900 data are different with respect to frequency-magnitude patterns but show similar fractal subsets (the close weave of the LA plot relates to earlier remarks concerning the frequency-magnitude pattern in Figure 15).

The pattern of BF+MJ in Figure 19A is either incomplete in a manner similar to SF and C, or is a combination of two sets, as mentioned above, one being a low-magnitude set of high fractal dimension, the other being represented mainly by the largest event. For example, the more "complete" set may primarily represent the geometry of the aftershock sequence of the 1952 event. If so, BF+MJ is more like OV+SS, both representing the least completed fractal arrays in the overall set. According to this interpretation, these regions are conspicuous candidates for occurrences of intermediate to large earthquakes. The plots in Figure 19B have fractal grid-lines with nodal intersections shown at decade values of recurrence times ranging from 0.1 to 10^6 years. These plots can be used to estimate recurrence times for any assumed distribution of missing events.

Another interesting feature of Figure 19A is the fact that the higher values of D^*_s are roughly the same for E+SH, CN+TC+NS, BJ+IP, and SF, whereas they are smaller for C and LA, and larger for BF+MJ. The subset in BF+MJ is probably local, as discussed above, so there may be different singularity spectra for northern and central to southern California that are also functions of more local geographic scales. In this respect, SF is geometrically more similar to the "cornerstone" regions than are the other subregions of interior California (values of D^*_s other than zero for OV+SS are not determined by the historic data at magnitudes above M = 5).

A value of $D^*_f = 2.4$ in Figure 19B would be consistent with D_f estimated by Shaw and Gartner (1986) for young faults statewide if it were assumed that a reduction by one dimensional unit is appropriate to convert seismically derived geometries to the epicentral plane (i.e., equivalent to $D_T = 2$ as the topologic limit for map data). Some of the steeper partial trends may reflect such a correspondence, but most most of the D^*_f subsets have much lower fractal dimensions. This is expectable if the geometries of the fault sets are not sampled in a fractally self-similar manner at all scales. The rough equivalencies found by Shaw and Gartner (1986) for fault fractals and values based on $D_s = 3b/c$ are based on average b-values and are generally consistent with the conversions indicated in Figure 19B (Inset); that is, D_s is roughly 1.6 x D^*_f for the same b-value, so the equivalent average is $D^*_f = 0.9$ for $D_s = 1.4$. This would correspond to a slightly positive slope for a general regression line through an aggregate array of the data shown in Figure 19B (the elongations of the trends shown seem to be consistent with this comparison).

INFORMATION THEORY AND COMPLETENESS OF SEISMIC DATA IN CALIFORNIA

An approach to questions of completeness independent of the concept of multifractal arrays can be based on relationships between the variables b, c, and $D_s = 3b/c$ expressed in terms of measures of uncertainty and redundancy calculated using Eqs. 1 and 2. The latter reduce the frequency distributions that give rise to the patterns of Figures 9 and 19 to a single scale of

measurement related to generalized values of b and c. Whether or not these values are characteristic still depends on interpretations of multifractal states, but these states are expressed by a model-dependent definition, in this case by $D_s = 3b/c$, rather than by fractal sets. Relations between idealized b-values, fractal D_s-values, information U_{obs}, relative uncertainty F, and redundancy R are illustrated in Figure 20. The values are

Figure 20 near here

calculated from hypothetical log f vs. M plots that are assumed to represent constant b-values ranging from 0 to 1.5. Reference fractal dimensions are based on the constant value c = 1.5 for illustration (if both b and c were assumed to vary between 0 and 3, then the ranges of b, c, and D_s would have the same limits; see Figure 12).

Values of U_{obs}, F, and R are calculated from Eqs. 1 and 2 using frequency ranges normalized to unit probability at constant b and for equal magnitude intervals of 0.1. The equal-probability reference state in this case is based on a set of 33 intervals; i.e., n = 33, $U^* = \log_2(33)$. This is the state where the liklihood of an event M_i is indistinguishable from that of any other event M_j in this range regardless of magnitude. The relative uncertainty is $F = U_{obs}/U^*$, and redundancy is R = (1-F) expressed as a percentage.

The relation between D_s and R is shown in Figure 20c. It is not linear even in this ideal case. In general there can be a range of R-values for nonuniform frequency distributions of the same average b-value, so apparently there is the possibility of regimes of variation of D_s with little change in R, and vice versa. This is indicated schematically by the sigmoidal line. This underscores an earlier observation that the entropy concept applied to hierarchical sets of data is relative to a presumed global sequence identified by an ordering parameter such as the b-value. Hence where the constant b-value represents an average of diverse hierarchies the relation to definitions of entropy is nonunique. One can then deal with concepts of local entropy states, but this becomes essentially equivalent to dealing with concepts of multifractal structures expressed in terms of singularity spectra as discussed above.

The largest calculated ideal value of redundancy is about 41 percent, which is approaching the fifty percent level said to be characteristic of written English (see Shannon and Weaver, 1949). According to Shannon this is the property of English that permits the construction of two-dimensional crossword puzzles. He also states that three-dimensional puzzles are possible at about 33 percent redundancy (Shannon and Weaver, 1949, p. 57). This is so because as R decreases the linguistic constraints are less specialized, and there are more degrees of freedom in linking words together.

If one attempts to relate the constructional dimensions of crossword puzzles to the relations of Figure 20, according to the above statements, there is apparently an inverse relationship between them and fractal dimensions expressed by constant b-value hierarchies. This can be visualized by assuming that a column-and-row puzzle (i.e., a two-dimensional puzzle) taps a fractal information source set arbitrarily at a particular fractal dimension, say D = 3. This choice expresses an assumed maximum degree of freedom for a reservoir from which to choose suitable words. It is like taking a volume-filling information set (in the sense of word-size distributions analogous to the seismic case discussed earlier) and projecting combinations onto a square matrix in a manner that weaves them together.

This convention defines a dimensional set $R_i = 50$ (Shannon's conditional

statement), $D_i = 3$, and $D_p = 2$, where the subscript i refers to the information context as in Figure 20, and subscript p refers to the hypothetical dimension of crossword puzzles. By this reasoning, an area-filling information set would project onto a cubic matrix ($R_i = 33$, $D_i = 2$, $D_p = 3$). That is, if Shannon is right about lower redundancy and higher puzzle dimension, then the R vs. D relation of Figure 20 indicates that the redundancy and information fractal dimension decrease as the puzzle dimension increases. Consequently, linear and point-like information sets would require puzzle hypercubes tending to infinite dimensionality at the limit defined by the pair ($R_i = 0$, $D_i = 0$). In this limit any conceivable combination of words can be used, so the concept of a puzzle has been lost in the process; that is, the idea of fitting special words together has vanished along with the redundancy and fractal dimension.

If redundancy were to increase beyond $R_i = 50$, then the possibilities for choosing words would have to be ordered in more than three information dimensions (e.g., by specifying sequences in time and space). In puzzle space, combinations would have to be restricted to less-than-square arrays, to linear arrays ($D_p = 1$) and finally to a condition where no sequences are possible at all (e.g., at $R_i = 100$ percent there is only one word of unit probability, so there are no possible word combinations). According to Figure 20, such a case would have no physically defined information dimension, and the concept of a puzzle again becomes meaningless.

The dimensional transformation that occurs with puzzles resembles the introductory comments regarding the contrast between the redundancy of optical receptors as a spatial array and the redundancy of neural perceptivity of coded color patterns. It is also analogous to the contrast already discussed between the spatial redundancies of fault patterns vis a vis the redundancies of earthquake distributions. It is interesting that the realizable range of seismic fractal dimensions in Figure 20 corresponds roughly to the range of realizable crossword puzzle constructions (i.e., seismic puzzles are analogous to crossword puzzles constructed in a puzzle space somewhere between two and three dimensions). This reinforces an earlier conjecture that if statistical properties of seismic frequencies were to resemble those of language, then there is some optimism for believing that we could solve seismic puzzles as well as we can solve word puzzles. To do so, however, requires that the seismic "dictionary" is as complete as a language dictionary in identifying repertoires of related words.

The concept of a seismic dictionary is considerably more involved than that of a seimic catalog, because it implies knowledge of relationships between events (words) of different types. This knowledge is represented statistically by such things as multifractal sets and singularity spectra illustrated in Figures 9, 12, and 19. Therefore, evaluations should be referred to a global catalog, and data limited to an area such as California are only suggestive of the available repertoire. With this caveate, however, we can examine relations between observed redundancy structures relative to the previously illustrated fractal structures.

Calculations of uncertainties and redundancies were made for the observed seismic frequencies in the eight subregions of Figure 14 as well as in the total California data set using the above procedures (i.e., Eqs. 1 and 2 were applied to the data as described below). Results are listed in Table 3 for different time intervals on three assumptions, where the corresponding redundancies are labeled R_1, R_2, and R_3. In each case the

Table 3 near here

frequency data are normalized to unit probability. Column (1) uses the observed frequencies of Figures 5 and 15 and refers F_1 and R_1 to the reference uncertainty $U^* = \log_2(n)$ for n = 33 as in Figure 20 (n = 33 means that the allowed number of magnitude symbols is limited to 33, in this case starting with M = 5 and ending at M = 8.2 at intervals of 0.1). Column (2) uses the observed frequencies, but F_2 and R_2 are based on U^* calculated from the actual number of different magnitude values in the record (at the magnitude resolution of 0.1) rather than basing them on the maximum allowed number n = 33. This is done to illustrate the effect of gaps in the record; for example, there are only two values of magnitude shown for BJ+IP and one for BF+MJ in Figure 15 during the pre-1900 time interval. Gaps reflect either the incompleteness of a continuous record, or conceivably that some magnitudes do not occur for structural reasons (e.g., if all earthquakes were of the maximum moment type and there was a limited repertoire of fault lengths for characteristic earthquakes, then the spectrum of magnitudes would also be limited). In Column (2), therefore, values of R_2 represent a variable equal-probability limit where n can range from 1 to 33. Column (3) assumes that each of the observed magnitudes has the same probability and calculates F_3 and R_3 relative to the maximum n = 33 as in R_1. This represents a situation where magnitude resolution is assumed to be good but where the earthquake counts are so poorly known that the incidence of any magnitude value can not be distinguished from that of any other (the simplistic extreme is where ignorance is so great in a region that it isn't known whether the next earthquake is likely to be a big one or a little one).

Relationships between the R-values of Table 3 are functions of the incompleteness of the seismic catalog whether that occurs from sampling deficiencies or from events that are structurally disallowed. Such alternatives indicate some of the problems with assumed reference states and maximum entropy models. It might be thought that Column (2) is the best estimate of redundancy because it is normalized to the actual magnitude values recorded. This could be true if a steady state were represented by clustering of events at magnitude intervals greater than 0.1. However, a spuriously high redundancy will be recorded if there are few magnitude values and the data gaps are not characteristic. In that case the equal-probability redundancy R_3 can approach R_1 because both are referred to n = 33 and the mean logarithmic probability is relatively small on either basis.

If clustering of magnitude intervals is sufficiently weak that there are no gaps at a level of 0.1 resolution in M for a sufficiently long record, then Table 3 provides some criteria by which to evaluate completeness. In that case, $F_3 = 1$ and $R_3 = 0$, so $(R_1-R_3) = R_1$. By the same token R_1 would be equal to R_2, so $(R_1-R_2) = 0$. In other words, in a complete catalog judged on this basis, Columns (3) and (4) would all be zeroes, and Column (5) would be the same as Column (1). This agrees with intuitive interpretation of Figures 5 and 15 in that the 1900-1982, and 1800-1982 intervals for the total data set in California (Region "ALL" in Table 3) come closest to satisfying these conditions. The 1979-1982 data set including magnitudes down to M = 4 is of approximately the same completeness by these criteria, whereas the 1975-1979 set is far from complete including the lower magnitudes. This tends to confirm earlier remarks that short-term fluctuations in b-values are important in the interpretation of teeter-totter oscillations about a steady state. It also suggests that the average redundancy for the more complete data sets of Table 3 are in the neighborhood of 25 to 30 percent on a basis equivalent to Figure 20 (roughly the

three-dimensional puzzle limit discussed above). This corresponds to mean values somewhat smaller than b = 1 and D_s = 2, respectively, in Figure 20. The high values of other entries in Columns (3) and (4), and low values in Column (5) in Table 3, however, indicate strong departures from the same steady-state assumption.

Incompleteness apparently reflects a lack of sufficient information to adequately evaluate details of the possible oscillatory ranges of D_s and b as illustrated for the composite data in Figures 9, 12, and 19. Therefore, predictability depends on such information in addition to the average values of parameters such as b, D_s, and R. This suggests, as concluded earlier, that the details of such multifractal variations may be more revealing when they are extended to lower magnitudes. In lieu of more extensive data sets, concepts of completeness can also be examined in terms of assumptions concerning such things as asperity and barrier counts. An idealized model is outlined in the next section to suggest how this approach to calculations of redundancies might be brought into alignment with the fractal data of Figures 9, 12, and 19.

REDUNDANCIES OF AN IDEALIZED BOND-COUNTING MODEL

Number frequencies based on the same data as the temporal frequencies can be counted differently by assuming that a given magnitude represents a multiple of a constant exponential unit of magnitude, which I will call unit bond strength. For example, if M = 5 is arbitrarily chosen to represent unit bond strength, then M = 6 is made up of ten units, and so on. This is equivalent to saying that a set of events represents the partitioning of these hypothetical "bonds", each defined as having the same strength (e.g., if seismic energy were represented as a function directly proportional to the exponential of magnitude, then these hypothetical bonds would correspond by definition to equal energy quanta in the rupture process regardless of how the energy function related to geometric and mechanical properties). This is also equivalent to saying that the unit bond is proportional to a standard unit of length L^*_o (taken literally, a given rupture event would represent a multiple of L^*_o, each length increment making an equal contribution to the seismic energy).

Relative to M = 5 as the reference state, one event per 0.1 magnitude interval corresponds to a partitioning of from 1 to 1585 equal-energy "bond-ruptures" between M = 5 and M = 8.2. The bond-count can be written in the same form as the Gutenberg-Richter relation, as follows:

$$\log N_B = A_c + B_c M \qquad (8)$$

where N_B is the number of bonds, B_c is the (positive) coefficient of proportionality, and A_c is a constant (the + sign is because N_B increases with M). In the idealized case B_c is equivalent to the Gutenberg-Richter b-value with the conversion $B_c = (1-b)$. For an arbitrary point of normalization at log f = -3.2 at M = 8.2, the relation of N_B and f is given by $\log N_B = \log f + 3.2$ (see Figure 21 below).

This basis inverts the probability distribution of bond-counts relative to that defined in terms of earthquake events without otherwise changing the form of the distribution. Physically, a large event represents many bond-counts, while a small event represents few. The rationale for such a convention can be made in terms of energy partitioning. For example, if the earth's crust were viewed as a network of bonds of approximately the same strengths, and these bonds were being broken individually and in contiguous groups at certain

intervals of time, we would want to know the numbers of breaks within stipulated volume domains both at short times and averaged over long times. It might also be assumed that for reasons of structural configurations (bond angles, etc), bond breaking is clustered in some regions and diffusely distributed in others. If this behavior were characteristic, the bond-counts per volume domain would also be characteristic in the same sense as are events in recurrence curves. However, the accounting of transient behavior would have a different probabilistic implication in that a single large event that normally represents a low event probability now represents a very large bond-count probability.

This idealized model resembles asperity and barrier concepts, but it relates more closely to the kinetic fracture model of Termonia and Meakin (1986). There crack distributions are expressed in terms of fractal dimensional sets that resemble those of fault dendrites. This relation will be cited later in discussion of a kinematic linkage model of California earthquakes. In the present context a bond-count model illustrates the effect of a different choice of reference states on the calculation of informational redundancies.

Relations between log f and log N_B are summarized in Figure 21 for b-values ranging from 0 to 1.5. The probability distribution is

Figure 21 near here

identical for the curves of event counts at b = 1 (B_c = 0) and curves of bond counts at B_c = 1 (b = 0). Note that the relative changes of counts with changes in b-values are different than those in Figure 20, as are the definitions of equal-probability distributions. On either basis, equal probability refers to constant frequencies at each magnitude; e.g., the same numbers of events or of bond counts occur for each and every magnitude interval. Normalized to a single event at M = 8.2, log N_B = 3.2 and the bond-count equal-probability basis has 1585 bond ruptures at each magnitude between 5 and 8.2. For the bond-count model this occurs by definition at the value B_c = 0 (b = 1); the equal-probability event basis occurs by definition at b = 0 (B_c = 1).

Although the equal-probability definition is the same in each case, it is physically quite different. For example, if the equal-probability event basis is represented by one event at each magnitude, then there is one bond count at M = 5 increasing to 1585 bond counts at M = 8.2. As a consequence, at low b-values, or for disproportionate numbers of large events, the total bond-rupture count will be very large and highly biased toward the larger magnitudes compared to the total event count. In other words, bond-count probabilities for large events are greatly enhanced, and this will be reflected in the corresponding calculations of information parameters U, F, and R. In order to distinguish these conventions, I indicate the event basis by the symbols U_E, F_E, R_E, and the bond-count basis by U_B, F_B, and R_B.

The relations between R_B, b, and D_s, analogous to Figure 20c, are shown in Figure 22. Also, the data of Table 3 for California event frequencies by subregion and time interval are recalculated on the bond-count basis in Table 4. The main difference between Figures 20c and 22 is the fact

Figure 22 and Table 4 near here

that there is now an extremum in the relation between D_s and R_B for monotonically increasing b-values. The highest b-value still corresponds to

the fractal value $D_S = 3$, but the greatest redundancy, R_B, occurs at $b = 0$ (this is because $b = 0$ gives the largest range of bond counts, hence the steepest variation in probabilities). Maximum uncertainty ($R_B = 0$) occurs at $b = 1$, the bond-count equal-probability condition. This coincides with $D_S = 2$ for the definition $D_S = 3b/c$ and $c = 1.5$. The simplest physical interpretation is that the most uncertain bond-count distribution is the area-filling one. That is, $b = 1$, $D_S = 2$ is the maximum entropy reference state for the idealized bond-count model.

The data from Figure 22 and Table 4 are plotted in Figure 23 in terms of relative uncertainties $F_B = 1-R_B$. This diagram emphasizes the inversion

Figure 23 near here

of redundancies relative to event counts and provides an alternative way to express measures of completeness. For example, if $b = 1$ happened to represent the steady state, then the steady state would coincide with the bond-count maximum uncertainty. This is the only state in the bond-count basis that has a unique relation to fractal dimension, $D_S = 2$ (i.e., the fractal dimension is not single-valued at other values of F_B and R_B). At greater bond-count redundancies (i.e., for frequency distributions that give net uncertainties below the maximum of Figure 23) more than a single b-value and fractal dimension are required to describe the geometry of the set. This plot, then, is another way to illustrate earlier conclusions concerning the existence of oscillations between low and high fractal dimensions within a set of frequency data, even though the set as a whole may be described by a regression relation of constant b-value (alternatively, Figure 23 can be viewed as a modified form of singularity spectrum analogous to Figure 12a).

A novel feature of Figure 23 is that the concept of completeness is expressed in a manner analogous to the filling of an "uncertainty reservoir". There is no requirement that a complete steady-state distribution must correspond to the maximum at $b = 1$, but it is evident that the more complete data sets for California earthquakes are closest to the maximum-entropy bond-count state at $F_B = 1$. It is emphasized, however, that the position of the maximum is an artifact of the convention $D_S = 3b/c$ as well as of the bond-count model. Obviously other singularity curves would be possible for redefinitions of the fault segmentation model of Aki (1981) and(or) redefinitions of the bond-strength model. This might be a useful way to "tune" the relationships among F_B, D_S, and b relative to constraints on fractal and topologic limits. For example, there are direct relationships between N_B, M, and the q-c relations in Figure 9 and Table 1 which can be used to test self-consistent bond-count distributions. However, these relations are not explored farther in the present paper. Instead Figure 23 is used primarily to summarize the relative completeness of different data sets.

It has already been observed that frequency distributions for the long-term data at magnitudes greater than 5 are similar to the short-term data for magnitudes greater than 4 from 1979 to 1982. This might suggest that a fractal oscillation in the neigborhood of b-values between 0.5 and 1.5 (the limiting range in Figure 23) is characteristic of the more complete data sets in California over differing time and magnitude ranges. The corresponding range of D_S is from about 1 to 3; these ranges of fractal and b-value variations are also similar to those found by Shaw and Gartner (1986) for comparisons of paleoseismicity in the U.S. and Japan.

These comparisons of fractal and b-value ranges in Figure 23 and Table 4 for California illustrate that most of the subregions fall far below the range of bond-count uncertainties characteristic of the more complete data sets,

implying variations of fractal dimensions exceeding the self-consistent range of the idealized model. This is another way of characterizing the transient frequency oscillations of the teeter-totter type sketched in Figure 4 and discussed in terms of frequency-magnitude diagrams. In the bond-count context, extremely low b-values and high R_B alternate with extremely high b-values and high R_B, rising and falling relative to the maximum entropy state ($R_B = 0$, $b = 1$); this is the fulcrum of the teeter-totter effect, represented here by the extremum of Figure 22. In agreement with earlier observations, the effect is greatest in Figure 23 for those subregions and time intervals associated with the largest earthquakes.

A SPACE-TIME FLOW MODEL OF SEISMICITY IN CALIFORNIA

The discussion of frequency distributions is concluded by examining the geographic variations in time of the events considered here, recognizing that a similar treatment is needed for the far larger data base at magnitudes below M = 5. To illustrate patterns, a set of linkage maps is constructed that represents the data at intervals of 5 years relative to the continuity of tielines between the arbitrary benchmarks in Figure 14. If an event of M = 5 or larger occurs in the subregion between two benchmarks, then that imaginary link is considered to be "broken". In this way a kinematic model of linkage breaks is constructed, as shown in Figure 24. Obviously, if smaller

Figure 24 near here

magnitudes were considered at the same map scale, all of the imaginary linkages would be broken and the maps would be blank. Therefore there is a relation between the geometric scale of the subregions, event magnitudes, and the patterns of mappable progressions of intermittent disappearances and reappearances of linkages. Although there are no dynamic parameters in this portrayal, if size scales are adjusted appropriately, then such maps show patterns that resemble the evolution of bond ruptures in experiments and dynamic simulations.

This idea resembles the bond-count model of the previous section, but it does not discriminate linkages according to strengths greater than a threshold value equivalent to M = 5 (i.e., any event in the data set occurring between contiguous benchmarks is considered a break). In order to bring the two concepts together with regard to geographic distributions of bond counts, the maps would have to be constructed according to a hierarchy of increasing threshold strengths. Such a construction should be informative as to kinematic motions of linkage breaks if it were based on a finer geograhic grid using data for events smaller than M = 5. For example, sets of linkage maps could be constructed for all breaks at magnitude intervals of 0.5 starting from M = 2 or lower. As the magnitude threshold increases, the map patterns of linkage breaks would converge toward the loci of the larger events. The patterns described below offer heuristic examples of this approach.

A computer simulation based on a kinetic fracture model with similar kinematic implications is given by Termonia and Meakin (1986). This study is of interest because it measures the fractal dimensions of the areas of coalescing ruptures in a way comparable to the fault dendrite fractal measurements of Shaw and Gartner (1986). The patterns of Figure 24 are too crude to permit such measurements, but the idea is consistent with dimensional scaling over different magnitude ranges. Given patterns at different scales a fractal linkage dimension could be evaluated in an analogous way. This sort of measurement would provide a direct relationship between geographic fractal

dimensions and the seismic fractal sets illustrated here. Combined with measurements of fractal dimensions of fault dendrites at comparable map scales, data would then be available to calibrate relationships between fractal sets defined on several different bases.

For example, if the patterns of Figure 24 were assumed to resemble those shown by Termonia and Meakin (1986), then the fractal dimension of active faults in California would be similar to their measurements. They give a fractal dimension of about 1.3 for such an array relative to the topologic limit $D_T = 2$. This is the approximate value estimated by Shaw and Gartner (1986) for the overall pattern of young faults in California. Thus there may be a reasonably close correspondence between ranges of fractal measurements based on seismic data, linkage models, and fault dendrites.

The space-time behavior of California seismicity shown in Figure 24 may be of interest from the forecasting viewpoint. In addition to the post-1926 "loosening up" of frequency-magnitude distributions previously noted (Figures 15 through 18) there is a suggestion that linkage breaks in central portions of California evolved by encroachment from the "cornerstone" subregions of Figures 2 and 3. In several instances the activation of events in coastal regions appears to have occurred by coalescence of propagating deformations that emanated from subregions E and BJ (representing loci of Plate Tectonic activity off the northwestern coast and in the Gulf of California). This behavior describes precursive histories of the 1857 and 1906 San Andreas breaks. Analogously, the 1872 Owens Valley event responded to coalescence of linkage breaks that emanated from E and CN (the offshore California and northwest Nevada cornerstones).

Given these suggestions of systematic evolution, future occurrences of similar events might be anticipated by watching the progressions of linkage patterns defined in terms of subregions an order of magnitude smaller and based on events greater than, say, magnitude 3 or 4. Current patterns of behavior described by Hill and others (1985) in the vicinity of the White Mountains, California may be an example of this effect converging toward the locus of a major event analogous to the 1872 Owens Valley event (see discussion of Figure 19).

Patterns of migrating activity are consistent with inferences concerning dendritic interactions over large geographic domains. They add weight to the implication that monitoring activities need to be expanded in range and simultaneously at reduced length scales to include smaller subregions than those considered here. Such a basis is required not only to evaluate the time-dependence of b-values and fractal dimensions discussed above, but also to permit the monitoring of deformation waves and seismic subcycles that wax and wane over time scales measured in decades.

A schematic summary of space-time evolution is summarized in Figure 25. The temporal continuity of each linkage is indicated in Figure 25a for the

Figure 25 near here

24 arbitrary links defined in Figure 14. The paragenetic sequence of broken lines is shown according to ascending number but could be rearranged according to different geographic scenarios using the information in Figure 24. Tests of alternative sequences, however, did not change the general patterns of convergences, so the numerical sequence in Figure 14 was retained (to correlate these sequences more specifically with geography they can be cross-referenced with Figure 14). Domains of persistently broken linkages are connected in Figure 25b; events greater than about M = 7 are also shown on this diagram by date and general position in the sequence.

The relative positions of continuously interrupted linkages and large events in Figure 25b suggest that these events tend to occur in the vicinities of space-time convergences of linkage-breaking events. This is another way to view patterns of transiently migrating deformations suggested by Figure 24. Figure 25c gives a schematic generalization of this idea, where the arrows show hypothetical trajectories of flow of seismic moment. Localities shown at the right margin represent subregions where future activity would be expected on the basis of such flow patterns (Note: At the time of this writing several additional events have occurred in the vicinites marked "Los Angeles-San Diego", "South Central", and "E. Sierra-Nevada"; the White Mountain seismic gap is within the latter vicinity. No events have occurred in north central California at M = 5 and above, but there has been low-level activity in the region between the northern Sierras and Eureka, and Mount Lassen has been experiencing long-period seismicity of types seen in active volcanic terrains).

From the perspective of fractal sets and information theory the kinematic patterns in Figure 25c can be interpreted according to concepts of convective motions in fluids. As demonstrated by the work of Halsey and others (1986) the concept of singularity spectra and fractal interpretations of Figures 9, 12, 19, and 23 can be related to thermodynamic as well as informational concepts of entropy production. According to the discussion given earlier here, the point-like sets which represent seismic gaps, and so on, act in an information sense something like a structural glue (analogous to gluons in particle physics and function words in language). This view is generally consistent with the patterns in Figure 25c. The large events identify the structural "backbone" of the multifractal sets, in accord with the high redundancies associated with such events in the bond-count model of Figure 23. At the same time these events represent both sources and sinks of energy dissipation and information flow according to ideas of phasic reentrant signaling discussed early in the paper (Edelman, 1978). They are information sources in the sense that they are loci of high uncertainty on the event basis of Table 3, Column (2) accompanied by low values of D and b; i. e., they feed information to subsequent states of neighboring regions. And they are information sinks in the sense that they occupy centroids of vicinities where previously generated information is flowing in the directions of the arrows in Figure 25c; i.e., they are loci of focused dissipation and high bond-count probabilities that correlate with low uncertainty values in Figure 23 and high bond-count redundancies in Table 4.

Large events in the space-time plot resemble whirlpools in a river that contains numerous islets (domains of unbroken linkages). Such structures are spatially redundant in that they are invariant sites in the channel geometry and represent high concentrations of vortical kinetic energy (analogous to high concentrations of bond counts and energy dissipation); in this respect they may also have high localized fractal dimensions, analogous to the right side of redundant distributions in Figure 23. At the same time, however, the distribution of whirlpool sites can also be of low fractal dimension and high informational uncertainty from the point-of-view of the overall geometry of flow in the river within which they are imbedded. These properties of simultaneous redundancy and uncertainty, depending on viewpoint, appear to be common to language, turbulent fluid motion, and the earthquake process.

DISCUSSION: CONJECTURES ON CORRELATIONS WITH GLOBAL SEISMIC MOMENT AND WITH GLOBAL MAGMA TRANSPORT

Given the suggestions of coordination over large geographic areas in

western North America, it is of importance to consider relations to phenomena on the global scale. Such correlations might be approached in terms of the concept of propagating deformation waves considered in the light of time variations of global moment. In examining Kanamori's (1978) compilation of the world's great post-1900 earthquakes (those with event moments greater than 10^{28} dyne cm, and usually with magnitudes greater than M_S = 8) it is striking that 12 of the 79 events listed by him (his Figure 4) occurred between 1904 and 1907. Several of these were distributed along the western margin of North and South America. This temporal concentration of events is even more interesting in that it was globally localized within a narrow swath that involves many plate boundaries of different types. These coincidences tend to be forgotten or to be overshadowed by the fact that the peak of Kanamori's moment distribution occurred between 1950 and 1970. This was a time when there was only one event in California that came anywhere near the magnitude of the 1906 event (the 1952 Kern County event of M_S = 7.7 in southern California, subregion BF).

Seen in another light, however, the post-1926 strain relaxation effect discussed earlier suggests the possibility that the absence of large events at a time of peak global moment might be expectable. If the strain states in California during the period 1950-1970 had been what they were during the 1850-1900 period, then it seems very likely that the probability of great earthquakes would have been very high during the time of peak global moment. As it is we have had a reprieve from that eventuality until the state-wide frequency pattern has again passed through a general maximum and falls into the pattern of high event uncertainty and high bond-count redundancy that was conspicuous during the late nineteenth and early twentieth centuries.

In order to illustrate these contrasts, the moment rate based on 5-year averages for California is shown in Figure 26 compared with the global curve from Kanamori (1978). Allowing for some phase shift, there is a suggestion

Figure 26 near here

of correlation (also compare with the curve in Figure 26 showing estimates of eruptive volcanic energy from Kanamori, 1983, Figure 5). These relations support a conjecture that concepts of monitoring already discussed may be enhanced by combining them with data on global moment flow. A physical connection is provided by the idea that patterns of moment flow in the Circumpacific seismic zones are signaled by the cornerstone regions of Figures 2, 3, and 14. According to this interpretation the seismic responses of interior portions of California result from areally more dispersed processes of crust-mantle interactions than are described by models of simple slip along linear zones taken to define the plate margin.

Updates of the compilation of global moment are given in Kanamori (1983, 1986). There have been 18 post-1970 events with M_S of about 8 or higher which occurred in Central and South America and in the western Pacific (Japan, Phillipines, Indonesia). This does not include the M_S = 8.1 Michoacan, Mexico event of 19 Septhember, 1985 that devastated Mexico City (Anderson, and others, 1986). The average event rate since 1970 has been about one great earthquake per year. For comparison there were about three per year during the 1903-1907 interval and as high as two per year during other 5-year intervals. The rate during the mid-1980's appears to be on an upswing. It is suggested that a pattern of high global rates of great earthquakes concentrated along coordinated portions of contiguous Circumpaciic plate margins is more important to the incidence of major earthquakes in California than is the amplitude of the maximum value of global moment (e.g., as existed

during the 1950-1970 period). This conjecture is consistent with a scenario where deformation waves propagate into central portions of the state from the cornerstone regions, as illustrated in Figures 24 and 25.

There is an analogous suggestion of an upswing in volcanic energy subsequent to the events cited in the compilation by Kanamori (1983) shown in Figure 26. This is manifested, in addition to the activity at Mount St. Helens, WA beginning with the May 18, 1980 eruption, by evidence of magmatic intrusion beneath Long Valley, CA (see Hill and others, 1985), and evidence of increasing uplift rates at Yellowstone, WY (see Smith and Braile, 1984). It is also conspicuous that Kilauea, HI, has been in a pattern of higher transport and eruption rates since the beginning of 1983 than any previously recorded (see discussion of Figure 6 in Shaw, in press). Numerous other volcanic regions of the Pacific and Circumpacific apparently are also close to unstable eruptive states. Thus, many phenomena point to the present time as one of a high and possibly increasing rate of global magmatic and seismic activity.

Moments per five-year interval from Table 2 are broken down by vicinity in Figure 27 in three different ways: (1) In Figure 27a the total data set is divided into two parts of roughly comparable areas; the northern part is defined by subregions E, SH, SF, NS, TC, CN, and C, and the southern part by subregions SS, OV, BF, MJ, LA, and BJ; (2) Figure 27b shows a more localized two-part comparison influenced by the central to northern San Andreas and the Garlock-southern San Andreas systems, represented respectively by (SF+C) and (LA+BF); (3) In Figure 27c the total data set is subdivided into five major regions within which there is likely to be coupled tectonic activity.

Figure 27 near here

The comparisons in Figure 27 combined with the patterns of Figures 24 and 25 support the conclusion that there is a high degree of regional coordination of strain states governing seismicity within California. During the latest hundred years highs and lows of seismic moment are roughly in phase between northern and southern California, possibly with inversely related amplitudes. During the 1950-1955 period four of the five major regions of Figure 27c represent moment maxima, notably excluding the southernmost region, and the 1979-1982 period represents an upward trend in moments in all five regions. Moment fluctuations in all these regions have similarities during the 20th century.

An interesting pattern of relative fluctuations of global volcanic energy and moment rates in California is shown by the curves in Figure 26. Highs in eruptive energy correspond to lows in moment rates during the latest hundred years. Although the compilation of data used by Kanamri (1983) to document volcanic energy dissipation is not a complete description of volcanic rates, the inverse correlation may not be totally coincidental. Shaw (1980) pointed out that there are resemblances between volume rates of global magma transport and global seismic moments. Including the cornerstone regions E, TC+CN, and BJ+IP, California is subjected to a nearly pervasive potential for magmatic injections from the mantle. In some instances, such as the activity in Hawaii described by Shaw (1980) and in the Long Valley Caldera beginning in May, 1980 (coincidently just after the Mount St. Helens, WA eruption), a correlation of high magma transport rates and high seismic moment rates is expected. On the global basis, however, there will also be an effect analogous to earlier descriptions of deviations from long-term steady states (the teeter-totter effect illustrated schematically in Figure 4).

For example, if globally high magmatic rates correlate with globally high

seismic moment rates, and if magmatic rates fluctuate along the Circumpacific plate margins, then both the local seismic moment rate and the local magma supply rate to California and vicinity can lag the global rates. In this respect magma transport acts to signal the interaction of mantle and crust that governs the distribution of crustal deformation states. This is proposed as one reason why the cornerstone regions of the present report seem to behave similarly and to act as indicators of seismic activity interior to California. The next few decades represent a crucial period during which to monitor possible global balances of this type.

In conclusion, some combination of the following five general factors seems crucial to the incidence of large earthquakes in California: (1) There is a rapid increase in global moment over a time interval of a few years (an event in California could represent the onset of any global fluctuation, but typically it would be expected to occur somewhere during the interval of a global maximum). (2) There is a period of increasing frequencies of large events along contiguous portions of the western margins of the Americas, perhaps involving two or more per year somewhere along an arc extending from the Aleutians to southern South America. (3) There is a premonitory pattern of numerous "unbroken linkages" at a magnitude level of 5 and above prevalent across California in the context of Figures 24 and 25 persisting for at least a few years (this is analogous to the concept of seismic gaps generalized over the linkage network of Figure 14). (4) On the basis of event counts there is a rapid premonitory drop (see local exceptions noted below) in average b-values and fractal D_s-values combined with a high informational uncertainty within one or more of the key subregions identified in Figure 14; rapid means during a time of the order of a decade (event redundancies are likely to be below 10 percent for data collected over a few decades at $M = 5$ and above, or over about a decade at $M = 4$ and above; improved forecasting is contingent on similar studies of smaller regions at lower magnitudes). (5) Related to the factors in (4) there is a pattern of b-value and D_s-value windows oscillating between extremum values with increasing magnitudes, ultimately resulting in extreme redundancies calculated on a bond-count basis such as that shown in Figures 22 and 23.

According to several of these criteria California has been experiencing the ebb tide of a generalized seismic cycle during the mid-twentieth century. Accompanying this effect, however, has been a sort of following wave of increasing seismic moments and frequencies that have locally culminated in large events such as the 1952 Kern County earthquake (see discussion of Table 2). This can be viewed as a subcycle, or as a phase of the overall seismic cycle in which large events occur on the ascending limb of seismic moment rates rather than at a maximum moment where the moment rate tends to fall abruptly. These projections would suggest that an earthquake as large as an 1857, 1872, or 1906 event would not be expected before the turn of the century. A transition in these patterns, however, appears to be occurring about now. If so, the apparent rapidity with which each of the above phenomena can change in amplitude, or even in sign, makes such a forecast unreliable in the absence of better documentation of all the factors mentioned. There is some optimism that extending criteria of multifractal analysis and the space-time variations of fractal seismic sets to lower magnitudes and smaller subregions will provide a premonitory network the patterns of which will be diagnostic for events greater than about $M = 7$ within time windows of one or two years over intervals of two or more decades into the future.

REFERENCES

Aki, Keiiti, 1981, A probabilistic synthesis of precursory phenomena: in D.W. Simpson and P.G. Richards, eds. Earthquake Prediction Washington, DC, American Geophysical Union, p. 566-574.

Aki, Keiiti, 1984, Asperities, barriers, characteristic earthquakes and strong motion prediction: Journal of Geophysical Research, v. 89, p. 5867-5872 .

Algermissen, S.T. and Perkins, D.M., 1976, A Probabilistic Estimate of Maximum Acceleration in Rock in the Contiguous United States: U.S. Geological Survey, Open-File Report 76-416, p. 1-45.

Algermissen S.T., Perkins, D. M, Thenhaus, P. C., Hanson, S. L., and Bender, B. L., 1982, Probabilistic Estimates of Maximum Acceleration and Velocity in Rock in the Contiguous United States: U.S. Geological Survey, Open-File Report 82-1033, p. 1-99.

Anderson, J.G. and others, 1986, Strong ground motion from the Michoacan, Mexico,Earthquake: Science 233, p. 1043-1049.

Berrill, J.B. and Davis, R.O., 1980, Maximum entropy and the magnitude distribution: Seis. Soc. Am., Bull. 70, p. 1823-1831.

Boore, D.M., 1984, Use of seismoscope records to determine M_L and peak velocities: Seis. Soc. Am., Bull. 74, p. 315-324.

Campbell, Jeremy, Grammatical Man, 1982: N.Y., Simon and Schuster,

Caputo, M., 1982, On the reddening of the spectra of earthquake parameters: Earthquake Prediction Research. v. 1, p 173-181.

Chouet, Bernard, Aki, K., and Tsujiura, M., 1978, Regional variation of the scaling law of earthquake source spectra: Seis, Soc. Am., Bull. v. 68, p. 49-79.

Dewey, Godfrey, Relativ Frequency of English Speech Sounds 1923, Cambridge, MA, Harvard University Press. (Author's note: The spelling of "relativ" is not a typographical error; it is an aspect of Dewey's phonetic method.)

Edelman, G.M., 1978 Group selection and phasic reentrant signaling: A theory of higher brain function: in G.M. Edelman and V.B. Mountcastle, The Mindful Brain, MIT Press, Cambridge, p. 51-100.

Ellsworth, W.L., Lindh, A.G., Prescott, W.H., and Herd, D.G., 1981, The 1906 SanFrancisco earthquake and the seismic cycle, in D.W. Simpson and P.G.Richards, Eds.: Earthquake Prediction American Geophysical Union, Washington, DC, p. 126-140.

Erickson, R.P., 1984, On the neural bases of behavior: Am. Sci., 72, p. 233-241.

Evernden, J.F., Isacks, B.L., and Plafker, G. Eds., 1978,Proceedings of Conference VI. Methodology for Identifying Seismic Gaps and Soon-To-Break Gaps U.S. Geological Survey, Open-File Report 78-943.

Gatlin, L.L., 1972,Information Theory and the living System: N.Y., Columbia University Press.

Gilbert, Walter, 1985, Genes-in-pieces revisited: Science, v. 228, p. 823-824.

Halsey, T.C., Jensen, M.H., Kadanoff, L.P., Procaccia, I., and Shraiman, B.I., 1986, Fractal measures and their singularitiy: The characterization of strange sets: Phys. Rev. A, v. 33, p. 1141-1151.

Hanks, T.C. and Kanamori, H., 1979, A moment magnitude scale: Jour. Geophys. Res., v. 84, p. 2348-2350.

Hanks, T.C. and Boore, D.M., 1984, Moment-magnitude relations in theory and practice: Jour. Geophys. Res., v. 89, p. 6229-6235.

Hanley, M.L., 1951, Word Index To James Joyce's Ulysses: Univ. Wisconsin Press, Madison.

Herd, D.G., 1978, Intracontinental plate boundary east of Cape Mendocino, California: Geology v. 6, p. 721-725.

Hill, D.P., Wallace, R.E., and Cockerham, R.S., 1985, Review of evidence on the potential for major earthquakes and volcanism in the Long Valley-Mono Craters-White Mountains regions of eastern California: Earthq. Pred. Res. v. 3, p. 571-594.

Howard, K. A., Aaron, J. M., Brabb, E. E., and Brock, M. R., 1978, Preliminary map of young faults in the United States as a guide to possible fault activity: U.S. Geological Survey, Miscellaneous Field Studies Map MF-916.

Jaynes, E.T., 1957, Information theory and statistical mechanics: Phys. Rev. v. 106, p. 620-630.

Jennings, P.C. and Kanamori, H., 1979, Determinations of local magnitude, M_L, from seismoscope records: Seis. Soc. Am., Bull., v. 69, p. 1267-1288 .

Jensen, M.H., Kadanoff, L.P., Libchaber, A., Procaccia, I , and Stavans, J., 1985, Global universality at the onset of chaos: Results of a forced Rayleigh-Benard experiment: Phys. Rev. Letts. v. 55, p. 2798-2801.

Kanamori, H., 1978, Quantification of earthquakes: Nature v. 271, p. 411-414.

Kanamori, H., 1981, The nature of sesmicity patterns before large earthquakes, in D.W. Simpson and P.G. Richards, eds.: Earthquake Prediction, Washington, D.C., American Geophysical Union, p. 1-19.

Kanamori, H., 1983a, Magnitude scale and quantification of earthquakes: Tectonophys. v. 93, p. 185-199.

Kanamori, H., 1983b, Global seismicity, in Earthquakes: Observation, Theory and Interpretation, Bologna, Italy, Soc. Italiana di Fisica, p. 596-608.

Kanamori, H., 1986, Rupture process of subduction-zone earthquakes: Ann. Rev. Earth Planet. Sci. v. 14, p. 293-322.

Khinchin, A.I., 1957, Mathematical Foundations of Information Theory, Dover, N.Y.

Konheim, A.G., 1981, Cryptography, A Primer, N.Y., Wiley-Interscience.

Lindh, A.G., 1983, Preliminary Assessment of Long-term Probabilities for large Earthquakes along selected fault segments of the San Andreas Fault system in California, U.S. Geological Survey, Open-File Report 83-63.

I.G. Main and P.W. Burton, 1984,Physical links between crustal deformation, seismic moment and seismic hazard for regions of varying seismicity: Geophys. J. R. Astr. Soc. 79, p. 469-488.

Main, I.G. and Burton, P.W., in press, Long-term earthquake recurrence constrained by tectonic seismic moment release rates: Seis. Soc. Am., Bull., v. 76.

Mandelbrot, B.B., 1961, On the theory of word frequencies and on related Markovian models of discourse, in R. Jakobsen, ed., Structure of language in its Mathematical Aspect.: Proc. 12th Symp. in Appl. Math. Providence, R.I., American MAthematical Society, p.190-219.

Mandelbrot, B.B., 1965, Information theory and psycholinguistics, in B.B. Wolman and E. Nagel, Scientific Psychology, N.Y. Basic Books, p. 550-562.

Mandelbrot, B.B., 1977, Fractals: San Francisco, W.H. Freeman,.

Mandelbrot, B.B., the Fractal Geometry of Nature, 1982, San Francisco, W.H. Freeman,.

Mark, R.K. and Bonilla, M.G., 1977, Regression analysis of earthquake magnitude and surface fault length using the 1970 data of Bonilla and Buchana: U.S. Geological Survey, Open-File Report 77-614.

Miller, G.A. and Chomsky, N., 1963, Finitary models of language users, in R.D. Luce, R.R. Bush, and E. Galanter, Eds., Handbook of mathematical Pshchology, Wiley, N.Y. v. II, Ch. 13, p. 419-491.

Miller, G.A. and Newman, E.B., 1958, Tests of a statistical explanation of the rank-frequency relation for words in written English: Am. Jour. Psychol. v. 17, p. 209-218.

Miller, G.A., Newman, E.B., and Friedman, E.A., 1958, Length-frequency statistics for written English: Information and Control, v. 1, p 370-389.

Mogi, K., 1985, Earthquake Prediction: Academic Press, Orlando FL.

Okubo, P.G. and Aki, K., in press, Fractal geometry in the San Andreas fault system: Jour. Geophys. Res.

Pierce, J.R., 1980, An introduction to information theory: Dover, N.Y.

Real, C.R., Toppozada, T.R., and Parke, D.L., 1980, Earthquake epicenter map of California: California Div. Mines and Geol., Map Sheet 39.

Ryall, A., Slemmons, D.B., and Gedney, L.D., 1966, Seismicity, tectonism, and surface faulting in the western United States during historic time: Seis. Soc. Am., Bull. v. 56, p. 1105-1135.

Schwartz, D.P., and Coppersmith, K.J., 1984, Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones: Jour. Geophys. Res. v. 89, p. 5681-5698.

Shannon, C.E., and Weaver, W., 1949, The Mathematical Theory of Communication: University of Illinois Press, Urbana.

Shaw, H.R., 1970, Earth tides, global heat flow, and tectonics: Science, v. 168, p. 1084-1087.

Shaw, H.R., 1980, The fracture mechanisms of magma transport from the mantle to the surface: in R.B. Hargraves, ed. Physics of Magmatic Processes, Princeton, NJ, Princeton University Press, p. 201-264.

Shaw, H.R., 1985, Links between magma-tectonic rate balances, plutonism, and volcanism: Jour. Geophys. Res. v. 90, p. 11275-11288.

Shaw, H.R., 1987, Uniqueness of volcanic systems: U.S., Geological Survey, Prof. Paper 1350. p.

Shaw, H.R., Gartner, A.E., and Lusso, F., 1981, Statistical Data for Movements on young Faults of the United States: U.S. Geological Survey, Open-File Report 81-946.

Shaw, H.R. and Gartner, A.E., 1984, Empirical Laws of Order Among Rivers Faults and Earthquakes: U.S. Geological Survey, Open-File Report 84-356.

Shaw, H.R. and Gartner, A.E., 1986, On the graphical Interpretation of Paleoseismic Data: U.S. Geological Survey, Open-File Report 86-394.

Sherburne, R.W., Boylan, R.T., and Parke, D.L. , 1985, Seismicity of California, April, 1979 through October, 1982: California Geology, v. 38, (April), p. 75-80.

Sieh, K.E., 1981, A review of geological evidence for recurrence times of large earthquakes: in D.W. Simpson and P.G. Richards, Eds., Earthquake Prediction, American Geophysical Union, Washington, DC, p. 181-207.

Sieh, K.E., 1984, Lateral offsets and revised dates of large prehistoric earthquakes at Pallett Creek, southern California: Jour. Geophys. Res., v. 89, p. 7641-7670.

Sieh, K.E., and Jahns, R.H., 1984, Holocene activity of the San Andreas fault at Wallace Creek, California: Geol. Soc. Amer. Bull., v. 95, p. 883-896.

Smith, R.B., and Braile, L.W., 1984, Crustal structure and evolution of an explosive silicic volcanic system at Yellowstone National Park: in Explosive Volcanism: Inception, Evolution, and Hazards, Washington, DC, National Academy Press, p. 96-109.

Smith, R.L., 1979, Ash-flow magmatism: Geol. Soc. Am., Spec. Paper 180, p. 5-26.

Smith, R.L. and Shaw, H.R., 1979, Igneous-related geothermal systems: in D.E. White and D.L. Williams, eds., Assessment of Geothermal Resources of the United States-1975, U.S. Geological Survey, Circular 726 p. 58-83.

Smith, R.L. and Luedke, R.G., 1984, Potentially active volcanic lineaments and loci in western conterminous United States: in Explosive Volcanism: Inception, Evolution, and Hazards, Washington, DC, National Academy Press, p. 47-66.

Sperry, R.W., 1974, The growth of nerve circuits: in Readings from Scientific American, Cellular and Organismal Biology, W.H. Freeman, San Francisco, p. 121-128.

Suppe, J., Powell, C., and Berry, R., 1975, Regional topography, seismicity, Quaternary volcanism, and the present-day tectonics of the western United States: Am. Jour. Sci., v. 275-A, p. 397-436.

Termonia, Y. and Meakin, P., 1986, Formation of fractal cracks in a kinetic fracture model: Nature v. 320, p. 429-431.

Toppozada, T.R., Parke, D.L., and Higgins, C.T., 1978, Seismicity of California 1900-1931: California Division of Mines and Geology, Special Report 135, p 1-39.

Toppozada, T.R., Real, C.R., Bezore, S.P., and Parke, D.L., 1979a, Compliation of Pre-1900 California Earthquake History, California Div. Mines and Geol., Open-File Report OFR 79-6 SAC.

Toppozada, T.R., Real, C.R., and Pierzinski, D.C., 1975, Seismicity of California, January, 1975 through March, 1979: California Geology, v. 32, (July), p. 139-142.

Toppozada, T.R., Real, C.R., and Parke, D.L., 1986, Earthquake history of California: California Geology, v. 39, (February), p. 27-33.

Tribus, M., and McIrvine, E.C., 1971, Energy and information: Sci. Am. v. 225, p. 179-188

von Seggern, David, 1980, A random stress model for seismicity statistics and earthquake prediction: Geophys. Res. Letts., v. 7, p. 637-640

Wallace, R.E., 1973, Surface fracture patterns along the San Andreas fault: in R.L. Kovachs and A. Nur, Eds., Proceedings of the Conference on Tectonic Problems of the San Andreas System (Stanford, CA, Stanford University Publication, Geological Sciences v. XIII, p. 248-250.

Wallace, R.E., 1984, Patterns and timing of late Quaternary faulting in the Great Basin province and relation to some regional tectonic features: Jour. Geophys. Res., v. 89, p. 5763-5769.

Weldon, R.J. and Sieh, K.E., 1985, Holocene rate of slip and tentative recurrence interval for large earthquakes on the San Andreas fault, Cajon Pass, southern California: Geol. Soc. Amer. Bull., v. 96, p. 793-812 .

Wesnousky, S.G., Scholz, C.H., Shimazaki, K., and Matsuda, T., 1983, Earthquake frequency distribution and the mechanics of faulting: Jour. Geophys. Res., v. 88, p. 9331-9346.

Wesson, R.L. and Wallace, R.E., 1985, Predicting the next great earthquake in California: Sci. Amer., v. 252, 35-43.

Wilson, K.G., 1979, Problems in physics with many scales of length: Sci. Amer. v. 241, p.158-179.

FIGURE 1. Schematic illustration of overlapping probability domains analogous to repertoires of word probabilities in a conversation among three people, or event probabilities in three faulting regions that have similar but not identical dynamical properties. The tree-like dendritic patterns indicate the effect of interaction on the channeling of information toward themes of common interest in the case of language, or toward faults of mutual influence in the case of earthquakes. Probabilities are progressively enhanced in regions of overlap, so that where there is a three-way overlap all of the subregions contribute to the same theme, whether it be a subject of focused conversation or a subject of focused seismic activity. Prior to the onset of feedback interactions, the probabilities in each subregion were different and totally independent except for limitations given by the possible numbers of letters and words in language or of the possible numbers of seismic mechanisms in the earthquake process. In subsequent diagrams, Circle 1 corresponds to subregion E (coastal and offshore northern California near Eureka), Circle 2 to subregion BJ (the general area of southernmost Imperial Valley and Baja California, including the Gulf of California and areas of convergence with the East Pacific Rise), and Circle 3 to subregion CN (the composite area consisting of eastern and northeastern California and western and northwestern Nevada).

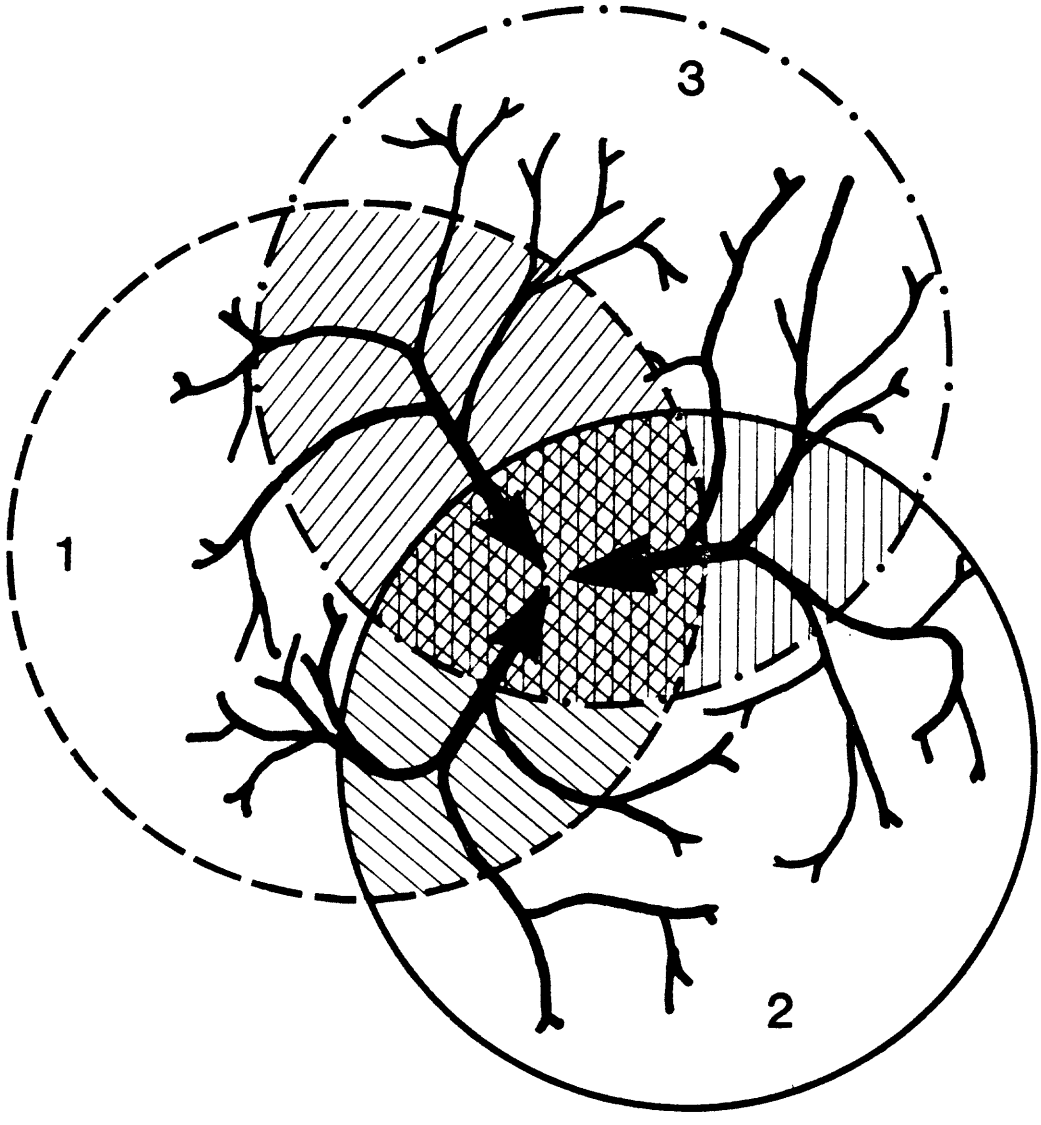

3
1
2

FIGURE 2. Schematic illustration of how the probability structure of Figure 1 might be mapped onto the tectonic framework of seismicity in California. There is still a three-way interaction of domains characterized by systems of deformationally interactive fault dendrites, but now the boundaries of the source subregions are open to other interactions at the global scale of Plate Tectonics; the probability logic is otherwise the same. Subregion E is essentially the Northern California Offshore Zone of Ryall and others (1966), CN corresponds roughly to their Ventura-Winnemucca Zone, while trends between E and BJ and between CN and BJ correspond to the converging areas between their Central and Southern California Zones (see discussion in text).

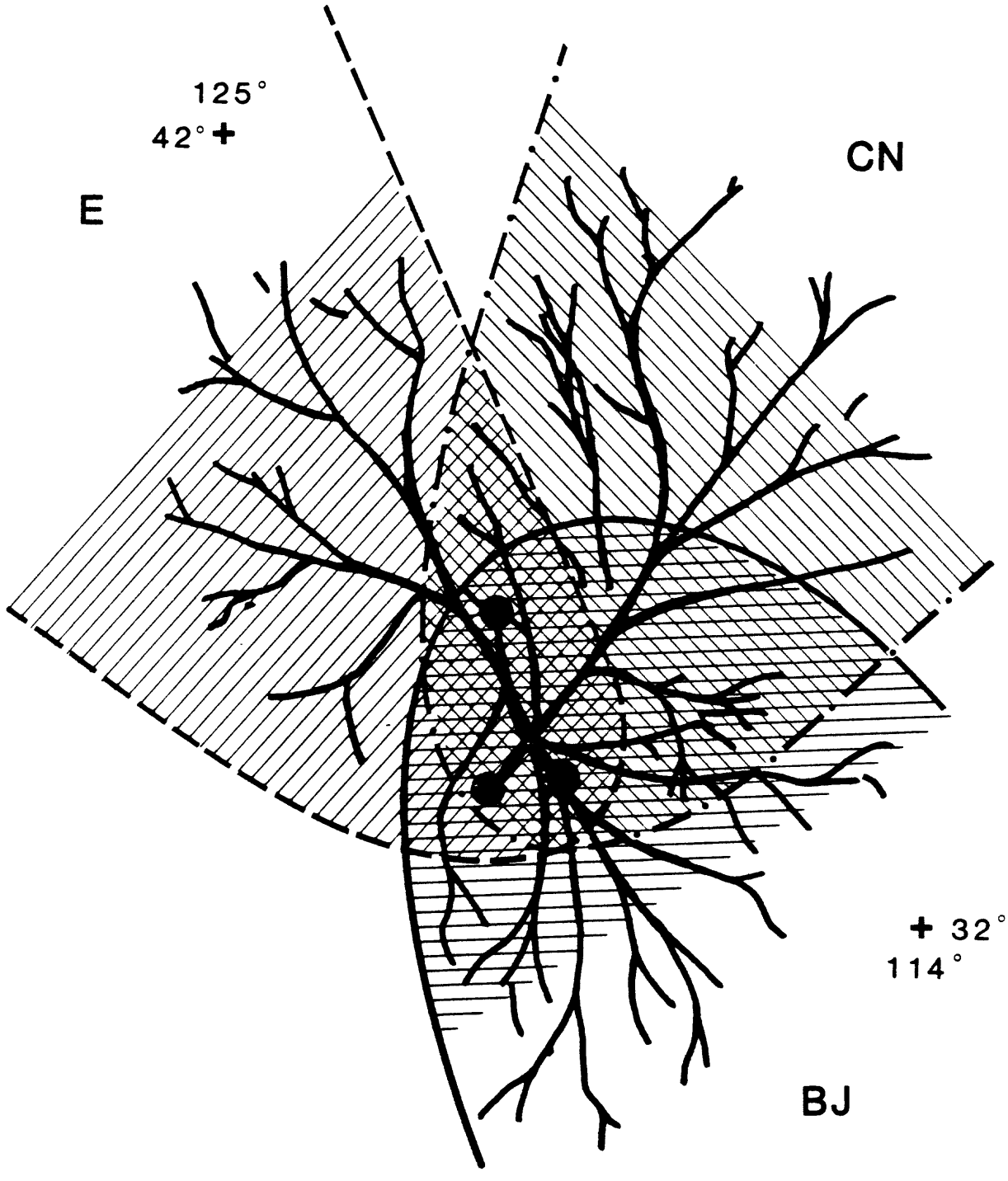

125°
42°
E
CN
32°
114°
BJ

FIGURE 3. A more specific version of Figure 2 showing the general distribution of major Quaternary faults traced from the index map of Toppozada and others (1978). Each of the subregions E, CN, and BJ is now shown as a fold-shaped or prong-shaped area, reflecting the complexity of interactions among long-lived features of crustal structure interacting with sources of mantle motions (for patterns of active faulting over time spans from the historic to roughly 15 Ma, see Howard and others, 1978; Shaw and others, 1981). Inset A: Historic fault breaks and epicenters, according to Ryall and others (1966, Figure 5). Inset B: Tectonic flux map evaluated by Ryall and others (1966, Figure 3) using earthquake data; map is modified to show areas inscribed within the lowest flux contour in their distribution. The inset map provides an independent seismic description of patterns that are assumed here to represent the convergence of three broadly interacting mantle-crust sources of deformation. See discussions in text, and Wallace (1984).

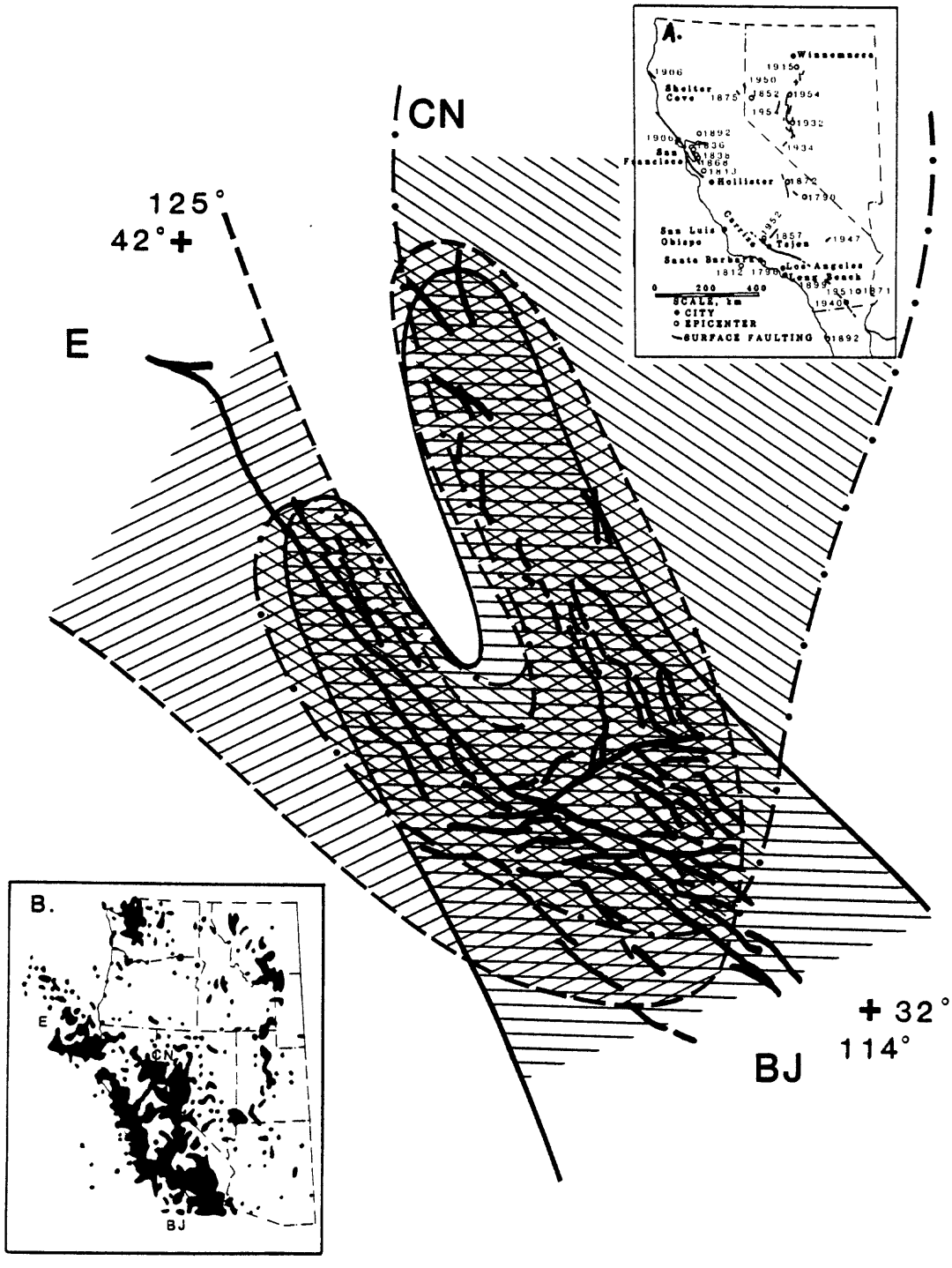

A.
Winnemucca
1915
1906
1950
Shelter Cove
1875
1852
1954
1954
1932
1892
1906
1836
1838
1868
1911
San Francisco
Hollister
1872
1906
1857
1952
San Luis Obispo
Tejon
1947
Santa Barbara
1812
1790
Los Angeles
Long Beach
1899
1951
1971
1940
0 200 400
SCALE, km
CITY
EPICENTER
SURFACE FAULTING
1892
CN
125°
42°
E
+ 32°
114°
BJ
B.
E
CN
BJ

FIGURE 4. Schematic frequency-magnitude diagrams illustrating relations between steady and transient behaviors of dendritic fault distributions having a variety of different bifurcation ratios (see discussion of length orders of branch sets in the text): (a) Transient excursions from an assumed steady state (heavy solid line) occur proportionally at a constant b-value (dotted lines), proportionally at b-values varying relative to a pivotal magnitude (heavy dashed lines), or occur disproportionally at variable b-values (combinations of the above). (b) Examples of dendrite types; shape and bifurcation ratio are theoretically independent variables, but the more compressed and elongated the shape the smaller are the lengths and numbers of the shorter branches (see exception in f below). (c) Families of independent fault dendrites having similar shapes and bifurcation ratios tend to have the same b-value; the frequency level is determined by the amount of deformational energy supplied to each one (analogous to variable discharge from stream dendrites of otherwise similar geometries). (d) Fault dendrites with dissimilar shapes and(or) bifurcation ratios tend to have highly dissimilar b-values averaged over the respective sets (any of the above can have subsets that resemble at a smaller scale the behavior of a different dendrite unless the dendritic structure is self-similar at all scales of sizes and rates of deformation). (e) Example of partial activations of subsets of a single dendrite having a relatively large bifurcation ratio. Depending on how the branches of different lengths, and subsegments of branches of different lengths, are sampled, the b-value oscillates between lower and higher values over different ranges of magnitudes which may or may not be characteristic. If characteristic, the oscillatory pattern represents a quasisteady state with constant average b-value and with generally constant ranges of alternating b-values (i.e., the zigzag line is essentially invariant over times that allow all subsets to be represented in the data; see inset of Figure 5, and examples illustrated by Shaw and others, 1986). (f) If the activation of subsegments of a single fault trace, whether continuously joined end-to-end as in a bamboo pole or in discontinuous intervals, influences increased rates of activation of longer and longer composite sections in a quasisteady manner, then the b-value can be significantly smaller than unity (i.e., the fractal dimension can be one or less and the b-value one or less; their relation depends on how the mechanisms of faulting are mapped onto the geometries of faulting). However, the same can be true for dendrites of compact form analogous to stream braiding; see text for discussion, and Shaw and others (1986).

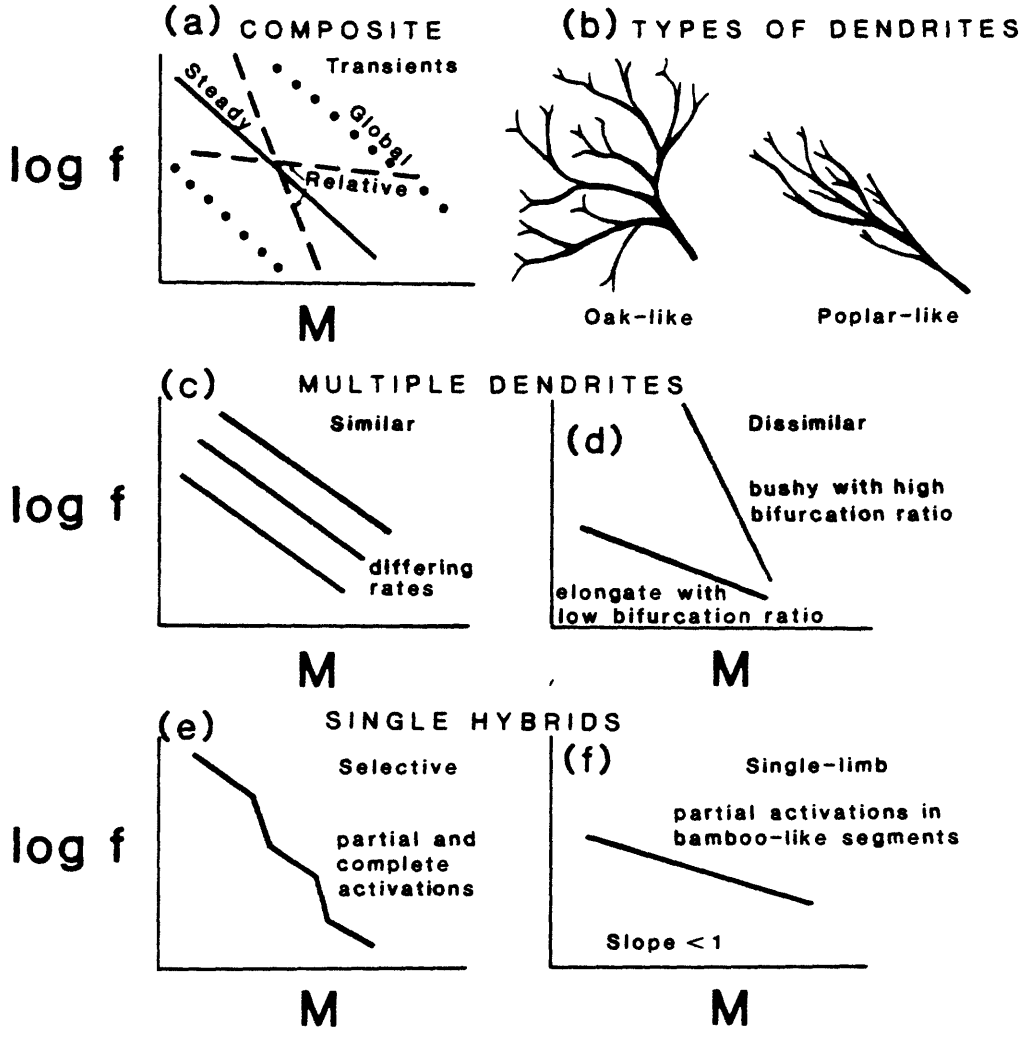

(a) COMPOSITE
Transients
Steady
Global
Relative
log f
M
(b) TYPES OF DENDRITES
Oak-like
Poplar-like
(c) MULTIPLE DENDRITES
Similar
differing rates
log f
M
(d)
Dissimilar
bushy with high bifurcation ratio
elongate with low bifurcation ratio
M
(e) SINGLE HYBRIDS
Selective
partial and complete activations
log f
M
(f)
Single-limb
partial activations in bamboo-like segments
Slope < 1
M

FIGURE 5. Incremental frequency-magnitude data for pre-1900 (crosses), post-1900 up to 1982 (dots), and total historic record 1800-1982 (open circles) of seismicity in California and immediate surroundings. The data represent all the earthquakes recorded in the following catalogs at magnitudes of 5 and above: Real and others (1978), Toppozada and others (1979a,b), Sherburne and others (1985). Toppozada and others (1986) give a list of earthquakes greater than magnitude 6 restricted to the immediate vicinity of California for events to 24 April, 1984; events more recent than 31 October, 1982 are not included in the figure. Inset: Cumulative recurrence curves for seismic zones in the western United States, modified from Ryall and others (1966, Figure 7). The regression lines are theirs (solid lines); the groupings of points, however, suggest a pattern where intervals of lower b-values are intermittently offset by steps or intervals of higher b-values. This effect is analogous to the more exaggerated oscillations seen in the incremental data; compare with zigzag patterns of frequency-magnitude data discussed by Shaw and Gartner (1986). Differences in patterns are explained in part by the different plotting bases (incremental or cumulative) and by regional effects and the methods of rate normalization used in the study of Ryall and others (1966, p. 1120 ff.). See later discussion of Figure 9 and seismically derived fractals in the text.

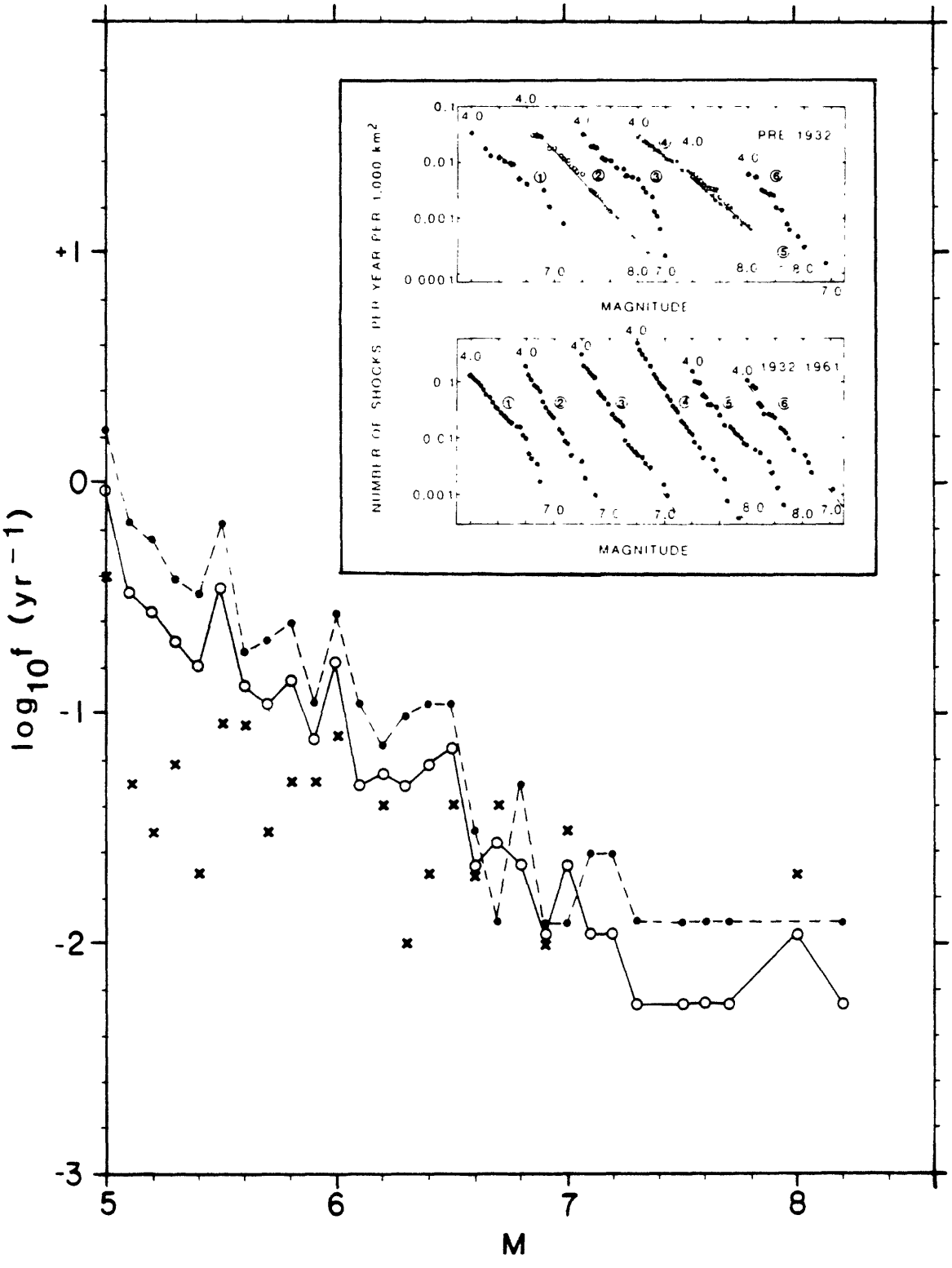

log10f (yr−1)
+1
0
-1
-2
-3
5
6
7
8
M
NUMBER OF SHOCKS PER YEAR PER 1,000 km²
0.1
0.01
0.001
0.0001
PRE 1932
1932 1961
MAGNITUDE
MAGNITUDE

FIGURE 6. The relation of symbol probabilities to their rank, comparing data for large samples of words in English (dots and crosses) with earthquake events from Figure 5 normalized to unit probability (open circles). For purposes of comparison each magnitude value at intervals of 0.1, as in Figure 5, constitutes a different word-like symbol and the frequency represents the relative numbers of times a given event symbol is used (i.e., an earthquake event is analogous to a word in a language statement involving variable rates of repetition of different words which also differ in "magnitude"; the most obvious measure of word magnitude is size measured in letters per word, but other criteria exist based on classifications of word types, etc.). Rank refers to the sequence of words or events arranged in order of decreasing probability (sequences at constant probability are included but relative rank is arbitrary in such intervals). The hyperbolic limit of direct inverse proportionality (solid line of slope -1) is often referred to as Zipf's law (see text). The trend of the earthquake data is roughly parallel to the language trend up to a rank of about 16 and then falls off abruptly. This sort of roll-off is a truncation effect that is also characteristic of samples of language limited to comparable total numbers of words. It is inferred that if an earthquake catalog were as extensive as these language catalogs, the probability-rank relations would be essentially indistinguishable (taken by itself this is not a strong test of comparable structural organization, as has been pointed out on several occasions; see Mandelbrot, 1961, 1965, and Miller and Chomsky, 1963). The dots were calculated from word counts in the listing of Hanley (1951) representing the 260,430 words in James Joyce's Ulysses, and the crosses were obtained similarly from the listing of words occurring 11 times or more in a 100,000-word sample of ordinary English given by Dewey (1923).

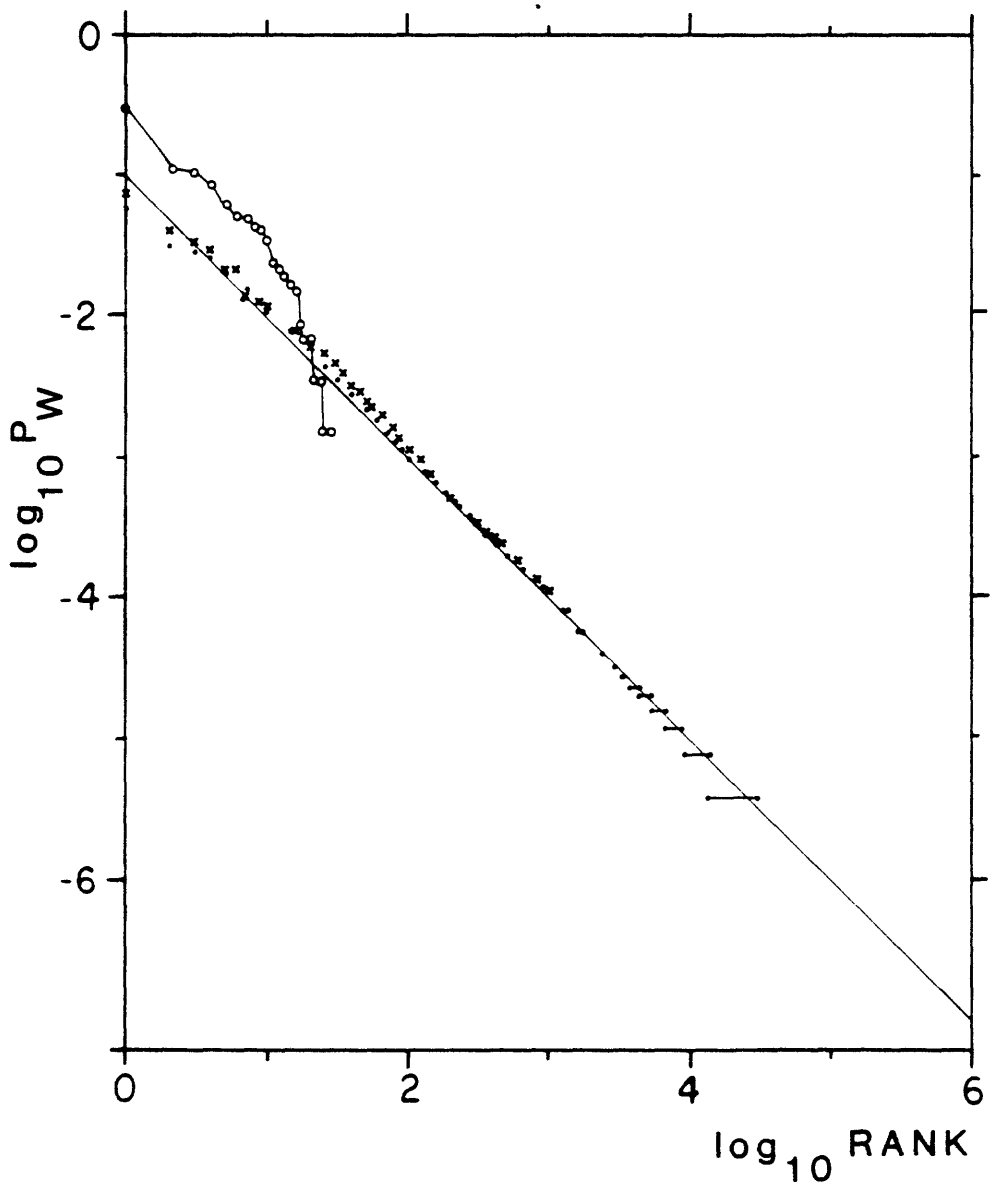

0
-2
-4
-6
0
2
4
6
log10 Pw
log10 RANK

FIGURE 7. Relation between symbol probability and size for word-symbols in English (circles and crosses), where size is expressed in letters per word; earthquake events are also plotted (dots), where the abscissa value of symbol size is taken to be equal to the magnitude (data of Figure 5). One set of data for English (circles) was obtained by making counts of the numbers of words at integer values of letters per word using the catalog of Hanley (1951); the other language set was taken directly from Miller and others (1958, Figure 2B). The similarity of these two data sets and the probability-rank relations of Figure 6 for three distinctly different samples of English usage indicates that statistical distinctions are more subtle than are demonstrated by gross probability-rank-size relations. These gross similarities are striking in view of the alleged semantic richness of the writing style of James Joyce (see Shannon and Weaver, 1949, p. 56) compared to the more ordinary forms of written English represented by the other data sets.

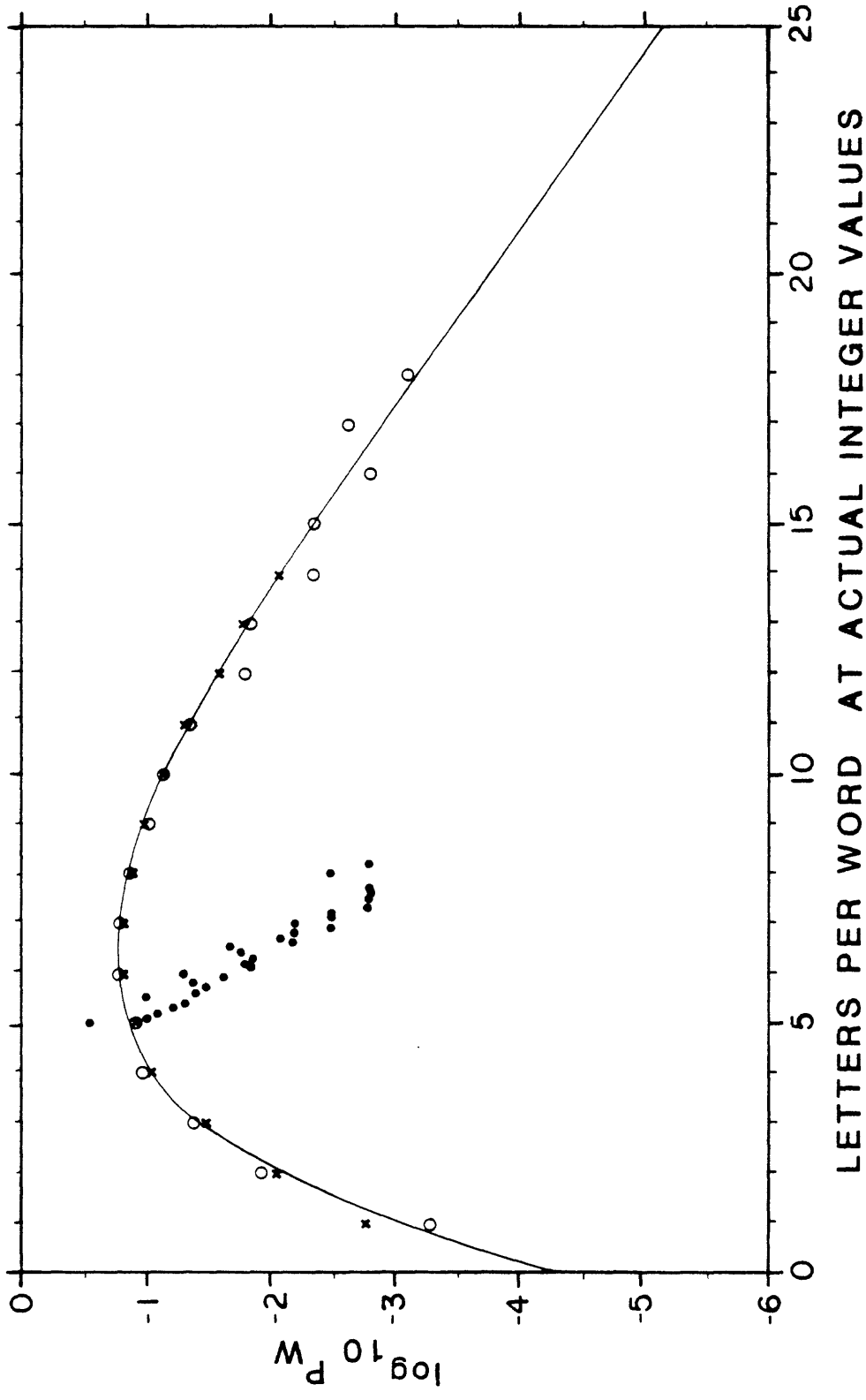

LETTERS PER WORD AT ACTUAL INTEGER VALUES
0
5
10
15
20
25
-6
-5
-4
-3
-2
-1
0
log10 Pw

FIGURE 8. Diagram illustrating basis for calculating fractional length scales (fractal dimensions) for language symbols based on an arbitrary convention for expressing symbol size in terms of a length dimension analogous to fault rupture lengths related to earthquake magnitudes (the data base is the Hanley set of Figure 7). For this purpose the length scale L* is defined according to the function $L^* = 2^m$, where m is the word size ("magnitude") in letters per word (the range is from 2 to 18 letters). Base two is used to emphasize the fact that log L* represents a potential family of length functions determined by a size or magnitude quantity; in this form, word magnitude measured in letters per word represents binary size units, or bits (a two-letter word is a two-bit word, etc.). If base ten were used, $\log_{10}L^*$ would be synonymous with letters per word. The ordinate gives total lengths expressed in two ways: (1) Circles represent incremental total lengths, where N is the number of symbols of length L* (this is the number of words of integer letters per word; N is normalized to a total vocabulary of 100,000 different words in this plot). (2) Crosses represent cumulative total lengths, where each point is the sum of NL* for all sets with lengths equal to or greater than L* in the sample. The latter is analogous to measuring the total length of all symbols using a yardstick of length L* applied to all symbols of equal or greater length (this convention is similar to methods of measuring coastline lengths using dividers set at varying lengths; see Mandelbrot, 1982; Shaw and Gartner, 1986). The slopes of curves in such diagrams are related to fractal dimension according to S = (1-D), where S is the slope and D is the fractal dimension (Mandelbrot, 1977, 1982). The incremental basis is limited only by the appropriate maximum topologic limits (typically 0 to 3), while the cumulative basis is necessarily limited to fractal dimensions between 1 and 2 because it is confined to the plane as in the coastline problem; see Figure 11 for comparison with data expressed in terms of mean word-length, and also for formal relations to a standard Cantor set expressed both in terms of the above construction and set theory.

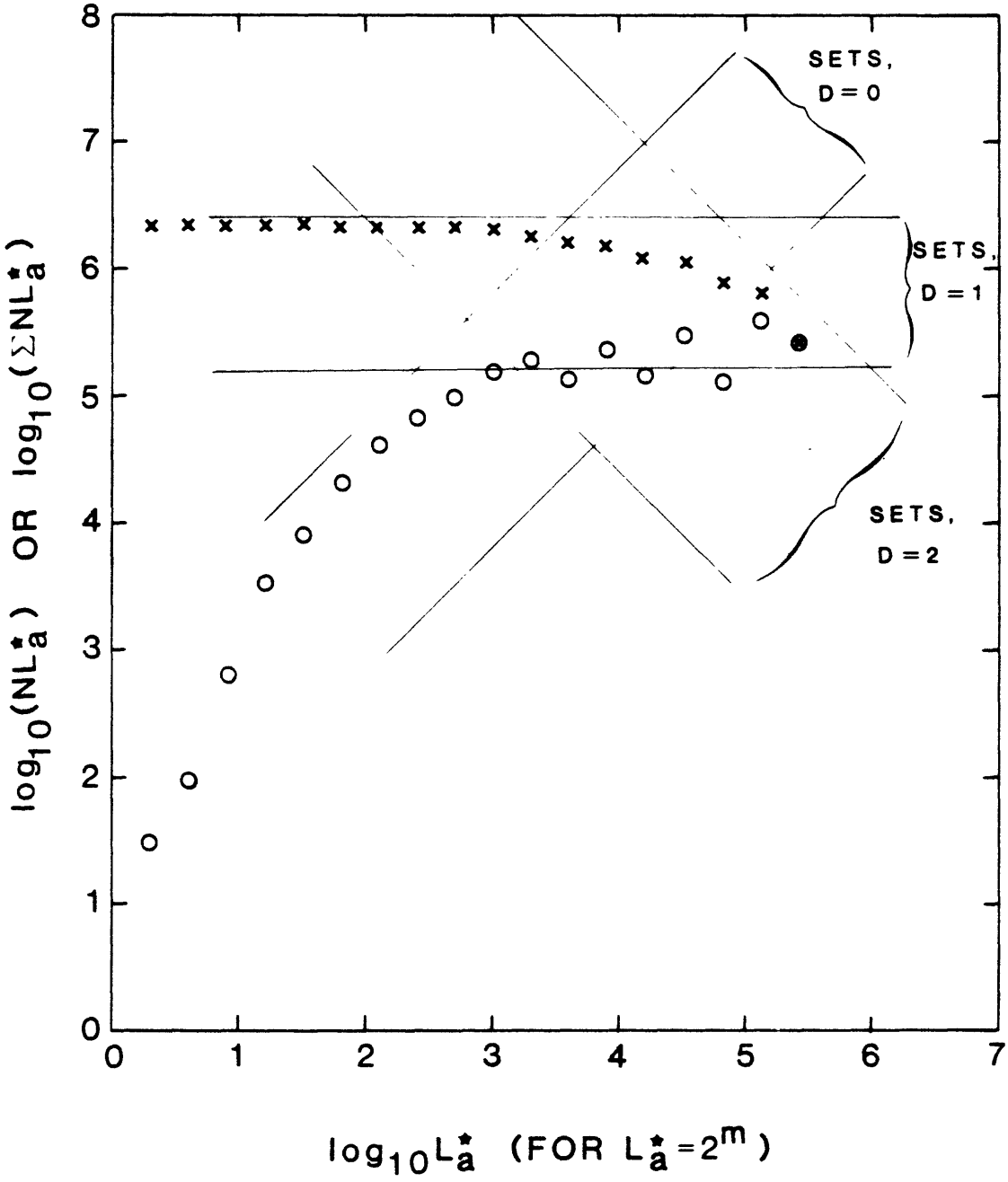

SETS,
D = 0
SETS,
D = 1
SETS,
D = 2
$\log_{10}(NL_a^*)$ OR $\log_{10}(\Sigma NL_a^*)$
$\log_{10} L_a^*$ (FOR $L_a^* = 2^m$)
0
1
2
3
4
5
6
7
8

FIGURE 9. Diagram comparing fractal constructions for an earthquake length scale $L^*_s = 2^M$ (open circles, incremental; crosses cumulative) and a revised language length scale $L^*_{a'} = 2^{m/3}$ (open squares, incremental; pluses, cumulative). Reference slopes and fractal D-values are otherwise constructed in the same way as in Figure 8; however, here the total number of different words is 29,899 based on the catalog of Hanley (1951). Except for multiplicative constants, and the roll-off at small word size reflecting the differing statistics of function and content words (see text), the seismic and language fractal subsets are similar (i.e., the seismic length scale is proportional to the cube root of the alphabetic length scale defined on the same logarithmic basis). Note that to compare the seismic set with empirical length data expressed to Base 10, length contributions are reduced by the factor $\log_{10}2$, resulting in a maximum average slope for the cumulative set equivalent to about D = 1; incremental sets may still have subtrends that define higher fractal dimensions (see comparisons in Figure 19).

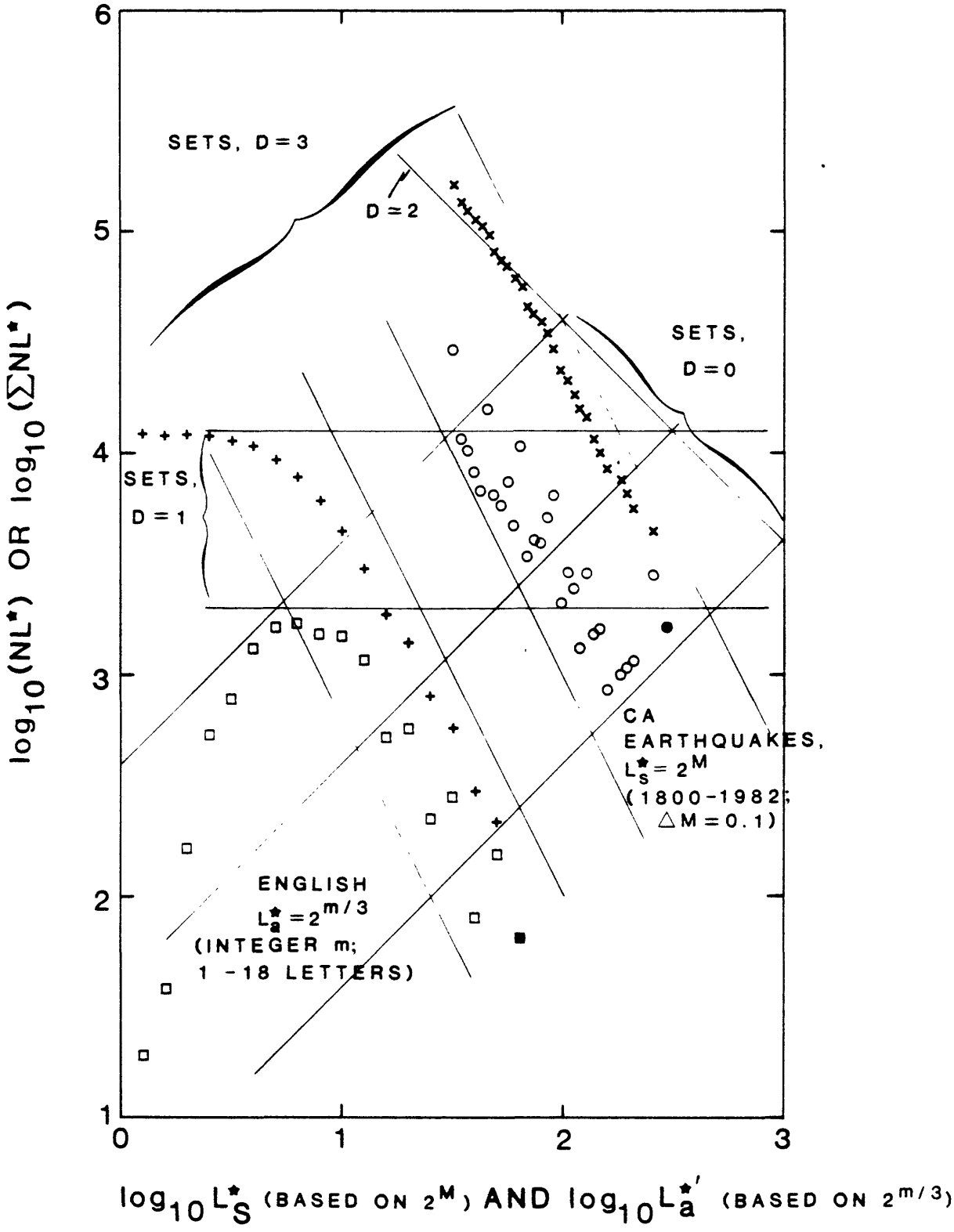

SETS, D = 3
D ≃ 2
SETS, D = 0
SETS, D = 1
CA EARTHQUAKES, $L_S^* = 2^M$ (1800-1982; ΔM = 0.1)
ENGLISH $L_a^* = 2^{m/3}$ (INTEGER m; 1 - 18 LETTERS)
$\log_{10}(NL^*)$ OR $\log_{10}(\Sigma NL^*)$
$\log_{10} L_S^*$ (BASED ON 2^M) AND $\log_{10} L_a^{*\prime}$ (BASED ON $2^{m/3}$)
0
1
2
3
1
2
3
4
5
6

FIGURE 10. Trends of average word size and event magnitude evaluated at a given frequency or frequency range. The frequency scale is that of the earthquake compilation of Algermissen and Perkins (1976) for the contiguous U.S., where the data points (open circles) represent frequency-magnitude values at the ends of their regression intervals (tielines omitted to show scatter); heavy dashed lines show averages and ranges of magnitudes for half log unit ranges of frequencies. Language data are represented by large and small dots and heavy bars. Frequencies of word counts were normalized to a range similar to that of the earthquake data by dividing by 400 (e.g., as though the counts represented a catalog generated over a 400-year time interval). Individual dots indicate average word length per frequency class in the word catalogs for the higher frequency values. Data sources are: small dots from Dewey (1923), and large dots from Hanley (1951). The heavy horizontal bars indicate averages taken over a frequency range indicated by the heavy vertical bars using the data of Hanley (1951), where the line lengths are estimates of the range of averages for the range of frequencies. The lumping was done to shorten the laborious process of extracting numbers from the Hanley catalog over small increments of frequency (there is a total of 260,430 words in the catalog and 29,899 different words, each word listed alphabetically with its count; a condensed tabulation is given only for the hundred most frequent words, which are represented here by the large dots). Light lines are reference slopes of -1, showing that the average trends (b-values of the averages) are somewhat less than unity for both data sets, apparently decreasing with decreasing average magnitudes. This roll-off is a truncation effect related to under-representation of the smaller words and event magnitudes; in the language data this reflects structural distinctions between function and content words (see text), but for the earthquake data it reflects truncation of the numerical regressions at intensities equivalent to about M = 4.3.

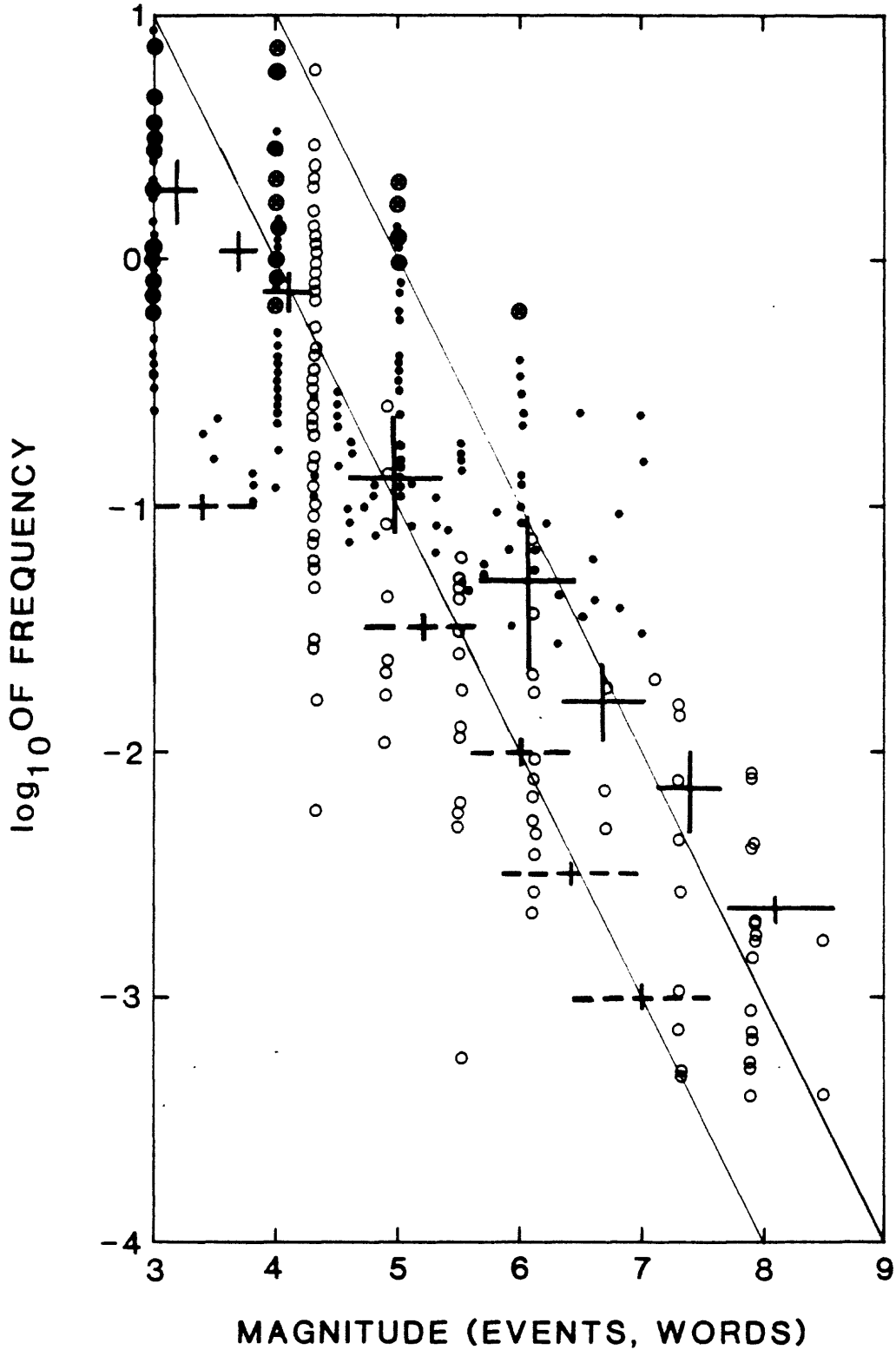

1
0
-1
-2
-3
-4
3
4
5
6
7
8
9
log10 OF FREQUENCY
MAGNITUDE (EVENTS, WORDS)

FIGURE 11. Fractal constructions as in Figure 8 but for word length expressed in terms of average numbers of letters per word. All data are taken from the catalog of Hanley (1951) but are averaged in two different ways. One set (dots and crosses) represents the data for averages shown by heavy bars in Figure 10. The other set (open circles and triangles) was obtained using the trend of data for the hundred most frequent words extrapolated to about the same word size of 8 letters per word and read at integer word lengths. The latter is heavily weighted by the shorter words, resulting in a compression analogous to the renormalized word set of Figure 9 (to compare the base 10 length scale with the base two length scale for the same letters per word, divide the former by 3.322; corresponding values are given at top of the diagram). The inset shows the same construction for a standard Cantor set, where the set theoretic fractal dimension is defined by $D = \log N/\log(1/r)$, where N is the number of partial line segments in the unit generator, and r is the ratio of segment length to unit length. The generator of the standard Cantor set is given by a unit interval in which the central third is cut out in each successive generation, giving $N = 2$, $r = 1/3$, hence $D = \log 2/\log 3 = 0.6309...$(see Mandelbrot, 1977, p. 98; Shaw and Gartner, 1986, Figures 15 and 16). In the present case this result is obtained graphically, where the slope is 0.37 for basis log(NL) giving $D = (1-.37) = 0.63$. The result for cumulative NL differs because it refers to an array of segments measured by successively smaller subsegments as the yardstick, hence it cannot be less than $D = 1$; i.e., when the Cantor set is measured as a dendritic set of line segments in a plane, as though it were a dissected fault set, it has a variable dimension between 1 and 2, hence in that context it is not a self-similar fractal (in real fault sets there is typically a self-similar interval where cumulative log-log plots are linear over a characteristic range of lengths; see Shaw and Gartner, 1986, Figures 5 and 14).

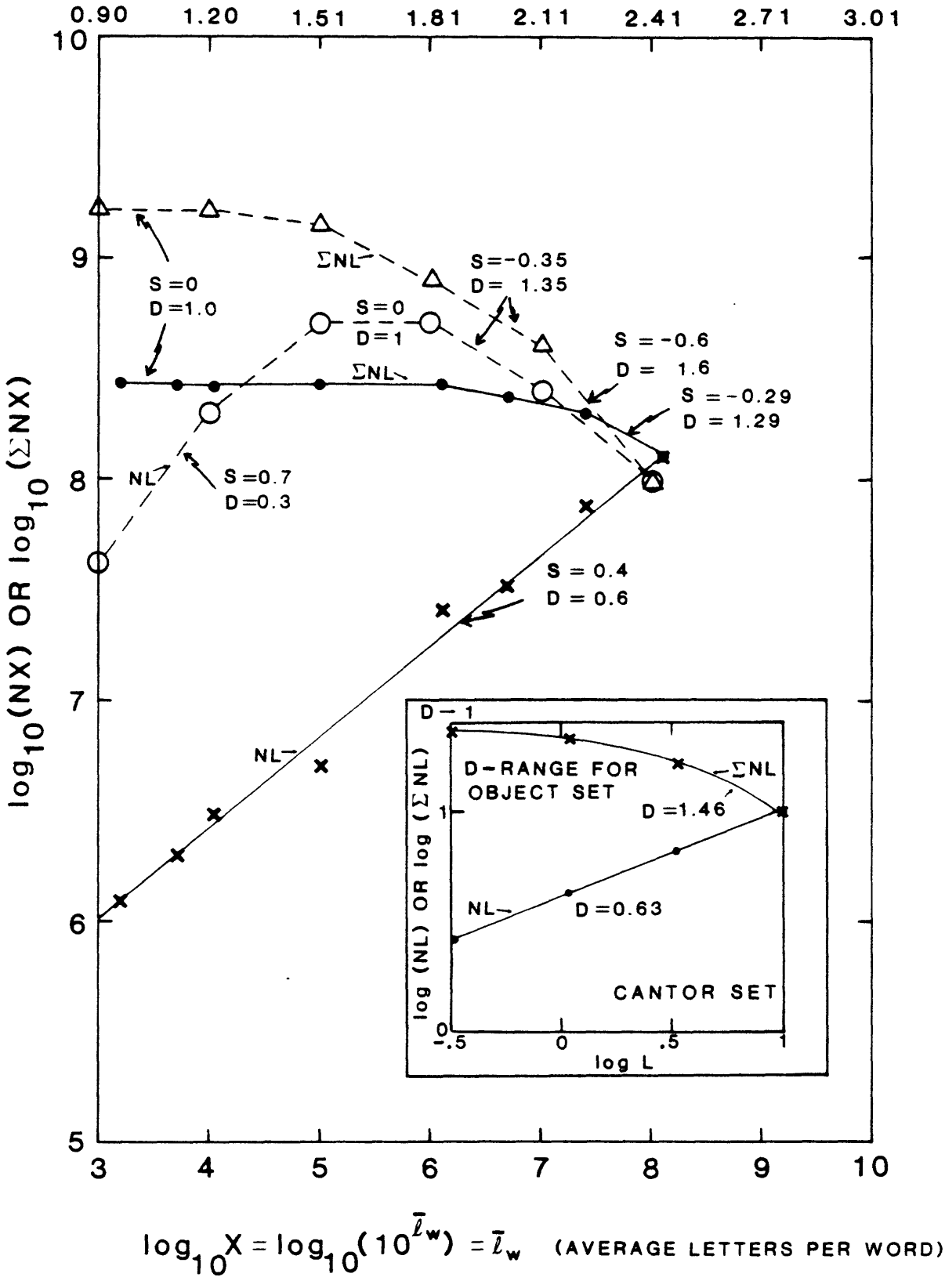
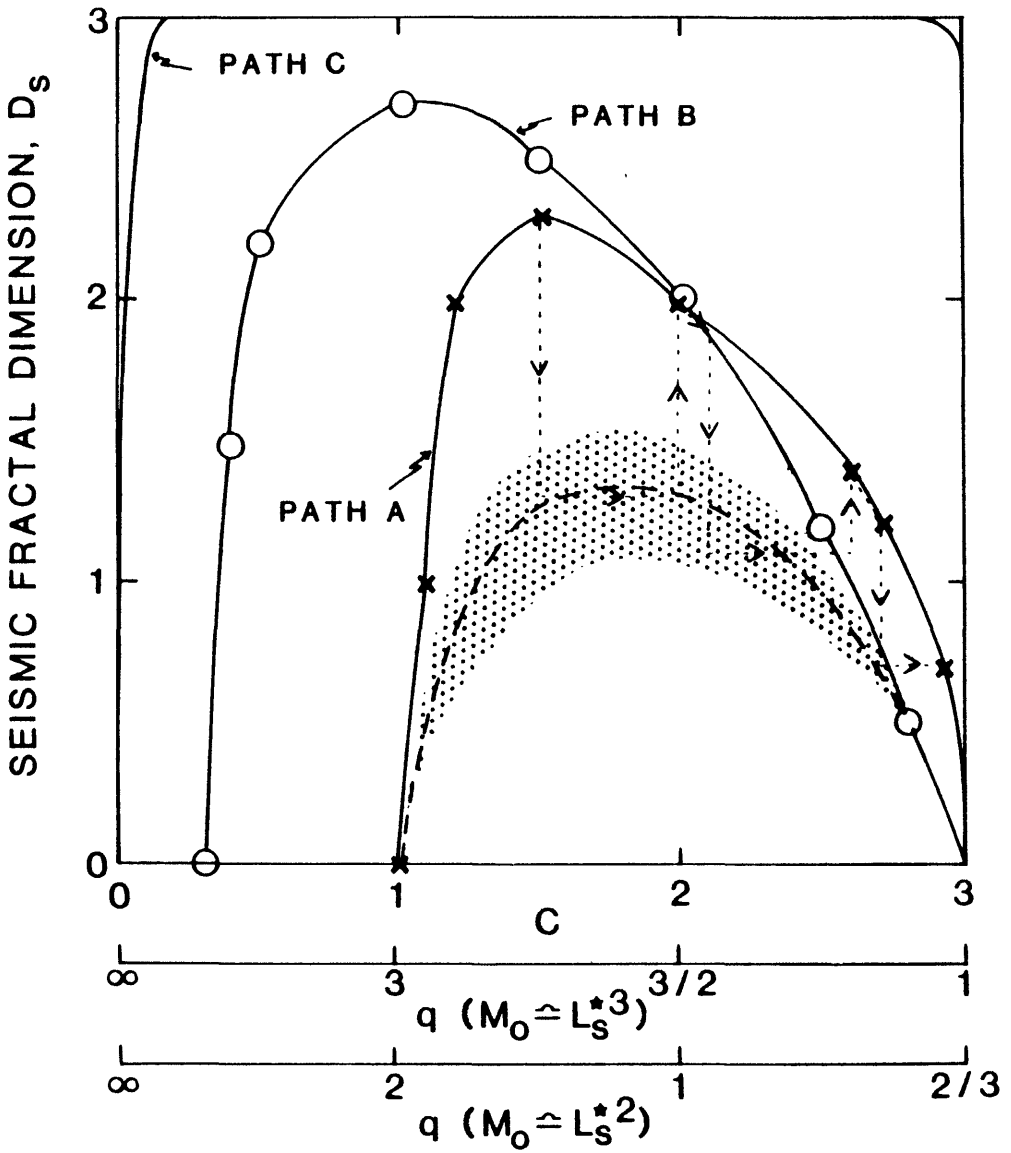

0.90
1.20
1.51
1.81
2.11
2.41
2.71
3.01
10
9
8
7
6
5
3
4
5
6
7
8
9
10
S =0
D =1.0
ΣNL
S=-0.35
D= 1.35
S=0
D=1
S = -0.6
D = 1.6
ΣNL
S = -0.29
D = 1.29
NL
S=0.7
D=0.3
S = 0.4
D = 0.6
NL
D = 1
D-RANGE FOR
OBJECT SET
ΣNL
D =1.46
NL
D =0.63
CANTOR SET
log (NL) OR log (ΣNL)
1
0
-.5
0
.5
1
log L
log10(NX) OR log10(ΣNX)
log10X = log10(10^lw) = lw (AVERAGE LETTERS PER WORD)

FIGURE 12. Seismic singularity routes through multifractal space: (a) Examples of regimes selected from possibilites illustrating nonuniqueness of seismic parameters D_s vs. c and(or) q. Path C in (a) expresses an extreme where both b- and c-values increase together toward the limit 3 at the upper right in (b) followed by abrupt decrease in b-values at c = 3. Any of the paths imply possible alternations with subsets at lower maximum b-values, as indicated by the dashed curve and stippled region in (a). Paths A through C do not represent any actual data sets, rather they suggest ranges of states over which frequency-magnitude data would describe continuous singularity spectra of types found in fluid turbulence (see Inset). Two q-value scales are shown at the bottom of (a), one assuming that seismic moment, M_o, is proportional to L^{*3}_s, the other that it is proportional to L^{*2}_s. The former corresponds to D_s defined for a topologic dimension of 3, and the latter corresponds approximately to equivalence of a fault length scale (symbol L_f in text) that is roughly proportional to $M_o^{1/2}$ (see empirical data discussed by Wesnousky and others, 1983, and by Shaw and Gartner, 1986). Magnitudes are as defined in Eq 6a, log L^{*q}_s = M. Inset: Path A in (a) rescaled to unit range of D_s and normalized c-values compared with the singularity spectrum of a two-scale Cantor set determined by Halsey and others (1986, Figures 2 and 4). The two-scale Cantor set is a mixed dissection set where there are two pairs of generating parameters N and r (see Figure 11). It is analogous to fractal measures of a hybrid tree in which there are two bifurcation ratios sampled according to a systematic mixing scheme, hence there are two limiting fractal dimensions on the abscissa at the distil limits f = 0 (representing the "pure end-members") converging to a common dimension for the set as a whole at maximum f (complete "end-member mixing"); note that the range of alpha exceeds the set dimension and in general may exceed unit value depending on the N-r ranges of the numerical generators or their dynamic equivalents (i.e., the function D = logN/log(1/r) for either of the generating functions of a generalized Cantor set is not restricted). There is a close resemblance to descriptions of statistical mechanical partition functions in critical phenomena of chemical phase equilibria and mixing trajectories in fluid convection (see discussions in Jensen and others, 1985; Halsey and others, 1986).

A


SEISMIC FRACTAL DIMENSION, D_S
PATH C
PATH B
PATH A
0
1
2
3
C
∞
3
3/2
1
q ($M_o \doteq L_s^{*3}$)
∞
2
1
2/3
q ($M_o \doteq L_s^{*2}$)

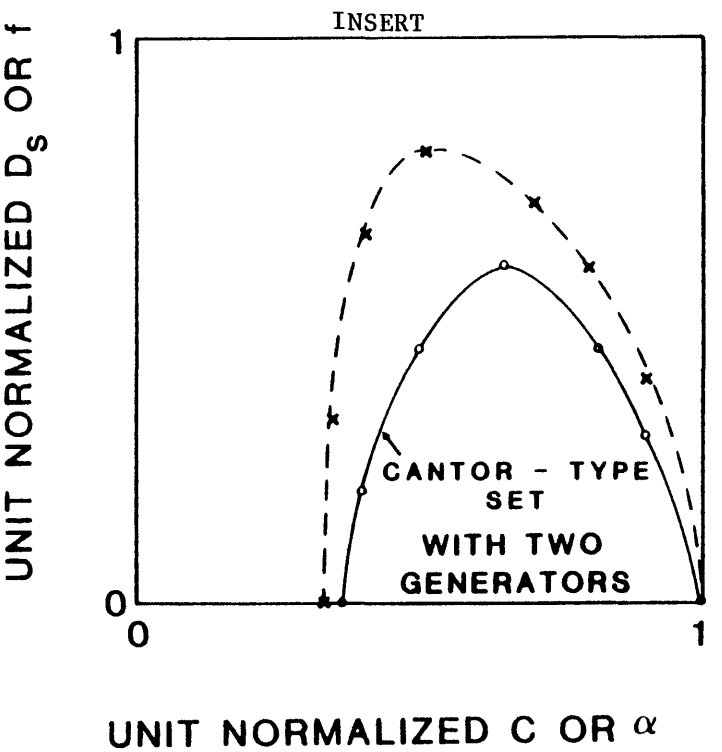

UNIT NORMALIZED D_S OR f
INSERT
1
0
CANTOR - TYPE SET
WITH TWO GENERATORS
0
1
UNIT NORMALIZED C OR α

FIGURE 12. Seismic singularity routes through multifractal space:
(b) Consistent values of parameters b, c, and fractal D_s based on $D_s = 3b/c$ (see text); given two parameters the third is read from a point on, or interpolated from, isopleths of D_s in Figure 18a of Shaw and Gartner (1986). Lines with arrows in (b) illustrate b-c-D_s routes consistent with particular patterns of b-value variations in frequency-magnitude diagrams of Shaw and Gartner (1986); curves A and B indicate generalized sets suggested by frequency-magnitude data with multiple b-values (and implicitly c-values) ranging between 0 and 3 (analogous to fractal arrays in Figure 9). The extrema in Paths A and B express a possibility of large variations in b-values at low or high c-values and relatively constant b-values at intermediate c-values.

B

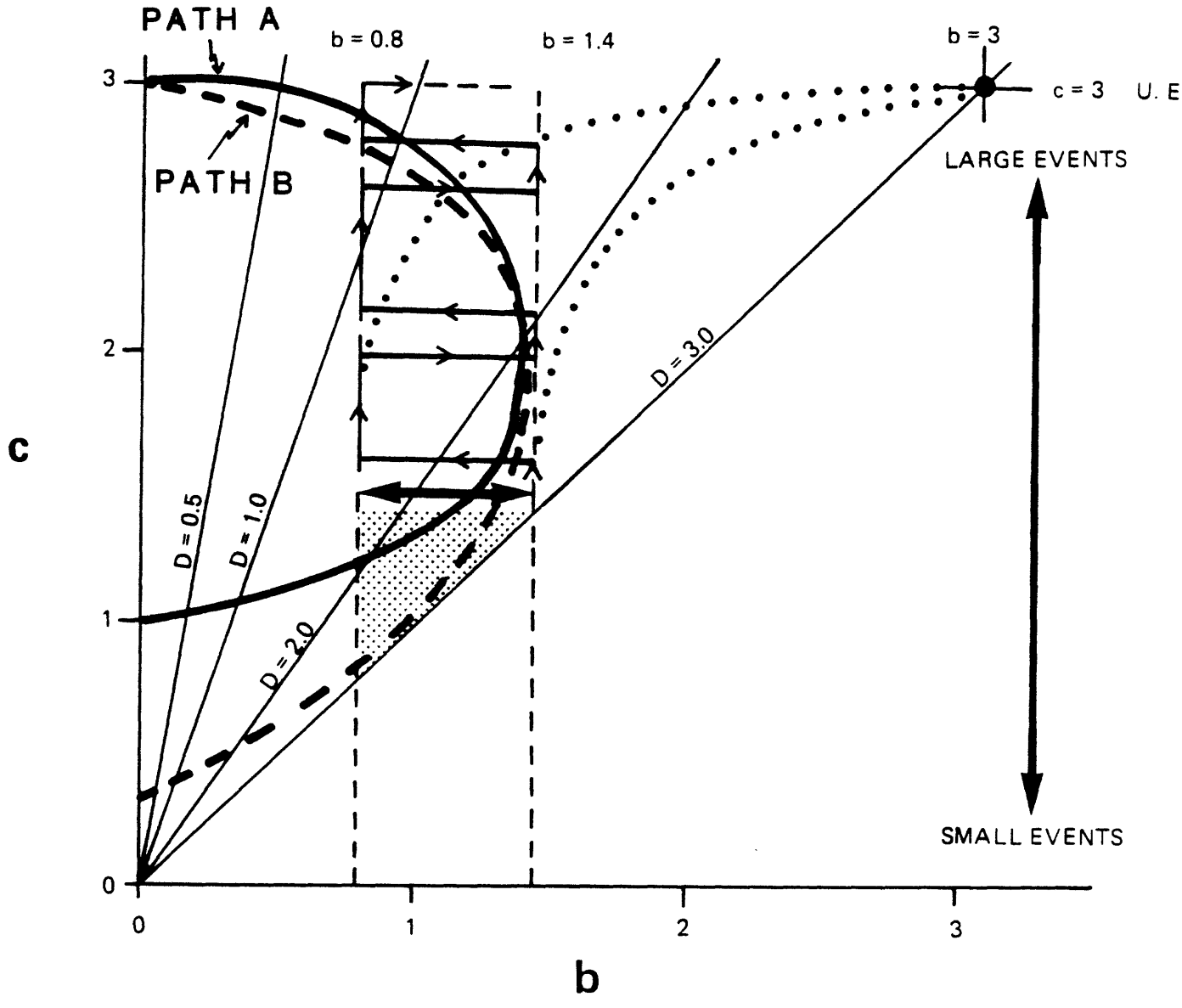
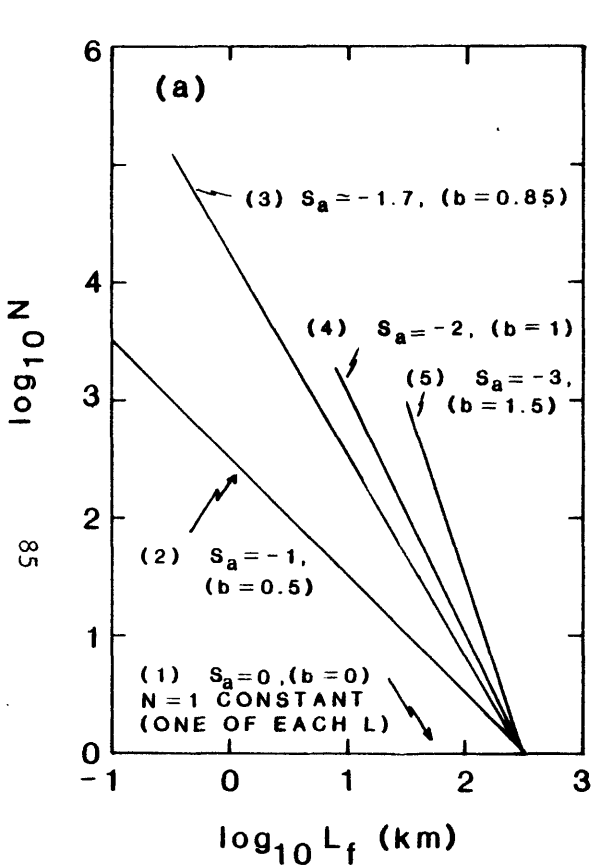
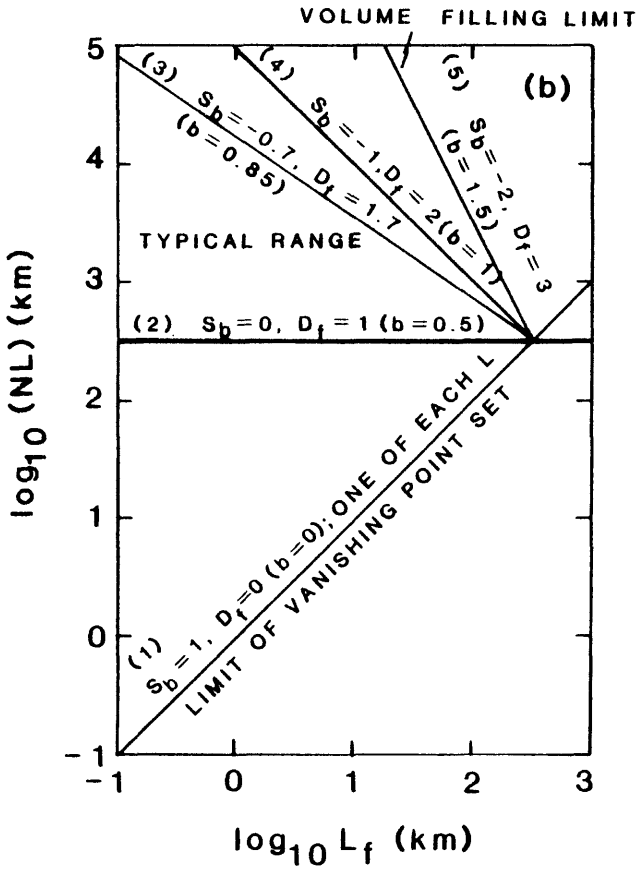

PATH A
PATH B
b = 0.8
b = 1.4
b = 3
c = 3 U. E.
LARGE EVENTS
SMALL EVENTS
D = 0.5
D = 1.0
D = 2.0
D = 3.0
c
b
0
1
2
3

FIGURE 13. Demonstration of relations between numbers (N), lengths (L_f), slopes (S), fault fractals (D_f), and recurrence parameters (b-values) for idealized fault dendrites scaled to N = 1 at a fault length roughly equivalent to an M = 8 earthquake using formula M = 4.964 + 1.243 log L_f for L_f in kilometers (converted from Eq. 1 in Shaw and Gartner, 1986): (a) log N *vs.* log L_f. (b) log (NL_f) *vs.* log L_f. (c) log (Sum NL_f) *vs.* log L_f (summation is for all segments of length equal to or greater than a chosen yardstick length). Lines of the same constant b-values are numbered 1 through 5 in each plot. Note that the absolute value of slope S_a in (a) is identical to the value of D_f determined from the relation D = (1-S_b) in (b), where S_i refers to the slope of a given plot. Therefore, D = -S_a in a plot of segment number against segment length for sets ordered as in (a); see number-length slopes of fault data given by Shaw and Gartner (1986, Figure 5). The formula D = (1-S) applied to (c) quickly converges to D_f = 1 for b-values below 0.5 (curve 2). Because of the summation convention for line segments, D_f in (c) can not be less than unity; however, D_f in (c) is similar to values derived from (a) and (b) over the more linear portions of the steeper curves (divergences near log N = 0 depend on summation increments; the dotted line is for intervals of 0.1 compared with intervals of 0.5 for the other lines). Graph (c) treats discontinuous line sets in a manner similar to that for connected fractals, where N would represent the number of divider intervals of length L measured along the irregular but unbroken line (as in measurements of coastlines). In that case it is obvious that the minimum fractal dimension of an unbroken line is D = 1, and the maximum is D = 2 for measurements in a plane; however, if the trajectories of connected lines can depart from the plane, as in Brownian motion, the fractal dimension can approach D = 3 (the same is true for disconnected sets such as dust clouds in which there is a hierarchical size distribution, or for debris clouds of platy fragments in which there is a wide range of aspect ratios). For connected lines, the quantities NL and Sum NL are automatically the same, but for broken, or "dissected" line sets the cumulative and incremental lengths are not the same because N refers to a partial measure representing a given subset in (b) which is then summed in (c); see alternative methods of measuring fault sets discussed by Shaw and Gartner (1986), and by Okubo and Aki (in press).


(a)
(3) $S_a = -1.7$, (b = 0.85)
(4) $S_a = -2$, (b = 1)
(5) $S_a = -3$, (b = 1.5)
(2) $S_a = -1$, (b = 0.5)
(1) $S_a = 0$, (b = 0) N = 1 CONSTANT (ONE OF EACH L)
$\log_{10} N$
$\log_{10} L_f$ (km)


VOLUME FILLING LIMIT
(b)
(3) $S_b = -0.7$, $D_f = 1.7$ (b = 0.85)
(4) $S_b = -1$, $D_f = 2$ (b = 1)
(5) $S_b = -2$, $D_f = 3$ (b = 1.5)
TYPICAL RANGE
(2) $S_b = 0$, $D_f = 1$ (b = 0.5)
(1) $S_b = 1$, $D_f = 0$ (b = 0); ONE OF EACH L
LIMIT OF VANISHING POINT SET
$\log_{10}$ (NL) (km)
$\log_{10} L_f$ (km)

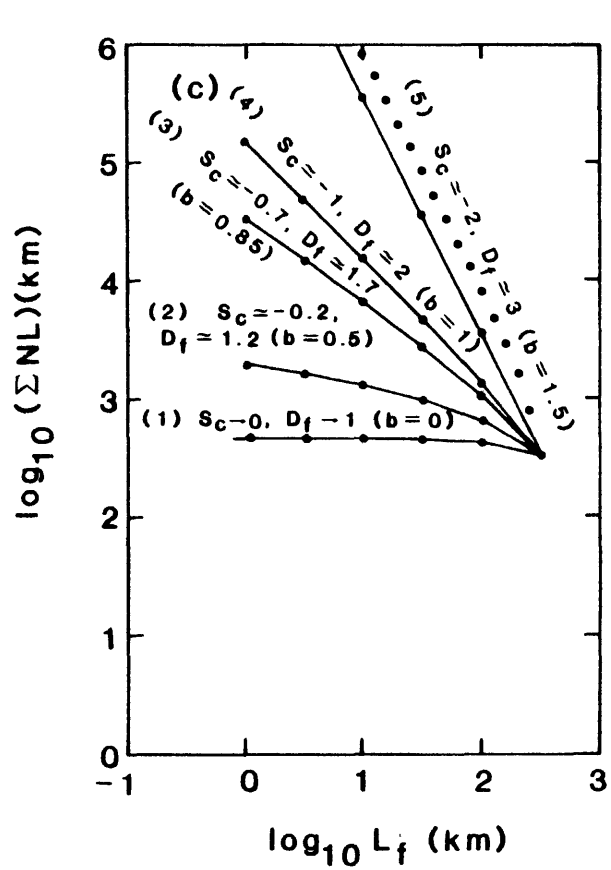

(c)
(3) $S_c = -0.7$, $D_f = 1.7$ (b = 0.85)
(4) $S_c = -1$, $D_f = 2$ (b = 1)
(5) $S_c = -2$, $D_f = 3$ (b = 1.5)
(2) $S_c = -0.2$, $D_f = 1.2$ (b = 0.5)
(1) $S_c = 0$, $D_f = 1$ (b = 0)
$\log_{10}$ (ΣNL)(km)
$\log_{10} L_f$ (km)

FIGURE 14. Index and "linkage" map of California and environs showing locations of subregions used for regional variations of frequency-magnitude patterns; "benchmarks" are positioned to approximately straddle the subregions. Numbered lines connecting the triangles are referred to descriptively as kinematic links for purposes of illustrating time variations of regional patterns for events equal to or greater than M = 5 (see Figure 24). The lines between the most northern and northeastern triangles are not numbered because data are not available in the catalogs to evaluate activity just beyond these limits (they are shown here and in Figure 24 to identify the perimeter of "California and environs"); in general, however, central Oregon, southwestern Idaho, eastern Nevada-western Utah, and western Arizona have had low seismicity during historic times (see Inset B of Figure 3; and Ryall and others, 1966). Events in SH and TC are generally south of 42^{o} north latitude; more northerly events offshore in E are referred to either lines 13 or 15. Events in the eastern parts of TC and CN are referred to lines 16, 17, and(or) 19, depending on coordinates. The arbitrary benchmarks are identified by nearby towns or geographic features.

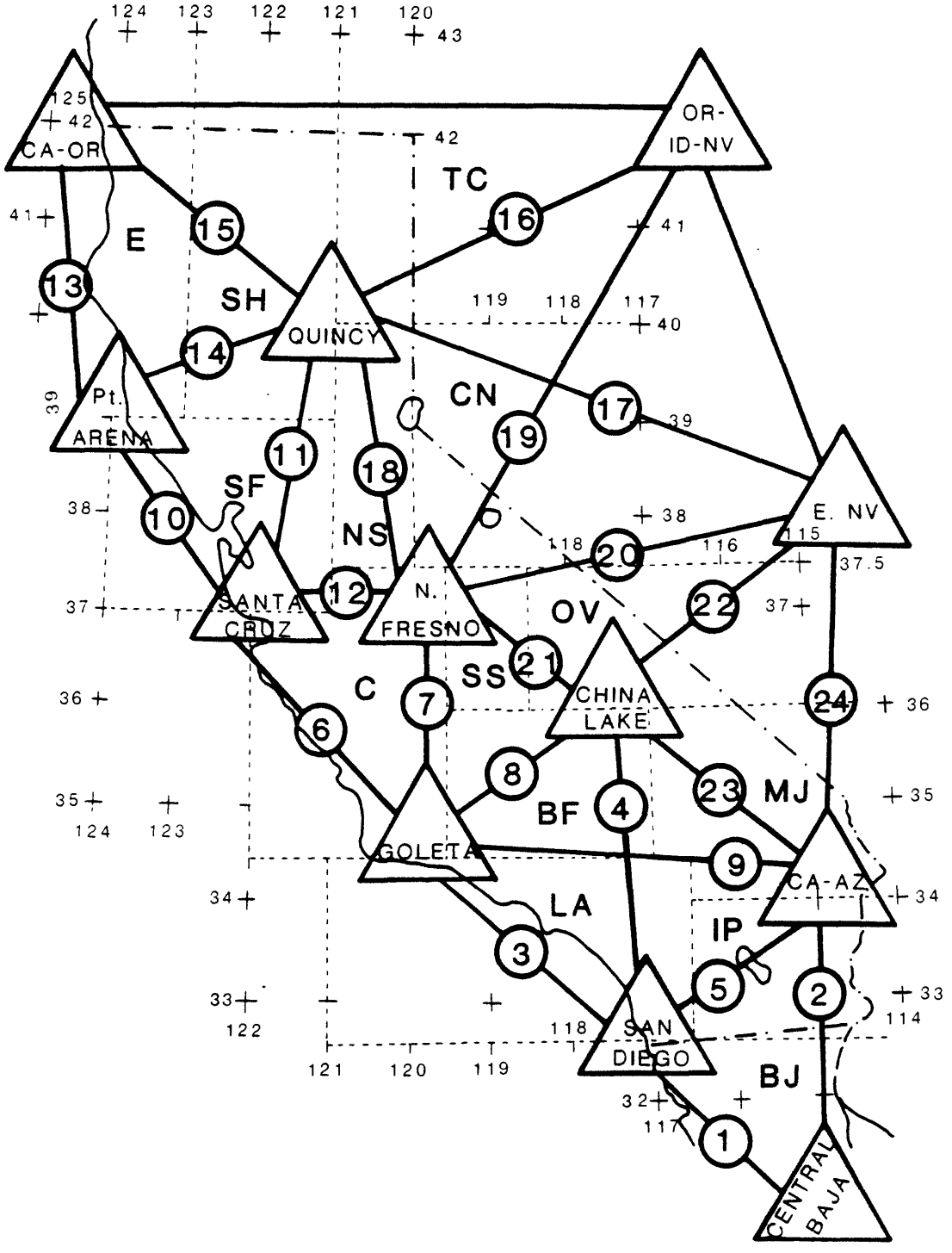

124
123
122
121
120
43
125
42
CA-OR
OR-
ID-NV
TC
E
SH
QUINCY
CN
Pt.
ARENA
SF
NS
E. NV
SANTA
CRUZ
N.
FRESNO
OV
SS
CHINA
LAKE
C
BF
MJ
GOLETA
CA-AZ
LA
IP
SAN
DIEGO
BJ
CENTRAL
BAJA
41
40
39
38
37
36
35
34
33
32
119
118
117
116
115
114
37.5

FIGURE 15. Frequency-magnitude data of Figure 5 subdivided into eight regional classes representing one or more of the subregions identified in Figure 14 (subregions were combined where there is a paucity of data or no specific reason to show individual sets; e.g., data for the northeast corner of California and northwestern Nevada were lumped). The three "cornerstone" areas are represented in this illustration by the groups E+SH (Eureka, Shasta), CN+TS+NS (California-Nevada, Three-Corner area, Northern Sierra), and BJ+IP (Baja California, Imperial Valley). The left column of diagrams gives the historic record; the right column is broken down according to pre- and post-1900 data.. A point to notice is the resemblance of frequency oscillations among the three cornerstone groups compared to other regions; e.g., there are relatively high frequencies in the vicinites M = 5, M = 5.5, M = 5.8, M = 6, and for E+SH and BJ+IP near M = 6.5. Also the oscillatory behavior often appears to reflect alternations between trends of similar b-values over roughly a decade in frequency (except LA which has a narrower bandwidth). Such effects are interpreted (see text) to correspond to shifts or "jumps" from one fractal subset to another among sets of the sorts illustrated in Figures 9 and 12. Comparisons with behavior of smoothed data over somewhat larger areas is indicated by the cumulative frequency curves of Ryall and others (1966) shown in the Inset of Figure 5.

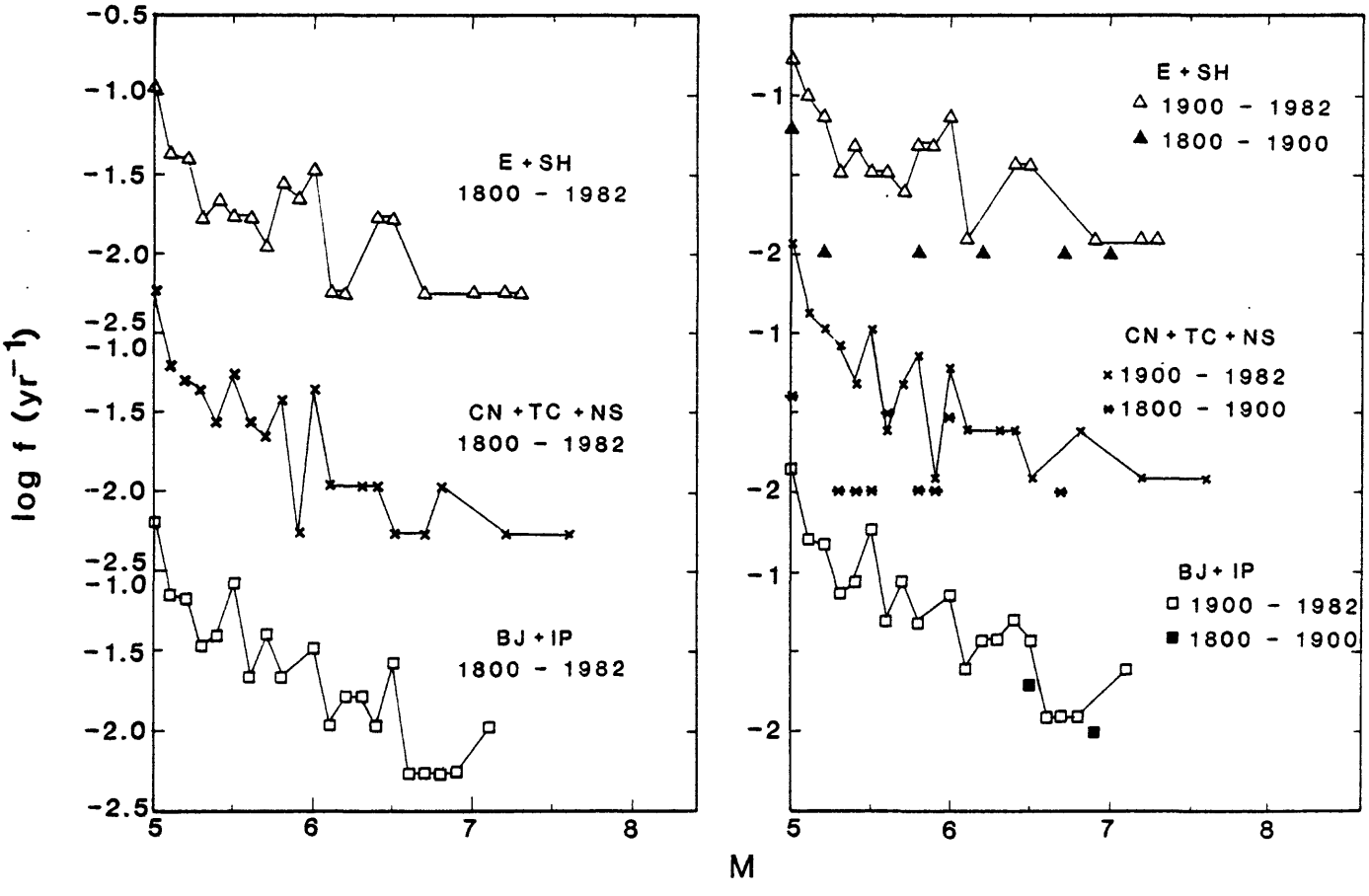


Figure 15. Continued.

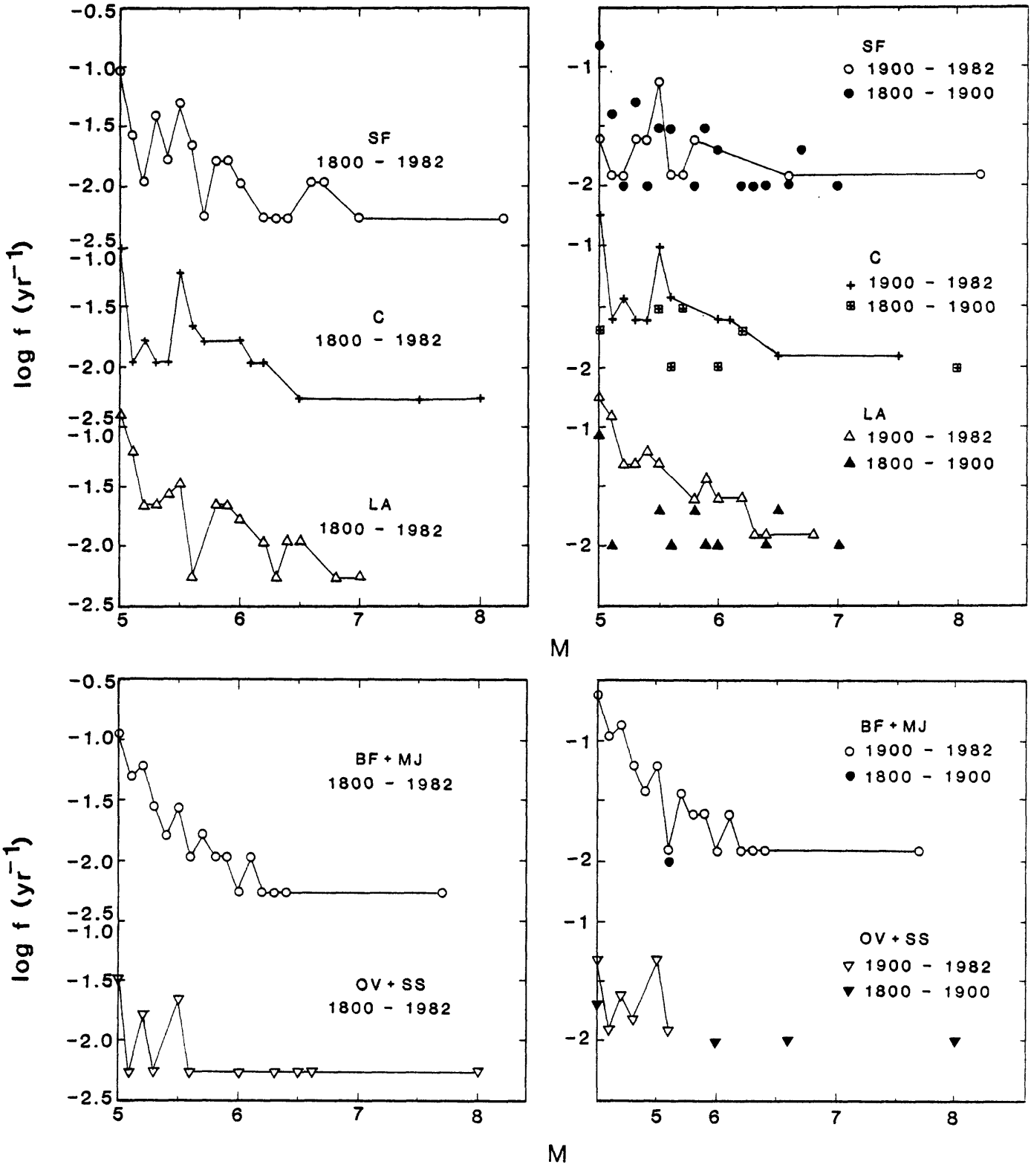
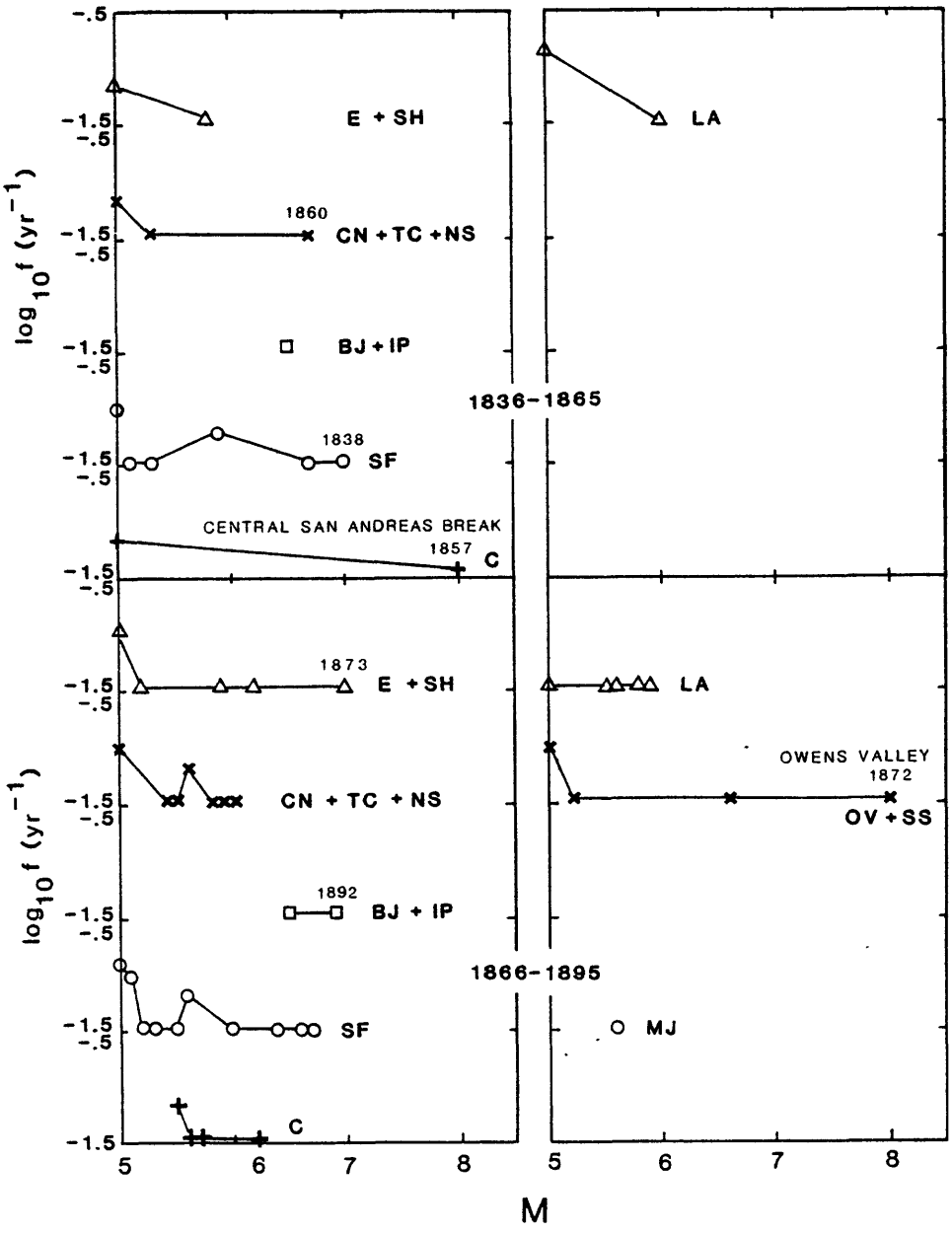


FIGURE 16. Frequency-magnitude data broken down into 30-year intervals for nine regional classes, from top left: E+SH, CN+TC+NS, BJ+IP, SF, C, LA, OV+SS, BF, MJ. The year of occurrence is given for some of the larger events in each class. Note that for the post-1926 data the patterns in the first three sets (the "cornerstone regions") are mutually more similar than they are to other sets; this marks the time when modern seismic networks began to be established, hence these resemblances probably apply to longer term behavior.



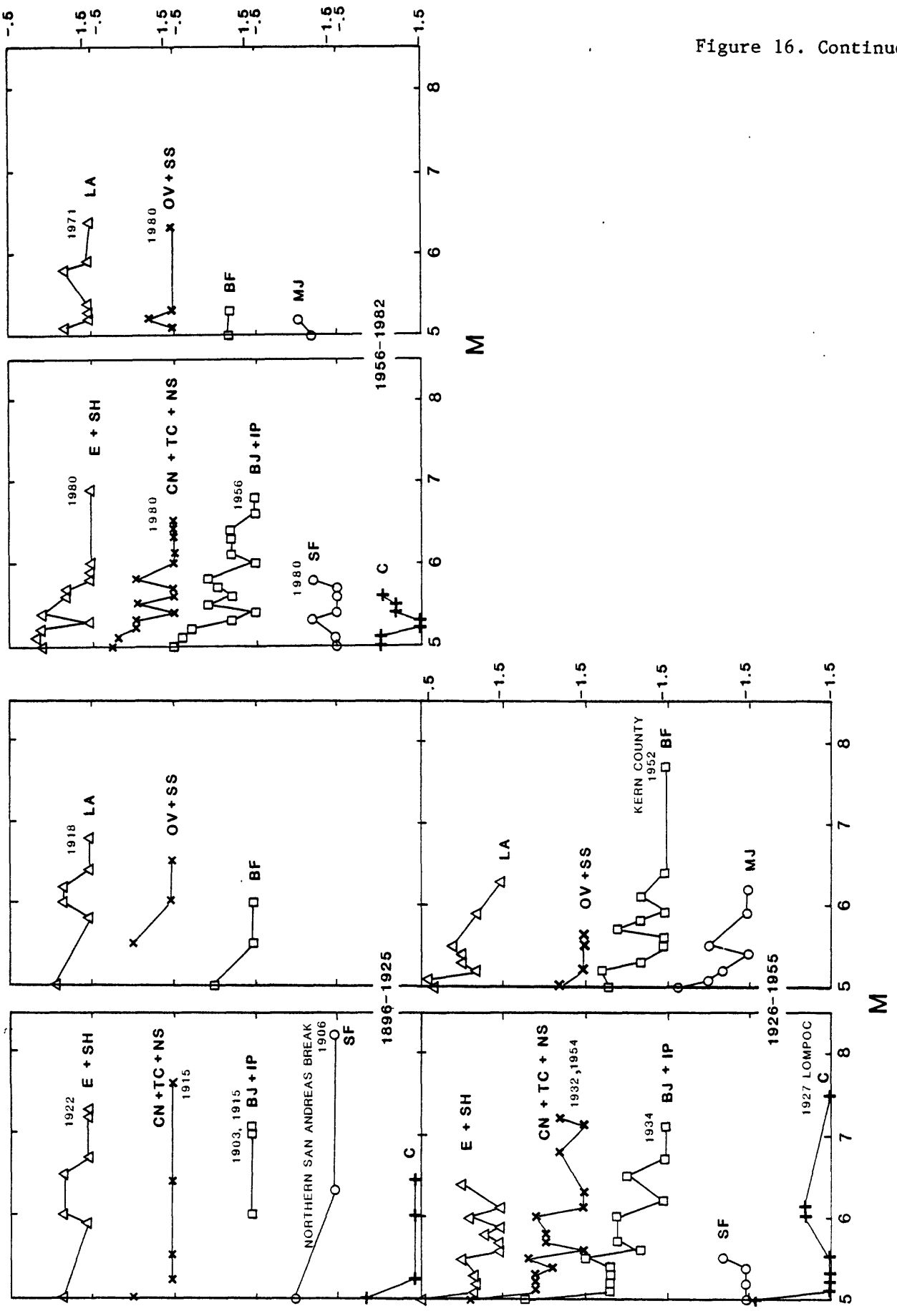


Figure 16. Continued

FIGURE 17. Synopsis of overall frequency-magnitude fluctuations at intervals of five years. The logarithmic frequency scale is compressed to emphasize contrasts in the distributions and ranges of magnitudes. Some of the larger events are identified by locality and year of occurrence (see Table 2 and Figure 18 for event counts and moments within subregions).

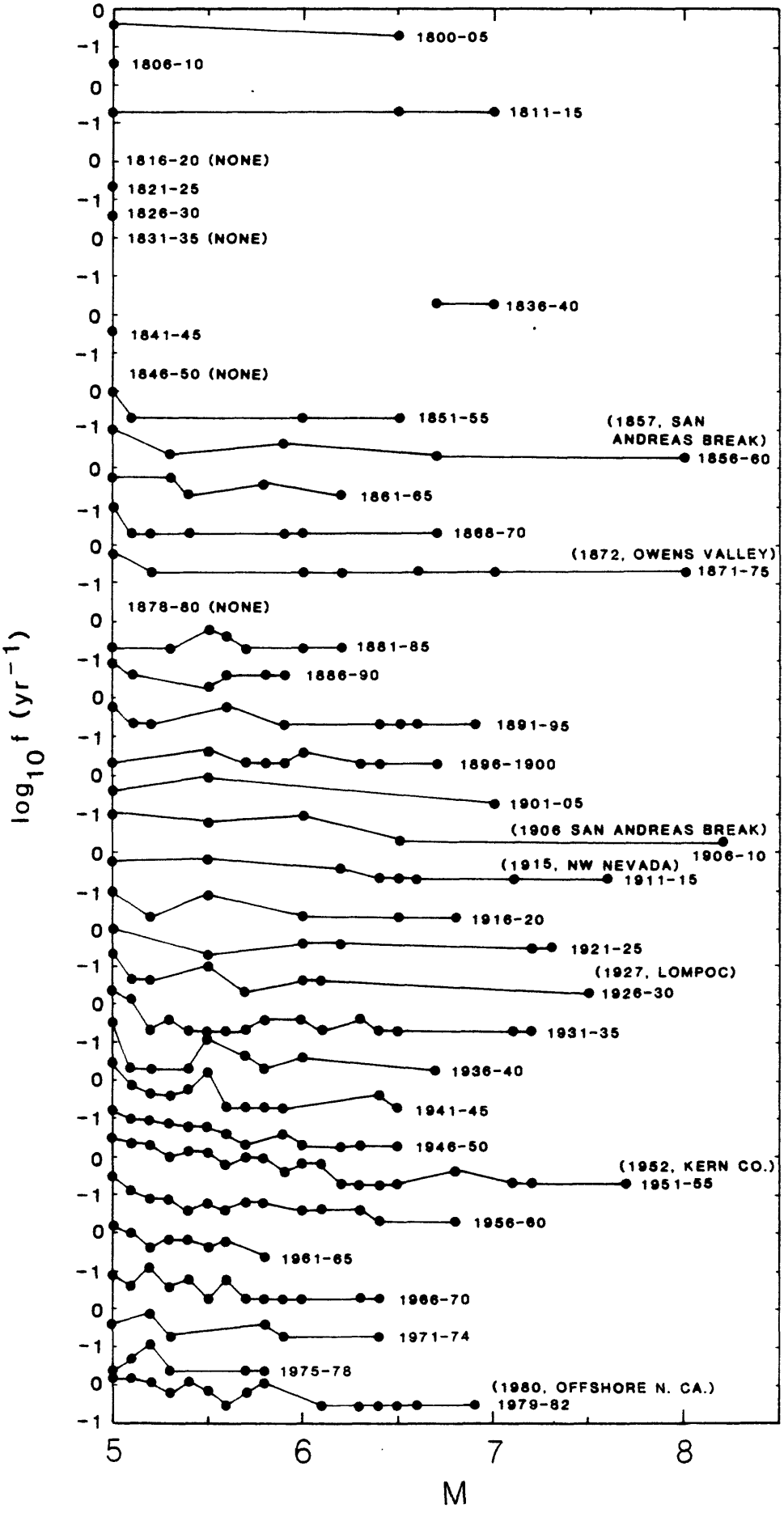

1800-05
1806-10
1811-15
1816-20 (NONE)
1821-25
1826-30
1831-35 (NONE)
1836-40
1841-45
1846-50 (NONE)
1851-55
(1857, SAN ANDREAS BREAK)
1856-60
1861-65
1866-70
(1872, OWENS VALLEY)
1871-75
1876-80 (NONE)
1881-85
1886-90
1891-95
1896-1900
1901-05
(1906 SAN ANDREAS BREAK)
1906-10
(1915, NW NEVADA)
1911-15
1916-20
1921-25
(1927, LOMPOC)
1926-30
1931-35
1936-40
1941-45
1946-50
(1952, KERN CO.)
1951-55
1956-60
1961-65
1966-70
1971-74
1975-78
(1980, OFFSHORE N. CA.)
1979-82
0
-1
5
6
7
8
M
log10 f (yr-1)

FIGURE 18. Summary histograms of earthquake counts and moments at five-year intervals from data of Table 2:

A. Total counts and counts per subregions of Figure 14. The numbers of large events (M = 7.0 $\pm$ 0.3) are shown by stars; major events (M = 7.5 and greater) are shown by solid dots. The sum of counts per histogram is given at right with symbol of subregions.

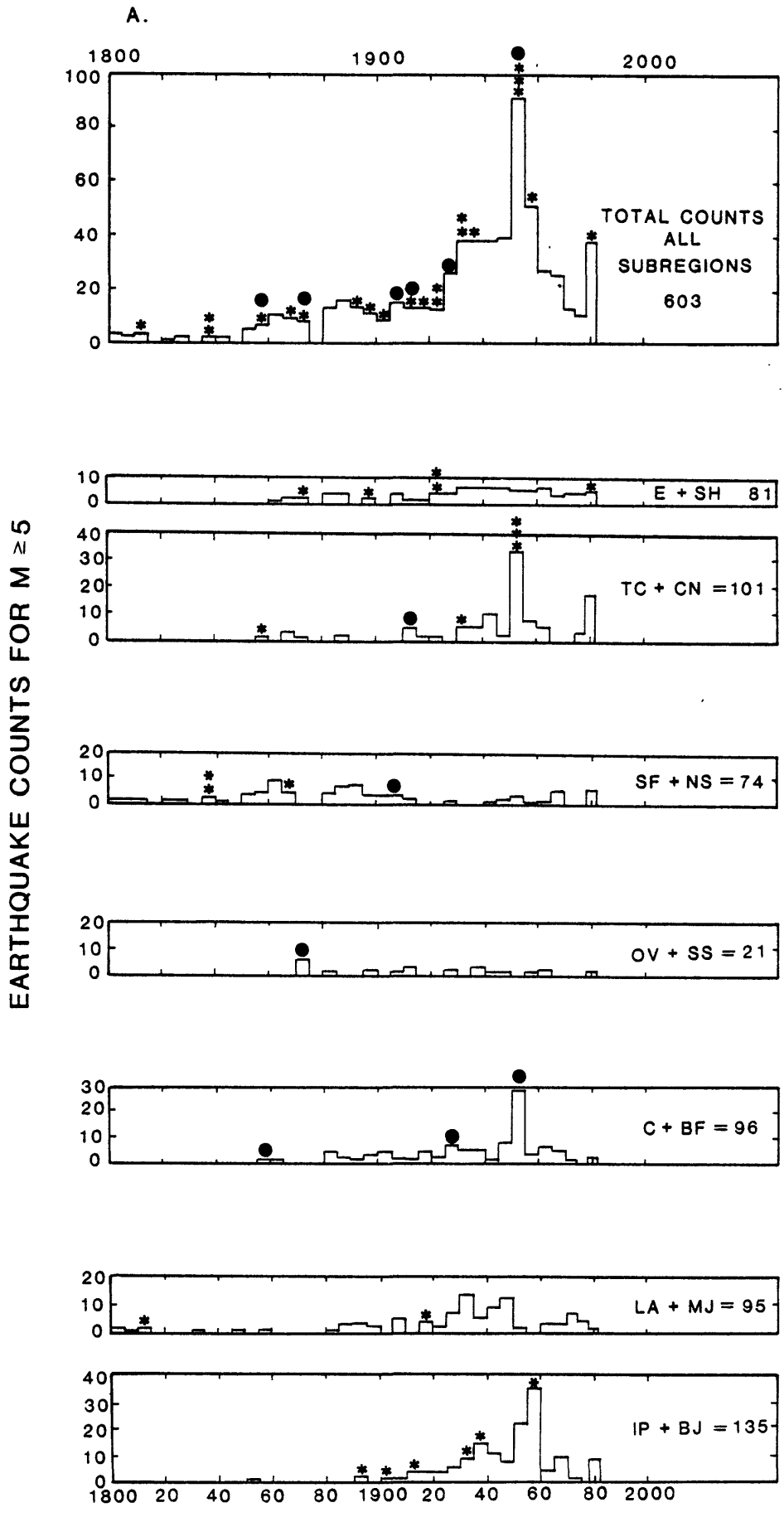

A.
1800
1900
2000
TOTAL COUNTS
ALL
SUBREGIONS
603
E + SH 81
TC + CN =101
SF + NS = 74
OV + SS = 21
C + BF = 96
LA + MJ = 95
IP + BJ = 135
EARTHQUAKE COUNTS FOR M ≥ 5
YEAR

FIGURE 18. Summary histograms of earthquake counts and moments at five-year intervals from data of Table 2:

B. Moments calculated using relation log M_o = 16.1 + 1.5 M (Hanks and Kanamori, 1979; Shaw and Gartner, 1986, Eq. 3). Ordinate gives the logarithm of moments summed within each five-year interval identified by region; stars and solid dots identify the largest events, as in A. Total moments per category in units of 10^{28} dyne-cm are given immediately below the subregion symbols (count numbers from A shown in parentheses above symbols). The moment magnitude equivalent to the total moment per category, designated EMM, is also given below the corresponding values of moment (e.g., EMM = 8.5 for M_o = 7.3×10^{28} dyne-cm, using the above equation).

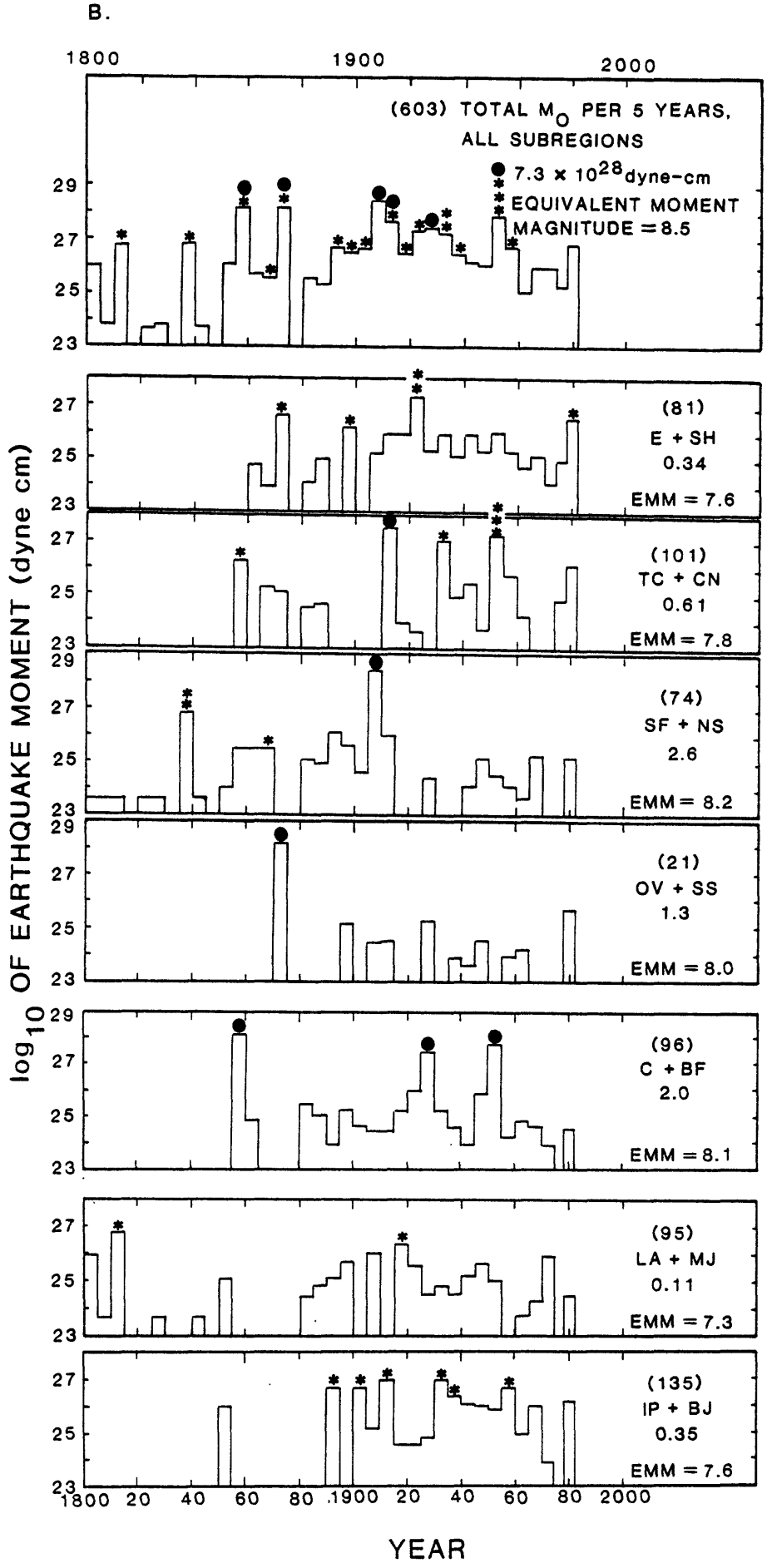

B.
1800
1900
2000
(603) TOTAL M_O PER 5 YEARS, ALL SUBREGIONS
7.3 × 10^{28}dyne-cm EQUIVALENT MOMENT MAGNITUDE = 8.5
(81) E + SH 0.34 EMM = 7.6
(101) TC + CN 0.61 EMM = 7.8
(74) SF + NS 2.6 EMM = 8.2
(21) OV + SS 1.3 EMM = 8.0
(96) C + BF 2.0 EMM = 8.1
(95) LA + MJ 0.11 EMM = 7.3
(135) IP + BJ 0.35 EMM = 7.6
$\log_{10}$ OF EARTHQUAKE MOMENT (dyne cm)
YEAR

FIGURE 19. Multifractal diagrams constructed as in Figure 9 from frequency-magnitude data in Figure 15 for California subregions of Figure 14:

A. Length scale based on relation $L^*_s = 2^{M/q}$ for $q = 4/3$. The Base 2 data are shown for comparison with Figure 9 and to illustrate approximate invariances of fractal subtrends with changes of logarithmic base. Frequencies are normalized to 1000 years. To compare with empirical correlations to Base 10 in Table 1 (and diagrams in B), values of $\log_{10} L^*_s$ should be multiplied by factor $\log_{10}2$. Differences in Base 2 and 10 plots are indicated by average slopes of overall sets on the different bases (i.e., regression lines through all data points of a set have negative slopes in A and positive slopes in B). Dashed lines indicate self-similar fractal subsets of constant D^*_s for specified basis (see Inset at right). Trends with slopes that are algebraically smaller than -2 (allowing for appropriate conversion of logarithmic base) exceed topologic range $D_T = 3$ and refer to higher order phase space (see Shaw and Gartner, 1986, for discussion of time as a fourth topologic dimension); slopes greater than +1 are imaginary because they have dimensions less than zero for any exponent of length scale. The full historic record is used in all cases except LA because the pre-1900 data are significant only in SF and LA. The latter subregion is shown for three time ranges because the pre- and post-1900 data are clearly separated (note that even though values are shifted the general patterns remain similar).

A.

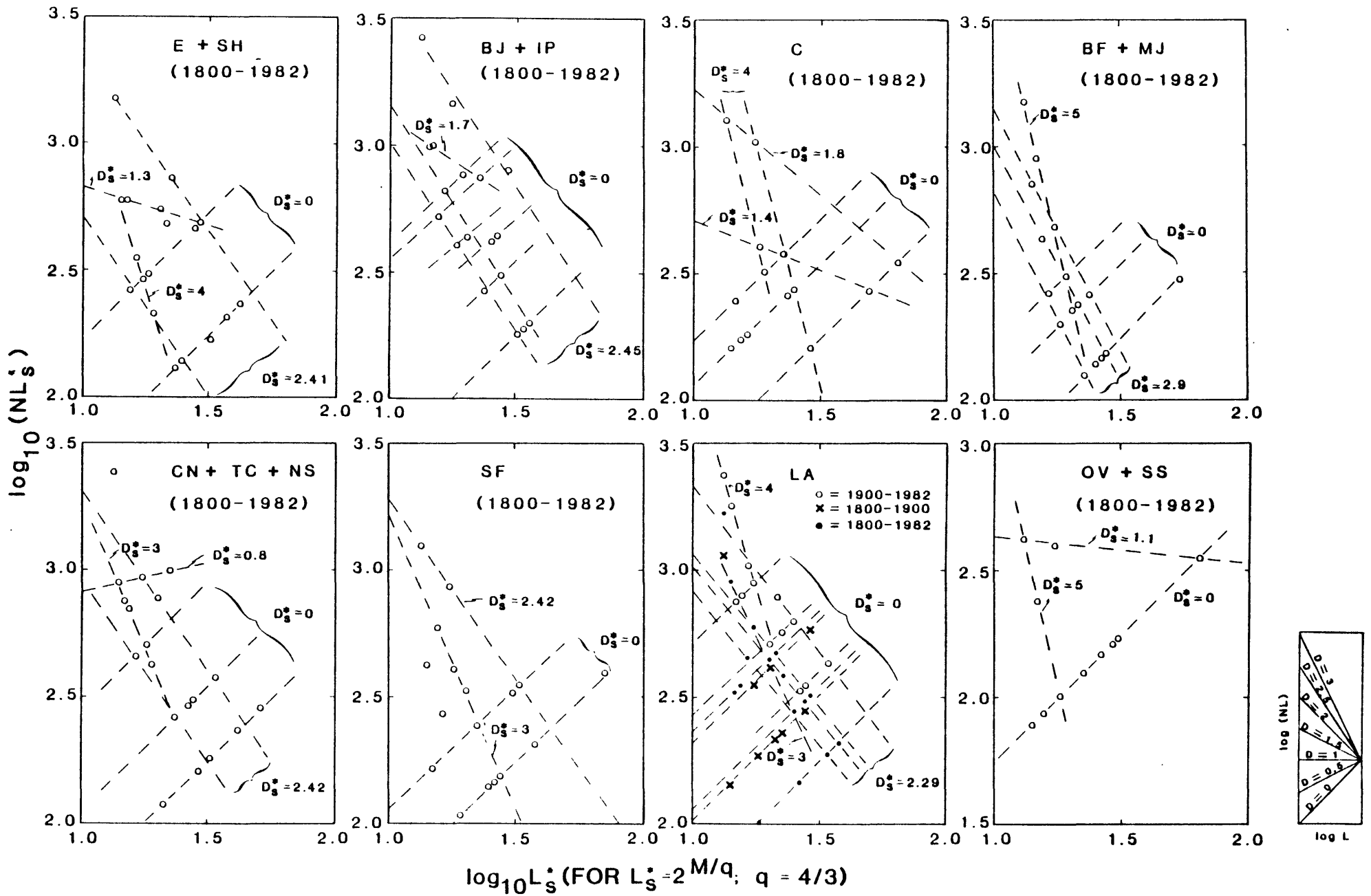

E + SH
(1800-1982)
D_S^* = 1.3
D_S^* = 0
D_S^* = 4
D_S^* = 2.41
BJ + IP
(1800-1982)
D_S^* = 1.7
D_S^* = 0
D_S^* = 2.45
C
(1800-1982)
D_S^* = 4
D_S^* = 1.8
D_S^* = 0
D_S^* = 1.4
BF + MJ
(1800-1982)
D_S^* = 5
D_S^* = 0
D_S^* = 2.9
CN + TC + NS
(1800-1982)
D_S^* = 3
D_S^* = 0.8
D_S^* = 0
D_S^* = 2.42
SF
(1800-1982)
D_S^* = 2.42
D_S^* = 0
D_S^* = 3
LA
D_S^* = 4
○ = 1900-1982
× = 1800-1900
● = 1800-1982
D_S^* = 0
D_S^* = 3
D_S^* = 2.29
OV + SS
(1800-1982)
D_S^* = 1.1
D_S^* = 5
D_S^* = 0
$\log_{10}(NL_S^*)$
$\log_{10} L_S^*$ (FOR $L_S^* = 2^{M/q}$; q = 4/3)
log (NL)
log L
D = 3
D = 2.5
D = 2
D = 1.5
D = 1
D = 0.5
D = 0

FIGURE 19. Multifractal diagrams constructed as in Figure 9 from frequency-magnitude data in Figure 15 for California subregions of Figure 14:

B. Length scale in A converted to dimensional exponents and units consistent with fault measurements in Shaw and Gartner (1986). Frequencies are renormalized to numbers of events occurring in 10,000 years (Holocene) at the temporal frequencies of Figure 15 (i.e., $\log_{10}N = \log_{10}f + 4$). The length L^*_f and fractal dimension D^*_f refer to length obtained by equating the following relations at the same value of M: $M = 4.964 + 1.243 \log_{10}L_f$ (for length in kilometers), and $\log_{10}L^*_s = (M/q) \log_{10}2$ (magnitude-defined units of length). Units defined by equating magnitude are called the seismically defined fault length, L^*_f; it is given by the empirical equation $\log_{10}L^*_f = 3.56 \log_{10}L^*_s - 4$. This defines magnitude scale at top as approximately that of Shaw and Gartner (1986) for faults of length defined on scale at bottom, where lengths are converted from values of L^*_s in A (i.e., these are geometrical relations calibrated by specified regression relation of fault-length vs. magnitude). Grid lines are drawn for fractal dimensions $D^*_f = (1-S)$. Nodes are shown for each decade of recurrence interval in years (i.e., at values of 1/f); e.g., for N = 1000, $\log_{10}f = -1$, giving average recurrence interval T = 10 years at M = 5 (converted fault length, L^*_f, is about 1 km). Large dots and heavy lines are converted data equivalent to A; estimated values D^*_f are listed below region symbols. Inset (at right): slope vs. fractal relation for constant b-values from 0 to 1.5 for arbitrary log N vs. M. Ratios of fractal dimensions on different bases D^*_s, D^*_f, and D_s are given in Inset; similarities of values suggest that multifractal trends are roughly invariant at integer powers of length scale (i.e., multifractals form crude universal sets in n-dimensional hyperspace up to at least $D^*_s = 10$, $D^*_f = 3$).

B.

APPROXIMATE MAGNITUDE

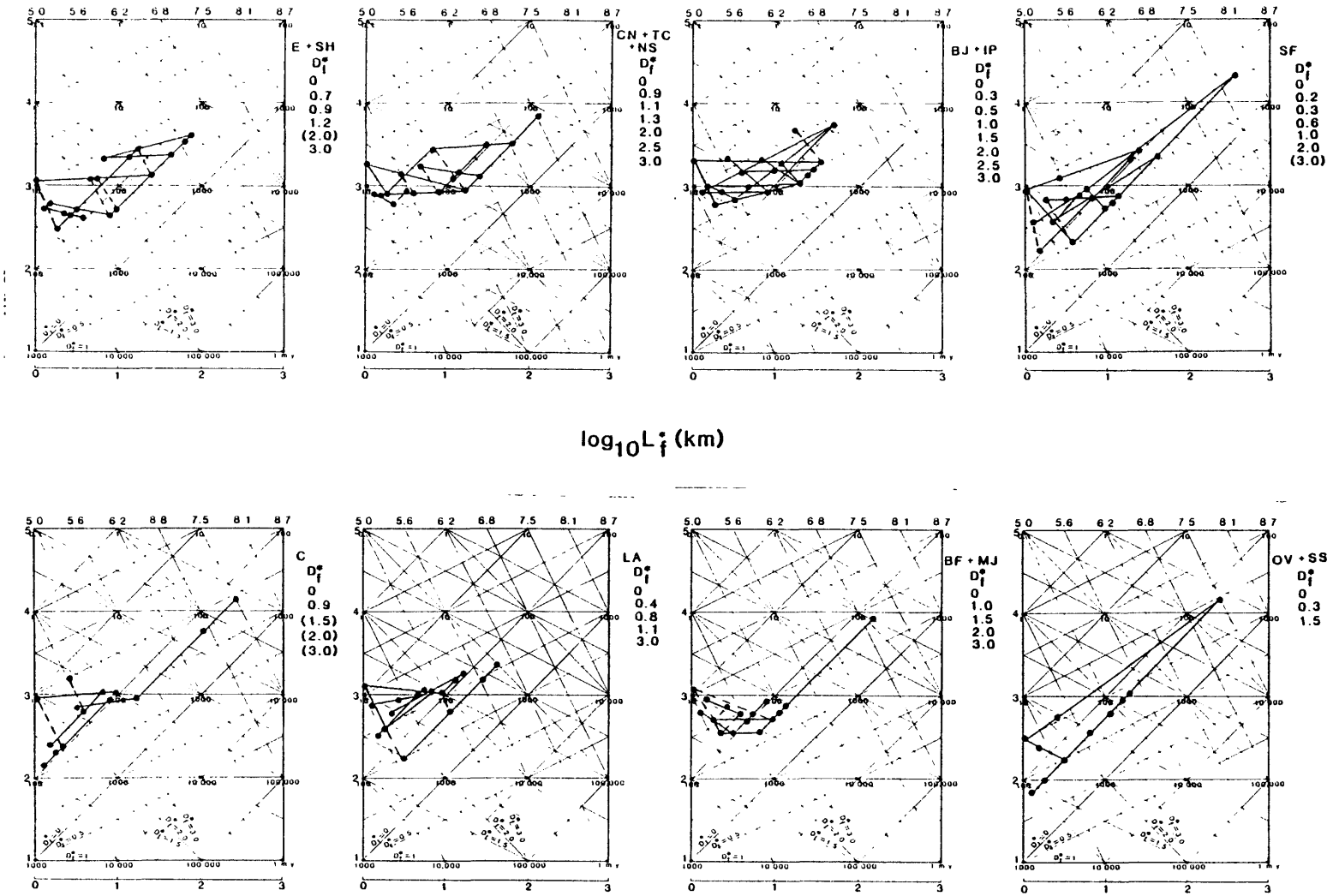
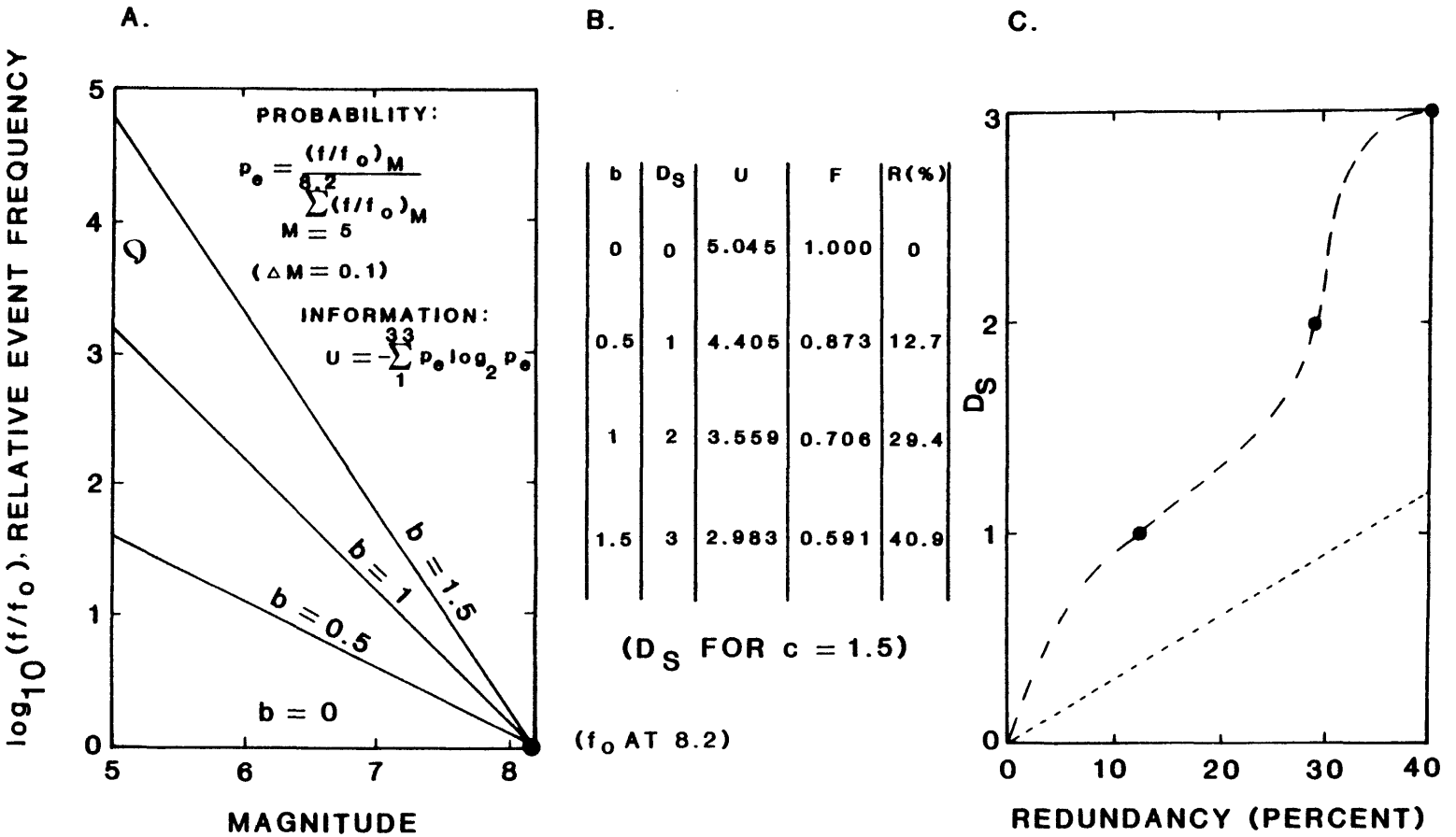


FIGURE 20. Idealized relations among b-values, fractal dimensions based on the definition $D_S = 3b/c$, and the information measures U (Shannon uncertainty), F (Shannon relative uncertainty), and R (Shannon redundancy) based on Equations (1) and (2) in text; $F = U_{obs}/U^*$, and $R = (1-F)x100$; R expressed in percent:

A. Idealized steady-state log f vs. M used to calculate values of U from Eq. (1).

B. Corresponding values of b, D_S, U, F, and R; i.e., each row shows a b-value in A. and the calculated values of the other parameters at that b-value. For this demonstration the value of the coefficient in the log M_O vs. M relation is assigned the constant value c = 1.5 (this is the usual assumption in the absence of specific evaluations); therefore in this special case $D_S = 2b$.

C. Trend of idealized fractal dimension, D_S, vs. the redundancy, R, from values in B, indicating possible deviations from idealized case. Light dashed line is trend if R were directly proportional to D_S in range 0 to 3.



b	D_S	U	F	R(%)
0	0	5.045	1.000	0
0.5	1	4.405	0.873	12.7
1	2	3.559	0.706	29.4
1.5	3	2.983	0.591	40.9

FIGURE 21. Conversion graph illustrating the relation between temporal frequencies and number frequencies in the idealized bond-count model for the same b-values used in Figure 20 (see text). Ordinate values at left give log N_B (for N_B the number of hypothetical "bonds") vs. M for the indicated b-values corresponding to temporal frequencies given by ordinate values on the right for events (as defined in previous illustrations) at the same b-values. Bond-count frequencies, N_B, are normalized so that there is one bond count at M = 5 for b = 0, and 1585 bond counts at M = 8.2 for b = 0; the reference points at M = 8.2 could be made the same for events and bond counts, but then there would be one bond count at M = 8.2 for b = 0 and 1/1585 count at M = 5 for b = 0. Other proportions vary according to the b-values of the corresponding curves; e.g., for b = 1 there is a constant count of N_b = 1585 for the normalization shown (if the reference at M = 8.2 were redefined so that log N_B = log f, there would then be one bond count at each magnitude for b = 1).

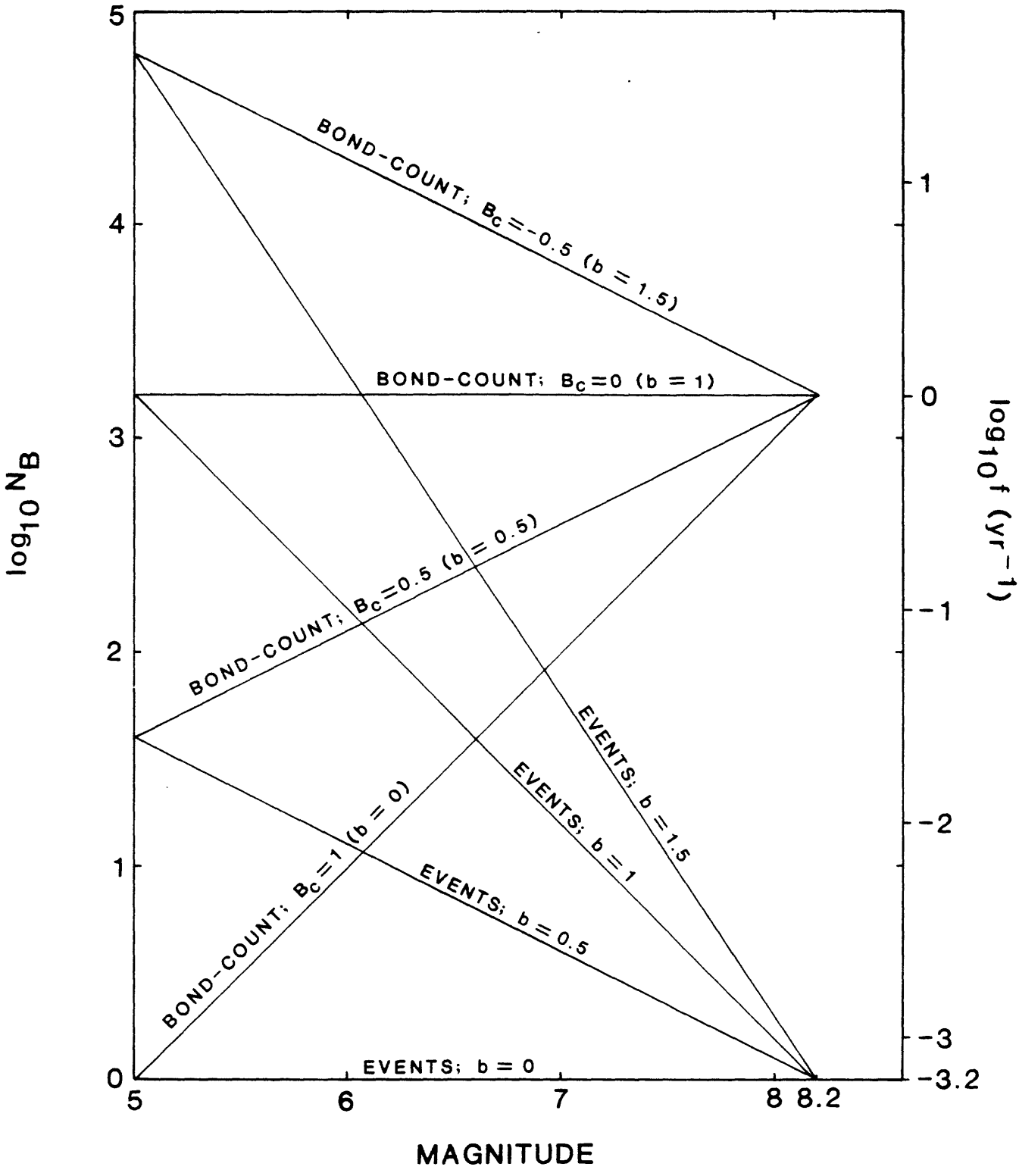

BOND-COUNT; B_c = -0.5 (b = 1.5)
BOND-COUNT; B_c = 0 (b = 1)
BOND-COUNT; B_c = 0.5 (b = 0.5)
BOND-COUNT; B_c = 1 (b = 0)
EVENTS; b = 1.5
EVENTS; b = 1
EVENTS; b = 0.5
EVENTS; b = 0
$\log_{10} N_B$
$\log_{10} f\ (yr^{-1})$
MAGNITUDE
5
6
7
8
8.2
0
1
2
3
4
5
1
0
-1
-2
-3
-3.2

FIGURE 22. Bond-count redundancy, R_B, plotted against fractal dimension, D_S, using the same relations employed in Figure 20 (c = 1.5) but based on probabilities calculated from bond counts in Figure 21 rather than on probabilities calculated from event frequencies as in Figure 20. The extremum at $D_S = 2$ at b = 1 occurs because this is where the slope of log N_B *vs*. M changes sign in Figure 21 (by contrast the plot of event redundancy, R_E, against D_S in Figure 20 is monotonic because the variations of log f *vs*. M have slopes of the same sign).

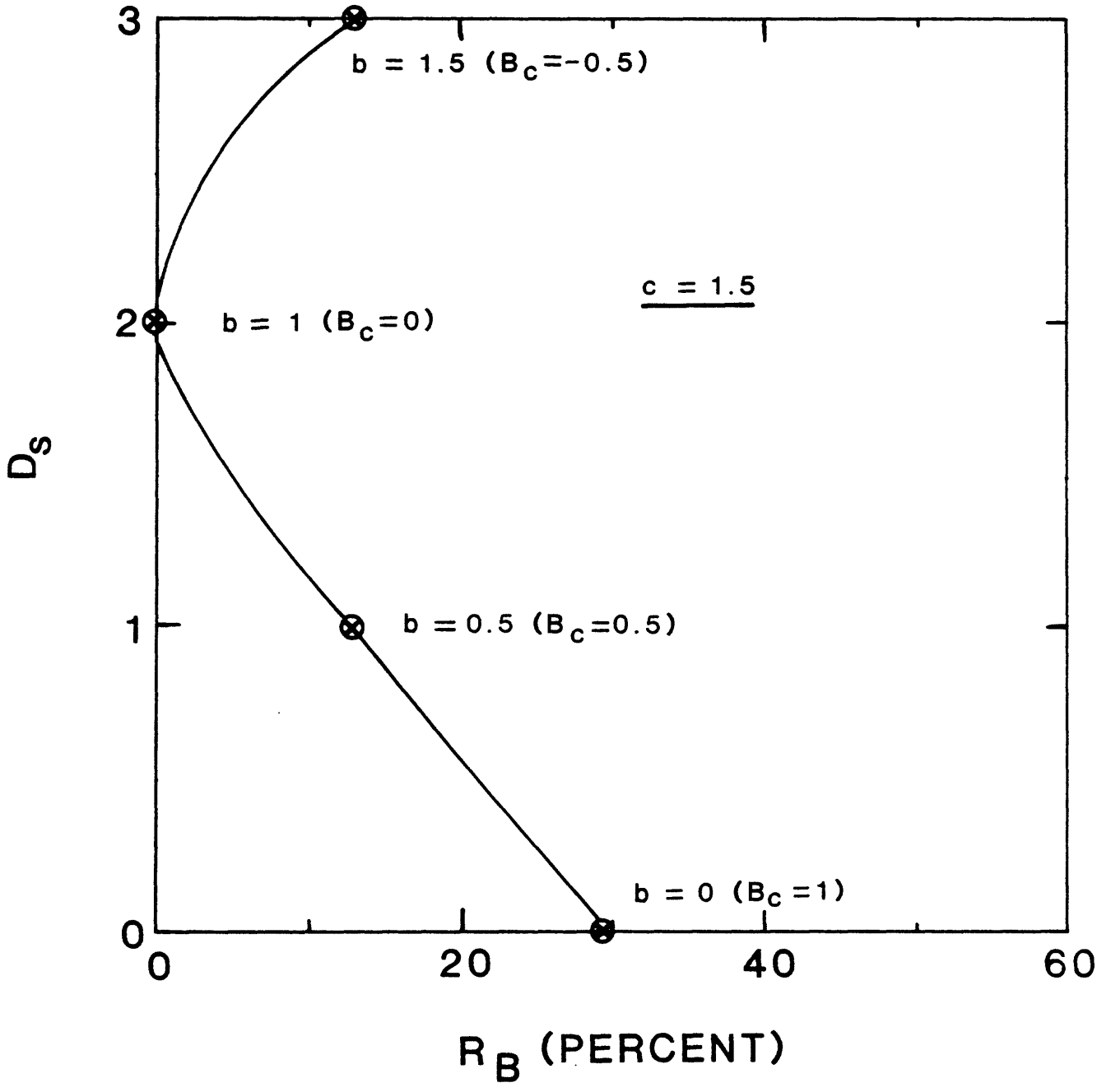

3
2
1
0
D_S
b = 1.5 (B_c=-0.5)
c = 1.5
b = 1 (B_c=0)
b = 0.5 (B_c=0.5)
b = 0 (B_c =1)
0
20
40
60
R_B (PERCENT)

FIGURE 23. Bond-count relative uncertainty, $F_B = (1- R_B/100)$, plotted against fractal dimension, D_s, for the same b-values used in Figures 20 and 22. F_B is shown on the ordinate so that the extremum at $R_B = 0$ in Figure 22 is now a maximum at $F_B = 1$, $b = 1$. This point defines the maximum uncertainty relative to a convention where there is equal probability that the same numbers of bond ruptures will occur at every magnitude within the 33 magnitude increments between M = 5 and M = 8.2 (i.e., this condition corresponds to the maximum entropy model based on bond counts for the delimitation of frequency-magnitude relations in Figure 20). Values of F_B calculated in the same way for the observed data within variously defined regions of Figure 14 and time windows are shown by horizontal lines (data from Table 4); time windows of areally total data sets are shown in parentheses. The fact that most of the partial data sets do not intersect the solid portions of the idealized curve is taken to reflect incompleteness of the seismic record and(or) deviations of the actual rupture statistics from the idealized definitions of F_B and D_s (i.e., a model might exist that would permit the full ranges of F_B and D_s for the observed ranges of average b-values found in frequency-magnitude plots). Note different informational implications for conventions based on the statistics of bond counts vis-a-vis events: on the event basis, high fractal D_s, high b-values, and high event redundancy, R_E, go together (this redundancy refers to the time domain because there is no spatial information in the frequency-magnitude data for events); the bond-count basis gives nonunique D_s and b (they are not single-valued) at a given value of uncertainty, F_B, but as a general tendency high uncertainty and fractal D_s go with higher b-values while the highest redundancy (lowest F_B) is at $b = 0$, $D_s = 0$. These quantities are also inferred from temporal data, but they contain geographic information related to implicit clustering of bond-count distributions at each magnitude. In this context the largest events have the greatest redundancies (lowest uncertainties, compared with highest uncertainties on the event basis) and the greatest possible paired ranges of D_s and b-values; these limits mimic the geographic invariance of major seismic gaps which simultaneously display high temporal uncertainty.

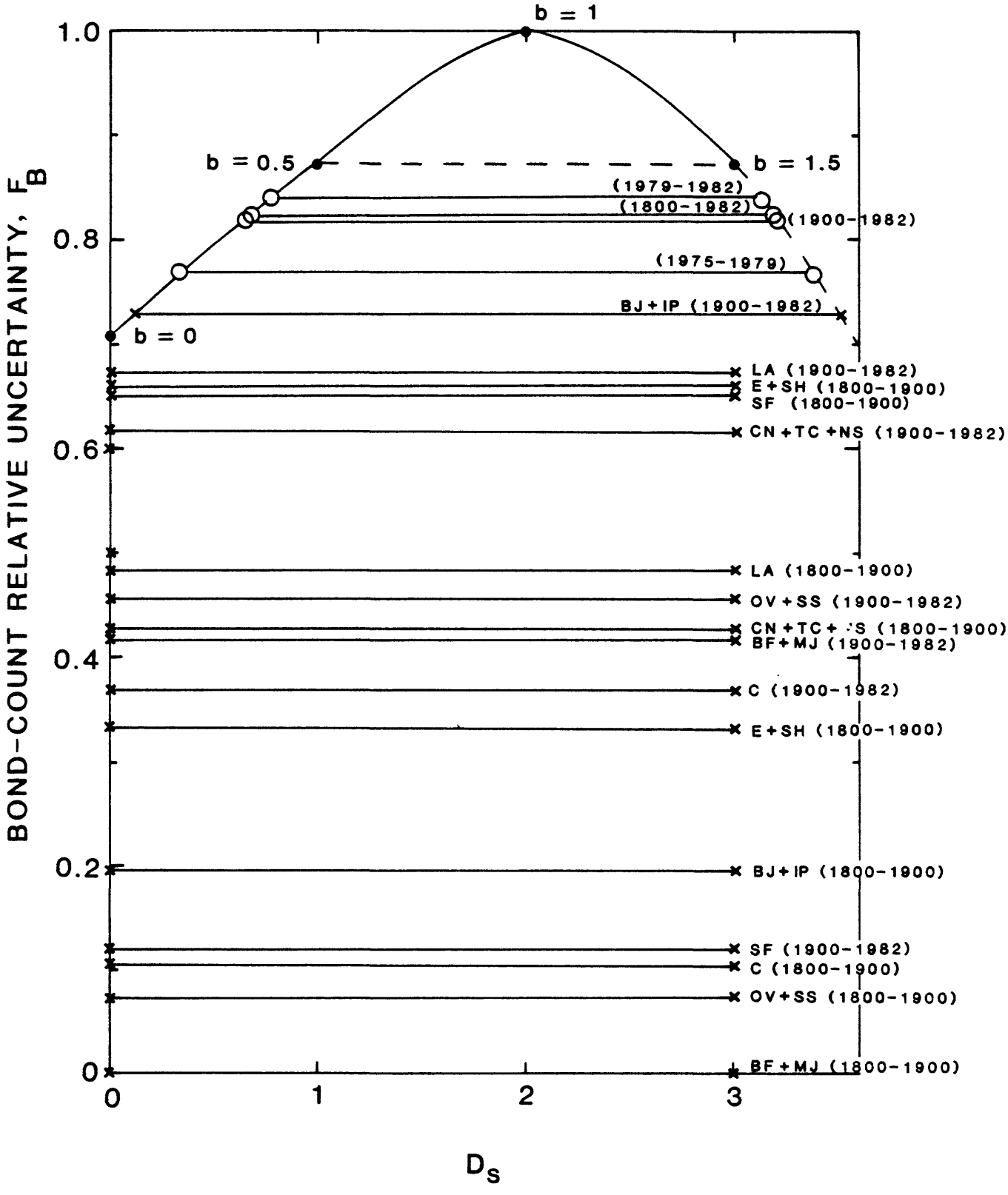
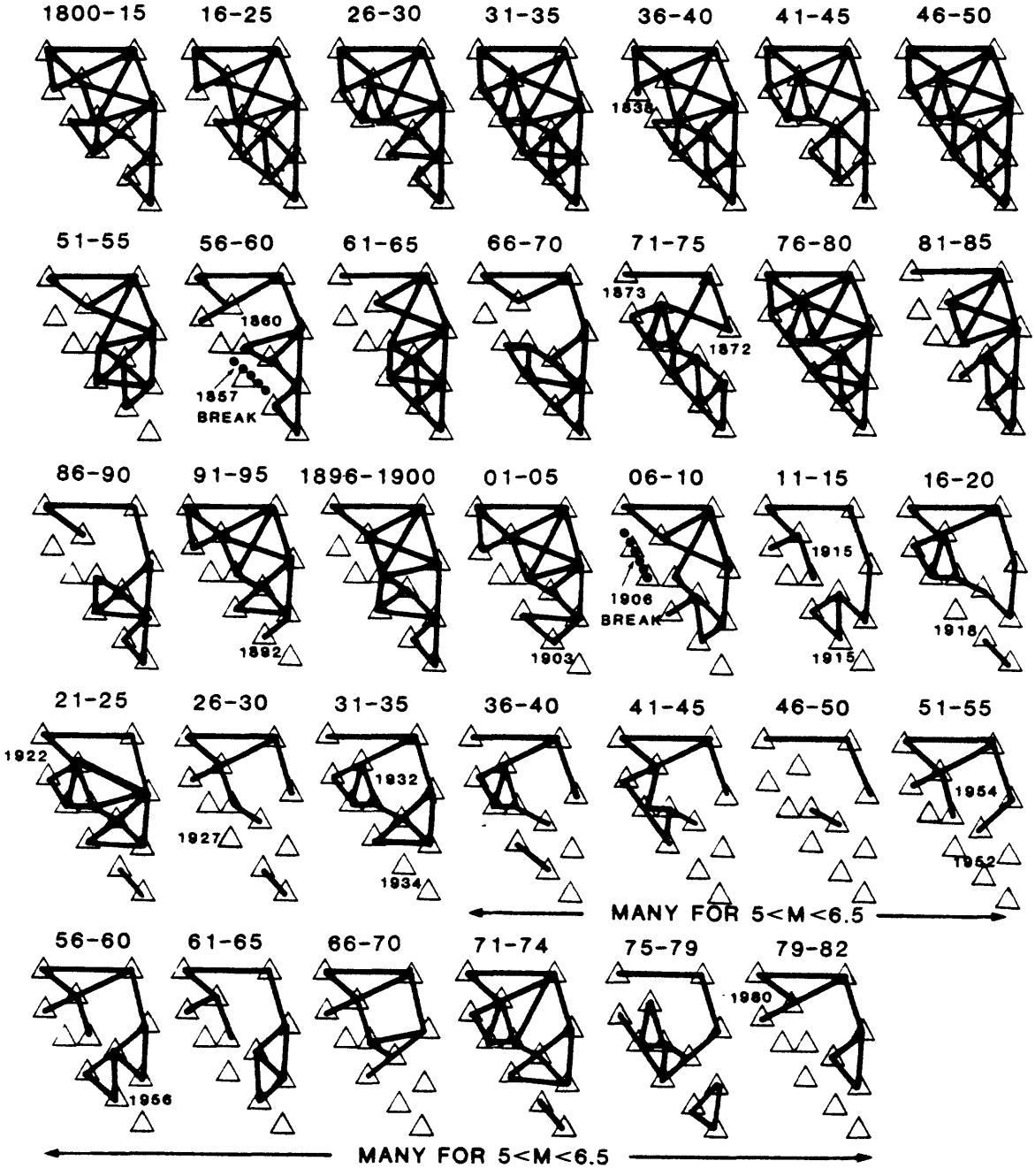

b = 1
b = 0.5
b = 1.5
b = 0
(1979-1982)
(1800-1982)
(1900-1982)
(1975-1979)
BJ+IP (1900-1982)
LA (1900-1982)
E+SH (1800-1900)
SF (1800-1900)
CN+TC+NS (1900-1982)
LA (1800-1900)
OV+SS (1900-1982)
CN+TC+NS (1800-1900)
BF+MJ (1900-1982)
C (1900-1982)
E+SH (1800-1900)
BJ+IP (1800-1900)
SF (1900-1982)
C (1800-1900)
OV+SS (1800-1900)
BF+MJ (1800-1900)
1.0
0.8
0.6
0.4
0.2
0
0
1
2
3
BOND-COUNT RELATIVE UNCERTAINTY, F_B
D_S

FIGURE 24. Kinematic linkage maps drawn at intervals of five years based on arbitrary subregions and benchmarks of Figure 14 and event statistics compiled in Table 2 and Figure 18. A "link" is defined as any one of the tielines between benchmarks of Figure 14 that remains "unbroken" by the occurrence of an event of M = 5 or larger in any part of a subregion lying near that line; the years of occurrence and approximate locations of the largest events having M = 6.7 or larger are also shown. Kinematically, the large events emanate from windows where there was prior development of linkage "breaks": the 1857 San Andreas fault reflects convergences of linkage breaks emanating from northwestern (subregion E) and southern sources (subregion BJ); the 1872 Owens valley fault break reflects convergences of northwestern (subregion E) and northeastern (subregion CN) sources; and the 1906 San Andreas fault break again reflects the northwestern and southern source effects (the different locations on the fault appear to be influenced by the relative contributions of these sources, possibly modified by differing contributions from the northeasterly source region). Inset (bottom): Patterns of coalescing breaks based on a kinetic model of physical bond ruptures developed by Termonia and Meakin (1986, Figure 2); although there are no direct correspondences of dynamic scaling, the implications of areally distributed deformations and coalescing fault breaks are analogous. The fractal dimension of the ruptures determined by these authors is about 1.27, a value similar to estimates of fault fractal D_f by Shaw and Gartner (1986) for active faults in California. Such comparisons suggest that more detailed linkage maps at progressively larger map scales, and based on occurrences of progressively smaller earthquakes, may provide significant tests for self-similarity and space-time patterns of migrating loci of fault ruptures.


1800-15
16-25
26-30
31-35
36-40
1838
41-45
46-50
51-55
56-60
1860
1857
BREAK
61-65
66-70
71-75
1873
1872
76-80
81-85
86-90
91-95
1892
1896-1900
01-05
1903
06-10
1906
BREAK
11-15
1915
1915
16-20
1918
21-25
1922
26-30
1927
31-35
1932
1934
36-40
41-45
46-50
51-55
1954
1952
MANY FOR 5<M<6.5
56-60
1956
61-65
66-70
71-74
75-79
79-82
1980
MANY FOR 5<M<6.5

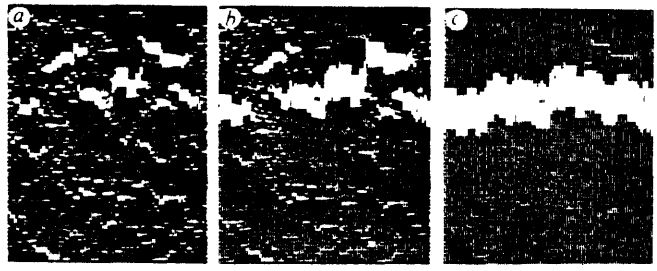

a
b
c

FIGURE 25. Space-time diagrams of migration of seismic activity in California based on patterns of. linkage continuity in Figure 24 arranged according to the numbering of tielines between the arbitrary benchmarks of Figure 14:

A. Paragenetic lines; blank portions represent times when the corresponding tielines have been broken by an event of M = 5 or larger. The times of occurrences of the six major events of Table 2 and Figure 18 are shown for reference.

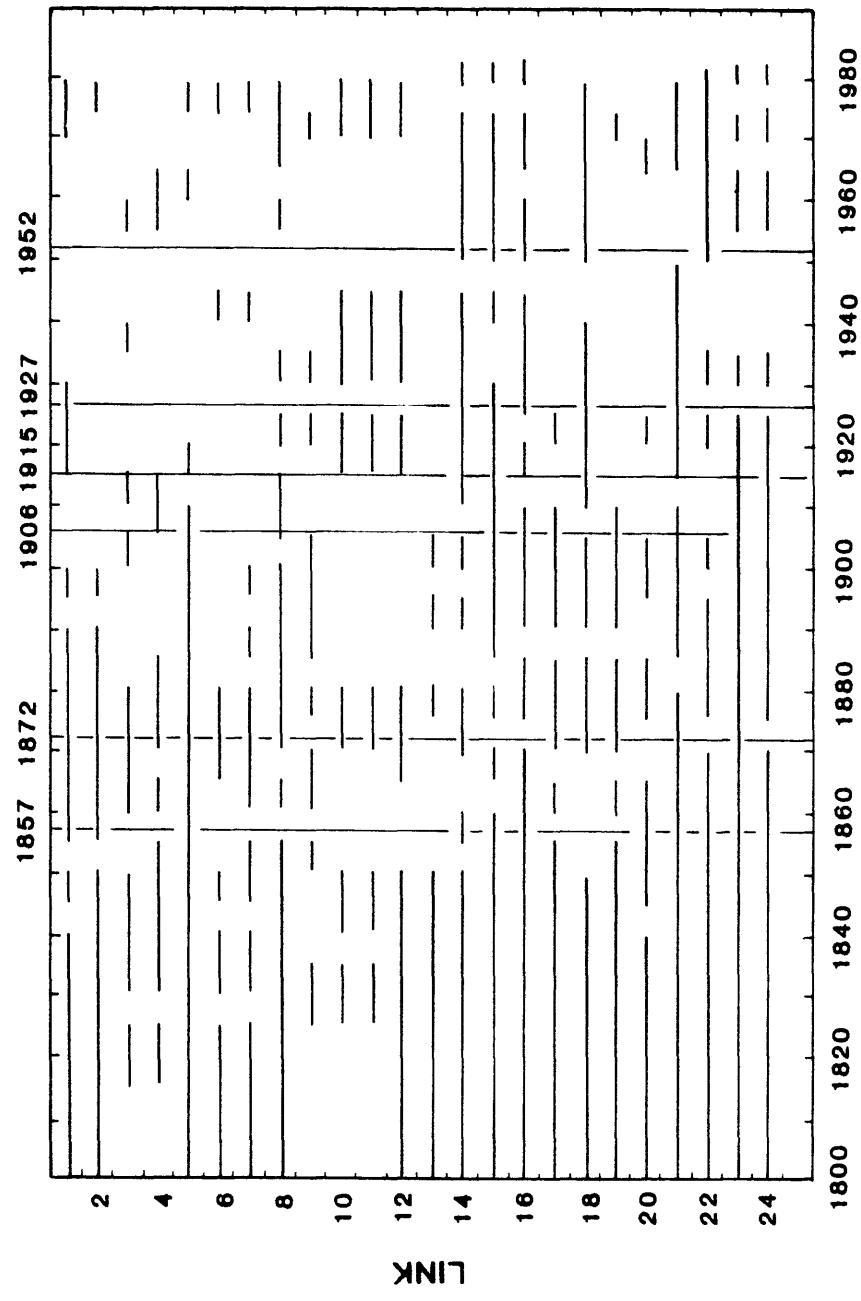

1857
1872
1906
1915
1927
1952
1800
1820
1840
1860
1880
1900
1920
1940
1960
1980
2
4
6
8
10
12
14
16
18
20
22
24
LINK

FIGURE 25. Space-time diagrams of migration of seismic activity in California based on patterns of linkage continuity in Figure 24 arranged according to the numbering of tielines between the arbitrary benchmarks of Figure 14:

B. Diagram in A modified by enclosing space-time regions within which linkage breaks were continuous. Events of M = 8 or larger are shown by large stars, events from M = 7.3 to M = 7.7 by solid dots, and events near M = 7 by open circles and dots.

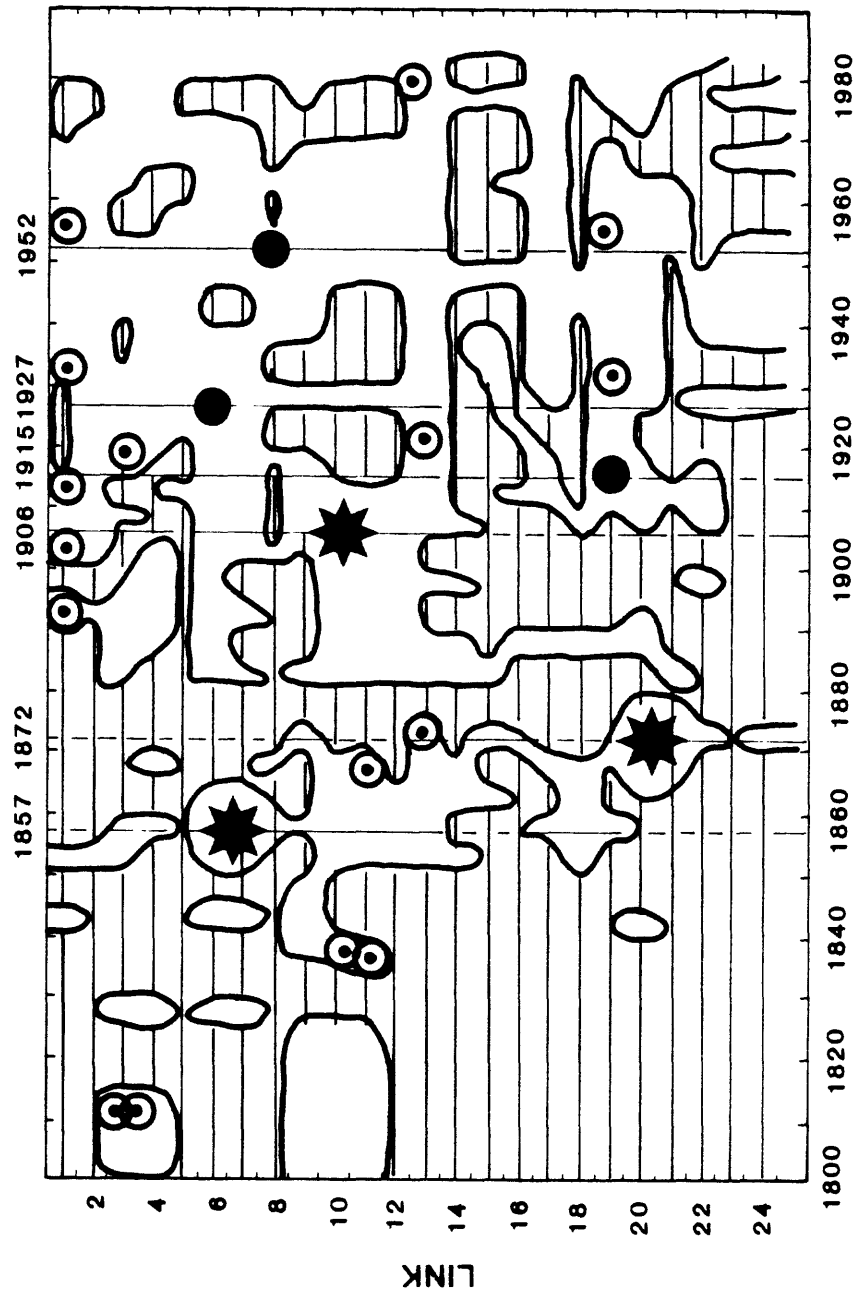

1857
1872
1906
1915
1927
1952
1800
1820
1840
1860
1880
1900
1920
1940
1960
1980
2
4
6
8
10
12
14
16
18
20
22
24
LINK

FIGURE 25. Space-time diagrams of migration of seismic activity in California based on patterns of linkage continuity in Figure 24 arranged according to the numbering of tielines between the arbitrary benchmarks of Figure 14:

C. Pattern in B modified to schematically illustrate trajectories of migrating seismic moments across subregion boundaries with time relative to distributions of the larger earthquakes. In a manner analogous to the effects of constrictions produced by boulders and promontories in rivers at low water, the large events represent loci of convergence relative to "upstream" trajectories and divergence relative to "downstream" trajectories (there is a physical analogy between such hydrodynamic stationary points and singular portions of faults where strain energy becomes concentrated). Inferences from these schematic trajectories for "regional forecasts" are shown at right margin (Note: Spatial relations can be inferred by comparison with Figure 14. Because two space dimensions are projected onto one space dimension represented by one of the numbered tielines, such diagrams can be ordered in a variety of ways. Trial constructions in which the sequences of tielines were rearranged, however, did not change the general patterns and inferences. Actual trajectories in three dimensions representing latitude, longitude, and time might be illustrated using stereo pairs as is done in three-dimensional representation of convective particle trajectories or structures of complex molecules.)

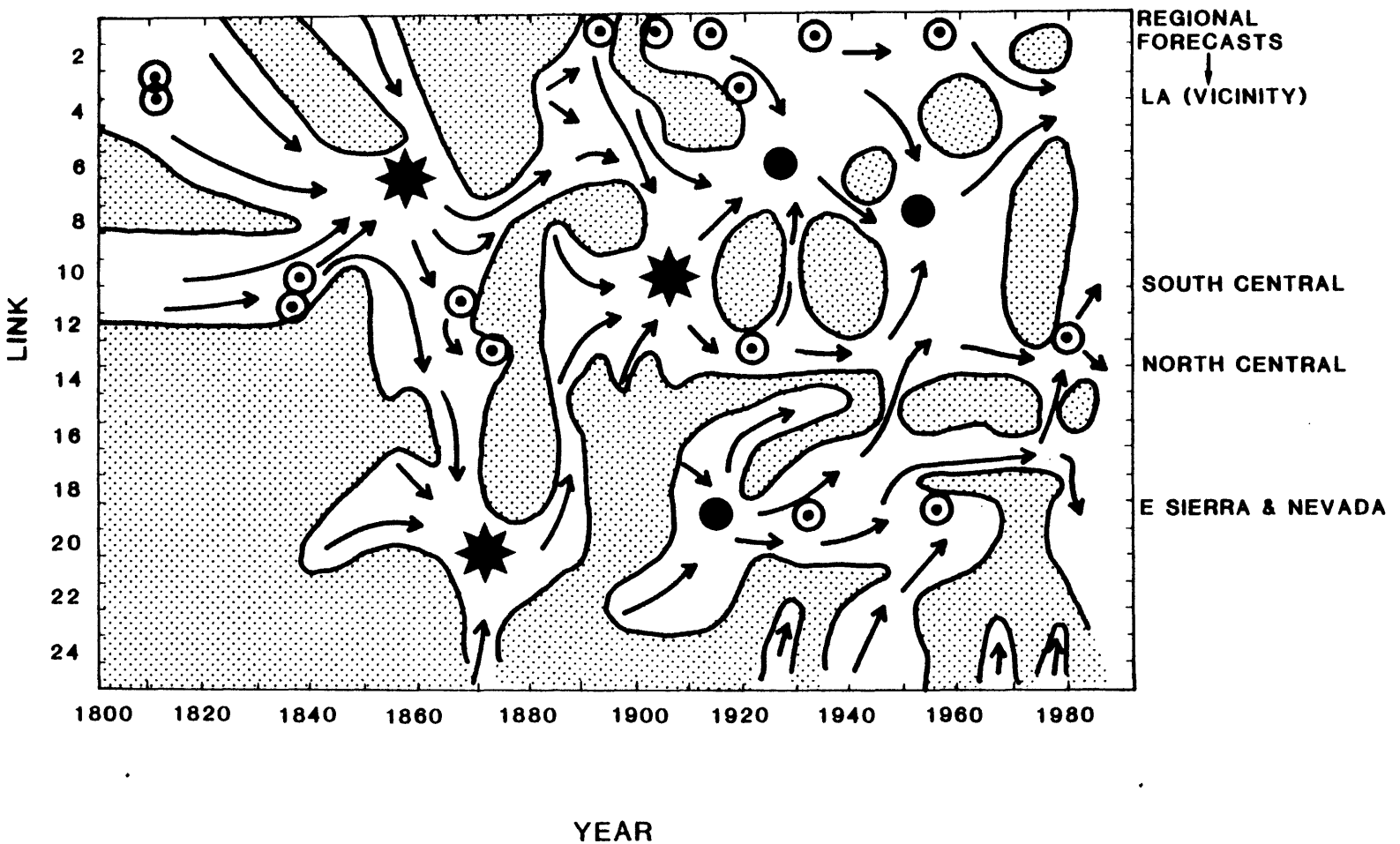

REGIONAL FORECASTS
LA (VICINITY)
SOUTH CENTRAL
NORTH CENTRAL
E SIERRA & NEVADA
LINK
2
4
6
8
10
12
14
16
18
20
22
24
1800
1820
1840
1860
1880
1900
1920
1940
1960
1980
YEAR

FIGURE 26. Time variation of moment rate for California earthquake history based on data compiled in Table 2. The rate is based on running sums of moments per five year interval divided by five. This is a histogram of discontinuous data shown connected by the heavy dashed lines to emphasize the variations (i.e., minima below 10^{23} dyne-cm yr^{-1} are not shown). The solid curve in upper right shows the trend of post-1900 Circumpacific moment rates based on the same time intervals and shown on the same ordinate scale (redrawn from Kanamori, 1978). Global seismic moment rates are more than an order of magnitude higher because they are dominated by convergent plate margin events any one of which may exceed the entire moment for historic seismicity in California; see Table 2B). Open circles and light lines show order-of-magnitude estimates of global thermal energy release of major volcanic eruptions, redrawn from Kanamori (1983, Figure 5); volcanic energy is in units of erg yr^{-1} on same ordinate scale (i.e., values are one-tenth of Kanamori's 10-year estimates). The vertical bar at right margin indicates orders of magnitude of thermal energy transport by magma (erg yr^{-1}): GMM is estimated global rate of magma production from Shaw (1970, 1980); GSM is estimated upper limit for secular rate of silicic magma production based on 10:1 ratio to mafic magma, after Smith and Shaw (1975) and Shaw (1985); HM is approximate secular average for Hawaiian volcanism from Shaw (1980, in press); HEM is 74-m.y. average rate for volcanic edifices of the Hawaiian-Emperor volcanic chain, from Shaw (1985); and SM is estimate of secular average rate of silicic magma production fed by a mafic source equivalent to HEM, from Smith (1979) and Shaw (1985). Note that rates of seismic energy release are more than four orders of magnitude smaller than moment rates (Kanamori, 1978, Figure 2), hence they are comparable to range of Kanamori's estimated eruptive energies, and smaller than either GMM or GSM by zero to four orders of magnitude. See text for remarks on post-1950 volcanicity.

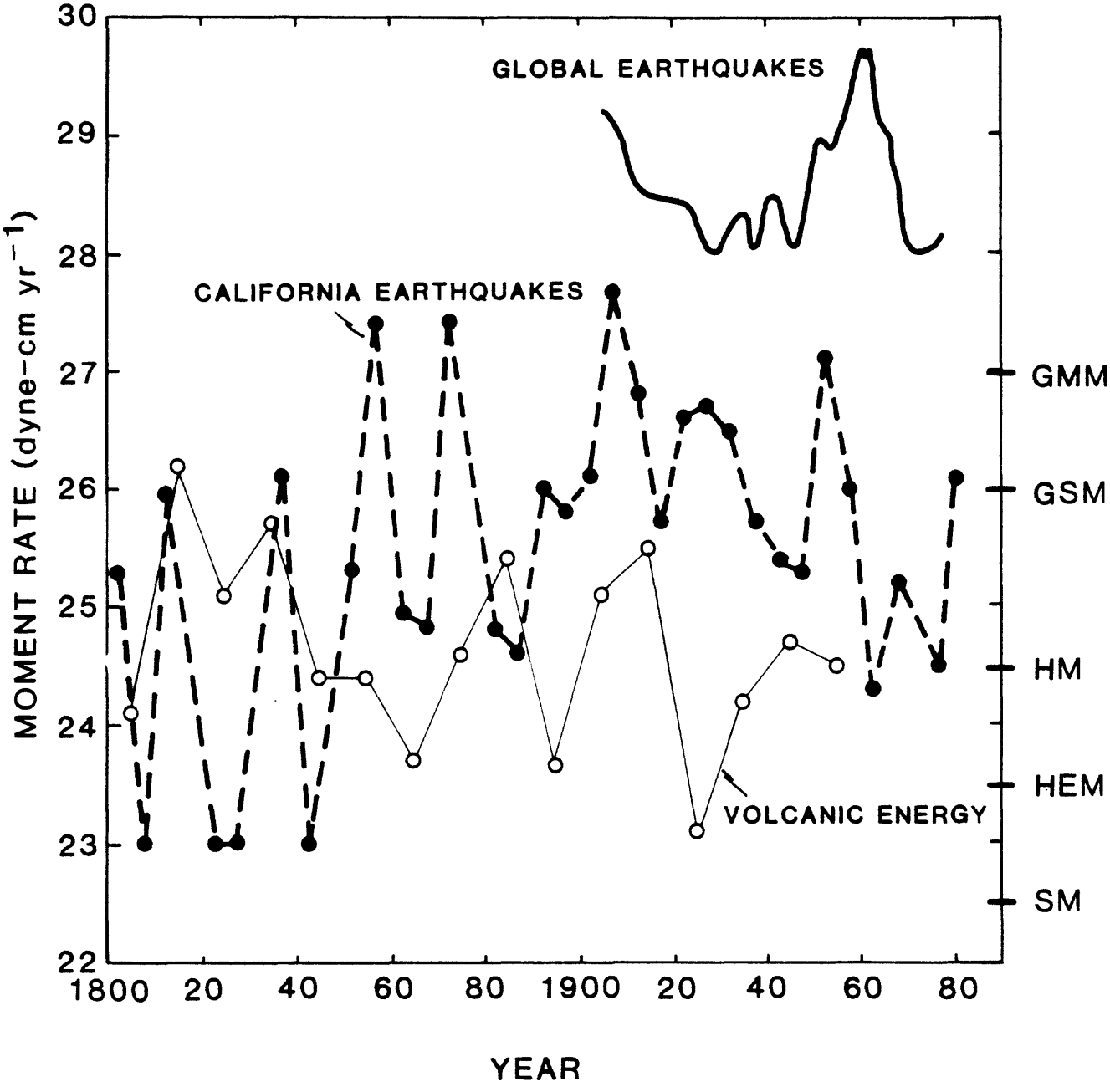

GLOBAL EARTHQUAKES
CALIFORNIA EARTHQUAKES
VOLCANIC ENERGY
GMM
GSM
HM
HEM
SM
MOMENT RATE (dyne-cm yr^{-1})
YEAR
30
29
28
27
26
25
24
23
22
1800
20
40
60
80
1900
20
40
60
80

FIGURE 27. Five-year increments of moments from Table 2 showing trends of moment rates for several different geographic distributions in Figure 14 (symbols are plotted at midpoints of five-year intervals up to 1975 and thereafter approximately at midpoints of the 1975-1979 and 1979-1982 intervals; absence of a symbol in any interval represents a gap of five years or longer in the data):

(a) northern vs. southern California, represented by subregions E+SH+SF+NS+TC+CN+C (open circles and dashed lines) vs. SS+OV+BF+MJ+LA+IP+BJ (dots and solid lines);

(b) San Francisco and neighboring vicinities of central and northern California represented by SF+C (open circles and dashed lines) vs. Los Angeles and neighboring regions represented by LA+BF (pluses and solid lines);

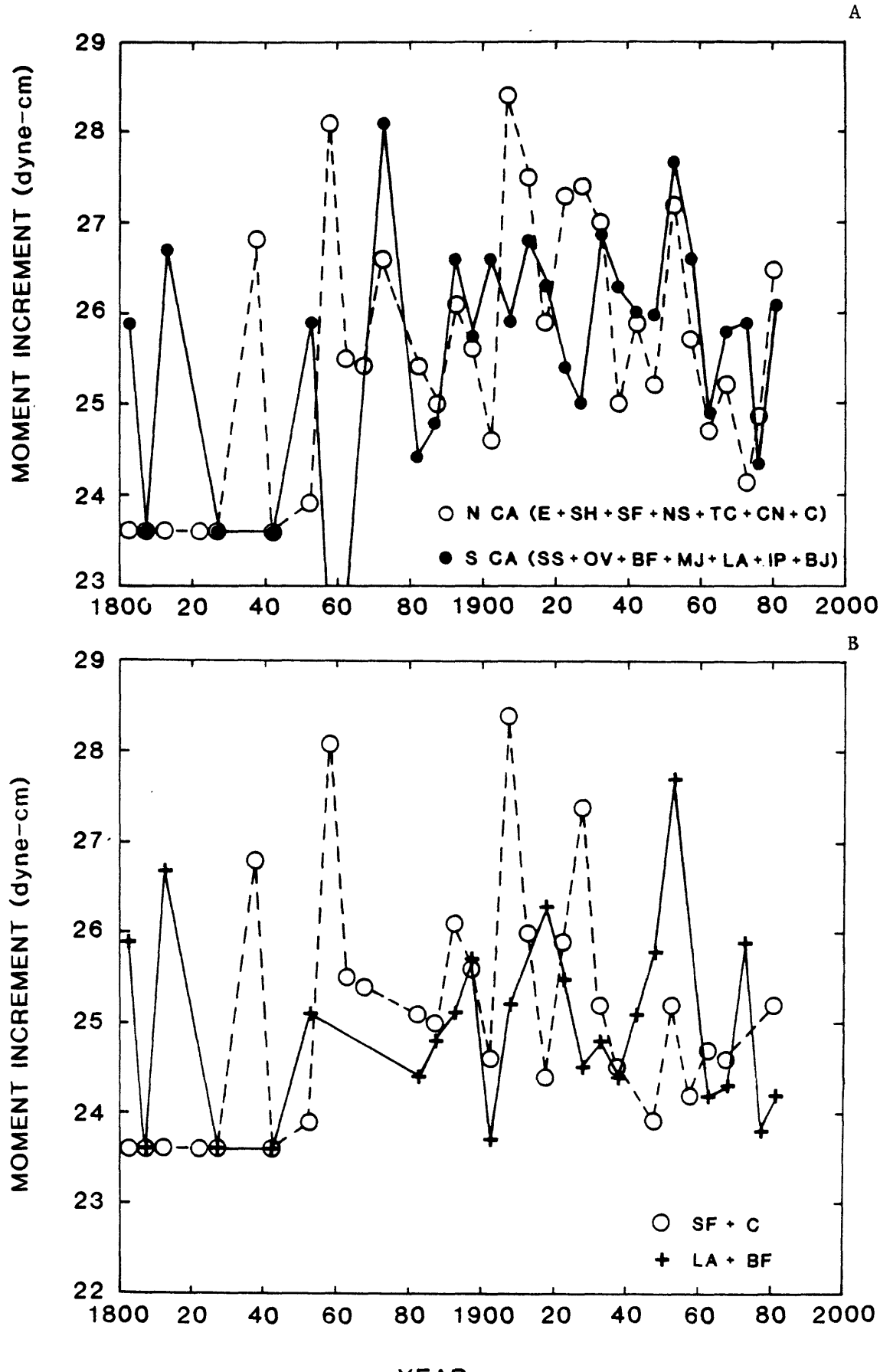

A
MOMENT INCREMENT (dyne-cm)
29
28
27
26
25
24
23
1800
20
40
60
80
1900
20
40
60
80
2000
N CA (E + SH + SF + NS + TC + CN + C)
S CA (SS + OV + BF + MJ + LA + IP + BJ)
B
MOMENT INCREMENT (dyne-cm)
29
28
27
26
25
24
23
22
1800
20
40
60
80
1900
20
40
60
80
2000
SF + C
LA + BF
YEAR

FIGURE 27. Five-year increments of moments from Table 2 showing trends of moment rates for several different geographic distributions in Figure 14 (symbols are plotted at midpoints of five-year intervals up to 1975 and thereafter approximately at midpoints of the 1975-1979 and 1979-1982 intervals; absence of a symbol in any interval represents a gap of five years or longer in the data):

(c) five composite regions of possible tectonic affinities consisting of subregions as labeled on the diagrams.

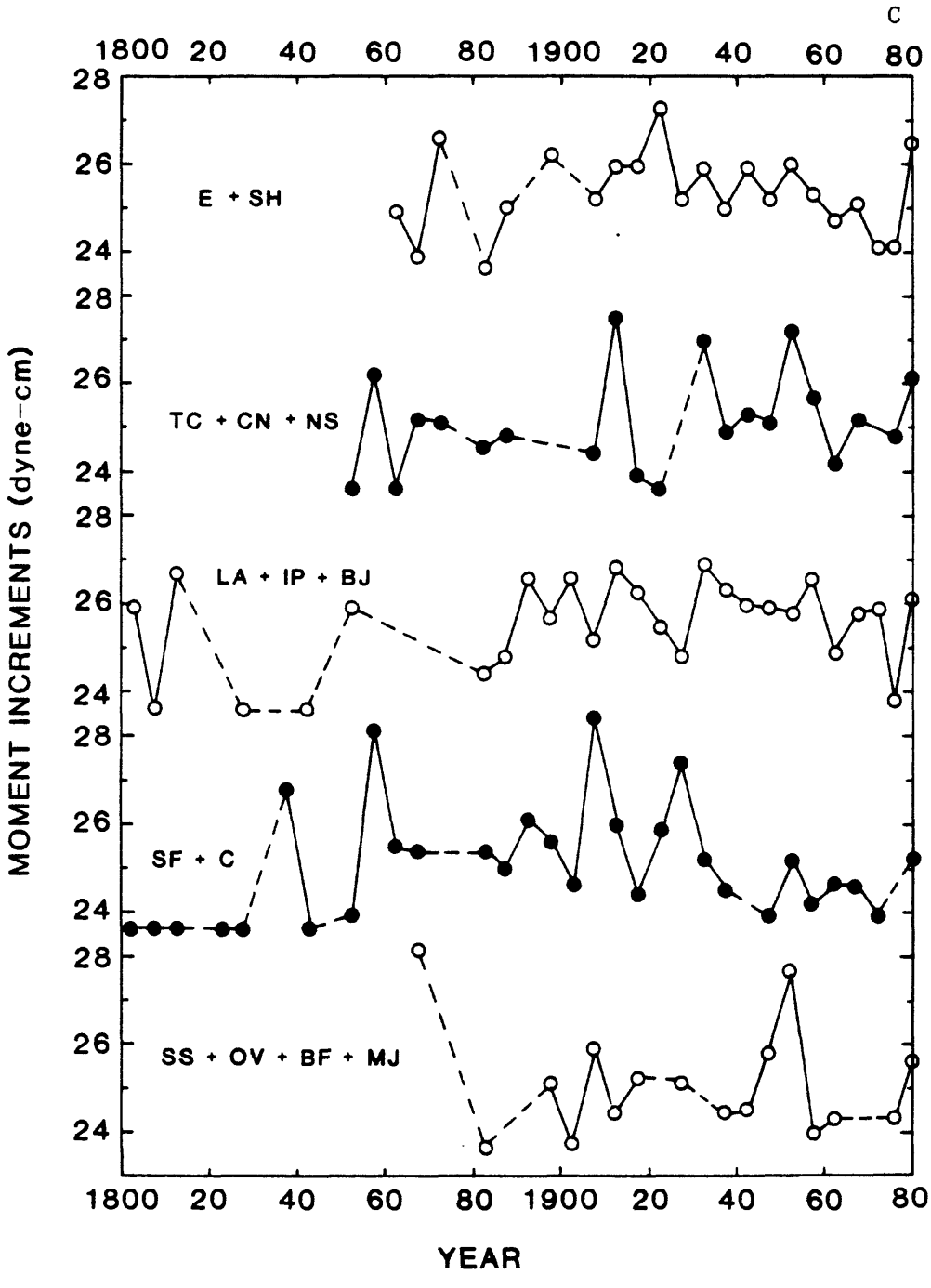

C
1800 20 40 60 80 1900 20 40 60 80
E + SH
TC + CN + NS
LA + IP + BJ
SF + C
SS + OV + BF + MJ
28
26
24
MOMENT INCREMENTS (dyne-cm)
YEAR

TABLE 1. Relations of magnitude and moment to the statistically derived seismic length scale L^*_s.

	q	c	qc	M_o (a)	log L^*_s (b)
(1)	2/3	3/2	1	L^*	M/q = cM = 3M/2
(2)	1	3/2	3/2	$L^{*3/2}$	M/q = M
(3)	4/3	3/2	2	L^{*2}	M/q = cM/2 = 3M/4
(4)	2	3/2	3	L^{*3}	M/q = cM/3 = M/2

(a) Entries give the function of L^*_s that is proportional to seismic moment, M_o, for the indicated values of q and c according to Eq. 6b for logarithms to Base 10.

(b) Entries give the equivalent functions of magnitude, M, that are proportional to a logarithmic length scale in Eq. 6a consistent with c in Eq. 6b and this table (i.e., q defined by $L^*_s = 10^{M/q}$).

TABLE 2. A. Counts at intervals of five years for earthquakes with magnitudes equal to or greater than M = 5 distributed by subregion in Figure 14.[a]
B. Logarithms of moments from $\log_{10} M_o = 16.1 + 1.5\ M$ (Hanks and Kanamori, 1979; Shaw and Gartner, 1986)) for events in A.

Time Interval[b]
Subregion[c]
A. Counts:

	E	SH	SF	NS	TC	CN	C	SS	OV	BF	MJ	LA	IP	BJ
1800-05	-	-	1	-	-	-	-	-	-	-	-	2	-	-
1806-10	-	-	1	-	-	-	-	-	-	-	-	1	-	-
1811-15	-	-	1	-	-	-	-	-	-	-	-	2*	-	-
1816-20	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1821-25	-	-	1	-	-	-	-	-	-	-	-	-	-	-
1826-30	-	-	1	-	-	-	-	-	-	-	-	1	-	-
1831-35	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1836-40	-	-	2**	-	-	-	-	-	-	-	-	-	-	-
1841-45	-	-	1	-	-	-	-	-	-	-	-	1	-	-
1846-50	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1851-55	-	-	2	1	-	-	-	-	-	-	-	1	-	1
1856-60	-	-	3	1	-	1*	1●	-	-	-	-	-	-	-
1861-65	1	-	7	1	-	-	1	-	-	-	-	-	-	-
1866-70	2	-	2*	2	-	3	-	-	-	-	-	-	-	-
1871-75	2*	-	-	1	-	-	-	-	5●	-	-	-	-	-
1876-80	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1881-85	1	1	4	-	1	-	4	1	-	-	-	1	-	-
1886-90	3	-	5	1	1	1	2	-	-	-	-	3	-	-
1891-95	-	-	7	-	-	-	1	-	-	-	-	3	-	2*
1896-1900	2*	-	3	-	-	-	3	-	1	-	-	2	-	-
1901-05	-	-	3	-	-	-	2	-	-	2	-	-	-	1*
1906-10	3	-	1●	2	-	-	2	-	1	-	1	4	-	1
1911-15	1	-	2	-	1●	3	1	1	1	-	-	-	2	1*
1916-20	1	-	-	-	-	1	1	-	-	3	1	3*	2	1
1921-25	4**	-	-	-	1	-	2	-	-	-	-	2	-	3
1926-30	4	-	1	-	-	-	6●	1	1	1	1	6	4	1
1931-35	6	-	-	-	-	5*	5	-	-	-	-	14	3	5*
1936-40	5	1	-	-	-	5	4	1	2	1	4	1	8*	6
1941-45	6	-	-	1	1	9	-	-	1	1	4	5	7	3
1946-50	4	2	1	1	1	1	1	-	1	7	10	3	4	3
1951-55[d]	5	-	2	1	-	33***	4	-	-	24●	-	1	7	14
1956-60	5	-	1	-	-	7	3	-	1	-	-	-	3	31*
1961-65	6	-	1	-	2	3	2	2	-	4	1	2	-	4
1966-70	3	-	3	2	-	-	5	-	-	-	-	3	3	6
1971-74[e]	4	-			-	-	1	-	-	-	-	7	1	-
1975-79[f]	1	3	-	-	1	2	-	-	-	-	3	1	-	-
1979-82[g]	5*	-	4	1	-	17	1	1	-	1	-	1	7	1
Sums: 603[h]	74	7	60	14	10	91	52	7	14	44	25	70	51	84

TABLE 2 B. Logarithms of moments from $\log_{10} M_o = 16.1 + 1.5\ M$ (Hanks and Kanamori, 1979; Shaw and Gartner, 1986)) for events in A.

B. $\log M_o^{(i)}$:

		E	SH	SF	NS	TC	CN	C	SS	OV	BF	MJ	LA	IP	BJ
1800-05	25.9	-	-	23.6	-	-	-	-	-	-	-	-	25.9	-	-
1806-10	23.7	-	-	23.6	-	-	-	-	-	-	-	-	23.6	-	-
1811-15	26.7	-	-	23.6	-	-	-	-	-	-	-	-	26.7	-	-
1816-20	---	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1821-25	23.6	-	-	23.6	-	-	-	-	-	-	-	-	-	-	-
1826-30	23.7	-	-	23.6	-	-	-	-	-	-	-	-	23.6	-	-
1831-35	---	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1836-40	26.8	-	-	26.8	-	-	-	-	-	-	-	-	-	-	-
1841-45	23.7	-	-	23.6	-	-	-	-	-	-	-	-	23.6	-	-
1846-50	---	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1851-55	26.0	-	-	23.9	23.6	-	-	-	-	-	-	-	25.1	-	25.9
1856-60	28.1	-	-	25.4	24.1	-	26.2	28.1	-	-	-	-	-	-	-
1861-65	25.6	24.8	-	25.5	23.6	-	-	24.8	-	-	-	-	-	-	-
1866-70	25.5	23.9	-	25.4	23.9	-	25.2	-	-	-	-	-	-	-	-
1871-75	28.1	26.6	-	-	-	25.1	-	-	-	28.1	-	-	-	-	-
1876-80	---	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1881-85	25.5	23.6	24.0	25.1	-	24.5	-	25.4	23.6	-	-	-	24.4	-	-
1886-90	25.3	25.0	-	24.5	24.8	24.4	24.5	25.0	-	-	-	-	24.8	-	-
1891-95	26.7	-	-	26.1	-	-	-	23.9	-	-	-	-	25.1	-	26.6
1896-1900	26.5	26.2	-	25.6	-	-	-	25.2	-	25.1	-	-	25.7	-	-
1901-05	26.6	-	-	24.6	-	-	-	24.5	-	-	23.7	-	-	-	26.6
1906-10	28.4	25.2	-	28.4	24.4	-	-	24.4	-	24.4	-	25.9	25.2	-	25.1
1911-15	27.6	25.9	-	26.0	-	27.5	25.7	24.4	23.6	24.4	-	-	-	25.5	26.8
1916-20	26.4	25.9	-	-	-	-	23.9	24.4	-	-	25.2	24.4	26.3	24.4	23.6
1921-25	27.3	27.3	-	-	-	23.6	-	25.9	-	-	-	-	25.5	-	24.4
1926-30	27.4	25.2	-	24.4	-	-	-	27.4	25.1	24.4	23.6	24.4	24.5	24.8	23.6
1931-35	27.2	25.9	-	-	-	-	27.0	25.2	-	-	-	-	24.8	23.8	26.9
1936-40	26.4	25.0	24.7	-	-	-	24.9	24.5	23.6	23.7	23.6	24.4	24.4	26.3	25.2
1941-45	26.1	25.9	-	-	24.1	23.6	25.3	-	-	23.6	23.9	24.5	25.1	26.0	24.5
1946-50	26.0	25.2	24.5	23.9	25.1	23.6	23.6	23.6	-	24.5	25.8	25.6	23.9	25.9	24.8
1951-55	27.8	26.0	-	24.5	23.6	-	27.2	25.2	-	-	27.7	-	25.0	25.4	25.8
1956-60	26.7	25.3	-	24.1	-	-	25.7	24.2	-	23.9	-	-	-	24.1	26.6
1961-65	25.0	24.7	-	23.6	-	24.2	-	24.7	24.2	-	24.2	23.6	23.7	-	24.9
1966-70	25.9	25.1	-	24.6	25.2	-	-	24.6	-	-	-	-	24.3	25.8	25.6
1971-74	25.9	24.1	-	-	-	-	-	23.9	-	-	-	-	25.9	23.8	-
1975-79	25.2	24.1	24.8	-	-	23.9	24.8	-	-	-	-	24.3	23.8	-	-
1979-82	26.8	26.5	-	25.2	23.8	-	26.1	24.5	25.6	-	24.1	-	24.2	26.1	25.3

Totals. Row (1): moments per subregion in units of 10^{28} dyne-cm;
Row (2): moment magnitudes of single event representing total moment and total moments per subregion$^{(h)}$:

		E	SH	SF	NS	TC	CN	C	SS	OV	BF	MJ	LA	IP	BJ
(1) M_o:	7.3	.34	.002	2.6	.003	.32	.29	1.5	.005	1.3	.50	.012	.10	.063	.29
(2) M:	8.5	7.6	6.1	8.2	6.3	7.6	7.6	8.1	6.4	8.0	7.7	6.7	7.3	7.1	7.6

TABLE 2 (footnotes) cont.

(a) Occurrences of large events (defined as range M = 7.0 $\pm$ 0.3) in a category is shown by an asterisk (one per event), and the occurrence of a major event (defined as exceeding M = 7.3) is shown by a large solid dot; no "large" events occur within the same space-time windows as the "major" events. There are six events designated as major, in chronological order: (1) central San Andreas, 1857, 8.0, (2) Owens Valley, 1872, 8.0, (3) northern San Andreas, 1906, 8.2, (4) N.W. Nevada, 1915, 7.6, (5) Lompoc vicinity, 1927, 7.5, (6) Kern County, 1952, 7.7. Restriction of the pre-1850 data to SF and LA reflects the fact that San Francisco and Los Angeles were the main sources of historical documentation at that time (see documentation and interpretation of historic reports by Toppozada and others, 1979a).

(b) Data sources are from Figure 5. The last three rows span about four years rather than five because of the chronological limits of the data sets.

(c) Counts may not agree exactly with frequency-magnitude plots in Figures 16 and 17 because of ambiguities where events occur on boundaries between subregions. Counts for the pre-1900 intervals are based on magnitudes interpreted from intensity data by Toppozada and others (1979a, last column in their table 2).

(d) Most of the events in CN for this period occurred during 1954, including two earthquakes (7.1 and 7.2) east of Fallon, Nevada; other 1954 events, however, occurred from the southernmost to northernmost subregions. By contrast, most of the events in BF were foreshock-aftershock events of the 1952 Kern County event (7.7); relatively few 1952 events occurred elsewhere.

(e) Several of the events in LA were associated with the 1971 San Fernando earthquake (6.4).

(f) Events occurring from 1 January, 1975 through 31 March, 1979 listed by Toppozada and others (1979b).

(g) Events occurring from 1 April, 1979 through 31 October, 1982 listed by Sherburne and others (1985). Many of the events in CN were in the southwest corner near the CN-SS boundary and represent the flurry of activity beginning in 1980 in the vicinity of Mammoth Lakes, California (there were earthquakes of magnitudes 6.4, 6.5, and 6.3, and many smaller shocks, within a period of a few days in May, 1980); these are thought to have been associated in part with magmatic transport in the Long Valley volcanic system. Events have occurred previously in that vicinity (see maps of Real and others, 1978; Toppozada and others, 1979a) though not as a conspicuous flurry within the Long Valley Caldera.

(h) The moment magnitude of an earthquake equivalent to the total historic moment in California and environs (8.5 in totals at end of table B.) is approximately equaled by any one of many post-1900 events in Circumpacific regions of subduction zone tectonics. Events of about this size occurred in Chile in 1906 and 1922; the 1960 event in Chile was an order of magnitude larger (see Kanamori, 1978).

(i) The column of moments immediately following the time interval gives the sums per time interval for all subregions.

TABLE 3. Redundancies for seismic event frequencies in California calculated from Eqs. (1) and (2).

Region and Time Interval[a]		Redundancies (percent)				
		R_1[b]	R_2[c]	R_3[d]	(R_1-R_2)[e]	(R_1-R_3)[f]
		(1)	(2)	(3)	(4)	(5)
E+SH:	1800-1900	59.3	20.6	48.7	38.7	10.6
	1900-1982	26.5	9.3	19.0	17.2	7.5
CN+TC+NS:	1800-1900	43.6	10.3	37.1	33.3	6.5
	1900-1982	33.1	17.4	19.0	15.7	14.1
BJ+IP:	1800-1900	71.9	8.5	69.3	63.4	2.6
	1900-1982	29.7	16.5	15.8	13.2	13.9
SF:	1800-1900	33.7	16.4	20.7	17.3	13.0
	1900-1982	37.7	9.1	31.5	28.6	6.2
C:	1800-1900	47.2	5.1	44.4	42.1	2.8
	1900-1982	43.2	17.2	31.4	26.0	11.8
LA:	1800-1900	45.6	17.3	34.2	28.3	11.4
	1900-1982	36.2	13.1	26.6	23.1	9.6
BF+MJ:	1800-1900	100.0	--.[g]	100.0	--.[g]	0.0
	1900-1982	35.3	18.3	20.8	17.0	14.5
OV+SS:	1800-1900	55.4	3.1	54.0	52.3	1.4
	1900-1982	46.4	9.9	40.5	36.5	5.9
ALL[h]:	1800-1900	28.9	17.0	14.3	11.9	14.6
	1900-1982	27.8	24.3	4.7	3.6	23.1
	1800-1982	26.0	23.1	3.7	2.9	22.3
	1975-1979	34.0	16.8	20.7	17.3	13.3
	1979-1982	24.1	26.1	2.6	2.0	21.5

(a) Regions identified in Figure 14.

(b) Calculated from Eqs. (1) and (2) for n = 33.

(c) Calculated from Eqs. (1) and (2) for n corresponding to the number of magnitude values for which there are actual events (magnitudes are read at intervals of 0.1 beginning from M = 5).

(d) Calculated from Eq. (2) for the observed n-values relative to the assumed limit n = 33 (i.e., relative to all possible M values between 5 and 8.2); the redundancy is zero for n_{obs} = 33 because this gives the maximum relative uncertainty, F = 1, for this assumption.

(e) The difference R_1-R_2 is zero if events occur at each of the 33 magnitude values. Dividing by $(1-R_2)$ gives R_3.

(f) The difference R_1-R_3 becomes equal to R_1 if events occur at all magnitude values, because when n_{obs} = 33, F = 1 and R_3 = 0

(g) This is the case n_{obs} = 1, U_{obs} = 0, F_1 = 0, R_1 = 100%, (R_1-R_3) = 0; but F_2 = 0/0, so R_2 and (R_1-R_2) are indeterminate.

(h) Data sources are the same as in Figure 5.

TABLE 4. Redundancies based on an idealized bond-count model*.

Region	Time Interval	R_B(percent)	Region	Time Interval	R_B(percent)
E+SH	1800-1900	66.6	LA	1800-1900	51.6
	1800-1982	34.1		1900-1982	32.8
CN+TC+NS	1800-1900	57.2	BF+MJ	1800-1900	100.0
	1900-1982	38.4		1900-1982	58.4
BJ+IP	1800-1900	80.4	OV+SS	1800-1900	92.5
	1900-1982	27.3		1900-1982	54.4
SF	1800-1900	35.0	ALL	1800-1900	50.6
	1900-1982	87.9		1900-1982	18.0
				1800-1982	17.8
C	1800-1900	89.7		1975-1979	23.2
	1900-1982	63.1		1979-1982	16.0

* The data sources and equations are the same as were used in Table 2 but are adjusted so that there is unit bond count at M = 5 increasing in direct exponential proportion to 1585 counts at M = 8.2. Probabilities, uncertainties, and redundancies in Eqs. (1) and (2) are calculated using the normalized bond-count frequencies.