Tectonophysics, 211 (1992) 61–83
Elsevier Science Publishers B.V., Amsterdam

Asperity distributions and large earthquake occurrence in subduction zones

Larry J. Ruff

Department of Geological Sciences, University of Michigan, Ann Arbor, MI 48109, USA

(Received June 14, 1991; revised version accepted March 2, 1992)

ABSTRACT

Ruff, L.J., 1992. Asperity distributions and large earthquake occurrence in subduction zones. In: T. Mikumo, K. Aki, M. Ohnaka, L.J. Ruff and P.K.P. Spudich (Editors), Earthquake Source Physics and Earthquake Precursors. *Tectonophysics*, 211: 61–83.

Plate tectonics and the seismic gap hypothesis provide the framework for long-term earthquake forecasting of plate boundary earthquakes. Unfortunately, detailed examination reveals that earthquake recurrence times and rupture length vary between successive earthquake cycles in the same subduction zone. Furthermore, larger coseismic slip is commonly associated with larger rupture length. Hence, large earthquake occurrence in subduction zones is characterized by variability in: (1) recurrence times, (2) rupture length, and (3) coseismic slip. These facts, plus many other observations, indicate that there are significant spatial variations in the "strength" of the plate interface. One simple description of these variations and their role in the earthquake cycle is the asperity model, where the large strong regions of the plate interface are called asperities, and the large earthquakes occur when the large asperities break. The asperity model of earthquake occurrence is able to qualitatively explain several features of large plate boundary earthquakes. To go beyond general qualitative notions, I pose the following scientific test: are the observed asperity distributions and a simple model of their interaction self-consistent with the above three observed features of large earthquake occurrence?

The distribution of the major asperities along plate boundary segments has now been determined for several subduction zones. Rupture process studies of adjacent large and great earthquakes have provided reliable estimates of the along-strike asperity lengths and separations for several adjacent asperities in the Kurile Islands, Colombia, and Peru subduction zones. The simplest mechanical model for asperity interaction is to idealize two adjacent asperities as frictional sliders that are connected by main springs to the upper plate, by a coupling spring to each other, and maintain frictional contact with a conveyer belt (the lower plate) that moves with a constant velocity. An "earthquake" occurs when the net force on the asperity frictional slider reaches some specified level. The failure force and spring constants are determined by the observed asperity distribution and simple models of elastic interaction. Two different macroscopic failure criteria are used. This simple mechanical model displays a remarkable range of behavior from simple to complex. When the two asperities are identical in all their properties, sequences of identical "earthquakes" are produced. For the more realistic case of non-identical asperities, "earthquake" sequences show great variety. Using system variables from the observed asperity distributions, the "earthquake" sequences typically display: (1) variable recurrence times, (2) variable rupture length, i.e. a combination of single-asperity and double-asperity failures, and for one of the failure criteria (3) larger coseismic slip for double-asperity failures. Statistical summaries of thousands of simulated "earthquake" sequences for asperity pairs in the Kuriles, Colombia, and Peru subduction zones are broadly consistent with the observed features of large earthquake occurrence in these subduction zones.

The main conclusion is that the asperity model provides a self-consistent explanation for: fault zone heterogeneity, the rupture process, and recurrence times and rupture mode of large earthquake sequences *via* a simple model for adjacent asperity interaction. In addition, a conclusion independent of any particular model for fault zone heterogeneity is that simple deterministic models of fault zone interaction can explain complex patterns of large earthquake occurrence in subduction zones.

1. Introduction

Plate tectonics provides the kinematic framework to explain why large earthquakes occur re-

Correspondence to: L.J. Ruff, Department of Geological Sciences, University of Michigan, Ann Arbor, MI 48109-1063, USA.

peatedly along plate boundaries. Soon after the advent of plate tectonics, several workers developed the seismic gap hypothesis (see, e.g., Fedotov, 1965; Mogi, 1968; Kelleher et al., 1973; McCann et al., 1979). The seismic gap hypothesis uses the previous history of earthquake occurrence along a plate boundary to make long-term forecasts of future earthquake occurrence. Unfortunately, detailed studies reveal that recurrence intervals between successive large earthquakes in the same subduction zone segment can vary by decades (see Rikitake, 1976; Sykes and Quittmeyer, 1981; Nishenko and Buland, 1987). In addition, the rupture length of successive large earthquakes in a subduction zone can vary by a hundred kilometers or more (Rikitake, 1976; Kanamori and McNally, 1982). While the seismic gap hypothesis has provided several successful long-term forecasts of earthquakes, we are still searching for a complete physical description of earthquake occurrence. Many investigators are using a variety of approaches to further understand the where, why, and when of large plate boundary earthquake occurrence. The approach presented in this paper can be viewed as an observational approach to directly determine those features of the earthquake rupture process that can be resolved from analysis of the seismic waves. Thus, the emphasis is to study large earthquakes and some of their characteristics, and then to use the simplest possible model to explain these observed characteristics.

The complexity of earthquake occurrence implies that the seismogenic plate interface must have spatial heterogeneities. The simplest mechanical property of the plate interface is the failure strength, and the simplest model for its variation is a binary classification of the seismogenic interface: strong and weak. The strong sub-regions can be called either barriers or asperities; in this paper I shall call them asperities. The asperity model then states that the largest earthquakes along the plate interface occur when the largest asperities fail. This asperity model is a very simple idea that has enough flexibility to explain some qualitative aspects of large earthquake occurrence. For example, the variable rupture length of large earthquakes is qualitatively explained as a consequence of asperity interaction. Adjacent asperities can interact strongly or weakly to produce either complicated sequences of earthquakes or uniform isolated earthquakes, respectively. It is now time to test some of these notions in a more quantitative fashion. Advances in waveform studies of large earthquakes now make it possible to identify subregions of higher slip within the fault area—these subregions are asperities. Furthermore, studies of adjacent great earthquakes along several subduction zones allow us to make maps of the asperity distribution. In this paper, I will use the simplest possible model of asperity interaction to test whether observed asperity distributions are consistent with several observed features of large earthquake occurrence.

2. Some observed features of large earthquake occurrence

Although we still do not understand the underlying physical mechanisms that determine the seismogenic potential of the plate interface, we observe that many subduction zones are characterized by the occurrence of great underthrusting earthquakes. Indeed, most of the largest earthquakes of the twentieth century are underthrusting events in subduction zones (see review by Kanamori, 1986). Large earthquakes do not occur uniformly throughout the world's subduction zones. Most of the large earthquakes (i.e. magnitude greater than 8) occur in the subduction zones in South America, Mexico, and northwestern Pacific Ocean (Japan to Alaska). These subduction zones are referred to as "strongly coupled" (Uyeda and Kanamori, 1979; Ruff and Kanamori, 1983). Within these subduction zones, the size of the characteristic largest earthquakes range in moment magnitude (M_W, Kanamori, 1978) from 8.1 (Mexico) to 9.5 (Southern Chile). In the case of great subduction zone earthquakes, larger earthquakes are typically associated with a larger fault length parallel to fault strike (see summary in Ruff, 1989). In addition, the depth extent of the seismogenic zone is nearly constant for most of the strongly coupled subduction zones (e.g. Tichelaar and Ruff, 1991). These observa-

tions imply that great earthquakes rupture the full width (depth range) of the seismogenic zone; the along-strike rupture length and average slip are the key determinants of overall seismic moment. In short, greater earthquakes result from greater along-strike rupture length. Therefore, much of the effort in characterizing great earthquake occurrence has been focussed on the two parameters: (i) along-strike rupture length, and (ii) recurrence time between great earthquakes. Aftershock areas can be used to estimate rupture length for most twentieth century great earthquakes. The along-strike rupture length can be estimated for older earthquakes by examination of intensity maps. The recurrence of great earthquakes over time spans of a few hundred years has been established in several subduction zones through the study of historical documents. Perhaps the best example of well-documented great earthquake occurrence over several hundred years is for the Nankai subduction zone along the coast of southern Honshu, Japan (e.g. Utsu, 1974; Ando, 1975; Ishibashi, 1981); several investigators found that the Nankai subduction zone could be divided into four or five seismic segments. Two important facts emerge from studies of earthquake occurrence along the Nankai and other subduction zones: (1) the recurrence time for any one seismic segment is variable, and (2) the rupture lengths of large earthquakes is variable, i.e. adjacent seismic segments will sometimes rupture as individual earthquakes, other times a larger "multiple-event" earthquake will rupture two or more adjacent segments. Much research effort is spent trying to understand the cause of observations (1) and (2).

Another observation about large earthquake occurrence is based on the seismic moments of earthquakes. Kanamori and McNally (1982) studied the two earthquakes sequences that ruptured the Colombia–Ecuador subduction zone. The entire 500 km subduction zone was ruptured by a great earthquake in 1906 (Kelleher, 1972). This subduction zone then reruptured in a sequence of three earthquakes that occurred in 1942, 1958, and 1979. Kanamori and McNally (1982) estimated the seismic moments of all earthquakes and found that the sum of seismic moments of the 1942, 1958, and 1979 events is less than about 1/2 the seismic moment of the 1906 earthquake. Since the total rupture length is the same for both sequences, then the seismic slip must vary in successive earthquake cycles if we assume that rupture width is constant. One implication of variable slip is that stress-drop may vary from one earthquake cycle to the next. Alternatively, one can allow both seismic slip and fault width to vary in such a fashion that stress-drop remains constant, as suggested by Kanamori and McNally (1982). While one empirical fact about plate boundary earthquakes is that stress-drop is approximately "constant" (Kanamori and Anderson, 1975), there is appreciable scatter in this "constant". The stress-drop of large subduction earthquakes can easily vary by a factor of two, possibly as much as a factor of ten (see, e.g., Kanamori and Anderson, 1975; Sykes and Quittmeyer, 1981). To summarize the above paragraph, a third observed feature of large earthquake occurrence is that seismic slip can also vary between successive earthquakes in a subduction zone segment—stress-drop may or may not vary in accordance with slip variations.

Figure 1 depicts the three basic observed features of large earthquake occurrence. The question that we will pursue is: can observed asperity distributions explain these three basic features of earthquake occurrence? But first, I shall briefly discuss a different approach to characterize complications in earthquake occurrence.

2.1. Time / slip predictable model

The time-predictable and slip-predictable models (Shimazaki and Nakata, 1980) were devised to provide a deterministic explanation for observation #1 of Figure 1. These two models are the end-member cases of a scheme to relate variations in recurrence times to variations in seismic slip (i.e., observation #3 of Fig. 1). The model makes no provision for observation #2 since each subduction zone is considered in isolation from its neighbors. Figure 2a shows how a hypothetical "uniform" seismic segment might rupture through time: the earthquake occurs when a constant failure stress is achieved, and the

stress-drop is constant so that the final stress is uniform. With a constant stress accumulation rate, the "ideal" earthquake behavior would be a sequence of identical earthquakes with a constant recurrence time. However, earthquakes do not behave this way. Figure 2 (b, c) shows a hypothetical sequence of three earthquakes; the two recurrence times are different. The time/slip-predictable model offers two end-member explanations for variable recurrence time: (i) the failure stress is always the same, but the final stress can be different (due to unknown factors), (ii) the failure stress can vary, but the final stress is always the same. While either end-member model can explain any sequence of recurrence times, it is possible to test which end-member model is preferred if we also know the stress-drop associated with each earthquake. Due to several technical problems in obtaining the stress-drop of earthquakes, it is probably more reliable to measure seismic slip, with the assumption that the seismic slip in any one segment is proportional to the stress-drop of that segment. Unfortunately, it is difficult to obtain slip estimates for older earthquakes. Several workers have tried to analyze sequences of earthquakes and slip estimates to decide whether the slip-predictable or time-predictable cases are preferred – but any study is hindered by the fact that we lack examples of sequences of at least three large earthquakes with known seismic slip. Perhaps the best case is the sequence of three earthquakes that occurred in a segment of the Nankai trough region. Shimazaki and Nakata (1980) used measurements of coastline uplift as a proxy for the displacements on the fault plane and found a slight preference for the time-predictable model. Overall, there is not enough statistical evidence to globally prefer either end-member model; earthquake sequences are too irregular (see Sykes and Quittmeyer, 1981; Nishenko, 1985). In a later section, we will see that a simple model of asperity interaction can explain irregular earthquake sequences.

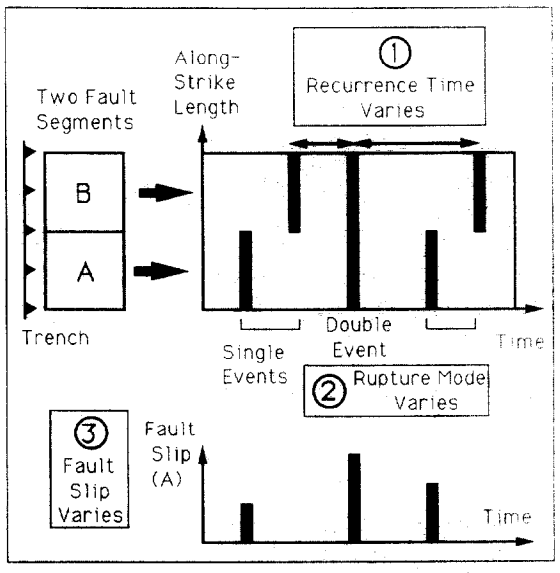


Fig. 1. Schematic illustration of three observed features of large earthquake occurrence in subduction zones. Two adjacent fault segments are considered. A hypothetical sequence of three earthquake cycles is shown; the bold lines indicate the time of occurrence and which fault segment ruptured. A double event ruptures both fault segments. As depicted, feature #1 is the fact that recurrence time varies between earthquakes in the same fault segment. The fact that rupture mode can change from single events to a double event is feature #2. The lower diagram shows coseismic slip of the three events for fault segment A. Feature #3 is the observation that coseismic slip can vary for double event and single event cycles.

At this point, one can either follow a statistical approach to the scatter in recurrence times, or probe the plate interface properties to follow a deterministic approach. The asperity model offers a physical basis for a deterministic explanation of the scatter in earthquake recurrence times.

3. Asperity model, rupture process, and asperity distributions

Mechanical properties within the seismogenic portion of plate interfaces are spatially heterogeneous. While many theoretical models of fault zone behavior assume that the mechanical properties and behavior of the fault zone are uniform, there is ample evidence that fault zones are highly non-uniform. Observations that can be quoted as evidence for spatial heterogeneity include: the occurrence of different size earthquakes within the same fault segment; foreshocks and aftershocks; earthquake clusters within a fault segment; changes in fault geometry during rupture; spatially varying coseismic displacements as observed by fault offsets or geodetic measurements;

and irregular rupture processes as observed by analysis of seismic waves. To understand the broad scope of earthquake occurrence, we must allow spatial heterogeneity in fault zone properties.

What is the best characterization of spatial heterogeneity? It is useful to first distinguish between "geometric heterogeneity" and heterogeneity within a continuous planar fault segment. Major changes in the along-strike plate boundary geometry are "geometric heterogeneities". For example, the cusp in the Japan trench between Honshu and Hokkaido is a major geometric heterogeneity that serves to segment the major earthquake rupture zones. Another example would be the northeastern truncation of the 1964 Great Alaskan earthquake rupture zone where the North America–Pacific plate boundary is changing from underthrusting to strike-slip motion. These major geometric boundaries are easily recognizable and provide the basic global segmentation of plate boundaries. At a finer scale, geometric offsets correlated with features in either the subducting or overriding plates may divide the plate interface into segments of a hundred kilometers or more in length.

Seismic segments of the plate interface are defined by the rupture zones of great earthquakes. One famous example of dividing a subduction zone into seismic segments is the Nankai subduction zone off southern Honshu (Utsu, 1974; Ando, 1975; Ishibashi, 1981). Global surveys have defined the major seismic segmentation of the world's coupled subduction zones (Rikitake, 1976; McCann et al., 1979). Given the lack of evidence to prove otherwise, it is commonly assumed that the plate interface is approximately planar within the seismic segments. Even so, there must be spatial heterogeneity within these seismic segments as evidenced by the seismicity and irregular rupture processes. Thus, we must introduce spatial heterogeneity into the mechanical behavior of a planar plate interface. Since we do not yet know the detailed physical laws that govern fault slip (see many other papers in this special volume that treat this topic), it is best to use the

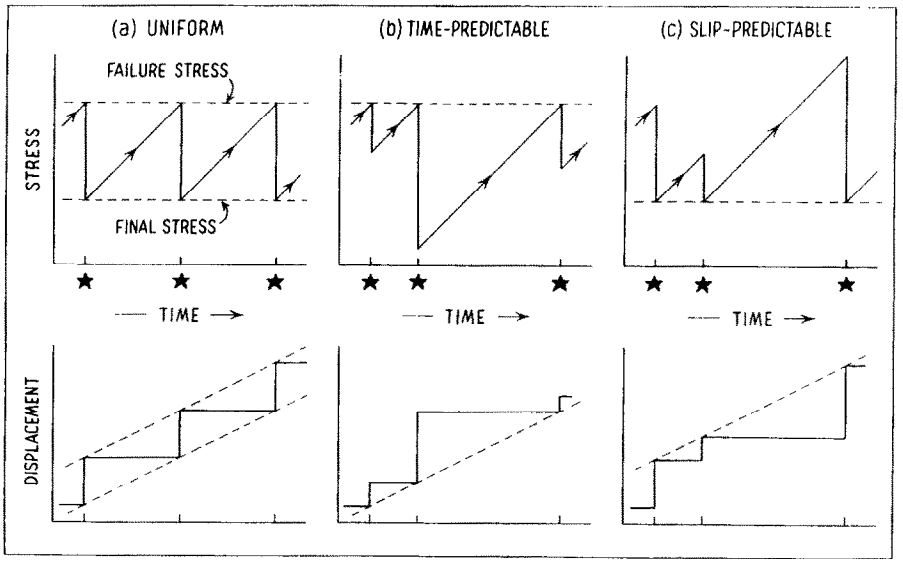


Fig. 2. Idealized earthquake sequences for a fault segment. Upper diagrams show stress as a function of time, occurrence of earthquakes is denoted by stars. The lower diagrams show the corresponding cumulative coseismic displacement. It is assumed that stress increases at a steady rate for all three models. For the uniform model (a), both the failure and final stress levels are constant (dashed lines). This model produces a sequence of identical earthquakes with constant recurrence times. The time-predictable (b) and slip-predictable (c) models both produce the same earthquake sequences in terms of recurrence times. The time-predictable model states that the failure stress is constant, but the final stress varies. The slip-predictable model states that the final stress is constant, but the failure stress varies.

simplest possible description of fault zone property. Perhaps the simplest characterization is the failure stress of the fault zone. The shear stress across any portion of the fault zone must increase to the level of the failure stress before slip can occur. The level of local shear stress then decreases due to slip. While it is likely that the failure stress is distributed over a range of values, the simplest characterization is to use a binary classification of failure stress: strong and weak. In this fashion, the fault zone segment can be divided into strong regions and weak regions. The strong regions can be referred to as either barriers (Das and Aki, 1977; Aki, 1979) or asperities (Kanamori, 1981). The distinction between the barrier or asperity model comes from the role that the strong regions play through the earthquake cycle.

It is important to state that failure stress is a macroscopic property that may represent several different disparate phenomena. While failure stress is idealized to be a frictional property of a planar fault interface, this property might be largely determined by small-scale geometric irregularities. Since we can only identify "strong" regions on the basis of seismic slip averaged over ten kilometers or more, there could be several smaller-scale mechanisms contributing to this spatially integrated macroscopic property of "strong".

The simple fault zone characterization of strong or weak does not specify whether slip occurs as an earthquake or as creep, nor whether stress-drop is complete or partial. Furthermore, the complete space–time history of slip depends on the spatial variations in failure stress in addition to how the tectonic shear stress is applied. Thus, while many seismologists might agree to the division of the fault zone into strong and weak regions, there are still many possibilities for fault zone behavior. The asperity model and the role of strong regions in the earthquake cycle has been previously described (e.g. Kanamori, 1981; Lay et al., 1982; Ruff, 1983, 1989), and is again presented in Figure 3. The key idea is that the largest earthquakes along a plate boundary segment occur when the largest strong regions (asperities) fail. Within the context of this model, the asperities can be found if we are able to determine the along-strike variation in seismic slip of the large earthquakes.

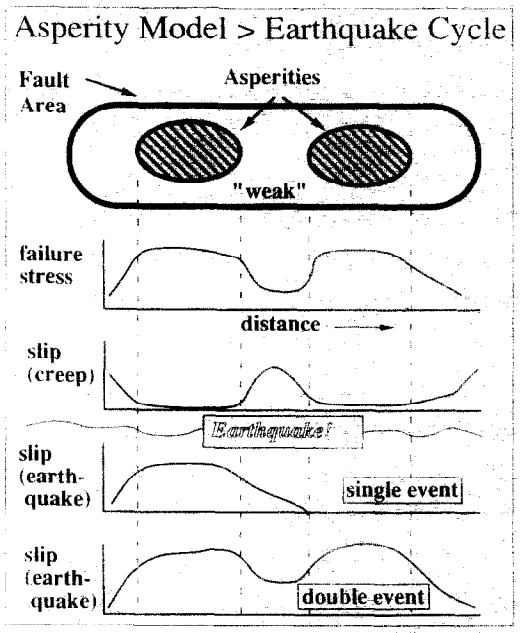


Fig. 3. Role of asperities in the earthquake cycle. Top diagram shows two adjacent large asperities along a portion of a plate interface (fault area). Diagram just below fault area depicts the interface strength as measured by the "failure stress"; the asperities are stronger than the surrounding weak regions. Lower diagrams show the cumulative slip (creep) just before a large earthquake (earthquake occurrence depicted by wavy line), and coseismic slip for two possible earthquakes, a single event in which one asperity ruptures, and a double event in which both asperities break. Coseismic slip is larger for the asperities than for the weak regions.

3.1. Asperity distributions in the Kurile Islands and South America subduction zones

Asperities are regions of larger seismic slip during great earthquakes. While it is likely that asperities come in a wide variety of sizes, it is the largest asperities that mostly influence great earthquake occurrence. The best way to locate asperities is to determine the variations in coseismic slip from a great earthquake that ruptures the plate boundary. There are two methods to determine coseismic slip within a rupture zone: (1) geodetic studies, and (2) analysis of the waves radiated by the earthquake. Both methods are

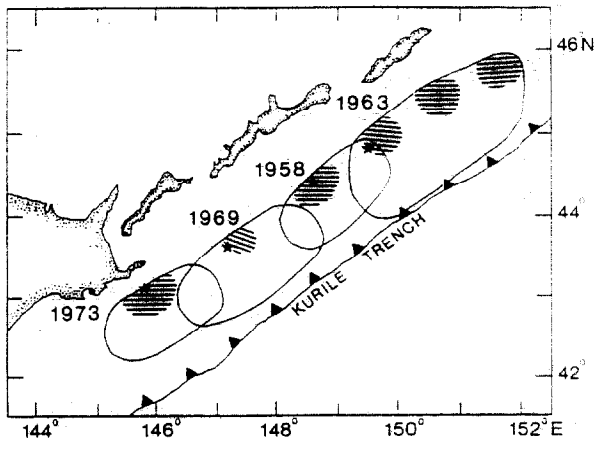
important, but unfortunately it is quite rare to have sufficient geodetic information to resolve any along-strike variations in slip for a subduction zone earthquake. Therefore, a systematic global study of subduction zone earthquakes must rely upon wave analysis. These studies are commonly referred to as "rupture process studies".

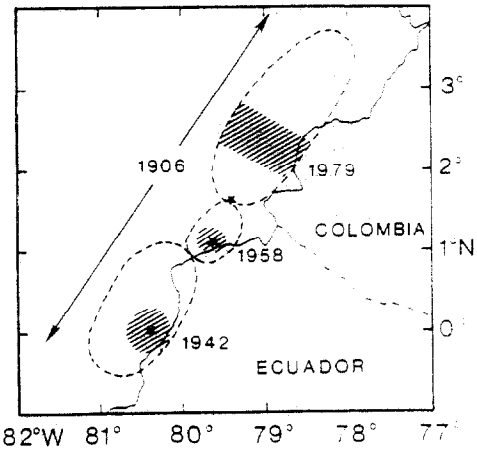
Rupture process studies follow from the advent of seismic source theory, rupture models, and the computational capability to construct synthetic seismograms and solve the inverse problem (see development in, e.g., Aki and Richards, 1980). While there were several pioneering studies in the 1970's, the systematic inversion of seismic waves for the rupture process of large subduction earthquakes occurred through the 1980's. Two independent research groups have studied the spatial–temporal distribution of moment release of great subduction zone earthquakes that occurred from 1963 to present. Summaries of these studies can be found in Kikuchi and Fukao (1987), and Beck and Ruff (1989). These rupture studies exploit the subtle time shifts in the moment release time function as seen by P waves at various azimuths to locate regions of higher moment release within the fault area (see above references for details). For the purpose of comprehensive global-scale rupture studies, the most reliable and highest resolution rupture studies are primarily based on long-period P waves. In special circumstances, other wave types offer some advantages. Particularly important for future studies of subduction earthquakes of all ages is the systematic use of tsunami waves to locate regions of higher moment release (Satake and Kanamori, 1991). Large earthquakes that have occurred since 1985 have received several independent rupture process studies using a variety of techniques including near-field waves; a review of all this work is beyond the scope of this paper. Since our focus here is on asperity interaction, we need to know the size and separation of adjacent asperities along subduction zones. This information is summarized in Beck and Ruff (1989) for the Kuriles, Colombia, and Peru subduction zones (Fig. 4). These asperity distributions can be related to several aspects of the large earthquakes and associated seismicity (see discussions in Schwartz and Ruff, 1987; Thatcher, 1990), but our main interest here is on sequences of large earthquakes in these subduction zones. Both the Peru and Colombia–Ecuador subduction zones have experienced a mixture of rupture modes, i.e. multiple events followed by a sequence of single events. Also, the Kurile Islands subduction zone displays a combination of single and multiple event ruptures in the most recent earthquake cycle. Of course, recurrence times vary for all of these subduction zone segments. Therefore, the three subduction zones for which we have asperity distributions display variability in (1) recurrence times, (2) rupture mode, and (3) seismic slip, between successive earthquake cycles.

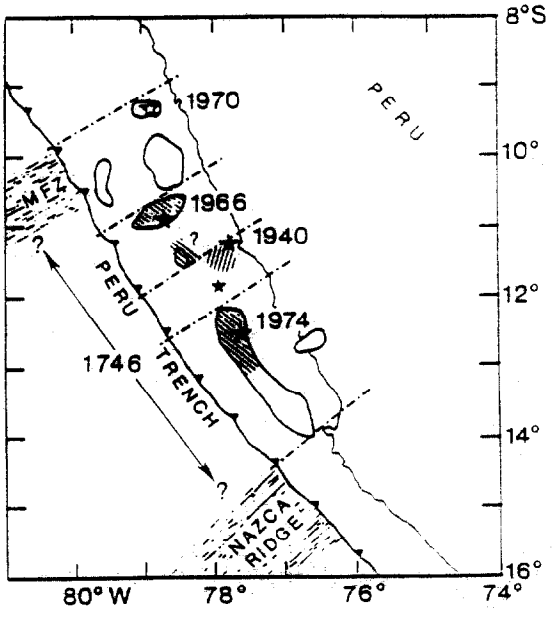
The rupture zones of large subduction earthquakes are elongated along the strike of the zone. The rupture length of great earthquakes seems to be more variable than the rupture width (see Sykes and Quittmeyer, 1981; Ruff, 1989). For this reason, plus the fact that our techniques for rupture process studies can better resolve asperity length than width, we will characterize asperity distributions by the along-strike lengths and separations of adjacent asperities. Asperity lengths and separations for the Kuriles, Peru, and Colombia–Ecuador subduction zones are summarized in Table 1. The interaction of any two adjacent asperities will depend on the ratio of average asperity length to separation ($\bar{a}/L$ in Table 1). As noted by Beck and Ruff (1987), the three asperities in the rupture zone of the great 1963 Kurile Islands earthquake (presumably a triggered multiple event earthquake) are closer together than the asperities in the southern Kuriles and Hokkaido regions, where isolated single-asperity ruptures occurred. This observation suggests that some features of large earthquake occurrence might be controlled by the basic geometric aspect of asperity distributions: ratio of asperity length to asperity separation. The next section will test this hypothesis.

4. Asperity interaction modeled by coupled frictional sliders

The slip and stress history of a fault segment will affect the stress state of neighboring seg-







ments. Interaction between fault segments is conveyed by the elastic behavior of the surrounding material. In detail, the geometry of the applied stresses, creep properties of the deeper portion of the plate interface, and asthenosphere viscoelastic behavior all affect the stress state of the fault segments (see e.g., Savage, 1983; Thatcher and Rundle, 1984; Matsu'ura and Sato, 1989). To gain insight as to the role of mechanical strength variations along the plate interface in affecting earthquake occurrence, we must simply characterize the elastic interactions and idealize all other effects. One explicit characterization of the elastic interactions between fault segments is the work of Rundle and Kanamori (1987), but their formulation is quite complicated. Perhaps the simplest characterization is to replace the three-dimensional continuum problem with a discrete element system that consists of frictional sliders connected by springs (Fig. 5). The frictional sliders are connected by main springs to the upper plate, and rest on the subducting lower plate, which acts as a frictional conveyer belt. The sliders are connected to each other by a spring that represents the along-strike elastic coupling between asperities in adjacent subduction zone segments. Elasticity of the lower plate is ignored. To simulate an entire subduction zone, we should connect several frictional sliders. However, at this point we want to first understand the behavior of the simplest interaction: two adjacent asperities.

The interface mechanical properties are idealized to be weak and strong to the extreme that all coupling between the upper and lower plates takes place through the asperities. Thus the frictional sliders have the area of the asperities and

Fig. 4. Great earthquake rupture zones and asperities (hachured regions) for three subduction zones. Kurile Islands subduction zone is shown at top; this zone was ruptured by great earthquakes that occurred from 1958 to 1973 (see Beck and Ruff, 1987; Schwartz and Ruff, 1987). Middle plot shows the Colombia–Ecuador subduction zone and asperities that were found for the most recent sequence of three large earthquakes from 1942 to 1979. Peru subduction zone (lower plot) was ruptured by three large earthquakes in 1940, 1966, and 1974 (see Beck and Ruff, 1989, for summary and discussion).

the influence of the weak portion of the interface is ignored. The dynamic aspects of the loading are ignored, we instead prescribe a steady rate of plate motion as the main springs are compressed by the frictional sliders stuck to the lower plate. Properties of this discrete element system that must be specified are the spring constants of the two main springs and the coupling spring, and the failure criteria for the two frictional sliders.

While "slip event" or "asperity failure" might be the best terms to describe earthquakes in the simple mechanical system, I shall frequently use "earthquake". Recall that the only real earthquakes in this paper are those shown in Figure 4.

4.1. Variables of the discrete element system

The three-dimensional elastic interactions are replaced with three springs acting on two frictional sliders. Two main springs represent the elastic compression of the upper plate as plate motions proceed, while the force due to the coupling spring depends on the offset in asperity positions. The net spring force acting on a frictional slider is the sum of the forces from the main and coupling springs. Let us keep track of asperity positions with X_1 and X_2 for asperities #1 and #2, respectively (Fig. 5). Define $X_1 = 0$ when the main spring #1 exerts no force, and similarly for X_2. The net force acting on asperity #1 is: $f_1 = -K_1X_1 - k(X_1 - X_2)$, where K_1 and k are the main and coupling spring constants, and negative force points opposite to positive motions. For asperity #2, the net force is: $f_2 = -K_2X_2 + k(X_1 - X_2)$. Note that for a given asperity offset, the coupling spring will act to increase the force on one asperity, and decrease the force on the other. Both X_1 and X_2 increase

TABLE 1

Observed asperity distributions and discrete element system parameters for the gaps between adjacent asperities, calculated for the rigid cube and rigid block models

Asperity I.D.	a_i (km)	L (km)	$\bar{a}/L$	Rigid cube			Modified coupling k	Rigid block		
				F_2/F_1	K_2/K_1	κ		F_2/F_1	K_2/K_1	κ
K63 c	60	30	1.83	1.44	1.20	0.40	1.21	1.20	1	0.33
K63 b	50	55	1.0	0.69	0.83	0.18	0.54	0.83	1	0.18
K63 a	60	60	1.08	0.73	0.86	0.20	0.60	0.86	1	0.17
K58	70	75	0.73	3.06	1.75	0.20	0.43	1.75	1	0.13
K69	40	95	0.58	0.33	0.57	0.09	0.23	0.57	1	0.11
K73	70									
C79	60	105	0.43	4.00	2.0	0.13	0.20	2.00	1	0.10
C58	30	105	0.33	0.56	0.75	0.06	0.09	0.75	1	0.10
C42	40									
P66	50	80	0.56	1.56	1.25	0.13	0.27	1.25	1	0.13
P40	40	30	1.50	0.64	0.80	0.27	0.88	0.80	1	0.33
P74 a	50	40	1.13	1.56	1.25	0.25	0.69	1.25	1	0.25
P74 b	40									

Asperity I.D.: K63 a, b, and c, three asperities of the 1963 Kurile Islands earthquake; K58, K69, and K73, asperities for the 1958, 1969, and 1973 Kurile Islands earthquakes (Schwartz and Ruff, 1987); C79, C58, and C42, asperities for the 1979, 1958, and 1942 Colombia–Ecuador earthquakes (Beck and Ruff, 1989); P66, P40, P74 a and b, primary asperities for the 1966 and 1940 Peru earthquakes, and the two asperities of the 1974 Peru earthquake (Beck and Ruff, 1989).

a = along-strike asperity length in kilometers.

L = along-strike asperity separation in kilometers.

$\bar{a}/L$ = ratio of average asperity length to separation.

System parameters (see text) are calculated for two models, rigid cube and rigid block with $W = 50$ km. Also, the modified coupling parameter is calculated for the rigid cube model.

with time due to plate velocity R. Define the frictional failure force for asperities #1 and #2 as F_1 and F_2. When f_1 obtains a value of $-F_1$, then the asperity "fails" as a slip event occurs and X_1 decreases. The details of failure criteria are discussed in a following section.

It will be convenient to measure asperity positions by non-dimensional values of X_1 and X_2. The maximum value of asperity position, X_{max}, is the greater of F_1/K_1 or F_2/K_2. Hence, the normalized asperity positions are $X_1 = X_1/X_{max}$ and $X_2 = X_2/X_{max}$, and they range between 0 and 1 (lower part of Fig. 5). The above equations for f_1 and f_2 show a linear dependence between X_1 and X_2 for a constant f_i ($i = 1$ or 2). Lines in the (X_1, X_2) plot of Figure 5 are for two important values of f_i: F_i and zero. Time can be measured in units of the maximum time to wait for a slip event: $T_{max} = X_{max}/R$. Thus all normalized recurrence times will range between 0 and 1. The physical parameters of the discrete element system can also be nondimensionalized. System behavior is described by the three ratios: F_2/F_1, K_2/K_1, and the coupling parameter $\kappa = k/K_1$. We now need to connect the observed asperity distributions to these ratios.

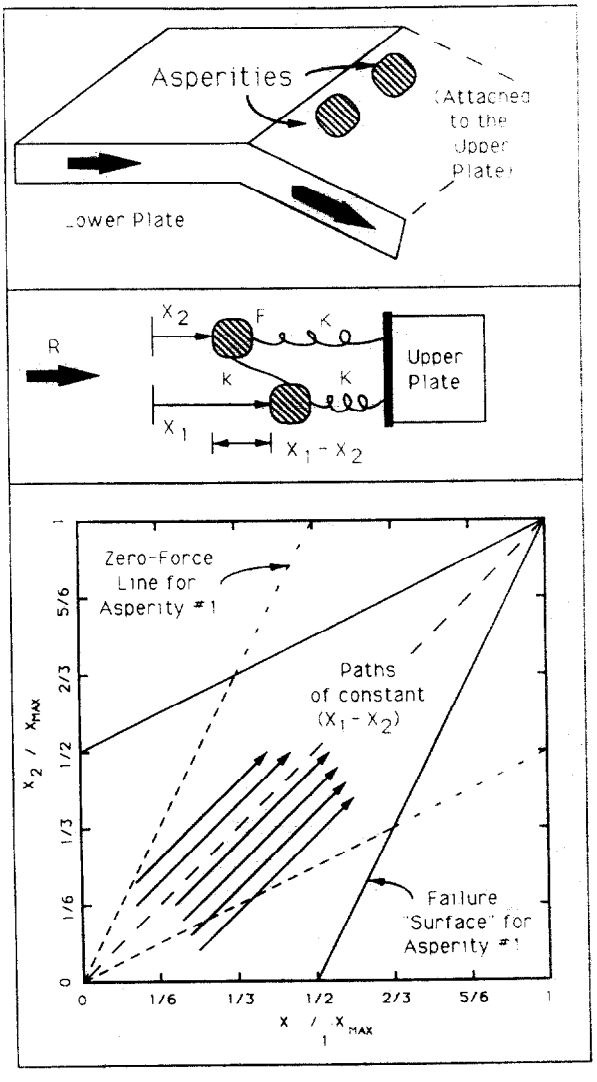


Fig. 5. Subduction zone asperities and coupled frictional sliders. Upper diagram depicts two adjacent asperities along the plate interface in a subduction zone. In the idealized model, all seismic coupling between the lower and upper plates takes place at the asperities. Interaction between asperities is modeled by the simple discrete element system (middle diagram) of two frictional sliders (the asperities, hachured) connected by springs (K) to the upper plate, connected to each other by a coupling spring (k), and in frictional contact with the lower plate. Failure force of the frictional contact is F, plate tectonic velocity is R. System status is given by asperity positions X_1 and X_2. Asperity offset $(X_1 - X_2)$ remains constant as plate motions proceed and the main springs are compressed. Eventually, the force on one of the asperities reaches F, and it slips. The other asperity may be triggered and also slip. The lower diagram plots X_2 versus X_1, both normalized to X_{max}. System status can be followed as a function of time in this diagram. For a given asperity offset, plate motions produce a path that parallels the main diagonal. Failure force F is achieved when a path reaches a failure surface (solid lines). Zero-force lines (dashed) are also shown. This diagram is constructed for identical asperities and $k = K$.

4.2. *Spring constants*

The spring constants depend on the basic equation of elasticity and the geometry of the asperity distribution, as shown in Figure 6. A simplified scalar form of elastic interactions is: $\sigma = C\epsilon$, where σ is the stress, ϵ is the strain, and C is an elastic constant. In the discrete element system, elastic interactions are modeled by Hooke's Law: $F = -KX$, where X is the displacement from equilibrium, K is the spring constant, and F is the force. Now, note that σ can be written as $\sigma = -F/A$, where F is the force and A is the cross-sectional area, and that ϵ can

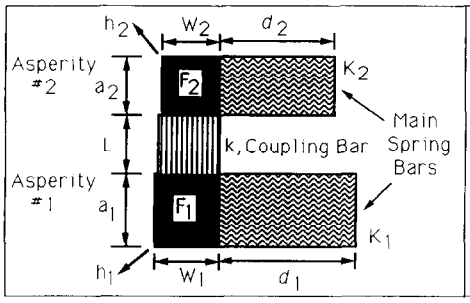
be written as $\epsilon = X/d$, where X is the displacement across the characteristic distance, d. With the above substitutions, the basic stress–strain relationship becomes:

$$F = -(CA/d)X = -KX$$

Thus the spring constant is: $K = CA/d$. There are still some choices that must be made to further specify K. Use of the simplified scalar elastic equation implies that we have idealized the elastic interactions as a homogeneous strain along an elastic bar of some dimension and cross-sectional area (Fig. 6). It is necessary to provide some geometric information in addition to asperity lengths and separations. The asperity is replaced by a rigid rectilinear block. Spring force is applied to two sides of the block: one side faces the adjacent asperity, while the other side faces the overlying plate. The sum of the forces on these two faces is balanced by the frictional force acting on the bottom side of the block. To fully describe each asperity block, we must specify the width (W) and height (h) of the block in addition to its along-strike length (a). For the moment, we let each block have different values for a, W, and h. We can now write down the dependence of the coupling and main spring constants on block dimensions.

Coupling spring constant. The basic stress–strain relationship for simple elastic shear is: $\sigma = \mu(X_1 - X_2)/2L$, where μ is the shear modulus, $(X_1 - X_2)$ is the asperity offset, and L is the asperity separation. To have a single value of k, we utilize an elastic coupling bar with a constant cross-sectional area that is the average of (W_1h_1) and (W_2h_2). Hence, the coupling spring constant is: $k = \mu(W_1h_1 + W_2h_2)/(4L)$.

Main spring constant. The basic stress–strain relationship for compression of an elastic bar is: $\sigma_i = EX_i/d_i = (5/2)\mu X_i/d_i$, where the subscript i is either 1 or 2 for the two asperities, d_i is elastic bar length, E is Young's modulus which equals $(5/2)\mu$ for a Poisson's ratio of 1/4. Since the area of the block face is a_ih_i, we can write the main spring constants as: $K_i = (5/2)\mu a_ih_i/d_i$. Unlike the situation for the coupling spring, it is not immediately obvious what is the best choice for the characteristic distance d_i. One might argue for several different dependencies on various combinations of system scale lengths, or possibly for a constant value based on some other dimension of subduction zones. Fortunately, it is possible to make a reasonable and simple choice for d_i guided by solutions to the classic elastic problem of half-space surface displacement due to a normal stress imposed over a rectangular area. This problem was solved by Love and is presented and discussed in Johnson (1985). For a pressure σ acting over a square region (i.e. $h_i = a_i$), the area-averaged displacement of the surface, X, can be related to σ as: $\sigma = cX/d$. Examination of the solution reveals that the best choice for d is a_i, and with a Poisson's ratio of 1/4, the numerical constant c is within 20% of $(5/2)\mu$. Therefore, one reasonable choice is to simply let $h_i = d_i = a_i$; and the spring constants would be: $K_i = (5/2)\mu a_i$.



	"Rigid Cube" Model	"Rigid Block" Model
F_2/F_1	$= (a_2/a_1)^2$	$= (a_2/a_1)$
K_2/K_1	$= (a_2/a_1)$	$= 1$
k/K_1	$= \frac{(1+a_2/a_1)^2}{20}\frac{a_1}{L}$	$= \frac{W}{5L}$

Fig. 6. Asperity geometry, spring constants, and system variables. Each frictional slider is a rigid block with length, width, and height of a, W, and h. Interaction of adjacent asperities is modeled by strains of elastic bars with various cross-sections and lengths. Formulas can relate elastic parameters and system geometry to spring constants (see text). System behavior depends on three non-dimensional ratios: F_2/F_1, K_2/K_1, and k/K_1. Specific formulas for the system ratios are given for two assumed geometries. See Table 1 for calculated values.

4.3. System parameters: F_2/F_1, K_2/K_1, and k

System behavior is governed not by individual values of the spring constants and failure forces, but by the ratios of F_2/F_1, K_2/K_1, and $k = k/K_1$. We shall use the above relationships and different block geometries to obtain values for these ratios for the asperity pairs in Table 1.

One could assume that failure stress varies between asperities, but we have no direct confirmation of this statement. Thus, we shall let the failure stress be the same for all asperities. With σ_f the failure stress, the failure force F_i is defined as: $\{F_i = \sigma_f(a_i W_i)$, and then $F_2/F_1 = (a_2 W_2)/(a_1 W_1)$. The most general form of the ratio of main spring constants is: $K_2/K_1 = a_2 h_2 d_1/(a_1 h_1 d_2)$. While the F_2/F_1 and K_2/K_1 ratios measure the asymmetry of the two asperities, the k/K_1 ratio measures the coupling between the asperities. For the most general form of k and K_1, this ratio is:

$$k/K_1 = \frac{(1/10)[d_1(W_1 h_1 + W_2 h_2)]}{(a_1 h_1 L)}$$

If we use the "square-face" specification of $d_i = h_i = a_i$, then the main spring ratio reduces to $K_2/K_1 = a_2/a_1$, and the coupling parameter becomes:

$$k/K_1 = \frac{(1/10)(W_1 a_1 + W_2 a_2)}{a_1 L}$$

At this point, it is now necessary to further specify the dimensions of the asperity blocks. We shall calculate the system ratios for two different models: the "rigid cube", and the "rigid block" with constant width and height.

Rigid cube. If we do not wish to introduce any length scales other than a_i and L, then we assume that the blocks are cubes, i.e. $W_i = h_i = a_i$. The formulas for the ratios of system variables are shown in Figure 6. Application of these formulas to the observed asperity distributions are listed in Table 1. The value of the coupling parameter, $k = k/K_1$, ranges from a low of 0.06 in the southern Colombia subduction zone to a high of 0.40 in the northern Kuriles subduction zone. Note that all asperity pairs show some asymmetry; that is, the K_2/K_1 and F_2/F_1 ratios differ from 1.

Rigid blocks with constant height and width. For this case, specify $W_1 = W_2 = W$ and $h_1 = h_2 = h$. The resultant formulas for the system ratios are shown in Figure 6. Note that h disappears from the ratios, but W plays a prominent role in the coupling parameter. The calculated values in Table 1 are for an asperity width of 50 km. While the width of the plate interface ruptured by large subduction zone earthquakes is typically on the order of 100 km, there is some evidence that the asperities occupy only the deeper half of the seismogenic width (see Schwartz and Ruff, 1987). While K_2/K_1 is 1 for all asperity pairs, F_2/F_1 causes asymmetry.

4.3.1. Further considerations of asperity interaction and system parameters

The above two model geometries do not exhaust the possibilities for system ratios, as will be illustrated in this short section. First of all, it is possible to increase the characteristic lengths of elastic interactions by allowing the asperity blocks to deform. The coupling parameter can either increase or decrease dependent upon the aspect ratio of asperity width to length. We shall not pursue this modification any further; rather we will consider a modified model that imparts a systematic shift to the coupling parameter.

The conceptual model for the main spring constants is two independent elastic bars (Fig. 6). Perhaps a more complete elastic model might have the two rigid asperity blocks pushing on the surface of an elastic half-space. The solution for a single square region pushing on a half-space has already been used to justify the choice for the elastic bar length for the main spring. If we examine this solution further, we see that the deflection of the surface returns back to zero over some distance outside of the square region. The characteristic length of this deflection is related to the asperity length. Thus, the displacement of one asperity tends to also displace the other asperity in the same direction, even without the direct coupling bar. This action serves as an additional coupling between asperities. If our model for elastic interaction were based on the

more complex system of blocks pushing on a half-space plus the direct coupling bar between the blocks, then the effective coupling would be stronger than the model of two independent elastic bars for the main springs. To provide a crude estimate of this effect, I have made several simplifying assumptions, including: approximate the functional dependence of surface deflection with a decaying exponential with asperity length as the scale length; and use an average asperity length for the two adjacent asperities. To then solve for the net force acting on asperity #1, the following expression can be obtained:

$$f_1 = -K_1(1-b)X_1 - (k + K_1 b)(X_1 - X_2)$$

where k and K_1 are the spring constants as previously described, and b is the non-dimensional interaction constant, $b = (\exp(-L/a) - \exp(-(L+a)/a))$ for the special case of $a_1 = a_2$. Note that the strength of the main spring is diminished by b, while the strength of the coupling spring is enhanced by b/κ. We can quantify this effect by forming a new ratio of the effective spring constants, $\kappa_{\text{mod}} = \kappa(1 + b/\kappa)/(1-b)$. Since b is always positive, κ_{mod} is always greater than κ. This modified coupling parameter is listed in Table 1, where we have modified the κ for the rigid cube model. The value of the coupling parameter is increased by a factor of 1.5 to 3 for the particular examples in Table 1.

To summarize all of the above considerations, Figure 7 shows the dependence of F_2/F_1, K_2/K_1, and κ upon $(\bar{a}/L)$ for the observed asperity distributions. Values for the coupling parameters correlate with the asperity length to separation ratio, as expected. The modified coupling parameter is distinctly larger than the other two cases, which produce similar values for κ. One could argue that the modified coupling parameter is a more appropriate representation of the elasticity, but I will show calculations for both the low and high coupling options. No two asperities are identical, thus there is some asymmetry for all asperity pairs. The rigid cube model produces more asymmetry in the F_2/F_1, K_2/K_1 ratios than the rigid block model for the same $(\bar{a}/L)$. The lower part of Figure 7 shows that most values of the system ratios range within a factor of two about 1.

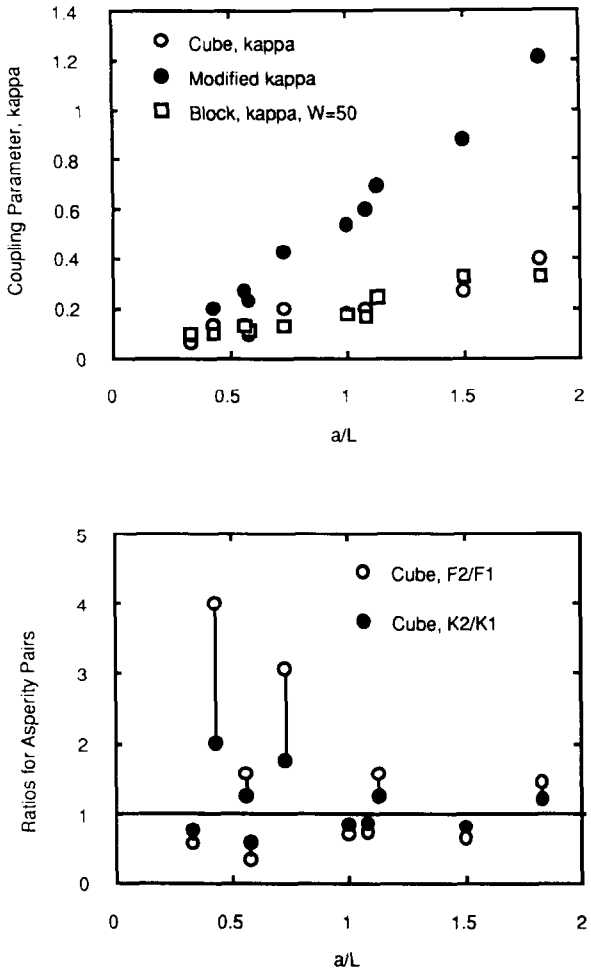


Fig. 7. System ratios for observed adjacent asperity pairs, plotted as a function of $\bar{a}/L$, ratio of average asperity length to separation. Upper diagram plots three different determinations of the coupling parameter. Lower diagram plots the other two system variables for the "Rigid Cube" model. These ratios would be 1 for symmetric identical asperities.

Hence, while no two adjacent asperities are identical, their sizes are comparable such that the larger asperity does not completely dominate the system behavior. The importance of non-identical asperities will be demonstrated later on with artificial "earthquake" sequences.

4.4. Failure criteria

Unlike the spring constants, failure criteria for the frictional sliders can not be specified from our observed asperity distributions. Recall that large asperities in subduction zones can be 50 km

or larger in their along-strike length. Therefore, it takes more than 15 s for the rupture front to sweep across a single asperity. It is quite possible that one end of the asperity is just beginning to slip while the other end may have already stopped slipping. Of course, in the discrete element system, a single frictional slider moves as one rigid block. Therefore, we can see that any failure criteria applied to a frictional slider in our discrete element system must represent the overall average slip and/or stress-drop of the asperity region. Other seismologists have used discrete element systems to represent the time-dependent nature of earthquake rupture, but this application requires a large number of coupled frictional sliders of sufficiently small dimension to adequately represent the phenomena of seismic waves and rupture fronts (e.g., Mikumo and Miyatake, 1978). For the present application, all these complicated unknown effects are replaced by an overall macroscopic rupture criterion. Indeed, the macroscopic failure criteria may be quite different from a detailed microscopic description of a small portion of the interface. We can use some features of large earthquake occurrence to guide us in our choice of macroscopic failure criteria. We will first focus on failure criteria for a single event, i.e. slip of a single frictional slider.

Single-asperity failure. A single-asperity failure refers to a slip event of the discrete element system where just one asperity, either #1 or #2, slips. Failure criteria can be applied to either asperity slip, asperity stress-drop, or asperity final stress. For single-asperity failure, I use the simplest criteria: the asperity fails at a constant value of failure stress, and final stress is zero. Since asperity stress is directly related to the net force acting on the asperity, $\sigma = F/A$ where σ is stress, F is force, and A is asperity area, we shall discuss failure in terms of the force. Specify the failure force on asperity #1 as F_1. Assume that the asperity offset $(X_1 - X_2)$ is positive such that asperity #1 will fail first as the main springs are compressed by the asperities resting on the lower plate. The total force, main spring plus coupling spring, acting on asperity #1 will eventually reach F_1 and an earthquake happens. Now, what is the position of the asperity after the earthquake? The simplest rule of zero final stress requires X_1 to slip until the net force on asperity #1 is zero. Because of the coupling spring, the final value of X_1 will be somewhere between 0 and X_2, position of asperity #2, since the coupling spring will act opposite to the main spring. This position is graphically represented by the zero-force line in Figure 5. Thus, for single-asperity failure, the final net force will be zero, the force-drop will be F_1, and the asperity slip will be $\delta X_1 = F_1/(K_1 + k)$. The same discussion for asperity #2 also defines $\delta X_2 = F_2/(K_2 + k)$. With this simple failure criterion, the stress-drop and slip will be the same for every single-asperity "earthquake" for a particular asperity. However, as we shall see in the artificial "earthquake" sequences, this does not imply that the recurrence time between single-asperity "earthquakes" is always the same.

Double-asperity failure. If the asperity offset is smaller than a critical value, then the slip of asperity #1 increases the net force on asperity #2 such that F_2 is exceeded (see Fig. 8). In that case, asperity #2 will fail and a "double event" earthquake will result, where the rupture length includes both fault segments. If we specify that the slip of asperity #1 is δX_1 (same as for single-asperity failure), then the final net force on asperity #1 will no longer be zero since asperity #2 also slips. Therefore, the simplest failure criteria of constant slip for the asperities regardless of whether the "earthquake" is a single or double event implies that: (1) the net force on both asperities will be non-zero after a double event, and (2) thus the net asperity force (stress)-drop of both asperities will be lower for a double event as compared to a single event. The above choices for double-asperity failure are referred to as failure mode I, displayed graphically in Figure 8. Other choices for the double event failure criteria can be made. For example, if one thinks that the triggered asperity should continue to slip until its net force is zero before it heals, then failure mode II would be: (i) the first-failed asperity slips the same amount as for a single event, then it heals and does not slip further; (ii) given the new position of the first-failed asperity, the other triggered asperity continues to slip until the net force is zero. In general, the slip of the triggered asper-

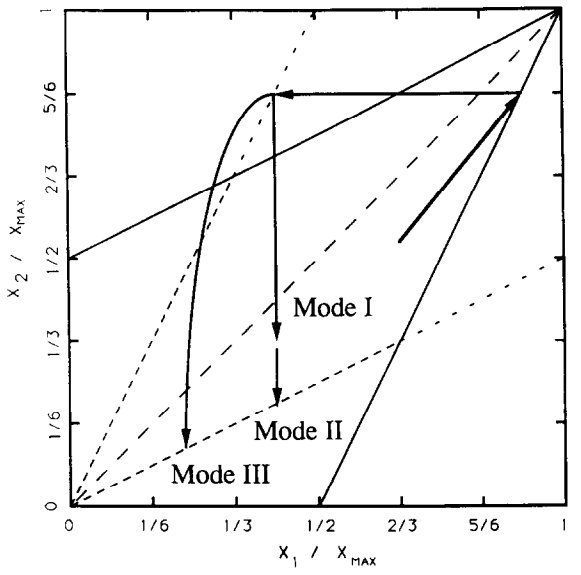


Fig. 8. Three modes for double event failure criteria. System behavior is shown on a (X_1, X_2) diagram; same parameters as for Fig. 5. A particular path is chosen for a small asperity offset that intersects the failure line for asperity #1 for X_1/X_{max} of 5/6. When asperity #1 slips to the zero-force line (dashed), the path crosses the failure line for asperity #2, which will then slip. Three different possibilities for the final system state are described in the lower part, with their paths depicted in the (X_1, X_2) diagram.

ity is greater for a double event than it would be for a single event (see Fig. 8), hence the average slip for a double event is greater than the average slip for single events in the two segments. If one thinks that the healing of rupture for the first-failed asperity is not completed before the other asperity slips, then failure mode II is modified to failure mode III, where both asperities slip more than they would in single events. Failure mode III then requires even further specification. Perhaps the simplest version of failure mode III is where the asperity positions are reset to zero. Several failure criteria could be defended as the best choice to represent the overall behavior of double event rupture. I will not argue that any one particular choice is best, I will simply show some of the properties and consequences of failure modes I and II in Figure 8 and the following section.

4.5. Modeling results: simulated "earthquake" sequences

Asperity distributions along several subduction zones show that the asperity length to separation ratios for adjacent asperities mostly fall in the range from 1/2 to 2 (Table 1 and Fig. 7). This represents moderate coupling between the subduction zone segments. If asperity length were an order of magnitude smaller than asperity separation, then adjacent asperities would be only weakly coupled. On the other hand, if asperity size were an order of magnitude larger than their separation, then adjacent asperities would be very strongly coupled, and therefore they would nearly always rupture together and act more like one larger asperity. The fact that the observed ratios of asperity size to separation is about 1, already implies that asperity interaction may produce interesting effects such as combinations of single and multiple events.

The discrete element system can be completely described by three non-dimensional ratios, plus the failure criteria, plus an initial offset of the asperities. It is remarkable that this simple system of two coupled frictional sliders displays a broad range of simple and complex behavior. Previous work on two coupled frictional sliders include: elucidation of some properties of the symmetric system of identical sliders by Nussbaum and Ruina (1987); a demonstration of complex behavior due to complex failure criteria by Huang and Turcotte (1990); and a comprehensive "global" analysis of the system behavior for the simplest version of failure mode III by Lomnitz-Adler and Perez Pascual (1989). The main contributions of my study to the mechanics of two coupled frictional sliders are of a practical nature: (a) I am trying to explain some observed features of large earthquake occurrence, where the frictional sliders represent asperities; (b) ratios of spring constants and failure forces are based on observed asperity distributions; (c) emphasis of the fact that simple macroscopic failure criteria can produce "earthquake" sequences that are irregular, especially when just a few earthquake cycles are considered. Rather than jumping immediately to complex system behavior, we will start with a

system specification that produces regular slip sequences.

4.5.1. Identical adjacent asperities with failure mode I

It is easy to see the complete system behavior for all initial conditions for the case of identical asperities and failure mode I. The repetitive system behavior is plotted on a diagram of X_2 versus X_1 (Fig. 9). Since the asperities are identical, $F_1 = F_2 = F$ and $K_1 = K_2 = K$. The ratio of k/K can still vary. Both X_1 and X_2 are normalized to $X_{max} = F/K$. The values of (X_1, X_2) that produce a net force of F on asperity #1 describe the "failure surface" for asperity #1; a straight line as plotted in Figure 9 for $\kappa = 1$. The failure surface for asperity #2 is also plotted. Given some initial starting position, the system follows a straight line in the (X_1, X_2) diagram with a slope of one as plate tectonic motions proceed. If the initial asperity offset $(X_1 - X_2)$ is positive, then asperity #1 will fail first; if it is negative, then asperity #2 will be the first to fail. Figure 9 shows the path followed by single event failures.

If the absolute value of the asperity offset is less than a certain value (Fig. 10), then the failure

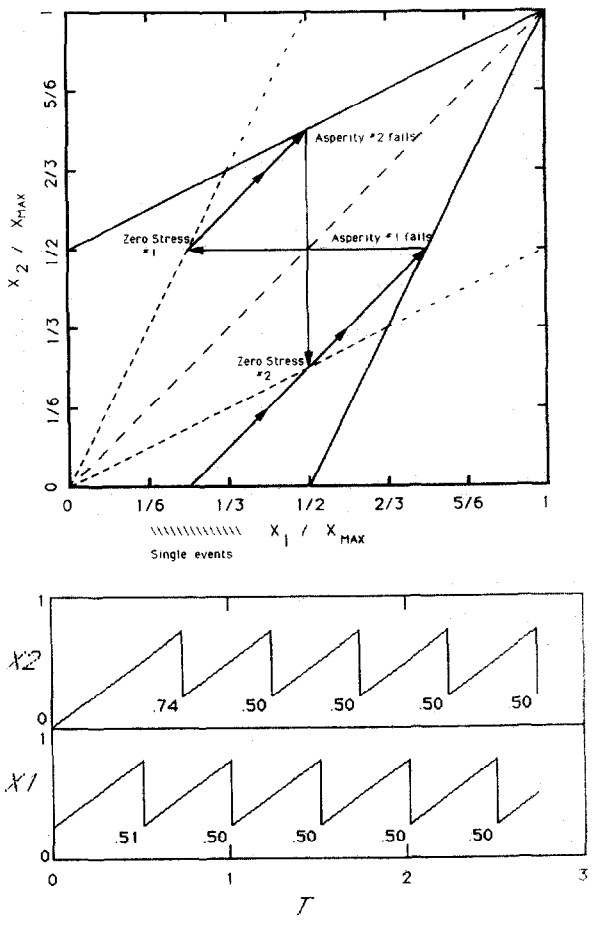


Fig. 9. Identical asperities produce identical alternating single-asperity failures. Top plot, for $\kappa = 1$, shows the system path in the (X_1, X_2) diagram. Symmetry of the failure and zero-force lines causes a two-cycle loop. Lower plot shows X_1 and X_2 as a function of time, all variables are nondimensionalized. X_1 and X_2 increase until a slip event occurs, then either X_1 or X_2 slips by $X_{max}/2$. The numbers just to the left of the slip events are the recurrence times. A sequence of alternating identical single events is produced, similar to the "uniform" model of Figure 2a.

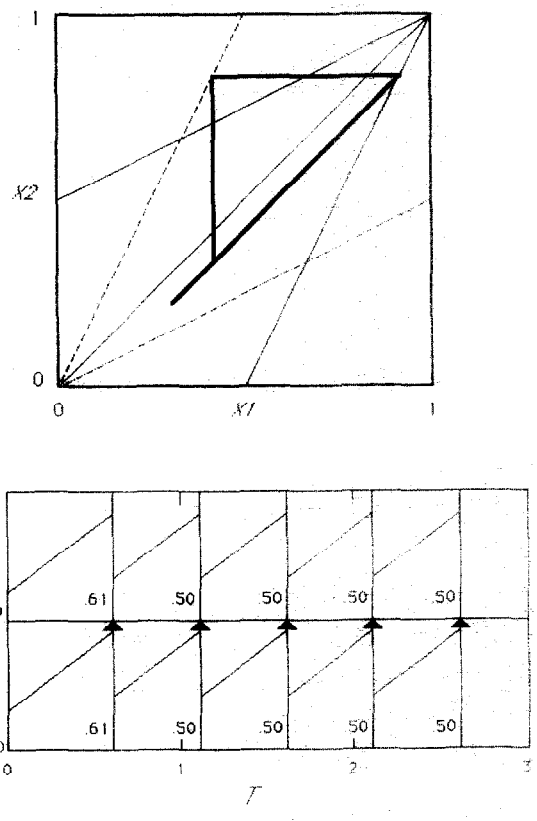


Fig. 10. Identical asperities produce identical double events for failure mode I. Same plots as Figure 9, except that initial asperity offset is in the double event range. Failure mode I specifies constant slip for both asperities for double event failure, thus the system path returns to the initial path. Lower plot shows a few of the endless sequence of identical double events. Double events are indicated by the continuous vertical line. The arrowhead points in the direction of rupture, i.e. toward the triggered asperity.

of one asperity triggers the failure of the other. Since failure mode I specifies that asperity slip will always be $F/(K+k)$, the system position after a double event failure is on the same trajectory. In other words, asperity offset has the same value after a double event as it did before. Therefore, a sequence of identical double events will be the result.

Because the two asperities are identical, the failure lines are symmetric about the central diagonal. This symmetry causes the regular behavior of slip sequences. The only remaining question is if the initial offset is in the range to yield either alternating single events, or double events. The lower part of Figure 10 shows that a regular sequence of identical double event "earthquakes" is produced for identical asperities with failure mode I.

4.5.2. Identical adjacent asperities with failure mode II

Failure mode II allows the second-failed asperity of a double event to slip until the zero-force line is achieved. Of course, if the initial asperity offset is within the single event regime, then an alternating sequence of identical single events will result, just as above. If the initial asperity offset is in the double event regime, then the first "earthquake" will indeed be a double event (Fig. 11). However, notice that the (X_1, X_2) value after the double event has a larger asperity offset than the initial offset. This shift in asperity offset is a direct result of failure mode II, and it places the system in the single event regime. There is only one special value of initial asperity offset that will produce a continuing sequence of double events; the trajectory of this offset hits the intersection of the failure and zero-force lines. Therefore, a system with identical asperities and failure mode II will, in general, produce only sequences of alternating identical single events.

As a general conclusion, it is quite apparent that two identical asperities with simple macroscopic failure rules will produce sequences of identical slip events. Recall that observed earthquake sequences do not behave in this manner. Adjacent identical asperities represent a special case of the discrete element system; any variation

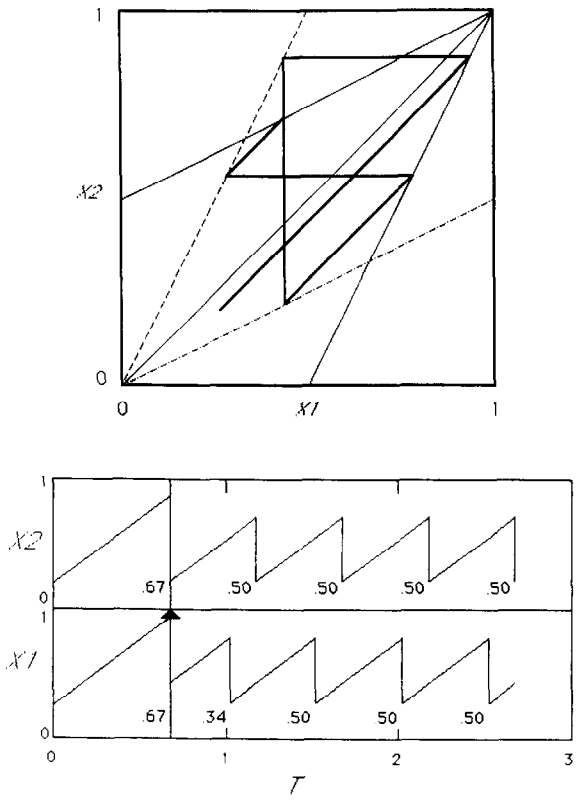


Fig. 11. Identical asperities and failure mode II. Same plots as Figures 9 and 10. Initial asperity offset is in the double event range, but failure mode II causes an increase in asperity offset such that the new path is in the alternating single event range. The sequence of slip events is shown below in the plot of X_1 and X_2 versus T.

in the ratios of F_1/F_2 and K_1/K_2 will break the symmetry. Of course, identical adjacent asperities would be an improbable occurrence in nature. Indeed, a glance at Table 1 shows that even the most basic parameter of asperities, their along-strike length, varies from one asperity to the next. This variation in the along-strike length of asperities is the primary cause of non-identical asperities in our model. We must now consider the more complicated, but more realistic, case of non-identical asperities.

4.5.3. Non-identical asperities with failure mode I

If the two asperities are not identical, then the F_2/F_1 and K_2/K_1 ratios differ from 1. The symmetry of the failure and zero-force lines in the (X_1, X_2) diagram is broken. One consequence of

the asymmetry is that the asperity offset changes after single or double event cycles. Thus, the system path in the (X_1, X_2) diagram will range over a more complicated pattern that includes both single and double events. The pattern of single and double events might eventually repeat itself (see Lomnitz-Adler and Perez Pascual, 1989), but the number of earthquake cycles in the repeating pattern can be quite large. In detail, the number of cycles in the repetitive pattern depends on the combination of F_2/F_1, K_2/K_1 and κ ratios. Certain combinations of these ratios can produce a small number of cycles in a pattern. However, recall that for our particular application we rarely observe more than two or three large earthquake cycles in a subduction zone segment. Therefore, the discrete element system can successfully produce variable recurrence times and rupture mode.

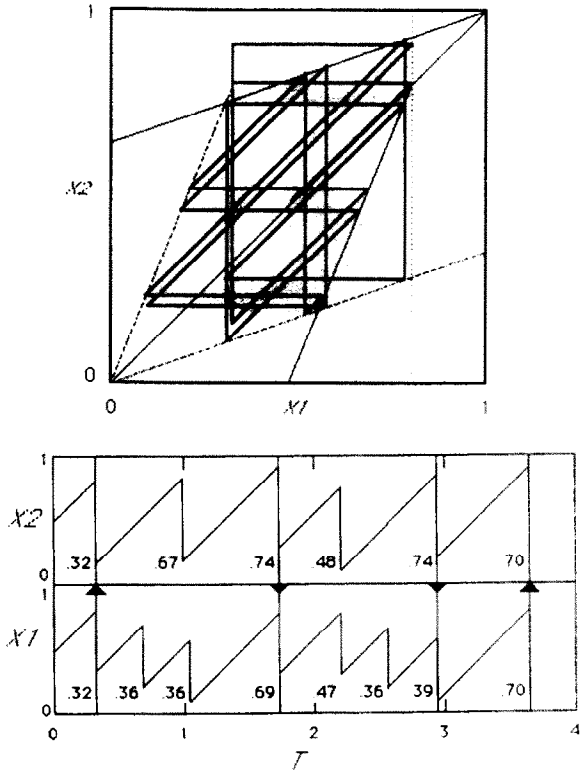


Fig. 12. Non-identical asperities with failure mode I, simulation for the asperity pair of the great 1974 Peru earthquake. System variables are: $F_2/F_1 = 1.56$, $K_2/K_1 = 1.25$, $\kappa = 0.69$; these values are for the (P74a P74b) asperity pair for the rigid cube model with high coupling (see Table 1). Plots follow the same scheme as Figures 9 through 11. Non-identical asperities produce asymmetry in the (X_1, X_2) diagram at top, and the system path wanders over a broad range of asperity offsets. Lower plot shows that recurrence times vary, as does rupture mode with a mixture of single and double events.

Since there is an infinity of choices for non-identical system ratios, we might as well choose system ratios from the asperity pairs listed in Table 1. The basic system behavior is illustrated for the (P74a P74b) asperity pair for the rigid cube model with high coupling. Several earthquake cycles for an arbitrary starting position are shown in Figure 12. Note that there is a mixture of single and double events, and the recurrence times vary between 0.36 and 0.74 in normalized time units (disregarding the initial time delay). If we adjust the average recurrence time, 0.55, to be 100 years (a typical value for subduction zones), then the recurrence times scatter from 65 to 134 years. Of course, different earthquake sequences might result from different starting values of asperity offset, but certain patterns occur that are characteristic for the system ratios. Clearly, non-identical asperities with failure mode I can produce "earthquake" sequences that match observed features #1 and #2: variable recurrence times and rupture mode.

4.5.4. Non-identical asperities with failure mode II

Failure mode II is designed to match observed feature #3, i.e. variable slip. Figure 13 shows several earthquake cycles for the same system ratios as Figure 12. Once again, the earthquake sequence consists of both single and double events, with variable recurrence times. In addition, we are guaranteed that the "seismic moment" of a double event is greater than the sum of "seismic moments" of single events in both segments. All three observed features of Figure 1 are satisfied by the artificial "earthquake" sequence of Figure 13. Note that asperity #2 always breaks in a double event; this characteristic seems to be true for all possible earthquake sequences for the particular values of the system ratios. Other choices of system ratios can produce a more irregular pattern of "earthquake" occurrence. Also note that the scatter in recurrence times for the larger asperity (#2) is less than the

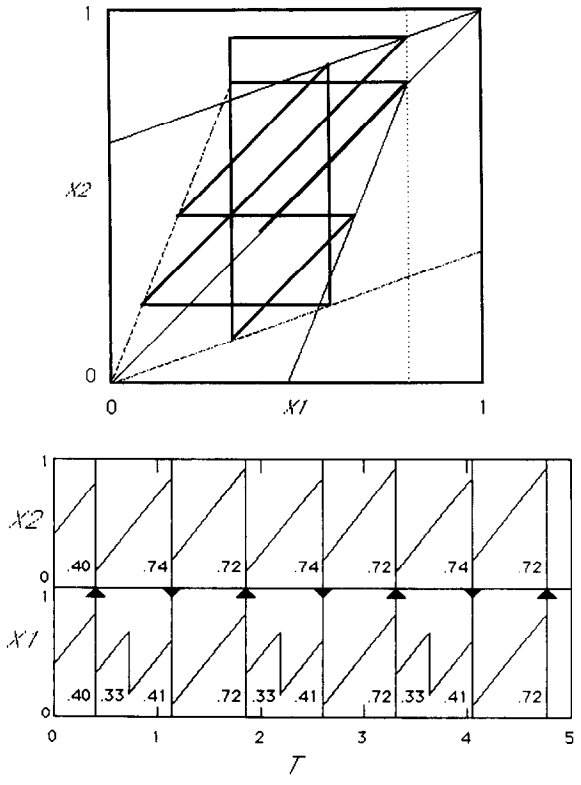


Fig. 13. Non-identical asperities with failure mode II, simulation for the asperity pair of the great 1974 Peru earthquake. Same as Figure 12, but rupture mode II is used for the double events. The slip event sequence in the lower plot satisfies all three observed features of large earthquake occurrence: (1) recurrence time varies, (2) rupture mode varies, and (3) average slip for double events is larger than for single events.

scatter in times for the smaller asperity (#1). This seems to be a somewhat general characteristic of simulated "earthquake" sequences.

4.5.5. Simulated "earthquake" sequences and the time / slip-predictable model

Before we move on to statistical summaries of the "earthquake" sequences, it is interesting to view the diagrams of X_1 and X_2 versus time in Figures 12, 13, and 14 in the context of the time/slip-predictable model. If a straight line could be drawn through the "peaks" of the asperity position just before slip occurs (see dotted lines in Fig. 14), then the time-predictable model could be successfully applied to the "earthquake" sequences in the individual "fault segments". On the other hand, if a straight line were able to fit all of the "troughs" in asperity position just after

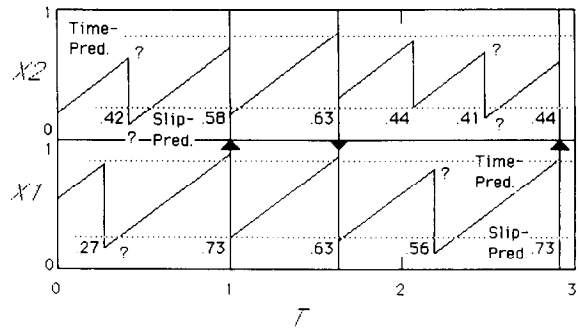


Fig. 14. Non-identical asperities with failure mode I, simulation for the Kuriles (K63a K63b) asperity pair. System ratios are for the rigid cube model with high coupling ($F_2/F_1 = 0.69$, $K_2/K_1 = 0.83$, $\kappa = 0.54$). Similar to sequences in Figures 12 and 13, a combination of single and double events with variable recurrence times is produced. Dotted lines are added to this plot to show that the above slip event sequence is more irregular than expected by the time-predictable or slip-predictable models.

slip events, then the slip-predictable model would successfully explain the "earthquake" sequences (see dotted lines in Fig. 14). A quick glance at Figures 12 through 14 shows that neither end-member model is applicable to these "earthquake" sequences, which are generated by the simplest interaction of two adjacent fault segments with the simplest failure criteria. Given the fact that the end-member cases of the time/slip-predictable model do *not* explain the scatter in recurrence times for these simplistic artificial "earthquake" sequences, it would be quite surprising if either end-member model were able to explain recurrence time scatter for real earthquake sequences.

4.6. Discussion

We now know that the discrete element system, with system ratios for observed asperity pair (P74a P74b) and failure mode II, can produce irregular "earthquake" sequences with variable recurrence time, rupture mode, and slip. There is not enough space here for hundreds of figures that would show all of the "earthquake" sequences for various starting positions for all ten asperity pairs for the different system ratios listed in Table 1. We need to find a greatly condensed representation of all the "earthquake" sequences.

Recall that our primary interest is in producing variable recurrence times and a mixture of single and double events. As for observed feature #3, "earthquake" sequences generated with failure mode II automatically satisfy this feature, while those with failure mode I do not. Thus, let us focus our attention on variable recurrence times and rupture mode. For each "earthquake" sequence, we can calculate average recurrence time and its standard deviation. The absolute value of standard deviation is not so important; we are more interested in the ratio of standard deviation to average recurrence time (see Nishenko and Buland, 1987, for discussion of this normalization relevant to observed recurrence times). With this ratio, the standard deviation can be scaled to a particular choice of average recurrence time, such as 100 years. Double event percentage is defined by dividing the number of double events ($\times 100$) by the total number of double and single events in any given "earthquake" sequence. Thus, "earthquake" sequences such as those in Figures 12, 13, and 14 can be replaced by two numbers: (1) normalized standard deviation of recurrence times, and (2) double event percentage. Closer scrutiny of Figures 12 through 14 reveals that average recurrence time and its scatter are different for the two asperities. Thus, the normalized recurrence time scatter is calculated separately for the two asperities, though a single value of double event percentage is retained. To mimic observations of real earthquake sequences, measurements of recurrence time scatter and double event percentage are based only on the last few slip events in an artificial sequence—a minimum of four events for each asperity. To produce better statistical estimates and to minimize the influence of starting values of (X_1, X_2), the last few earthquakes of several different sequences are averaged together for each asperity.

Figure 15 shows three plots for three asperity pairs, one each from the Kuriles, Colombia, and Peru subduction zones. The hachured regions show the range of statistical values produced for a given set of system ratios. The statistical values for the rigid block model overlap those for the low-coupling rigid cube model, probably because their coupling parameters are similar, and hence

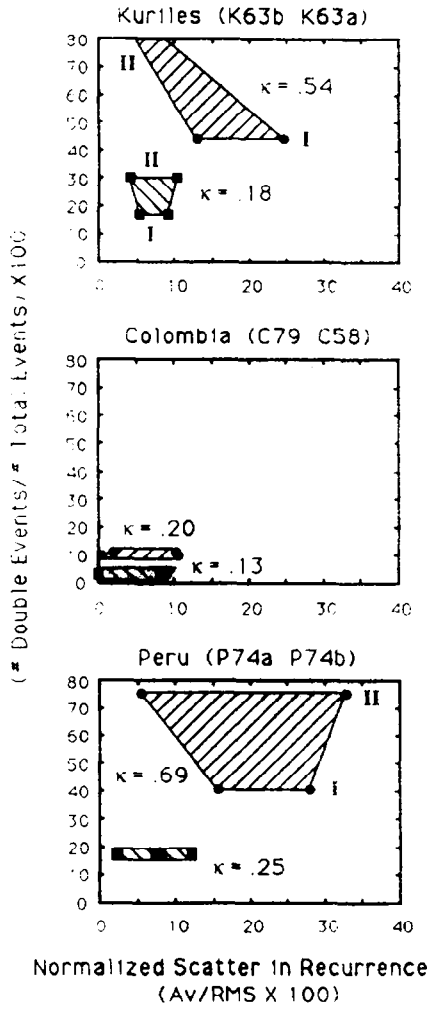


Fig. 15. Summary statistics of artificial "earthquake" occurrence for asperity pairs in the Kuriles, Colombia, and Peru subduction zones. Asperity pairs are identified for each plot of double event percentage versus recurrence time scatter. If the average recurrence time is 100 years for all cases, then the numbers listed for recurrence time scatter would be the standard deviation in years. All earthquake sequences are for the rigid cube model (see Table 1) for both low and high values of coupling, κ. The four corners of the hachured regions show the range in recurrence time scatter and double event percentage for both asperities and failure modes. As an overall statistical conclusion, higher values of κ cause a higher percentage of double events and a greater scatter in recurrence times (see text).

the rigid block results are not plotted. It is possible to make some general statements based on the results for these three arbitrary, yet typical, asperity pairs. First of all, it is no surprise that increasing the coupling parameter systematically

increases the percentage of double events. For coupling parameters of around 1/5, the double event percentage ranges from about 10 to 30%. In detail, one curious observation is that the asymmetry of the asperity pairs increases from (K63b K63a) to (P74a P74b) to (C79 C58), while the ranking of double event percentage would be the inverse of this order for comparable coupling parameters (see Fig. 15). I do not know if this observation is a general rule. The two high values of coupling for the Kuriles and Peru asperity pairs produce double event percentages of 40% and higher. Indeed, for the Kuriles asperity pair with $\kappa = 0.54$ and failure mode II, the "earthquake" sequences would always evolve to 100% double events with identical recurrence times, hence recurrence time scatter is zero. Certain combinations of non-identical system ratios with failure mode II can produce this behavior.

Another general characteristic is that failure mode II tends to produce a higher double event percentage than failure mode I. However, for low values of coupling, the double event percentage seems to be about the same for both failure modes. This might be due to the fact that the difference in slip between failure modes I and II decreases for decreasing κ. Normalized scatter in recurrence times tends to increase for a large κ. On the other hand, the special situation for the Kuriles asperity pair with high coupling and failure mode II produces a normalized scatter of zero.

Given the range in the coupling parameters and asymmetry of the asperity pairs in Table 1 and the options of failure modes I and II, simulated "earthquake" sequences can produce double event percentages between 0% and 100%, and recurrence time scatter up to 30% of average recurrence time. The simple discrete element system can produce a great range of irregular "earthquake" behavior. One curious observation is that the discrete element system apparently does *not* produce "earthquake" sequences with a low percentage of double events and a large percentage scatter in recurrence times.

It is tempting to plot "observed" values of double event percentage and recurrence time scatter in the plots of Figure 15. However, there are several reasons not to do so, for example: (1) the limited number of large earthquake cycles in subduction zones does not allow us to make reliable estimates of statistical parameters; (2) plotting "data points" in Figure 15 might give the impression that it is possible to specify system ratios that mimic particular earthquake sequences in particular subduction zones. With respect to the latter reason, it is important to remember the simplifications employed to go from subduction zone asperities to a pair of coupled frictional sliders. Even with this simple system, it appears that different combinations of system ratios can produce comparably irregular "earthquake" sequences. The extent of this non-uniqueness is difficult to explore, but it is certainly possible to produce a particular short sequence of "earthquakes" with different model descriptions. Despite the above objections, perhaps we can briefly mention the overall character of multiple event percentage and recurrence time scatter for the Kuriles, Colombia, and Peru subduction zones. Recurrence time scatter ranges from less than 10% to more than 40% of average recurrence times, and *multiple* event percentage ranges from 25% to about 50% for these subduction zones. While these numbers are reasonably compatible with the model statistical results presented in Figure 15, I do not wish to make any quantitative conclusions based on this comparison. However, I am willing to conclude that the asperity model for large earthquake occurrence passes the scientific test of self-consistency. The simplest model of asperity interaction with system variables based on observed asperity distributions can produce artificial "earthquake" sequences that display three of the observed features of large earthquake occurrence: variable recurrence time, rupture mode, and coseismic slip.

5. Conclusions

Observations of large underthrusting earthquakes in subduction zones indicate that the seismogenic plate interface is heterogeneous in its mechanical properties. Three important fundamental observations of large earthquake occurrence are: (1) recurrence time in any one segment

will vary between successive earthquake cycles; (2) rupture length of great earthquakes will vary between successive earthquake cycles; and (3) the coseismic slip in any one segment can be larger if the rupture length is greater. The asperity model provides for a simple first-order classification of the along-strike failure strength of the plate interface, and the role that the strong regions (asperities) play during the earthquake cycle. Rupture process studies have determined the first-order asperity distribution along several subduction zones. These results indicate that the geometric ratio of along-strike asperity length to separation distance for ten asperity pairs varies between 1/2 and 2. The fact that this ratio is close to 1 for these subduction zones implies that interaction of adjacent asperities might control some aspects of large earthquake occurrence. To test this idea, I use the simplest model of asperity interaction.

The simplest representation of two adjacent interacting asperities is a discrete element model of two frictional sliders coupled to the upper plate and to each other by elastic springs, with a frictional contact on the subducting lower plate. All coupling between the plates takes place through the area of the asperities, i.e. the frictional sliders. The spring constants are specified by the observed asperity distributions. The macroscopic failure criteria for single and double events are not well-constrained, but simple reasonable choices can be specified. Sequences of "earthquakes", i.e. frictional slider slip events, can then be produced. It is shown that this two-asperity discrete element model with non-identical asperities and simple failure criteria can easily produce artificial "earthquake" sequences with: (1) variable recurrence times in each asperity segment, (2) variable rupture length, i.e. a mixture of single events and double events, and (3) variable seismic slip between "earthquake" cycles. Thus, the asperity model for large underthrusting earthquakes passes the scientific test posed by the above basic observations of large earthquake occurrence. In contrast, the time-predictable and slip-predictable models fail to explain the variable recurrence times of the artificial "earthquake" sequences generated by the simple model of asperity interaction. On the other hand, many other models of strength heterogeneity and interaction could probably be constructed to explain the above basic observations. An optimistic conclusion, that is independent of any particular model, is that complicated patterns of earthquake recurrence may have a simple underlying deterministic explanation.

Acknowledgements

I thank the hosts and participants of the International Symposium on Earthquake Source Physics and Earthquake Precursors; this symposium provided exciting and stimulating "interaction". Earthquake studies at the University of Michigan are supported by the National Science Foundation (EAR-8720955 and EAR-9019003).

References

Aki, K., 1979. Characterization of barriers on an earthquake fault. J. Geophys. Res., 84: 6140–6148.

Aki, K. and Richards, P.G., 1980. Quantitative Seismology. Freeman, San Francisco, 932 pp.

Ando, M., 1975. Source mechanisms and tectonic significance of historical earthquakes along the Nankai Trough. Tectonophysics, 27: 119–140.

Beck, S. and Ruff, L., 1987. Rupture process of the great 1963 Kurile Islands earthquake sequence: asperity interaction and multiple event rupture. J. Geophys. Res., 92: 14123–14138.

Beck, S., and Ruff, L., 1989. Great earthquakes and subduction along the Peru trench. Phys. Earth Planet. Inter., 57: 199–224.

Das, S. and Aki, K., 1977. Fault planes with barriers: A versatile earthquake model. J. Geophys. Res., 82: 5658–5670.

Fedotov, S., 1965. Regularities of the distribution of strong earthquakes in Kamchatka, Kurile Islands, and northeast Japan. Tr. Inst. Fiz. Acad. Nauk SSSR, 36: 66–93.

Huang, J. and Turcotte, D., 1990. Are earthquakes an example of deterministic chaos? Geophys. Res. Lett., 17: 223–226.

Ishibashi, K., 1981. Speculation of a soon-to-occur seismic faulting in the Tokai district, central Japan, based on seismotectonics. In: D. Simpson and P. Richards (Editors), Earthquake Prediction—An International Review. AGU, Washington D.C., pp. 297–332.

Johnson, K.L., 1985. Contact Mechanics. Cambridge University Press, Cambridge, U.K., 452 pp.

Kanamori, H., 1978. Quantification of earthquakes. Nature, 271: 411–414.

Kanamori, H., 1981. The nature of seismicity patterns before large earthquakes. In: D. Simpson and P. Richards (Editors), Earthquake Prediction—An International Review. AGU, Washington D.C., pp. 1–19.

Kanamori, H., 1986. Rupture processes of subduction zone earthquakes. Annu. Rev. Earth Planet. Sci., 14: 293–322.

Kanamori, H. and Anderson, D.L., 1975. Theoretical basis of some empirical relations in seismology. Bull. Seismol. Soc. Am., 65: 1073–1095.

Kanamori, H. and McNally, K.C., 1982. Variable rupture mode of the subduction zone along the Ecuador–Colombia coast. Bull. Seismol. Soc. Am., 72: 1241–1253.

Kelleher, J., 1972. Rupture zones of large South American earthquakes and some predictions. J. Geophys. Res., 77: 2087–2103.

Kelleher, J., Sykes, L. and Oliver, J., 1973. Possible criteria for predicting earthquake locations and their application to major plate boundaries of the Pacific and Caribbean. J. Geophys. Res., 78: 2547–2585.

Kikuchi, M. and Fukao, Y., 1987. Inversion of long-period P waves from great earthquakes along subduction zones. Tectonophysics, 144: 231–247.

Lay, T., Kanamori, H. and Ruff, L., 1982. The asperity model and the nature of large subduction zone earthquakes. Earthquake Pred. Res., 1: 3–71.

Lomnitz-Adler, J. and Perez Pascual, R., 1989. Exactly solvable two-fault model with seismic radiation. Geophys. J. Int., 98: 131–141.

Matsu'ura, M. and Sato, T., 1989. A dislocation model for the earthquake cycle at convergent plate boundaries. Geophys. J., 96: 23–32.

McCann, W., Nishenko, R., Sykes, L. and Kraus, J., 1979. Seismic gaps and plate tectonics: seismic potential for major plate boundaries. Pure Appl. Geophys., 117: 1087–1147.

Mikumo, T. and Miyatake, T., 1978. Dynamical rupture process on a three-dimensional fault with non-uniform frictions and near-field seismic waves. Geophys. J., 54: 417–438.

Mogi, K., 1968. Some features of recent seismic activity in and near Japan. Bull. Earthquake Res. Inst., Tokyo, 46: 1225–1236.

Nishenko, S., 1985. Seismic potential for large and great interplate earthquakes along the Chilean and southern Peruvian margins of South America: a quantitative reappraisal. J. Geophys. Res., 90: 3589–3615.

Nishenko, S. and Buland, R., 1987. A generic recurrence interval distribution for earthquake forecasting. Bull. Seismol. Soc. Am., 77: 1382–1399.

Nussbaum, J. and Ruina, A., 1987. A two degree-of-freedom earthquake model with static and dynamic friction. PAGEOPH, 125: 629–656.

Rikitake, T., 1976. Recurrence of great earthquakes at subduction zones. Tectonophysics, 35: 335–362.

Ruff, L., 1983. Fault asperities inferred from seismic body waves. In: H. Kanamori and E. Boschi (Editors), Earthquakes: Observation, Theory, and Interpretation. North-Holland, Amsterdam, pp. 251–276.

Ruff, L., 1989. Do trench sediments affect great earthquake occurrence in subduction zones? In: L. Ruff and H. Kanamori (Editors), Subduction Zones, II. Pure Appl. Geophys., 129: 263–282.

Ruff, L. and Kanamori, H., 1983. Seismic coupling and uncoupling at subduction zones. Tectonophysics, 99: 99–117.

Rundle, J. and Kanamori, H., 1987. Application of an inhomogeneous stress (patch) model to complex subduction zone earthquakes: a discrete interaction matrix approach. J. Geophys. Res., 92: 2606–2626.

Satake, K. and Kanamori, H., 1991. Use of tsunami waveforms for earthquake source study. Nat. Hazards, 4: 193–208.

Savage, J.C., 1983. A dislocation model of strain accumulation and release at a subduction zone. J. Geophys. Res., 88: 4984–4996.

Schwartz, S. and Ruff, L., 1987. Asperity distribution and earthquake occurrence in the southern Kurile Islands arc. Phys. Earth Planet. Inter., 49: 54–77.

Shimazaki, K. and Nakata, T., 1980. Time-predictable recurrence model for large earthquakes. Geophys. Res. Lett., 7: 279–282.

Sykes, L.R. and Quittmeyer, R.C., 1981. Repeat times of great earthquakes along simple plate boundaries. In: D. Simpson and P. Richards (Editors), Earthquake Prediction —An International Review. AGU, Washington D.C., pp. 217–247.

Thatcher, W., 1990. Order and diversity in the modes of circum-Pacific earthquake recurrence. J. Geophys. Res., 95: 2609–2623.

Thatcher, W. and Rundle, J., 1984. A viscoelastic coupling model for the cyclic deformation due to periodically repeated earthquakes at subduction zones. J. Geophys. Res., 89: 7631–7640.

Tichelaar, B. and Ruff, L., 1991. Seismic coupling along the Chile subduction zone. J. Geophys. Res., 96: 11,997–12,022.

Utsu, T., 1974. Space-time pattern of large earthquakes occurring off the Pacific coast of the Japanese islands. J. Phys. Earth, 22: 325–342.

Uyeda, S. and Kanamori, H., 1979. Back-arc opening and the mode of subduction. J. Geophys. Res., 84: 1049–1061.