BIBLIOGRAPHIC INFORMATION

PB95-183968

Report Nos:

Title: Generalized Semi-Markov Process for Modeling Spatially and Temporally Dependent Earthquakes.

Date: cJul 93

Authors: K. A. Lutz and A. S. Kiremidjian.

Performing Organization: Stanford Univ., CA. John A. Blume Earthquake Engineering Center.

Sponsoring Organization: *National Science Foundation, Arlington, VA.*Electric Power Research Inst., Palo Alto, CA.

Contract Nos: EPRI-RP2356-91, NSF-EID-9024032

Supplemental Notes: Also pub. as Stanford Univ., CA. John A. Blume Earthquake Engineering Center rept. no. REPT-104.

NTIS Field/Group Codes: 48F (Geology & Geophysics)

Price: PC A06/MF A02

Availability: Available from the National Technical Information Service, Springfield, VA. 22161

Number of Pages: 102p

Keywords: *Site characterization, *Earthquakes, *Markov processes, Earthquake engineering, Tectonics, Faults(Geology), Forecasting, Earth movements, Dynamic response, Structural geology, Mathematical models, Seismic hazards.

Abstract: Site-specific hazard estimation requires the modeling of the occurrences of earthquakes on any faults with the potential to impact the site. Previous earthquake occurrence models have assumed either spatial independence or temporal independence or both. However, for large magnitude earthquakes (approximately moment magnitude 6.5 and above) occurring infrequently on long faults, evidence indicates that the assumptions of temporal and spatial independence are not valid. A new fault behavior model incorporating temporal and spatial dependence is needed to estimate site-specific hazard in areas subject to such earthquakes. This research develops an earthquake occurrence model that is a generalized semi-Markov process (GSMP) and allows for the simulation of the fault behavior through time. The size of each simulated earthquake is related to the amount of slip that is released. In order to apply the model to a fault, the following data must be known for each cell along the entire length of the fault: the slip rate, the mean and standard deviation of the earthquake interarrival times, and the time of the last earthquake. Additionally, the time of the last earthquake that ruptured the entire fault must be known. The model can then simulate the sizes and locations of earthquakes occurring along the fault for the time period of interest.

PB95-183968

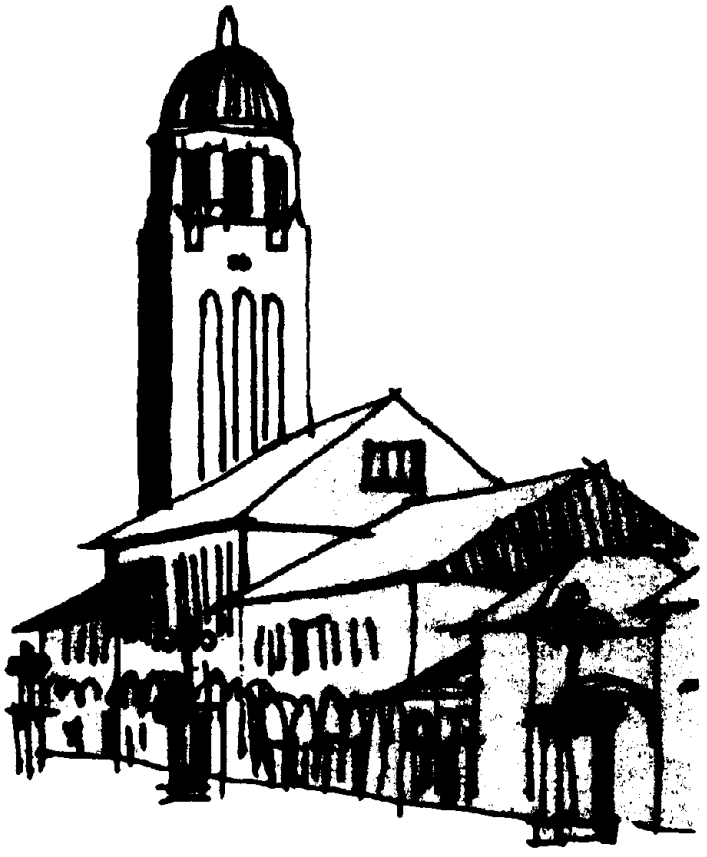
The John A. Blume Earthquake Engineering Center

Department of Civil Engineering
Stanford University

A GENERALIZED SEMI-MARKOV PROCESS FOR MODELING SPATIALLY AND TEMPORALLY DEPENDENT EARTHQUAKES

by

Kimberly A. Lutz
and
Anne S. Kiremidjian



This report was partially supported by the

National Science Foundation
Graduate Student Fellowship
and
Grant No. EID 9024032

Electric Power Research Institute
Contract No. RP2356-91
and
The John A. Blume
Earthquake Engineering Center

Report No. 104

July 1993

REPRODUCED BY:
U.S. Department of Commerce
National Technical Information Service
Springfield, Virginia 22161

Reproduced from best available copy

PB95-183968

A GENERALIZED SEMI-MARKOV PROCESS FOR MODELING SPATIALLY AND TEMPORALLY DEPENDENT EARTHQUAKES

by

Kimberly A. Lutz and Anne S. Kiremidjian

This report was partially supported by

The National Science Foundation Graduate Student Fellowship
and
Grant No. EID 9024032,

Electric Power Research Institute
Contract No. RP2356-91,

and

The John A. Blume
Earthquake Engineering Center

Report No. 104

July 1993

NTIS is authorized to reproduce and sell this report. Permission for further reproduction must be obtained from the copyright owner.

© Copyright by Kimberly Anne Lutz 1993

All Rights Reserved

ABSTRACT

PB95-183968

Site-specific hazard estimation requires the modeling of the occurrences of earthquakes on any faults with the potential to impact the site. Previous earthquake occurrence models have assumed either spatial independence or temporal independence or both. However, for large magnitude earthquakes (approximately moment magnitude 6.5 and above) occurring infrequently on long faults, evidence indicates that the assumptions of temporal and spatial independence are not valid. A new fault behavior model incorporating temporal and spatial dependence is needed to estimate site-specific hazard in areas subject to such earthquakes.

This research develops an earthquake occurrence model that is a generalized semi-Markov process (GSMP) and allows for the simulation of the fault behavior through time. The fault is discretized into short cells; the model traces through time the slip accumulated on each cell and the amount of slip release on each cell due to earthquake occurrences. The size of each simulated earthquake is related to the amount of slip that is released. In order to apply the model to a fault, the following data must be known for each cell along the entire length of the fault: the slip rate, the mean and standard deviation of the earthquake interarrival times, and the time of the last earthquake. Additionally, the time of the last earthquake that ruptured the entire fault must be known. The model can then simulate the sizes and locations of earthquakes occurring along the fault for the time period of interest.

Application of the model to the northern San Andreas fault (the portion of the fault that ruptured in 1906) implies that there are two distinct processes at work. The North Coast section generates large earthquakes (approximately moment magnitude 7.7 to 8.1), and the South Santa Cruz Mountains segment generates somewhat smaller earthquakes (approximately moment magnitude 6.8 to 7.4). The San Francisco Peninsula segment represents a transition between these two behaviors.

The model is relatively insensitive to the cell size chosen, to the distribution chosen to model the times between earthquakes triggering at a given place on the fault, and to the choice of a segmentation model that subdivided the San Francisco Peninsula segment. The moment magnitude of the largest earthquakes simulated are sensitive to the slip rate. The results for individual segments are highly sensitive to the mean interarrival times, but the aggregate results are much less sensitive.

This research develops an earthquake occurrence model that is appropriate for estimating hazard due to large, spatially and temporally dependent earthquakes. Because smaller magnitude earthquakes can also be important in seismic hazard analysis, however, this model must be combined with another designed to model lower magnitude seismicity (perhaps a Poisson model) in order to estimate the total site-specific hazard.

ACKNOWLEDGMENTS

Many individuals provided valuable input through the development of this research project. Gerry Shedler lent his expertise on the Generalized Semi-Markov Process, making a complex subject more easily grasped. Professors Haresh Shah, Allison Smith, and Ronnie Borja suggested various avenues that should be more completely explored. We are greatly indebted to them for their advice and recommendations.

Many thanks to the graduate students at the Blume Center who provided critical appraisals of the ideas developed in this research. They include Hjortur Thrainsson, Stephanie King, Sara Wadia, Mohsen Rahnama, and Renate Fruchter.

This report is based on the doctoral dissertation of Kimberly Lutz. The research work was supported partially by the National Science Foundation Graduate Student Fellowship and Grant No. EID 9024032. Additional financial support provided by the Electric Power Research Institute Contract No RP2356-91 is gratefully acknowledged. The report was published with funds provided by the John A. Blume Earthquake Engineering Center. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation or that of Electric Power Research Institute.

For the foolishness of God is wiser than man's wisdom, and the weakness of God is stronger than man's strength. *I Corinthians 1:25*

PB95-183968

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 OVERVIEW

The goal of seismic hazard analysis is to estimate the risk to a structure at a given site due to the occurrence of earthquakes. The steps involved in seismic hazard analysis include identifying earthquake sources, modeling the occurrences of earthquakes on these sources, determining the bedrock motion at the site due to an earthquake's occurrence, evaluating the soil's amplification of the motion at the site, and determining the structural response. This research is concerned with the second step outlined above; namely, the modeling of earthquake occurrences.

Specifically, this research models the occurrences of earthquakes that exhibit both temporal and spatial dependence. Such earthquakes typically have large seismic moments (corresponding to approximately moment magnitude 6.5 or greater) and occur infrequently on long faults. They usually have rupture zones on the order of tens or even hundreds of kilometers and are frequently associated with surface rupture. Though there is a pattern in time and in space to the rupture zones associated with these earthquakes, it is not always deterministic. Characteristic earthquakes, which repeatedly rupture the same section of fault with earthquakes of similar magnitudes, often display temporal and spatial dependence. Since temporally and spatially dependent earthquakes are a significant class of earthquakes capable of inflicting serious damage to structures, the ability to model them is an important tool for seismic hazard analysis. There is a need for a new earthquake occurrence model because previous models have assumed either time independence or space independence or both.

This research develops an earthquake occurrence model that includes both temporal and spatial dependence which can be applied to specific faults that generate earthquakes displaying such behavior. The types of information that can be obtained by using this model to simulate the occurrences of earthquakes on a fault can be divided into two main categories: short-term results and long-term results. Short-term results are those that estimate the probabilities of earthquakes occurring within a short time frame (such as the economic life of a structure) starting from the present time. Long-term results characterize the fault's behavior over many thousands of years in order to estimate, for example, the largest earthquake that the fault can generate and the pattern of

seismicity on the fault. Both types of results are important in assessing the earthquake hazard at a site.

1.2 SEISMIC HAZARD ANALYSIS

As previously noted, earthquake occurrence modeling is only one component of seismic hazard analysis. To clarify the role of earthquake occurrence models, the five basic steps in seismic hazard analysis are explained below and illustrated in Figure 1.1.

1.2.1 IDENTIFICATION OF SOURCES

The first step in seismic hazard analysis is to identify all sources of earthquakes that can affect the chosen site. Earthquake sources can be modeled in three ways. Point sources repeatedly generate earthquakes from exactly the same point. Since the epicenters of repeated earthquakes in a region usually display some scatter, such sources are rare. Earthquake faults are frequently modeled as two-dimensional line sources; they generate earthquakes whose epicenters lie in a narrow band along a line. In some areas, there is definite seismicity that is not attributable to a well-defined fault. Such seismicity can be modeled as arising from a three-dimensional area source, in which the earthquakes' epicenters are scattered within a well-defined area. Area sources can also model the scatter of earthquakes' epicenters occurring along a subduction zone.

1.2.2 EARTHQUAKE OCCURRENCE MODELING

The second step in seismic hazard analysis is to model the occurrence of earthquakes on each source. Many different earthquake occurrence models have been proposed; the following briefly discusses the major classes of models. A more complete review of earthquake occurrence models can be found in Anagnos and Kiremidjian (1988).

Poisson Models

Poisson models (e.g., Cornell, 1968; Der Kiureghian and Ang, 1977), which assume that earthquakes occur independently in time, space, and magnitude, form the simplest class of earthquake occurrence models. The rate of earthquake occurrences is uniform and estimated from the Gutenberg-Richter equation

$$\log N(m) = a - bm \tag{1.1}$$

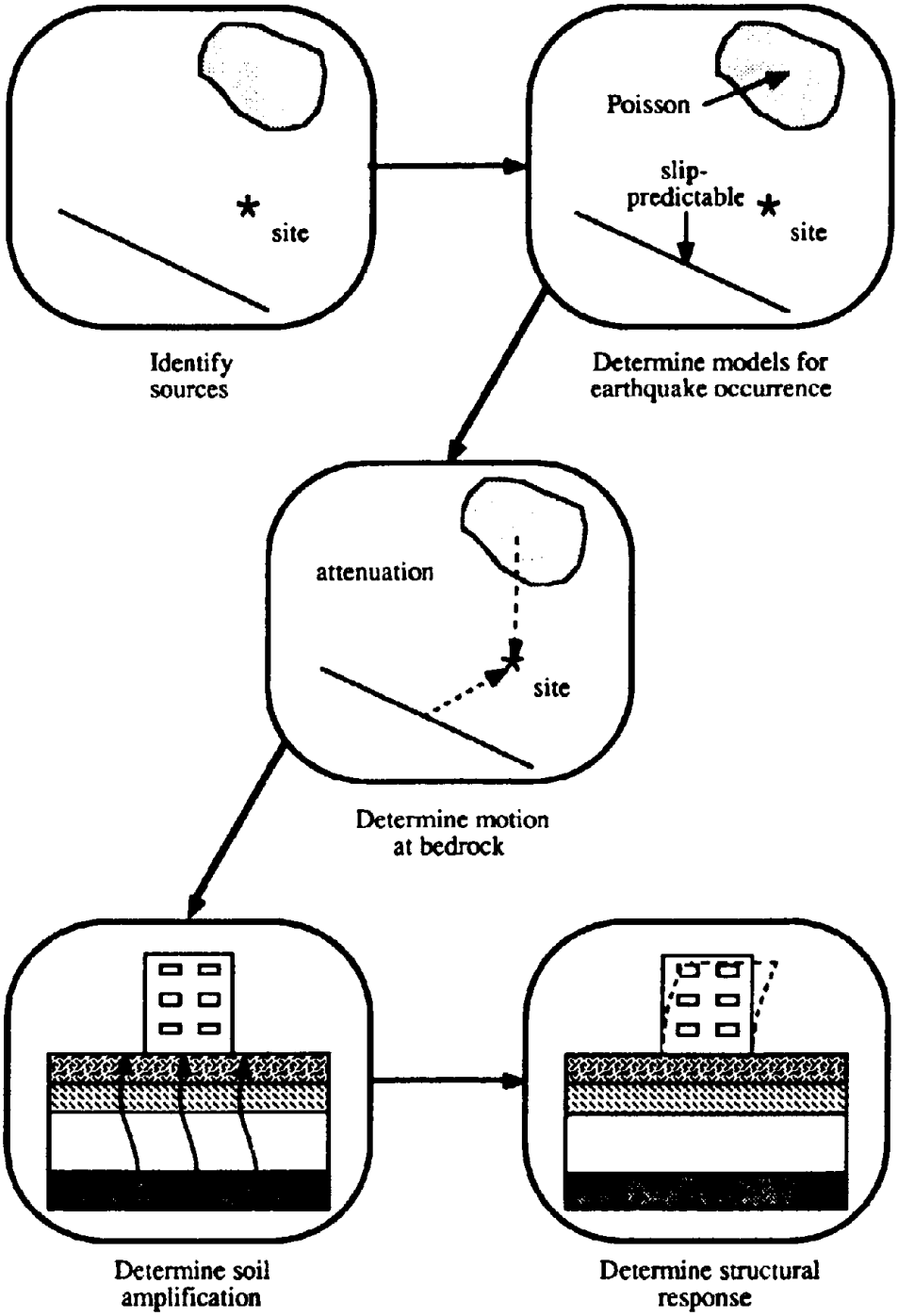


FIGURE 1.1 Steps in seismic hazard analysis

where

m = the magnitude of the earthquake

$N(m)$ = the number of events with a magnitude $\geq m$ during a given time period

a, b = constants.

Poisson models have a constant hazard rate, which is the probability of an earthquake in a time period $(t, t + \Delta t)$ given that there were no earthquakes in the time period $(0, t)$.

Poisson models are applicable to regions characterized by frequent, smaller magnitude earthquakes which display neither temporal nor spatial dependence. Gardner and Knopoff (1974) examined the earthquake record for Southern California and found that when aftershocks are removed, the sequence of earthquakes with magnitudes smaller than 5.8 is Poissonian. However, cyclic patterns, indicating dependence between earthquakes, have been noted by many authors (e.g., Bufe, et. al., 1977; Shimazaki and Nakata, 1980; Sykes and Quittmeyer, 1981). Regions displaying such patterns with large, rare earthquakes are not adequately described by Poisson models.

Time-Predictable Models

Time-predictability (Shimazaki and Nakata, 1980) is the correlation of the size (measured by seismic displacement) of an earthquake with the amount of time until the occurrence of the next earthquake. Models based upon time-predictability introduce time dependence into earthquake occurrence modeling (Anagnos and Kiremidjian, 1984).

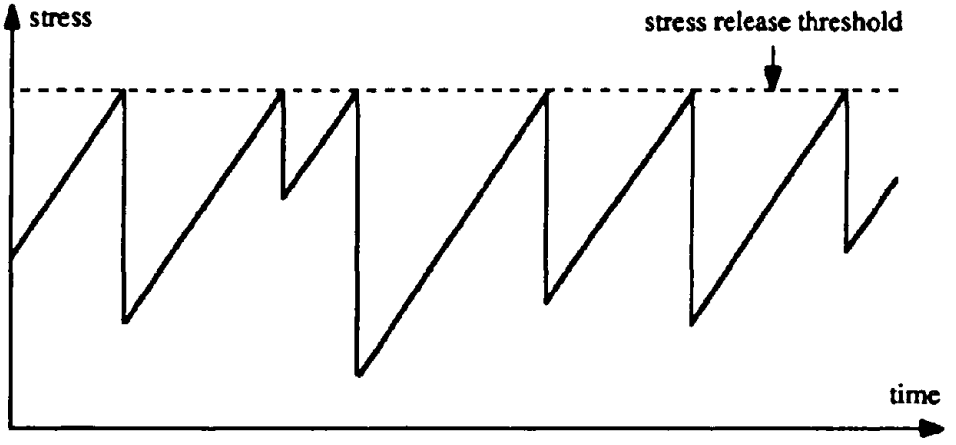


FIGURE 1.2 Time-predictable model

They describe a mechanism for earthquake occurrences involving a fixed stress threshold at which earthquakes occur. The rate at which stress accumulates is assumed to be constant. When an earthquake occurs, a random amount of stress is released. Since the stress release threshold is fixed, knowing the amount of stress released during an earthquake allows prediction of the time, though not the size, of the next earthquake. Figure 1.2 shows a sample time history of stress accumulation and release for the time-predictable model.

Several regions, primarily along plate boundaries, have been noted to display time-predictability (Bufe, et. al., 1977; Shimizaki and Nakata, 1980; Mogi, 1981; Sykes and Quittmeyer, 1981). It is important to realize, however, that time-predictable models describe earthquakes repeatedly rupturing the same section of fault but do not attempt to describe any spatial pattern in the earthquake sequence.

Slip-Predictable Models

Slip-predictability is the correlation of the time between earthquakes with the size of the earthquake at the end of the time interval (Shimazaki and Nakata, 1980). Models based upon slip-predictability also incorporate time dependence (Kiremidjian and Anagnos, 1984). As in the time-predictable model, the rate of stress accumulation is assumed to be constant. Earthquakes occur at random times, with all stress accumulated above a given threshold being released during each earthquake; the state of stress on the fault following each earthquake is the same. Given the time between earthquakes, the

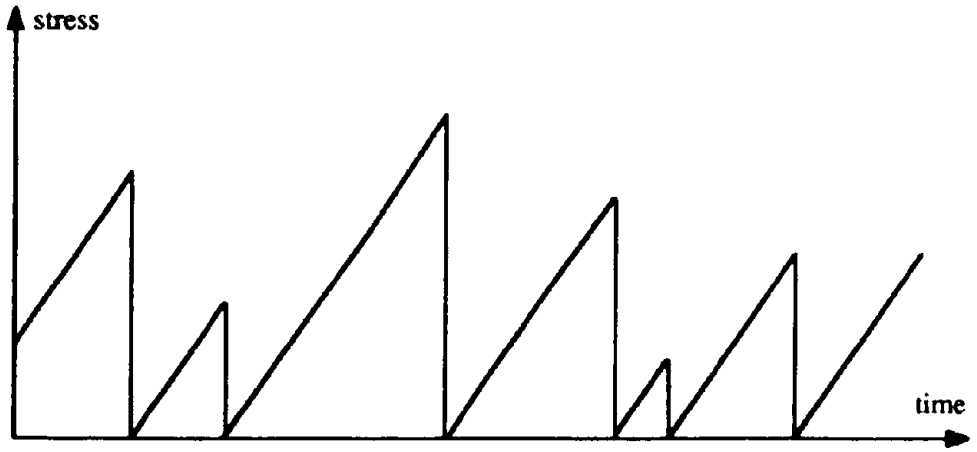


FIGURE 1.3 Slip-predictable model

amount of stress released, though not the interarrival time, can be predicted. Figure 1.3 illustrates a time history of stress accumulation and release for a slip-predictable model.

Earthquakes along the Middle America Trench, Mexico, display a slip-predictable pattern (Wang, et. al., 1982). Similar to time-predictable models, slip-predictable models address temporal dependence but not spatial dependence. Therefore, neither class of models completely describes the true behavior of large, rare earthquakes.

Random Slip Rate Models

Both the time-predictable and slip-predictable models assume a constant slip rate. However, a study of large earthquakes in Alaska (Li and Kisslinger, 1985) and experimental studies of rock fractures (Price, 1981; Ranalli, 1987) suggest that non-linear slip accumulation mechanisms may be important, particularly in long-term earthquake prediction models.

Random slip rate models include the effects of a non-uniform stress accumulation rate as well as inhomogeneous fault properties. Suzuki and Kiremidjian (1988) assume a random slip accumulation rate for each successive earthquake on a given section of fault. Though the value of the slip accumulation rate is random after each earthquake, it is constant between earthquakes. Figure 1.4 shows a sample time history of stress accumulation and release for the random slip rate model. As with time-predictable and slip-predictable models, random slip rate models incorporate temporal dependence but not spatial dependence.

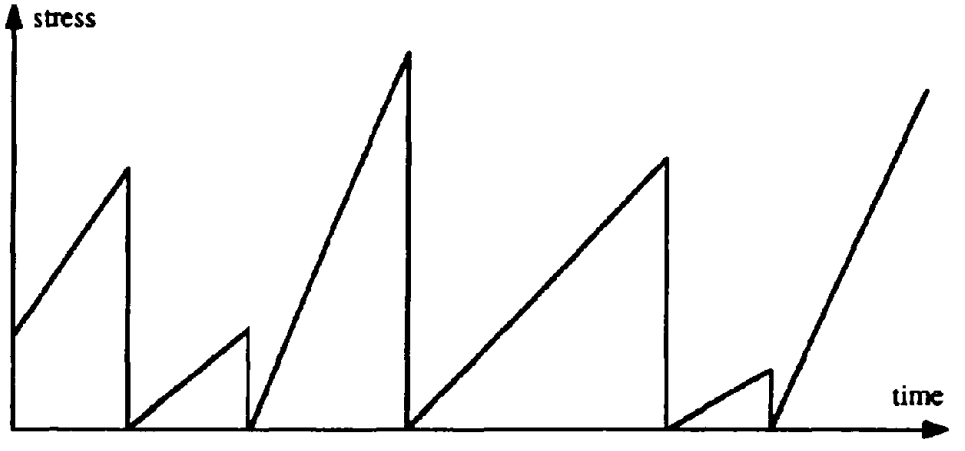


FIGURE 1.4 Random slip rate model

1.2.3 MOTION AT BEDROCK

Once a suitable earthquake occurrence model has been chosen, the next step is to determine the motion at bedrock resulting from a given earthquake. There are two main ways to do this: attenuation relationships (e.g., Joyner and Boore, 1988) and geophysical models (e.g., Suzuki and Kiremidjian, 1988). Attenuation functions are developed by fitting curves to available data. They have the limitations of being biased toward the data and of being applicable only to areas with similar geology. Attenuation relationships frequently predict ground motion parameters directly, rather than separating the prediction of bedrock motion from the soil amplification.

Geophysical models are based upon source mechanisms and wave-propagation theory. They have an advantage over attenuation relationships in that they are specific to the area under study. However, they require extensive data to describe the geology not only of the site, but also of the area between the site and the earthquake source. Data describing the source characteristics, such as the rupture area, seismic moment distribution, and stress drop, are also needed for these models. In addition, they are computationally more complex than attenuation functions, making them difficult to implement in a site hazard analysis.

1.2.4 SOIL AMPLIFICATION

Local soil conditions play a crucial role in determining the ground motions experienced at a site (e.g., Seed and Idriss, 1969; Seed, et. al., 1976; Idriss, 1991). During the Loma Prieta earthquake of 1989, waterfront areas of San Francisco Bay experienced significant damage due to amplification of ground motion and liquefaction of the soil (Borcherdt, 1990; USGS Circular 1045, 1989). In order to estimate the effects of soil amplification at a specific site, knowledge of the soil profile from bedrock to the surface is needed. Methods for predicting the surface motion from the bedrock motion that incorporate information on the soil properties in depth are presented in Kiremidjian, et. al., 1991.

1.2.5 STRUCTURAL RESPONSE

One purpose of seismic hazard analysis is to determine the response of the structure at the site. Many methods for evaluating structural response due to earthquakes are based on one parameter, such as the peak ground acceleration. Current methods, however, use response spectra to represent ground motion content at different frequencies

in order to more completely describe the structural response. Seismic hazard analysis allows determination of the structure's response, which is part of seismic risk evaluation.

1.3 IMPETUS FOR A NEW EARTHQUAKE OCCURRENCE MODEL

Mounting evidence suggests that many large earthquakes display not only temporal dependence, but spatial dependence as well. Since the previously mentioned models do not include both temporal and spatial dependence, a new model is needed.

1.3.1 TEMPORAL DEPENDENCE

Temporal independence is the property of the distribution of earthquake interarrival times being the same throughout time. Implicit in the concept of temporal independence is the idea that no matter what the earthquake history, the probability of an earthquake occurring in a small time increment is the same. Deduced from the elastic rebound theory is the idea that immediately following an earthquake, the probability of another earthquake rupturing the same area is small, and the probability of another earthquake increases as time passes. Thus, the elastic rebound theory supports the idea of temporal dependence because the distribution of earthquake interarrival times changes through time.

Figure 1.5 shows a space-time plot of seismicity of magnitude 5 or greater along the northern part of the San Andreas fault system. Following each of the large magnitude earthquakes in 1857 and in 1906, there is a period of low earthquake activity along the rupture zone, which is denoted by a vertical line. This pattern of activity, with few earthquakes following a major one, lends credence to the elastic rebound theory. A model that attempts to describe seismicity due to faults displaying behavior similar to that of the northern San Andreas fault must then incorporate temporal dependence.

1.3.2 SPATIAL DEPENDENCE

Spatial independence is the property of earthquake rupture zones being randomly distributed along the length of the fault. It has been recognized that long faults do not rupture completely during a single earthquake. This has given rise to the concept of fault segmentation, which divides long faults into segments, each of which is capable of rupturing independently. There are studies suggesting that physical controls in the fault zone define the ends of segments, and that these segments persist through many seismic

cycles (Schwartz, 1988). Large earthquakes on long faults exhibit spatial dependence because the rupture zones depend upon the physical controls governing the segmentation rather than being uniformly distributed over the length of the fault.

Figure 1.6 shows one proposed segmentation model of the San Andreas fault, suggesting that it generates spatially dependent earthquakes. Furthermore, Figure 1.5 also demonstrates that earthquakes along the northern San Andreas fault are spatially dependent, as the rupture zones are not evenly scattered along the fault's length. Figures 1.5 and 1.6 together demonstrate that earthquakes on the northern San Andreas fault are both temporally and spatially dependent. The presented model will not only describe the behavior of the northern San Andreas fault, but also the behavior of other faults whose earthquakes are temporally and spatially dependent.

1.4 SCOPE OF THIS RESEARCH

In this dissertation, a model is developed for temporally and spatially dependent earthquake occurrences that utilizes the generalized semi-Markov process. The model will be applied to the northern San Andreas fault to estimate probabilities of earthquake occurrences in specified time periods and to describe the long-term behavior of the fault. Sensitivity of the model to input data will be studied. Finally, extensions to this research and future areas of investigation will be suggested.

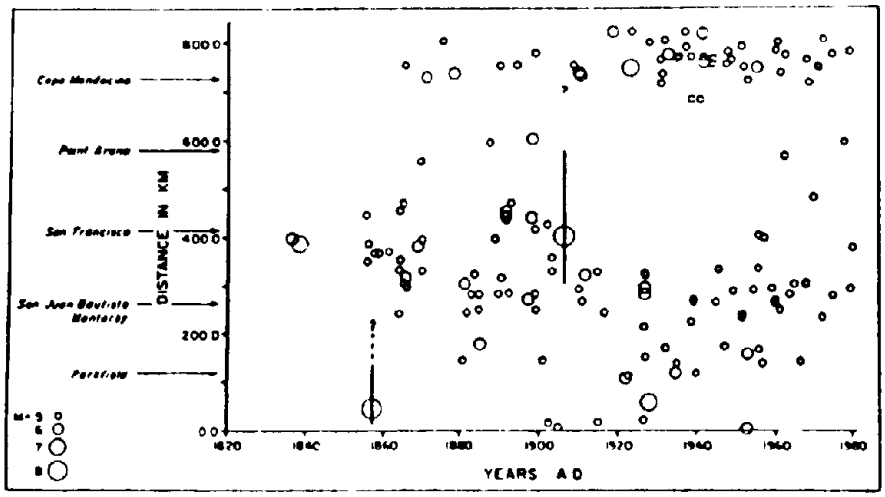


FIGURE 1.5 Space-time plot of $M \geq 5$ seismicity along the northern San Andreas system (from Ellsworth, et. al., 1981)

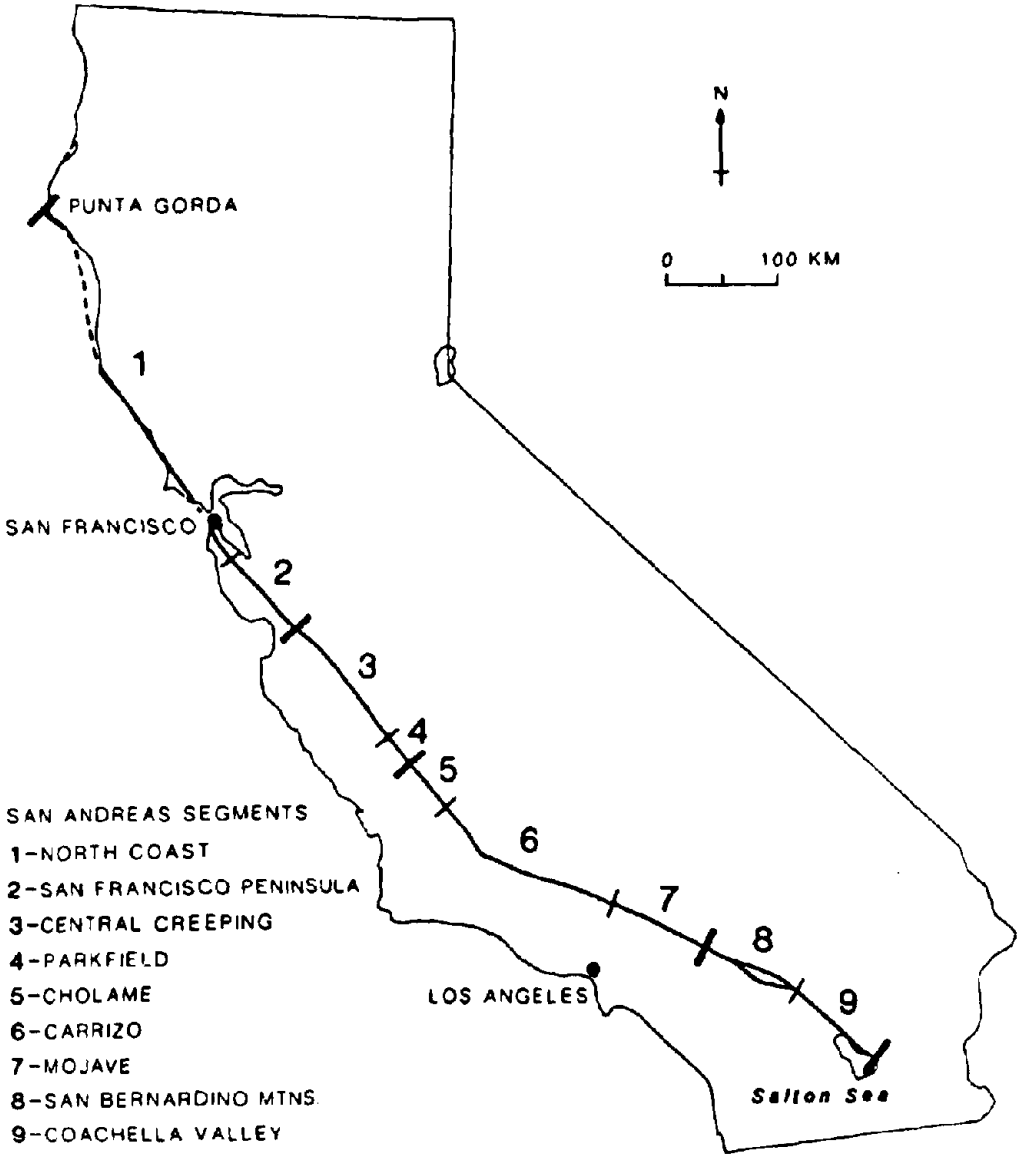


FIGURE 1.6 Segmentation model for the San Andreas fault (from Schwartz, 1988)

1.5 ORGANIZATION OF PRESENTATION

Chapter 2 discusses the generalized semi-Markov process (GSMP), which is the mathematical process that underlies the earthquake occurrence model. The model formulation is described in Chapter 3. The data required for the simulation of the model is presented in Chapter 4. The model's application to the northern San Andreas fault is the subject of Chapter 5. In Chapter 6, the sensitivity of the results is investigated. Finally, Chapter 7 summarizes the thesis and presents conclusions, as well as suggesting areas of future study.

CHAPTER 2
GENERALIZED SEMI-MARKOV PROCESSES

2.1 INTRODUCTION

Generalized semi-Markov processes (GSMPs) are a class of stochastic processes that are well-suited to a variety of problems. Like the semi-Markov process, the GSMP moves from state to state at random times. In the GSMP, there are several events that can occur in a given state. The first of those events to occur is called the trigger event, and it determines the time until the state transition and the probability distribution for the next state. The semi-Markov process is a special case of the GSMP in which there is only one event associated with each state and hence with each state transition.

The GSMP is well-studied within the field of Operations Research (e.g., Shedler, 1987; Whitt, 1980). Because of the complexity inherent in the GSMP, it can be used to model many different phenomena, including queues and machine availability. It can also be used to model fault behavior and provides a convenient framework for simulating the underlying stochastic process.

2.2 SPECIFICATION OF A GSMP

In order to uniquely describe a GSMP, six different components must be specified: the state space, the event set, the event set mapping, the event scheduling mechanism, the clock speeds, and the state transition mechanism. These components can be thought of as "building blocks" that are used to construct a GSMP. Each of these building blocks will be discussed within the context of modeling fault behavior in order to clarify them and describe how they fit together. The fault whose behavior is to be modeled will be discretized into several cells of uniform length, each of which accumulates strain due to the passage of time and periodically releases it in the form of an earthquake. Since strain accumulation and release are difficult to measure, the model is developed to track coseismic slip.

A GSMP is composed of one or more state variables, whose values characterize the process. The values of the state variables, called the state of the process, change with time, reflecting changes in the process itself. In modeling fault behavior, each cell has an associated state variable whose value characterizes the amount of slip accumulated at that point on the fault. A state variable of zero means that there is currently no slip

accumulated at that location; an earthquake has just occurred and ruptured that portion of the fault. A state variable of two means that, at the present time, there is enough slip accumulated at that point on the fault to rupture a length of two cells should an earthquake trigger on that cell. This is not a prediction of an earthquake, but merely a characterization of how much slip is accumulated. The building blocks of a GSMP tell what values the state variables can assume and detail the mechanisms that change the state of the process.

2.2.1 STATE SPACE

The set of all values that the state variables can assume forms the state space, S. Because the state space must have a finite or countable set of states, the state variables cannot take on a range of real numbers, but instead are restricted to discrete values. In the fault model, each state variable can assume any value in the set composed of zero and the integers up to and including the number of cells on the fault. The minimum amount of slip that can be released during an earthquake corresponds to zero, i.e., no rupture is possible. The maximum amount of slip that can be released corresponds to a rupture of the entire fault.

2.2.2 EVENT SET

All the events that can occur and cause a state transition form the event set, E. The event set must contain all events that can occur in any state of the process, though it is not necessary for all events to occur in every state. The event set must be finite. The event that occurs first in the current state, called the trigger event e^*, will cause the state transition. If more than one event occurs simultaneously, then the trigger event is a set called E^*.

In the fault behavior model, each cell has two events associated with it. The first is that the value of the variable representing the amount of accumulated slip on the cell is incremented. This event takes place when the amount of elapsed time allows the accumulation of enough slip to rupture a portion of the fault that is longer by the length of one cell. The second event occurs when the cell triggers an earthquake. The time of occurrence of an earthquake is governed by a probability distribution describing the time between earthquakes originating on the cell. Each time there is an earthquake, the value of the state variable for the triggering cell tells how many cells can rupture.

2.2.3 EVENT SET MAPPING

As mentioned above, not all events necessarily occur in every state of the process. The event set mapping, $E(s)$, describes which events can occur in each state. In the fault behavior model, the event that the amount of accumulated slip on a cell is incremented will be scheduled unless that cell can already rupture the entire fault (i.e., unless its state variable is already equal to the number of cells on the fault). In addition, each cell will have the event that it triggers an earthquake scheduled in all states.

2.2.4 EVENT SCHEDULING MECHANISM

In each state, the events are scheduled by determining the time until they occur. The event scheduling mechanism can depend on the event being scheduled, the trigger event, the old state, and the new state. Since the GSMP is a stochastic process, the events will not always be scheduled deterministically. In many cases, the time until the event occurs is simulated by choosing a random number from a probability distribution that describes the interarrival times for that event.

In the fault behavior model, the rate at which slip accumulates is assumed to be constant. The relationship between slip accumulated and length of fault ruptured is described by an empirical equation. Given the slip rate, the length of a cell, and the current amount of accumulated slip, the amount of time it takes for each cell to increment its accumulated slip is deterministic. Therefore, the events that a cell increments its amount of accumulated slip are scheduled deterministically.

The time until a given cell triggers an earthquake is random. Such an event is scheduled by simulating a random number from a probability distribution that describes the times between a given cell triggering an earthquake.

2.2.5 CLOCK SPEEDS

Associated with each scheduled event is a clock that tells the amount of time remaining until that event occurs. The clocks run at a finite, deterministic speed that must be specified and be greater than or equal to zero. A clock speed equal to unity means that the rate at which the clock runs is the same as the rate at which time passes. Thus, a clock showing 62 days on it would take exactly 62 days to run to zero. By convention, the clock speeds of events that are not scheduled are zero.

While it is possible for a GSMP to utilize clock speeds other than unity for scheduled events, this is not the usual case. In the fault behavior model, the clock speeds for all scheduled events are unity.

2.2.6 STATE TRANSITION MECHANISM

A state transition occurs when the clock of an event (or set of events) reaches zero. When this trigger event or trigger event set occurs, the GSMP moves into a new state. The state transition mechanism gives the probability function of the process moving into each possible new state, given the old state and the trigger event.

In the fault behavior model, assume that the trigger event is an event that the amount of accumulated slip on a cell is incremented. The new state would be the same as the old state except that the state variable corresponding to that cell would be increased by one. If the trigger event is that a cell causes an earthquake, then the state variable of that cell tells how many cells must rupture. Rules that will be detailed in the next chapter determine exactly which cells rupture and what the new values of their state variables will be. The new state is then the same as the old state except for the cells that rupture during the earthquake.

2.3 FORMULATION OF THE GSMP

A GSMP provides a convenient framework for simulating the underlying stochastic process. Since GSMPs are used to describe complex phenomena, there is usually no closed form equations for the probabilities associated with the various states of the process. For example, the probability that a certain cell ruptures during a specified time period could be a quantity of interest in the fault behavior model. These probabilities can be estimated by simulation.

Figure 2.1 shows how a GSMP can be simulated. To begin the simulation, all the state variables must be set to some beginning configuration. In the fault behavior model, this beginning configuration could be all zeroes, meaning that there has just been an earthquake that ruptured the entire length of the fault. It is also possible for the beginning configuration to represent some state of accumulated slip, such as the current state.

The event set mapping tells which events can occur in the current state of the process. After determining these events, they are scheduled using the event scheduling mechanism. The event with the shortest time scheduled on its clock is the trigger event.

(This assumes that all the clocks run at the same rate. If this is not the case, then the trigger event is determined by dividing the time on each clock by its speed. The event with the smallest ratio will be the trigger event.) When some of the events are scheduled deterministically, more than one event can occur simultaneously. The trigger event can then be a set of all events occurring at the shortest time scheduled on any of the clocks.

The time of the simulation is advanced to the time at which the trigger event (or trigger event set) occurs, and the new state is determined from the old state and the trigger state using the state transition mechanism. If the simulation is to continue, the events to be scheduled in the new state are determined from the event set mapping, and the cycle begins again.

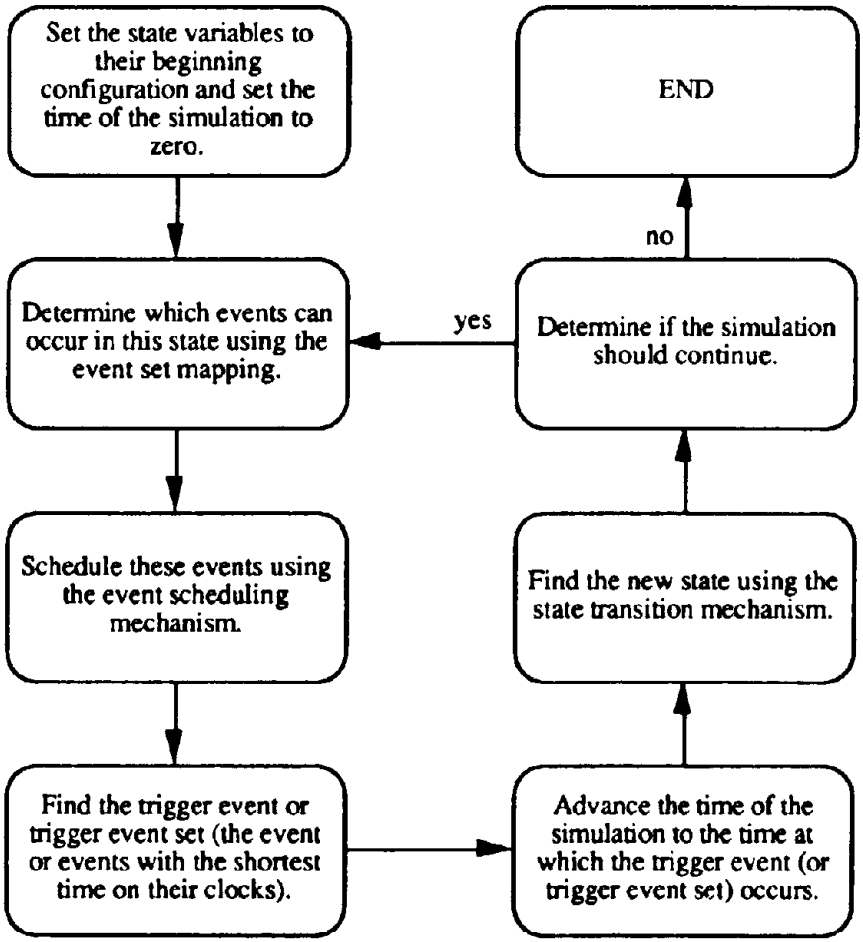


FIGURE 2.1 Using the GSMP framework to simulate the underlying process

The probabilities of interest can be estimated through this simulation procedure. For example, consider the probability that a given cell on the fault ruptures during a given time period. The behavior of the fault will be simulated for the given time period for many cycles. This is an example of the independent, identically distributed (i.i.d.) sampling problem. For each cycle, a random variable will either have a value of 0 (the given cell did not rupture during the cycle) or 1 (the cell did rupture). These random variables are i.i.d. if the same initial conditions are used each time. The probability of the given cell rupturing during any given cycle is estimated as the ratio of the number of times that the cell did rupture during a cycle divided by the total number of cycles simulated. In this manner, the simulation of the GSMP can yield estimates of desired probabilities. Simulation will be the tool used to estimate probabilities that describe the behavior of faults generating temporally and spatially dependent earthquakes.

CHAPTER 3
THE FAULT BEHAVIOR MODEL

3.1 INTRODUCTION

The GSMP framework is useful for modeling spatially and temporally dependent earthquakes. The random time between state transitions provides the temporal dependence needed for the model. In order to introduce spatial dependence into the model, a complex state space is utilized. The basic physical quantity that the fault behavior model tracks is accumulated slip. Since it can be calculated from the slip rate and elapsed time, the amount of accumulated slip is a convenient way to describe the state of the fault. In addition, the slip released during an earthquake can be directly related to the moment release and thus the moment magnitude of the earthquake. The fault is discretized into short cells, and the amount of slip accumulated on each of these cells is represented by one state variable. The array of state variables represents the configuration of accumulated slip on the fault.

As time progresses, the model keeps track of the slip accumulated and released through earthquakes. There are two different processes at work--constant slip accumulation and occasional slip release. In addition, the slip can be released (i. e., an earthquake can nucleate) at any point along the length of the fault. The GSMP framework allows for different events to "compete" to trigger the next state transition and thus can be used to describe the behavior of the fault.

3.2 SPECIFICATION OF THE GSMP UNDERLYING THE FAULT BEHAVIOR MODEL

The GSMP framework will be used to describe the stochastic process underlying the fault behavior model. Define $\{X(t), t \geq 0\}$ to be a stochastic process where

$$X(t) = \{(A_1(t), B_1(t)), (A_2(t), B_2(t)), \ldots, (A_N(t), B_N(t))\} \tag{3.1}$$

Each cell j has two state variables associated with it, $A_j(t)$ and $B_j(t)$. At time t, $A_j(t) = k$ if cell j is capable of rupturing k cells. This means that at time t, there is enough slip accumulated on cell j to cause a rupture with a length of k cells. Since j is an index on the cell number, it can assume values from one to N, where N is the maximum number of cells. Since k refers to the number of cells that cell j can rupture,

it can assume values from zero to N, where zero corresponds to no rupture possible and N corresponds to rupturing the entire fault. The relationship between accumulated slip and rupture length is determined from an empirical equation and will be discussed in Chapter 4.

3.2.1 STATE SPACE

All the values that the process $X(t)$ can assume form its state space S.

$$S = \left\{ \left((a_1, b_1), (a_2, b_2), \ldots, (a_N, b_N) \right) \in \left(\{0, 1, \ldots, N\} \times \{0, 1\} \right)^N \right\} \tag{3.2}$$

Equation 3.2 states that each of the a_j variables may take on values in the set $\{0, 1, \ldots, N\}$, and each of the b_j variables may take on values in the set $\{0, 1\}$. (The lower case a_j and b_j are used, rather than the upper case $A_j(t)$ and $B_j(t)$ used above, to denote specific values of the state variables, rather than their values through time.) There are no restrictions on the permissible values of the state space within the $2N$-dimensional space defined above.

3.2.2 EVENT SET

In order to determine how and when the process $X(t)$ moves between states, it is necessary to enumerate the event set E, which contains all the different events that can occur in the process. The event set E for the fault behavior process is

$$E = \left\{ e_{11}, \ldots, e_{N1}, e_{12}, \ldots, e_{N2}, e_{13}, \ldots, e_{N3} \right\} \tag{3.3}$$

For each cell j, the event e_{j1} is the event that cell j increments by one the number of cells that it is capable of rupturing. This occurs when, due to the passage of time, cell j has accumulated sufficient slip to rupture a length of fault that is longer by the length of one cell.

For each cell j, the events e_{j2} and e_{j3} are events that cell j triggers an earthquake. Duplicate events are used (rather than only one event) to circumvent the fact that the GSMP framework requires that the only events that can be rescheduled after a state transition are the trigger events. In the fault behavior model, the cell that triggers the earthquake ruptures. In addition, neighboring cells may rupture because the triggering cell had a value of a_j greater than one, meaning that a length of fault longer than one cell ruptures during the earthquake. The events that each of the ruptured cells

triggers an earthquake now need to be rescheduled. However, for the cells that did not trigger the earthquake but did rupture, these events are old events and cannot be rescheduled under the GSMP framework. To get around this problem, event e_{j3} is scheduled whenever it is necessary to reschedule event e_{j2} and vice-versa. The $B_j(t)$ variables in the state transition are used to tell whether event e_{j2} or e_{j3} is scheduled at time t. When $b_j = 0$, event e_{j2} is scheduled; when $b_j = 1$, event e_{j3} is scheduled.

3.2.3 EVENT SET MAPPING

When the process $X(t)$ is in a given state s (that is, of necessity, a subset of the entire state space), the events that can occur and trigger a transition into the next state must be determined. This set of events $E(s)$ is a subset of the entire event set E defined above. The event set mapping tells how to determine which events are in $E(s)$ for any state s.

$$\begin{aligned} &\text{For } s = \{(a_1, b_1), (a_2, b_2), \ldots, (a_N, b_N)\} \in S: \\ &\qquad e_{j1} \in E(s) \text{ if and only if } a_j \neq N \\ &\qquad e_{j2} \in E(s) \text{ if and only if } b_j = 0 \\ &\qquad e_{j3} \in E(s) \text{ if and only if } b_j = 1 \end{aligned} \tag{3.4}$$

Equation 3.4 states that for each cell j, the event e_{j1} (that cell j increments the amount of slip accumulated on it) is scheduled unless there is already enough accumulated slip on cell j to rupture the entire fault. As mentioned above, event e_{j2} (that cell j triggers an earthquake) is scheduled when $b_j = 0$, and the duplicate event e_{j3} is scheduled when $b_j = 1$.

3.2.4 EVENT SCHEDULING MECHANISM

After determining which events are to be scheduled in the current state, the clocks of each of the events must be set. For event e_{j1} (incrementing the amount of slip accumulated on cell j), the clock is set deterministically according to an equation of the form:

$$\log(d) = a + b\log(l) \tag{3.5}$$

where

d = the average surface displacement in meters

l = the surface rupture length in kilometers

a, b = constants.

For example, suppose that the amount of time that slip must accumulate to cause a rupture length of one cell during an earthquake is desired. The average displacement during the earthquake is calculated from the regression equation using a rupture length of one cell. Dividing the average displacement by the slip rate yields the desired quantity. In order to determine the additional amount of time it takes to accumulate enough slip to rupture n cells when currently $n-1$ cells can rupture, the above calculation is done for rupture lengths of n cells and $n-1$ cells. The time increment is the desired quantity.

The above procedure makes several assumptions. The first is that the average surface displacement will be the same as the slip released on the cell on which the earthquake originates. The second is that the surface rupture length is the same as the length of cells that rupture. While these assumptions will not always be true in every earthquake, they allow the rupture length to be calculated deterministically from the time elapsed since the last earthquake. It would be possible to specify probability distributions relating surface rupture length and true rupture length and relating average slip displacement and slip accumulated at the epicenter of the earthquake. The rupture length would then be determined probabilistically from the elapsed time since the last earthquake. While this could be more representative of the actual situation, the simpler deterministic method is used in this research.

To set the clock corresponding to event e_{j2} (the event that cell j triggers an earthquake), a random number is simulated from D_j, the distribution of times between earthquakes triggering on cell j. This distribution differs from the distribution of interarrival times for cell j rupturing during an earthquake because cell j can break due to rupture "spilling over" from an adjacent cell. As will be shown in Chapter 4, the mean and standard deviation of D_j are determined by trial and error based upon the estimated mean and standard deviation of the earthquake interarrival times.

Weibull distributions (e.g., Kiremidjian and Anagnos, 1984; Anagnos and Kiremidjian, 1984) and lognormal distributions (e.g., USGS Working Group, 1990) have been used to represent interarrival time distributions. There are no data describing the distributions of triggering earthquakes, D_j. For this reason, it is suggested that D_j be modeled by either the Weibull distribution or the lognormal distribution, in an analogous manner with the interarrival time distributions.

The clock corresponding to event e_{j3} is set using the same procedure outlined above for event e_{j2}. When cell j breaks and $b_j = 0$, event e_{j3} is scheduled by simulating its time from D_j. If $b_j = 1$, the event e_{j2} is scheduled in the same way. Events e_{j2} and e_{j3} are alternately scheduled each time cell j ruptures, regardless of whether cell j triggered the earthquake.

3.2.5 CLOCK SPEEDS

In the GSMP describing the fault behavior process, the clock speeds for all events are unity.

3.2.6 STATE TRANSITION MECHANISM

The final component that must be specified to completely describe the fault behavior process is the state transition mechanism, which is summarized by a flowchart in Figure 3.1. The state transition mechanism tells how the new state s' of the process $X(t)$ is determined when the old state s and the trigger event set E^* causing the state transition are known. In Figure 3.1, events of the form e_{j1} are referred to as type 1 events (that is, events in which the amount of slip accumulated on a cell is incremented). Events of the forms e_{j2} and e_{j3} are referred to as type 2 events (that is, events that a cell triggers an earthquake).

Figure 3.1 refers to the "rules for rupture" in determining which cells rupture during an earthquake. When an earthquake triggers on cell j at time t, the value of $A_j(t)$ is the number of cells that rupture during that earthquake. The following rules for determining which cells rupture were developed based upon observations from past earthquake ruptures. It is assumed that rupture begins on the cell that originates the earthquake and is confined to one segment whenever possible. The rupture is assumed continuous, and it does not jump over any cells. It is further assumed that cells with more accumulated slip will break before cells with less accumulated slip. The following cases demonstrate the rules for determining which cells rupture. Figures 3.2-3.5 illustrate these cases with a fault composed of five cells divided into two segments.

Case #1: Triggering cell j has $a_j = 1$ at time of earthquake

In this case, only one cell, the triggering cell, can rupture during the earthquake. The position of the cell within the segment and the amount of slip accumulated on surrounding cells do not affect the rupturing cell.

Figure 3.2(a) shows the state of accumulated slip on the fault directly before an earthquake triggering on cell 4. The only cell that ruptures is cell 4, as seen in Figure 3.2(b). Because all slip accumulated on cell 4 is released during the earthquake, $a_4 = 0$ in the new state.

Case #2: Triggering cell j has $a_j = m$, the number of cells on its segment

In this case, the earthquake triggers on cell j, which is capable of rupturing m cells, where m is the number of cells on the segment containing cell j. The rationale for

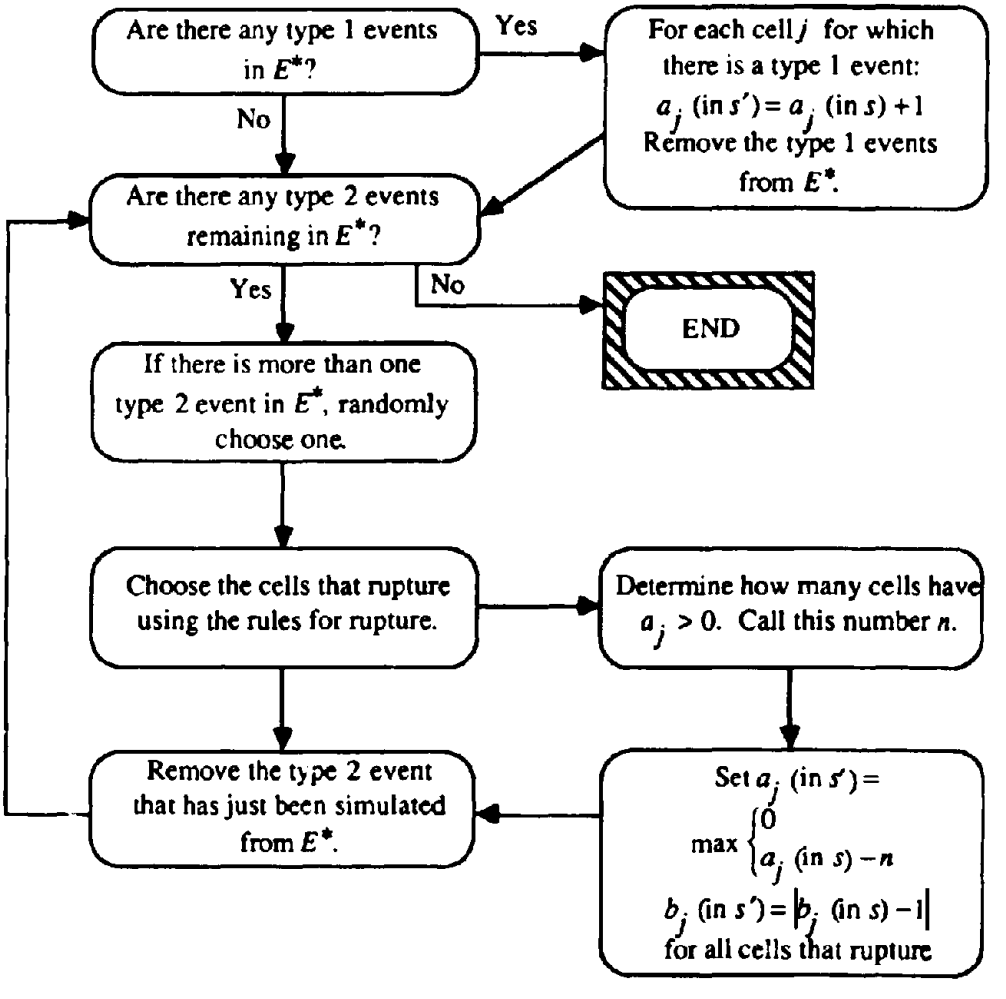


FIGURE 3.1 State transition mechanism

(a) State of slip before earthquake (b) State of slip after earthquake

FIGURE 3.2 Illustration of case #1 demonstrating rules for rupture

segmenting a fault is that the ends of the segments are delineated by some physical change in the fault, such as a change in strike, and that each segment is capable of rupturing independently. Therefore, whenever possible, the model limits rupture to a single segment.

Figure 3.3 illustrates this case. Figure 3.3(a) shows the state of accumulated slip directly before the earthquake. Since $a_3 = 3$, three cells must rupture during the earthquake. Because the segment containing cell 3 has exactly three cells on it, those three cells rupture. Figure 3.3(b) shows that all slip is released on cells 1, 2, and 3 during the earthquake. Therefore, in the new state, $a_1 = 0$ and $a_2 = 0$ and $a_3 = 0$.

Case #3: Triggering cell j has $a_j = m$, fewer than the number of cells on its segment

The earthquake triggers on cell j, which is capable of rupturing m cells, where m is a number that is smaller than the number of cells on the segment containing cell j. In this case, the choice of cells to rupture is governed by their relative accumulated slip. Whenever possible, the cell with more accumulated slip will rupture.

Figure 3.4(a) shows the state of accumulated slip directly before the earthquake. In this case, $a_2 = 2$, meaning that two cells must rupture during the earthquake. Because the segment contains three cells, a choice must be made between rupturing cell 1 or cell 3. Since $a_1 > a_3$, cells 1 and 2 rupture during the earthquake. Figure 3.4(b) shows the new state. If $a_1 = a_3$, there is no reason to favor rupturing one cell over the other as they

(a) State of slip before earthquake (b) State of slip after earthquake

FIGURE 3.3 Illustration of case #2 demonstrating rules for rupture

are on the same segment and have the same amount of accumulated slip. In that case, one of the cells is randomly chosen to be the one that ruptures along with the trigger cell.

These three cases can be used to determine the rupture whenever an earthquake occurs. If the state of the triggering cell is greater than the number of cells on the segment, then the segment containing the triggering cell ruptures first. Then the next cell to rupture is chosen using case #3. In other words, the cells at the ends of the rupture are examined and the one with the greater accumulated slip ruptures. Then the entire segment containing the chosen cell must rupture before rupture can continue to another segment. In this way, the rupture propagates from segment to segment until the number of cells indicated by the state of the trigger cell have ruptured.

Because all the cases discussed above involve the total release of all accumulated slip on the rupture cells, the new state for all ruptured cells is zero. This is not always the case, however, as illustrated in Figure 3.5. In this case, cell 2 triggers the earthquake. Since $a_2 = 5$, the entire fault must rupture. At the time of the earthquake, $a_4 = 0$, indicating that recently an earthquake has released all the slip on cell 4. By the time the current earthquake occurs, there is not even enough slip accumulated on cell 4 to rupture a length of one cell. Therefore, cell 4 releases very little slip even though it ruptures during the earthquake. For this reason, only four of the cells release slip. The new state of the process is determined by counting the number of rupturing cells with a state greater than zero. This number is subtracted from the old state of each of the rupturing cells, and the larger of this number and zero is the new state of the process. In Figure 3.5(b), cells 2 and 3 retain some accumulated slip after the earthquake while the other cells release all their accumulated slip.

The reasoning behind this procedure is that there is enough energy at the triggering cell to rupture five cells. Since one of the cells has very little accumulated slip, the rupture actually releases slip on only four cells. There must then be some accumulated slip left on the fault after the rupture. Cells with a large amount of accumulated slip will retain some of it after the earthquake.

(a) State of slip before earthquake (b) State of slip after earthquake

FIGURE 3.4 Illustration of case #3 demonstrating rules for rupture

As mentioned above, these "rules for rupture" are based upon intuition about how fault rupture can propagate. One assumption made by the rules presented here that may be debatable is that a cell in the rupture zone with no accumulated slip will not terminate the earthquake's rupture. As was illustrated in Figure 3.5, the number of cells that rupture is determined solely by the state of the variable describing the amount of slip accumulated on the trigger cell at the time rupture begins. The cells that break are chosen according to the scheme outlined above. The fact that a cell in this rupture zone has a state of zero when the earthquake occurs means only that some of the cells will retain some accumulated slip after the earthquake.

The implication of this assumption in the model is that large earthquakes rupturing the same section of fault tend to cluster together in time. One large earthquake occurs but does not rupture the entire fault. Another earthquake then triggers shortly afterward on a part of the fault that did not rupture during the first earthquake. According to the assumptions made, the section of fault that broke in the first earthquake can break again.

It is possible, however, that the presence of a cell with no accumulated slip will serve to terminate the rupture of an earthquake. In that case, the rupture zone would be smaller, and more cells would retain accumulated slip after the earthquake. The implication of this assumption is that the fault will tend to generate a greater number of earthquakes that break the entire fault and fewer smaller magnitude earthquakes. Suppose an earthquake triggers on a cell with enough accumulated slip to rupture the entire fault, and that the cells adjacent to the trigger cell recently experienced an earthquake and have no slip accumulated. Following the given assumption, the cells with no accumulated slip would be sufficient to terminate the rupture even though a large amount of slip is accumulated on the trigger cell.

If this is indeed the case, then the rules of rupture should be changed. The logic of the model remains the same; the only change is in the method for deciding which cells

(a) State of slip before earthquake (b) State of slip after earthquake

FIGURE 3.5 Illustration of a non-zero new state

break and how much slip is accumulated after the earthquake occurs. In other words, the state transition mechanism of the GSMP would be modified.

The state transition mechanism describes how the state of the process changes when slip is accumulated and released. Slip accumulation causes the state variables to be incremented by one. Slip release, in the form of an earthquake, causes certain cells to rupture and release some or all of their accumulated slip. The "rules for rupture" determine which cells rupture, and the state of accumulated slip on the rupturing cells determines the new state of the process. The state transition mechanism is the last component that must be specified in order to completely describe the GSMP underlying the fault behavior model.

CHAPTER 4
MODEL SIMULATION

4.1 INTRODUCTION

The purpose of developing fault behavior models is to characterize the occurrences of earthquakes along actual faults, allowing the estimation of seismic hazard at a particular site. In order to do this, data describing the physical fault are used to determine the parameters that define the fault behavior model. In theory, closed form solutions could be developed for any desired quantity, such as the probability of an earthquake of a given magnitude or greater originating at a certain place on the fault during a given period of time. However, the complexity of the model presented in Chapter 3 makes closed form solutions either non-existent or very difficult to obtain.

For this reason, simulation will be used to apply this model to actual faults. In simulation, the initial conditions (the state of slip accumulated when the simulation begins) and the time horizon are selected. Many repetitions of the model are then simulated for the time horizon of interest and the average results (or the distribution of results) reported. The advantages to this scheme for analyzing fault behavior include the ability to apply a complex model not amenable to closed form solutions and the observation of earthquakes' occurrences through time.

4.2 DATA REQUIRED FOR MODEL APPLICATION

In order to apply this model, four different classes of information are needed. These are: data describing the fault's configuration, data describing the slip accumulation and release mechanisms, the earthquake history of the fault, and equations relating slip release with other quantities of interest.

The fault whose behavior is to be modeled must generate earthquakes displaying temporal and spatial dependence and be capable of completely rupturing during one earthquake. This requirement allows for the simulation of the fault's behavior from a "time zero," that is, from the time of the last earthquake that ruptured the entire length of the fault and released all accumulated slip. This, in turn, allows the current state of accumulated slip to be estimated (based on the earthquake history) so that the probabilities of earthquakes within time frames beginning at the present can be assessed.

If a fault is not capable of completely rupturing in one earthquake but does display spatial and temporal dependence, it is possible that the model may apply to a portion of the fault. For example, the entire San Andreas fault does not rupture in one earthquake. However, the northern portion of the San Andreas ruptured completely during the 1906 earthquake and generates earthquakes displaying temporal and spatial dependence, as was discussed in Chapter 1. In Chapter 5, the model is applied to this section of the San Andreas fault.

4.2.1 DATA DESCRIBING THE FAULT'S CONFIGURATION

The data describing the fault's configuration include the length of the chosen fault, its depth, and its segmentation. Each segment is assumed to have homogeneous properties; that is, each cell on a segment has the same slip rate, the same interarrival time mean and standard deviation, and the same distribution of times between triggering earthquakes. The lengths and locations of any segments on the fault must be known. If the entire fault has homogeneous properties, no segments need be specified.

In applying the model, the fault's length is discretized into short units called cells. Each segment (and the fault as a whole) is made up of an integer number of cells. Theoretically, the smallest magnitude earthquake that the model generates corresponds to a rupture length of one cell. In reality, however, a model with very short cells does not generate earthquakes with much smaller minimum magnitudes than a model with very long cells provided that the earthquakes' interarrival time statistics remain constant, as will be discussed more completely in Chapter 6. Because the amount of time required to complete each simulation increases with the number of cells, the choice of cell length must balance this consideration with the need to accurately represent the length of each segment with an integer number of cells.

Figure 4.1 shows the segmentation suggested by the USGS Working Group (1990) for the portion of the San Andreas fault to which the model is applied. This segmentation divides the northern San Andreas into three segments: the North Coast segment, the San Francisco Peninsula segment, and the South Santa Cruz Mountains segment. The 340 km North Coast is by far the longest segment, stretching from offshore in the northwest to San Andreas Lake in the southeast, and it is judged by USGS to be capable of generating earthquakes with a maximum magnitude of 8.

The San Francisco Peninsula segment begins at San Andreas Lake and runs 61 km to the southeast; it has a maximum magnitude of 7. This segment can be considered to have two subsegments: the Mid-Peninsula subsegment and the North Santa Cruz

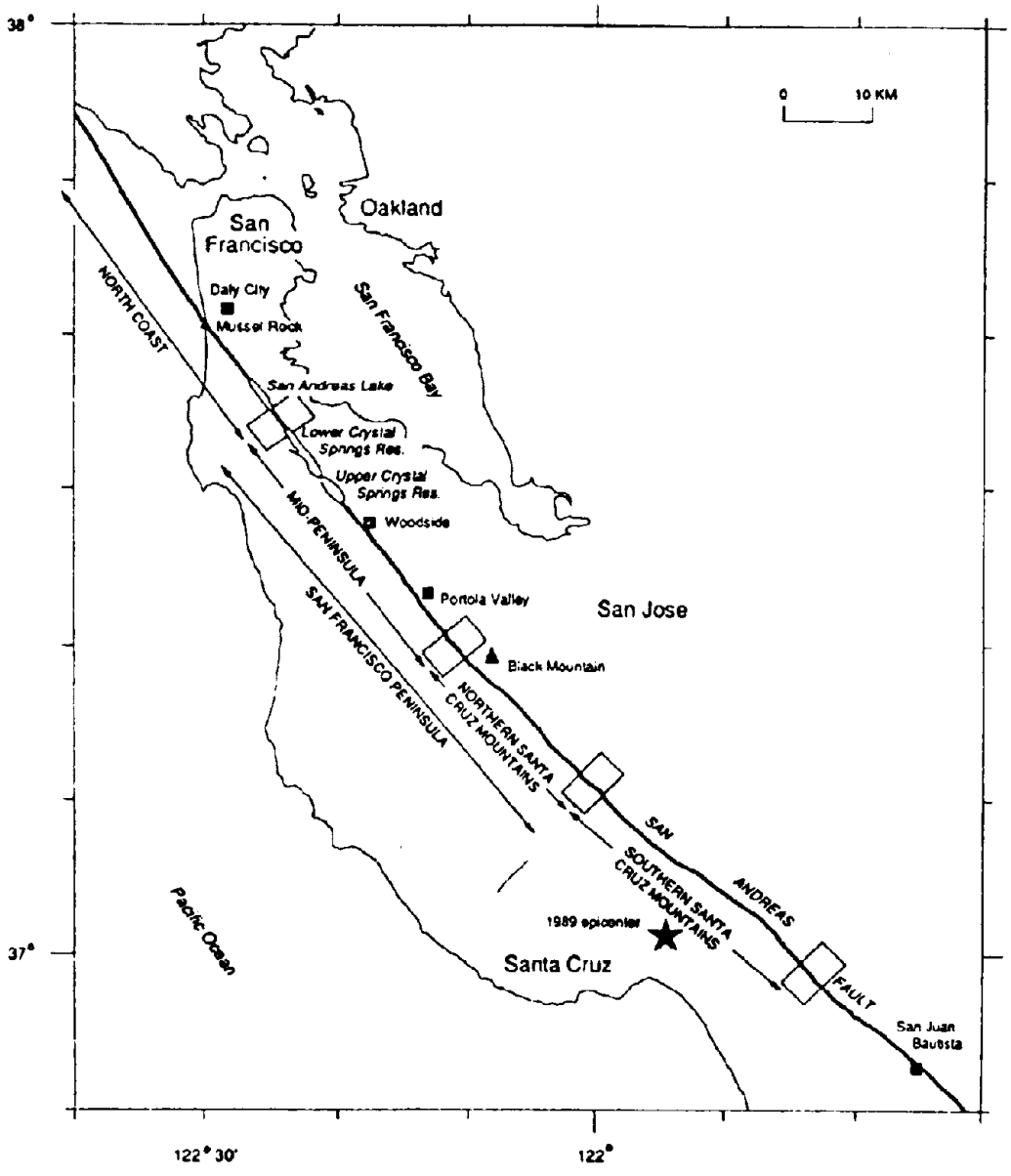


FIGURE 4.1 Northern San Andreas fault segmentation

Mountains subsegment. USGS considers two separate scenarios; one in which the Peninsula segment is undivided and one in which it is divided into the two subsegments. The results presented in Chapter 5 consider the segmentation in which the San Francisco Peninsula segment is undivided. As part of the sensitivity analysis presented in Chapter 6, the alternate segmentation is considered.

The South Santa Cruz Mountains segment is 39 km long and lies at the southeast end of the northern San Andreas fault. The 1989 Loma Prieta earthquake originated on this segment and generated a magnitude 7 earthquake.

The choice of cell length determines how quickly the simulation can be carried out and how accurately the lengths of the segments can be represented. In this research, results using cell lengths of 10 km, 20 km, and 30 km will be obtained and compared. The fault depth is taken as 20 km.

4.2.2 DATA FOR THE SLIP ACCUMULATION AND RELEASE MECHANISMS

The slip rates, the interarrival time statistics, and the form of the distribution of the times between triggering earthquakes are the data that describe the slip accumulation and release mechanisms. The slip rate is a measure of how quickly the sides of the fault are moving in relation to each other and therefore of how quickly stress is accumulating in the fault zone. In the same way that segmentation determination is inexact, slip rate estimates differ from researcher to researcher. The fault behavior model requires that each cell within a given segment have the same slip rate, though the slip rate may vary from segment to segment. This research assumes that the slip rates remain constant throughout time, though this may not be the case.

The mean and standard deviation are the interarrival time statistics and describe how often earthquakes occur. They are based on scarce historical data and sometimes also on geologic occurrence data, which can be obtained by trenching to date prehistoric earthquakes. Another method for estimating recurrence time is to divide the slip released coseismically by the slip rate. All these methods introduce uncertainty because of lack of data. Despite the problems in determining slip rate and interarrival time statistics, these data have been estimated for many faults capable of producing large earthquakes displaying spatial and temporal dependence.

Even if there were no uncertainty in determining the interarrival time statistics of earthquakes, the interarrival times would not be deterministic. The form of the

distribution of the times between triggering earthquakes is another piece of data required. As mentioned in Chapter 3, lognormal and Weibull distributions are possible choices. In the results presented in Chapter 5, it is assumed that the times between triggering earthquakes have a lognormal distribution. A comparison with results obtained from using the Weibull distribution will be made in Chapter 6. The slip rate estimated by the USGS Working Group (1990) for the entire northern portion of the San Andreas fault is 19 $\pm$ 4 mm/year. In Chapter 6, results will be computed and compared for slip rates of 15, 19, and 23 mm/year.

The interarrival times are estimated by USGS to be 237 $\pm$ 73 years for the North Coast segment, 138 $\pm$ 40 years for the San Francisco Peninsula segment, and 84 $\pm$ 24 years for the South Santa Cruz Mountains segment. However, the fault behavior model requires as input not the interarrival time statistics, but instead the statistics of the times between earthquakes triggering on each cell of the fault. These statistics are referred to as the input mean and input standard deviation, while the interarrival time statistics are referred to as the output mean and output standard deviation.

Tables 4.1, 4.2, and 4.3 show the input statistics and the output statistics for each scenario considered. The third column, titled "Other Info.," tells when the scenario considers interarrival time means other than those estimated by USGS or distributions other than the lognormal for the trigger times. "Seg 1 -" refers to segment 1 having a mean interarrival time equal to the mean value listed above minus one standard deviation. The coefficient of variation is the same for the new mean as it was for the unchanged mean. Similarly, "Seg 1 +" refers to the mean listed above plus one standard deviation, with an unchanged coefficient of variation. "Weibull" refers to using the Weibull distribution to model the trigger times.

Since there has been no attempt to directly estimate the input statistics from the fault itself, it is necessary to determine them by trial and error to make the output statistics close to the estimated interarrival time statistics. As there is no requirement that all cells on a segment break during an earthquake, each cell can have unique output statistics. The value reported in the tables for the output mean for a given segment is the average of the output means for all the cells on that segment. Thus, some parts of the segment will rupture more often than the output mean would suggest while others will break less often. The reported output standard deviation for a given segment is the average of the output standard deviations of all the cells. It is thus a measure of the average amount of variability on the segment, but it is not a true standard deviation.

TABLE 4.1 Input and output statistics for segment 1 (North Coast)

Cell Size	Slip Rate	Other Info.	Input Mean	Input St. Dev.	Output Mean	Output St. Dev.	USGS Mean	USGS St. Dev.
km	mm/yr		yr	yr	yr	yr	yr	yr
10	15		420	50	236	103	237	73
10	19		430	50	238	125	237	73
10	23		416	80	238	77	237	73
20	15		430	75	234	101	237	73
20	19		443	75	235	119	237	73
20	19	Seg 1 -	300	25	164	95	164	51
20	19	Seg 1 +	445	50	308	105	310	96
20	19	Seg 2 -	417	75	238	81	237	73
20	19	Seg 2 +	425	75	236	112	237	73
20	19	Seg 3 -	435	75	236	82	237	73
20	19	Seg 3 +	420	75	237	105	237	73
20	19	Weibull	465	75	238	116	237	73
20	23		395	80	238	78	237	73
30	15		430	75	238	109	237	73
30	19		418	75	237	115	237	73
30	23		400	90	237	83	237	73

TABLE 4.2 Input and output statistics for segment 2 (San Francisco Peninsula)

Cell Size	Slip Rate	Other Info.	Input Mean	Input St. Dev.	Output Mean	Output St. Dev.	USGS Mean	USGS St. Dev.
km	mm/yr		yr	yr	yr	yr	yr	yr
10	15		240	50	136	54	138	40
10	19		380	50	140	80	138	40
10	23		2000	0	108	56	138	40
20	15		220	50	138	56	138	40
20	19		350	50	139	77	138	40
20	19	Seg 1 -	285	50	139	64	138	40
20	19	Seg 1 +	395	50	139	80	138	40
20	19	Seg 2 -	185	50	96	48	98	28
20	19	Seg 2 +	700	30	176	103	178	52
20	19	Seg 3 -	250	50	139	70	138	40
20	19	Seg 3 +	2000	0	127	58	138	40
20	19	Seg 3 +	2000	0	127	58	138	40
20	23	Weibull	350	50	137	73	138	40
30	15		240	75	139	66	138	40
30	19		600	100	133	73	138	40
30	23		2000	0	96	43	138	40

TABLE 4.3 Input and output statistics for segment #3 (South Santa Cruz Mountains Segment)

Cell Size	Slip Rate	Other Info.	Input Mean	Input St. Dev.	Output Mean	Input St. Dev.	USGS Mean	USGS St. Dev.
km	mm/yr		yr	yr	yr	yr	yr	yr
10	15		125	45	83	23	84	24
10	19		123	40	83	23	84	24
10	23		127	45	83	23	84	24
20	15		100	30	83	22	84	24
20	19		105	30	84	24	84	24
20	19	Seg 1 -	105	30	84	24	84	24
20	19	Seg 1 +	105	30	86	23	84	24
20	19	Seg 2 -	103	30	85	22	84	24
20	19	Seg 2 +	103	30	83	24	84	24
20	19	Seg 3 -	76	30	60	21	60	17
20	19	Seg 3 +	132	35	108	30	108	31
20	19	Weibull	100	20	84	23	108	31
20	23		104	30	84	24	84	24
30	15		86	20	84	21	84	24
30	19		88	20	85	24	84	24
30	23		87	22	84	25	84	24

4.2.3 EARTHQUAKE HISTORY OF THE FAULT

One of the requirements for applying the fault behavior model is that the fault be capable of completely rupturing during one earthquake. This allows the simulation to begin at a specified time after the last earthquake that ruptured the entire fault. When estimating the earthquake hazard within a relatively short period of time compared to the interarrival times of the earthquakes, it is important to take into account where in the earthquake cycle the specified time period lies. For example, the hazard immediately following a large earthquake is relatively small and increases as time passes. Knowing the earthquake history of the fault allows an estimate of hazard beginning at a specified time.

In addition to estimating hazard within a given time frame, it is also possible to use the fault behavior model to determine the long term behavior of the fault. For example, by simulating the model for a long period of time, the number of earthquakes per year as a function of magnitude can be estimated. This gives information about the relative numbers of earthquakes of different magnitudes that can be expected and also about the maximum magnitude earthquake that can be expected on the fault. For this

type of information, the earthquake history of the fault from the last earthquake rupturing the entire fault is not necessary.

For the northern San Andreas fault, the last earthquake rupturing the entire fault occurred in 1906. The only sizable earthquake since then is the Loma Prieta earthquake of 1989, which ruptured segment #3 of the fault. These data will be used in the results presented in the next chapter.

4.2.4 RELATIONSHIPS BETWEEN RELEASED SLIP AND OTHER QUANTITIES OF INTEREST

The basic quantity tracked in this model is accumulated slip. Other quantities of interest are calculated from it. The relationship between slip released and rupture length will be used to determine how long it takes to increment the amount of slip accumulated on each cell. When an earthquake occurs, the amount of slip released will be related to the amount of moment released in order to determine the moment of the earthquake. The moment magnitude is then calculated from the moment.

As noted in Chapter 3, an empirical equation is used to relate slip released and rupture length. The specific equation used in this research is from Wells and Coppersmith (1991):

$$\log(d) = -1.33 + 0.84 * \log(l) \tag{4.1}$$

where

d = the average surface displacement in m

l = the surface rupture length in km.

Table 4.4 shows the application of this equation in determining the transition times for slip accumulation events. $T(j,k)$ is the amount of time it takes for cell j to accumulate enough slip to rupture k cells when currently there is enough slip accumulated to rupture $k-1$ cells. As the model simulation proceeds, each cell accumulates slip. At any given time, it is possible to calculate the length of fault that would rupture should an earthquake occur. Since only an integer number of cells can break, the length of fault that ruptures must be rounded to the nearest integer number of cells. Thus, $T(j,1)$ corresponds to the amount of time it takes to accumulate enough slip to break a length of fault equal to 0.5 of a cell. If there is enough slip accumulated to break a rupture length of 0.5 of a cell, rounding to the nearest integer number of cells would yield a rupture length of one cell.

$T(j,2)$ is the amount of time it takes to accumulate enough slip to break a length of 1.5 cells minus $T(j,1)$, etc.

The slip released coseismically can be related to the seismic moment of the earthquake using the equation

$$M_0 = \mu \bar{u} A = \mu \bar{u} L D \tag{4.2}$$

where

M_0 = the seismic moment of the earthquake in dyne-cm
μ = the modulus of rigidity for rock (3×10^{11} dyne / cm^2)
$\bar{u}$ = the average slip released coseismically in cm
L = the length of rupture in km
D = the depth of the fault in km.

TABLE 4.4 Transition times for slip accumulation events (20 km cells)

# of Cells	Rup. Len. (km)	Disp. (m)	Incre-mental Disp. (m)	Slip Rate (mm/yr)	k	$T(j,k)$ (yr)	Cum. Time (yr)
0.5	10	0.32	0.32	19	1	17.0	17.0
1.5	30	0.81	0.49	19	2	25.8	42.9
2.5	50	1.25	0.44	19	3	23.0	65.8
3.5	70	1.66	0.41	19	4	21.5	87.3
4.5	90	2.05	0.39	19	5	20.5	107.8
5.5	110	2.43	0.38	19	6	19.8	127.6
6.5	130	2.79	0.37	19	7	19.2	146.9
7.5	150	3.15	0.36	19	8	18.8	165.6
8.5	170	3.50	0.35	19	9	18.4	184.0
9.5	190	3.84	0.34	19	10	18.0	202.0
10.5	210	4.18	0.34	19	11	17.7	219.7
11.5	230	4.51	0.33	19	12	17.5	237.2
12.5	250	4.83	0.33	19	13	17.2	254.4
13.5	270	5.16	0.32	19	14	17.0	271.4
14.5	290	5.48	0.32	19	15	16.8	288.2
15.5	310	5.79	0.32	19	16	16.6	304.8
16.5	330	6.10	0.31	19	17	16.4	321.2
17.5	350	6.41	0.31	19	18	16.3	337.5
18.5	370	6.72	0.31	19	19	16.1	353.6
19.5	390	7.02	0.30	19	20	16.0	369.6
20.5	410	7.32	0.30	19	21	15.9	385.5
21.5	430	7.62	0.30	19	22	15.7	401.2

When an earthquake breaks more than one cell, the amount of moment released by each cell is calculated separately, as the cells can release different amounts of slip. The seismic moment of the earthquake is the sum of the moment released by each cell rupturing during the earthquake.

The seismic moment calculated from equation 4.2 is related to the moment magnitude using the relationship proposed by Hanks and Kanamori, 1979:

$$M_w = \frac{2}{3}\log(M_0) - 10.7 \tag{4.3}$$

where

M_w = the moment magnitude
M_0 = the seismic of the earthquake in dyne-cm.

Equations 4.2 and 4.3 allow the estimation of the seismic moment and moment magnitude of earthquakes whose occurrence is simulated with the model.

4.3 IMPLEMENTATION OF THE MODEL

In order to implement the fault rupture model, a simulation procedure is developed. Figure 4.2 outlines the procedure for simulating the sequences of earthquakes in time and for computing various quantities of interest. Each of the steps in this figure is explained in the following discussion.

The data required for the simulation of the fault rupture model were described in Section 4.2. These data included the fault length, the fault depth, the segmentation model, the slip rate for each segment, the interarrival time statistics for each segment, the distribution type of the times between each cell triggering an earthquake, and the earthquake history of the fault.

In addition to the data, several functional relationships need to be defined. These include the relationships between accumulated slip and rupture length, between slip released and seismic moment, and between seismic moment and moment magnitude. The specific relationships were given in equations 4.1, 4.2, and 4.3.

The cell size must be chosen such that an integer number of cells will adequately characterize the length of each segment and yet the computational effort will not be prohibitive due to the large number of cells. The input mean and standard deviation for each segment are the input parameters to the model. They are determined by trial and error, as noted above.

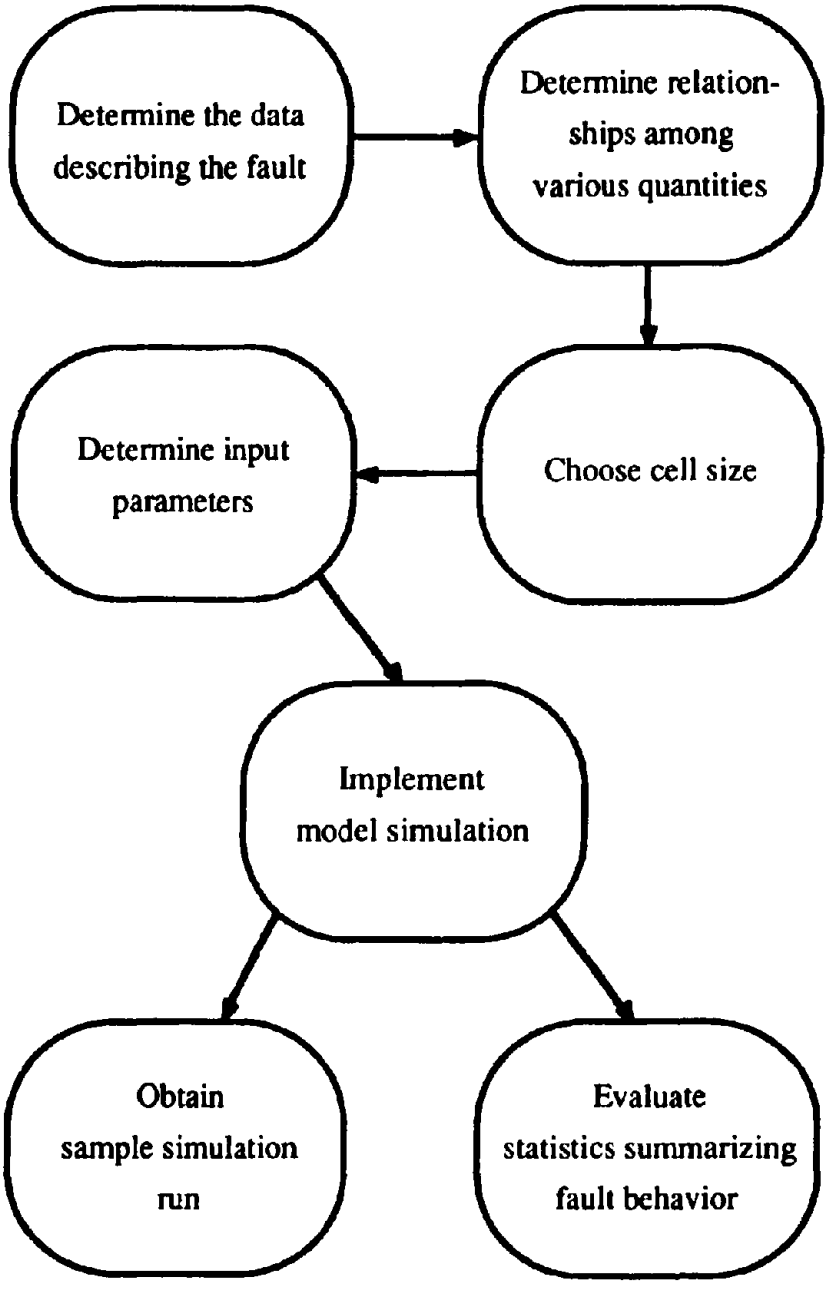


FIGURE 4.2 Steps in the implementation of model simulation

Once these data have been estimated, sequences of earthquakes can be simulated and their properties and statistics studied. Appendix A contains a chart describing how the various portions of the computer program used to simulate this model relate to each other.

Several different types of output are of interest when studying the fault behavior. By simulating a sample of earthquake sequences in time and in space, the times at which each state transition (incrementing the amount of accumulated slip or triggering an earthquake) occurs and the state of the variables after each transition can be traced. The number of times each cell ruptures and the output mean and standard deviation for each cell provide additional information. The average and standard deviation of the output means and of the output standard deviations of the cells on each segment are reported in order to compare their values to the input parameters. Other quantities calculated include:

- the probability of any part of the fault rupturing during an earthquake
- the probability of any part of a given segment rupturing during an earthquake
- the probability of a given cell rupturing during an earthquake
- the probability of an earthquake triggering on a given segment
- the probability of an earthquake triggering on a given cell
- the probability of more than 50% of a given segment rupturing during one earthquake

In addition to the above, other quantities reported include different distributions for each cell, for each segment, and for the entire fault. The probabilities of observing different interarrival times are obtained by multiplying the distribution of interarrival times by the probability of occurrence of an earthquake. The distribution of rupture lengths multiplied by the probability of occurrence of an earthquake gives the probability of observing different rupture lengths. The moment release distribution multiplied by the probability of occurrence gives analogous information. Finally, the number of earthquakes per year of a given magnitude or greater shows not only the relative numbers of earthquakes of different magnitudes but also the expected cutoff magnitude.

In Chapter 5, the results from the application of the model to the northern San Andreas fault are presented. Some of the output information described above is reported in order to analyze the behavior of the fault.

CHAPTER 5
RESULTS OF MODEL APPLICATION

5.1 INTRODUCTION

This chapter presents the results obtained from applying the fault behavior model to the northern San Andreas fault. The data used in applying this model were discussed in Chapter 4. Results are presented for the case of 20 km cells, a slip rate of 19 mm/yr, and interarrival time means as estimated by the USGS Working Group (1990). The lognormal distribution is assumed in modeling the earthquakes' trigger times. The sensitivity of the results to the cell size, to the slip rate, and to the interarrival times will be the subject of Chapter 6. In addition, an alternate segmentation and the use of the Weibull distribution for modeling the earthquakes' trigger times will be considered in Chapter 6.

The distinction was drawn in Chapter 4 between input statistics and output statistics. The input statistics are adjusted so that the output means averaged over each segment are close to the estimated mean interarrival times. Figure 5.1 shows the input and output means for each cell for the base case. Note that the variability in the output means over the cells in segment #1 (North Coast) and in segment #3 (South Santa Cruz Mountains) is relatively small. However, segment #2 (San Francisco Peninsula) shows significant variability among the cells. This suggests that segment #2 is a transition segment in which the character of the fault changes from the longer interarrival times to the north to the shorter interarrival times to the south.

Figure 5.2 shows the input and output standard deviations for each cell. Adjusting the input parameters allows the average of the output means over each segment to correspond closely to the USGS estimated mean interarrival times. However, the output standard deviations cannot be adjusted in this way. Because there are several cells on each segment, there is only a loose correlation between the input standard deviation and the output standard deviation. For example, the North Coast segment contains 17 cells. Each of these cells generates earthquakes according to a probability distribution characterized by the input mean and input standard deviation. When the simulation begins, the time until each cell triggers an earthquake is simulated from this distribution. The time until the next earthquake triggers on the segment is the smallest of these trigger times. In effect, there are 17 samplings (because there are 17 cells) from the distribution,

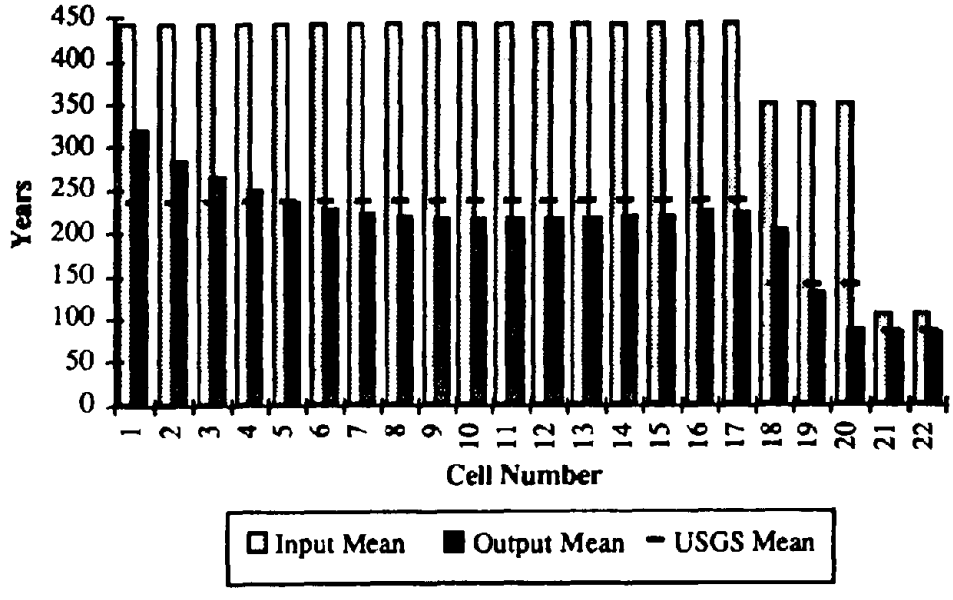


FIGURE 5.1 Input and output means for the base case

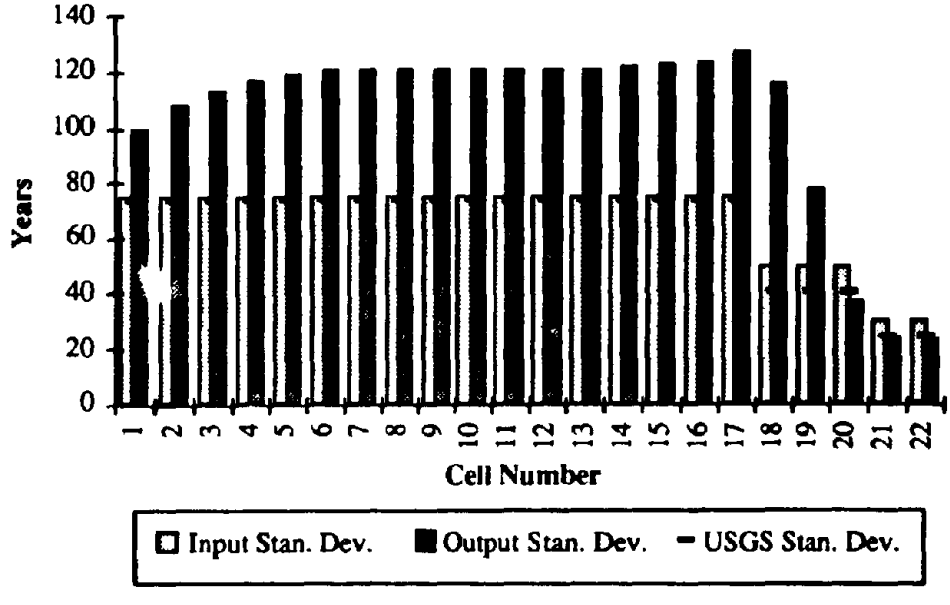


FIGURE 5.2 Input and output standard deviations for the base case

and the smallest number controls the time until the next earthquake. Raising or lowering the mean of this distribution will then affect the time until the next earthquake. Given enough simulations, the time until the next earthquake will be longer when the mean of the distribution is higher.

When the input standard deviation is changed, however, the major effect is on the output mean, not the output standard deviation. This is because a greater variability in the distribution, characterized by the larger standard deviation, increases the chances that one of the 17 samplings will be low and thus decreases the time until the next earthquake. Given enough simulations, increasing the standard deviation lowers the time until the next earthquake.

These observations imply that the output standard deviation is largely intrinsic to the model and has more to do with rupture spilling over from adjacent cells than with the input standard deviation. Note that the cells on segment #3 are an exception because the mean interarrival time of earthquakes on segment #3 is short compared to the mean interarrival times of segments #1 and #2. Most of the earthquakes in the simulation therefore initiate on segment #3, making the input parameters more closely associated with the output parameters than on the other segments.

5.2 SAMPLE SIMULATION

One of the strengths of simulation is that it allows analysis of the statistics of many simulations and also of the sample path of individual simulations. This can provide clues to the behavior of the fault that are sometimes obscured by looking only at summary statistics. In Figure 5.3, a typical simulation of the base case, the time and moment magnitude of each earthquake are listed on the left. The heavy lines denote the cells that ruptured during each earthquake; the asterisk above each heavy line shows the cell on which the earthquake originated. Figure 5.3 shows the earthquake history from time 0 years, at which there is no slip accumulated anywhere on the fault, to time 3309 years, at which there is an earthquake that ruptures the entire fault.

Most of the earthquakes displayed in Figure 5.3 (67%) originated on segment #3. 23% of the earthquakes originated on segment #1, with the remaining 10% originating on segment #2. The earthquakes originating on segment #3 typically have a moment magnitude between 7.0 and 7.4 (though larger or smaller magnitude earthquakes do occur). Larger magnitude earthquakes originate on segments #1 and #2 at longer intervals, and their typical magnitude range is 7.8 to 8.1. Some clustering of these larger

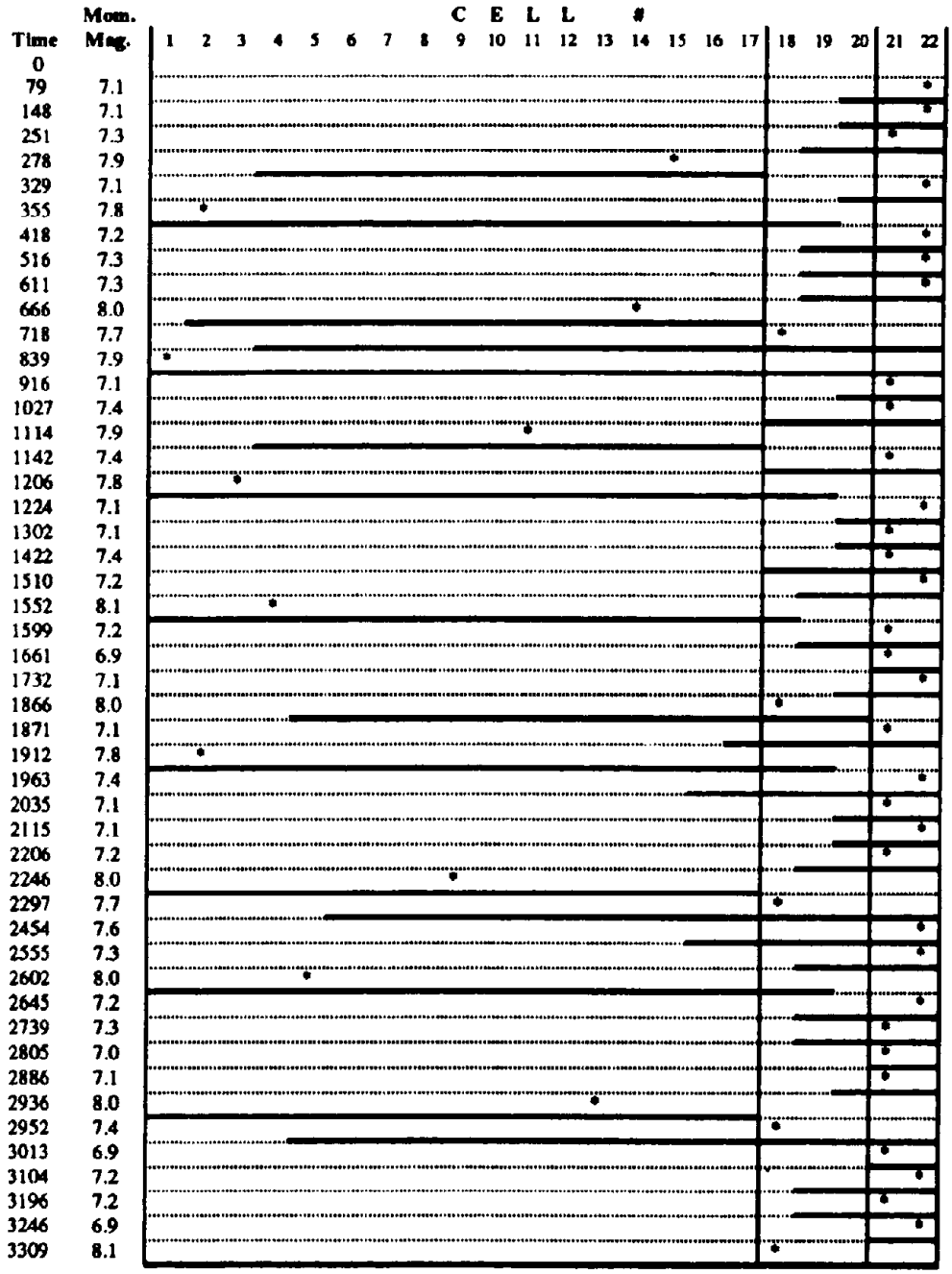


FIGURE 5.3 Sample simulation for 3309 years

earthquakes occurs; if a large earthquake does not rupture all of segment #1, there is frequently another earthquake shortly thereafter. Most of the earthquakes rupturing segment #2 spill over from another segment.

These observations suggest that earthquakes similar to the Loma Prieta earthquake occur rather regularly and more frequently than larger earthquakes. While earthquakes similar to the 1906 San Francisco earthquake are less common, they do occur and rupture a longer portion of the fault. The San Francisco Peninsula segment does not often originate an earthquake, but does frequently rupture when other earthquakes spill over onto it.

5.3 FREQUENCY OF EARTHQUAKES

Figure 5.4 shows the number of earthquakes per year based upon a simulation of 1,000,000 years. The x-axis is the moment magnitude of the earthquakes, while the y-axis is the number of earthquakes per year with a given moment magnitude or greater. Most of the smaller magnitude earthquakes originate on segment #3, while the larger magnitude earthquakes originate on segments #1 and #2. This confirms the behavior observed in the sample simulation in Section 5.2.1.

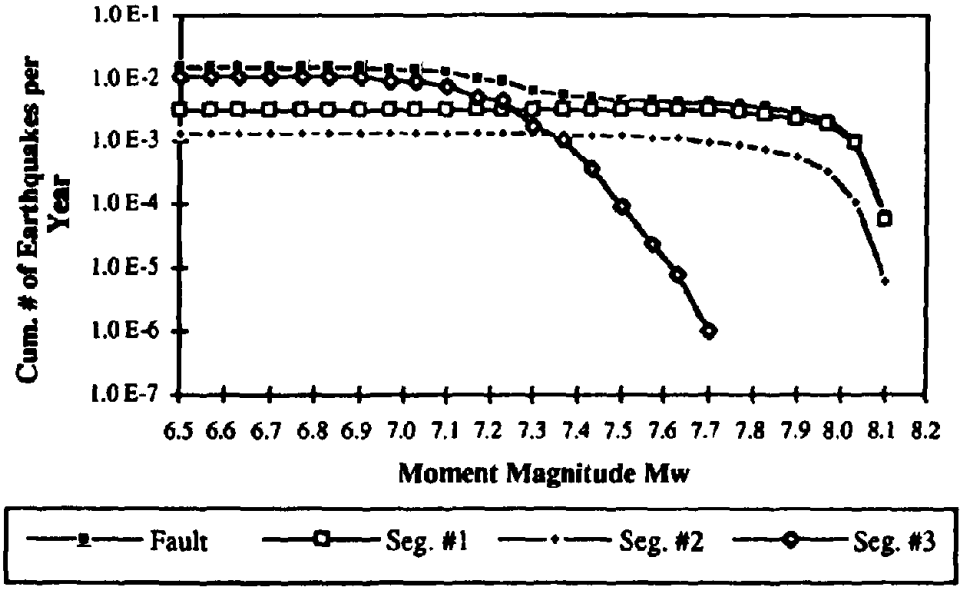


FIGURE 5.4 Number of earthquakes per year for the base case

The average number of earthquakes per year with a moment magnitude exceeding 6.5 is 0.0155, or one earthquake every 64.5 years. The annual rate of occurrence of earthquakes with a moment magnitude of 7.9 or greater is 0.0029, or one earthquake every 345 years. Earthquakes falling into this moment magnitude range are of a size that is on the order of the 1906 earthquake. The 345 year recurrence interval predicted by this model is higher than the approximately 150-200 year recurrence interval estimated by others (Ellsworth, et. al., 1981; Thatcher and Lisowski, 1987). One possible reason may be that the other recurrence interval estimates were made for earthquakes rupturing the San Francicso Peninsula section, which may not necessarily be of the same magnitude as the 1906 earthquake.

5.4 PROBABILITY OF EXCEEDING A GIVEN MAGNITUDE DURING A GIVEN TIME HORIZON

An important quantity for site hazard analysis is the probability of exceeding a given magnitude level during a given time period. The given magnitude level will usually be related to a ground motion parameter, such as peak ground acceleration, at the site. The probability of exceeding a given magnitude level can be used to estimate the probability of exceeding a given ground motion level at a site. The probability of exceeding given magnitude levels will be estimated for fixed time intervals as a function of moment magnitude. The same probability will also be estimated for fixed magnitude levels as a function of time.

5.4.1 FORECASTS FOR FIXED TIME INTERVALS

Figure 5.5 shows the probability of experiencing an earthquake with a moment magnitude m or greater during time intervals of 50 and 100 years. The results are based on a simulation of 10,000 cycles of the fixed time interval. It is assumed that at time equal to zero, an earthquake occurs that ruptures the entire fault and releases all accumulated slip. Review of the simulations revealed that all the earthquakes that occurred originated on segment #3. The earthquake history of earthquakes triggering on segment #3 is then the same as the history of earthquakes triggering anywhere on the fault. As a consequence, Figure 5.5 applies to segment #3 and to the entire fault as a whole.

The probability of earthquakes with a moment magnitude greater than or equal to 6.5 is 0.0117 for a time period of 50 years and 0.735 for a time period of 100 years. Note

that for a time period of 50 years, the largest earthquakes observed have a moment magnitude between 6.8 and 6.9. In a time period of 100 years, the largest earthquakes have moment magnitudes of 7.2 to 7.3. Since the simulation begins at the zero slip level, it is not possible for enough slip to accumulate within 50 or 100 years to cause earthquakes with larger magnitudes than this.

5.4.2 FORECASTS FOR FIXED MAGNITUDE LEVELS

Figures 5.6-5.9 show the probability, as a function of time, of the occurrence of an earthquake exceeding certain magnitude levels for the entire fault as a whole and for earthquakes originating on each segment separately. These graphs are based on 10,000 simulations of each time period (20 years, 40 years, etc.) In these graphs, it is again assumed that an earthquake rupturing the entire fault and releasing all accumulated slip occurs at time zero.

Note that in Figure 5.6, the graphs corresponding to moment magnitudes 6.5 and 7.0 differ only by a probability of approximately 0.1. Thus, there are not many earthquakes whose moment magnitudes fall into the range from 6.5 to 7.0. The graph

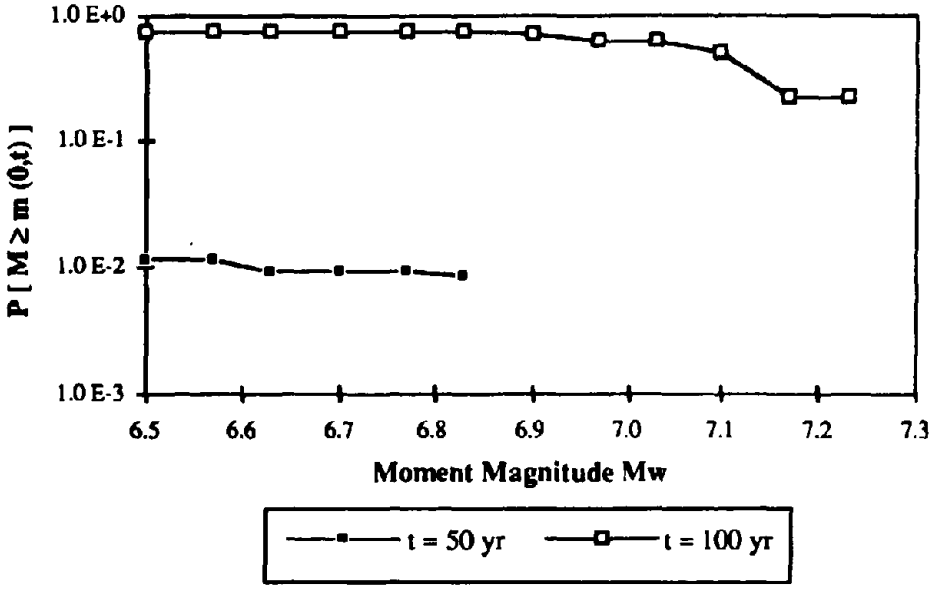


FIGURE 5.5 P[$M \geq m$] during the time intervals (0,50) and (0,100) for the base case

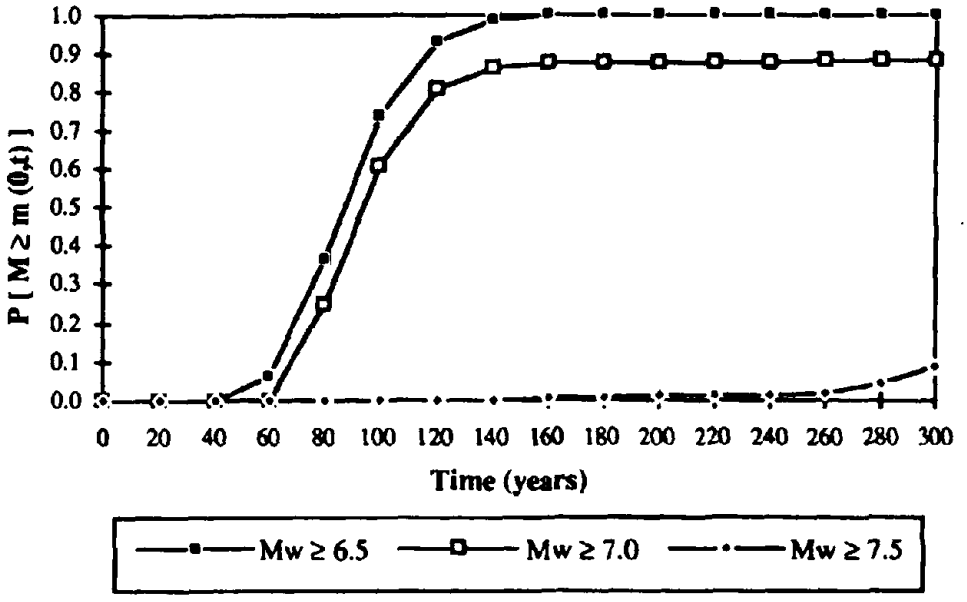


FIGURE 5.6 P[$M \geq m$] for fixed magnitude levels for the base case--Entire fault

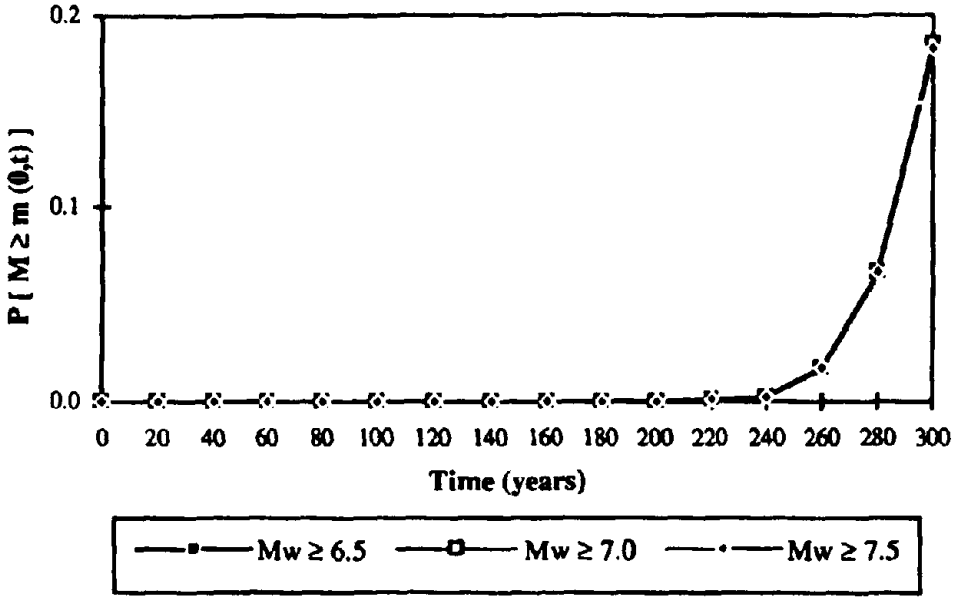


FIGURE 5.7 P[$M \geq m$] for fixed magnitude levels for the base case--Segment #1

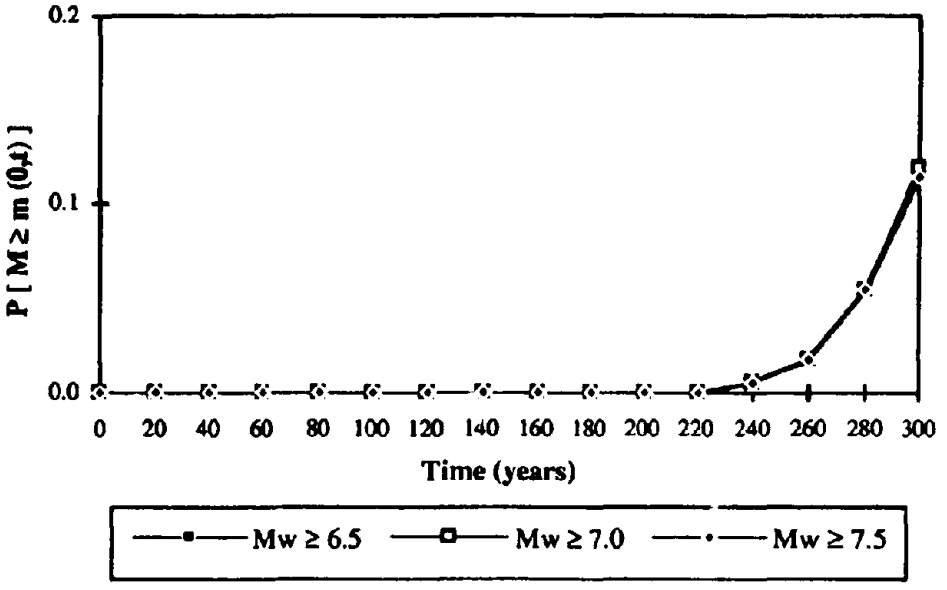


FIGURE 5.8 P[$M \geq m$] for fixed magnitude levels for the base case--Segment #2

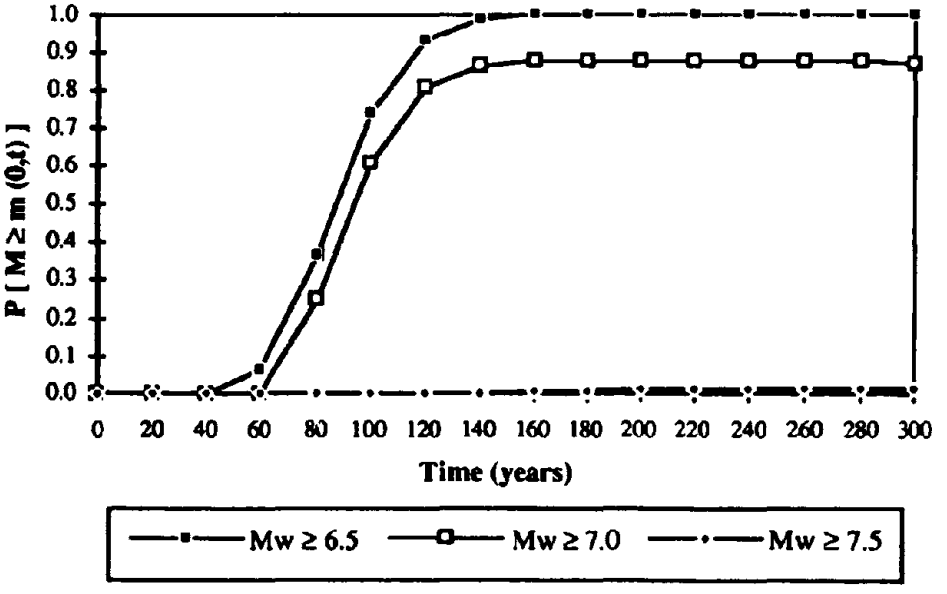


FIGURE 5.9 P[$M \geq m$] for fixed magnitude levels for the base case--Segment #3

corresponding to moment magnitude 7.5 does not begin to rise until approximately 140 years and reaches a final probability of 0.083 at time 300 years.

In Figures 5.7 and 5.8, the graphs corresponding to all three magnitude levels are the same because only earthquakes with a moment magnitude level of 7.5 or higher are actually observed. Comparing the two figures shows that segment #1 triggers more earthquakes than segment #2 as it has a higher probability of experiencing an earthquake. In a time period of 300 years, an earthquake originates on segment #1 with a probability 0.182 whereas the probability of an earthquake originating on segment #2 is 0.114.

Figure 5.9 shows that the earthquakes with magnitude less than 7.5 originate on segment #3. While this segment is responsible for the greatest number of earthquakes, they are almost all of smaller magnitude than those originating on the other segments. Figures 5.6-5.9 confirm the observations of the behavior of the northern San Andreas fault noted in previous sections.

5.5 HAZARD RATE

The hazard rate is defined as the probability that an earthquake exceeding a given magnitude level occurs in the time period $(t, t + \Delta t)$ given that there were no earthquakes in the period $(0, t)$. Figure 5.10 shows the hazard rate for the entire fault using $\Delta t = 10$

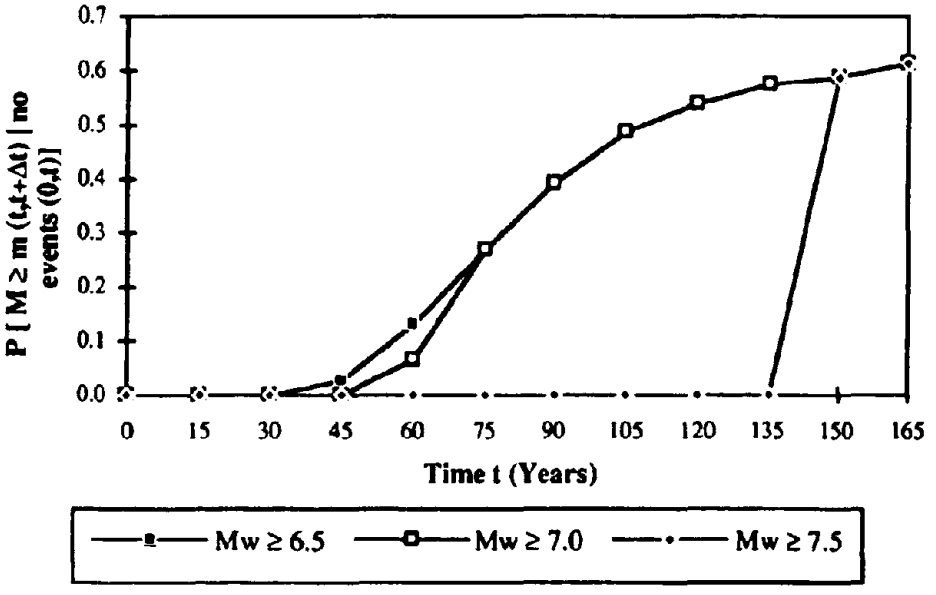


FIGURE 5.10 Hazard rate for the base case ($\Delta t = 10$ years)

years. All the earthquakes whose occurrences are reflected in this graph triggered on segment #3 because Δt is chosen to be short enough that only one earthquake is observed in each time period $(t, t + \Delta t)$. As segment #3 has the shortest mean interarrival time, it is the segment that triggers each earthquake in this simulation. If the behavior of the fault were simulated for longer than the 10 year period, earthquakes would be observed on other segments as well.

Note that for earthquakes with a moment magnitude of 6.5 or greater, the hazard rate begins to rise at 30 years. At 75 years, the hazard rate is the same for earthquakes with magnitudes exceeding 6.5 as it is for earthquakes with magnitudes exceeding 7.0. This reflects the fact that at $t = 75$ years, all the earthquakes that occur within the additional 10 years have a moment magnitude of 7.0 or greater. The same phenomenon occurs at $t = 150$ years; all the earthquakes occurring within the additional 10 years have a moment magnitude of 7.5 or greater. This is a reflection of the fact that the longer the time since the last earthquake, the greater the amount of slip that will be released when there is an earthquake. As discussed above, all earthquakes in this simulation triggered on segment #3. These factors combine to cause the magnitudes of the simulated earthquakes to fall into narrow ranges.

Because the mean time for earthquakes originating on segment #3 (105 years) is significantly smaller than for those originating on segment #2 (350 years) and segment #1 (443 years), the hazard rate calculated above is difficult to interpret. As the amount of time since the last earthquake increases past the mean interarrival time for segment #1, the hazard rate as calculated above begins to lose its meaning. The probability that a gap time of 150 years would be observed in this model is small. For the model using 20 km cells and a slip rate of 19 mm/yr, the mean trigger time for cells on segment #3 is 105 years, with a standard deviation of 30 years. Assuming the lognormal distribution, the probability of no earthquake being triggered by a cell in segment #3 during the 150 year time period following a large earthquake releasing all accumulated slip is less than 0.05%. However, a gap time of 150 years for earthquakes originating on segment #1 and segment #2 would be expected, since the mean trigger times for cells on those segments are much higher.

The most probable scenario then would involve an earthquake originating on segment #3 during a 150 year time period, with no earthquakes triggering on the other segments during those time frames. This scenario, however, does not lend itself to computing a quantity that resembles the hazard rate as defined above. The hazard rate for

the entire fault is driven by the hazard rate for the segment with the smallest interarrival time for triggering earthquakes. Due to the spill over of earthquake from segment to segment, it is difficult to treat each segment separately and compute a hazard rate for each one. In this model, the hazard rate does not have a meaning when it is calculated assuming a gap time that is unrealistic for the segment generating earthquakes most frequently.

A Poisson model has a constant hazard rate, which would be a horizontal line on the graph. Figure 5.10 demonstrates that this model has an increasing hazard rate for gap times up to 105 years. (As discussed above, the hazard rate is not a meaningful quantity as the gap time increased much past the trigger time mean of the segment most often triggering earthquakes.) The elastic rebound theory implies that immediately after an earthquake, the probability of another earthquake is low. For large magnitude earthquakes, this model incorporates that idea.

5.6 COMPARISON WITH USGS ESTIMATES

The USGS Working Group (1990) has estimated the probabilities of earthquakes rupturing more than 50% of each of the northern San Andreas fault segments during the time period 1990-2020. The Working Group assumed a renewal model based on a lognormal distribution of interarrival times. Their estimates will be compared to the estimates obtained with this model. Both the Working Group and this model application result in the probabilities of an earthquake on the South Santa Cruz Mountains segment (segment #3) that are virtually zero during this time period. The Working Group estimates the probability of an earthquake on the North Coast (segment #1) to be 0.02 during this time frame. The corresponding probability obtained from this model is virutally zero. The Working Group states that they do not consider differences of probability less than 0.1 to be significant. Therefore, this model and the Working Group are in agreement regarding segment #1 as well as segment #3.

The Working Group and this model differ in the assessment of the probability of an earthquake on segment #2, the San Francisco Peninsula segment. The Working Group estimates this probability to be 0.23 because this segment is thought to have ruptured on its own in 1838. Since it is known to have ruptured 68 years later in 1906, it is quite possible that is will rupture again during the 114 years between 1906 and 2020. The presented model views the San Francisco Peninsula segment quite differently, however. Nearly all the simulated earthquakes that rupture this segment originated on a different

segment. Because this model finds the probability of an earthquake originating on either segment #1 or segment #3 to be virtually zero, the corresponding probability of rupturing more than 50% of segment #2 is also very small.

5.7 CONCLUSIONS

The results based on 20 km cells, a slip rate of 19 mm/yr, and interarrival time means equal to the USGS estimates support the idea of two distinctly different behaviors on the northern San Andreas fault. The South Santa Cruz Mountains segment generates earthquakes with a moment magnitude that is usually in the range of 6.8 to 7.4. The North Coast segment generates earthquakes less frequently, but the moment magnitudes of those earthquakes are usually in the range of 7.7 to 8.1. The San Francisco Peninsula section acts as a transition between those two segments; it more often breaks due to an earthquake spilling over from an adjacent segment than due to an earthquake originating on it. If this is true, then the San Francisco Peninsula segment has a greater capacity for stress accumulation than either the North Coast segment or the South Santa Cruz Mountains segment. This contrasts the view expressed by the Working Group, which felt that this segment has a greater, rather than smaller, chance of rupturing during the time period 1990-2020.

This model predicts a mean interarrival time of earthquakes with a moment magnitude of 7.9 or greater of 345 years. Earthquakes in this magnitude range are similar in size to the 1906 earthquake. In addition, this model shows an increasing hazard rate for gap times of 165 years or less. For larger gap times, the hazard rate is difficult to define for this model and is therefore not estimated.

CHAPTER 6
SENSITIVITY OF RESULTS

6.1 INTRODUCTION

In Chapter 5, the results for the base case model for the northern San Andreas fault were presented. The base case considered 20 km cells, a slip rate of 19 mm/yr, and mean interarrival times equal to the USGS estimates of the mean interarrival times. In the base case, it was assumed that the times between earthquakes triggering on a given cell were lognormally distributed. In reality, though, neither the slip rate nor the mean interarrival time values are known precisely. Furthermore, the choice of cell size is made to allow accurate representation of the segment lengths and could easily be changed. Each of these quantities will be varied in order to observe their effect on the results.

In addition, the times between earthquakes will be modeled as a Weibull, rather than a lognormal, distribution. Finally, an alternate segmentation model, which divides the San Francisco Peninsula segment into two segments, will be considered.

6.2 SENSITIVITY TO CELL SIZE

The cell size is a parameter that is chosen based upon the geometry of the fault. Table 6.1 gives the length of each segment when modeled with 10 km cells, 20 km cells, and 30 km cells. It should be noted that any of the three cell lengths enable the accurate representation of the entire fault. In addition, both 10 km cells and 20 km cells allow the lengths of the segments to be accurately represented. However, the length of segment #3, and to a lesser extent, the length of segment #1 are not accurately represented by 30 km cells. In Chapter 5, the number of earthquakes per year and the probability of an earthquake exceeding a given magnitude during a given time horizon (for fixed time intervals) were estimated by using 20 km cells. These results will be compared to those obtained by considering cell sizes of 10 km and 30 km, keeping the slip rate and the mean interarrival times constant.

Figures 6.1-6.4 show the number of earthquakes per year for each segment and for the entire fault for each cell size. For the entire fault and for segment #1, the results are not sensitive to the cell size. On segment #2, the results are similar for 10 km cells and for 20 km cells. Using 30 km cells, however, leads to cumulative numbers of earthquakes an order of magnitude lower than those obtained from using the other cell

sizes. The maximum moment magnitude observed for earthquakes originating on segment #3 differs depending upon the cell size.

Figures 6.5 and 6.6 show the probability of an earthquake occurring that exceeds a given magnitude during time periods 50 and 100 years long, respectively. Each of these time periods begins when a large earthquake ruptures the entire fault and releases all accumulated slip. In these figures, all the earthquakes that occurred originated on segment #3. Figure 6.5 shows that using 30 km cells leads to earthquakes of only two different magnitude levels during the 50 year time horizon. This is due to the fact that only discrete numbers of cells can rupture during an earthquake. The larger the cell size, the more the moment magnitudes of the earthquakes tend to cluster together. The results for 10 km cells and 20 km cells do not differ radically. Figure 6.6 again shows similar results for all cell sizes except at the very largest magnitudes.

TABLE 6.1 *Lengths of each segment using different cell sizes*

		10 km Cells	20 km Cells	30 km Cells
Segment #1 (North Coast)	# of Cells	34	17	12
	Model Length (km)	340	340	360
	USGS Length (km)	340	340	340
	% Diff. in Model	0.0 %	0.0 %	5.9 %
Segment #2 (SF Peninsula)	# of Cells	6	3	2
	Model Length (km)	60	60	60
	USGS Length (km)	61	61	61
	% Diff. in Model	-1.6 %	-1.6 %	-1.6 %
Segment #3 (S SC Mountains)	# of Cells	4	2	1
	Model Length (km)	40	40	30
	USGS Length (km)	39	39	39
	% Diff. in Model	2.6 %	2.6 %	-23.1 %
Entire Fault	# of Cells	44	22	15
	Model Length (km)	440	440	450
	USGS Length (km)	440	440	440
	% Diff. in Model	0.0 %	0.0 %	2.3 %

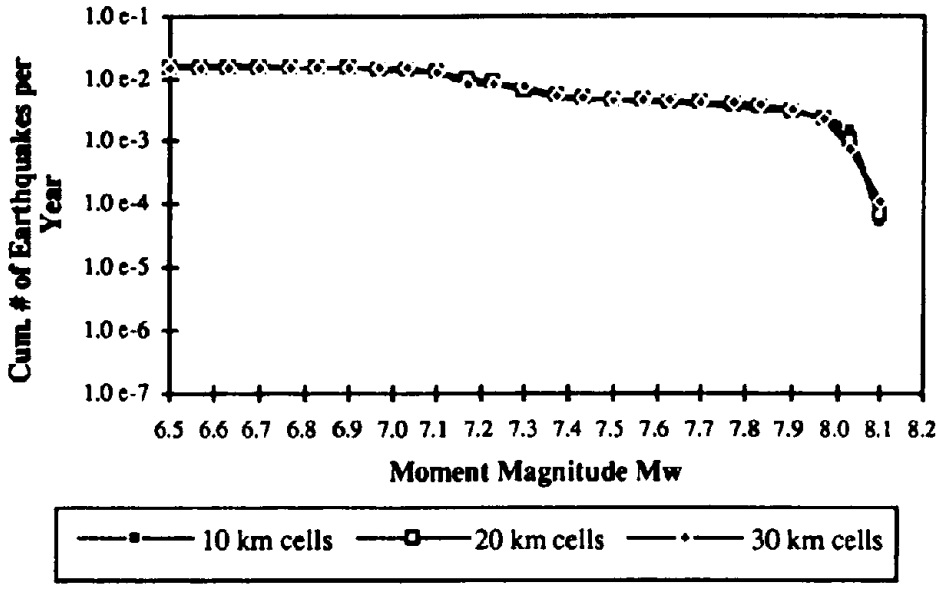


FIGURE 6.1 Number of earthquakes per year for varying cell sizes--Entire fault

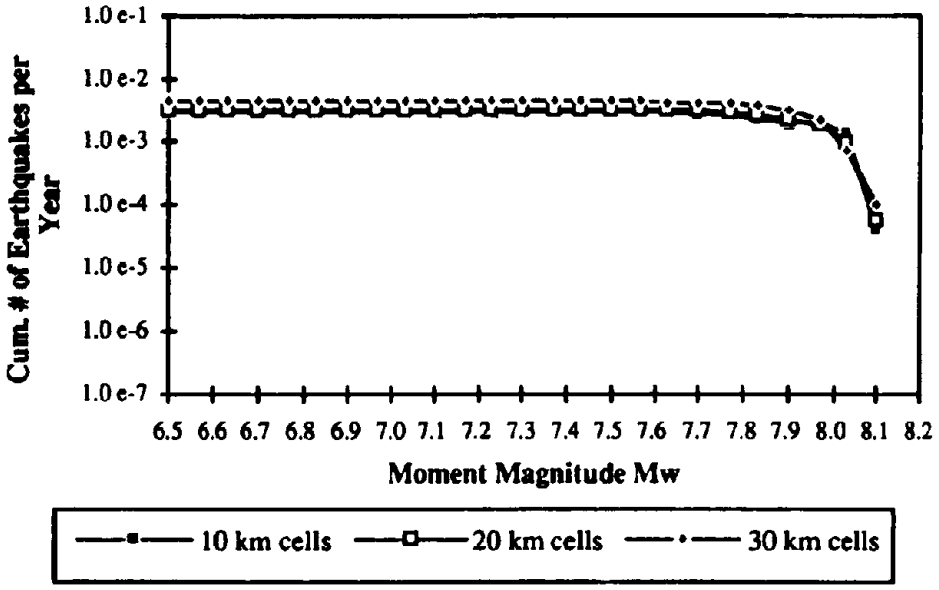


FIGURE 6.2 Number of earthquakes per year for varying cell sizes--Segment #1

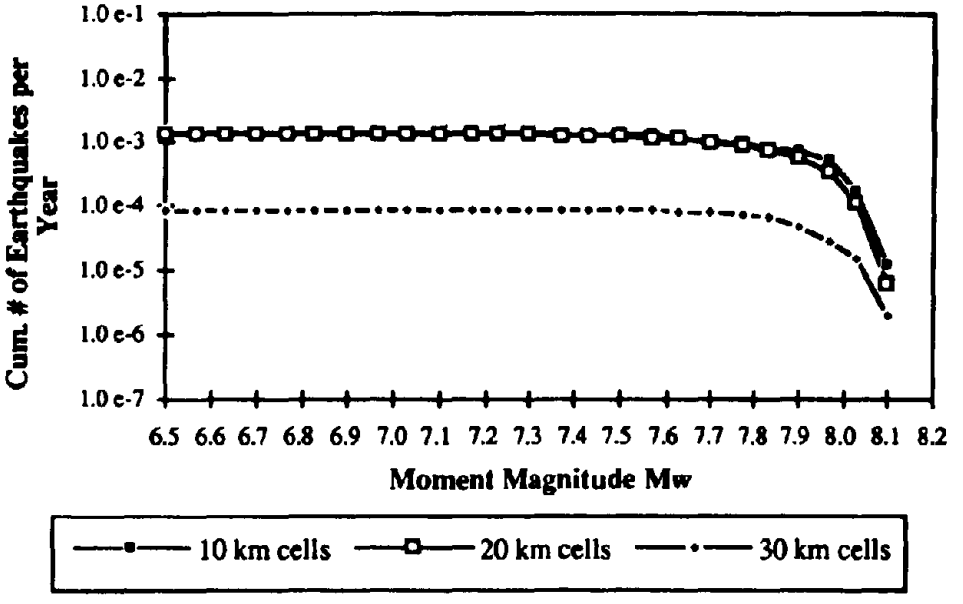


FIGURE 6.3 Number of earthquakes per year for varying cell sizes--Segment #2

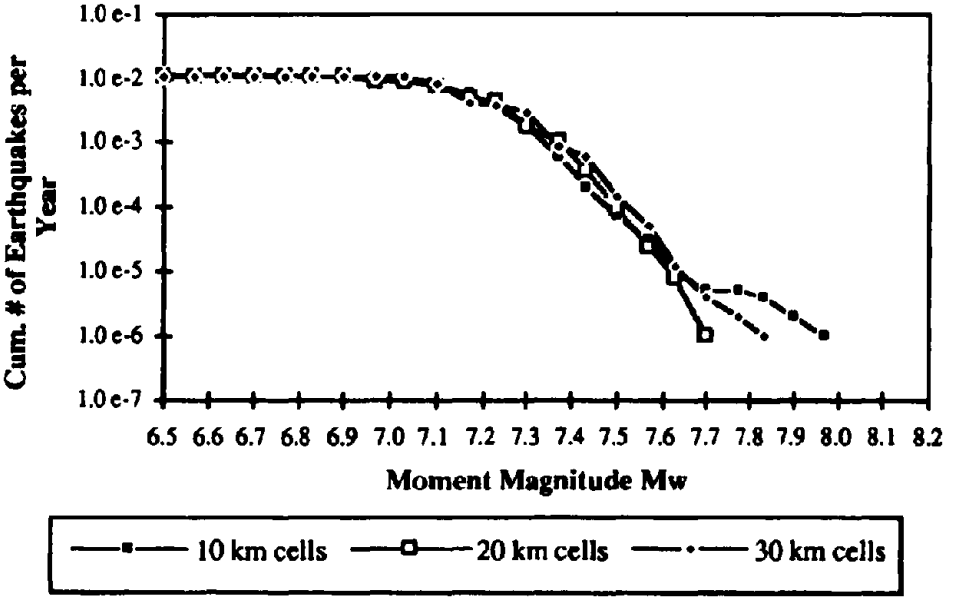


FIGURE 6.4 Number of earthquakes per year for varying cell sizes--Segment #3

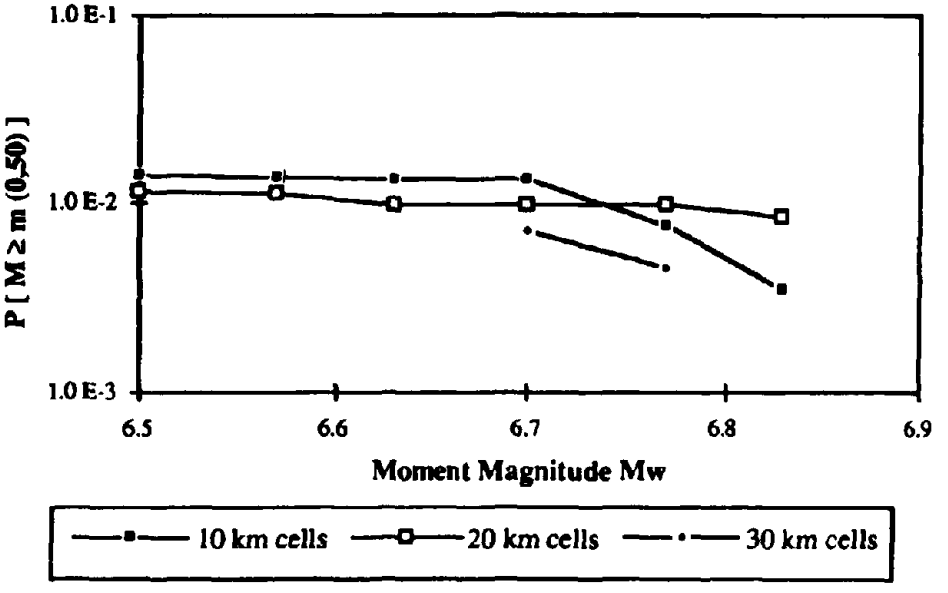


FIGURE 6.5 P[$M \geq m$] during the time interval (0,50) for varying cell sizes

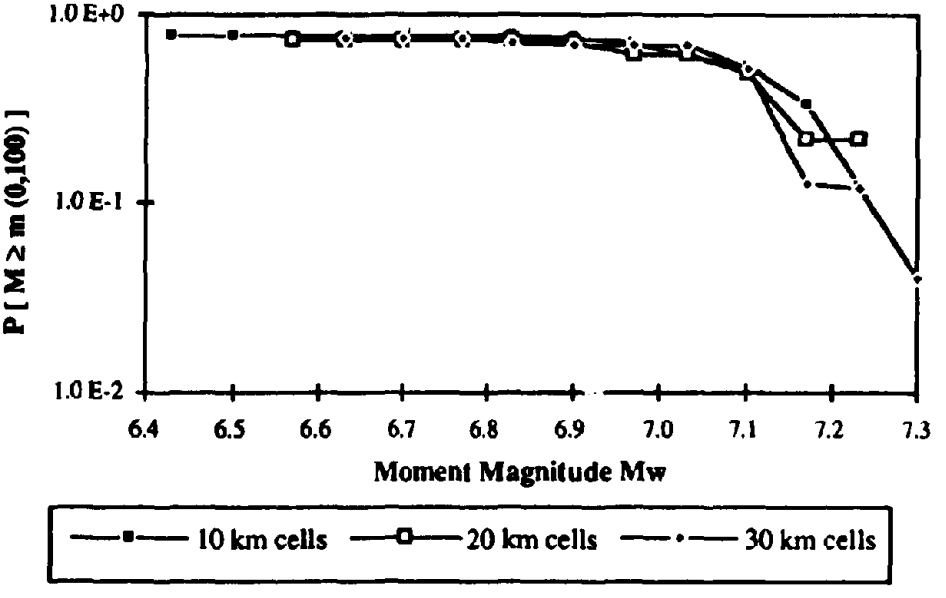


FIGURE 6.6 P[$M \geq m$] during the time interval (0,100) for varying cell sizes

A cell size of 30 km does not appear to be a good choice for modeling the northern San Andreas fault because this cell size does not accurately represent the lengths of the segments. Based upon the results presented here, there is no clear reason to prefer either 10 km cells or 20 km cells. The results from these two cell sizes are similar and differ only at the largest magnitude levels, at which the estimates are the most uncertain due to the relatively few observations of the largest earthquakes.

The length of time it takes to simulate the model depends in part upon the number of cells, as the number of events in the model is directly related to the number of cells. Each cell has two events associated with it--incrementing the amount of accumulated slip and triggering an earthquake. When the cell size is halved, twice as many events must be scheduled and tracked. This causes the computation time to roughly double for the smaller cell size.

Based on these observations, the 20 km cell size appears to be best suited for application to the northern San Andreas fault. This cell size models the lengths of the segments closely. It produces results similar to those obtained with the 10 km cells but avoids the additional computation time.

6.3 SENSITIVITY TO SLIP RATE

The results in Chapter 5 were computed assuming that the slip rate is 19 mm/yr. Slip rates, however, are difficult to estimate from the available information and reported values can vary greatly. The USGS Working Group (1990), for example, assigned a slip rate on the northern San Andreas fault of 19 ± 4 mm/yr.

In order to study the effect of slip rate on the results obtained by this model, results computed with slip rates of 15 mm/yr, 19 mm/yr, and 23 mm/yr are compared, assuming that the cell size is 20 km and that the mean interarrival times are the same as the USGS estimates of these mean values.

Figure 6.7 shows that for the entire fault, the results are not sensitive to the slip rate except at the highest magnitudes. The highest cutoff magnitudes were observed for the case when the slip rate is 23 mm/yr, and the lowest cutoff magnitudes were observed for the case when the slip rate is 15 mm/yr. When the interarrival times are kept constant, as they were in this example, the larger slip rate will have more accumulated slip when an earthquake occurs and hence a larger moment magnitude. Figure 6.8 shows similar behavior for earthquakes originating on segment #1.

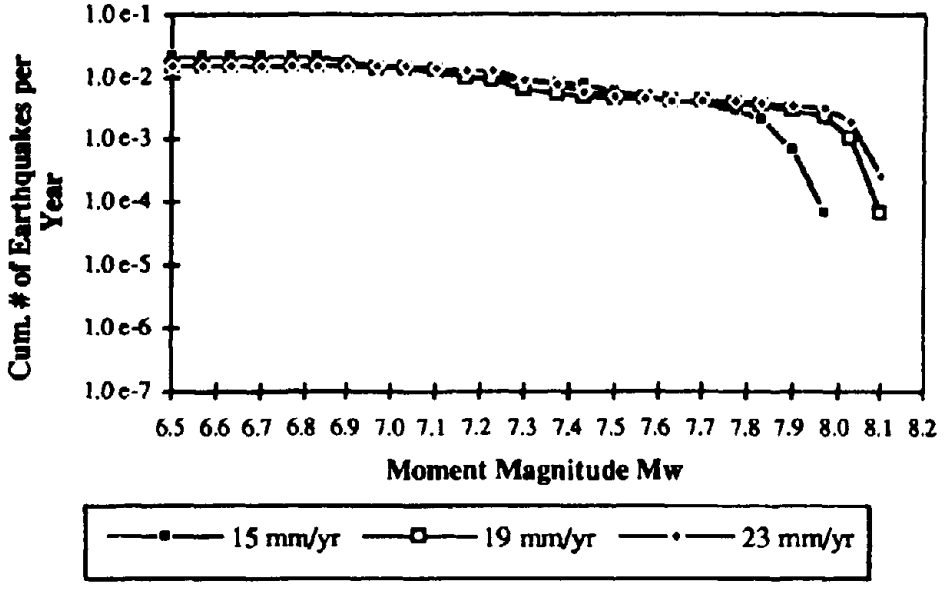


FIGURE 6.7 Number of earthquakes per year for varying slip rates--Entire fault

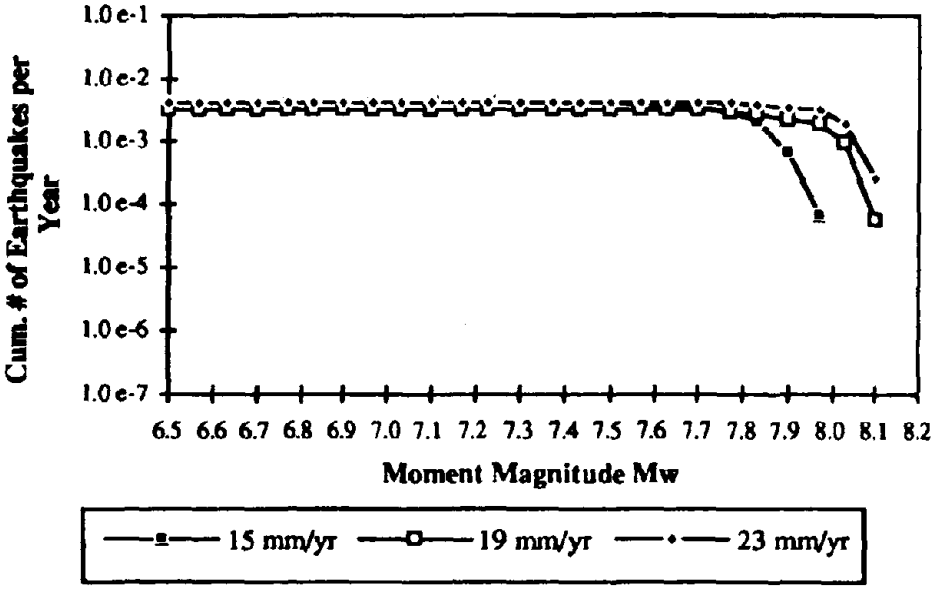


FIGURE 6.8 Number of earthquakes per year for varying slip rates--Segment #1

Figure 6.9 shows that the number of earthquakes per year triggering on segment #2 is sensitive to slip rate. When the slip rate is 23 mm/yr, no earthquakes triggered on segment #2 during the simulation time frame (1,000,000 years). This is due to the fact that at that high slip rate, segment #2 breaks frequently due to rupture spilling over from adjacent segments. Even without triggering any earthquakes itself, segment #2 has an average interarrival time of 107 years, which is smaller than the USGS estimate of 138 years.

When the slip rate is 15 mm/yr or 19 mm/yr, segment #2 does trigger some earthquakes. The cutoff magnitude is again influenced by the choice of slip rate, with the larger cutoff magnitude corresponding to the larger slip rate. In addition, there is a larger number of earthquakes per year with the smaller slip rate. Since a smaller slip rate is associated with smaller magnitude earthquakes, fewer cells rupture during each earthquake. Thus, when the slip rate is smaller, there must be a greater number of earthquakes to rupture all the cells with the same mean interarrival time.

Figure 6.10 shows that segment #3 is also sensitive to the slip rate. As noted previously for the other segments, the size of the cutoff magnitude is dependent upon the choice of slip rate.

Figures 6.11 and 6.12 show the probability of an earthquake occurring that exceeds a given magnitude during time periods of 50 and 100 years, respectively. (As noted before, an earthquake occurred at time zero that ruptured the entire fault and released all accumulated slip.) In these figures, all the earthquakes originated on segment #3. Figure 6.11 demonstrates that within 50 years, there are few earthquakes observed, and they have a narrow magnitude range. As before, the smallest slip rate is associated with the largest probability of earthquake occurrence and with the smallest moment magnitude. The largest slip rate is associated with the largest moment magnitude.

Figure 6.12 shows similar behavior for the 100 year time horizon. The cutoff magnitude is highest when the slip rate is 23 mm/yr and lowest when the slip rate is 15 mm/yr.

The results display considerable sensitivity to slip rate. When the slip rate is higher, the largest observed earthquakes have a higher moment magnitude. In addition, the lowest slip rate is associated with more earthquakes occurring, though they are of somewhat smaller magnitude. The average recurrence interval for earthquakes with a magnitude of 7.9 or greater is 1495 years for a slip rate of 15 mm/yr, 345 years for a slip

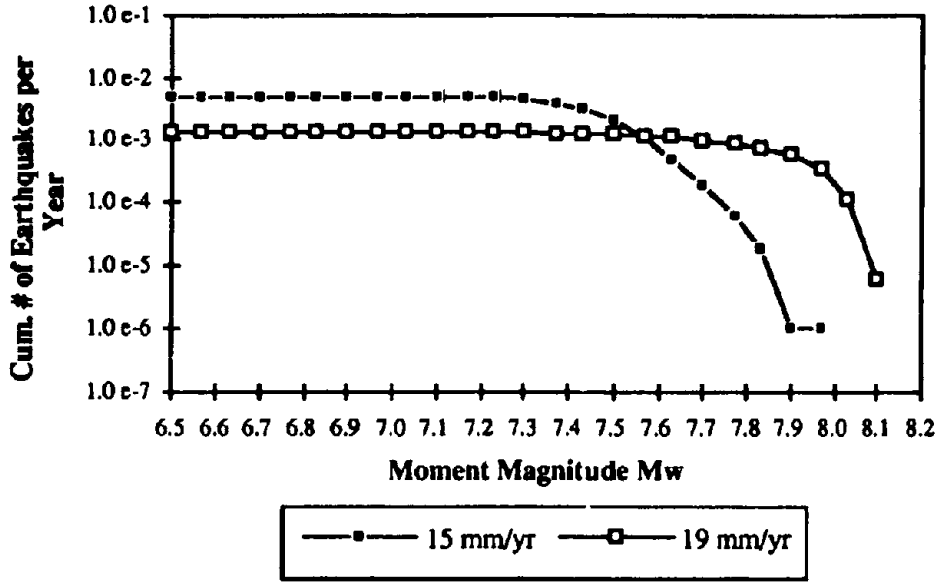


FIGURE 6.9 Number of earthquakes per year for varying slip rates--Segment #2

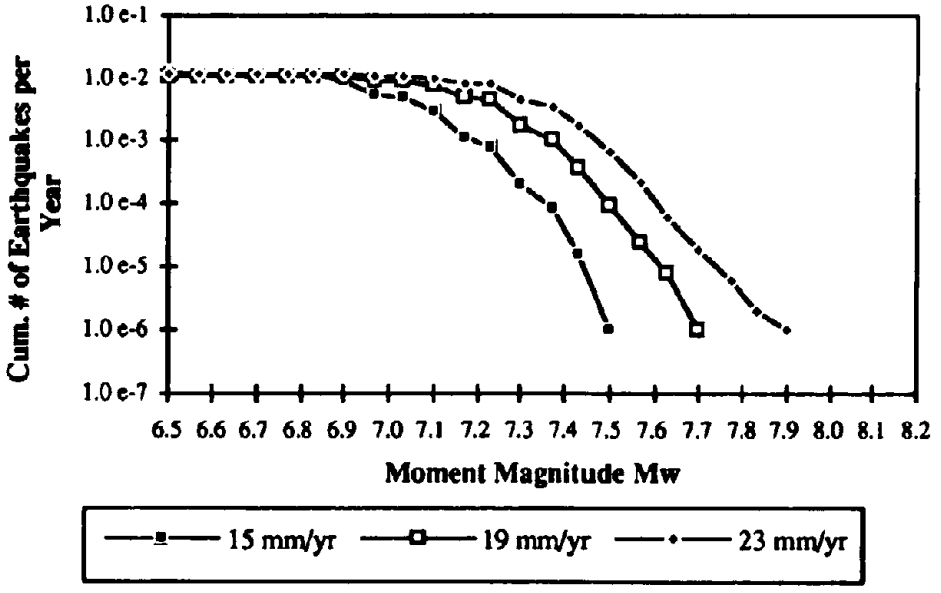


FIGURE 6.10 Number of earthquakes per year for varying slip rates--Segment #3

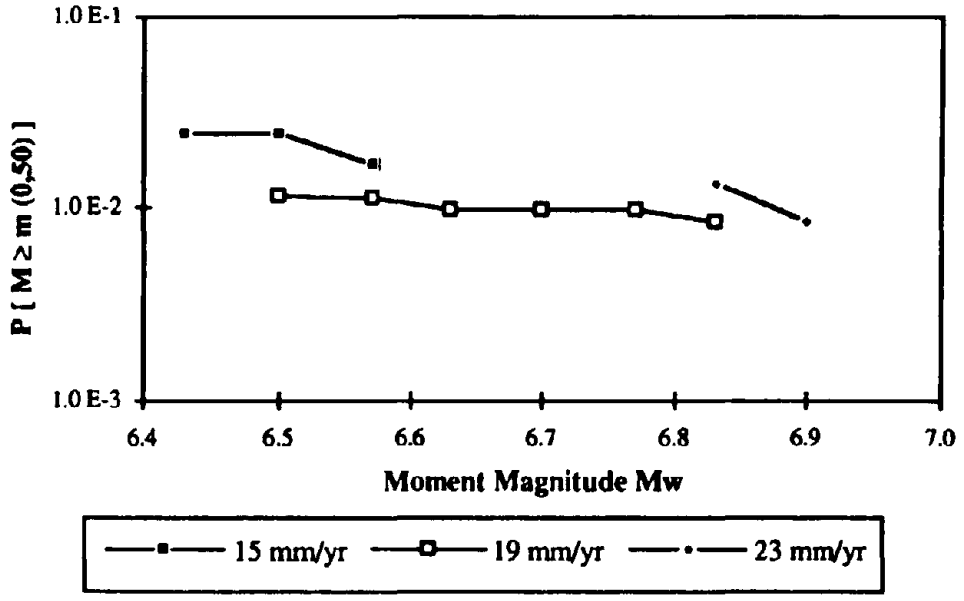


FIGURE 6.11 P[$M \geq m$] during the time interval (0,50) for varying slip rates

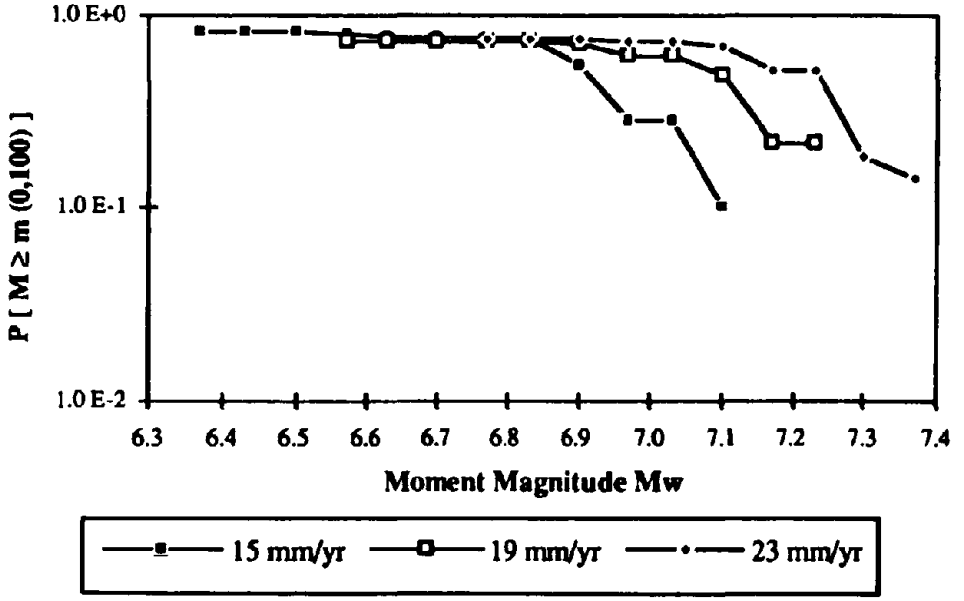


FIGURE 6.12 P[$M \geq m$] during the time interval (0,100) for varying slip rates

rate of 19 mm/yr, and 278 years for a slip rate of 23 mm/yr. The extremely large recurrence interval for the 15 mm/yr case is a reflection of the fact that at the lower slip rate, the largest earthquakes simulated have a lower magnitude. The difference in recurrence interval between the 19 mm/yr and 23 mm/yr cases is much smaller. These results suggest that the major effect noted when the slip rate is varied is the change in the cutoff magnitude of the largest earthquakes.

6.3 SENSITIVITY TO INTERARRIVAL TIMES

All the results presented thus far are for the mean interarrival times as close as possible to those estimated by the USGS Working Group (1990). Figures 6.13-6.18 demonstrate the effect of varying the mean interarrival times. In the legends of these figures, "Mn" refers to all segments having mean interarrival times as estimated by USGS. "1 -" refers to segment #1 having a mean equal to the mean estimated by USGS minus one standard deviation, or 164 years. The mean values of segments #2 and #3 are kept constant under the scenario designated as "1 -". Similarly, "1 +" refers to segment #1 having a mean interarrival time equal to the USGS mean plus one standard deviation, or 310 years, while the mean values of the other two segments are kept constant. Segment #2 has a mean of 98 years in the scenario "2 -" and a mean of 178 years in the scenario "2 +." Segment #3 has a mean of 60 years in "3 -" and a mean of 108 years in "3 +."

Figures 6.13 and 6.14 demonstrate that for the entire fault and for earthquakes originating on segment #1, the choice of mean interarrival time does not have a large impact on the results. The cutoff magnitude varies slightly depending on the scenario, but that effect is quite small. In addition, the cumulative number of earthquakes per year varies from scenario to scenario, but this variance is less than an order of magnitude.

Figure 6.15 shows that the cumulative number of earthquakes originating on segment #2 is highly dependent on the interarrival time scenario selected. When scenario "3 +" is chosen, there are no earthquakes triggering on segment #2. Under scenarios in which earthquakes do trigger on segment #2, "2 +" corresponds to the fewest cumulative number of earthquakes. In this scenario, the mean interarrival time of segment #2 is higher, while the mean values of the other segments are kept constant. Therefore, there should be fewer earthquake triggering on segment #2 in scenario "2+" than in other scenarios. Scenario "2 -" is associated with the largest number of earthquakes triggering on segment #2 because its mean interarrival time is decreased. The cutoff magnitudes are

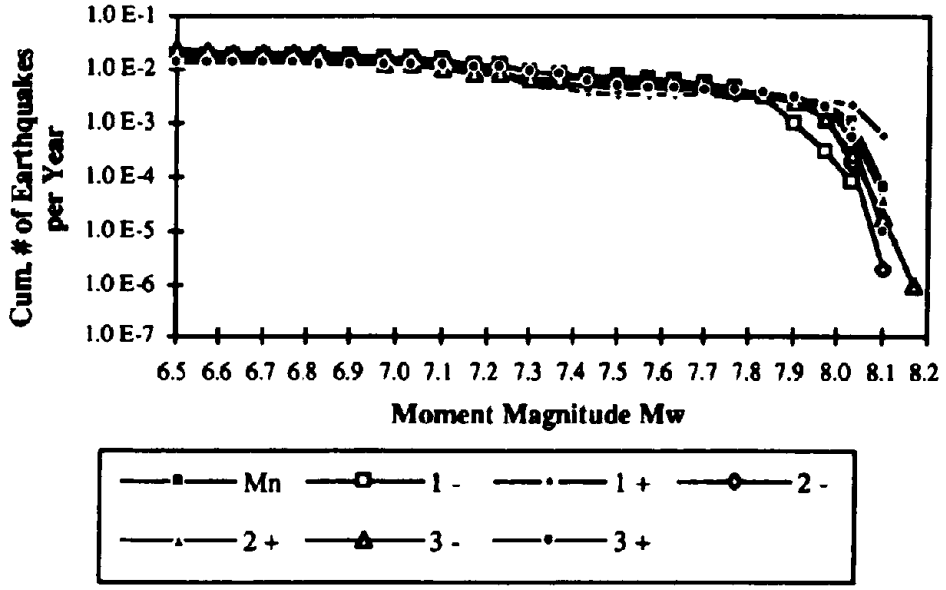


FIGURE 6.13 Number of earthquakes per year for varying interarrival times--Entire fault

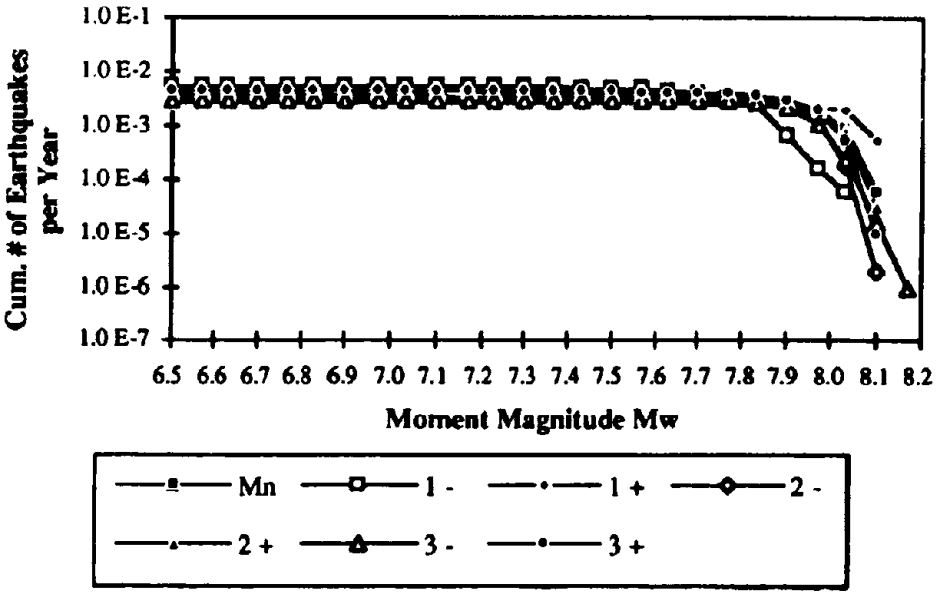


FIGURE 6.14 Number of earthquakes per year for varying interarrival times--Segment #1

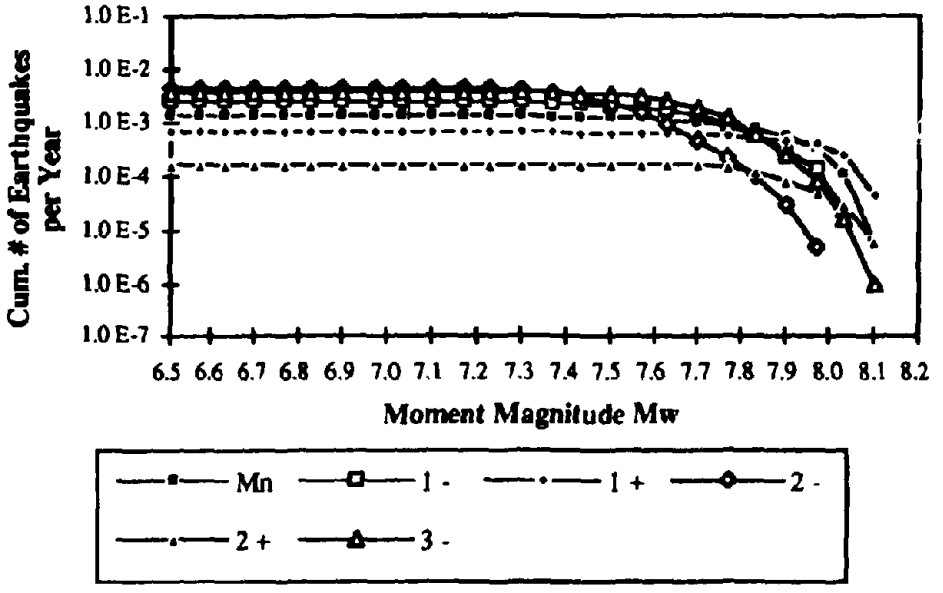


FIGURE 6.15 Number of earthquakes per year for varying interarrival times--Segment #2

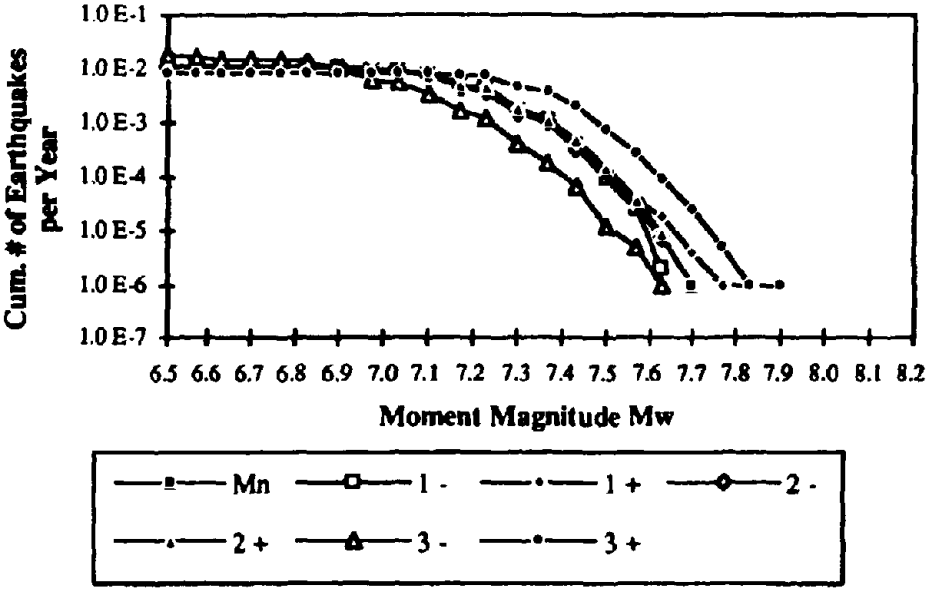


FIGURE 6.16 Number of earthquakes per year for varying interarrival times--Segment #3

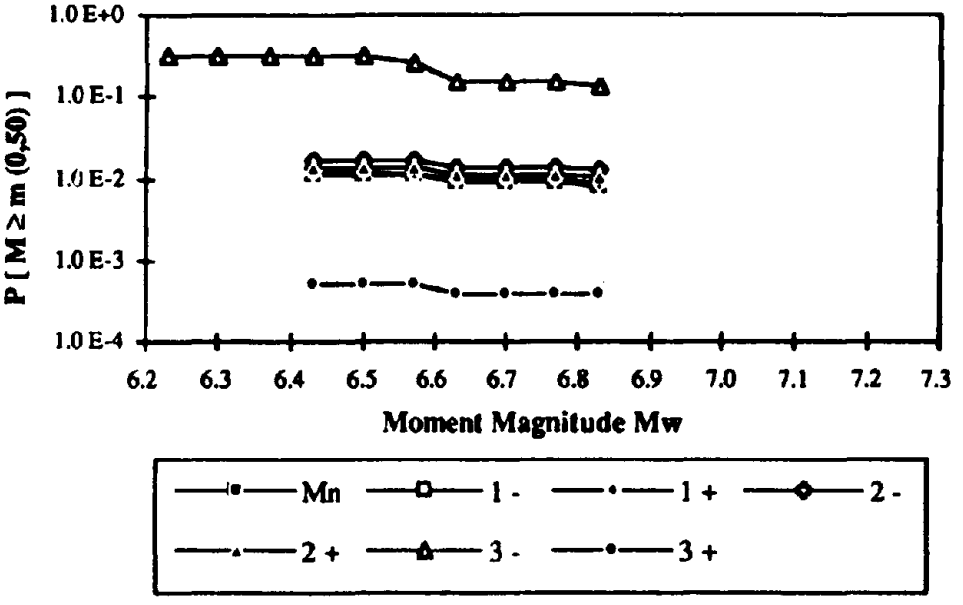


FIGURE 6.17 P[$M \geq m$] during the time interval (0,50) for varying interarrival times

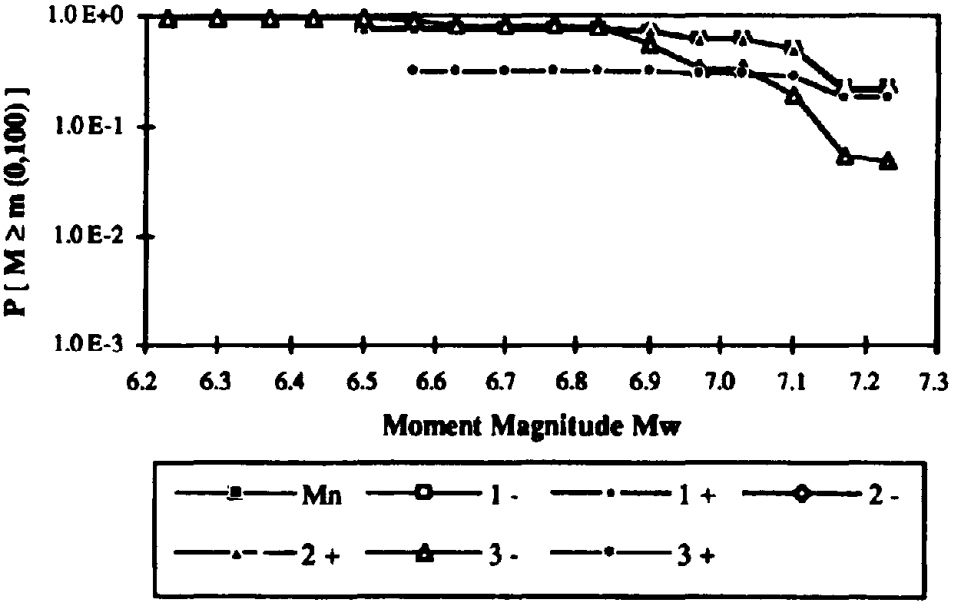


FIGURE 6.18 P[$M \geq m$] during the time interval (0,100) for varying interarrival times

also slightly smaller under the scenario "2 -" as the faster triggering earthquakes do not have as much slip accumulated and are therefore smaller.

Figure 6.16 shows that the primary effect on earthquakes triggering on segment #3 is on the cutoff magnitude. The scenario "3 -" has the smallest cutoff magnitude since the earthquakes occur more frequently and are therefore smaller. As would be expected, the scenario "3 +" is associated with the highest cutoff magnitude. The cumulative number of earthquakes per year is highest under the scenario "3 -" and lowest under the scenario "3 +", but that is not a large effect.

Figure 6.17 shows the results for a 50 year time horizon starting at time zero. For this short time horizon, all observed earthquakes originated on segment #3. The results are highly dependent upon the scenario chosen. The highest probabilities of an earthquake occur when the mean interarrival time on segment #3 is lowered. When the mean interarrival time on segment #3 is increased, the probability of an earthquake decreases dramatically. For all other scenarios, the probabilities are comparable.

Figure 6.18 shows a similar graph but for a time horizon of 100 years. In this case, lowering the mean interarrival time on segment #3 results in a smaller probability of earthquakes with a moment magnitude above 7.0 and with a higher probability of earthquakes with smaller moment magnitude.

These results point out how the behavior of each segment influences the others. Even though the results on segments #2 and #3 vary widely depending upon the chosen interarrival time scenario, the results for the entire fault do not vary greatly. This is because lowering the number of earthquakes that trigger on one segment tends to raise the number of earthquakes that trigger on the other segments. The statistics of the overall fault may not be changed significantly, but the spatial distribution of the earthquakes will be altered by the choice of mean interarrival time.

6.4 SENSITIVITY TO TRIGGER TIME DISTRIBUTIONS

As discussed in Chapter 3, earthquake interarrival times have been modeled in the past with both the lognormal distribution and the Weibull distribution. The results presented in Chapter 5 assume that the times between earthquakes triggering on a given cell are lognormally distributed. Figures 6.19-6.24 compare those results with the results obtained by assuming that the earthquake trigger times are Weibully distributed.

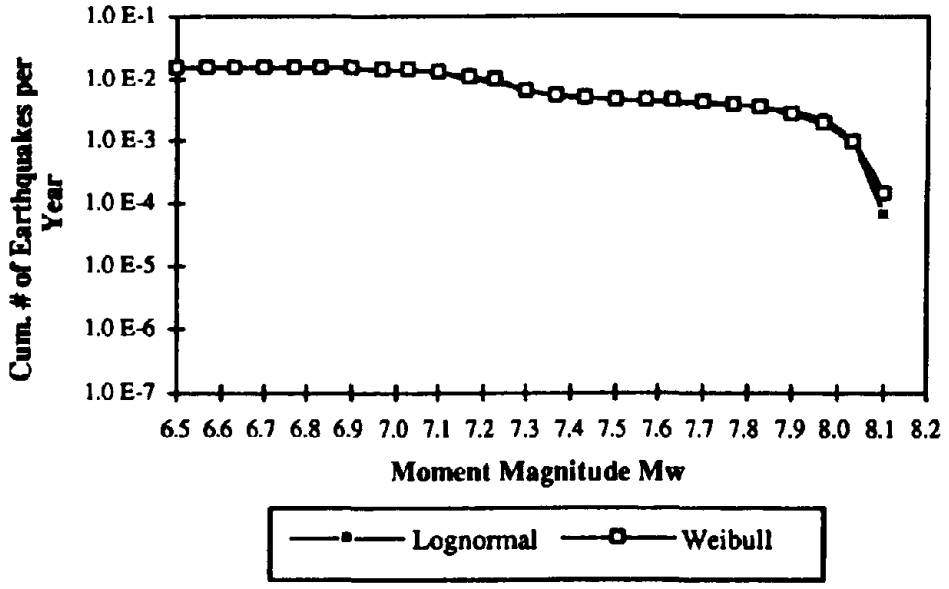


FIGURE 6.19 Number of earthquakes per year for varying trigger time distributions--Entire fault

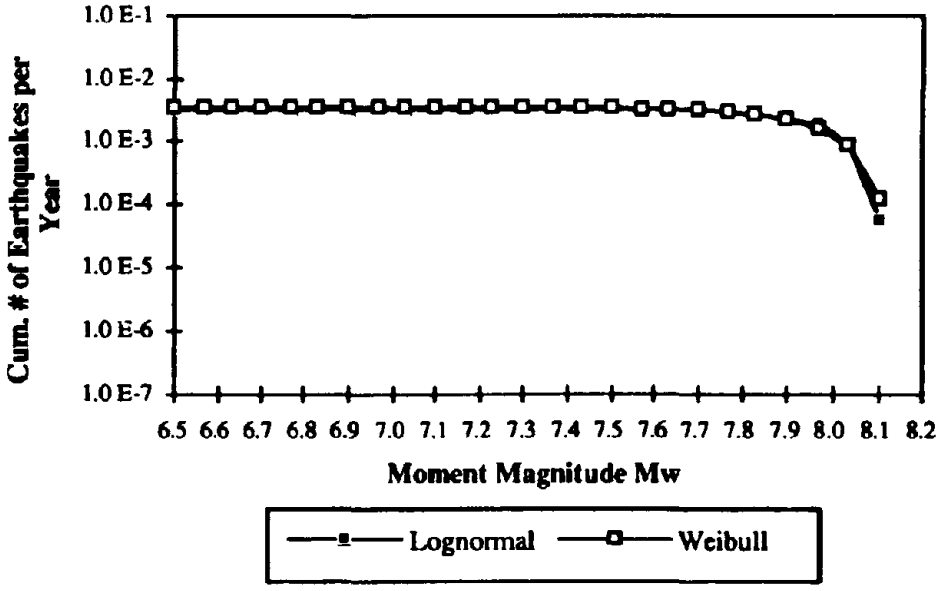


FIGURE 6.20 Number of earthquakes per year for varying trigger time distributions--Segment #1

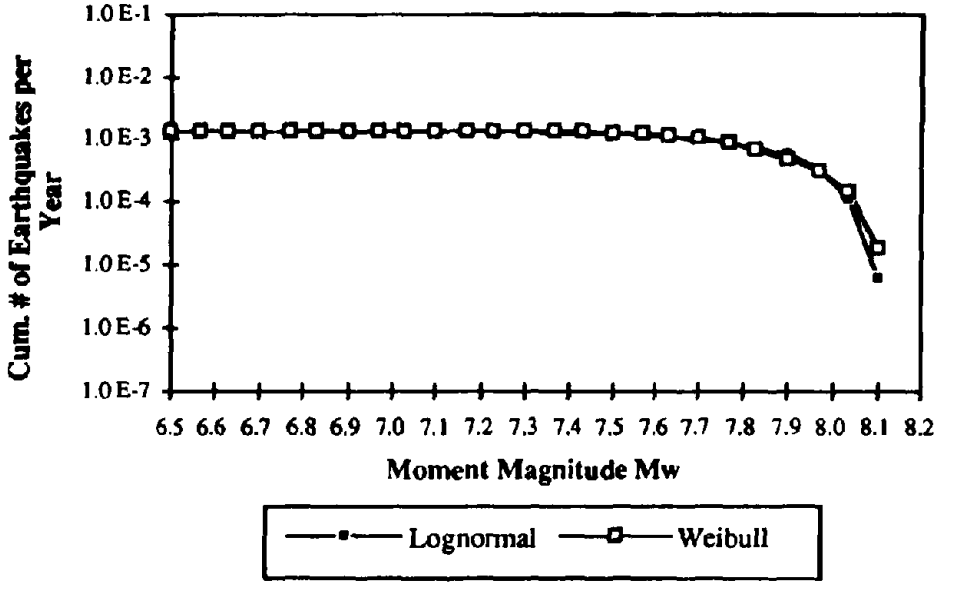


FIGURE 6.21 Number of earthquakes per year for varying trigger time distributions--Segment #2

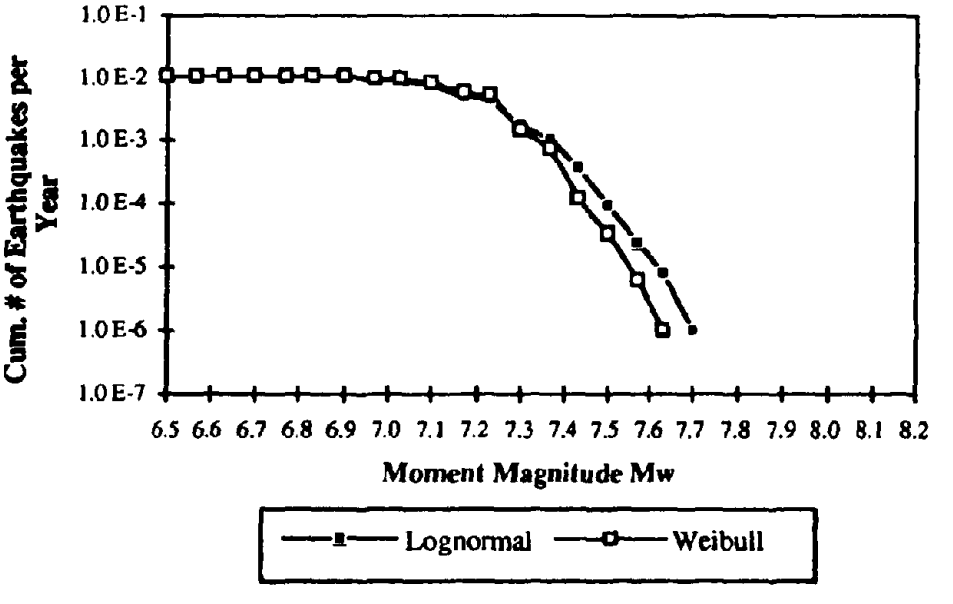


FIGURE 6.22 Number of earthquakes per year for varying trigger time distributions--Segment #3

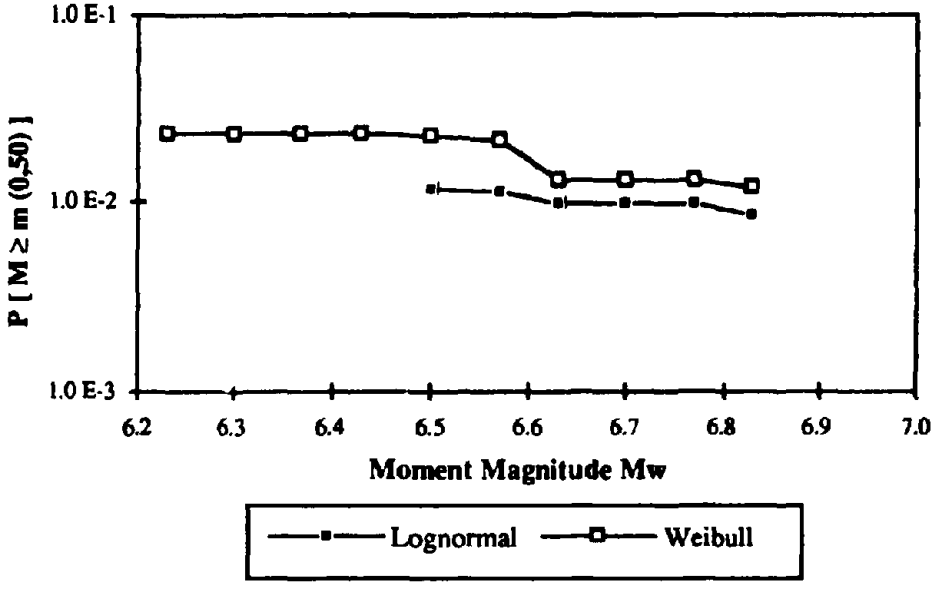


FIGURE 6.23 P[$M \geq m$] during the time interval (0,50) for varying trigger time distributions

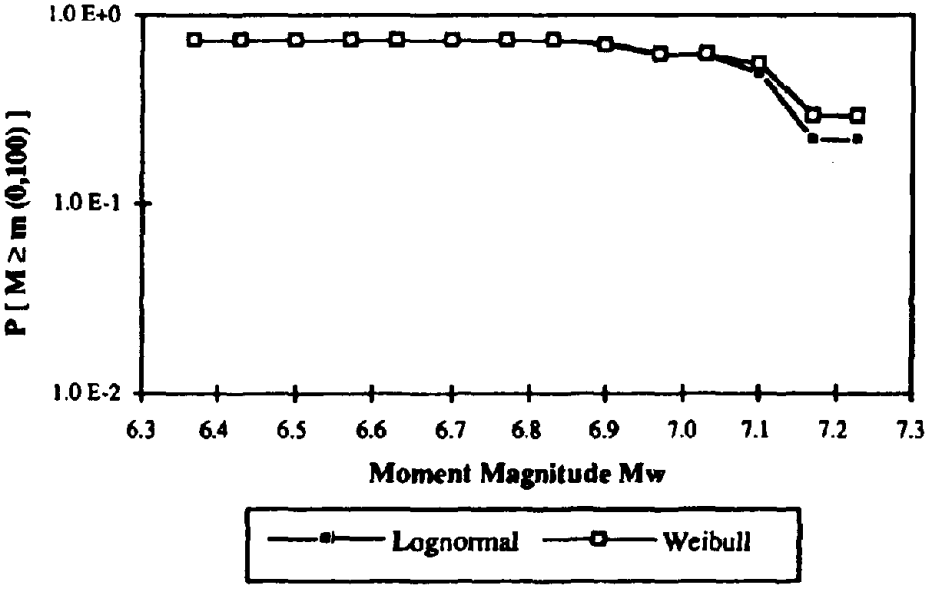


FIGURE 6.24 P[$M \geq m$] during the time interval (0,100) for varying trigger time distributions

Figures 6.19-6.21 demonstrate that the number of earthquakes per year triggering anywhere on the fault, triggering on segment #1, and triggering on segment #2 are relatively insensitive to the choice of trigger time distribution. Figure 6.22 shows a slight sensitivity to the trigger time distribution chosen for earthquakes that initiated on segment #3. On that segment, the lognormal distribution shows a cutoff magnitude that is slightly higher than the cutoff magnitude associated with the Weibull distribution.

Figures 6.23 and 6.24 show the probability of an earthquake exceeding a given magnitude level during time periods of 50 years and 100 years after an earthquake that ruptures the entire fault and releases all accumulated slip. All the earthquakes in these graphs triggered on segment #3. (This is due to the short time horizons selected. Earthquakes would trigger on other segments if the time horizon were longer.) Figure 6.23 shows that the Weibull distribution yields a higher probability of occurrence. Figure 6.24 shows little difference between the results based upon the distribution chosen, but the Weibull distribution is associated with slightly higher probabilities of occurrence.

The Weibull distribution is broader than the lognormal distribution for the same mean and standard deviation. There is therefore more probability in the tails of the Weibull distribution than in the lognormal distribution. That implies that the chance of simulating a low number from the Weibull distribution is greater than the chance of simulating a low number from the lognormal distribution. Each segment is composed of many cells, each of which has its trigger time set by simulating from the distribution. The smallest trigger time initiates the next earthquake. The Weibull distribution, with its wider tails, will tend to have a shorter trigger time than the lognormal distribution. This, in turn, leads to smaller interarrival times with smaller magnitude earthquakes, confirming the behavior observed above.

6.5 SENSITIVITY TO SEGMENTATION

Adopting the segmentation proposed by the USGS Working Group (1990), it was assumed in the results presented in Chapter 5 that the Northern San Andreas fault is composed of three segments: the North Coast segment, the San Francisco Peninsula segment, and the South Santa Cruz Mountains segment. The Working Group also considered an alternate segmentation in which the San Francisco Peninsula segment was composed of two subsegments, the Mid-Peninsula segment and the North Santa Cruz Mountains segment. Table 6.2 shows the length of each segment in the alternate

TABLE 6.2 Lengths of segments in alternate segmentation model

Segment Name	Number of Cells	USGS Length (km)	Model Length (km)	% Diff. in Model
North Coast	17	340	340	0.0 %
Mid-Peninsula	2	41	40	2.5 %
N. Santa Cruz Mountains	1	20	20	0.0 %
S. Santa Cruz Mountains	2	39	40	2.6 %

segmentation as estimated by USGS and as modeled in this dissertation. Table 6.3 shows the input and output parameters for the alternate segmentation.

Figure 6.25 shows the input and output means for the alternate segmentation. The data for the base case scenario was given in Figure 5.1. For the North Coast segment, the alternate segmentation is associated with more variability in the means than is the base case. The mean interarrival times of the San Francisco Peninsula segment can be more closely modeled by the alternate segmentation, which divides the segment into two subsegments. Both segmentation models allow the mean interarrival time for earthquakes on the South Santa Cruz Mountains segment to be accurately modeled.

Figure 6.26, which is analogous to Figure 5.2, shows the input and output standard deviations for the alternate segmentation. In general, the standard deviations are more closely modeled in the alternate segmentation than in the base case. This is especially true of the North Coast segment, which shows an average output standard deviation of 90 years, as contrasted with the base case's value of 119 years.

Figures 6.27-6.30 show the number of earthquakes per year for the base case and for the alternate segmentation. For the entire fault, for the North Coast segment, and for

TABLE 6.3 Input and output statistics for the alternate segmentation

Segment Name	Input Mean	Input St. Dev.	Output Mean	Output St. Dev.	USGS Mean	USGS St. Dev.
	yr	yr	yr	yr	yr	yr
North Coast	436	75	237	90	237	73
Mid-Peninsula	229	50	130	70	129	49
N. Santa Cruz Mountains	500	50	91	36	95	44
S. Santa Cruz Mountains	103	30	85	22	84	24

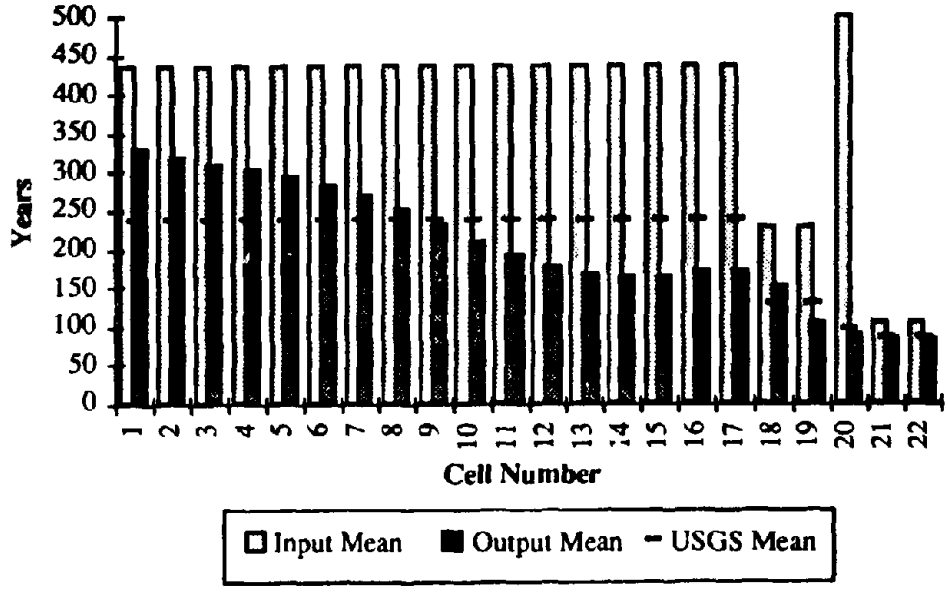


FIGURE 6.25 Input and output means for the alternate segmentation

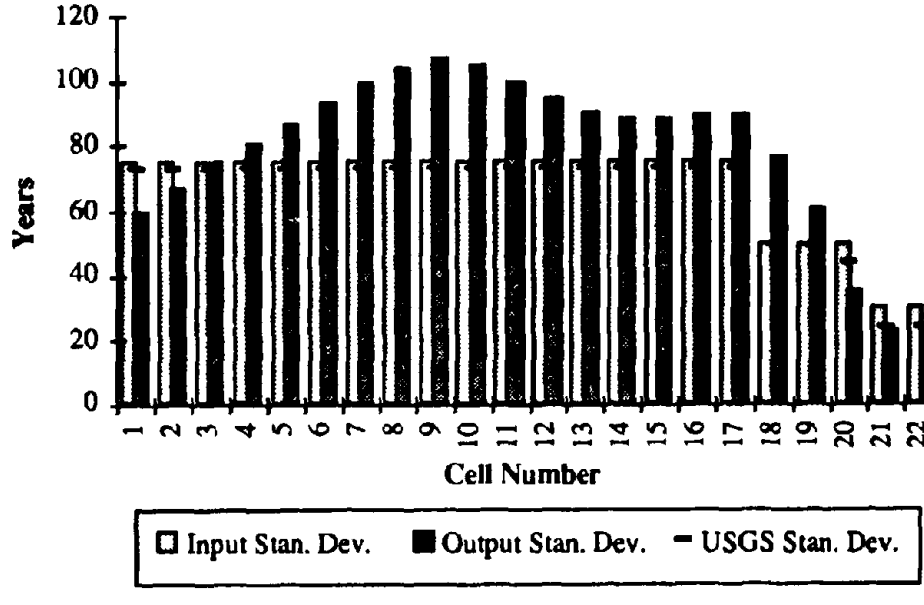


FIGURE 6.26 Input and output standard deviations for the alternate segmentation

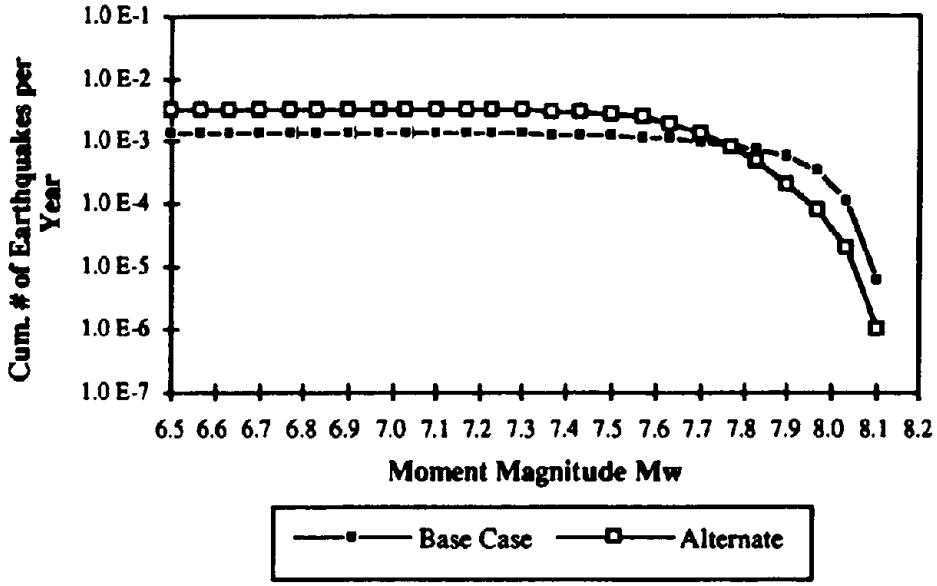


FIGURE 6.27 Number of earthquakes per year for varying segmentations--Entire fault

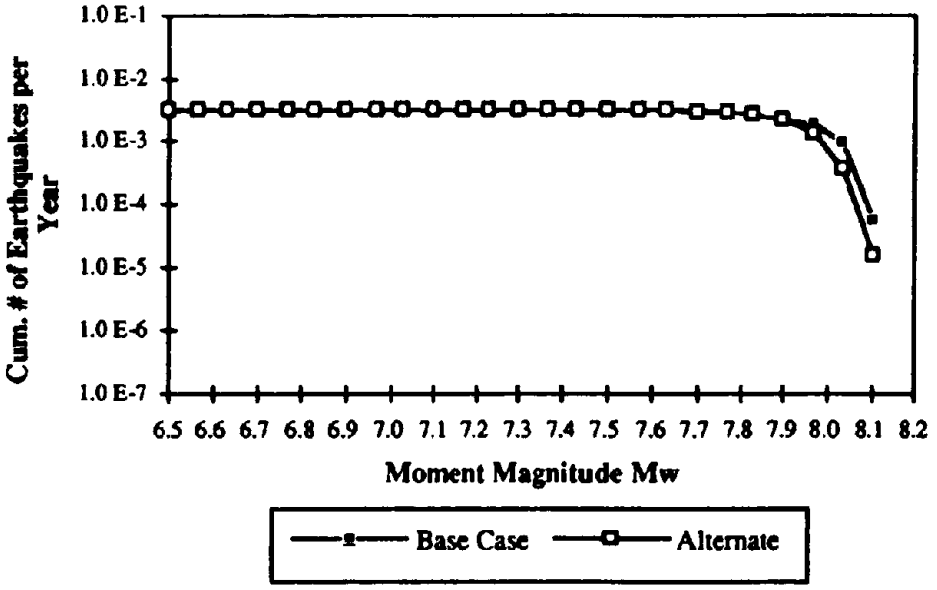


FIGURE 6.28 Number of earthquakes per year for varying segmentations--North Coast

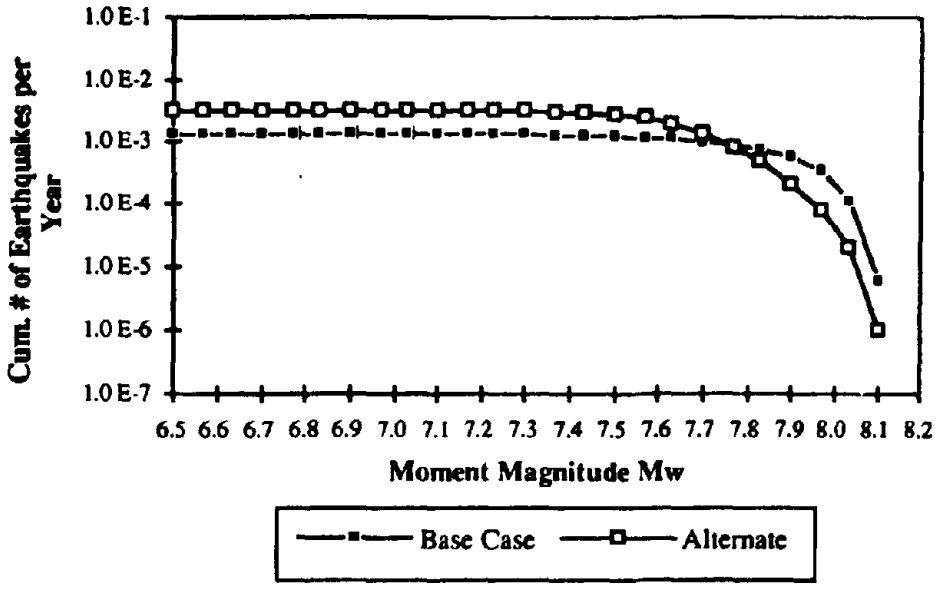


FIGURE 6.29 Number of earthquakes per year for varying segmentations--San Francisco Peninsula

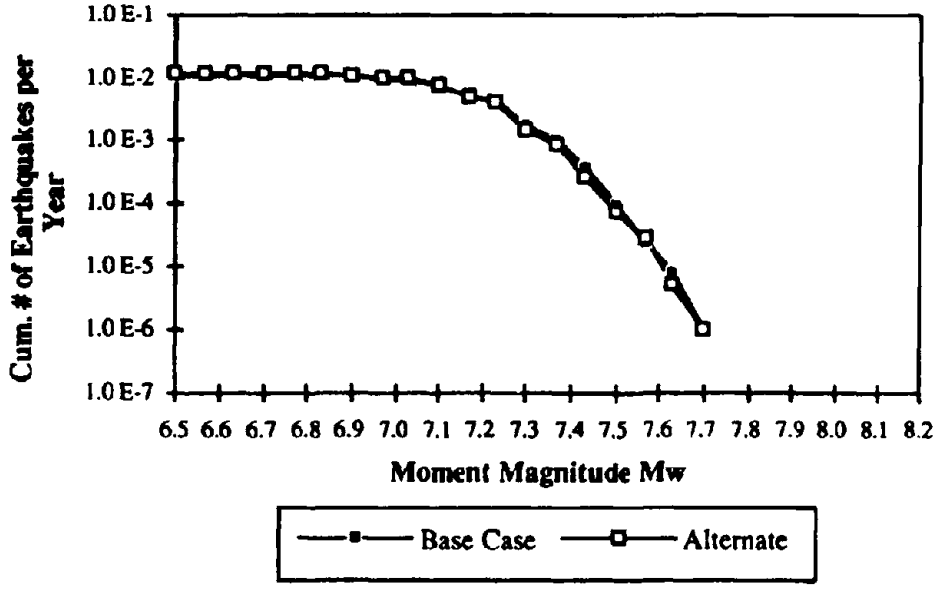


FIGURE 6.30 Number of earthquakes per year for varying segmentations--South Santa Cruz Mountains

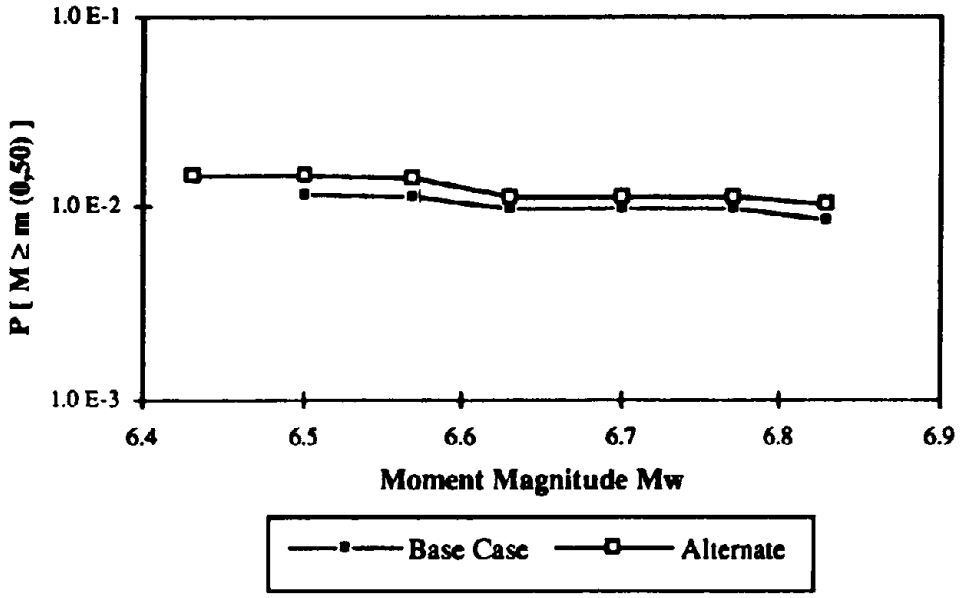


FIGURE 6.31 P[$M \geq m$] during the time interval (0,50) for varying segmentations

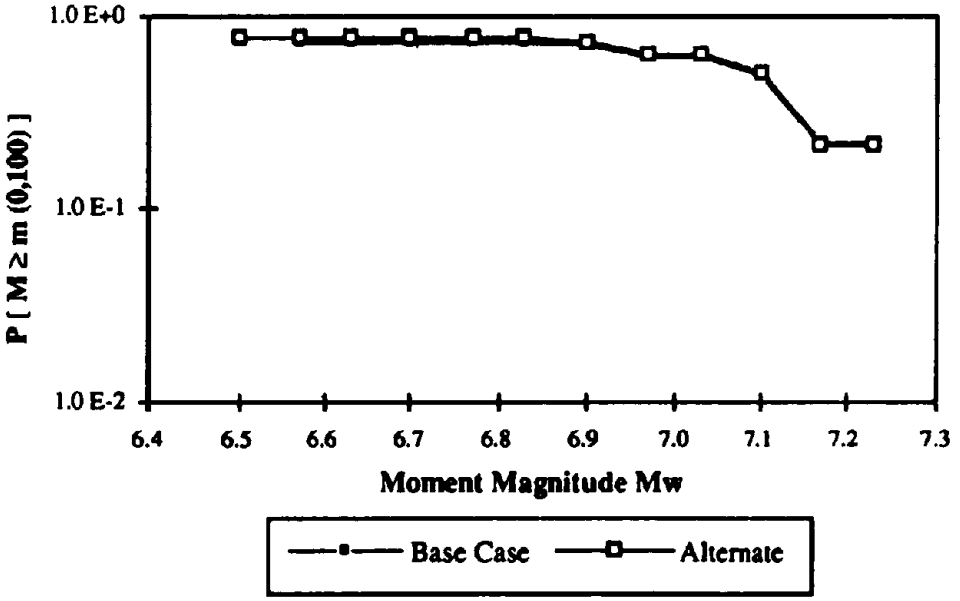


FIGURE 6.32 P[$M \geq m$] during the time interval (0,100) for varying segmentations

the South Santa Cruz Mountains segment, the results from the two segmentations are comparable. Figure 6.29 shows that for the San Francisco Peninsula segment, the alternate segmentation yields smaller maximum magnitudes and more frequent earthquakes. All the earthquakes reflected in the alternate segmentation part of the graph triggered on the Mid-Peninsula subsegment. Because the North Santa Cruz Mountains segment ruptures frequently when rupture spills over from another segment, its input mean is very high, which prevents it from triggering earthquakes.

The USGS Working Group estimates the mean interarrival time on the Mid-Peninsula segment to be 129 years and the mean interarrival time on the North Santa Cruz Mountains segment to be 95 years. For the base case, however, these two segments were both considered to be part of the San Francisco Peninsula segment, for which the Working Group estimated the mean interarrival time to be 138 years. Since the alternate segmentation has a smaller mean interarrival time on the San Francisco Peninsula than the base case, the number of earthquakes per year should be larger for the smaller magnitudes and smaller for the larger magnitudes. This confirms the behavior shown in Figure 6.29.

Figures 6.31 and 6.32 show the probability of earthquakes within time periods of 50 years and 100 years, respectively. Due to the short time periods, the earthquakes shown on these graphs triggered on the South Santa Cruz Mountains segment. The figures show that the results are similar for the two segmentation. It is not surprising that the results for the two segmentations are comparable since the alternate segmentation does not involve changing the configuration of the segment that is generating the earthquakes.

CHAPTER 7
CONCLUSIONS

7.1 SUMMARY

This research develops an earthquake occurrence model that applies to faults that generate earthquakes displaying temporal and spatial dependence. These are generally large magnitude earthquakes (approximately moment magnitude 6.5 and greater) and occur infrequently. However, because they have a large magnitude, they are potentially catastrophic when they do occur. The ability to accurately model the occurrences of these earthquakes is an important tool for analysis of site specific seismic hazard.

The presented model is a generalized semi-Markov process, which is a stochastic process that has been previously used in applications such as queue theory. In this dissertation, the GSMP is brought into the field of earthquake engineering with its application to earthquake occurrence modeling.

The application of this model to the northern San Andreas fault predicts that the fault can be expected to repeatedly generate earthquakes similar in size and space to the Loma Prieta earthquake with an interarrival time that is on the order of 80-100 years. Larger earthquakes similar to the 1906 earthquake can also be expected, but with a much larger interarrival time (on the order of several hundred years). The San Francisco Peninsula segment appears to be a transition segment in this model, which repeatedly ruptures as part of an earthquake that initiates on an adjacent segment but does not often trigger an earthquake on its own.

This model describes temporally and spatially dependent earthquakes, which usually have moment magnitudes of approximately 6.5 and greater. Smaller magnitude earthquakes can also cause damage, particularly near the epicenter. In order to completely describe the earthquake hazard in areas subject to temporally and spatially dependent earthquakes, the presented model must be combined with a model designed to describe the lower magnitude seismicity.

7.2 CONCLUSIONS

The generalized semi-Markov process is a well-studied stochastic process that is applicable to the problem of modeling temporally and spatially dependent earthquakes.

The advantages of using the GSMP include its ability to have a complex state space, its random time between state transitions, and its convenient framework for simulation.

7.2.1 MODEL APPLICATION

Several conclusions can be drawn from the application of this model to the northern portion of the San Andreas fault. The predictions made include:

- Segment #1 (North Coast) generates primarily large earthquakes with a moment magnitude of approximately 7.8 or greater.
- Segment #3 (South Santa Cruz Mountains) generates somewhat smaller earthquakes with a moment magnitude of approximately 6.9 to 7.2.
- Segment #2 (San Francisco Peninsula) does not frequently trigger an earthquake but ruptures due to spill-over from the two adjacent segments.
- The largest earthquakes observed have a moment magnitude of 8.1 to 8.2.

These results imply that the North Coast segment has a large capacity for stress accumulation and therefore generates earthquakes less frequently. Because the slip rate is the same for all three segments, the North Coast earthquakes have the most stress accumulated and hence the largest magnitudes. The earthquakes on the South Santa Cruz Mountains segment have the smallest magnitudes since this segment has the least capacity for stress accumulation.

The San Francisco Peninsula segment does not generate many earthquakes in the model due to its relatively short average interarrival time (138 years) and proximity to a segment with a 84 year interarrival time (South Santa Cruz Mountains). If this were true, it would imply that the segment has a high potential for stress accumulation (because it does not trigger earthquakes) that is rarely realized. The behavior of this segment is critically important given its proximity to the highly populated San Francisco Bay region, and it deserves further study.

7.2.2 MODEL SENSITIVITIES

In order to understand which input data most influence the results, the cell size, slip rate, and interarrival time statistics were varied. Similarly, the Weibull distribution was assumed for the trigger time distribution instead of the lognormal distribution used in

the base case. An alternate segmentation model was also considered as part of these sensitivity studies. The types of results that were compared were the number of earthquakes per year and the probability of an earthquake with a magnitude exceeding a given level during fixed time intervals. The major conclusions are:

- The results are insensitive to cell size except at the highest magnitudes provided that the chosen cell size accurately represents the lengths of the segments.
- The moment magnitude of the largest observed earthquakes is most sensitive to slip rate.
- The results for individual segments are highly sensitive to interarrival times, but the aggregate results for the entire fault are much less sensitive to this parameter.
- The use of the Weibull distribution, rather than the lognormal distribution, affects the results only minimally.
- The results based upon the alternate segmentation were similar to those obtained using the base case.

The lack of sensitivy to the cell size is a desirable property of the model. Since the model is a mathematical construct developed to explain observed behavior, changing the parameters of the model itself should not have a major impact on the results.

The aggregate results for the entire fault are not very sensitive to any of the data varied in the sensitivity study. Since the three segments influence each other, increasing the activity on one segment (by decreasing its mean interarrival time) decreases the activity on the other segments. The results for the entire fault are then rather insensitive to changing the input parameters, except at the largest magnitudes.

The results for each individual segment are much more sensitive than the results of the entire fault. The most important parameters for the results on each segment are the slip rate and the interarrival time statistics. Obviously, changing the mean interarrival time for a segment will affect the results obtained for that segment. It is also logical that the size of the earthquakes simulated is closely related to the assumed slip rate. What is a little more obscure, however, is how changing the behavior of one segment affects the others. For example, if the mean interarrival time of the South Santa Cruz Mountains segment is increased, the San Francisco Peninsula segment begins to trigger more earthquakes because there is not as much rupture spilling over from the other segments.

This model treats the fault as one entity made up of several segments. While modifying the behavior of one of the segments will change the results for each segment, the overall results for the entire fault do not change very much.

7.3 FUTURE WORK

Several questions still remain regarding the model and its application. In addition, the model can be extended to include other earthquake characteristics not considered in this research.

7.3.1 EXTENSION TO SITE HAZARD

As mentioned in Chapter 1, one of the reasons for developing earthquake occurrence models is to enable site-specific hazard to be estimated. This model could easily be extended to do this. An empirical attenuation function or other ground motion propagation model for estimating the hazard at a site is needed. Many attenuation functions require as input the distance from the rupture zone to the site and the size (for example, the magnitude) of the earthquake. In the current model, each time that an earthquake is simulated, these data are known. The space-time model can then be extended to include this capability relatively easily.

For site hazard estimation purposes it is necessary to specify a forecast time period. Thus, the length of each simulation period can be chosen to coincide with the time period for which the hazard estimate is desired. When the model simulates the occurrence of an earthquake, the value of the ground motion parameter at the site can be calculated. The largest value of the ground motion parameter at the site during each simulation should be recorded. After many simulations, it is possible to construct a plot of the probability of exceeding a given value of the ground motion parameter as a function of the parameter's value. This plot is then a measure of the site specific hazard.

While the preceding discussion assumed that the ground motion parameter desired at the site was a single value, it is also possible to use a similar methodology to develop the spectra at the site.

7.3.2 VARIABLE SLIP RATE

Through development of this dissertation, it was assumed that the slip rate for each cell was constant throughout time. As noted in Chapter 1, however, this may not be

the case. Following Suzuki and Kiremidjian (1988), the slip rate for each cell could be simulated from a distribution and then kept constant until the next earthquake, at which time it would again be simulated. It would also be possible to specify the slip rate as a function of time, which would be useful if the slip rate were to vary between the occurrence of two earthquakes. Either of these schemes for considering a variable slip rate could be incorporated into the model. The only thing that would change is the calculation of the amount of time until a cell accumulates enough slip to increment the number of cells that it is capable of rupturing.

7.3.3 VARIABLE CELL SIZES

In this research, the chosen cell size is uniform for all segments. For the northern San Andreas fault, the North Coast segment is much longer than the other two segments. Using a uniform cell size requires the long segment to be composed of many cells. In order to reduce computation time, it would be desirable to vary the cell size from segment to segment. Short segments generating earthquakes more often could utilize a smaller cell size than longer segments that generate earthquakes less often.

In order for this model to use variable cell sizes, the way in which the state of each cell characterizes the amount of slip accumulated must be modified. Currently, the state of the cell tells how many cells would rupture if an earthquake were to occur. If non-uniform cell sizes were used, the state of the cell would need to give information about the length of rupture possible in units of kilometers, rather than in units of cells. This is a straightforward modification that would allow the determination of the cells that rupture even if they have differing lengths.

7.3.4 TWO-DIMENSIONAL MODEL

One assumption made in this model is that the entire depth of the fault ruptures in each earthquake. This can be an unrealistic assumption, as demonstrated by the lack of surface faulting in the 1989 Loma Prieta earthquake (Plafker and Galloway, 1989). In order to model the situation in which only part of the fault's depth ruptures during an earthquake, a two-dimensional model is needed. In such a model, not only would the length of the fault be discretized into cells, but so would the depth. The place on the fault at which the rupture begins would be specified by its distance along the fault and by its depth below the surface. Rules for rupture would have to be developed that determine

when rupture propagates toward the fault's surface and away from the fault's surface as well as determining when rupture propagates along the length of the fault.

Such a two-dimensional model would be much more computationally intensive than the one-dimensional model developed here. In addition, it may be difficult to develop rules for rupture owing to a lack of understanding about how rupture propagates. Nevertheless, a two-dimensional model has the potential to more realistically represent the rupture zones of the earthquakes.

7.3.5 APPLICATION TO OTHER FAULTS

The fault behavior model is developed to apply to faults that generate earthquakes displaying temporal and spatial dependence. In order to test the model's general applicability, it should be applied to faults other than the northern San Andreas. The general criteria for choosing faults to which to apply the model is that the fault generate temporally and spatially dependent earthquakes. In addition, the data described in Chapter 4 must be available.

One candidate for model application is the Alaska-Aleutian island arc, which marks the subduction of the Pacific Plate under the North America Plate. Temporal and spatial patterns have been noted for great (magnitude 7.4 and above) earthquakes occuring in subduction zone (Sykes, et. al., 1981). Since rupture on subduction zones is, by nature, at depth and does not usually cause surface rupture, the two-dimensional model discussed above would be particularly applicable to this area.

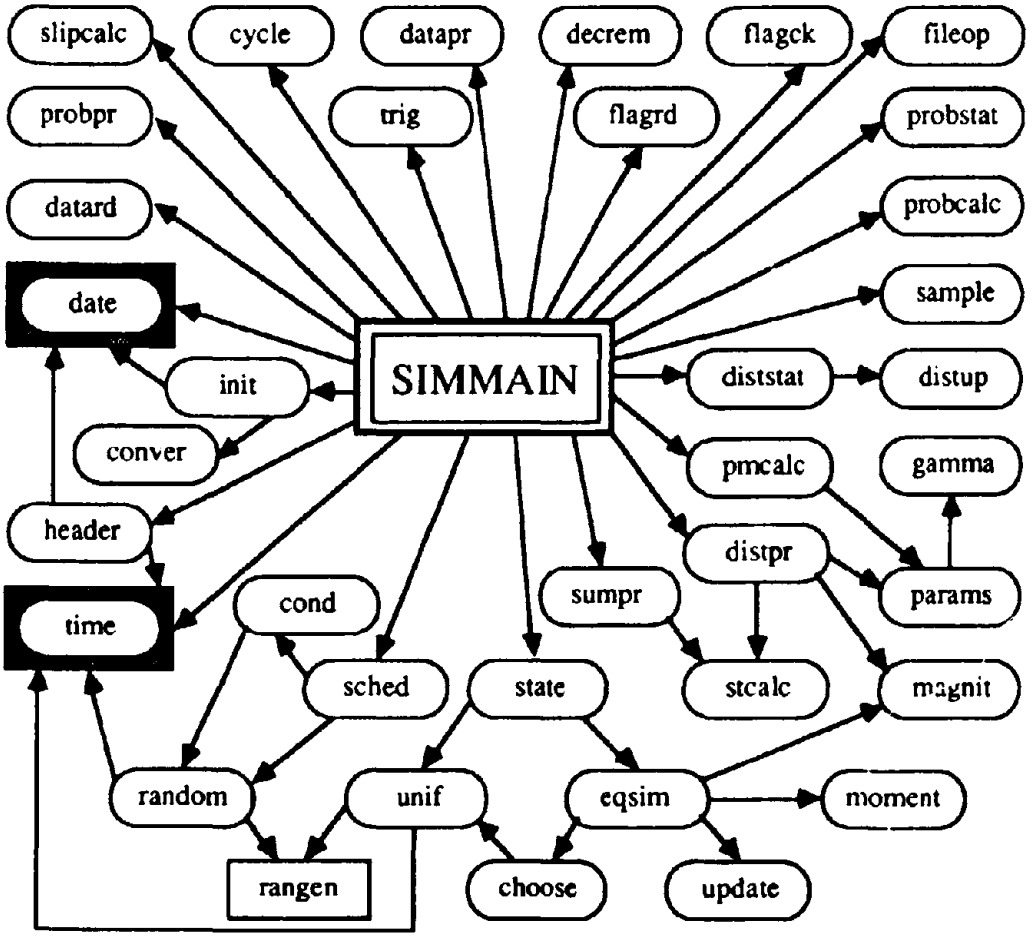
7.3.6 INTERACTION BETWEEN FAULTS

The northern San Andreas fault is not the only fault that ruptures along the boundary between the Pacific plate and the North American plate. Other faults in the San Andreas fault system include the Hayward and Calaveras faults. It is possible that a rupture on any of the three faults could affect the state of stress on the other two. If this is the case, then modeling the earthquakes occurring on each fault separately is not sufficient. A more complex model would be needed to model the behavior and interactions of all the faults simultaneously. Unfortunately, data describing the interaction between faults is very limited, making the development of such a model difficult.

APPENDIX A
DIAGRAM OF THE COMPUTER PROGRAM

This chart shows how the various portions of the computer program used to simulate this model relate to each other. An arrow pointing from program module A to program module B indicates that module A calls module B. The comments in the listing of the computer program (Appendix B) explain the function of each program module.

The program module SIMMAIN, located in the center of the diagram, is the main program. The modules whose names are contained in rectangles with rounded corners are subroutines. The module RANGEN, whose name is in a rectangle, is a function. The modules DATE and TIME are FORTRAN-supplied subroutines.



REFERENCES

Anagnos, T. and A. S. Kiremidjian (1984), "Stochastic Time-Predictable Model for Earthquake Occurrences," *Bull. Seis. Soc. Am.* **74**, No. 6, pp. 2593-2611.

Anagnos, T. and A. S. Kiremidjian (1988), "A Review of Earthquake Occurrence Models for Seismic Hazard Analysis," *Prob. Eng. Mech.* **3**, No. 1, pp. 3-11.

Benjamin, J. R. and C. A. Cornell (1970), *Probability, Statistics, and Decision for Civil Engineers*, McGraw-Hill.

Borcherdt, R. D. (1990), "Influence of Local Geology in the San Francisco Bay Region, California, on Ground Motions Generated by the Loma Prieta Earthquake of October 17, 1989," *Proceedings of the International Symposium on Safety of Urban Life and Facilities*, Tokyo, Japan, November 1-2.

Bufe, C. G., Harsh, P. W., and R. O. Burford (1977), "Steady-State Seismic Slip--A Precise Recurrence Model," *Geophys. Res. Lett.* **4**, pp. 91-94.

Cornell, C. A. (1968), "Engineering Seismic Risk Analysis," *Bull.. Seis. Soc. Am.* **58**, No. 5, pp. 1583-1606.

Der Kiureghian, A. and A. S. Ang (1977), "A Fault-Rupture Model for Sesimic Risk Analysis," *Bull. Seis. Soc. Am.* **67**, No. 4, pp. 1173-1194.

Ellsworth, W.L., A. G. Lindh, W. H. Prescott, and D. G. Herd (1981), "The 1906 San Francisco Earthquake and the Seismic Cycle," *Earthquake Prediction: An International Review*, A. G. U., Maurice Ewing Series **4**, pp. 126-140.

Gardner, J. K. and L. Knopoff (1974), "Is the Sequence of Earthquakes in Southern California, with Aftershocks Removed, Poissonian?" *Bull. Seis. Soc. Am.* **64**, No. 5, pp. 1363-1367.

Hanks, T. C. and H. Kanamori (1979), "A Moment Magnitude Scale," *J. Geophys. Res.* **84**, No. B5, pp. 2348-2350.

Idriss, I. M. (1991), "Earthquake Ground Motions at Soft Soil Sites," *Proceedings of the Second International Conference on Recent Advances in Geotechnical Engineering and Soil Dynamics*, St. Louis, Missouri, March 11-15.

Joyner, W. B. and D. M. Boore (1988), "Measurement, Characterization, and Prediction of Strong Ground Motion," *Earthquake Engineering and Soil Dynamics II--Recent Advances in Ground-Motion Evaluation*, Proceedings of the specialty conference on June 27-30, Park City, Utah, pp. 43-102.

Kiremidjian, A. S. and T. Anagnos (1984), "Stochastic Slip-Predictable Model for Earthquake Occurrences," *Bull. Seis. Soc. Am.* **74**, No. 2, pp. 739-755.

Kiremidjian, A. S., King, S., Sugito, M., and H. C. Shah (1991), "Simple Site-Dependent Ground Motion Parameters for the San Francisco Bay Region," Report No. 97, The John A. Blume Earthquake Engineering Center, Department of Civil Engineering, Stanford University.

Li, C. V. and C. Kisslinger (1985), "Stress Transfer and Non-Linear Stress Accumulation at Subduction-Type Plate Boundaries--Applications to the Aleutians," *Earthquake Prediction*, K. Shimazaki and W. Stuart (editors), Birkhauser, Basel, pp. 812-839.

Mogi, K. (1981), "Seismicity in Western Japan and Long-Term Earthquake Forecasting," *Earthquake Prediction: An International Review*, A. G. U., Maurice Ewing Series **4**, pp. 43-51.

Plafker, G. and J. P. Galloway, editors, (1989), *Lessons Learned from the Loma Prieta, California, Earthquake of October 17, 1989*, U. S. Geological Survey 1045.

Price, J. N. (1981), "Brittle Fracture," *Fault and Joint Development in Brittle and Semibrittle Rock*, Chaper 1, Pergamon Press, Oxford.

Ranalli, G. (1987), "Strength, Fracture, and Plasticity," *Rheology of the Earth: Deformation and Flow Processes in Geophysics and Geodynamics*, Chapter 5, Allen & Unwin, Inc., Boston, MA.

Schwartz, D. P. (1988), "Geologic Characterization of Seismic Sources: Moving into the 1990s," *Earthquake Engineering and Soil Dynamics II--Recent Advances in Ground-Motion Evaluation*, Proceedings of the specialty conference on June 27-30, Park City, Utah, pp. 1-42.

Seed, H. B. and I. M. Idriss (1969), "Influence of Soil Conditions on Ground Motions during Earthquakes," *J. Soil Mech. Found. Eng.* **91**, No. SM1, pp. 99-137.

Seed, H. B., Ugas, C., and J. Lysmer (1976), "Site-Dependent Spectra for Earthquake-Resistant Design," *Bull. Seis. Soc. Am.* **66**, No. 1, pp. 221-243.

Shedler, G. S. (1987), *Regeneration and Networks of Queues*, Springer-Verlag, Berlin, Heidelberg, New York.

Shimazaki, K. and T. Nakata (1980), "Time-Predictable Recurrence for Large Earthquakes," *Geophys. Res. Lett.* **7**, pp. 279-282.

Suzuki, S. and A. S. Kiremidjian (1988), "A Stochastic Ground Motion Forecast Model with Geophysical Considerations," The John A. Blume Earthquake Engineering Center Report No. 88, Stanford University.

Sykes, L. R., Kisslinger, J. B., House, L., Davies, J. N., and K. H. Jacob, "Rupture Zones and Repeat Times of Great Earthquakes along the Alaska-Aleutian Arc, 1784-1980," *Earthquake Prediction: An International Review*, A. G. U., Maurice Ewing Series **4**, pp. 73-80.

Sykes, L. R. and R. C. Quittmeyer (1981), "Repeat Times of Great Earthquakes along Simple Plate Boundaries," *Earthquake Prediction: An International Review*, A. G. U., Maurice Ewing Series **4**, pp. 217-247.

Thatcher, W., and M. Lisowski (1987), "Long-Term Seismic Potential of the San Andreas Fault Southeast of San Francisco, California," *J. Geophys. Res.* **92**, No. B6, pp. 4771-4784.

USGS Circular 1045 (1989), "Lessons Learned from the Loma Prieta, California, Earthquake of October 17, 1989."
USGS Circular 1053 (1990), "Probabilities of Large Earthquakes in the San Francisco Bay Region, California."

Wang, S.-C., McNally, K. C., and R. J. Geller (1982), "Seismic Strain Release Along the Middle America Trench, Mexico," *Geophys. Res. Lett.* **9**, No. 3, pp. 182-185.

Wells and Coppersmith (1991), personal communication.

Whitt, W. (1980), "Continuity of Generalized Semi-Markov Processes," *Math. Oper. Res.* **5**, No. 4, pp. 494-501.