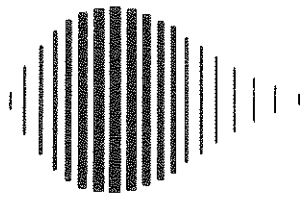


NATIONAL CENTER FOR EARTHQUAKE
ENGINEERING RESEARCH
State University of New York at Buffalo

MODELLING EARTHQUAKE GROUND MOTIONS IN SEISMICALLY ACTIVE REGIONS USING PARAMETRIC TIME SERIES METHODS

by

Glenn W. Ellis and Ahmet S. Cakmak
Princeton University
School of Engineering and Applied Science
Department of Civil Engineering
Princeton, NJ 08544

Technical Report NCEER-87-0014

August 25, 1987

This research was conducted at Princeton University and was partially supported by the National Science Foundation under Grant No. ECE 86-07591.

NOTICE

This report was prepared by Princeton University as a result of research sponsored by the National Center for Engineering Research (NCEER). Neither NCEER, associates of NCEER, its sponsors, Princeton University or any person acting on their behalf:

a. makes any warranty, express or implied, with respect to the use of any information, apparatus, method, or process disclosed in this report or that such use may not infringe upon privately owned rights; or

b. assumes any liabilities of whatsoever kind with respect to the use of, or the damage resulting from the use of, any information, apparatus, method or process disclosed in this report.

MODELLING EARTHQUAKE GROUND MOTIONS IN SEISMICALLY ACTIVE REGIONS USING PARAMETRIC TIME SERIES METHODS

by

Glenn W. Ellis[1] and Ahmet S. Cakmak[2]

August 25, 1987

Technical Report NCEER-87-0014

NCEER Contract Number 86-2011

NSF Master Contract Number ECE 86-07591

1 Graduate Student, Department of Civil Engineering, Princeton University
2 Professor, Department of Civil Engineering, Princeton University

NATIONAL CENTER FOR EARTHQUAKE ENGINEERING RESEARCH
State University of New York at Buffalo
Red Jacket Quadrangle, Buffalo, New York 14261

ABSTRACT

The application of the Autoregressive-Moving Average (ARMA) process is an efficient method to model strong motion accelerograms, after processing by a variance and frequency stabilizing transformation. This report presents two methods for modelling accelerograms.

The first method was developed to model individual accelerogram components. From this modelling procedure, parameters describing the change in variance during the record, the change in the dominant frequency during the record, and the correlation structure of the stabilized series were estimated. This univariate procedure was used to calculate modelling parameters for 148 accelerogram components recorded in California. These parameters were related to physical variables, such as earthquake magnitude, epicentral distance, and site geology, allowing simulations to be generated for sites where no ground motion records are available.

A second procedure was developed to model the three acceleration components together as a group. This multivariate procedure was used to calculate modelling parameters from accelerograms recorded in Mexico and Taiwan, with particular emphasis placed on the accelerograms recorded from the 1985 Michoacan earthquake. By relating the modelling parameters to physical parameters, it is possible to generate realistic three-dimensional simulations for sites in these regions.

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

1.1. The need for strong motion accelerograms

The damage incurred by structures during an earthquake depends upon both the nature of the structure and the properties of the ground motion. However, relatively few sites have recordings of strong ground motion. To provide input motions to structural models for sites for which no strong ground motion data exist, it is necessary to simulate accelerograms. These simulations must have realistic duration, frequency content, and intensity for the physical conditions of the site. Also, because the accelerograms which have been recorded show different behavior under similar conditions, it is useful to measure the variability in ground motion. This makes it possible to generate many simulations with varying properties within the range of the expected ground motion.

Many procedures have been developed to simulate the frequency spectra of earthquake ground motion. These simulations are useful in predicting the response of linear structural models, but do not work well to predict nonlinear response. Artificial accelerograms can be used as input motion for linear and nonlinear structural models. Also, from the accelerogram the Fourier spectrum, response spectrum, velocity time history, and displacement time history may be computed. Thus it is appropriate to simulate the accelerogram, since each type of input motion may be derived from it.

1.2. Background

Because seismic waves are initiated by irregular faulting and then travel through complex ground formations with random properties, resulting in many reflections, refractions, and attenuations before reaching the recording station, a stochastic approach has been taken to model the accelerograms. The application of Autoregressive Moving Average (ARMA) processes to model earthquake accelerograms has proven effective in numerous studies (Jennings et al., 1968; Chang et al., 1979; Kozin, 1977). The difficulty in modelling accelerograms using ARMA models is the nonstationarity of the variance and frequency content of the records. Three basic approaches have been taken to handle this nonstationarity: (1) fitting time-varying ARMA parameters to the

original records (Kozin, 1977; Jurkevics and Ulrych, 1979; Gersch and Kitagawa, 1985), (2) fitting time-invariant ARMA parameters to a series stabilized by transformations (Polhemus and Cakmak, 1981; Cakmak et al., 1985; Ellis et al., 1987), and (3) fitting time-invariant ARMA parameters to short sections of the original records (Chang et al., 1979).

In this report, two modelling procedures which have been developed to fit time-invariant ARMA parameters to a series stabilized by transformations are presented. The first method is used to model individual accelerogram components independently (univariate), while the second method models the three components simultaneously (multivariate). Although for some purposes repeated realizations of a given record are useful, there is greater need to relate the parameters estimated from recorded accelerograms to the physical parameters affecting the ground motion. Thus the modelling procedures were applied to the large databases of free-field strong motion accelerograms recorded in California, Mexico, and Taiwan. By relating the modelling parameters generated from the two procedures to physical variables such as earthquake magnitude, epicentral distance, and geological site conditions, it is possible to generate simulations for any size earthquake at any given site in the regions studied.

1.3. Organization of report

The report can be divided into two parts: univariate modelling in Chapters 2 and 3 and multivariate modelling in Chapters 4, 5, and 6. In Chapter 2 the procedure for modelling one component of an accelerogram is described. The application of the univariate modelling procedure to accelerograms recorded in California is then presented in Chapter 3.

The multivariate modelling procedure is presented in Chapter 4. Chapters 5 and 6 show the results of its application to accelerograms recorded in Mexico, particularly the 1985 Michoacan earthquake, and also to data from the SMART-1 array in Taiwan.

References

Cakmak, A. S., R. I. Sherif, and G. W. Ellis (1985). "Modelling Earthquake Ground Motions in California Using Parametric Time Series Methods," *Soil Dynamics*

and Earthquake Engineering, Vol. 4, No. 3. 124-131.

Chang, M. K., J. W. Kwiatkowski, R. F. Nau, R. M. Oliver and K. S. Pister (1982). "ARMA Models for Earthquake Ground Motions," *Earthquake Engineering and Structural Dynamics*, Vol. 10, 651-662.

Ellis, G. W., A. S. Cakmak and J. Ledolter (1987). "Modelling Earthquake Ground Motions in Seismically Active Regions Using Parametric Time Series Methods," *Third International Conference on Soil Dynamics and Earthquake Engineering*, 551-566.

Gersch, W. and G. Kitagawa (1985). "A Time Varying AR Coefficient Model for Modelling and Simulating Earthquake Ground Motion," *Earthquake Engineering and Structural Dynamics*, Vol. 13, 243-254.

Jennings, P. C., G. W. Housner, and N. C. Tsai (1968). "Simulated Earthquake Ground Motion," *California Institute of Technology Technical Report*, Pasadena, CA.

Jurkevics, A. and T. J. Ulrych (1979). "Autoregressive Parameters for a Suite of Strong Motion Accelerograms," *Bulletin of the Seismological Society of America*, Vol. 69, 2025-2036.

Kozin, F. (1977). "Estimation and Modelling of Non-stationary Time Series," *Proceedings of the Symposium on Computer Methods in Engineering*, University of Southern California.

Polhemus, N. W. and A. S. Cakmak (1981). "Simulation of Earthquake Ground Motions Using Autoregressive Moving Average (ARMA) Models," *Earthquake Engineering and Structural Dynamics*, Vol. 9, 343-354.

2. UNIVARIATE MODELLING PROCEDURE

2.1. Introduction to ARMA Models

In the processing of accelerograms the records are digitized at uniformly spaced time intervals, normally 0.02 seconds in length. Such a sequence of n observations (Z_t , t = 1, 2, ... , n) is called a discrete time series. If the time series is regarded as a realization from a stochastic process, then it is possible to generate many realizations from the same process having the same probabilistic structure as the original record.

Many stochastic processes can be approximated by autoregressive (AR) models. In the autoregressive model the current deviation of the process from its mean value μ, $(Z_t - \mu)$, is expressed as a function of previous deviations $(Z_{t-1} - \mu), (Z_{t-2} - \mu), ..., (Z_{t-p} - \mu)$, and a shock a_t by

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \phi_2(Z_{t-2} - \mu) + \cdots + \phi_p(Z_{t-p} - \mu) + a_t \qquad (2.1.1)$$

where

p = order of the AR model

ϕ_k = autoregressive parameter at lag k, k=1,2,...,p

μ = mean level

a_t = white noise sequence with variance σ_a^2

Because accelerograms have a zero mean, μ is assumed to be zero. The autoregressive parameters and white noise variance are estimated from the data.

Another important model for stochastic processes is the moving average (MA) model. In the moving average model, the current value of the process Z_t is related to the past value of the shocks a_t by

$$Z_t - \mu = -\theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} + a_t \qquad (2.1.2)$$

where

q = order of the moving average process

θ_k = moving average parameter for lag k, k=1,2,...,q

Again, μ is assumed to be zero and the moving average parameters and the white noise variance are estimated from the data.

By including both autoregressive and moving average terms one obtains a mixed autoregressive-moving average (ARMA) model of order (p,q):

(2.1.3)

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \phi_2(Z_{t-2} - \mu) + \cdots + \phi_p(Z_{t-p} - \mu) + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

ARMA models allow greater flexibility in fitting time series than either the AR or MA models alone. The procedures for identifying, fitting, and validating ARMA models are discussed in detail in Box and Jenkins (1976).

2.2. Nonstationarity of Accelerograms

Due to the nonstationarity of accelerograms, it is not appropriate to fit an ARMA model directly to the time series. As can be seen in Fig. 2.1b, this nonstationarity manifests itself most conspicuously in the large changes of variance over time. However, in most records examined in California the frequency content is also variable. Typically, the predominant frequency decreases with time.

Several approaches exist for handling the nonstationarity of the time series. One approach is to divide the time series into short segments, each segment short enough to be considered stationary. In one study five-second segments were used (Chang et al., 1982). It was found that ARMA models of order (2,1) or (4,1) usually fit the data well. Although the ARMA parameters changed for each segment, the form of the model tended to remain constant.

Another approach is to allow the ARMA parameters to vary with time. Such an approach was taking by Kozin (1977) who used an autoregressive model of the form

$$Z_t = \phi_1(t)Z_{t-1} + \phi_2(t)Z_{t-2} + \cdots + \phi_p(t)Z_{t-p} + g(t)a_t \qquad (2.2.1)$$

where g(t) is estimated by fitting an envelope to the observed series using a cubic

spline technique. The series is then transformed to make the variance constant, and the time-varying AR parameters are modeled as a sum of discrete Legendre polynomials. More recently, Jurkevics and Ulrych (1979) and Gersch and Kitagawa (1985) have also modeled accelerograms using an autoregressive model with time varying parameters. Jurkevics and Ulrych fit an AR (2) model to the accelerogram in a time-adaptive manner resulting in two time-dependent AR functions and a time-dependent innovation variance. Gersch and Kitagawa used a smoothness priors-time varying AR coefficient model to model accelerograms and allowed the model order to change with time. Their model also resulted in time-dependent AR parameters and an innovation variance function. However, in all of these methods it is difficult to relate the AR functions to the physical variables affecting ground motion. This is because these methods produce time-varying functions. A modelling procedure producing single-valued parameters to describe the ground motion is easier to relate to physical variables.

2.3. Univariate Modelling Procedure

Because one of the primary goals of this study is to relate modelling parameters to physical variables affecting the ground motion, a modelling procedure using single-valued parameters based on the methodology in Polhemus and Cakmak (1981) and Cakmak et al. (1985) was used. Several modifications, however, were made to the stabilization procedure used in these studies, including: a frequency envelope allowing the frequency content to change during the duration of a simulation; a constrained ARMA model allowing negligible response at zero frequency; and physical principles relating the modelling parameters to physical variables. The modified procedure was used in Ellis et al. (1987).

The first step in the modelling procedure is to determine the duration of the earthquake, T, by eliminating the first 1% and final 2% of the cumulative energy of the record. This is shown in Fig. 2.1 for a horizontal component of the Kern County earthquake (7/21/52) recorded at the Taft Lincoln School (USGS No. 95). The shortened accelerogram is then divided by a standard deviation envelope, thus stabilizing the variance of the series to 1.0 as shown in Fig. 2.2. However, most accelerograms still exhibit nonstationary frequency content. The frequency content is stabilized by multiplying the time increment by a frequency envelope. This results in a time series

that is stationary in both variance and frequency content as shown in Fig. 2.3. An ARMA model can now be fitted to the stabilized series.

From the transformation and modelling procedure, parameters are generated to describe the standard deviation envelope, frequency envelope, and the ARMA model for each record analyzed. These parameters were related to physical variables such as earthquake magnitude, epicentral distance, geographic location, and soil type. Also, relationships among the parameters were found, making it possible to reduce the complexity of the model.

Using the physical relationships developed, it is possible to reverse the procedure to generate simulations for any set of physical variables. First the modelling parameters are calculated from the physical variables. From these modelling parameters a time series, standard deviation envelope, and frequency envelope are generated. By transforming the time scale of the time series and multiplying by the standard deviation envelope, a realistic simulation can be generated possessing similar characteristics to an actual accelerogram.

The validity of the results was assessed by comparing the acceleration time histories, response spectra, and Fourier spectra of the original earthquake and several simulations.

The procedure can be summarized by the following steps.

1. Shorten the accelerogram
2. Calculate a standard deviation envelope
3. Stabilize the variance
4. Calculate the frequency envelope
5. Stabilize the frequency content
6. Fit an ARMA model to the stabilized series
7. Relate the ARMA and transform parameters to physical variables
8. Generate simulations from parameters fitted to the original series

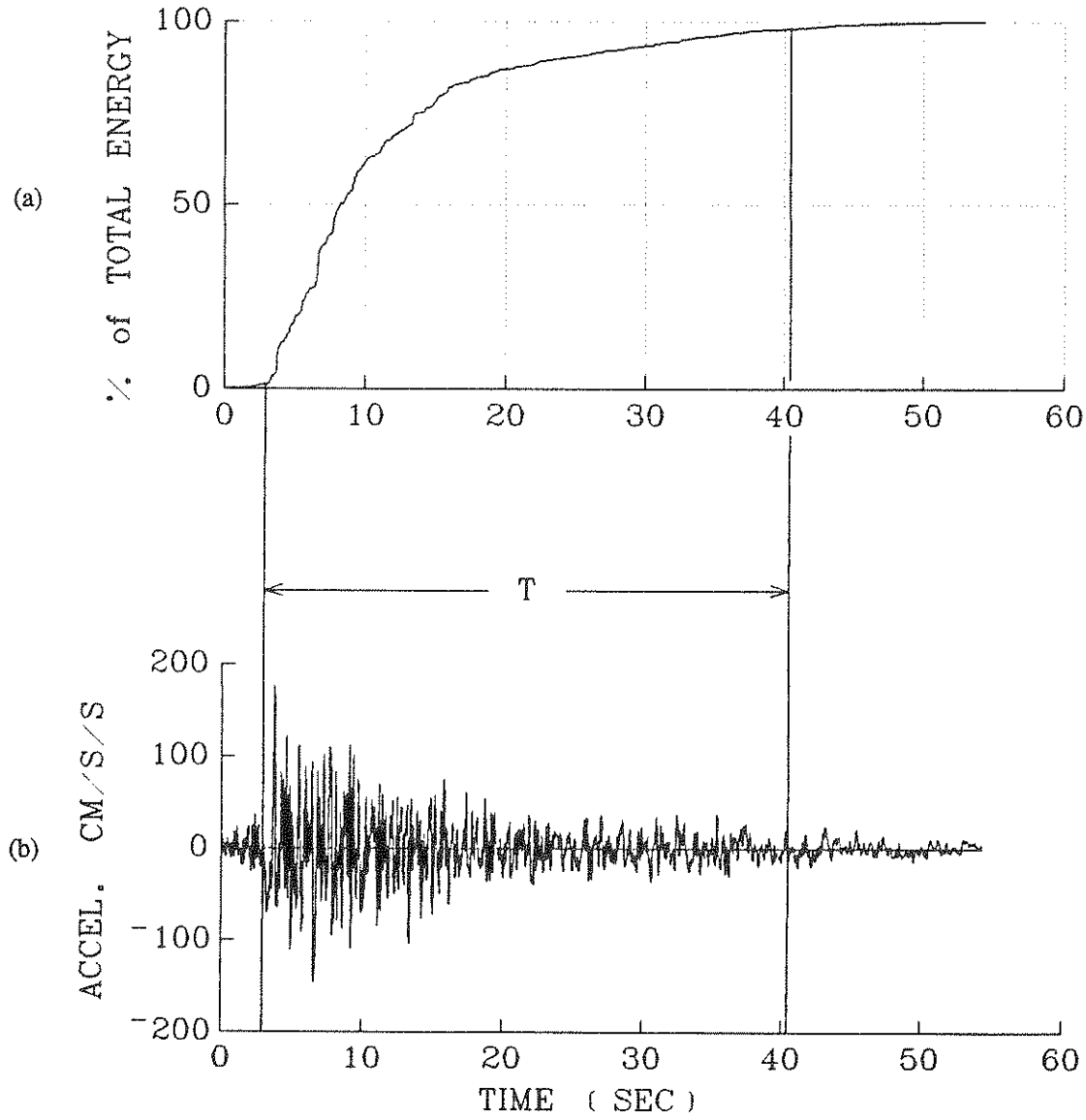


Fig. 2.1 Shortening the accelerogram.

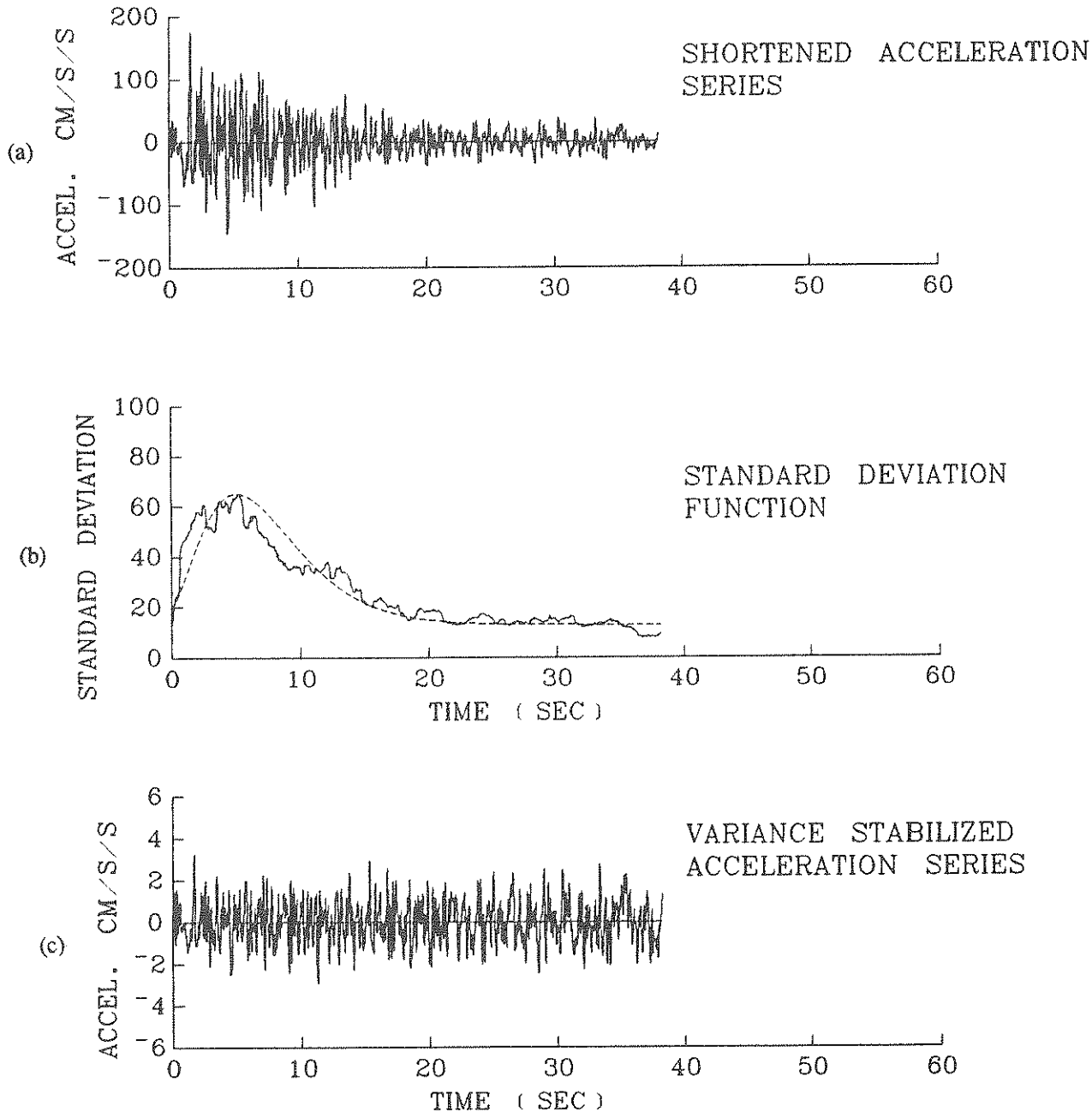


Fig. 2.2 Stabilizing the variance.

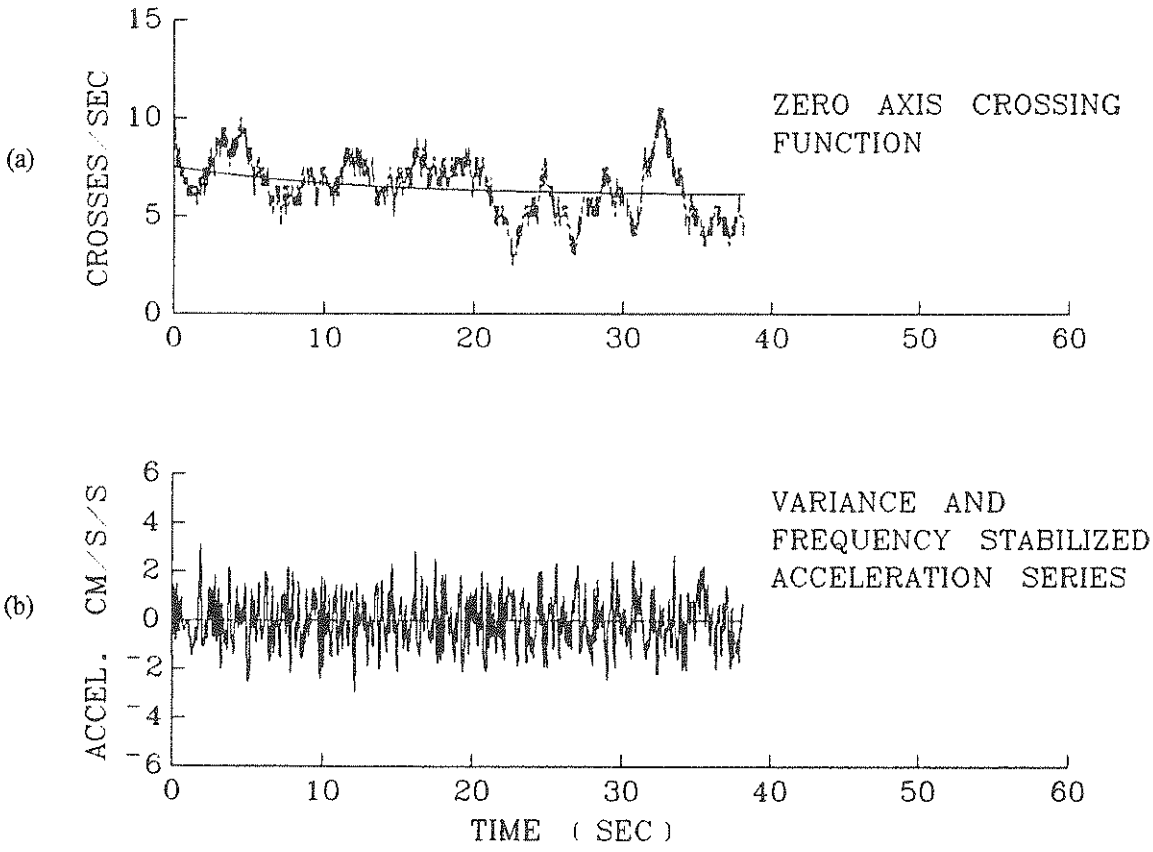


Fig. 2.3 Stabilizing the frequency of zero crossings.

and from parameters calculated as a function of physical variables.

9. Compare the time histories and frequency spectra of the original and simulated series to validate the procedure.

2.4. Shortening the Accelerogram

As shown in Fig. 2.1b the accelerogram normally exhibits weak shaking in the beginning of the record and at the end. To determine the duration of the earthquake, T, several different methods may be used to judge the amount of the record to be eliminated. For example, a threshold acceleration in which all of the record before the first occurrence and after the last occurrence of the threshold acceleration is eliminated may be used. However, to select one threshold acceleration which would eliminate enough of the acceleration for strong records but not too much for weak records would be difficult. Secondly, since extreme peaks in acceleration can occur at random times in the weak part of the shaking, the shortening of the record will vary for accelerograms that are different realizations of the same underlying statistical process.

Another approach is to use the cumulative energy of the accelerogram, I_0, measured by

$$I_0 = \int_{t=0}^{t=T} Z^2(t)\, dt \tag{2.4.1}$$

to shorten the record. For the accelerograms recorded at discrete time intervals this relation is

$$I_0 = \sum Z_t^2 \Delta t \tag{2.4.2}$$

By dividing the amount of energy the accelerogram has recorded at any time, t, by the total energy of the accelerogram, the cumulative energy function may be plotted as shown in Fig. 2.1a. Because this is a normalized quantity, the function can be used for both weak and strong acceleration records. Also, since the accelerogram has been integrated, the effect of extreme peaks is smoothed out and thus the underlying statistical process is being measured. By examining numerous earthquakes in California it was found that eliminating the first 1% and final 2% of the energy isolated the strong

shaking well for most records.

2.5. Estimation of the Standard Deviation Envelope

The variance of a random variable Z_t is given by

$$\sigma_z^2(t) = E[(Z_t - \mu_t)^2] \tag{2.5.1}$$

For acceleration time histories with a mean of zero, this relation reduces to

$$\sigma_z^2(t) = E[Z_t^2] \tag{2.5.2}$$

Thus by squaring the acceleration and calculating its running average, an estimate of the variance of the series at any given time may be found. Using an equally weighted two-second time window, the variance envelope is calculated by

$$\sigma_z^2(t) = \frac{1}{101} \sum_{i=t-50}^{t+50} Z_i^2 \tag{2.5.3}$$

where increments of t are 0.02 seconds in length. By using various window sizes to estimate the variance envelope for a number of accelerograms, the two-second window was found to be the smallest window size to estimate a smooth variance envelope. The square root of the variance envelope provides an estimate of the standard deviation envelope. By dividing the shortened accelerogram by the standard deviation envelope, a time history with a stationary variance of approximately 1.0 is obtained.

To relate the standard deviation envelope to physical variables, a smooth function, $\hat{\sigma}_z$, used by Polhemus and Cakmak (1981) of the form

$$\hat{\sigma}_z(t) = c_1(\alpha - k_1)\left(\frac{t}{\tau}\right)^3 e^{-(c_2/\tau)t} + k_1 \tag{2.5.4}$$

where

$$c_1 = \frac{8e^3}{3\sqrt{3}}$$

$$c_2 = 2\sqrt{3}$$

is fitted as shown in Fig. 2.2.b. The maximum intensity of the strong shaking is measured by α. The standard deviation of the weak shaking, k_1, is estimated as the average of the standard deviation envelope during the final 1/3 of the record. Finally, the duration of strong shaking, τ, is estimated such that the energy as measured by the

standard deviation function is equal to the energy of the function fitted to it in Eq. (2.5.4)

2.6. Estimation of the Frequency Function

Because the arrival of high frequency P-waves precedes the arrival of S-waves and surface waves, the variance stabilized series has a nonstationary frequency content. Most records have initially high predominant frequencies which quickly decrease with time. One measure of this phenomenon is the number of zero axis crossing per second, $F_c(t)$, as shown in Fig. 2.3a. This envelope was calculated using an equally weighted two-second window as

$$F_c(t) = \frac{\textit{no. of zero axis crossings between } t \pm 1 \textit{ second}}{2 \textit{ seconds}} \tag{2.6.1}$$

In Fig. 2.3.a, a smooth function $F'_c(t)$, is fitted to the zero axis crossings.

$$F'_c(t) = c_0 e^{-bt} + k_2 \tag{2.6.2}$$

where

c_0 = The initial value of the zero axis crossings

b = The rate of decay

and k_2 = The zero axis crossings of the weak shaking.

$F'_c(t)$ is fitted by first assigning k_2 as the average value of the zero axis crossings during the final 1/3 of the record. A least squares fit may then be calculated by

$$\ln[F_c(t) - k_2] = a - bt \tag{2.6.3}$$

where a and b are the regression coefficients. The value of c_0 is then

$$c_0 = k_2 + e^a \tag{2.6.4}$$

The frequency function, $F'_c(t)$, is then used to change the time increments by

$$\Delta t'(t) = (\Delta t) F'_c(t) \tag{2.6.5}$$

where

Δt = The original time increment of 0.02 seconds

$\Delta t'(t)$ = The new time increment at time t

The transformed record is reduced to the same duration as the initial record by

$$\Delta t'' = \Delta t' \times \frac{duration\ of\ the\ original\ record}{duration\ of\ the\ transformed\ record} \tag{2.6.6}$$

and digitized to 0.02 second intervals using linear interpolation. Physically, Eqs. (2.6.5) and (2.6.6) expand the time increment in the beginning of the record, where the higher frequencies occur, and decrease the time increment at the end of the record, which is dominated by the lower frequencies. This results in a time history with a constant predominant frequency.

2.7. Constrained ARMA Model

With the variance and frequency content stabilized, it is now possible to fit an ARMA model to the series. To choose the appropriate order model, a group of about 20 California records was studied. Using the principles outlined by Box and Jenkins (1976), ARMA processes of order (2,1), (3,1), and (4,1) were fit to the stabilized acceleration time series. It was found that an ARMA (2,1) model fit most of the data well. The goodness of fit of a model is found by examining the residuals of the fit. If the residuals are correlated, a higher order model is necessary. Although higher order models of order (3,1) and (4,1) were found to reduce the correlation among the residual series in some cases, the reduction in the variance of the residual series was small. The slight improvement in the fit of a higher order model did not justify the the added complexity and the greater difficulty in relating the extra coefficients to physical variables.

When generating simulations with the ARMA (2,1) model it was found that the Fourier spectra and response spectra of the simulations had a consistently higher value at low frequencies than the original records. Because the original accelerograms were filtered to allow no response at zero frequency, a model was sought which had no response at zero frequency and fit the data well.

A constrained ARMA (2,2) model was found to meet these requirements. By setting the equation for the Fourier spectrum of an ARMA (2,2) model

$$F(f) = \left\{ 2\sigma_a{}^2 \frac{1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1-\theta_2)cos(2\pi f) - 2\theta_2 cos(4\pi f)}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1-\phi_2)cos(2\pi f) - 2\phi_2 cos(4\pi f)} \right\}^{1/2} \qquad (2.7.1)$$

σ_a^2 = white noise variance

equal to zero at zero frequency ($f = 0$), the constraint $\theta_1 + \theta_2 = 1$ was found. This constraint violated the invertibility requirement

$$\theta_1 + \theta_2 < 1 \qquad (2.7.2)$$

for the ARMA (2,2) model. It was found that the constraint

$$\theta_1 + \theta_2 = 0.99 \qquad (2.7.3)$$

allowed negligible response at zero frequency, while also satisfying Eq. (2.7.2).

To fit the constrained ARMA model $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$ to the data, a nonlinear constrained optimization program written by Quandt and Goldfeld (1985) was utilized. The parameters ϕ_1, ϕ_2, θ_1, and θ_2 were estimated to minimize the variance of the residuals, σ_a^2, while also satisfying the imposed constraint and the invertibility and stability requirements of an ARMA (2,2) model.

Of the five parameters estimated, only three can be chosen independently: ϕ_1, ϕ_2, and θ_1. Equation (2.7.3) relates θ_1 and θ_2. of the time series Z_t. Because the variance of the time series was stabilized to be 1.0, σ_a^2 can be expressed as a function of ϕ_1, ϕ_2, and θ_1 as derived in Appendix A.1.

$$\sigma_a^2 = \frac{(1-\phi_2)(1-\phi_2^2) - \phi_1^2(1+\phi_2)}{(1-\phi_2)c_1 + (1+\phi_2)\phi_1 c_2 + (1-\phi_2)\phi_2 c_3}\sigma_z^2 \qquad (2.7.4)$$

where

$$c_1 = 1 - \theta_1(\phi_1 - \theta_1) - \theta_2[\phi_1(\phi_1 - \theta_1) + \phi_2 - \theta_2]$$

$$c_2 = -\theta_1 - \theta_2(\phi_1 - \theta_1)$$

$$c_3 = -\theta_2$$

Because the variance of the time series was stabilized to be 1.0, Eq. (2.7.4) expresses

σ_a^2 as a function of ϕ_1, ϕ_2, and θ_1.

2.8. Generating Simulations

ARMA modelling has a distinct advantage over other schemes in the ease with which simulations can be generated by reversing the modelling process. The modelling parameters needed to generate the simulations may be found by one of two methods. In the first method, parameters estimated from the modelling procedure may be used. Simulations generated from these parameters will be multiple realizations of the original accelerogram. In the second method, the modelling parameters for a given site may be calculated from equations relating the modelling parameters to physical variables. These equations are presented in Chapter 3.

Given all the model parameters, it is possible to reverse the modelling procedure to create simulations. First, a series with stable variance and frequency content is generated at 0.02 second increments by

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} - \theta_1 a_{t-1} - \theta_2 a_{t-2} + a_t \,. \tag{2.8.1}$$

A zero crossing frequency envelope, $F'_c(t)$, is then calculated from Eq. (2.6.2) to rescale the time axis of Z_t as

$$\Delta t'_s(t) = \frac{0.02}{F'_c(t)} \,. \tag{2.8.2}$$

After changing the time scale, it is necessary to reduce the record to the original duration by

$$\Delta t''_s(t) = \Delta t'_s(t) \times \frac{\textit{duration of the original simulation}}{\textit{duration of the transformed simulation}} \,.$$

The series is then digitized into equal increments of 0.02 seconds by linear interpolation. From Eq. (2.5.4) a standard deviation envelope, $\hat{\sigma}(t)$, is calculated. Multiplying the time series by this envelope, a simulation nonstationary in both variance and frequency content is created.

References

Box, G. E. P. and G. M. Jenkins (1976). *Time Series Analysis Forecasting and*

Control, Holden-Day, San Francisco.

Cakmak, A. S., R. I. Sherif, and G. W. Ellis (1985). "Modelling Earthquake Ground Motions in California Using Parametric Time Series Methods," *Soil Dynamics and Earthquake Engineering,* Vol. 4, No. 3. 124-131.

Chang, M. K., J. W. Kwiatkowski, R. F. Nau, R. M. Oliver and K. S. Pister (1982). "ARMA Models for Earthquake Ground Motions," *Earthquake Engineering and Structural Dynamics,* Vol. 10, 651-662.

Ellis, G. W., A. S. Cakmak and J. Ledolter (1987). "Modelling Earthquake Ground Motions in Seismically Active Regions Using Parametric Time Series Methods," *Third International Conference on Soil Dynamics and Earthquake Engineering,* 551-566.

Gersch, W. and G. Kitagawa (1985). "A Time Varying AR Coefficient Model for Modelling and Simulating Earthquake Ground Motion," *Earthquake Engineering and Structural Dynamics,* Vol. 13, 243-254.

Jurkevics, A. and T. J. Ulrych (1979). "Autoregressive Parameters for a Suite of Strong Motion Accelerograms," *Bulletin of the Seismological Society of America,* Vol. 69, 2025-2036.

Kozin, F. (1977). "Estimation and Modelling of Non-stationary Time Series," *Proceedings of the Symposium on Computer Methods in Engineering,* University of Southern California.

Polhemus, N. W. and A. S. Cakmak (1981). "Simulation of Earthquake Ground Motions Using Autoregressive Moving Average (ARMA) Models," *Earthquake Engineering and Structural Dynamics,* Vol. 9, 343-354.

Quandt, R. E. and S. M. Goldfeld (1985). *GQOPT4/I Reference Manual,* Princeton University, Princeton, NJ.

3. CALIFORNIA STUDY

3.1. Introduction

The San Andreas fault system forms the boundary between the North Pacific plate to the West and the North American plate to the East. The North Pacific plate is moving northwest relative to the North American plate, forming a transform fault between the two plates. Because of the large population centers located in this seismically active region, estimating the ground motion expected from a major earthquake is an important goal to increase the safety of the structures located in this region.

The database of strong motion accelerograms recorded in California contains more records than any other region in the world. The univariate modelling procedure was applied to nine seismic regions shown in Fig. 3.1. From these regions 99 horizontal and 49 vertical records were examined. The earthquake magnitude for these records varied from 3.0 to 7.7, while the epicentral distance varied from 5 to 125 kilometers. More information on these records can be found in Hudson (1976).

3.2. Relating the Model Parameters to Physical Variables

For each record that was modelled, nine parameters were calculated: α, τ, k_1, c_0, b, k_2, ϕ_1, ϕ_2, and θ_1. The results were first examined for relationships among the model parameters. The parameters which were found to be unrelated to other parameters were then related to physical variables. The estimated functional relationships are shown in Table 3.1.

The properties of the ground motion and therefore of the modeling parameters are affected by the source mechanism, the distance and the properties of the material through which the waves travel, and the geological conditions at the recording site. These properties are described by physical variables such as the magnitude of the earthquake, M, epicentral distance, d, geographical area, and site conditions. To find the form of the functional relationship, the following procedure was followed. First, the parameters were plotted versus functions of physical variables as suggested by physical considerations. These scatterplots were then linearized and a linear

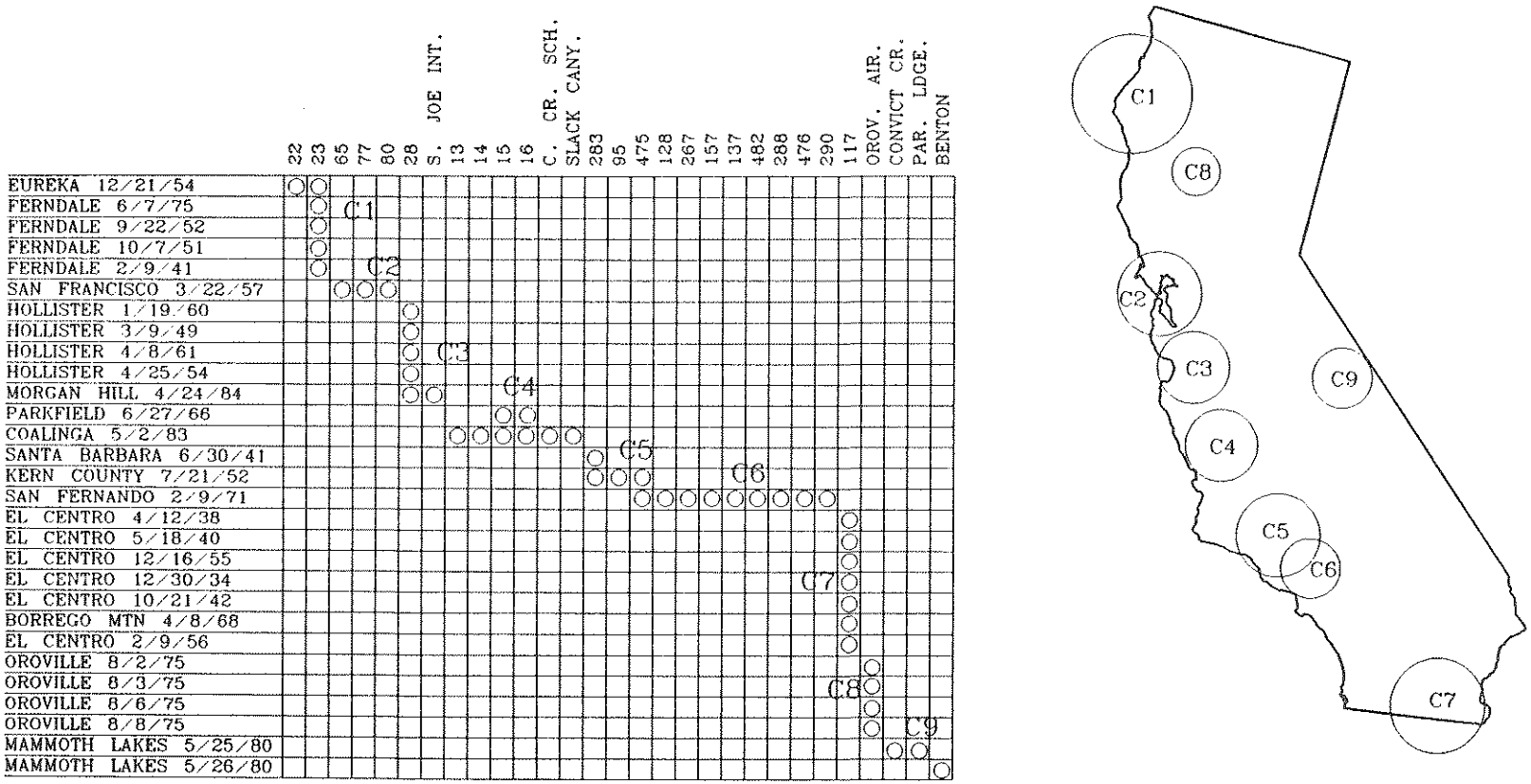


Fig. 3.1 California records analyzed.

Table 3.1 Parametric Relations for California

H = Horizontal Components V = Vertical Components
M = magnitude
d = distance
k_ϕ = location parameter

Log α	$= 0.60 + 0.23 \text{ Log } \frac{10^M}{k_\phi d^2}$	H	$r^2 = 0.38$
	$= 0.52 + 0.19 \text{ Log } \frac{10^M}{k_\phi d^2}$	V	$r^2 = 0.27$
τ	$= 0.55 + 0.73 \sqrt{d}$	H	$r^2 = 0.39$
	$= 1.05 + 0.121 \sqrt{d}$	V	$r^2 = 0.62$
Log k_1	$= -0.60 + 0.85$ Log α	H	$r^2 = 0.59$
	$= -0.54 + 0.80$ Log α	V	$r^2 = 0.56$
$1/c_0$	$= -0.15 + 0.014 \frac{M}{k_\phi^{0.33}}$	H	$r^2 = 0.36$
	$= -0.01 + 0.010 \frac{M}{k_\phi^{0.33}}$	V	$r^2 = 0.18$
b	$= 0.18$	H	
	$= 0.19$	V	
k_2	$= 0.89 + 11.6\ k_\phi$	H	$r^2 = 0.55$
	$= 0.50 + 14.4\ k_\phi$	V	$r^2 = 0.56$
ϕ_1	$= 2 - \frac{k_\phi}{d^{0.32}}$	H	$r^2 = 0.72$
	$= 2 - 0.91 \frac{k_\phi^{1.1}}{d^{0.13}}$	V	$r^2 = 0.55$
ϕ_2	$= 0.80 - 0.91\ \phi_1$	H	$r^2 = 0.97$
	$= 0.99 - 0.97\ \phi_1$	V	$r^2 = 0.99$
θ_1	$= 0.22$	H	
	$= 0.33$	V	

regression model was fitted. The t-ratios of the regression coefficients were calculated to see if the relationship was significant at the 95% confidence level. The coefficient of determination, r^2, was also calculated. In the simple linear regression model, r^2 is the square of the correlation coefficient, expressing the percentage of variation that is explained by the fitted regression relationship. Examining Table 3.1, it can be seen that of the nine parameters estimated for each component, five could be related to physical variables, two were found to be constants, and two were related to other parameters.

3.2.1. ARMA Parameters

To understand the physical meaning of the ARMA parameters, the theoretical Fourier spectrum of the ARMA model was plotted using Eq. (2.6.1). By varying the values of ϕ_1, ϕ_2, and θ_1, the importance of each of these variables was found.

The relative amounts of high and low frequencies were controlled mainly by ϕ_1 for the values calculated in this study. As ϕ_1 approached a theoretical maximum of 2, the relative amount of lower frequencies increased. Because high frequencies damp out more quickly with distance than low frequencies (Trifunac and Brady, 1975), ϕ_1 was modelled as a function of distance. However, due to the physical properties of the soil at the recording station and the geology of the path through which the seismic waves travel, it is expected that local conditions of each area studied will influence the value of ϕ_1 as well. Thus, ϕ_1 was modelled as a function of epicentral distance and also of a site parameter estimated for each group of sites as

$$Log\,(2 - \phi_1) = a_s - b\ Log\ d \qquad (3.2.1.1)$$

where

a_s = constant estimated for each group of sites

b = constant estimated for all California .

By setting $k_\phi = 10^{a_s}$, Eq. (3.2.1.1) can be simplified to

$$\phi_1 = 2 - \frac{k_\phi}{d^b} . \qquad (3.2.1.2)$$

The value of b was estimated to be 0.32 for the horizontal components, and the value of k_ϕ estimated for each site is shown in Fig. 3.2. For the vertical case, ϕ_1 was regressed on distance and k_ϕ calculated from the horizontal component. The results for ϕ_1 are shown in Table 3.1.

In Fig. 3.2, contour lines are drawn for k_ϕ values in order to estimate k_ϕ for areas not included in this study. High k_ϕ values estimated at Oroville and Mammoth Lakes areas indicate more high frequency content in these records compared to accelerograms recorded along the San Andreas fault. An explanation of this tendency is that the medium in which the waves propagate is fractured at the San Andreas fault due to the crushing effect of the relative plate movements, while away from the plate boundary the earth remains more uniform.

The site conditions at many recording stations in California have been classified by Trifunac and Brady (1975) into three types shown in Table 3.2. A regression relationship was estimated between k_ϕ and the site classification.

$$k_\phi = 0.287 + 0.230(site\ classification) \qquad r^2 = 0.26 \tag{3.2.1.3}$$

Equation (3.2.1.3) indicates higher frequency content at stiffer sites.

In examining the other ARMA parameters, it was found that the second autoregressive parameter, ϕ_2, varied linearly with ϕ_1. The value of θ_1 used in Eq. (2.6.1) was found to have virtually no effect upon lower frequencies and only a small effect upon the frequencies between about 15 Hz and 25 Hz. Because of this insensitivity, θ_1 was estimated to be a constant. Thus the model was found to be insensitive to a variation in the value of the moving average parameters, but their existence as constants in the model was still required to reduce the correlation of residuals and to constrain the Fourier amplitude at zero frequency to zero.

3.2.2. Standard Deviation Envelope

Because α and the maximum acceleration are closely related, the functional relationship for α was based upon the definition of local magnitude developed by Richter (1935) as

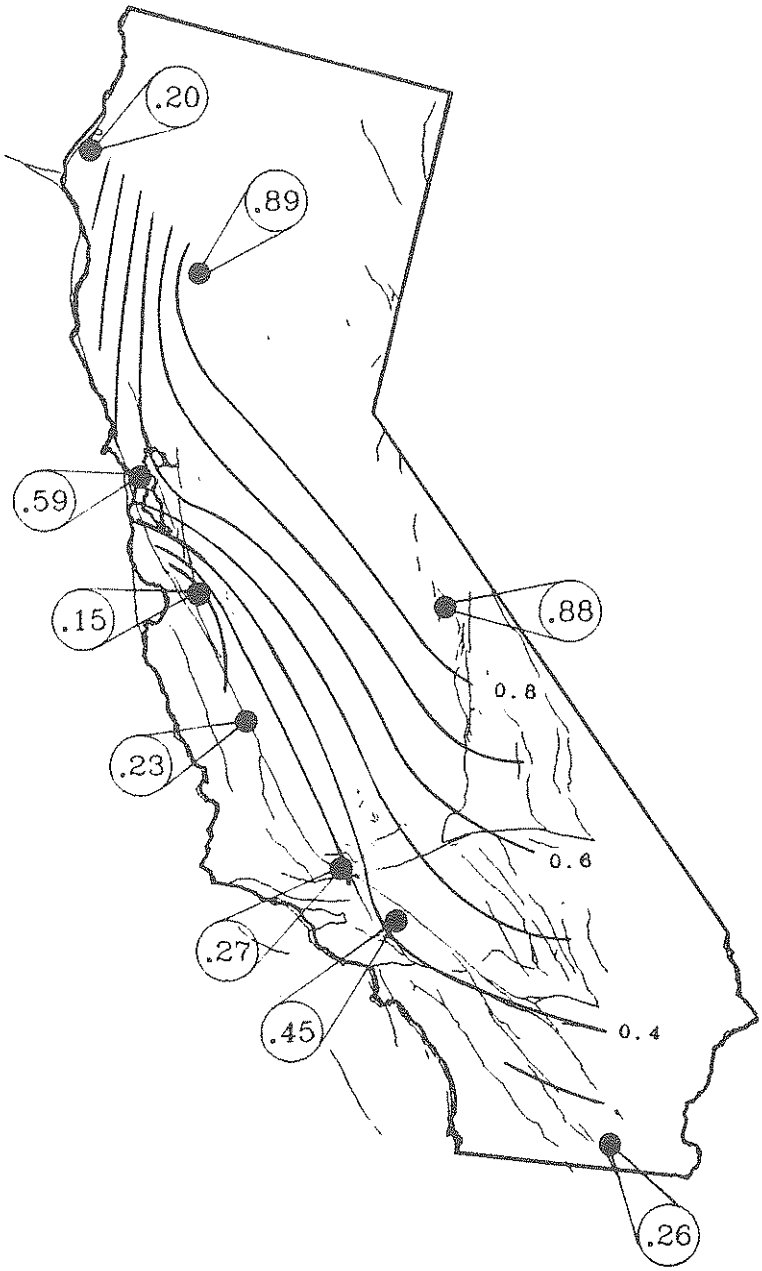


Fig. 3.2 Contour map of k_ϕ values.

Table 3.2 Classification of Site Conditions

Classification	Site Geology
0	soft alluvial deposits
1	hard sedimentary rock or an intermediate site between 0 and 2
2	basement or crystalline rock

Table 3.3 Horizontal-Vertical Relations

$\alpha^V = 0.41\ \alpha^H$	$r^2 = 0.60$
$\tau^V = 1.6 + 0.90\ \tau^H$	$r^2 = 0.58$
$c_o^V = 3.1 + 1.2\ c_o^H$	$r^2 = 0.42$
$b^V \sim b^H$	
$k_2^V = 1.12\ k_2^H$	$r^2 = 0.74$
$\phi_1^V = -0.99 + 1.5\phi_1^H$	$r^2 = 0.73$
$\theta_1^V = 1.5\ \theta_1^H$	

$$M = Log(a) + 3Log(d) - 2.92 \qquad (3.2.2.1)$$

where

a = maximum acceleration

d = epicentral distance .

Thus α was expected to be a function of the ratio of 10^M divided by a power of d. It was found that d^2 best described the data, suggesting body-wave attenuation of α. Also, the site parameter, k_ϕ, was found to be highly correlated to the amplitude of the strong shaking.

Because the velocities of wave propagation for P, S, and Surface waves differ, the time between their arrival and the duration of shaking depends upon the distance the waves travel. Therefore, the length of time of strong shaking, τ, was related to the epicentral distance. Taking the square root of the distance made the relation linear.

3.2.3. Frequency Envelope

Finally, the parameters of the frequency envelope were examined. The initial frequency of zero axis crossings, c_0, was found to be a function of k_ϕ and magnitude. A study by Terashima (1968) relating the spectral peak of P-waves to magnitude suggested the functional form. Because the dynamic receptance of the site affects the response, k_ϕ was also included. The effect of P-waves is significant only during the early shaking; the frequency of zero axis crossings during the weak part of the shaking, k_2, was found to be a function of the site only. The rate of decay, b, was found to be a constant.

3.3. Relationship between Vertical and Horizontal Parameters

Linear relationships between the independent vertical and horizontal parameters are shown in Table 3.3. The amplitude of strong shaking for the vertical case is slightly less than half that of the horizontal case. The duration of strong shaking, τ, was found to be approximately equal for the vertical and horizontal case. Thus the standard deviation envelope for the vertical case varies only in amplitude from the horizontal case.

In examining the frequency envelope, it was found that both the initial zero crossing frequency, c_0, and weak shaking frequency, k_2, were greater for the vertical case. The decay rate, b, was found to be approximately the same. This indicates that the frequency envelope for the vertical case has the same shape as the horizontal case and varies only in magnitude.

Finally, the value of ϕ_1 was found to be lower for the vertical component. This again indicates the higher frequency content in the vertical direction. The greater frequency content in the 15 Hz to 25 Hz range of the vertical component is indicated by a higher value of θ_1.

3.4. Simulations of the Original Records

The modelling parameters needed to generate simulations may be found by one of two methods. First, the parameters estimated by the modelling procedure may be used. Simulations generated from these parameters will be multiple realizations of the original accelerogram (having the same statistical properties). Second, for a given site k_ϕ may be found from Fig. 3.2 or Eq. (3.2.1.3). Using this value along with the earthquake magnitude and epicentral distance for which the simulation is to be generated, the modelling parameters may be found from the parametric relations summarized in Table 3.1.

To assess the validity of the model, simulations were generated from the parameters fitted to the original series (Fig 3.3) and from parameters calculated as a function of physical variables (Figs. 3.4 and 3.5). Because the white noise sequence a_t is random, each generated simulation will be different. Thus several simulations were generated to compare with the original record.

A statistical analysis of six sets of simulations is shown in Table 3.4. By examining the average of two simulations generated from the fitted parameters and two from the parametric relations in Table 3.1, the error in modelling the accelerogram and in relating the modelling parameters to physical variables was assessed. One value compared was the maximum acceleration. To measure the accuracy in modelling the dominant frequency of the accelerogram, the maximum amplitude of the Fourier spectrum and its corresponding frequency were calculated for both cases. A measure of the error

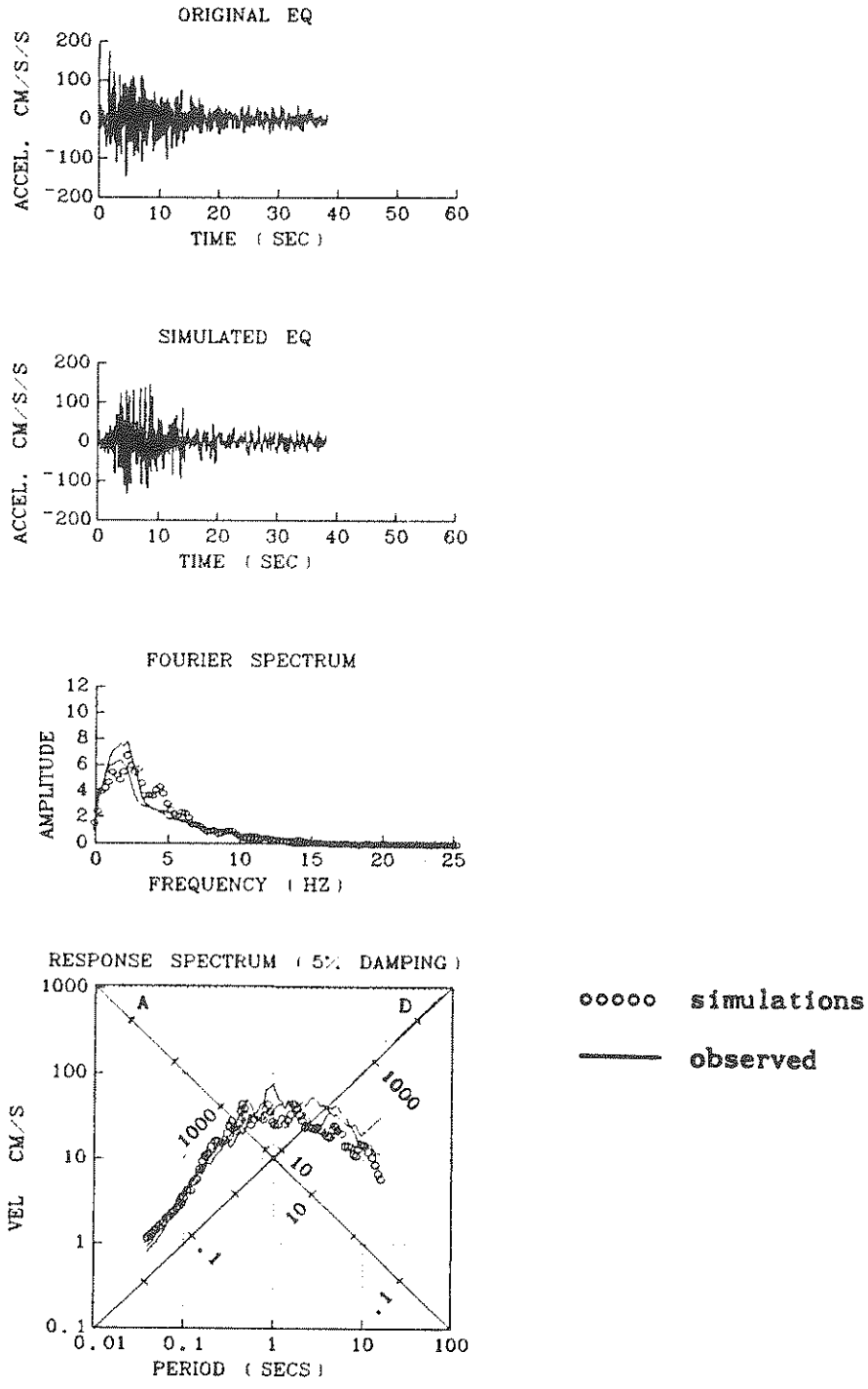


Fig. 3.3 A simulated record.

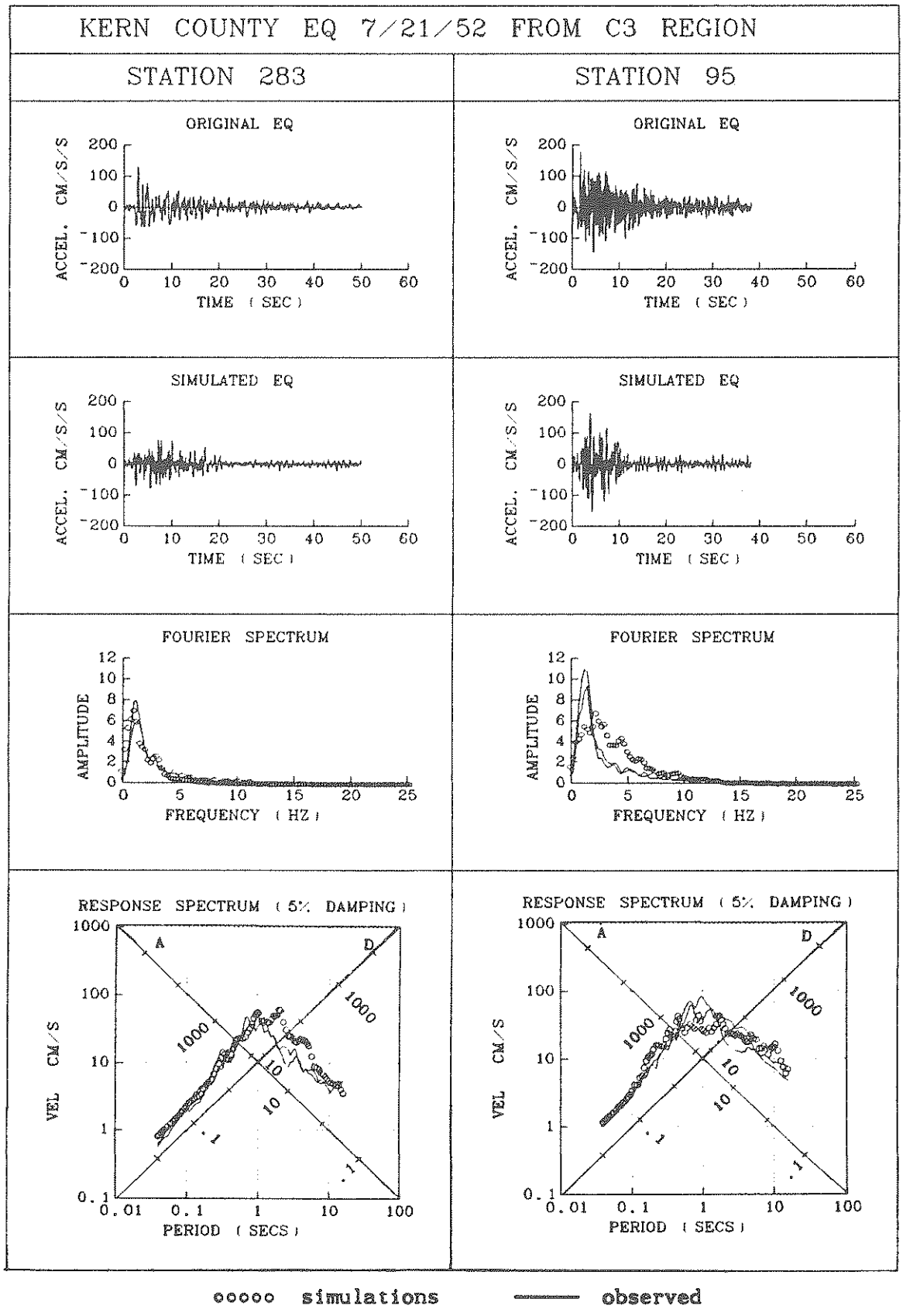


Fig. 3.4 Simulations of one earthquake recorded at two stations.

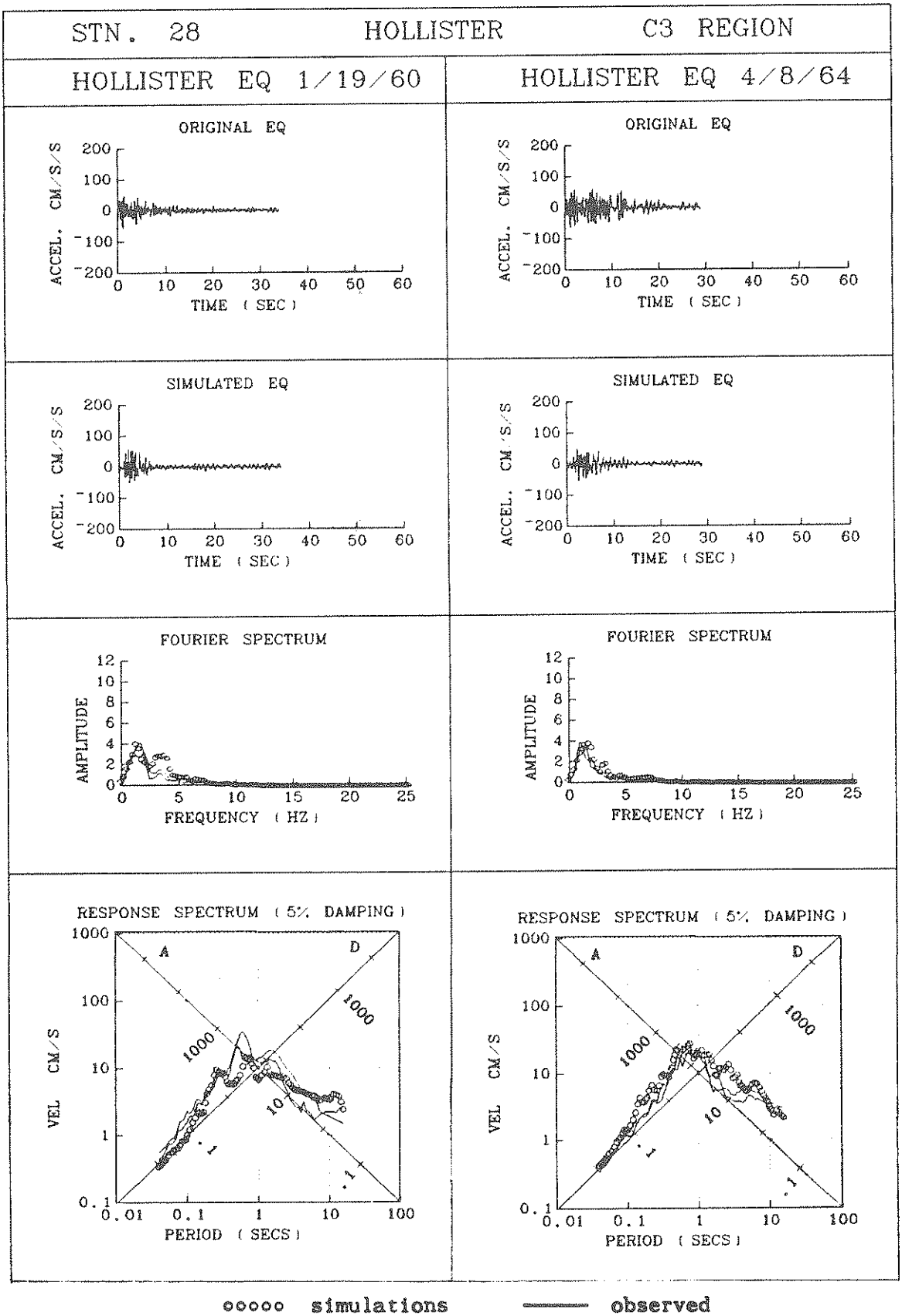


Fig. 3.5 Simulations of two earthquakes recorded at one station.

Table 3.4 California Error Analysis

ACCELEROGRAM		ACTUAL	FROM FITTED PARAMETERS	FROM PARAMETRIC RELATIONS
Kern County	Max. accel.	129	126	84
Eq. 7/21/52	Max F(f)	6.8	8.7	6.8
recorded at	f_{max}	1.00	0.88	1.25
Station 283	σ Fourier Spec.	--	0.40	0.37
	σ Log(SV)	--	0.12	0.19
Kern County	max. accel.	176	136	155
Eq. 7/21/52	Max F(f)	6.6	7.0	9.9
recorded at	f_{max}	2.25	2.00	1.38
Station 95	σ Fourier spec.	--	0.40	0.58
	σ Log(SV)	--	0.16	0.16
Hollister Eq.	max. accel.	56	50	68
1/19/60	Max F(f)	4.0	3.7	3.6
recorded at	f_{max}	1.25	1.63	1.88
Station 28	σ Fourier spec.	--	0.37	0.42
	σ Log(SV)	--	0.52	0.33
Hollister Eq.	max. accel.	63	67	60
4/8/64	Max F(f)	3.8	4.4	3.9
recorded at	f_{max}	1.75	1.50	1.25
Station 28	σ Fourier Spec.	--	0.32	0.33
	σ Log(SV)	--	0.23	0.21
San Francisco	max. accel.	35	42	48
Eq. 3/22/57	Max F(f)	1.8	1.9	1.8
recorded at	f_{max}	3.75	1.63	2.13
Station 65	σ Fourier Spec.	--	0.47	0.58
	σ Log(SV)	--	0.54	0.70
San Francisco	max. accel.	64	56	68
Eq. 3/22/57	Max F(f)	2.4	2.2	2.1
recorded at	f_{max}	4.00	2.62	2.00
Station 77	σ Fourier Spec.	--	0.50	0.65
	σ Log(SV)	--	0.38	0.49

over the entire frequency range is the standard deviation of the Fourier spectrum calculated as

$$\sigma = \left[\frac{\sum (F_{actual} - F_{simulated})^2}{\sum (F_{actual})^2} \right]^{1/2} . \qquad (3.4.1)$$

A useful test for records to be used as input to structural models is a comparison of response spectra. To accurately assess the error over the entire frequency range of the response spectrum, the logarithmic standard deviation of the velocity spectrum was calculated.

$$\sigma = \left[\frac{\sum (Log\ V_{actual} - Log\ V_{simulated})^2}{\sum (Log\ V_{actual})^2} \right]^{1/2} \qquad (3.4.2)$$

By examing Figs. 3.3, 3.4, and 3.5 and Table 3.4, it can be seen that the simulations have similar properties to the original records. In all cases the maximum accelerations are in close agreement. Some variability between the simulations and the original records can be seen in the length of strong shaking. In general, the Fourier spectra and the response spectra closely match. Both the maximum of the Fourier spectrum and the frequency at which the maximum occurs is simulated well with the exception of the Kern County earthquake recorded at station 95. As expected, the simulations that were generated from parameters fitted to the original data matched the original records more closely than the simulations generated from parameters restricted to the relations in Table 3.1.

REFERENCES

Hudson, D. E. (ed.) (1976). *Strong Motion Earthquake Accelerograms - Index Volume,* Report No. EERL 76-02, EERL, California Institute of Technology, Pasadena, CA.

Richter, C. F. (1935). "An Instrumental Earthquake Magnitude Scale," *Bulletin of the Seismological Society of America,* Vol. 5, 1-32.

Terashima, T. (1968). "Magnitude of Microearthquakes and Spectra of Microearthquake Waves," *Bulletin of the International Institute of Seismology and Earthquake Engineering,* Vol. 5, 31-108.

Trifunac, M. D. and A. G. Brady (1975). "On the Correlation of Seismic Intensity Scales with the Peaks of Recorded Strong Ground Motion," *Bulletin of the Seismological Society of America,* Vol. 65, 307-321.

4. MULTIVARIATE MODELLING PROCEDURE

4.1. Introduction

Figure 4.1 shows a typical set of acceleration components. Several observations are obvious and occur in most accelerograms. First, the vertical component initially is stronger due to the arrival of P-waves early in the record, but eventually the stronger horizontal components dominate. Second, the horizontal components are similar in severity of shaking and in the shape of their standard deviation envelopes. Third, the change in zero axis crossings over the length of the accelerogram is similar for all three components. Using these characteristics a multivariate procedure was devised to accurately model the three acceleration components as a group. The advantage of this model is that the relationships between the three components are measured. The study of these relationships resulted in a more efficient modelling procedure which required fewer parameters to describe the accelerograms and produced more accurate three-dimensional simulations.

4.2. Stabilization Procedure

The stabilization procedure is similar to the univariate procedure in that the accelerograms are first shortened, their variance and frequency content is stabilized, and finally an ARMA model is fitted to the stabilized series. The differences are

(1) Each record is shortened by the same amount depending on the cumulative energy of the vector magnitude of the three components.

(2) Instead of three separate standard deviation envelopes being used to stabilize the variance of the three components, a standard deviation envelope calculated from the vector magnitude and an envelope measuring the average vertical angle of the vector magnitude are used to construct standard deviation envelopes to stabilize the horizontal components and the vertical component.

(3) The zero axis crossings of all three components are added together to stabilize the frequency content by changing the time scales of all three components equally.

(4) A multivariate ARMA model is fitted to the stabilized acceleration time series to measure the cross-correlation among components.

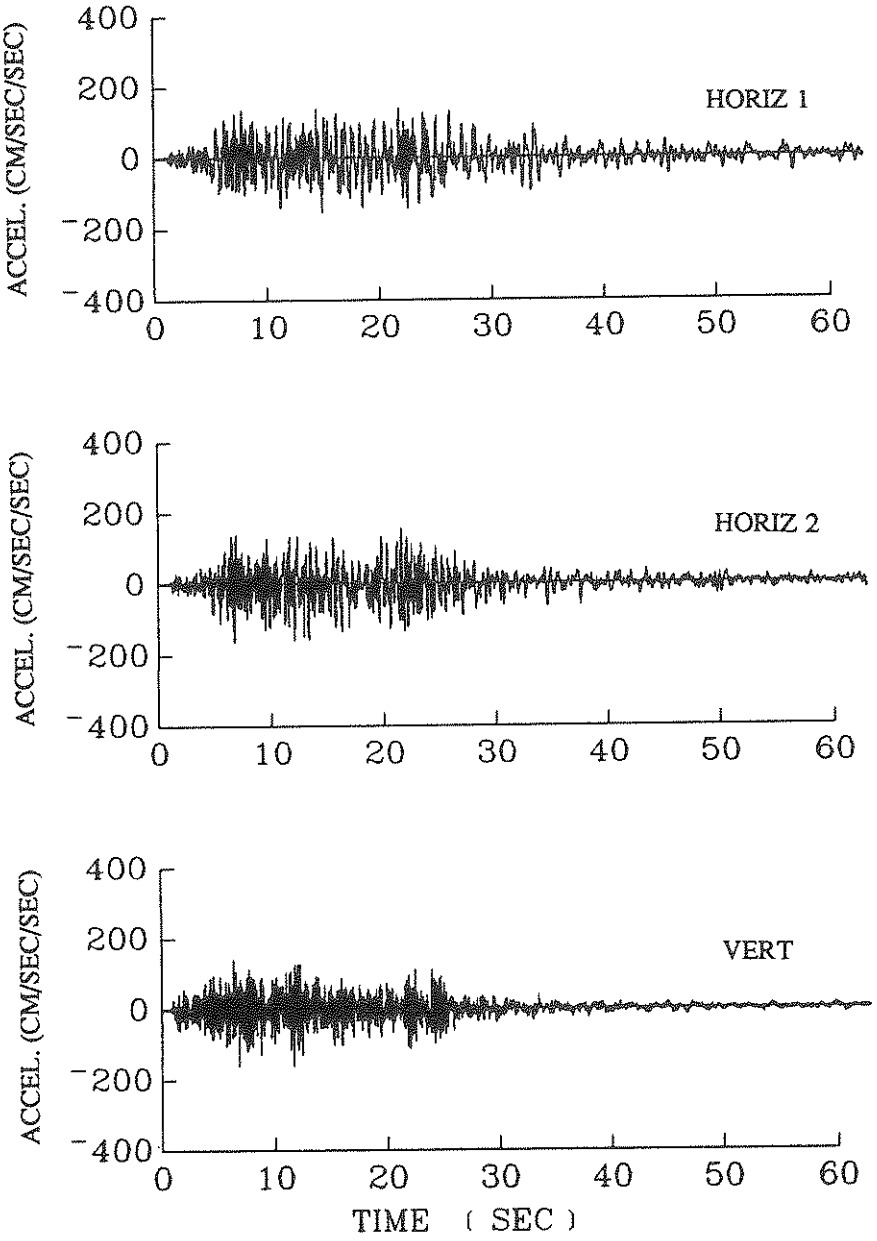


Fig. 4.1 Three components of the October 15, 1979 Mexicali Valley earthquake recorded at station Cerro Prieto.

(5) The ARMA parameters are transformed into parameters describing the frequency spectra.

4.2.1. Changing Coordinate Systems

The first step in stabilizing the accelerograms is to convert the acceleration components from the Cartesian coordinate system (H_1, H_2, and V) into spherical coordinates (ρ, a, and γ) as shown in Fig. 4.2 by

$$H_1(t)=\rho(t)\cos\gamma(t)\cos a(t) \tag{4.2.1.1}$$

$$H_2(t)=\rho(t)\cos\gamma(t)\sin a(t) \tag{4.2.1.2}$$

$$V(t)=\rho(t)\sin\gamma(t)\,. \tag{4.2.1.3}$$

The vector magnitude, $\rho(t)$, and the vertical angle, $\gamma(t)$, will be used to stabilize the acceleration time series in Cartesian coordinates.

4.2.2. Shortening the records

The procedure for shortening the records is similar to that used in Section 2.3 for the univariate modelling procedure. The only difference is that instead of shortening each component separately, the vector magnitude, $\rho(t)$, is used to shorten each component by an equal amount. Thus the energy is calculated as

$$I_0=\sum \rho_t^2\Delta t\,. \tag{4.2.2.1}$$

Dividing the amount of energy of the vector magnitude at any time, t, by the total energy of the vector magnitude results in the cumulative energy function. Using the cumulative energy function, the acceleration components in both Cartesian and spherical coordinates are shortened by eliminating the beginning of the components corresponding to 1% energy and the end of the components corresponding to 2% energy. The shortened records are shown in Figs. 4.3 and 4.4 respectively.

4.2.3. Variance Stabilization

To stabilize the variance of the acceleration components, two envelopes are used. One is the vertical angle envelope, $\hat{\gamma}(t)$, estimated from the shortened vertical angle time series, $\gamma(t)$. The other is the standard deviation envelope, $\hat{\sigma}_\rho(t)$, estimated from

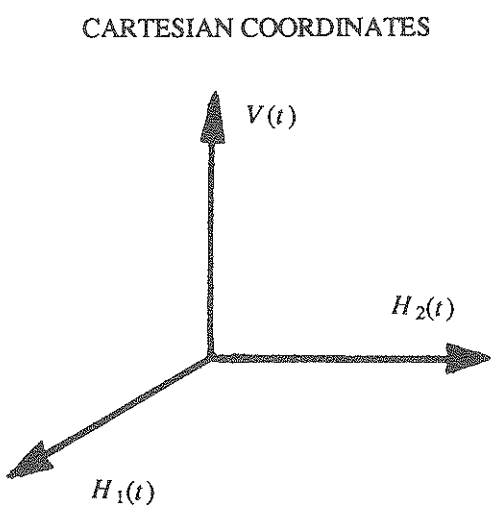


Fig. 4.2 Coordinate systems used in the multivariate analysis.

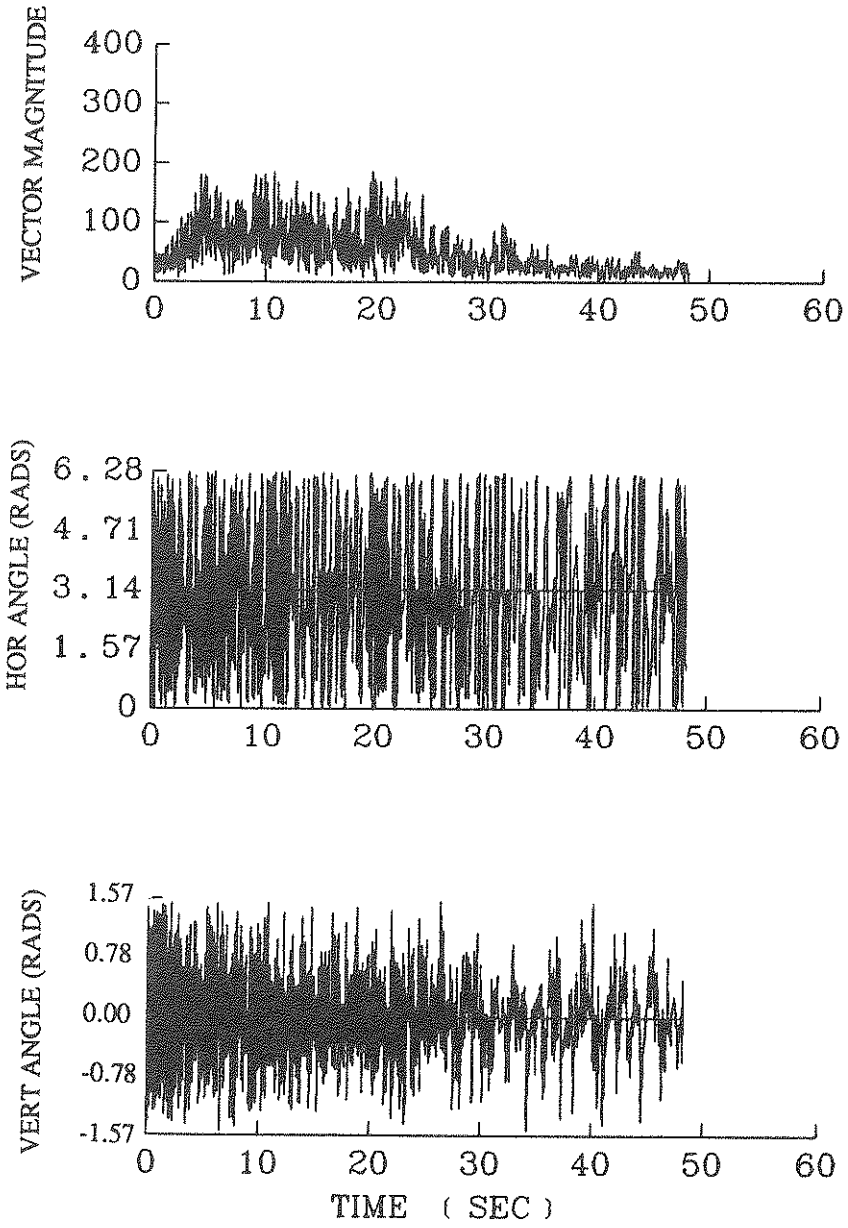


Fig. 4.3 The acceleration time series in spherical coordinates.

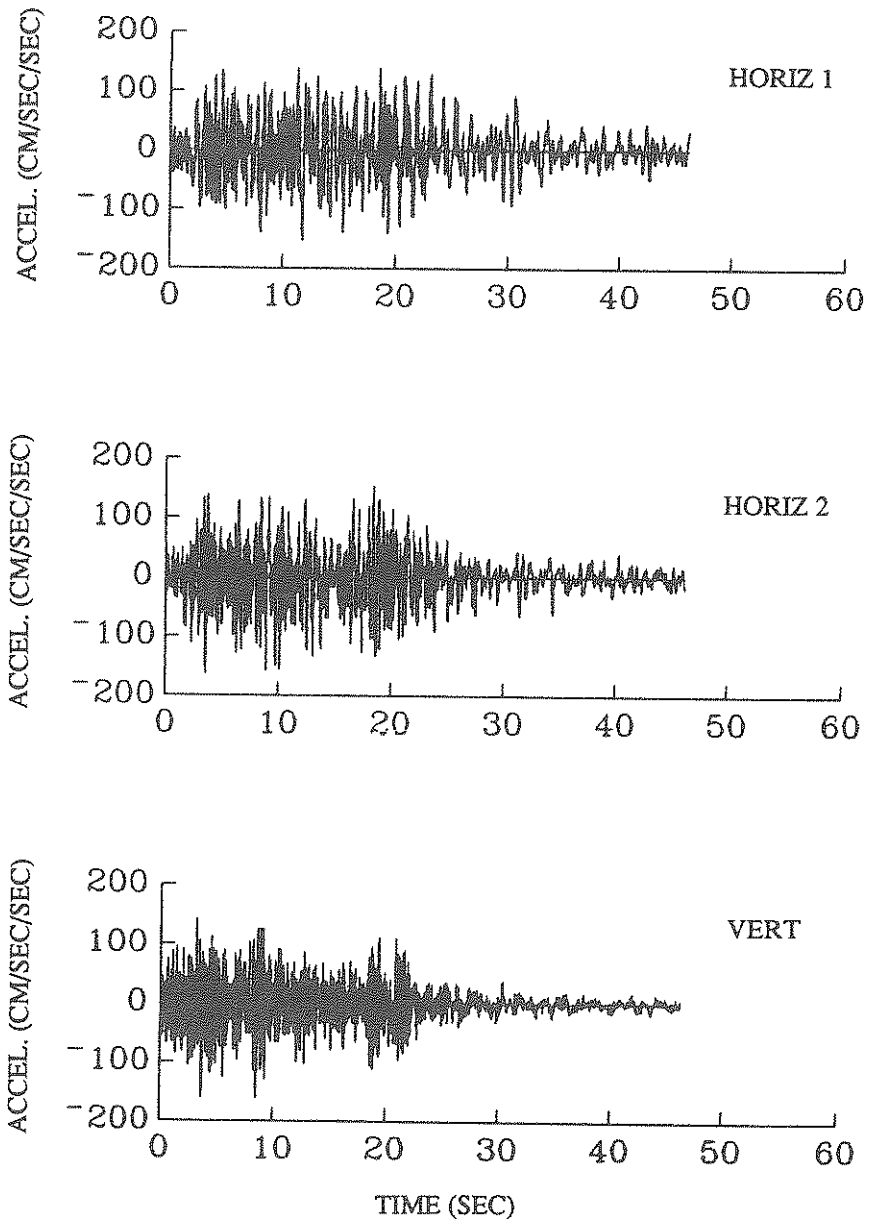


Fig. 4.4 The shortened accelerogram in Cartesian coordinates.

the shortened vector magnitude, $\rho(t)$.

To calculate the vertical angle envelope, the first step is to calculate the running average of the absolute value of the vertical angle, $\bar{\gamma}(t)$, by

$$\bar{\gamma}(t) = \frac{1}{101} \sum_{i=t-50}^{i=t+50} |\gamma_i| . \qquad (4.2.3.1)$$

As shown in Fig. 4.5a a smooth function, $\hat{\gamma}(t)$, is then fit to $\bar{\gamma}(t)$.

$$\hat{\gamma}(t) = (c_3 - k_3)(1 + \frac{t}{b_3})e^{-\frac{t}{b_3}} + k_3 \qquad (4.2.3.2)$$

where

c_3 = the initial value of the function

b_3 = the rate of decay

and

k_3 = the lower limit of the function.

The value of c_3 is estimated as the mean value of $\bar{\gamma}(t)$ during the first 10% of the record. The lower limit of the function, k_3, is estimated as the mean value of $\bar{\gamma}(t)$ during the final 1/3 of the record. Finally, the value of b_3 is estimated so that the areas under $\hat{\gamma}(t)$ and $\bar{\gamma}(t)$ are equal.

$$\sum \hat{\gamma}_i \Delta t - \sum \bar{\gamma}_i \Delta t = 0 \qquad (4.2.3.3)$$

To relate these parameters to physical variables the ratio, $r_3 = \frac{c_3}{k_3}$, was calculated.

The standard deviation envelope is calculated from the vector magnitude, $\rho(t)$. Because $\rho(t)$ has a zero mean, an estimate of the variance of the series may be obtained by

$$\sigma_\rho^2(t) = \frac{1}{101} \sum_{i=t-50}^{t+50} \rho_i^2 . \qquad (4.2.3.4)$$

The square root provides an estimate of the standard deviation envelope, $\sigma_\rho(t)$.

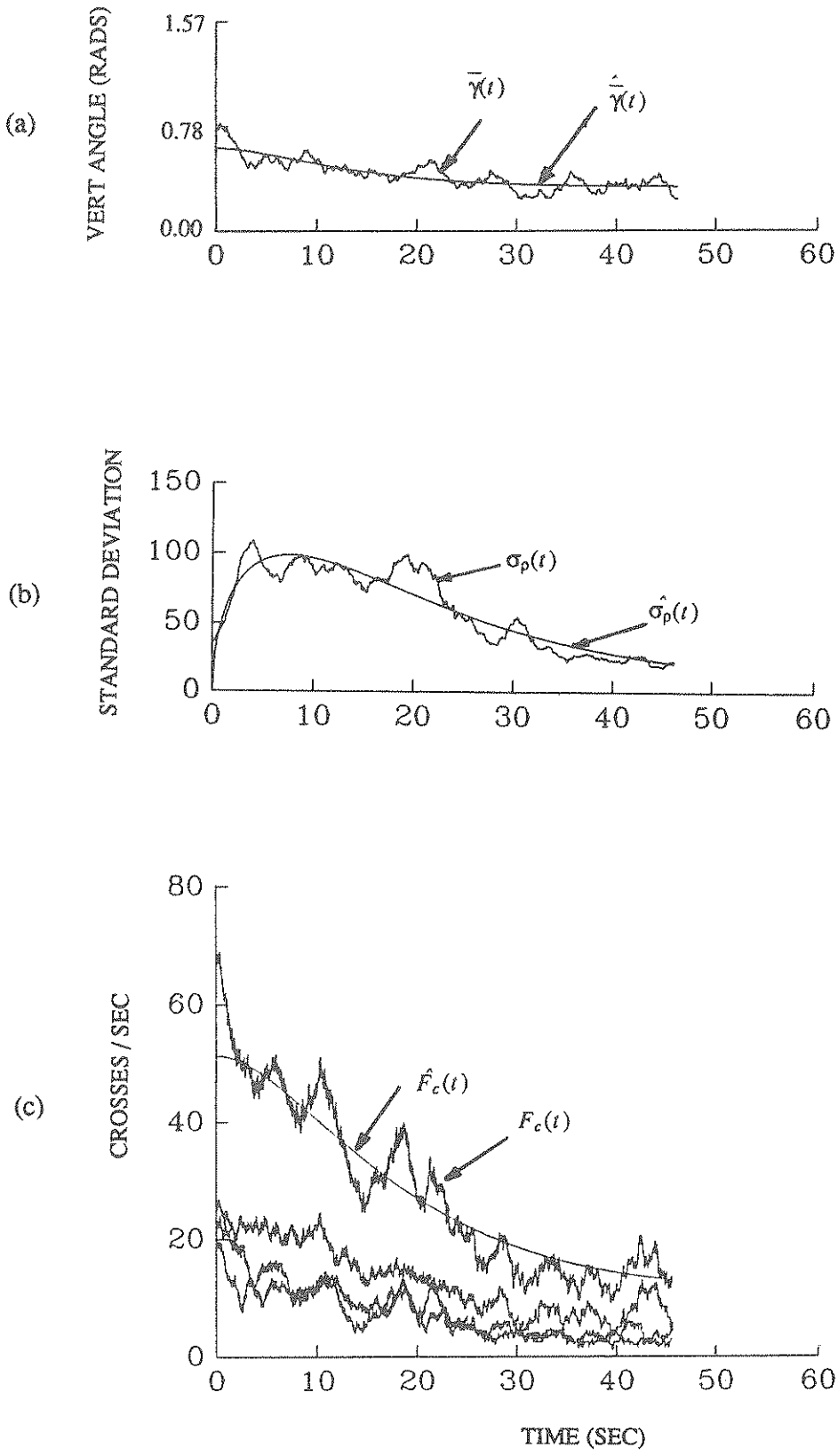


Fig. 4.5 Functions used to stabilize the accelerograms: (a) vertical angle envelope, (b) standard deviation envelope, and (c) frequency envelope.

A smooth function, $\hat{\sigma}_p(t)$, of the form

$$\hat{\sigma}_p(t) = c_1(\alpha - k_1)(\frac{t}{\tau})^p e^{-(\frac{c_2}{\tau})t} + k_1 \tag{4.2.3.5}$$

where

$$c_1 = \frac{8e^3}{3\sqrt{3}}$$

$$c_2 = 2\sqrt{3}$$

can now be fit to the standard deviation envelope. The value of α is a measure of the maximum of the strong shaking segment. The weak shaking is measured by k_1. The duration of strong shaking is measured by τ and the product of τ, and p measures the time to the maximum of the function, t_{max}. The parameters of the function are estimated by minimizing the error between the standard deviation envelope and the function to be fit to the envelope

$$error = \sum [\,(\hat{\sigma}_p)_i - (\sigma_p)_i]^2 \tag{4.2.3.6}$$

using subroutine ZXMIN in the IMSL subroutine library (1977).

By using the functions $\bar{\gamma}(t)$ and $\sigma_p(t)$, standard deviation envelopes for stabilizing the variance of each acceleration component in Cartesian coordinates may be calculated as

$$\sigma_v(t) = \sigma_p(t)\,\sin\bar{\gamma}(t) \tag{4.2.3.7}$$

$$\sigma_h(t) = \frac{1}{\sqrt{2}}\sigma_p(t)\,\cos\bar{\gamma}(t) \tag{4.2.3.8}$$

where σ_v is the standard deviation envelope for stabilizing the variance of the vertical component and σ_h is the standard deviation envelope for stabilizing the horizontal components. The variance stabilized time series are calculated as

$$\hat{V}(t) = \frac{V(t)}{\sigma_v(t)} \tag{4.2.3.9}$$

$$\hat{H}_1(t) = \frac{H_1(t)}{\sigma_h(t)} \tag{4.2.3.10}$$

$$\hat{H}_2(t) = \frac{H_2(t)}{\sigma_h(t)}\,. \tag{4.2.3.11}$$

The three components are now stabilized to have a variance of approximately 1.0.

4.2.4. Frequency Stabilization

As in the univariate model, it is still necessary to stabilize the frequency content over the length of the record. This is done by using a frequency envelope to change the time scale of the record. The frequency envelope is computed by calculating the crosses per second of each component of the accelerogram, adding them together to get the total number of crosses, and then fitting a smooth function to the total number of crosses. This is shown in Fig. 4.5c with the crosses of the components shown on the bottom of the graph, their sum labeled as $F_c(t)$, and the smooth function labeled as $\hat{F}_c(t)$.

The sum of zero axis crossings is calculated using a two second time window as

$$F_c(t) = \frac{\textit{no. of zero axis crossings of all components between } t \pm 1 \textit{ second}}{2 \textit{ seconds}} . \qquad (4.2.4.1)$$

The same smooth function that was used to model the vertical envelope was fit to the zero axis crossings.

$$\hat{F}_c(t) = (c_2 - k_2)(1 + \frac{t}{b_2})e^{-\frac{t}{b_2}} + k_2 \qquad (4.2.4.2)$$

The value of c_2 is estimated as the mean value of $F_c(t)$ during the first 10% of the record. The lower limit of the function, k_3, is estimated as the mean value of $F_c(t)$ during the final 1/3 of the record. Finally, the value of b_3 is estimated so that the areas under $\hat{F}_c(t)$ and $F_c(t)$ are equal.

The frequency function, $\hat{F}_c(t)$, may then be used to change the time increments of the variance stabilized acceleration components by

$$\Delta t' = (\Delta t)\hat{F}_c(t) \qquad (4.2.4.3)$$

where

Δt = the original time increment of 0.02 seconds

$\Delta t'$ = the new time increment.

The transformed records are then reduced to the same length of time as the initial records by

$$\Delta t'' = \Delta t' \times \frac{duration\ of\ the\ original\ records}{duration\ of\ the\ transformed\ records} \qquad (4.2.4.4)$$

and digitized at 0.02 second intervals using linear interpolation. Since all of the components are stabilized by the same frequency function, the relative time among the components remains the same.

4.3. Constrained ARMA Model

The stabilization procedure results in three stabilized time series to be modelled by an ARMA model as shown in Fig. 4.6. However, it is no longer appropriate to use the univariate model presented in Section 2.6 since it will model each time series independently without capturing any cross-correlation among the stabilized series. To include the cross-correlation in the model, a multivariate ARMA model is required. The multivariate ARMA model allows feedback among the time series, and thus may reduce the modelling parameters necessary to model the time series.

4.3.1. Full Multivariate Model

A full multivariate time series model for time series with zero mean may be written as

$$\mathbf{Z}_t - \phi_1 \mathbf{Z}_{t-1} - \phi_2 \mathbf{Z}_{t-2} - \cdots - \phi_p \mathbf{Z}_{t-p} = \mathbf{a}_t - \theta_1 \mathbf{a}_{t-1} - \theta_2 \mathbf{a}_{t-2} - \cdots - \theta_q \mathbf{a}_{t-q} \qquad (4.3.1.1)$$

where

n = the number of time series

p = the order of the AR component

$\phi_k = n \times n$ autoregressive parameter matrix at lag k, $k=1, 2,..., q$

$\mathbf{Z}_k$ = a vector of length n of the component of each time series at lag k

q = order of MA component

$\theta_k = n \times n$ moving-average parameter matrix at lag k, $k=1, 2,..., p$

$\mathbf{a}_k$ = a vector of length n of the component of each white noise series at lag k .

Thus in the multivariate case the ARMA parameters become matrices. To estimate the values of the parameter matrices, the SCA Statistical System (1986) was used.

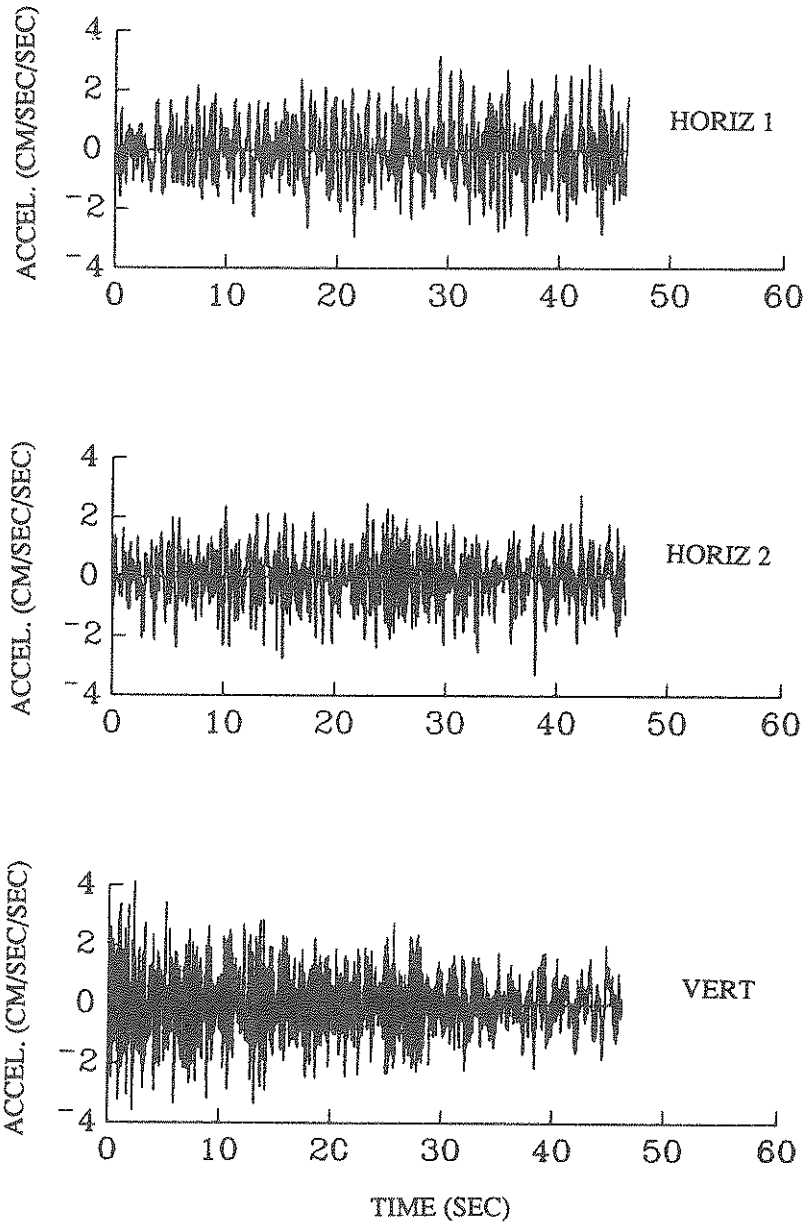


Fig. 4.6 Stabilized accelerograms.

To apply the multivariate model to the stabilized accelerogram components, an ARMA model must be found that meets the zero response at zero frequency condition (F(0) = 0). Because the only constraint allowed in the SCA Statistical System is that matrix elements may be set equal to constants, a constraint of the form used in Section 2.6 ($\theta_1 + \theta_2 = 0.99$) is not possible. The solution was to use a moving average component of the form

$$\theta_1 = \begin{bmatrix} 0.99 & 0 & 0 \\ 0 & 0.99 & 0 \\ 0 & 0 & 0.99 \end{bmatrix}. \qquad (4.3.1.2)$$

Since all the off-diagonal terms are zero, this is simply a set of three separate univariate moving-average terms. The Fourier spectrum of a first order moving average model is

$$F(f) = [\, 2\sigma_a^2(1 + \theta_1^2 - 2\theta_1 \cos 2\pi f \,)\,]^{1/2}. \qquad (4.3.1.3)$$

For $f = 0$, the Fourier amplitude is equal to zero when $\theta_1 = 1$. This violates the invertibility requirement

$$|\theta_1| < 1. \qquad (4.3.1.4)$$

Therefore, the value 0.99 was used to satisfy the invertibility requirement of Eq. (4.3.1.4) while allowing negligible response at zero frequency.

To find the order of the AR parameter matrices to be used, six accelerograms from northern and central Mexico were modelled and multivariate ARMA (2,1) and ARMA (3,1) parameters were estimated with the moving-average matrix constrained as in Eq. (4.3.1.2). The quality of the fit of each model was measured by comparing the correlation of the residual series, the variance of the residual series, and the Fourier spectrum calculated from the ARMA parameters. It was found that the ARMA (3,1) model fit the data better than the ARMA (2,1) model. The need for the higher order model was greatest for time series containing a broad frequency content. Although both processes can model a time series with a Fourier spectrum containing a single peak, the extra coefficient in the ARMA (3,1) model allows much more flexibility in the shape of the Fourier spectrum of the ARMA coefficients fit to the time series. For cases in which the Fourier spectrum of the time series has a very narrow peak, the ARMA (2,1) model was found to fit equally well as the ARMA (3,1) model. In no cases was a model of higher order than the ARMA (3,1) found to be necessary to model the correlation structure of the time series.

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \phi_3 Z_{t-3} = a_t - \theta_1 a_{t-1} \qquad (4.3.1.5)$$

where

Z_k = a vector of length 3 of the stabilized acceleration components at lag k

ϕ_1 , ϕ_2 , ϕ_3 = 3×3 autoregressive matrices

θ_1 = the diagonalized matrix shown in Eq. (4.3.1.2) .

The estimated parameters are the ϕ_1 , ϕ_2 , and ϕ_3 matrices which have a total of 27 elements. It was found that only the diagonal terms were always significant at the 95% confidence level. Because the matrices were not symmetric and no patterns were found in the off-diagonal elements, it was impossible to simplify the model. Because the complexity of the model made it impossible to relate the ARMA parameters to physical variables, and because the off-diagonal terms were very small even when they were significant at the 95% confidence level, a simplified model was investigated.

4.3.2. Diagonalized Multivariate Model

To simplify the model, all of the off-diagonal ARMA parameters were set equal to zero, reducing the number of modelling parameters from 27 to 9. Thus the diagonalized ARMA model is of the form

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \phi_3 Z_{t-3} = a_t - \theta_1 a_{t-1} \qquad (4.3.2.1)$$

where

Z_k = a vector of length 3 of the stabilized acceleration components at lag k

$$\phi_1 = \begin{bmatrix} \phi_{1,1} & 0 & 0 \\ 0 & \phi_{1,2} & 0 \\ 0 & 0 & \phi_{1,3} \end{bmatrix}$$

$$\phi_2 = \begin{bmatrix} \phi_{2,1} & 0 & 0 \\ 0 & \phi_{2,2} & 0 \\ 0 & 0 & \phi_{2,3} \end{bmatrix}$$

$$\phi_3 = \begin{bmatrix} \phi_{3,1} & 0 & 0 \\ 0 & \phi_{3,2} & 0 \\ 0 & 0 & \phi_{3,3} \end{bmatrix}$$

$$\theta_1 = \begin{bmatrix} 0.99 & 0 & 0 \\ 0 & 0.99 & 0 \\ 0 & 0 & 0.99 \end{bmatrix}$$

a_k = a vector of length 3 of the component of each white noise series at lag k .

This model is the equivalent of three univariate ARMA models. The only correlation among time series is measured by the covariance matrix of the residuals. Thus the model is a compromise between the simplicity of the unvariate model which treats each time series independently and the full multivariate model which captures the cross-correlation among components but does not result in a parsimonious model.

Because the covariance matrix of the residuals is symmetric, it is composed of only six independent elements. The three diagonal elements of the matrix, σ_1^2 , σ_2^2 and σ_3^2 represent the variance of the series and can be calculated as a function of the ARMA parameters by

$$\sigma_i^2 = \frac{c_1 - c_2 - c_3 - c_4 - c_5}{c_6 - c_7} \qquad (4.3.2.2)$$

$c_1 = 1 - \phi_2 - \phi_3(\phi_1 + \phi_3) - \phi_1(\phi_1 + \phi_2\phi_3)$

$c_2 = \phi_2[\phi_1(\phi_1 + \phi_3) + \phi_2(1 - \phi_2)]$

$c_3 = \phi_1\phi_3[\phi_1(\phi_1 + \phi_3) + \phi_2]$

$c_4 = \phi_3(1 - \phi_2)(\phi_1\phi_2 + \phi_3)$

$c_5 = \phi_3^2[\phi_2^2 - \phi_3(\phi_1 + \phi_3)]$

$c_6 = [1 - \theta_1(\phi_1 - \theta_1)][1 - \phi_2 - \phi_1\phi_3 - \phi_3^2]$

$c_7 = \theta_1[\phi_1 + \phi_2(\phi_1 + \phi_3) + \phi_3(\phi_1^2 + \phi_1\phi_3 + \phi_2)]$.

This relationship is derived in Appendix A.3. From the three off-diagonal terms, the cross-correlation, $\rho_{i,j}$, between the residuals is calculated as

$$\rho_{i,j} = \frac{\sigma_{i,j}^2}{\sigma_i \sigma_j} \; . \qquad (4.3.2.3)$$

Thus the cross-correlation among the time series is expressed only by the three correlation coefficients of the residuals. By averaging the cross-correlation between the first horizontal component and the vertical component with the second horizontal

component and the vertical component, the cross-correlation can be expressed as

$$\rho_{H,V} = \frac{1}{2}[\rho_{1,3} + \rho_{2,3}] \tag{4.3.2.4}$$

$$\rho_{H,H} = \rho_{1,2} \tag{4.3.2.5}$$

where $\rho_{H,V}$ is the correlation between the horizontal and vertical residual series and $\rho_{H,H}$ is the correlation between horizontal components of the residuals. This model was used in the Mexico and Taiwan study. It was found that the cross-correlation among the residuals was usually negligible (less than 0.10).

To relate the ARMA parameters to physical variables the three sets of auto-regressive coefficients were factored into a multiplicative model. If the original model is expressed using the backshift operator, B, as

$$(\mathbf{I} - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)\mathbf{Z}_t = (\mathbf{I} - \theta_1 B)\mathbf{a}_t \tag{4.3.2.6}$$

where $\mathbf{I}$ is the identity matrix and B operates such that $B\mathbf{Z}_t = \mathbf{Z}_{t-1}$, $B^2\mathbf{Z}_t = \mathbf{Z}_{t-2}$, etc., then the left hand side of the equation can be factored into an AR(1) and an AR(2) term as

$$(\mathbf{I} - \phi_3' B)(\mathbf{I} - \phi_1' B - \phi_2' B^2)\mathbf{Z}_t = (\mathbf{I} - \theta_1 B)\mathbf{a}_t \tag{4.3.2.7}$$

where

$$\phi_3' = \frac{1}{r_1}$$

$$\phi_2' = -\frac{\phi_3}{\phi_3'}$$

$$\phi_1' = \phi_1 - \phi_3'$$

r_1 = the only real root for B of $I - \phi_1 B - \phi_2 B^2 - \phi_3 B^3$.

For the multiplicative model shown in Eq. (4.3.2.7), the maximum of the Fourier spectrum occurs at the peak of the AR(2) factor. The frequency of the maximum Fourier amplitude of an AR(2) model can be calculated as

$$f_{\max} = \frac{50}{2\pi}\cos^{-1}\left[\frac{|\phi_1'|}{2\sqrt{-\phi_2'}}\right]. \tag{4.3.2.8}$$

The maximum of the Fourier spectrum can then be calculated by substituting $f_{\max}$ computed in Eq. (4.3.2.8) into the equation of the Fourier spectrum of an ARMA (3,1)

model.

(4.3.2.9)

$$F(f) = \left[\frac{2\sigma_a^2(1 + \theta_1^2 - 2\theta_1 \cos 2\pi \frac{f}{50})}{[50(2\pi)][1 + \phi_1'^2 + \phi_2'^2 - 2\phi_1'(1 - \phi_2')\cos 2\pi \frac{f}{50} - 2\phi_2' \cos 4\pi \frac{f}{50}][1 + \phi_3'^2 - 2\phi_3' \cos 2\pi \frac{f}{50}]} \right]^{1/2}$$

Also it was found that the summation of the auto-regressive terms was slightly less than one for all of the time series examined.

$$\phi_1 + \phi_2 + \phi_3 \approx 0.99 \qquad (4.3.2.10)$$

This can be expressed for the multiplicative model as

$$(\phi_1' + \phi_3') + (\phi_2' - \phi_1'\phi_3') - (\phi_2'\phi_3') \approx 0.99\,. \qquad (4.3.2.11)$$

The exact value of the summation varied for each region studied. Thus the ARMA parameters estimated for each of the series, ϕ_1, ϕ_2, and ϕ_3, were transformed into ARMA parameters representing an equivalent multiplicative model, ϕ_1', ϕ_2', and ϕ_3'. From the multiplicative parameters the maximum Fourier amplitude, $F(f_{max})$, and the frequency at which the maximum occurs, f_{max}, are calculated. The relationship among the AR parameters in Eq. (4.3.2.10) or Eq. (4.3.2.11) completes the specification of the model.

Originally a multivariate stabilization procedure and a multivariate ARMA process were used to model the accelerograms. It has been shown that the off-diagonal ARMA coefficients and the cross-correlation of the residual series calculated for most time histories was not statistically different from zero. Thus the procedure was modified to a multivariate stabilization procedure and a univariate ARMA process. The advantage of using the multivariate stabilization procedure is that fewer functions were needed. Modelling three accelerogram components with the univariate procedure requires three standard deviation envelopes and three frequency envelopes. The multivariate procedure requires only three envelopes: the standard deviation envelope, the frequency envelope, and the vertical angle envelope. Thus it is easier to relate the modelling parameters estimated from the multivariate procedure to physical

variables.

4.4. Generating Simulations

To generate simulations using the multivariate procedure, the modelling procedure is reversed. Given f_{max} and $F(f_{max})$ and Eq. (4.3.2.11) the parameters for the multiplicative ARMA model may be calculated by solving Eqs. (4.3.2.8), (4.3.2.9) and (4.3.2.11) for ϕ_1', ϕ_2', and ϕ_3'. From the multiplicative ARMA parameters, the standard ARMA parameters may be calculated by

$$\phi_1 = \phi_1' + \phi_3' \tag{4.4.1}$$

$$\phi_2 = \phi_2' - \phi_1'\phi_3' \tag{4.4.2}$$

$$\phi_3 = -\phi_2'\phi_3' . \tag{4.4.3}$$

From these parameters three series with stable variance and frequency content are generated by

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} - \theta_1 a_{t-1} + a_t \tag{4.4.4}$$

where the terms are defined as in Eq. (4.3.2.1).

To introduce the nonstationary frequency content to the series, a zero crossing frequency envelope, $\hat{F}_c(t)$, is computed from Eq. (4.2.4.2). Because only the shape and not the magnitude of the frequency envelope is important in rescaling the time scale, the ratio $r_2 = \frac{c_2}{k_2}$ and the rate of decay, b_2, were related to physical variables. Thus to calculate the frequency envelope in Eq. (4.2.4.2), the value of k_2 can be chosen as 1.0 and the value of c_2 as r_2. The time axis of each component is rescaled by

$$\Delta t'_s = \frac{0.02}{F_c(t)} . \tag{4.4.5}$$

After changing the time scale of each component, the records are reduced to their original duration by

$$\Delta t''_s = \Delta t'_s \times \frac{\textit{duration of the original simulation}}{\textit{duration of the transformed simulation}} . \tag{4.4.6}$$

The three components are then digitized into equal increments of 0.02 seconds.

To introduce the nonstationary variance into the simulations, the standard deviation envelope and vertical angle envelope are calculated from Eqs. (4.2.3.5) and (4.2.3.2). The standard deviation envelopes for the vertical and horizontal components may be calculated by

$$\sigma_v(t) = \sigma_\rho(t)\sin\hat{\gamma}(t) \tag{4.4.7}$$

$$\sigma_h(t) = \frac{1}{\sqrt{2}}\sigma_\rho(t)\cos\hat{\gamma}(t)\,. \tag{4.4.8}$$

By multiplying the vertical component by $\sigma_v(t)$ and the two horizontal components by $\sigma_h(t)$, a set of simulations nonstationary in both variance and frequency content is created.

References

International Mathematical and Statistical Library 1 (1977). *Reference Manual,* Houston, Texas.

SCA Statistical System (1986). *Reference Manual for Forecasting and Time Series Analysis, Version III,* Scientific Computing Associates, DeKalb, Illinois.

5. MEXICO STUDY

5.1. Introduction

Mexico is a seismically active region situated on the southern edge of the North American Plate. Along the southern and central Pacific coast of Mexico, the Cocos plate is being forced under the North American plate causing subduction earthquakes in that part of the country. In the North, the Pacific Plate is sliding past the North American Plate causing transform faulting in that region. Accelerograms from both areas are included in this study.

Table 5.1 shows a matrix of the accelerograms which were modelled. These accelerograms are from four groups located as shown in Figure 5.1. The group in Mexico City recorded one earthquake and each accelerograph was located at approximately the same distance from the epicenter; thus only the soil conditions vary. The accelerograms in the Guerrero array recorded the same event and were all located on rock, with just the epicentral distance varying. In the two other groups the epicentral distance, soil conditions, and earthquake magnitude all vary.

By analyzing these accelerograms with the multivariate procedure presented in Chapter 4, three models describing the ground motion in Mexico were developed.

(1) A model for various soil conditions in Mexico City for ground motion expected from a magnitude 8+ earthquake along the Pacific coast.

(2) A model for sites located on rock at varying distances from a magnitude 8+ earthquake along the Pacific coast.

(3) A model for any site condition, epicentral distance, or magnitude earthquake.

The development of these models is described in the following sections.

5.2. The Michoacan Earthquake

Table 5.1 Matrix of accelerograms modelled in Mexico.

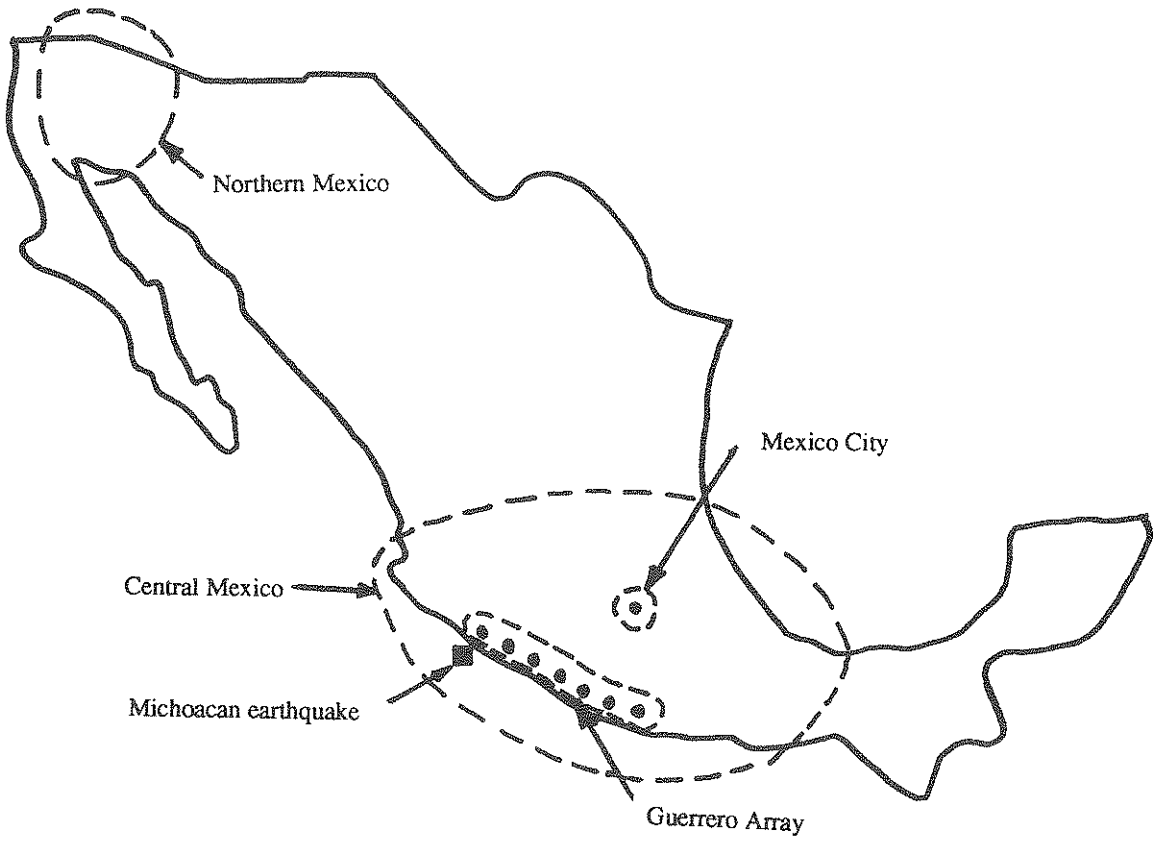


Fig. 5.1 Regions studied in Mexico.

The September 19, 1985 Michoacan earthquake occurred along the Pacific coast of Mexico and recorded a magnitude of 8.1 on the Richter Scale. The depth of the earthquake was estimated at approximately 20 to 30 kilometers. Numerous accelerograms were recorded along the coast by the Guerrero array and in Mexico City. The damage near the epicenter was relatively small, but in Mexico City at a distance of more than 350 kilometers the damage was severe. Together, the accelerograms recorded in Mexico City and the Guerrero array provide a detailed record of the effects of a magnitude 8+ earthquake.

The occurrence of a major earthquake in the Michoacan and Guerrero region was anticipated by seismologists. The subduction of the Cocos Plate beneath Mexico is the most active subduction thrust fault in the Western Hemisphere (Anderson et al., 1986). In this century 42 earthquakes of magnitude 7 or greater have occurred. Before the Michoacan earthquake, the Michoacan seismic gap was identified as an area with high seismic potential due to the fact that it had remained seismically dormant for an unusually long period of time. The Guerrero gap is still considered to have high seismic potential, possibly causing an even greater risk to Mexico City in the future.

Damage along the coast was small. Liquefaction and a 2 to 3 meter tsunami caused damage to a few tourist attractions and industrial estates along the coast. The damage inland was limited to particularly weak buildings with the exception of Mexico City.

The Guerrero array is located along the Pacific coast (see Fig. 5.2) with one station located inland at Teacalco. All of the accelerographs except one were placed on competent rock formations. The values of the compressional wave velocities in this area vary from 1.6 km/sec to 5.6 km/sec with an average of 4 km/sec (Anderson et al., 1986).

A typical set of accelerograms recorded in the region is shown in Fig. 5.3. Unlike the accelerograms recorded in Mexico City, these records have a broad range of frequency content and also shorter duration than those recorded in Mexico City. The accelerograms recorded at two sites, Caleta de Campas and La Villita, were located directly above the fault zone between the two major sources where energy was

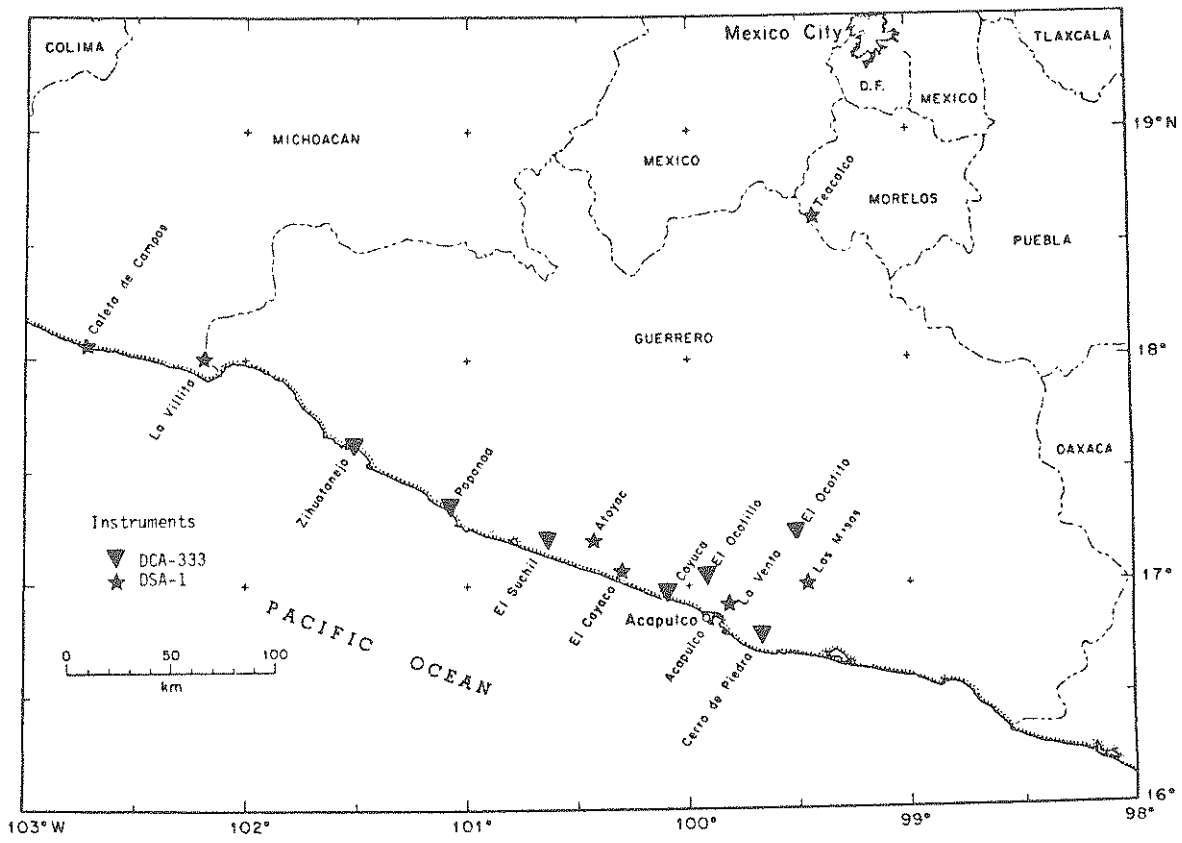


Fig. 5.2 Location of recording stations in the Guerrero array. (after Anderson et al., 1986).

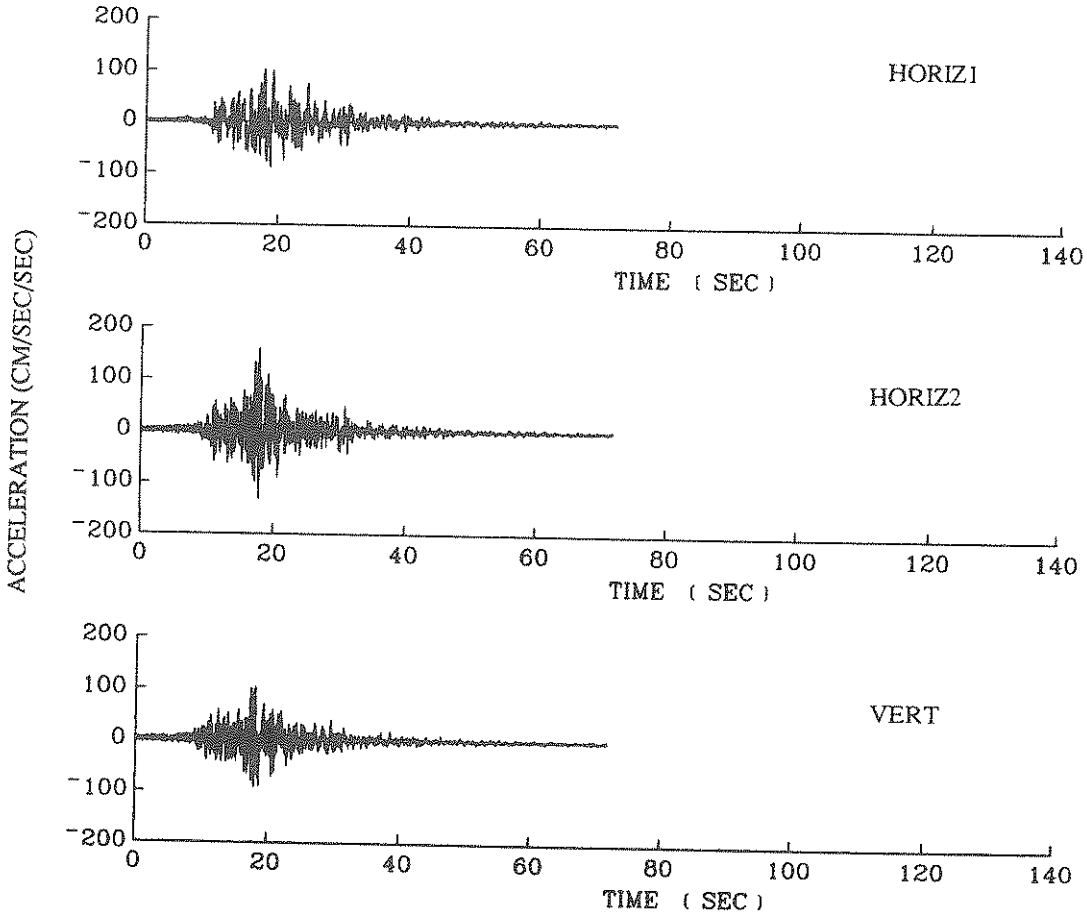


Fig. 5.3 Accelerogram components recorded at station Zihuatenejo.

released. These stations recorded the earthquake as two separate events and were not included in the study. At stations farther to the south, the energy from the two separate subevents arrived at about the same time.

In Mexico City the damage was severe. At least 8000 people were killed or are missing and 30,000 were injured. The number of buildings destroyed or badly damaged numbered 300 to 3300 causing $4 billion dollars in damage (Anderson et al.,1986).

Mexico City is located on a lakebed which has been dried out during the last few centuries. The large amount of damage in Mexico City is mainly due to the amplification of the seismic waves in the soft sediments below the city. Figure 5.4 shows a map of Mexico City with the location of damaged buildings. The city is divided into three zones: the hill zone, the transition zone, and the lake zone. Most of the buildings that collapsed were tall structures located in the lakebed zone. Selected frequencies of the ground motion were greatly magnified by the lakebed causing damage to those buildings with resonance frequencies of about 0.5 Hz.

Accelerograms recorded on rock in the hill zone (station UNAM) and on soft sediments in the lake zone (station SCTI) are shown in Figs. 5.5 and 5.6. A comparison of the two records shows the magnification which occurred due to the amplification of the low frequencies in the soft sediments. The record at SCTI has a single peak in the Fourier spectrum at 0.5 Hz. In addition to the amplification of low frequencies, the duration of strong shaking of the Mexico City records is much longer than those recorded in the source region.

5.3. Mexico City Study

Records from nine stations located in Mexico City and listed in Table 5.1 were analyzed. Because the records are all located at approximately the same distance from the epicenter, the attenuation effects should be negligible among the stations. The differences in the accelerograms, and thus in the modelling parameters, are caused by the varying site conditions.

It was found for all stations except TLHD that a strong correlation exists between the type of soil and the final angle of shaking between the vertical and horizontal

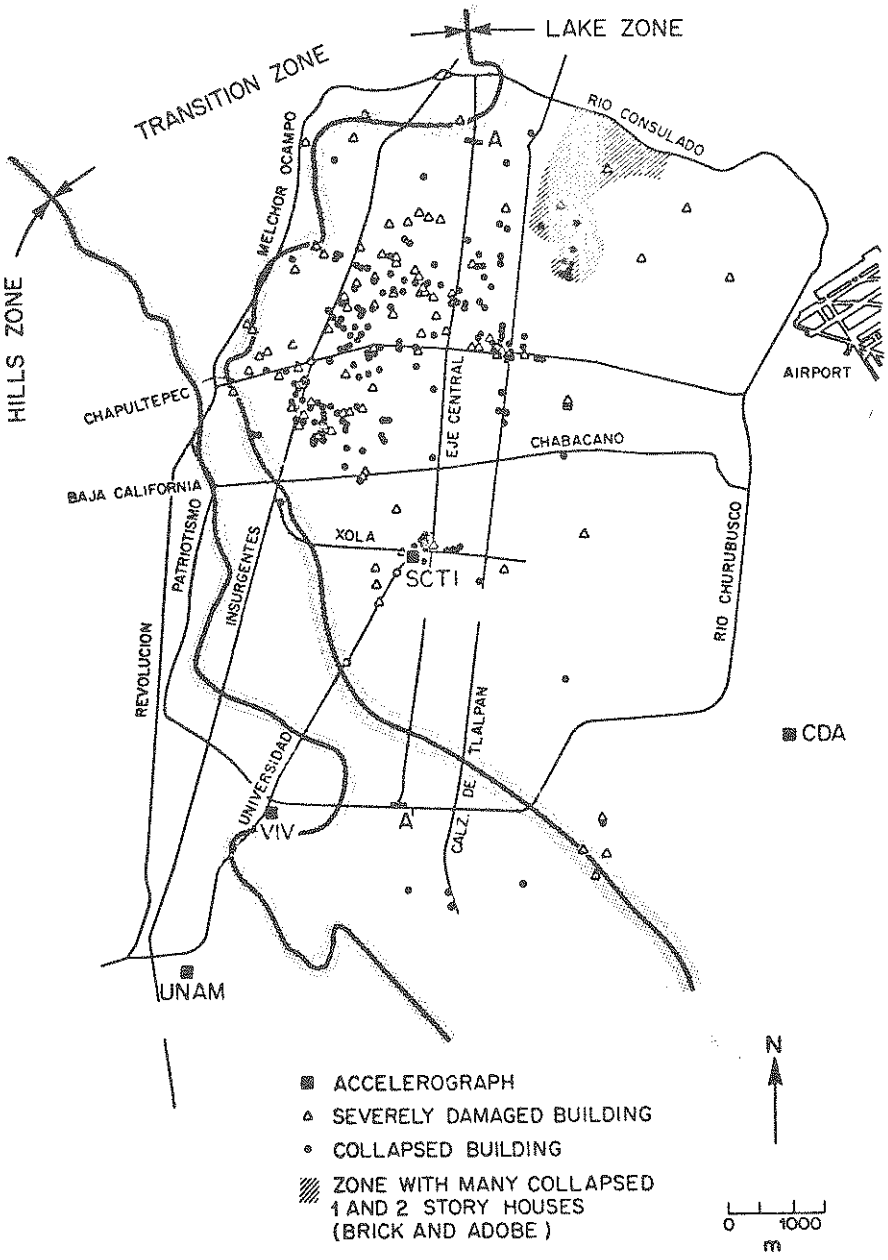


Fig. 5.4 Map of damage in Mexico City caused by the September 19, 1985 Michoacan earthquake.

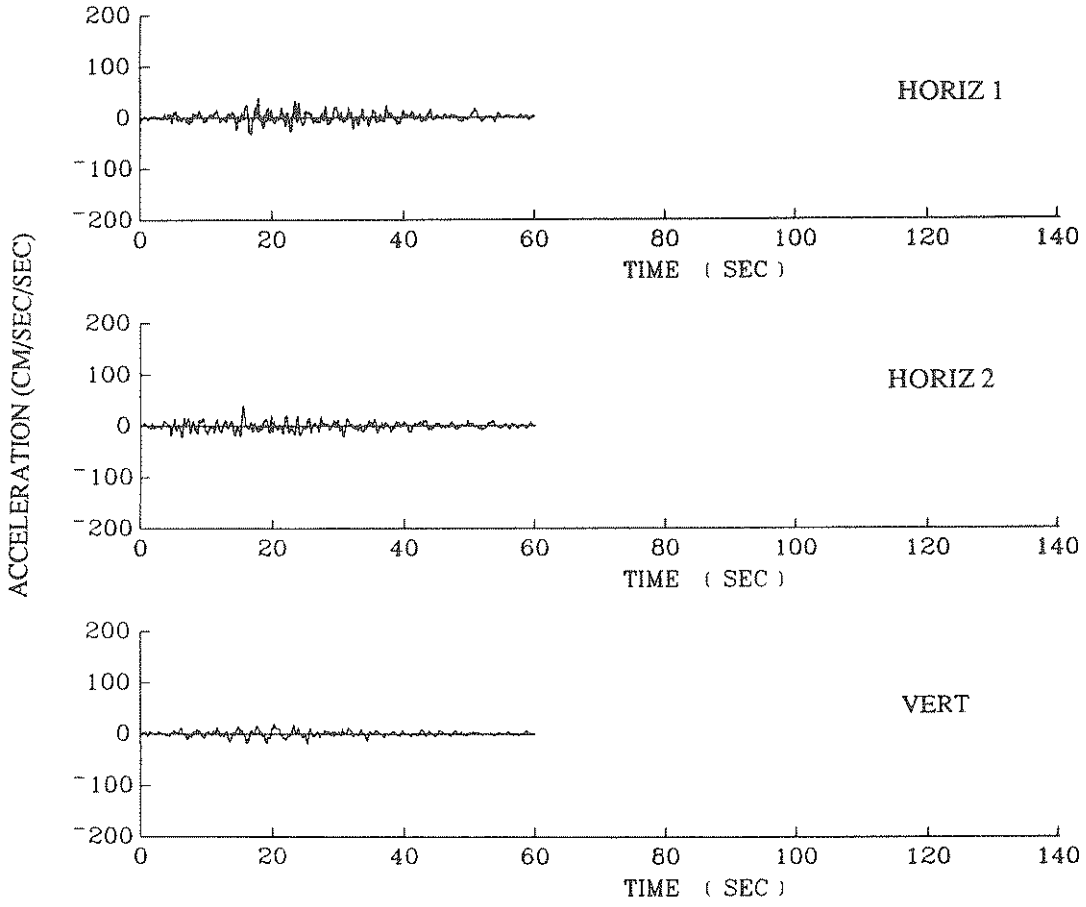


Fig. 5.5 Accelerogram components recorded at station UNAM.

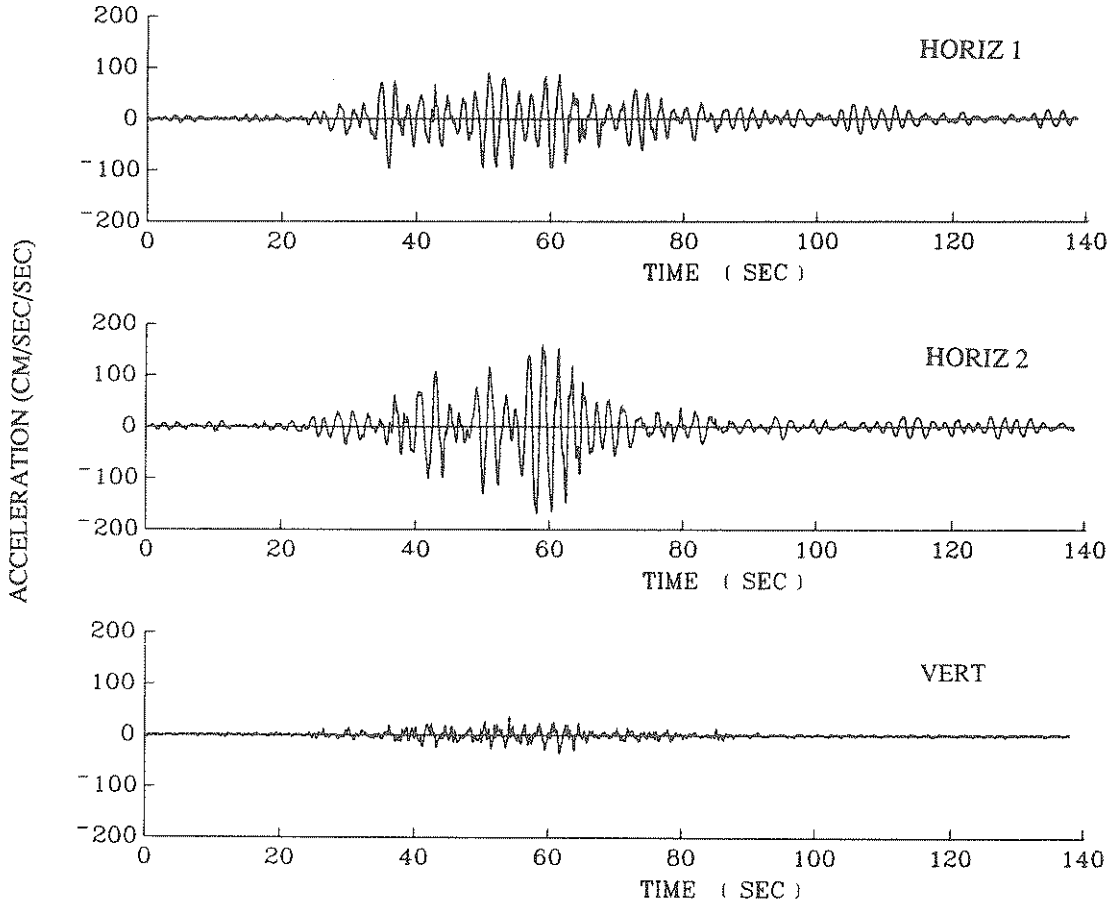


Fig. 5.6 Accelerogram components recorded at station SCTI

components, k_3. Figure 5.7 shows an explanation of the phenomenon. The incoming waves are refracted more at the bedrock-soil interface for soft soil deposits than they are for stiffer deposits. This results in the shear waves oscillating predominantly in the horizontal plane for soft soil deposits and results in low values of k_3. Thus the value of k_3 is small for soft deposits and large for stiffer deposits and rock. In Fig. 5.9 the values of k_3 are plotted on a map of Mexico City. In the hill zone the values of k_3 are largest, while in the lake zone where the soil is soft they are small. The two stations located very close together, and thus possessing similar site conditions, at the Universidad Nacional Autonoma de Mexico (CUMV and CUIP), have similar k_3 values.

Because k_3 was found to be a measure of the site conditions, all of the parameters calculated were regressed versus k_3. The functional relations are shown in Table 5.2. The standard errors of the regression coefficients and the coefficient of determination, r^2, are also shown. The value of α, which measures the maximum of the fitted standard deviation envelope, was found to be highly correlated with k_3. This relationship is shown in Fig 5.8. Because of the amplification due to the resonance of soft soil deposits, as measured by low values of k_3, the value of α is highest when k_3 is low and falls off rapidly as the value of k_3 increases. The one exception to this trend is station TLHD. No explanation for this difference has been found, although it may be caused by irregular geology at the site.

The other parameters which are strongly affected by k_3 are those which describe the frequency content of the variance and frequency stabilized time history. The frequency at which the maximum of the Fourier spectrum occurs, f_{max}, was found to increase linearly with k_3 as shown in Fig 5.10. Thus for stiffer soil deposits the predominant frequency was measured to be higher. Because the resonance frequency of a soil deposit increases with the stiffness of the soil layer, the relationship between f_{max} and k_3 follows physical theory.

The second measure of frequency content is the maximum value of the ARMA Fourier spectrum, $F(f_{max})$, of the stabilized series. A high value of $F(f_{max})$ indicates a high, narrow peak and a low value indicates a lower, broader peak and thus a wider range of frequencies. Although the value of $F(f_{max})$ shows little dependence upon k_3 in Fig. 5.11b for the vertical components, it was found to decrease with increasing

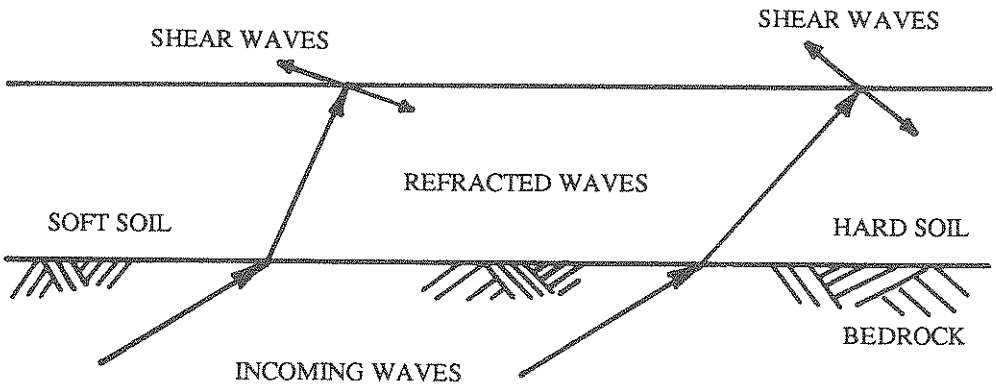


Fig. 5.7 Refraction of shear waves at bedrock-soil interface.

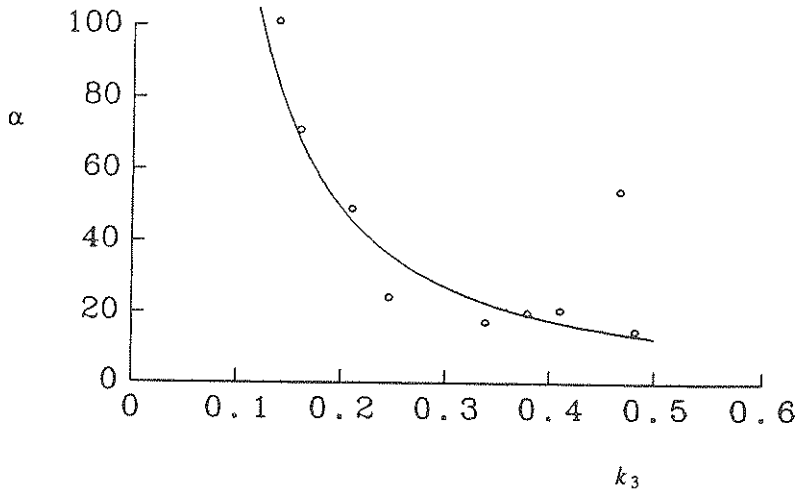


Fig. 5.8 Relationship between the maximum of the fitted standard deviation envelope, α, and k_3 for Mexico City.

Table 5.2 Parametric Relations for Michoacan Earthquake (Mexico City)

H = Horizontal Components

V = Vertical Components

$\ln(\alpha) = (1.52\pm0.26) - (1.48\pm0.19)\ln(k_3)$		$r^2 = 0.91$
$\tau = 22.3\pm10.9$		
$k_1 = (2.54\pm2.18) + (0.22\pm0.044)\alpha$		$r^2 = 0.79$
$t_{max} = 21.0\pm8.7$		
$r_2 = 1.46\pm0.44$		
$b_2 = 8.24\pm10.9$		
$r_3 = 1.16\pm0.18$		
$b_3 = 17.0\pm23.1$		
$k_3 = f$ (*geographic location*)		
$f_{max} = (0.233\pm0.21) + (1.26\pm0.62)k_3$	H	$r^2 = 0.20$
$f_{max} = (-0.022\pm0.08) + (1.84\pm0.87)k_3$	V	$r^2 = 0.39$
$F(f_{max}) = (1.410\pm0.17) - (2.14\pm0.49)k_3$	H	$r^2 = 0.54$
$F(f_{max}) = 0.592\pm0.15$	V	
$\phi_1 + \phi_2 + \phi_3 = 0.9988\pm0.00213$		

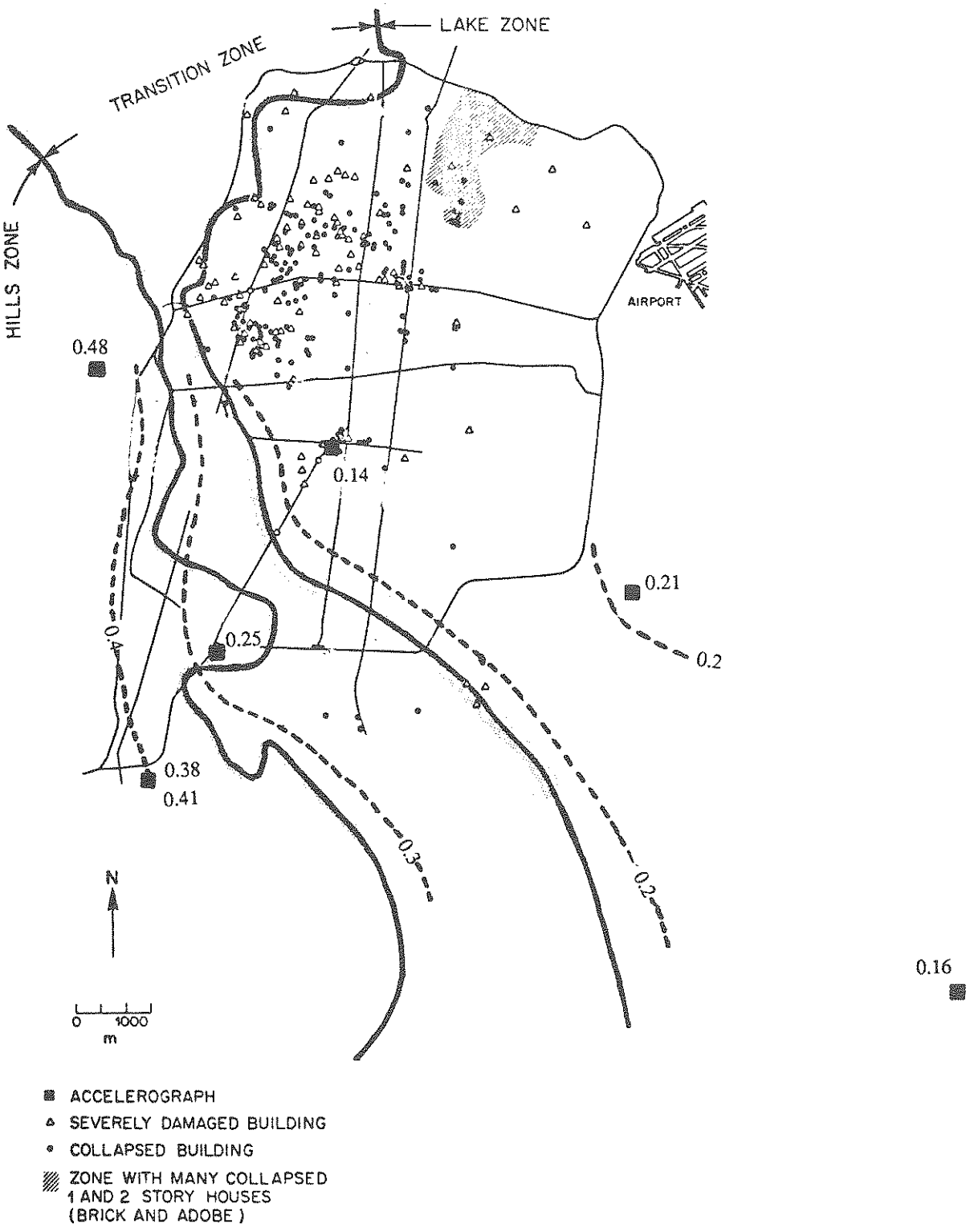


Fig. 5.9 Values of k_3 calculated for Mexico City.

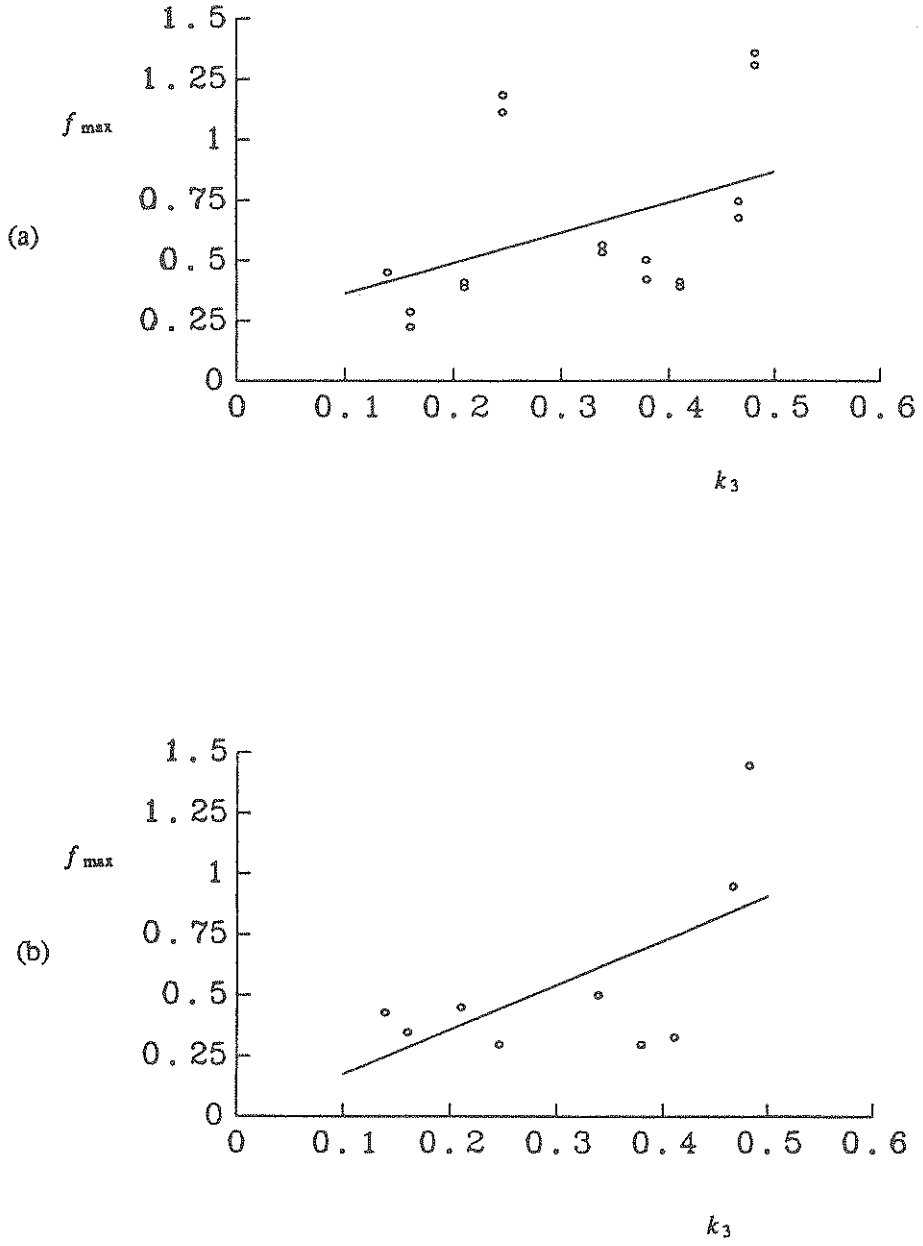


Fig. 5.10 Relationship between the dominant frequency of the ARMA Fourier spectrum of the stabilized acceleration series for (a) horizontal components and (b) vertical components recorded in Mexico City.

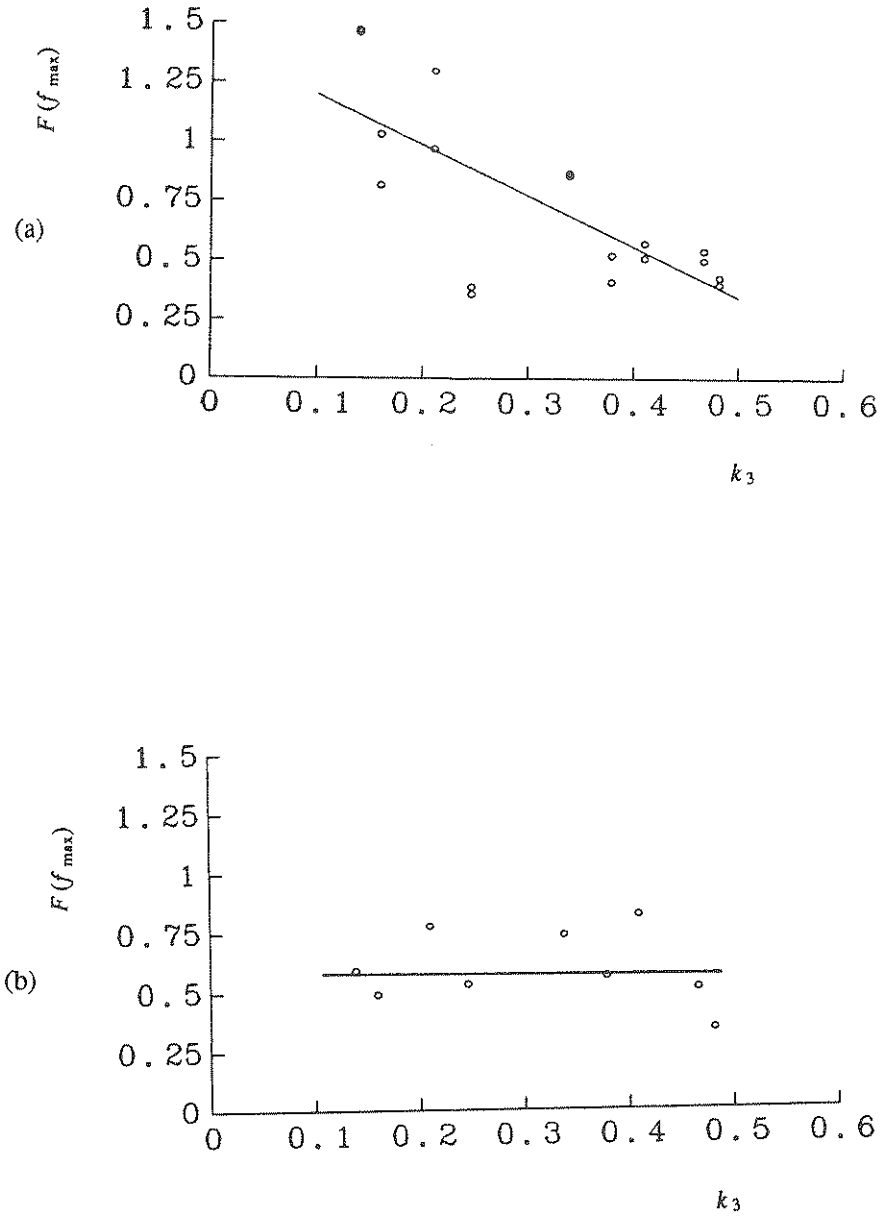


Fig. 5.11 Relationship between the maximum of the ARMA Fourier spectrum of the stabilized acceleration series for (a) horizontal components and (b) vertical components recorded in Mexico City.

values of k_3 in Fig 5.11a for the horizontal components studied. This indicates that sites located on stiff soils and rock have a broader range of frequency content than soft soil deposits which have a narrow peak. Because the amplification of the ground motion near the resonance frequency of the soil deposit increased only the frequency components around 0.5 Hz, a much more narrow peak for low values of k_3 agrees with physical theory.

5.4. Guerrero Array Analysis

Records from eleven stations in the Guerrero array listed in Table 5.1 were analyzed. Because these stations were located on competent rock sites, the differences among the accelerograms may be attributed to the effect of varying distance. Included in the analysis of the Guerrero array are the two stations in the Mexico City region located on rock, stations CUMV and CUIP.

All of the modelling parameters calculated for the Guerrero array were regressed versus distance. The results are shown in Table 5.3. The value of α attenuated with distance is shown in Fig 5.12. The values recorded at rock sites in Mexico City were only about one quarter of those recorded near the fault zone. The value of τ measuring the duration of strong shaking was found to increase with distance. This may be explained by the spreading of the arrival time of waves travelling at different velocities because of the longer distance they must travel.

The value of k_3 was found to be higher with a narrower range of values than k_3 calculated for Mexico City. Figure 5.13.a illustrates this behavior. Using an ANOVA analysis, it was found that the mean level of k_3 recorded on rock varied from the mean level of k_3 recorded on softer soils at the 95% confidence level. Recording uniformly higher values of k_3 for rock provides further evidence that it is a good measure of the soil type.

Figs. 5.14a, 5.14b, 5.15a, and 5.15b show the relationship between f_{max} and $F(f_{max})$ and epicentral distance. The decrease of f_{max} with distance and the increase of $F(f_{max})$ with distance both reflect the greater attenuation rate of high frequency with distance as reported by Trifunac and Brady (1975). When the high frequency content decreases, the dominant frequency is lowered and the peak of the Fourier spectrum

Table 5.3 Parametric Relations for Michoacan Earthquake (rock sites)

H = Horizontal Components

V = Vertical Components

d = epicentral distance (km)

$\ln(\alpha) = (14.98\pm2.37) - (2.08\pm0.41)\ ln(d)$		$r^2 = 0.71$
$\ln(\tau) = (0.176\pm1.35) + (0.00503\pm0.0044)d$		$r^2 = 0.39$
$k_1 = (2.16\pm1.66) + (0.268\pm0.040)\alpha$		$r^2 = 0.78$
$t_{max} = 11.8\pm6.4$		
$r_2 = 1.17\pm0.11$		
$b_2 = 4.14\pm3.79$		
$r_3 = 1.16\pm0.27$		
$b_3 = 3.81\pm1.85$		
$k_3 = 0.487\pm0.042$		
$f_{max} = (5.37\pm1.44) - (0.011\pm0.0047)d$	H	$r^2 = 0.17$
$f_{max} = (6.02\pm1.48) - (0.014\pm0.0048)d$	V	$r^2 = 0.42$
$F(f_{max}) = (-0.0029\pm0.094) + (0.000984\pm0.00031)d$	H	$r^2 = 0.30$
$F(f_{max}) = (-0.0050\pm0.13) + (0.000970\pm0.00042)d$	V	$r^2 = 0.32$
$\phi_1 + \phi_2 + \phi_3 = (0.999\pm0.0031) - (0.0051\pm0.0010)f_{max}$		$r^2 = 0.44$

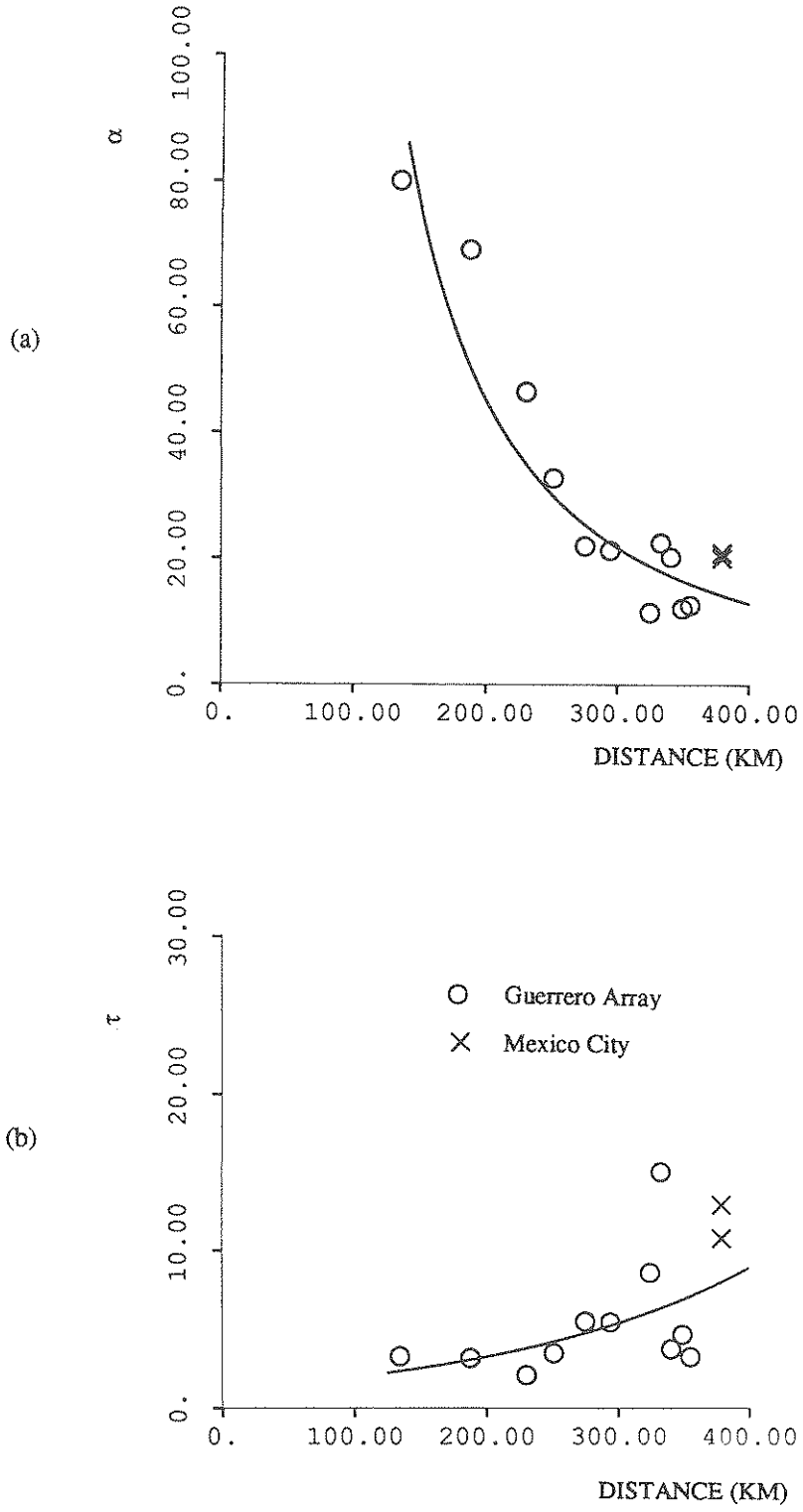


Fig. 5.12 Relationship between (a) α and epicentral distance and (b) τ and epicentral distance for rock sites.

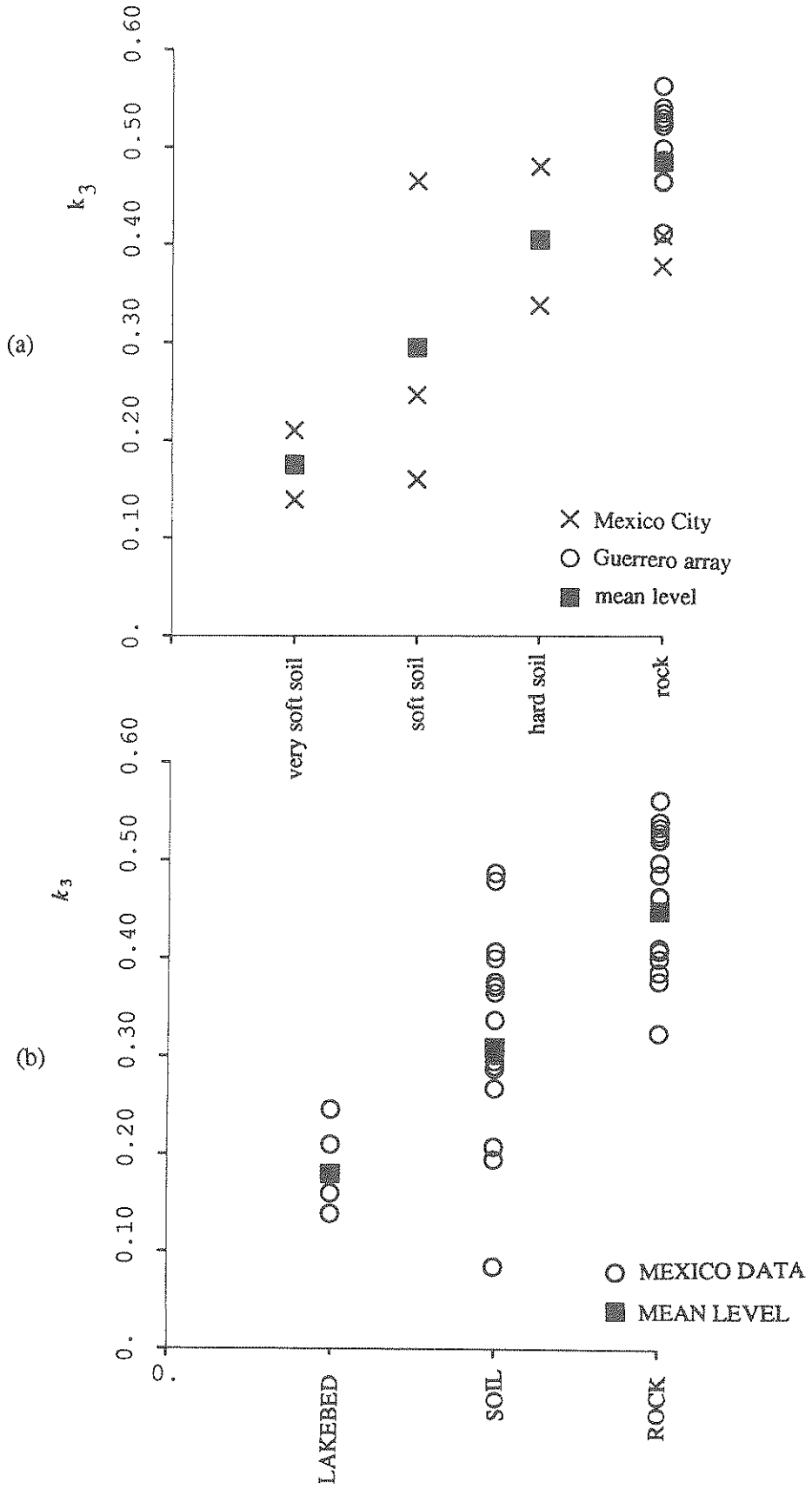
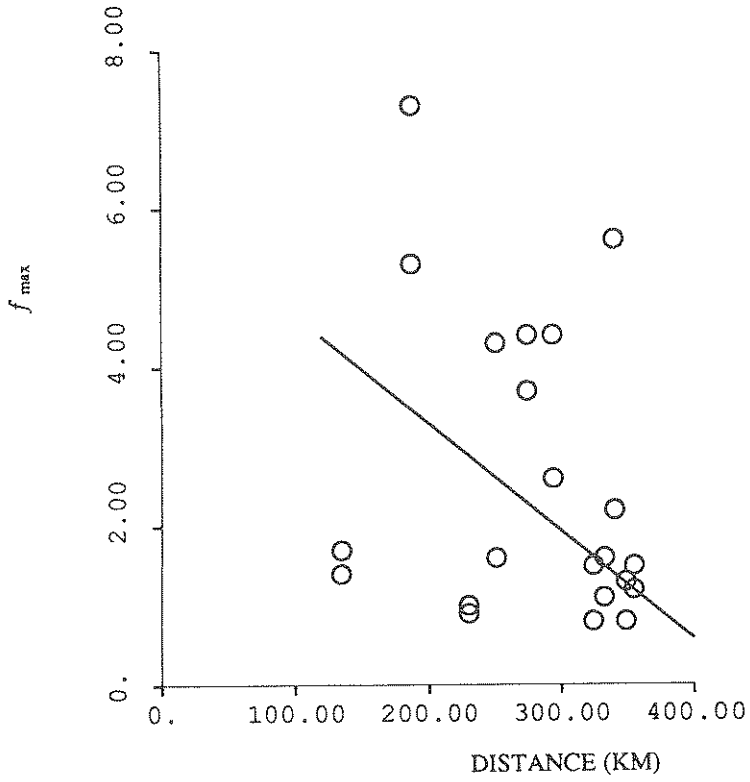


Fig. 5.13 Relationship between k_3 and soil type for (a) the September 19, 1985 Michoacan earthquake and (b) all Mexico regions.

(a)



(b)

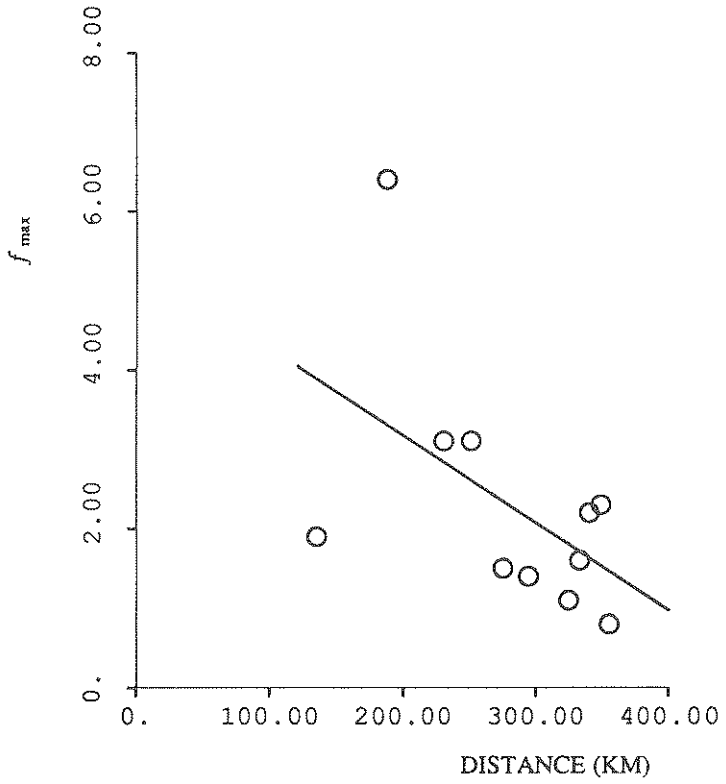
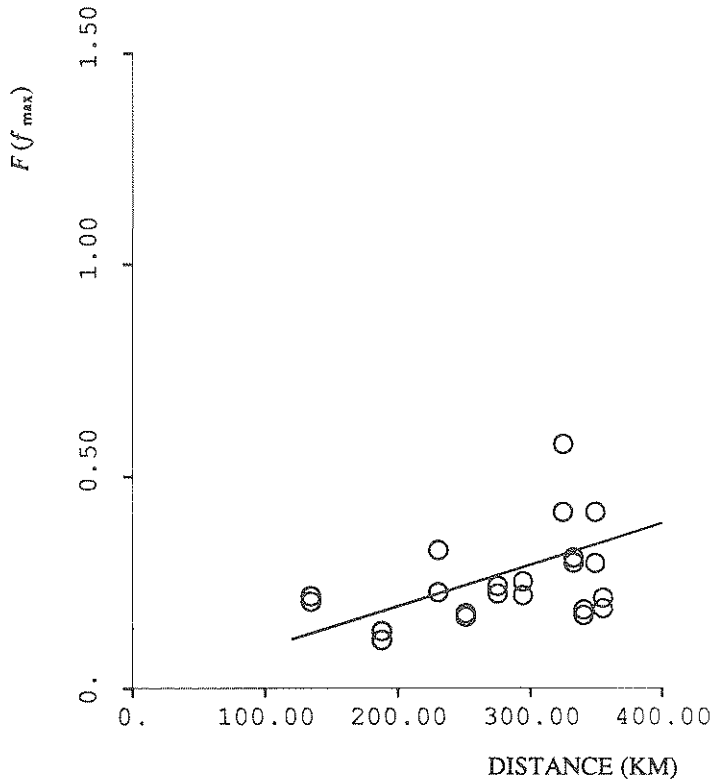


Fig. 5.14 Relationship between the dominant frequency of the ARMA Fourier spectrum of the stabilized acceleration series for (a) horizontal components and (b) vertical components recorded on rock sites.

(a)



(b)

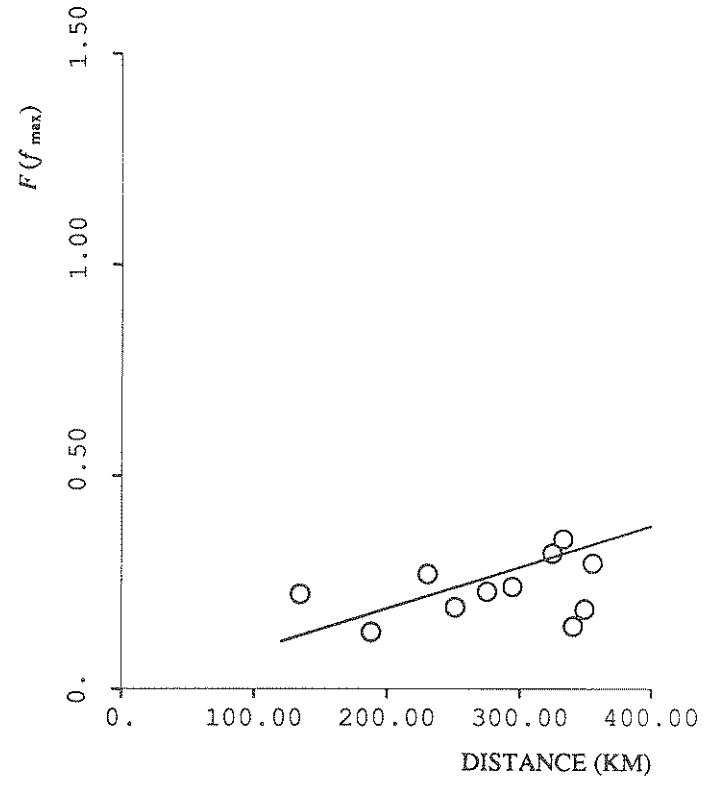


Fig. 5.15 Relationship between the maximum of the ARMA Fourier spectrum of the stabilized acceleration series for (a) horizontal components and (b) vertical components recorded on rock sites.

becomes narrower.

5.5. Mexico Analysis

In addition to the September 19, 1985 Michoacan earthquake, other earthquakes have been recorded throughout Mexico. Although many of these records are missing at least one component and thus are not suited for analysis by the multivariate modelling procedure, enough records were available to increase the database upon which the parametric relations were based. To form a comprehensive model for Mexico, the modelling parameters calculated from all the records in Table 5.1 were analyzed. The resulting modelling parameters were then related to distance, magnitude, and site conditions as measured by k_3.

The parametric equations for the modelling parameters are given in Table 5.4. As shown in Fig. 5.13b, k_3 was found to be a function of site conditions. As expected from the behavior of seismic waves refracted at the bedrock-soil interface shown in Fig. 5.7, k_3 is highest for rock sites and decreases for softer soil conditions. In Chapter 6, k_3 will be shown to be a function of shear wave velocity based on data from California, Mexico, and Taiwan.

Figure 5.16 shows the functional relations between the parameters describing the standard deviation envelope function and physical variables. Figures 5.16a and 5.16b compare the recorded data with the value predicted from the multiple regression analysis. The value of α was found to be a function of distance, magnitude, and k_3. The functional form of this relationship is similar to that used for α in the California study. The effect of k_3 in this function indicates that α is larger for soft soil deposits.

The duration of strong shaking, τ, was found to vary with magnitude and k_3. The increase of τ with magnitude can be explained by the relationship between magnitude and length of fault rupture. When the earthquake magnitude is higher, the length of the fault is longer. Thus the site receives waves from one part of the fault before waves from the rest of the fault can reach the site. Since the arrival of the waves is spread out in time, the duration of the strong shaking is increased. The decrease of τ with increasing k_3 indicates the shaking dies out more slowly for sites with soft soil

Tabel 5.4 Parametric Relations for Mexico

d = epicentral distance (km)

M = Magnitude

$$\alpha = (-1.3\pm13.1) + (13.24\pm2.22)\ln\left(\frac{10^M}{d^3 k_3^3}\right) \qquad r^2 = 0.49$$

$$\ln(\tau) = (0.512\pm0.75) + (0.392\pm0.11)M - (3.63\pm1.17)k_3 \qquad r^2 = 0.34$$

$$k_1 = (2.04\pm2.0) + (0.247\pm0.023)\alpha \qquad r^2 = 0.74$$

$$\ln(t_{max}) = (1.249\pm0.22) + (0.00377\pm0.00086)d \qquad r^2 = 0.34$$

$$r_2 = 1.16\pm0.37$$

$$b_2 = 3.29\pm6.76$$

$$r_3 = 1.60\pm1.48$$

$$b_3 = 5.43\pm12.6$$

$$k_3 = f\ (soil\ type)$$

$$\ln(f_{max}) = (2.31\pm0.51) + (3.14\pm0.58)k_3 - (0.321\pm0.09)M - (0.00262\pm0.0027)d \qquad r^2 = 0.54$$

$$\ln[F(f_{max})] = (-1.08\pm0.11) - (0.0031\pm0.00023)d - (2.25\pm0.29)k_3 \qquad r^2 = 0.64$$

$$\phi_1 + \phi_2 + \phi_3 = (1.010\pm0.0081) - (0.0159\pm0.0019)f_{max} \qquad r^2 = 0.38$$

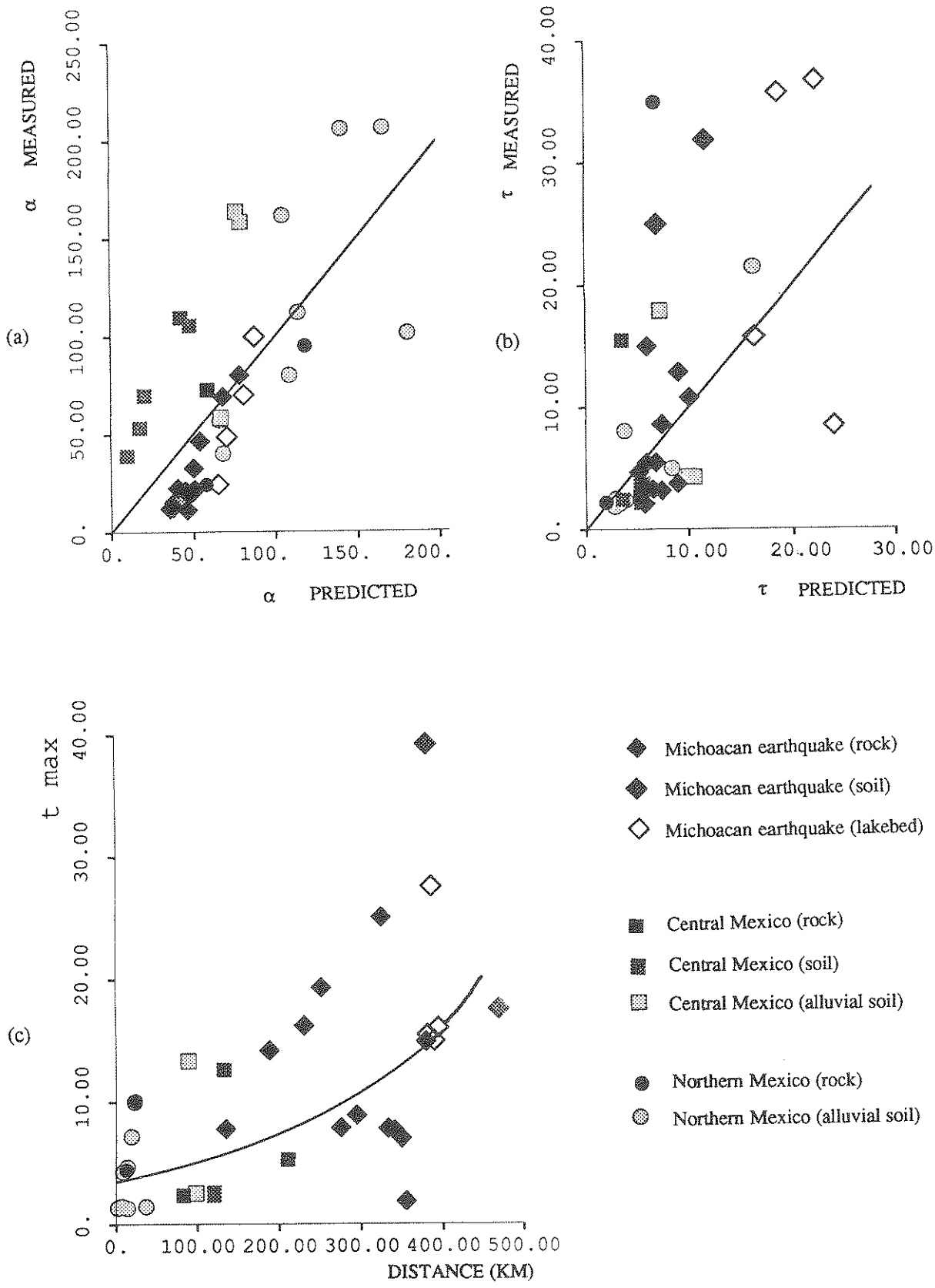


Fig. 5.16 Comparison of (a) α measured and predicted and (b) τ measured and predicted and (c) relationship between t_{max} and epicentral distance for all Mexico regions studied.

deposits.

The time for the standard deviation envelope to reach a maximum, t_{max}, was found to vary with distance. As the distance between the source and the recording site is increased, the difference in the time of arrival of the P-waves which trigger the accelerograph and the S-waves and surface waves increases; thus t_{max} increases.

In Figure 5.17 the parameters describing the Fourier spectrum are plotted. The form of these functions indicates broad peaks with much high frequency content for sites near weak earthquakes. As the epicentral distance or the strength of the earthquake increases, the amount of high frequency that is filtered out also increases. The result is narrow peaks and lower dominant frequencies. Research by Terashima (1968) has found similar results for the effect of earthquake magnitude on frequency content. The dependence upon distance results from the more rapid attenuation of high frequency waves with distance (Trifunac and Brady (1975)). The effect of k_3 indicates that softer deposits amplify the lower frequencies, thus lowering the dominant frequency and increasing the narrowness of the peak of the spectrum.

5.6. Spatial Correlation of Time Series

In Chapter 4 it was found that the cross-correlation among the three accelerogram components is very small. In this section the cross-correlation between components located at two different sites is investigated. In Mexico City there are six recording stations located close enough together to attempt to measure their cross-correlation. These stations are CUMV, CUIP, TACY, SXPU, CDAF, and SCTI.

To calculate the cross-correlation between stations, the variance stabilized accelerograms were used. No frequency correction was used so that all of the time series were measured at the same time increment. The cross-correlation between components at different stations was then measured at time lags between -20 to 20 seconds. The results are shown in Figs. 5.18 and 5.19.

The only pair of stations for which the time lag to maximize the cross-correlation was the same for each component were stations CUMV and CUIP. The cross-correlation between these stations is shown in Fig. 5.18a. Both stations are located on the Ciudad

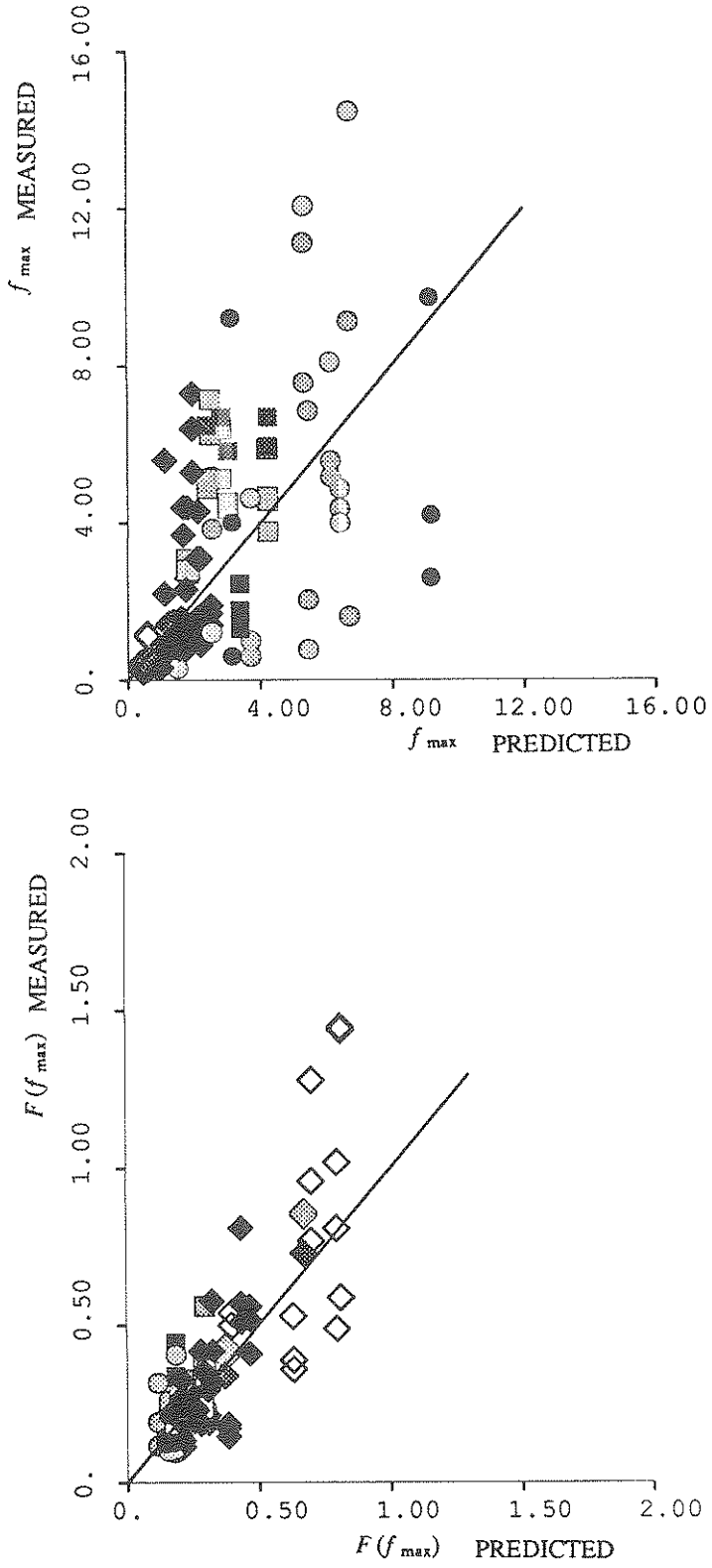


Fig. 5.17 Comparison of the measured and predicted values of (a) the dominant frequency and (b) the maximum of the ARMA Fourier spectrum of the stabilized acceleration series for all Mexico regions studied.

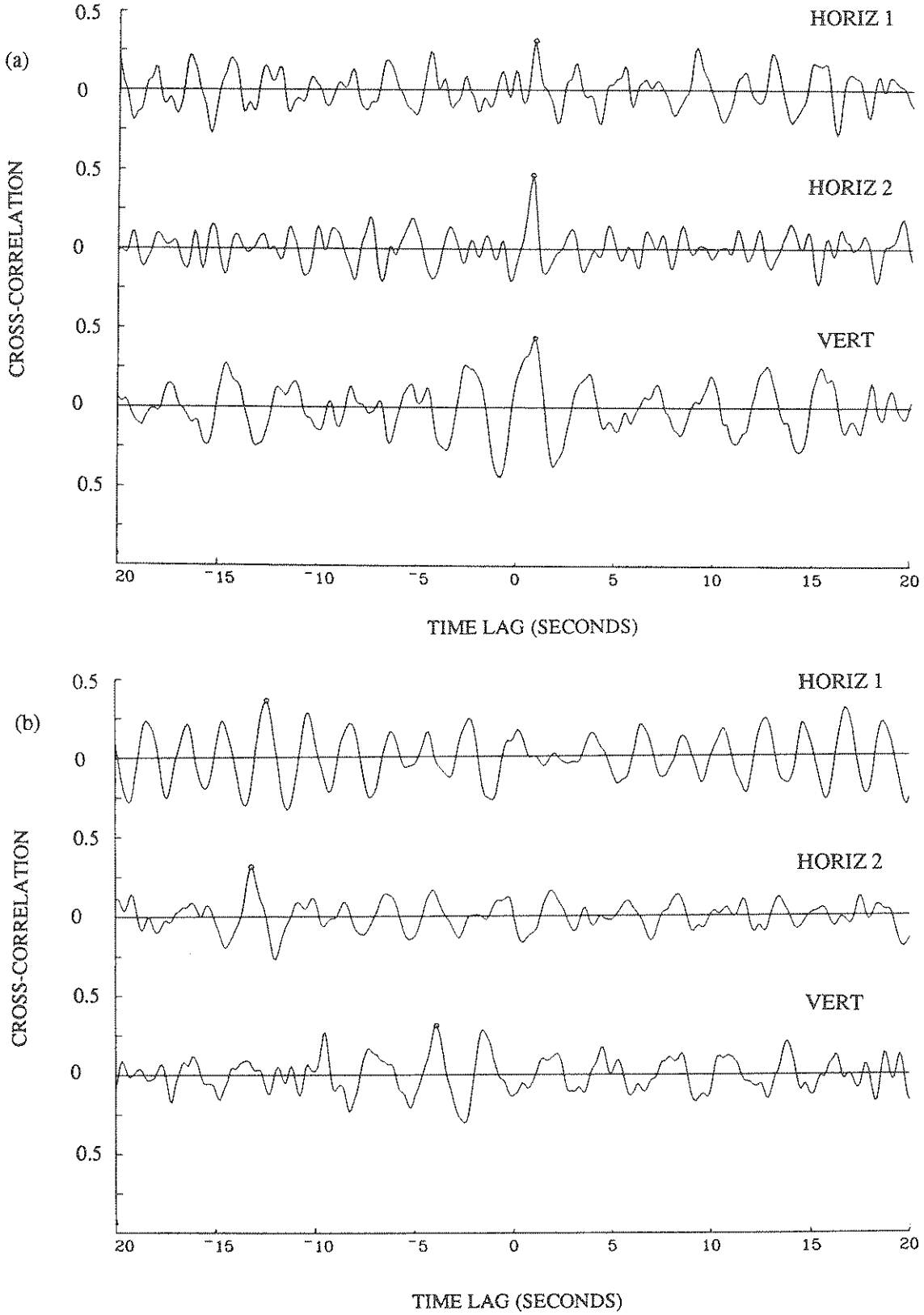


Fig. 5.18 Cross-correlation between (a) stations CUMV and CUIP and (b) stations CUMV and CDAF.

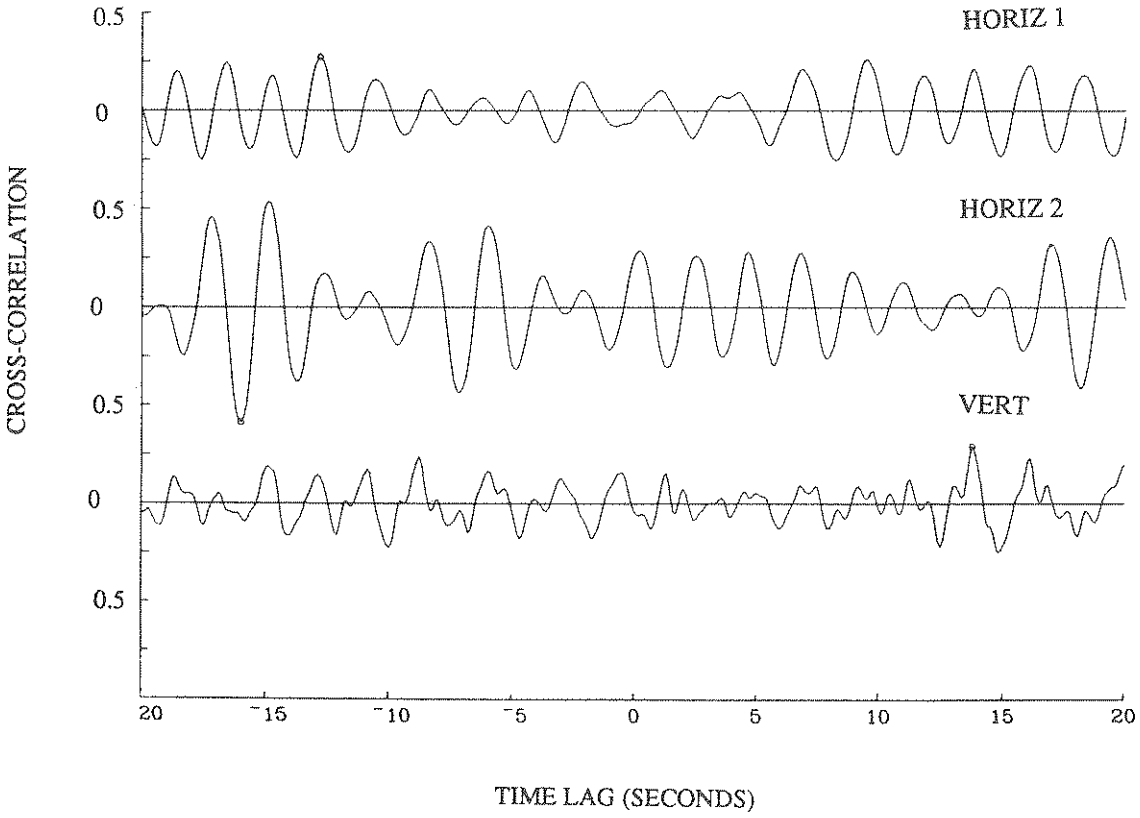


Fig. 5.19 Cross-correlation between stations CDAF and SCTI.

Universitaria campus separated by a distance of less than 1000 meters. This distance was much shorter than the other separation distances which ranged from 2500 to 9000 meters. Thus it is correct to assume that the ground motion of the recording stations located in Mexico City is independent with the exception of stations CUMV and CUIP. This result is in agreement with earlier research by Harichandran and Vanmarcke (1984) who found that in the Smart-1 array in Taiwan the cross-correlation among series was high for separation distances up to 1000 meters.

5.7. Simulating the Original Records

At this point it is possible to generate simulations using the procedure outlined in Section 4.4. The parametric relations in Table 5.2 were used to calculate parameters from which simulations were generated for four sites located in Mexico City. These sites are UNAM (located on rock), TACY (located on stiff soil), TLHB (located on soft soil), and SCTI (located on very soft soil). In Figures 5.20 to 5.30 the results for simulations generated for station UNAM are shown. The original accelerogram record is shown in Fig. 5.20. In Fig. 5.21 a simulation generated from modelling parameters calculated from Table 5.2 is shown. The differences between the original record and the simulation are due to the error in the modelling procedure and in the parametric relations. Good agreement in the duration and intensity of strong shaking exists between the two records.

From the standard errors of the regression coefficients listed in Table 5.2, confidence intervals for the modelling parameters can be calculated from

$$\sigma = [\sum_i \sigma_{\eta_i}^2 \eta_i^2]^{\frac{1}{2}} \tag{5.7.1}$$

where

σ = standard error of the modelling parameter

η_i = the independent variable

σ_{η_i} = the standard error for the independent variable i .

For example, for k_3 equal to 0.38 the mean value of f_{max} for the horizontal components can be calculated as

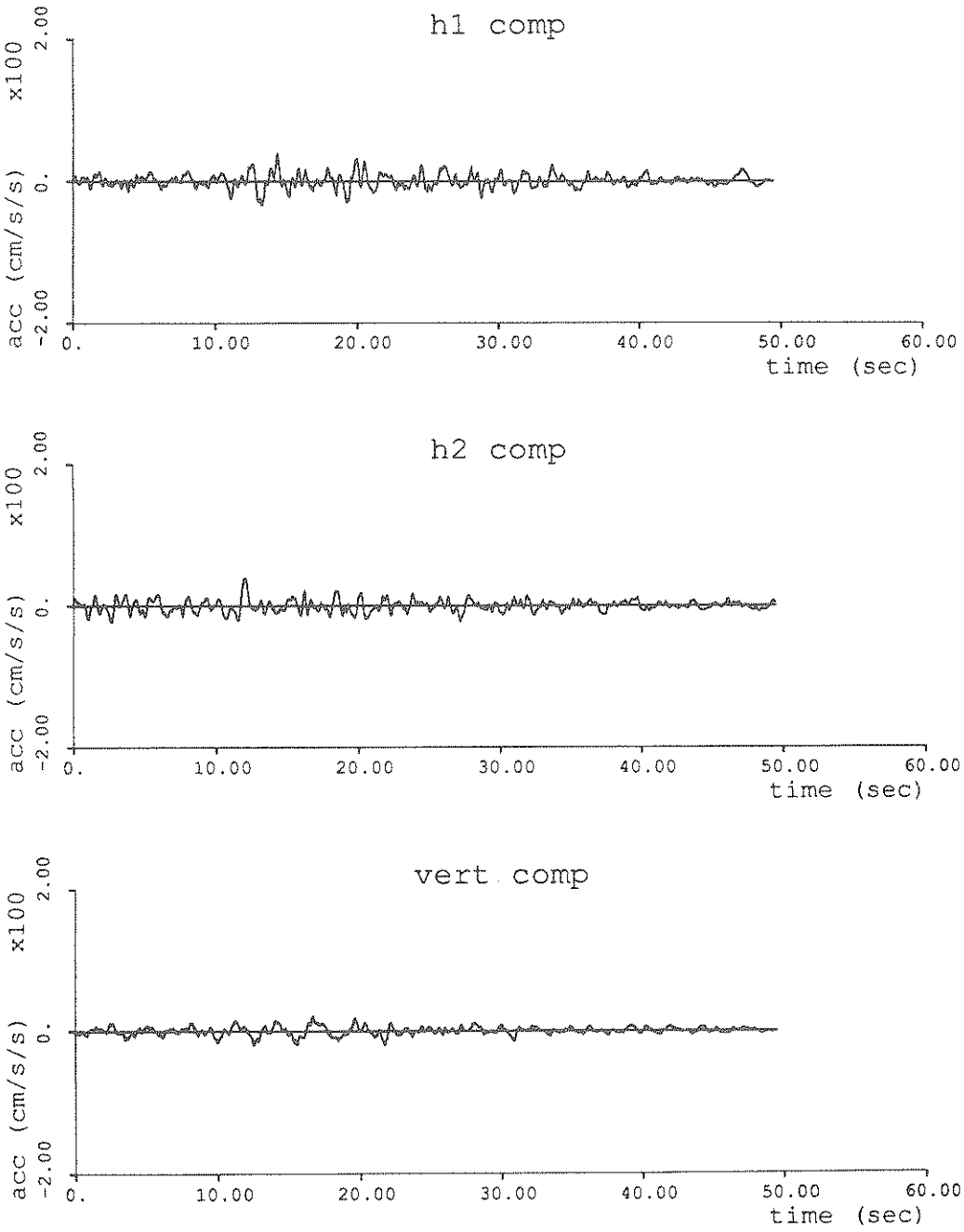


Fig. 5.20 Shortened accelerogram recorded at station UNAM.

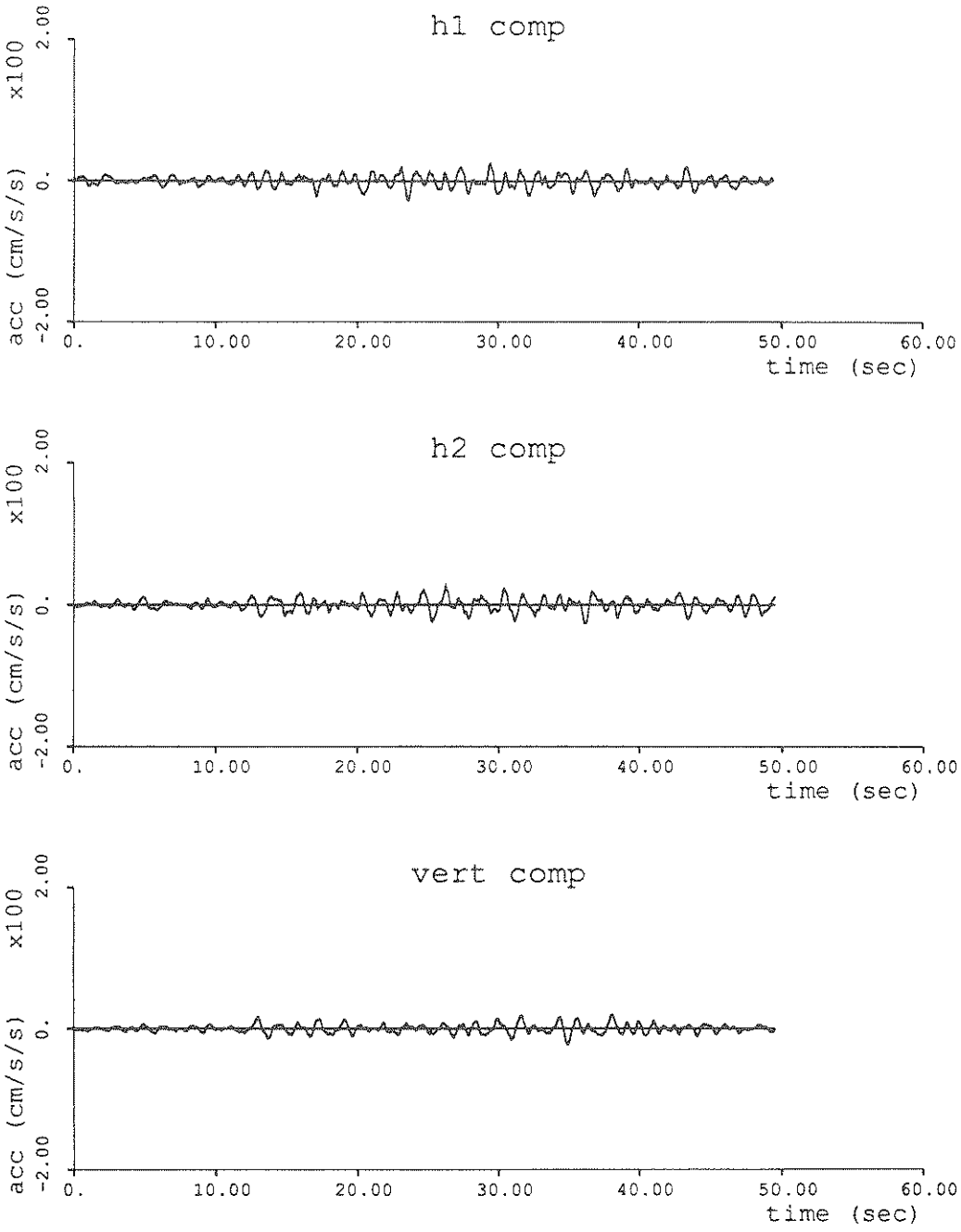


Fig. 5.21 Simulation of UNAM accelerogram using the modelling parameters calculated from Table 5.2.

$$\bar{f}_{max} = 0.233 + 1.26 \times 0.38 \qquad (5.7.2)$$

$$= 0.71 .$$

The standard error of f_{max} is then

$$\sigma = [0.21^2 + (0.62 \times 0.38)^2]^{\frac{1}{2}} \qquad (5.7.3)$$

$$= 0.31 .$$

The 95% confidence interval for f_{max} is then

$$f_{max} = 0.71 \pm 2(0.31) . \qquad (5.7.4)$$

Thus f_{max} for station UNAM can be expected to fall between 0.09 Hz and 1.33 Hz with a mean of 0.71 Hz. Figures 5.22 and 5.23 show simulations for UNAM using the upper and lower limit of α. In Fig. 5.24 a comparison is made of the horizontal Fourier and response spectra of the original records, the mean simulation, and the simulations for α equal to two standard errors above the mean value predicted.

In Fig. 5.25 a simulation with the highest dominant frequency expected is shown. Fig. 5.26 shows the shift in the dominant frequency in both the Fourier and response spectra. Similarly, in Figs. 5.27 and 5.28 a simulation with $F(f_{max})$ equal to the maximum expected value is shown.

To use the simulations as input motion for structural models, an envelope of the most damaging motion is useful in addition to the mean expected motion. As illustrated in the preceding paragraphs, it is quite simple to produce simulations with varying statistical properties. By generating simulations for the maximum and minimum expected value of each modelling parameter, an envelope of the expected ground motion may be generated. Figures 5.29 and 5.30 show the envelopes for the ground motion expected at station UNAM. The envelopes are a combination of the maximum and minimum values of the simulations.

Comparisons between the original accelerograms and the simulated accelerograms for stations TACY, TLHB, and SCTI are shown in Figs. 5.31 to 5.39. By comparing the original acceleration time histories with the simulated ones, it is apparent that the simulations capture the severity and length of shaking well. The model also captures

the amplification of the horizontal components for the sites located in the lakebed (TLHB and SCTI), while the vertical component remains small. The frequency content is compared using the Fourier spectrum and the response spectrum. Again the model does quite well reproducing the original frequency content. For the stations located in the lakebed region the Fourier spectrum has a narrow peak located at about $f = 0.5\,Hz$. The response spectra also match well with a very strong response above a period of one second and very little response at higher frequencies. Thus the linear response of a structure to the simulated motion will be similar to the response to actual ground motion. In general, the envelopes capture most of the variations from the mean expected motion.

References

Anderson, J. G., P. Bodin, J. N. Brune, J. Prince, S. K. Singh, R. Quaas, M. Onate (1986). "Strong Ground Motion from the Michoacan, Mexico Earthquake," *Science*, Volume 233, 1043-1049.

Harichandran, R. and E. Vanmarcke (1984). "Space-Time Variation of Earthquake Ground Motion", *Massachusetts Institute of Technology Technical Report*, Cambridge, Massachusetts.

Terashima, T. (1968). "Magnitude of Microearthquakes and Spectra of Microearthquake Waves," *Bulletin of the International Institute of Seismology and Earthquake Engineering*, Vol. 5, 31-108.

Trifunac, M. D. and A. G. Brady (1975). "On the Correlation of Seismic Intensity Scales with the Peaks of Recorded Strong Ground Motion," *Bulletin of the Seismological Society of America*, Vol. 65, 307-321.

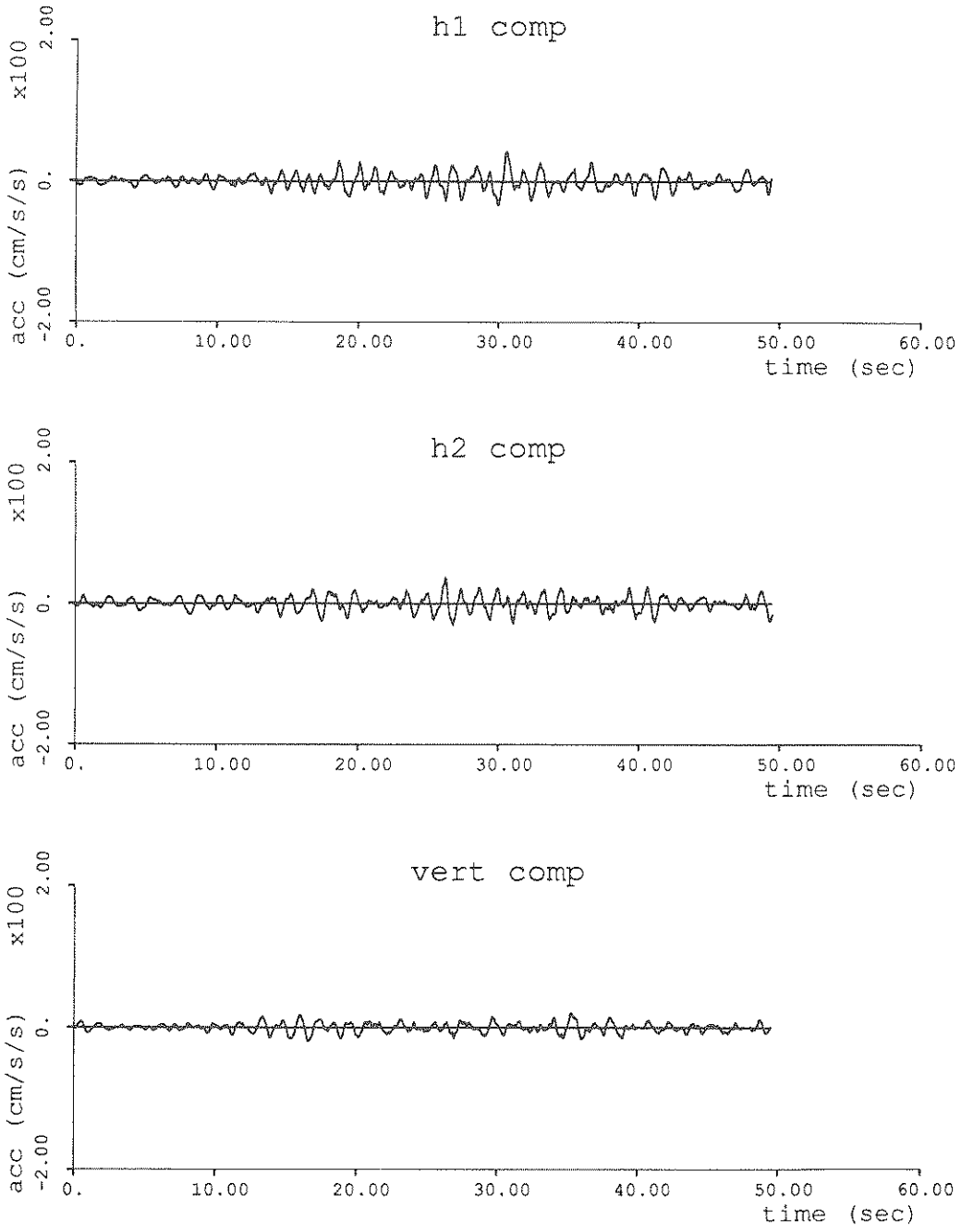


Fig. 5.22 Simulation of UNAM accelerogram using α equal to two standard deviations above the value calculated from Table 5.2.

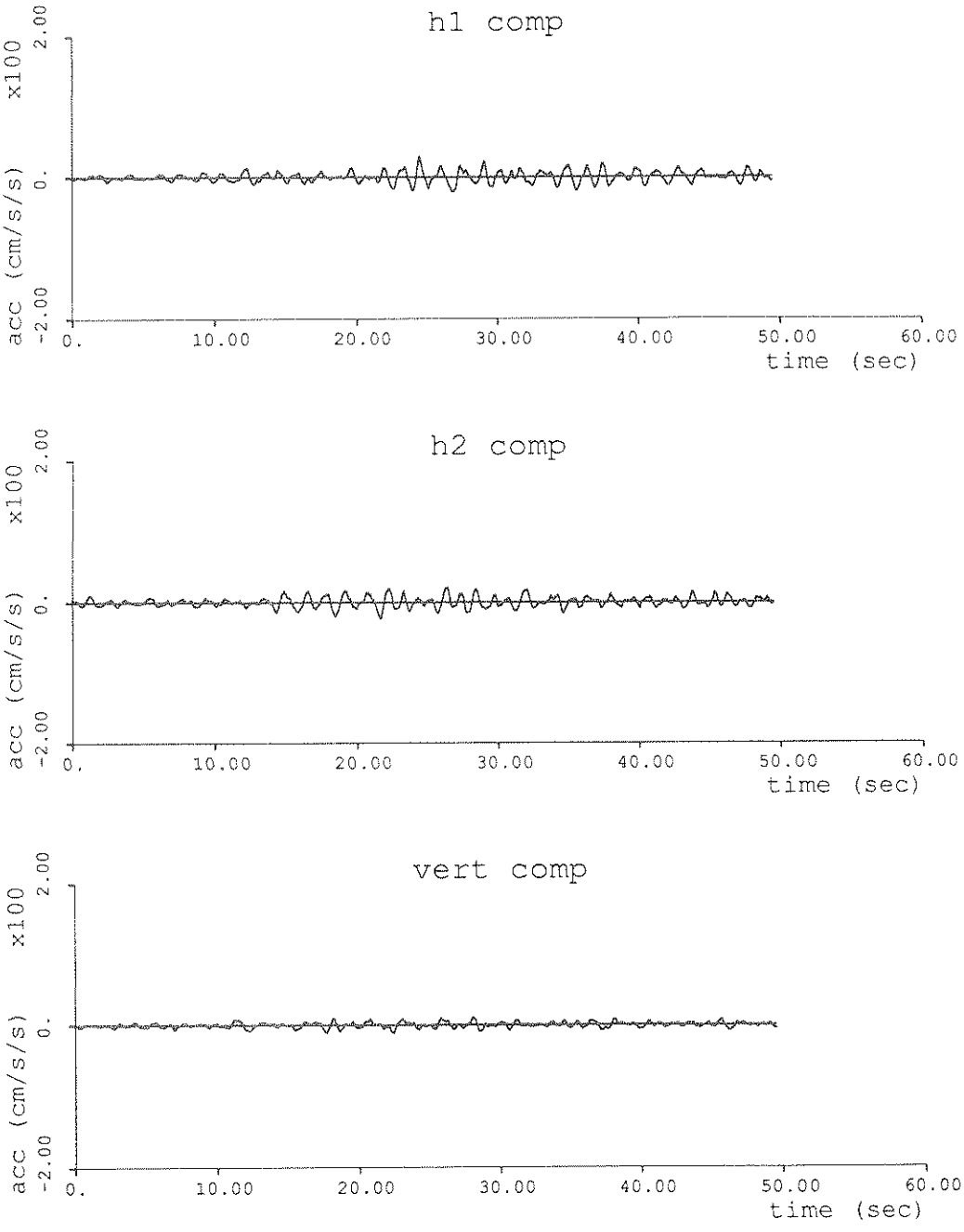


Fig. 5.23 Simulation of UNAM accelerogram using α equal to two standard deviations below the value calculated from Table 5.2.

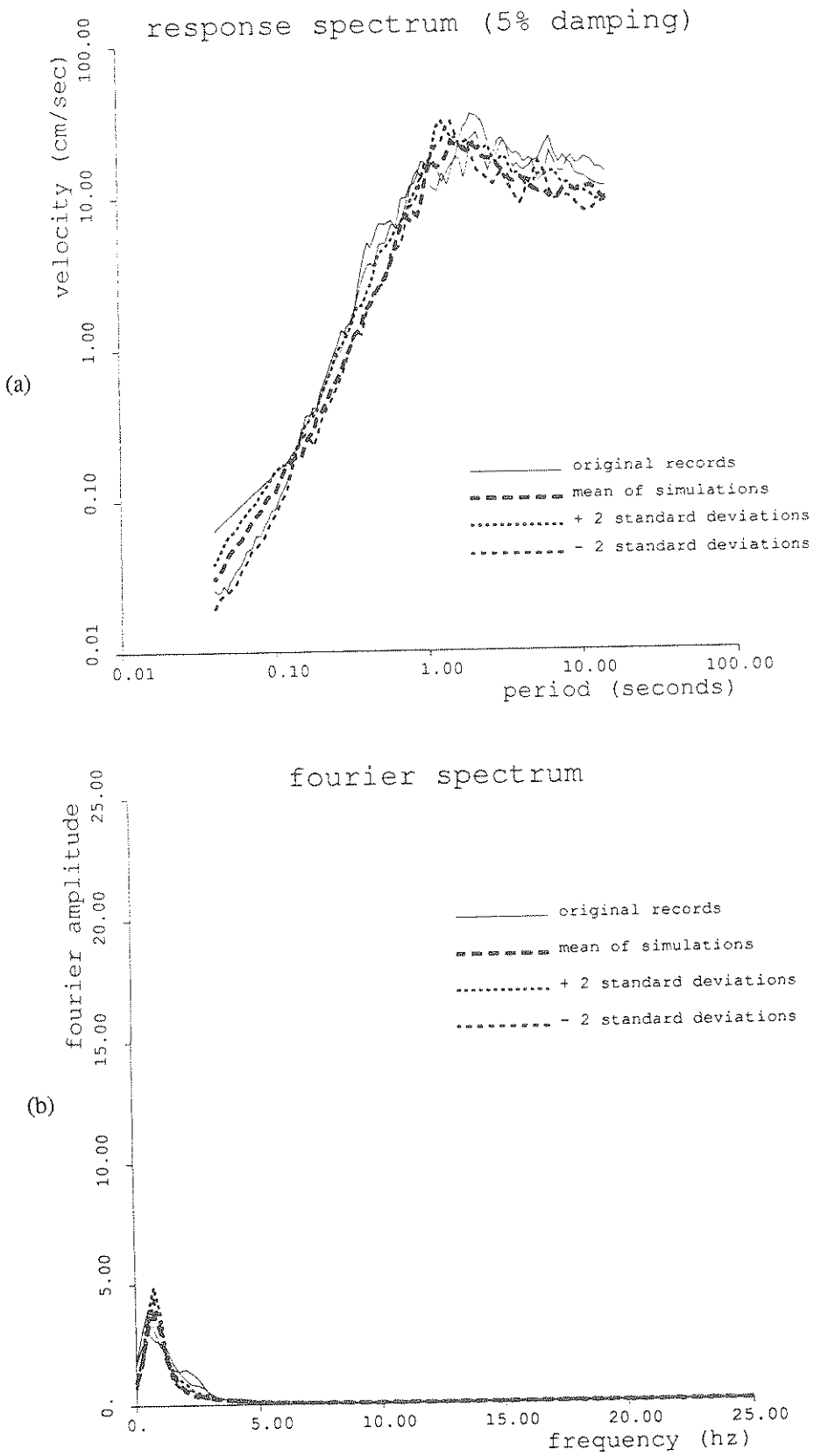


Fig. 5.24 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original UNAM accelerogram, simulations using parameters calculated from Table 5.2, and simulations using α equal to $\pm$ two standard deviations from the value calculated from Table 5.2.

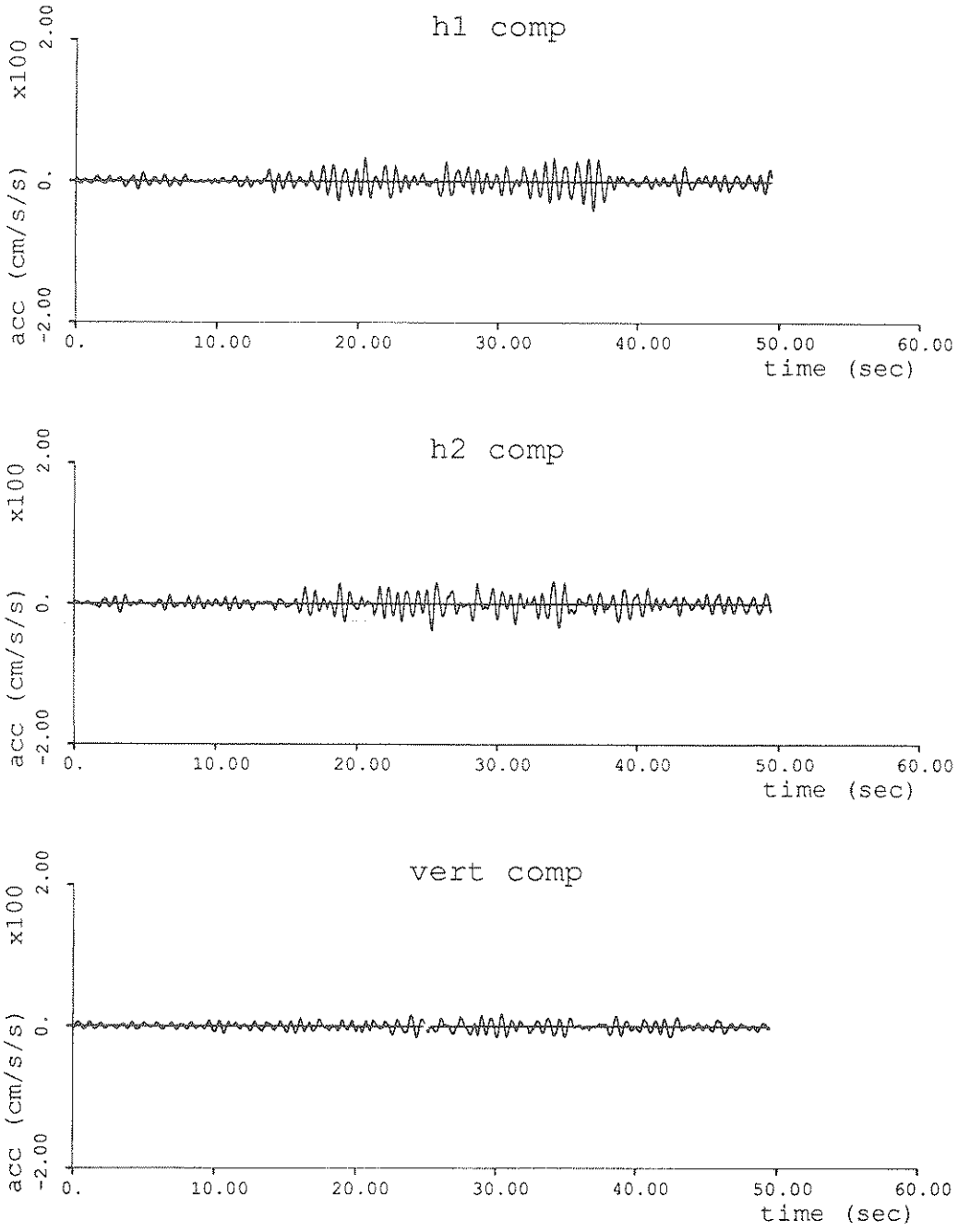
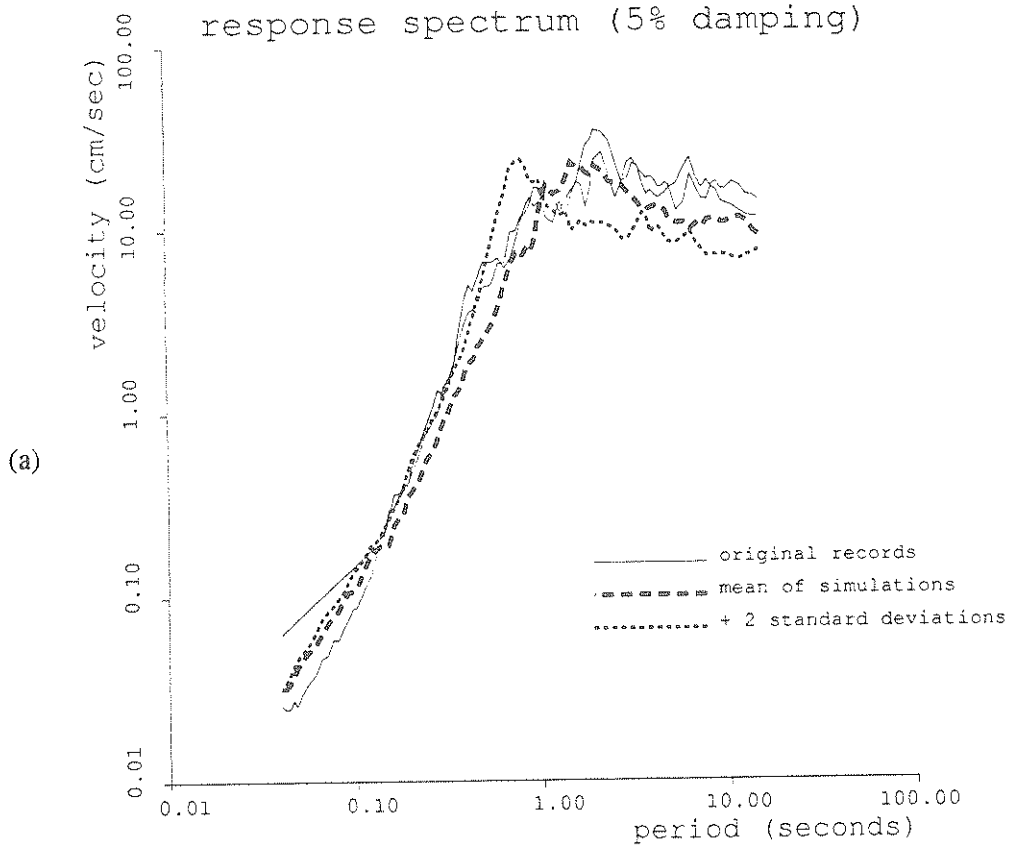
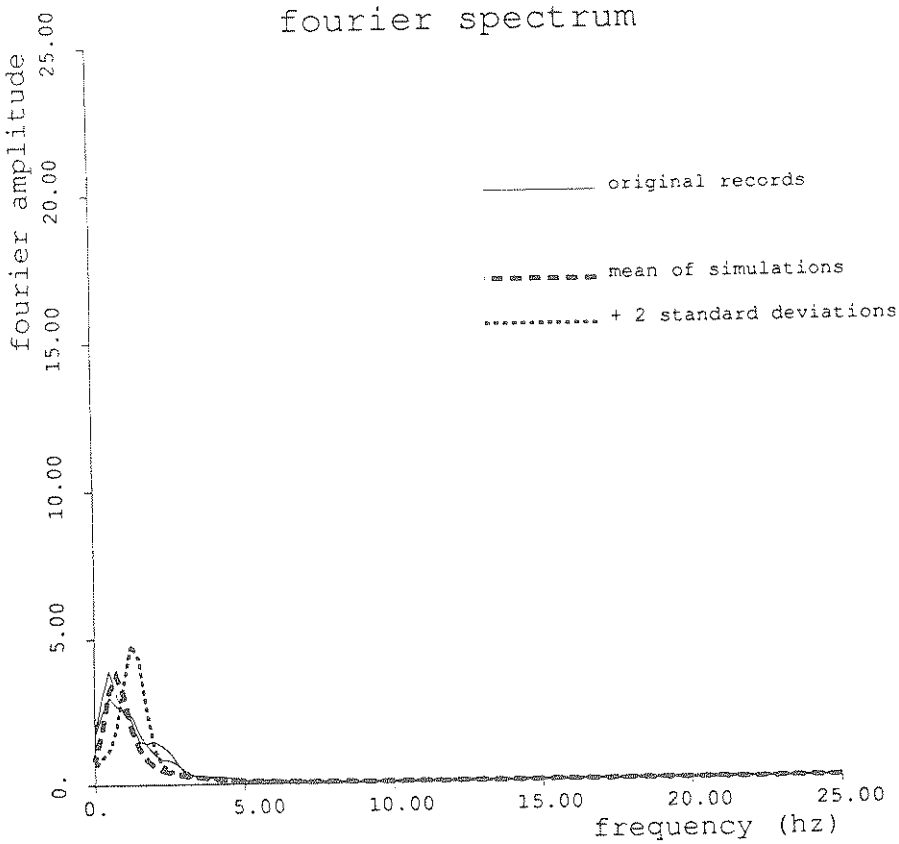


Fig. 5.25 Simulation of UNAM accelerogram using f_{max} equal to two standard deviations above the value calculated from Table 5.2.



(a)



(b)

Fig. 5.26 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original UNAM accelerogram, simulations using parameters calculated from Table 5.2, and simulations using f_{max} equal to two standard deviations above the value calculated from Table 5.2.

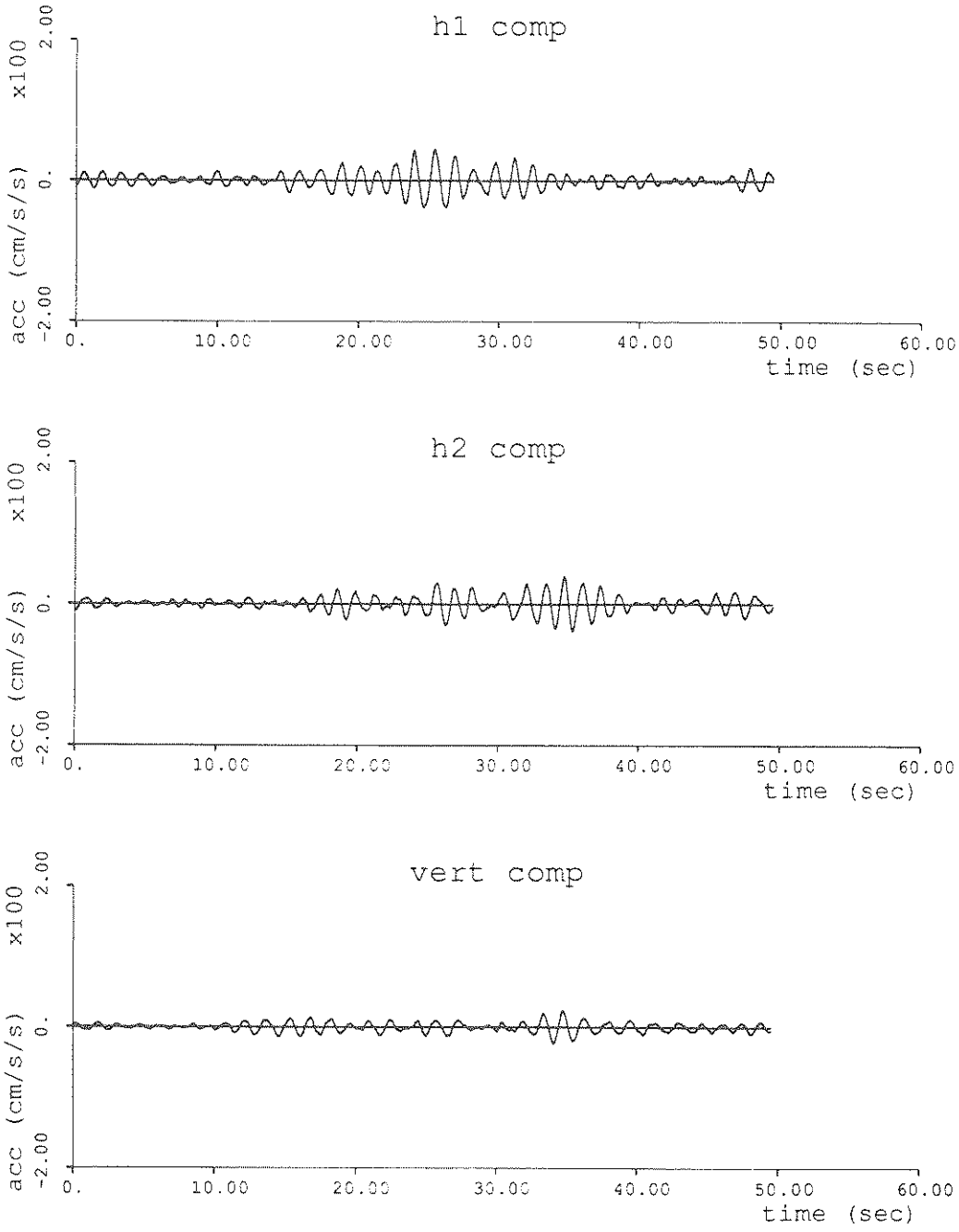


Fig. 5.27 Simulation of UNAM accelerogram using $F(f)$ equal to two standard deviations above the value calculated from Table 5.2.

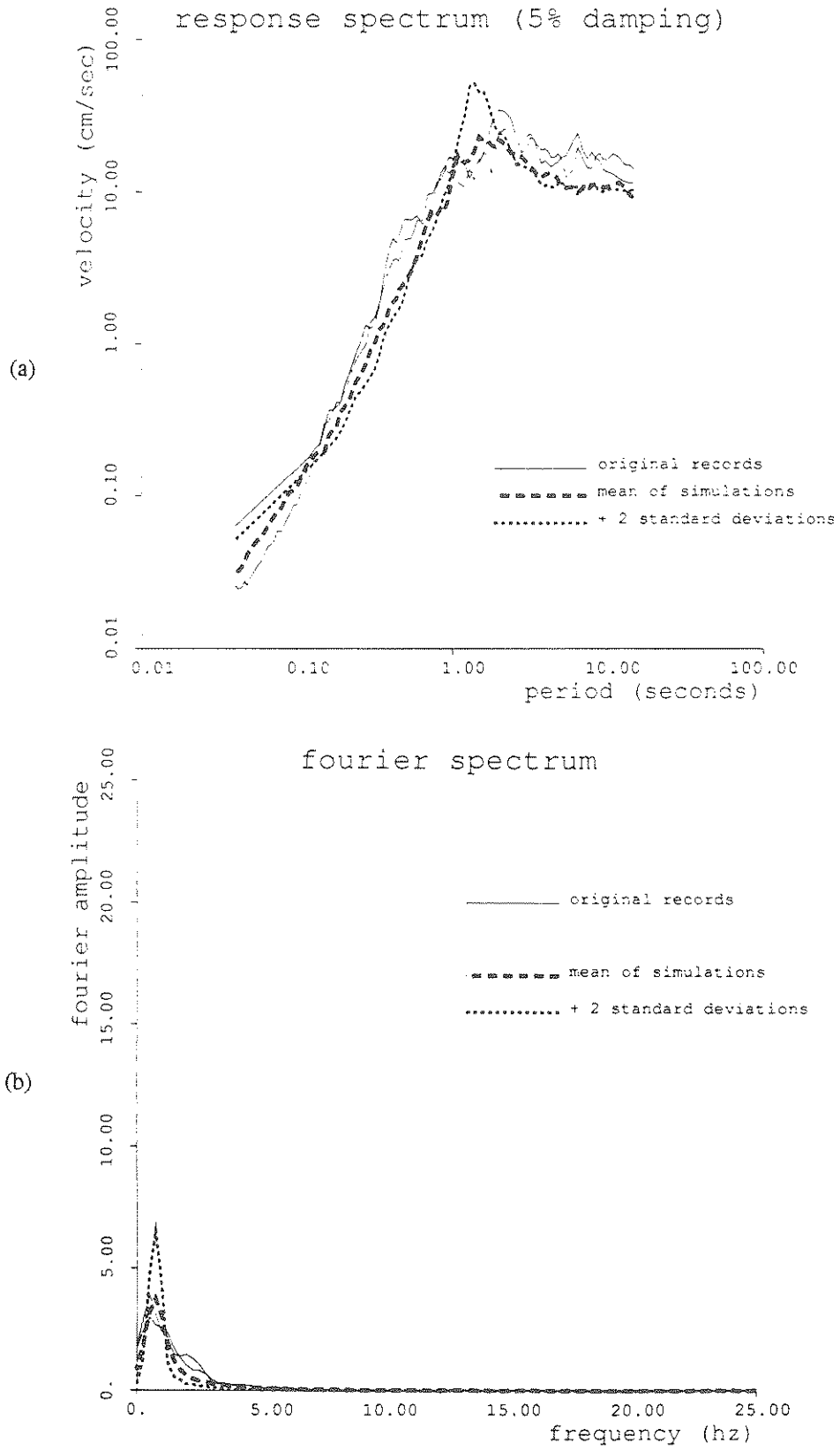


Fig. 5.28 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original UNAM accelerogram, simulations using parameters calculated from Table 5.2, and simulations using $F(f)$ equal to two standard deviations above the value calculated from Table 5.2.

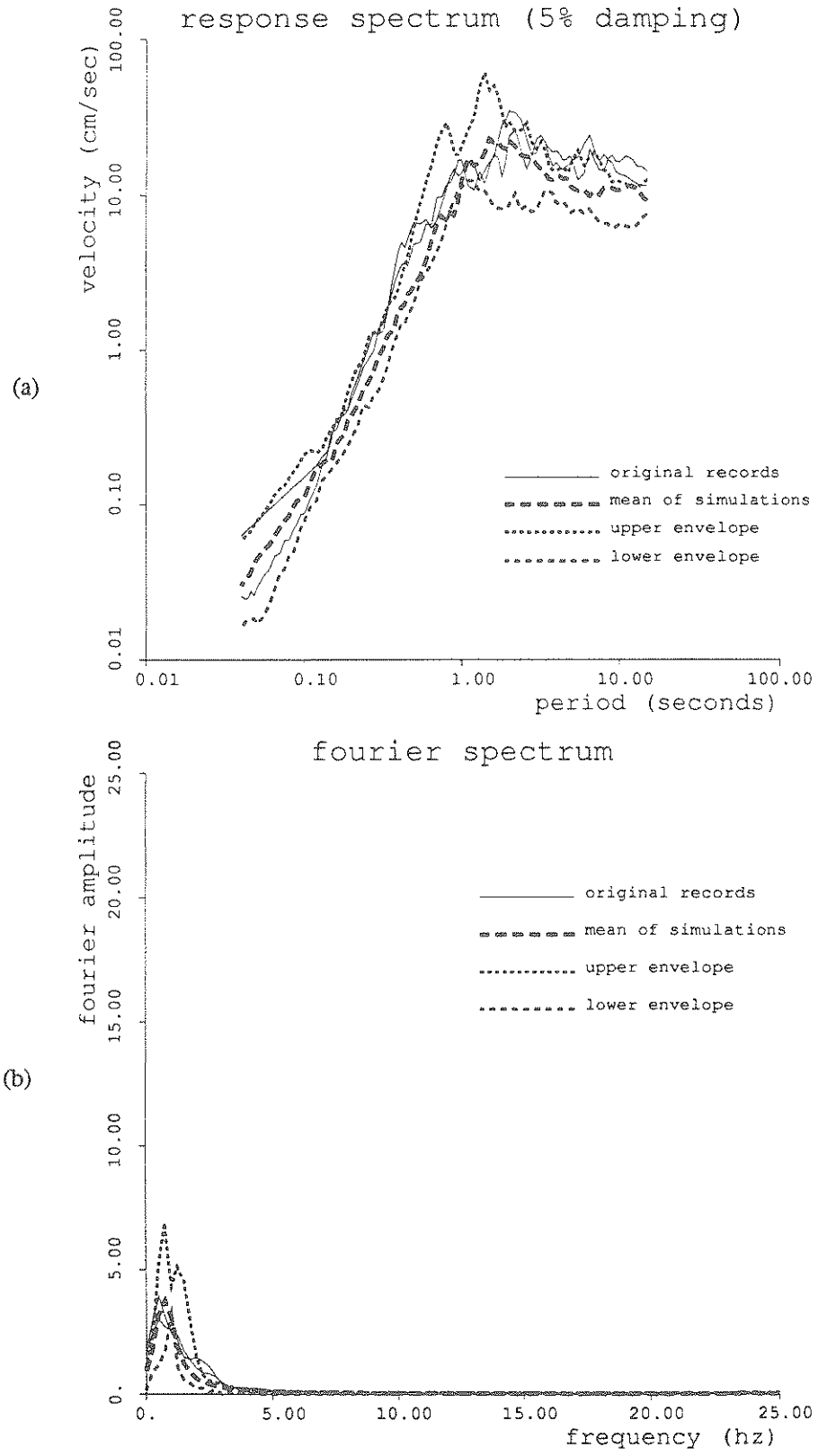


Fig. 5.29 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original UNAM accelerogram, simulations using parameters calculated from Table 5.2, and upper and lower envelopes for the simulations.

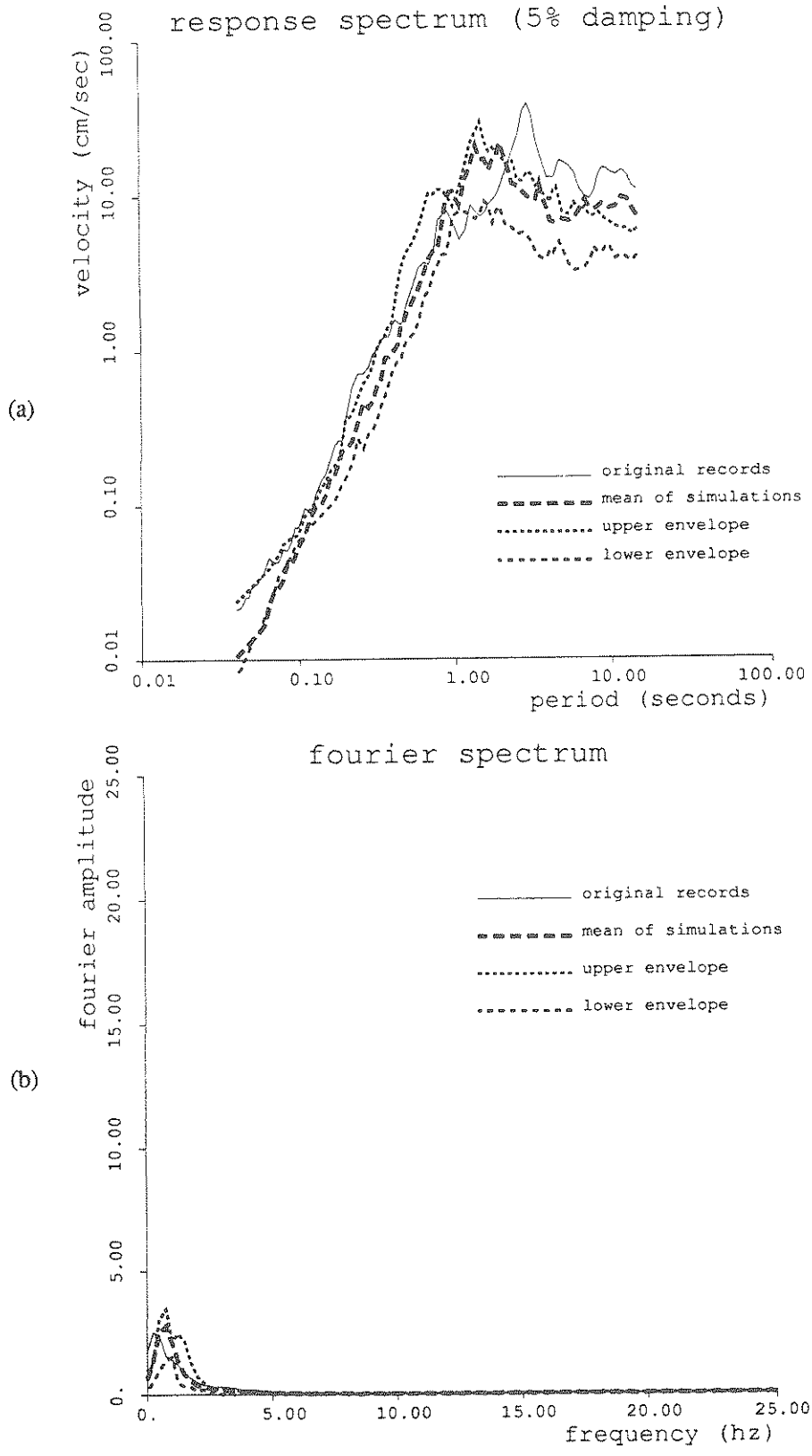


Fig. 5.30 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original UNAM accelerogram, simulations using parameters calculated from Table 5.2, and upper and lower envelopes for the simulations.

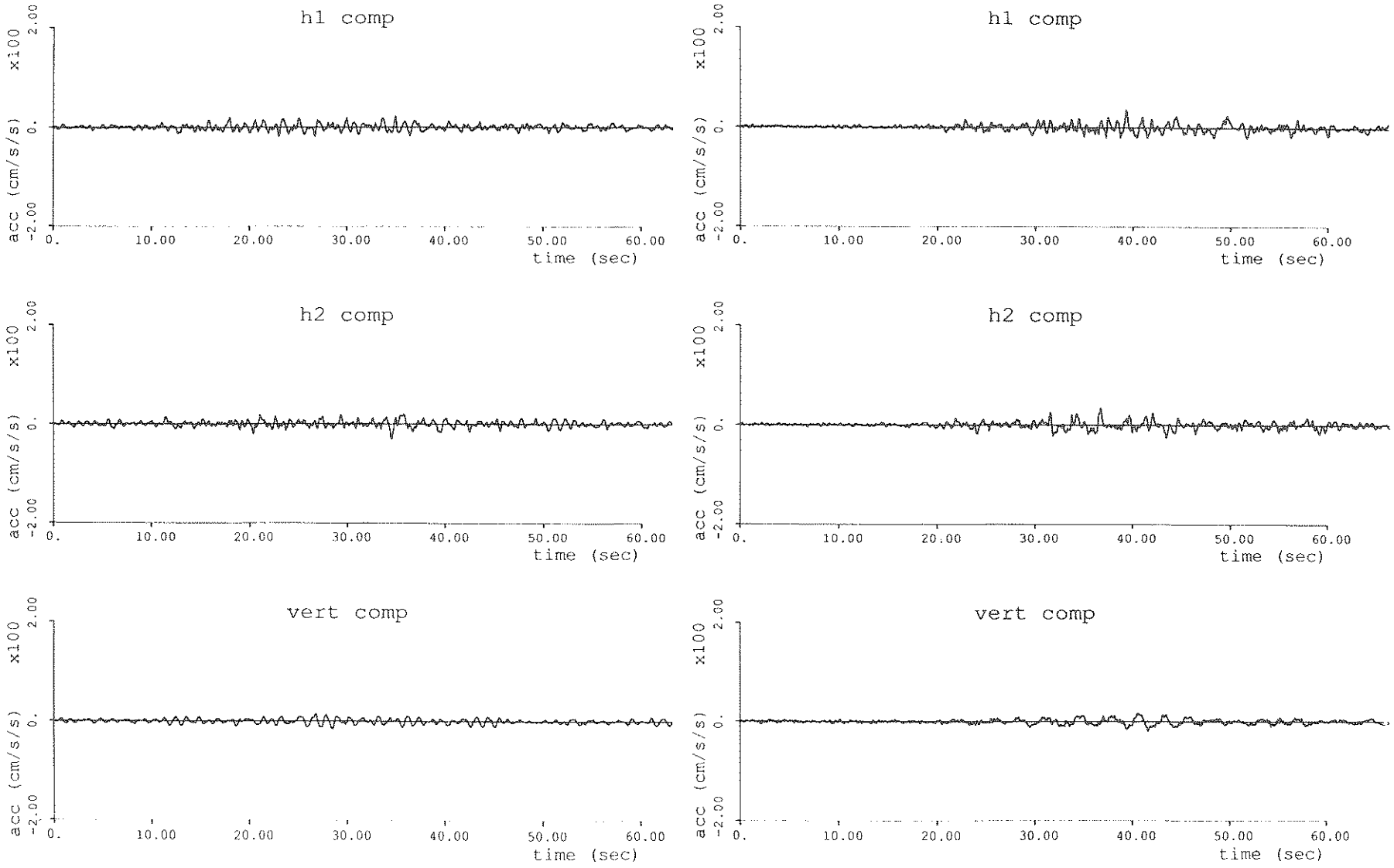
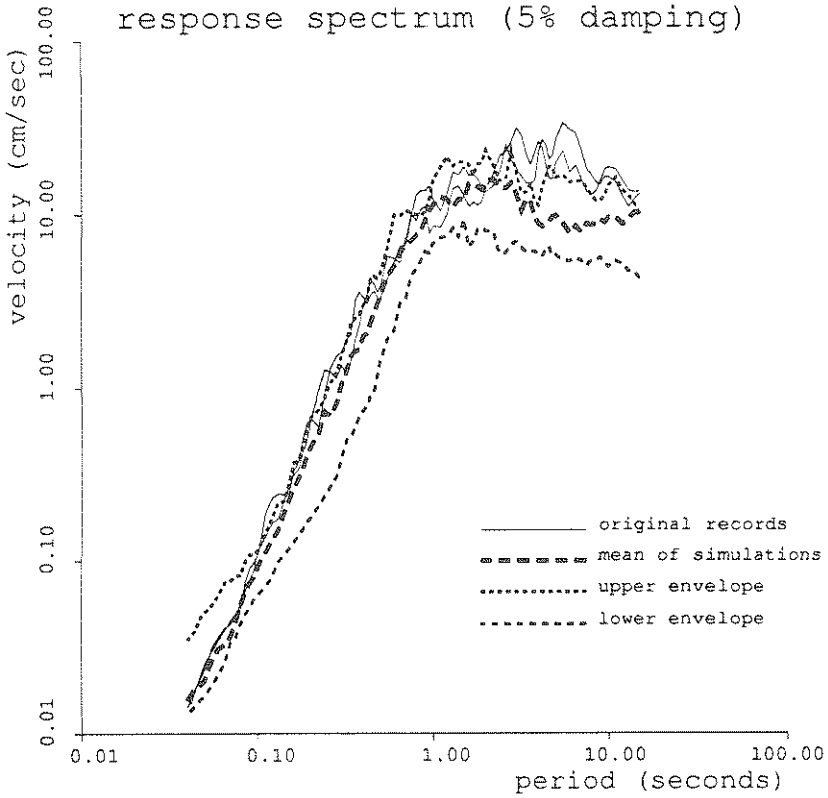


Fig. 5.31 Comparison of (a) simulation of TACY accelerogram using parameters calculated from Table 5.2 and (b) original TACY accelerogram.

(a)



(b)

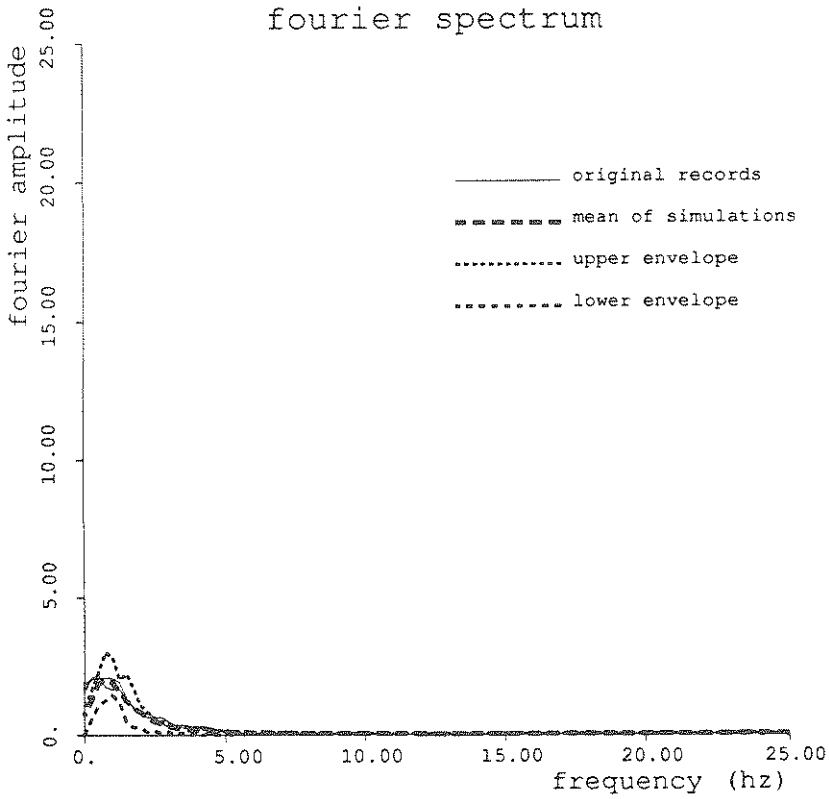
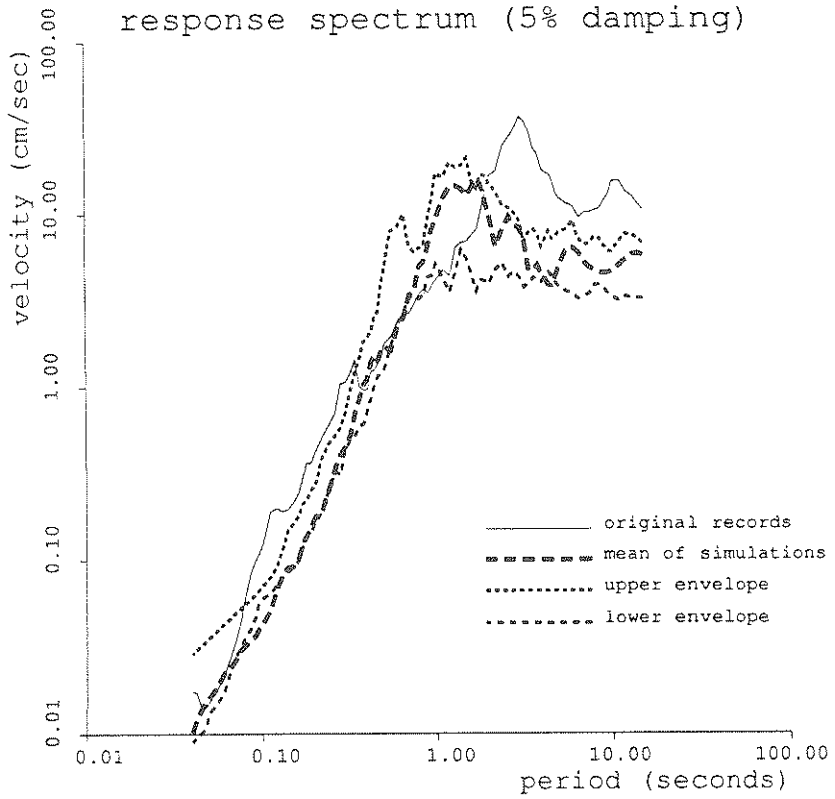


Fig. 5.32 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original TACY accelerogram and simulations using parameters calculated from Table 5.2.

(a)



(b)

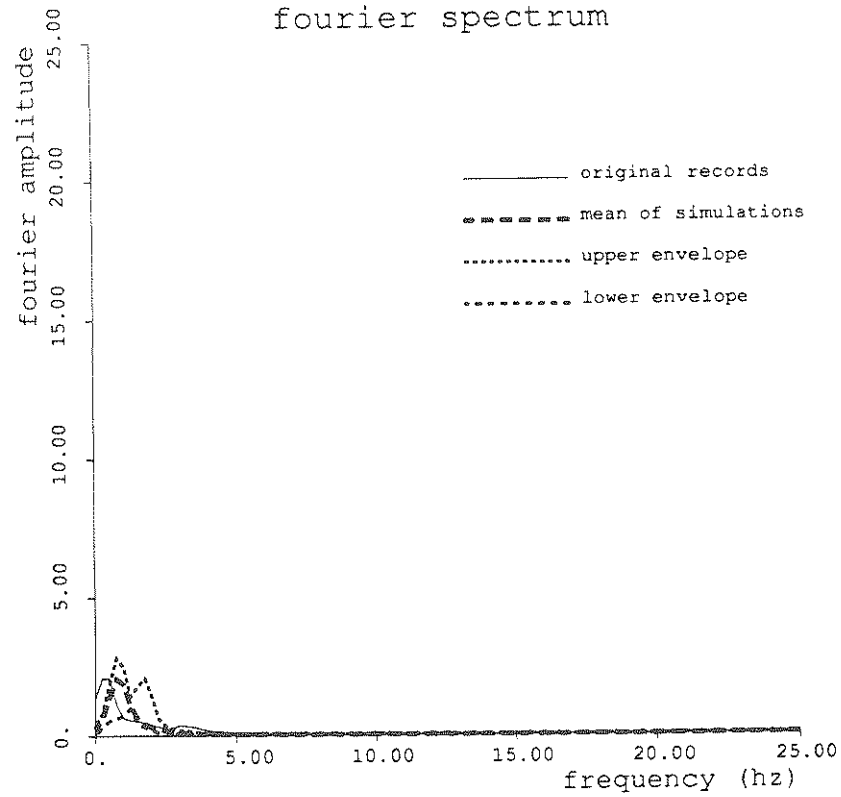


Fig. 5.33 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original TACY accelerogram and simulations using parameters calculated from Table 5.2.

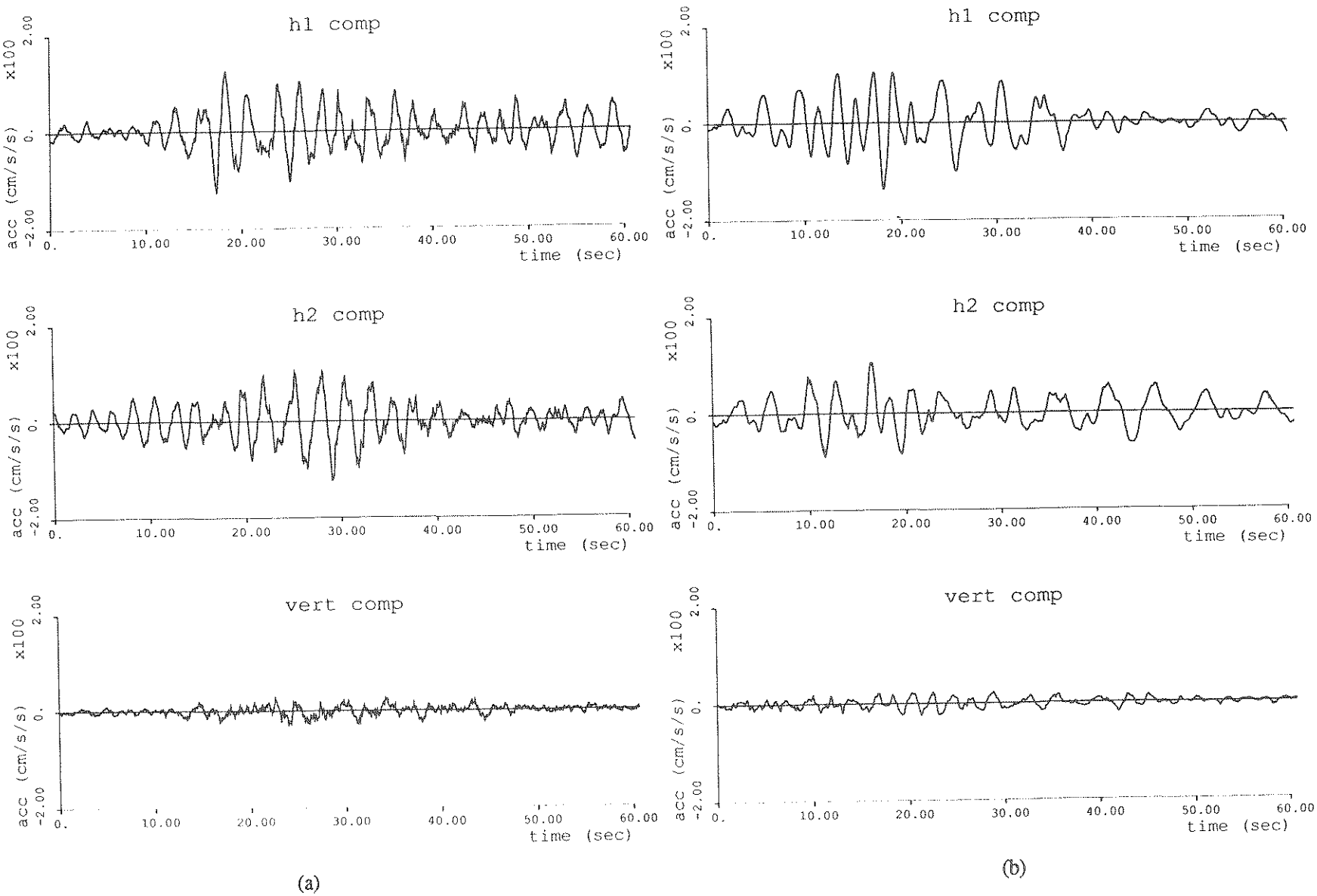


Fig. 5.34 Comparison of (a) simulation of TLHB accelerogram using parameters calculated from Table 5.2, and (b) original TLHB accelerogram.

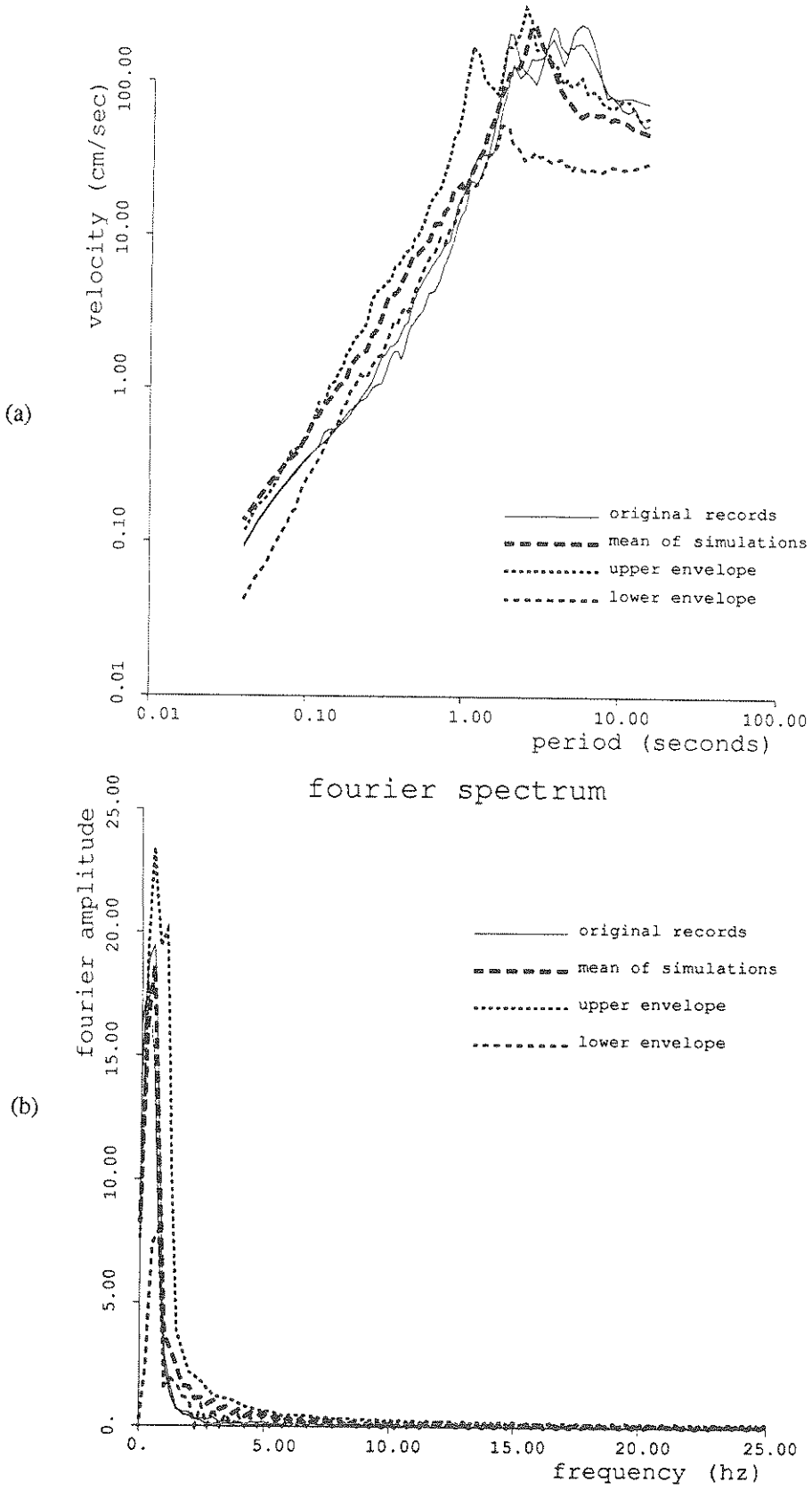


Fig. 5.35 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original TLHB accelerogram and simulations using parameters calculated from Table 5.2.

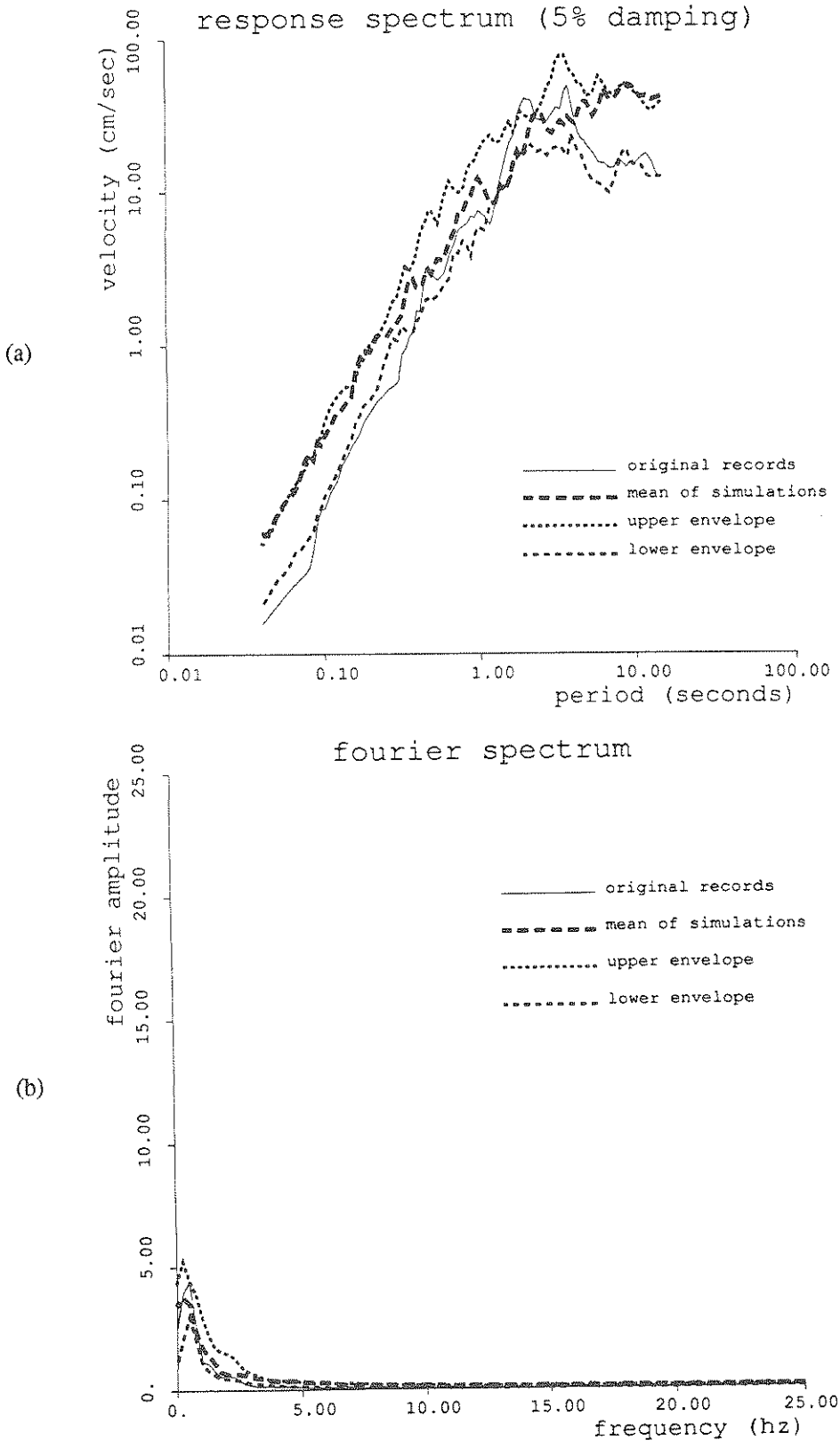


Fig. 5.36 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original TLHB accelerogram and simulations using parameters calculated from Table 5.2.

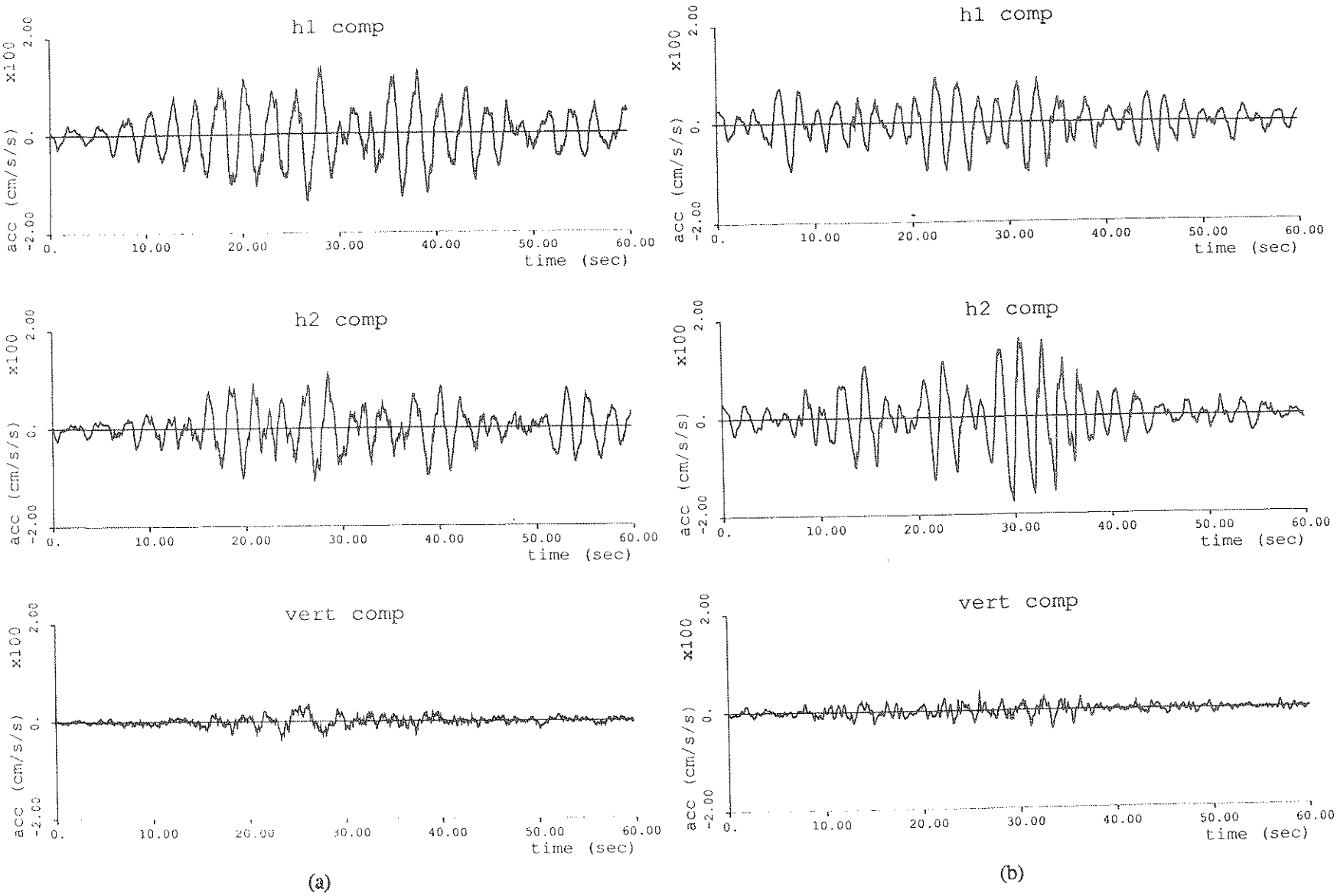
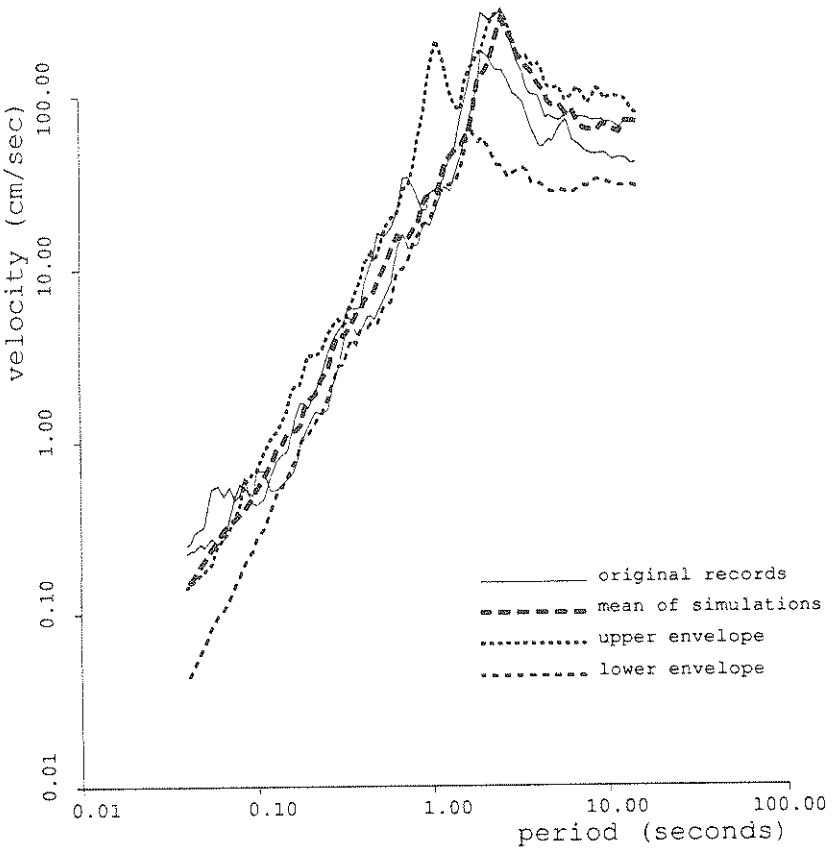


Fig. 5.37 Comparison of (a) simulation of SCTI accelerogram using parameters calculated from Table 5.2, and (b) original SCTI accelerogram.

(a)



(b)

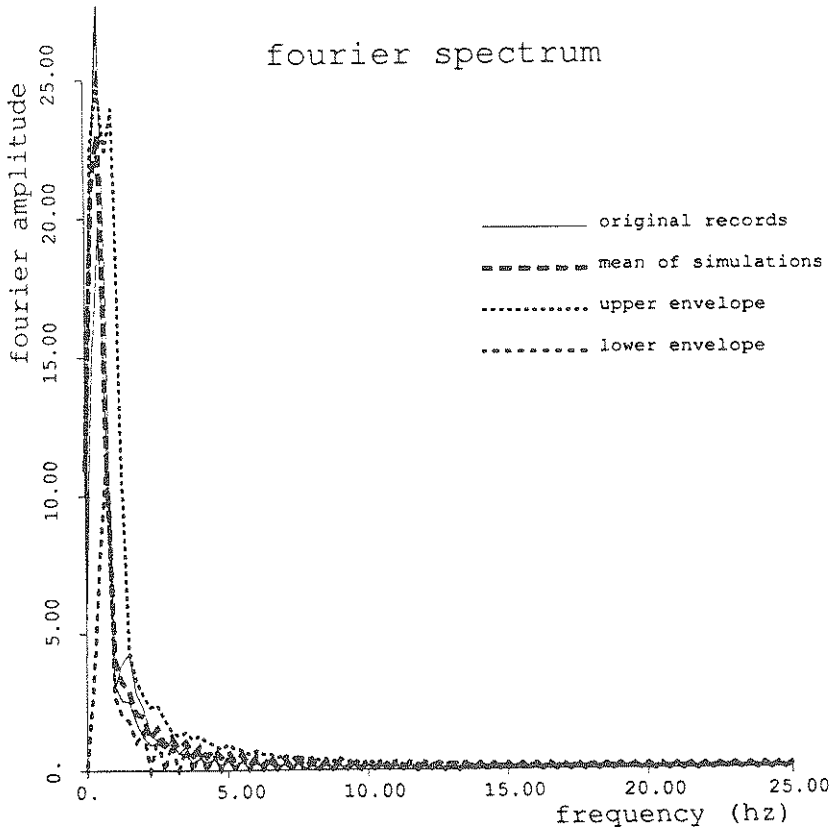
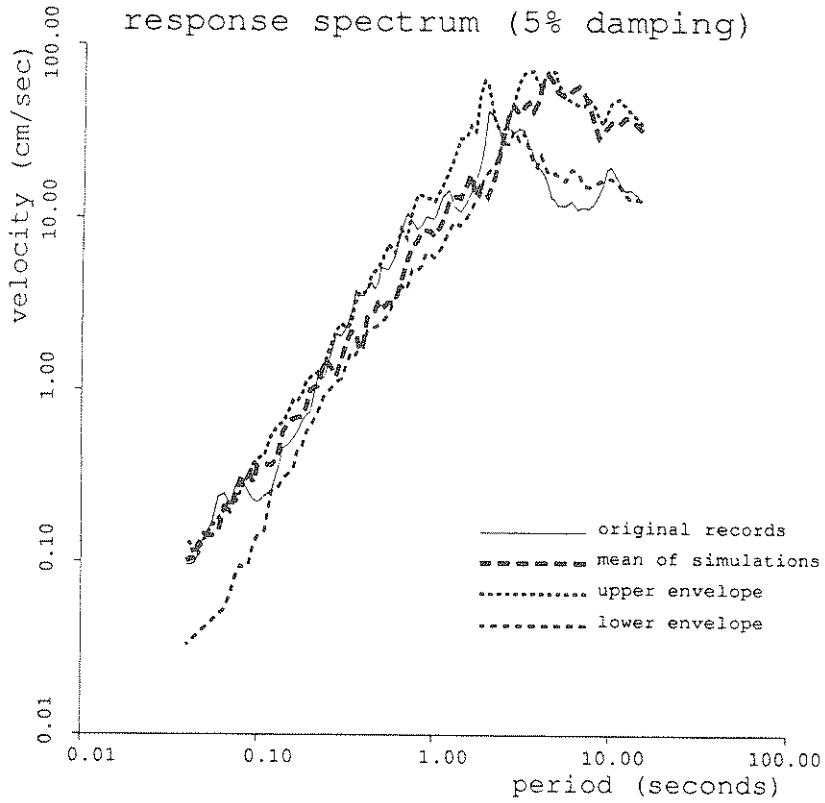


Fig. 5.38 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original SCTI accelerogram and simulations using parameters calculated from Table 5.2.

(a)



(b)

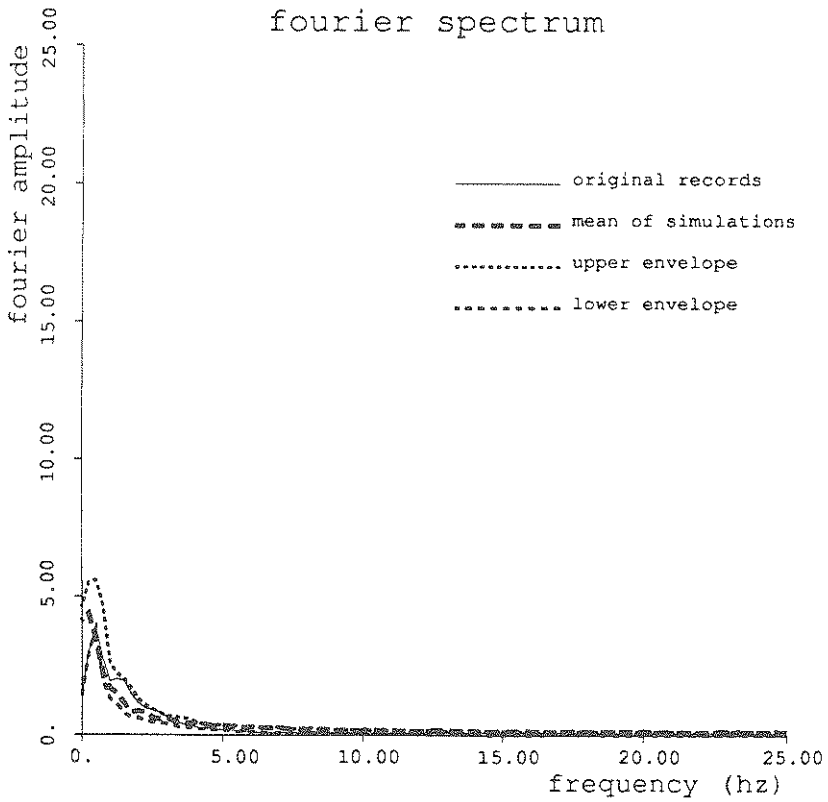


Fig. 5.39 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original SCTI accelerogram and simulations using parameters calculated from Table 5.2.

6. TAIWAN STUDY

6.1. Introduction

The SMART 1 array is a densely instrumented array located in Lotung, Taiwan. As shown in Fig. 6.1a, the array consists of a center instrument surrounded by three concentric circles of accelerographs. The radii of these circles are 200, 1000, and 2000 meters. The array is located on top of an alluvial soil layer about 500 meters deep with a P-wave velocity of 500 to 1000 meters per second. The P-wave velocity in the bedrock is approximately 3300 meters per second. Numerous earthquakes have been recorded by this array since its first implementation in 1980 (Bolt et al., 1982).

In this chapter the four events recorded in Table 6.1 will be modelled. A total of 92 accelerograms recorded the four events. To assess the validity of the model, the stations recording the four events were randomly divided into two groups. The accelerograms in the first group were modelled using the multivariate procedure developed in Chapter 4. Functional relations between the modelling parameters and physical variables were then developed. To assess the validity of the model, simulations generated from the functional relationships were used to predict the ground motion recorded by the accelerograms in the second group.

6.2. Comparison with Mexico Results

To compare the modelling parameters estimated in Taiwan with the parametric relations estimated for Mexico, the Mexican model shown in Table 5.4 was used to predict the Taiwan modelling parameters. In Fig. 6.2a the results are shown for α. The Mexican model applied to the SMART 1 data predicts a value of α which is too high. One possible explanation is that for some earthquakes in Taiwan the epicentral distance varies greatly from the hypocentral distance. For example, an earthquake may have an epicenter located very close to a recording station, but because the earthquake occurred deep below the ground surface it will have a large hypocentral distance. In Mexico the epicentral distance was usually very large and therefore did not vary much from the hypocentral distance. In Fig. 6.2b the hypocentral distance rather than the epicentral distance was used to predict the values of α. Although this

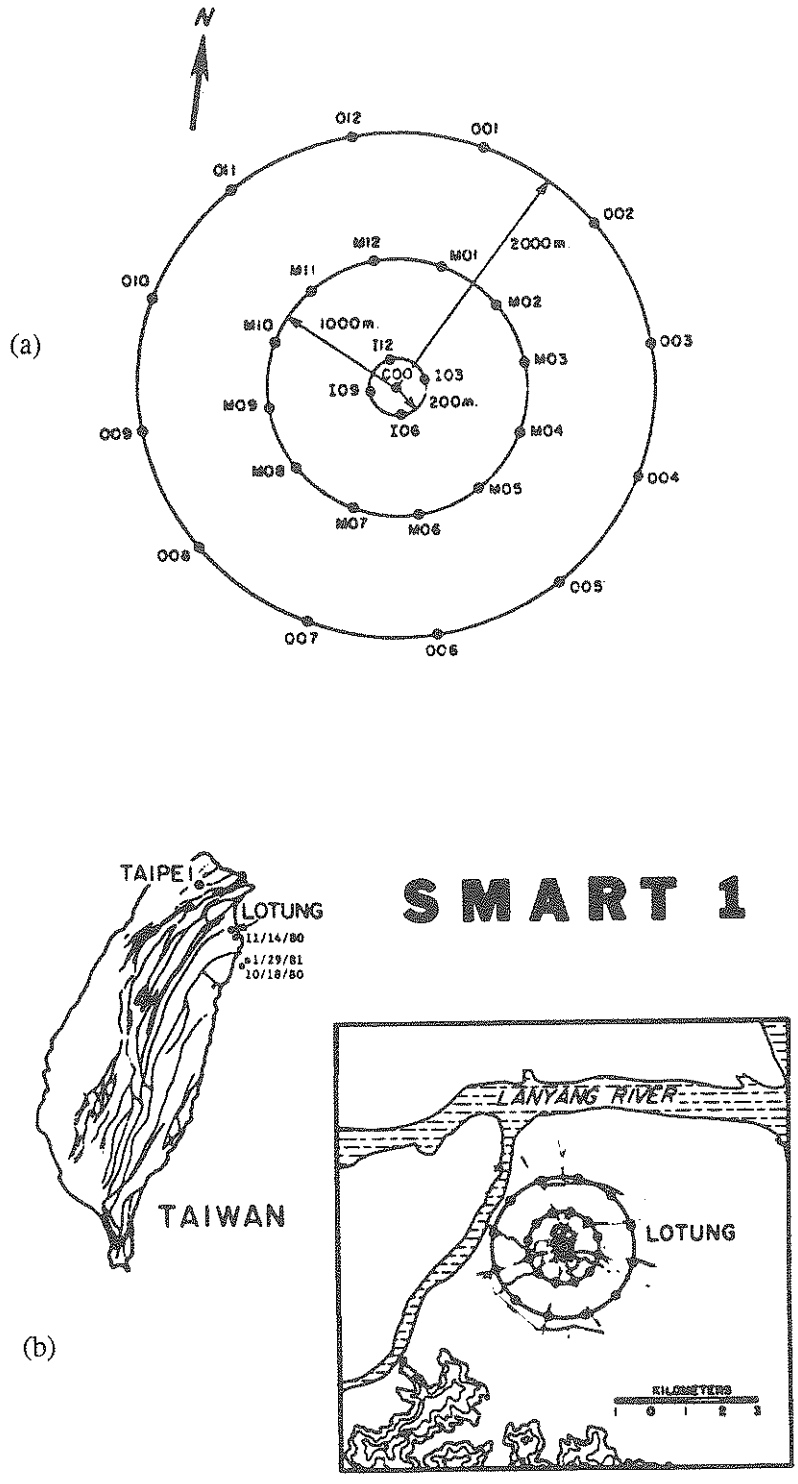
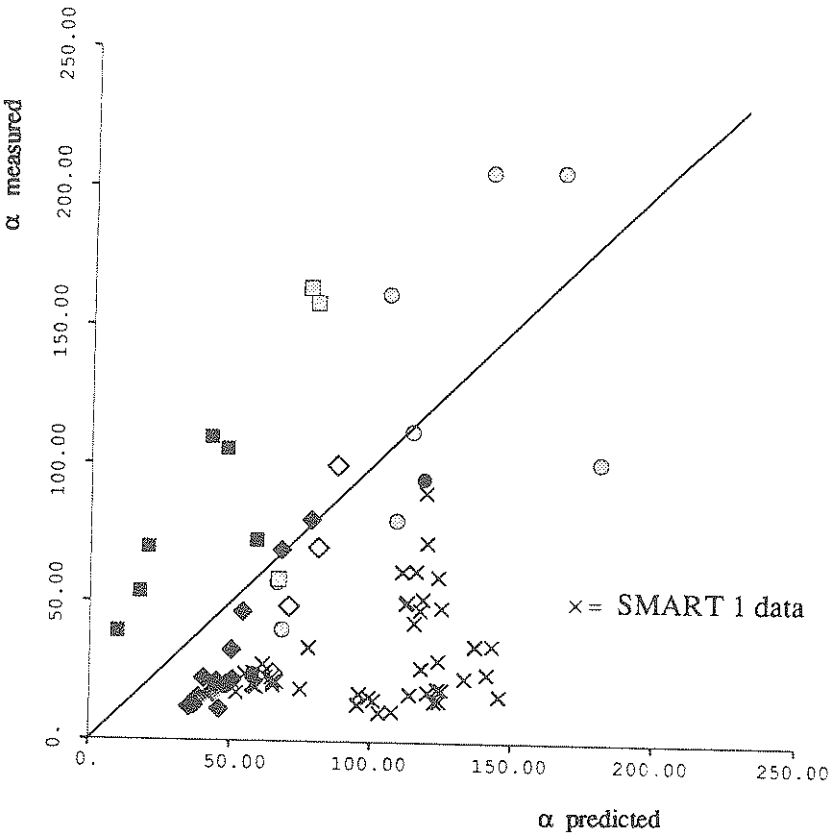


Fig. 6.1 Location of (a) stations in SMART 1 array and (b) SMART 1 array in Taiwan (after Harichandran and Vanmarcke (1984)).

Table 6.1 Taiwan Records analyzed

Event No.	Date	Depth (km)	M_L	dist (km)
2	11/14/80	62.1	5.9	10.0
5	1/29/81	11.1	6.9	30.0
18	2/28/82	10.0	4.9	8.0
20	12/17/82	30.6	5.9	116.6

(a)



(b)

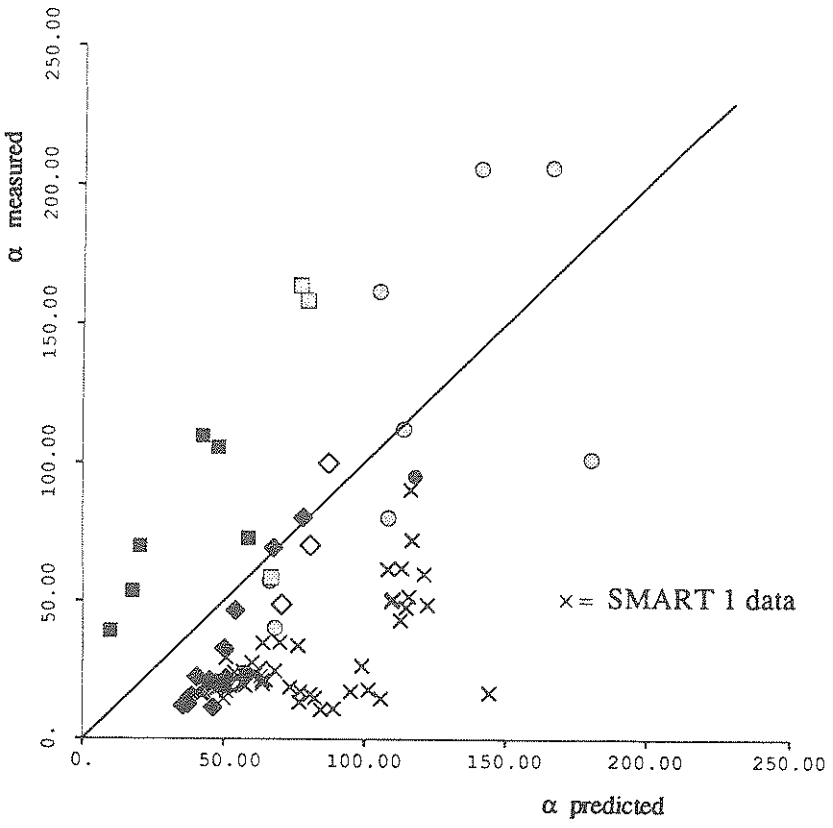


Fig. 6.2 Comparison of α measured and α predicted using the Mexican model in Table 5.4. Distance used to calculate α for the Taiwan data is (a) epicentral and (b) hypocentral distance.

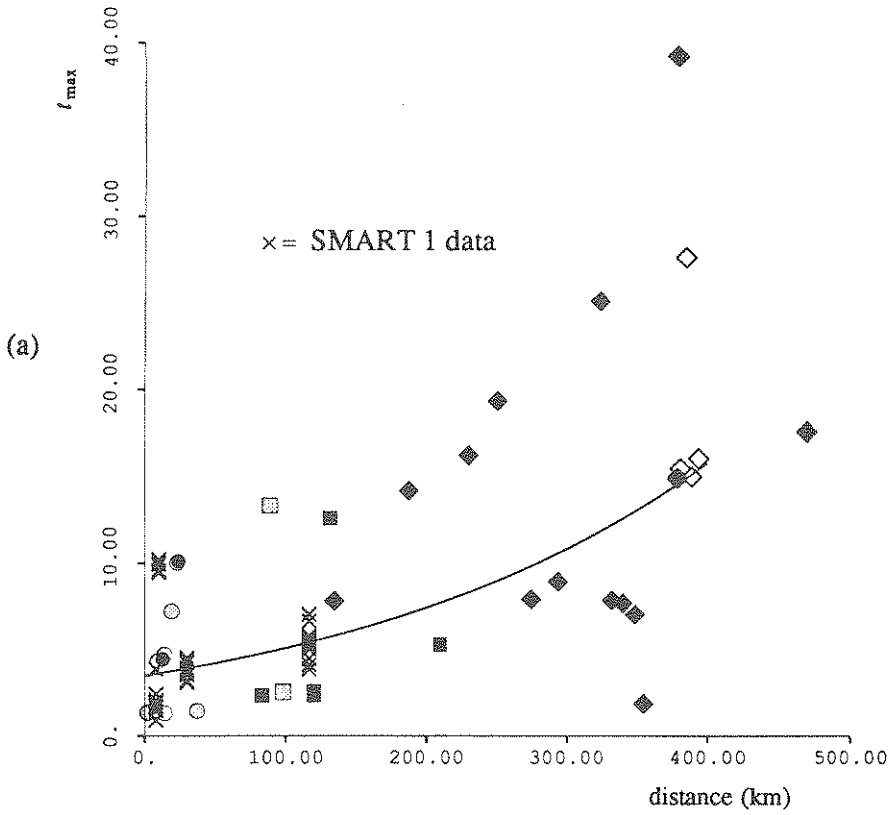
improves the accuracy of the predicted values of α, the Mexican model still predicts values of α which are higher than the measured values for the SMART 1 data. Thus a new relationship between α and the physical variables must be estimated for Taiwan.

In Fig. 6.3 the Mexican model is used to predict values of t_{max} and τ. The values for Taiwan are consistent with those predicted by the Mexican model. In Fig. 6.4 the values describing the frequency content predicted by the Mexican model are shown. Both f_{max} and $F(f_{max})$ are consistent with the Mexican results. Thus the main difference between the functional relations estimated for Mexico and the Taiwan results is that for similar physical conditions a lower value of α is estimated for Taiwan.

6.3. Parametric Relations

A multiple regression analysis was performed on the parameters estimated for the SMART 1 array and the physical variables listed in Table 6.1. The resulting parametric relations are shown in Table 6.2. As in the Mexican study the standard errors of the estimated regression coefficients and the coefficient of determination are also reported. Although parameters from 45 records were included in the regression analysis, the physical variables affecting the ground motion consisted of only four groups. Thus the functional relations presented in the Taiwan study can not be considered as reliable as the results of the Mexican study.

The form of the functional relation between α and the physical variables is similar to the one used in the Mexican study. Because of the varying depth of the earthquakes in Taiwan, the hypocentral distance was used in the regression analysis instead of the epicentral distance. The coefficient of determination, r^2, was highest when the hypocentral distance was used. The value of τ was found to vary with distance as shown in Fig. 6.5b. As the distance between the epicenter and the recording station increases, the difference in the arrival time of the seismic waves also increases resulting in a longer duration of strong shaking. In Fig. 6.6a the functional relation for t_{max} is shown. This parameter was found to increase linearly with depth. The difference in the arrival time of waves traveling at different velocities increases the time to the maximum shaking.



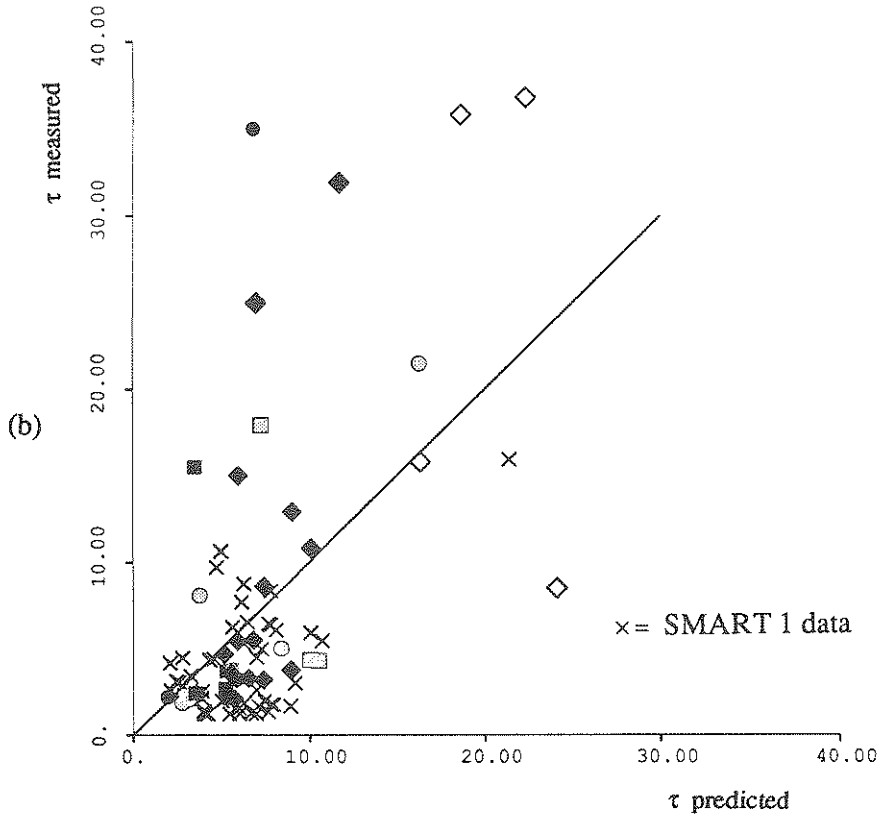


Fig. 6.3 (a) Relationship between t_{max} and epicentral distance and (b) comparison between τ measured and τ predicted using the Mexican model in Table 5.4.

Table 6.2 Parametric Relations for Taiwan

H = horizontal components

V = vertical components

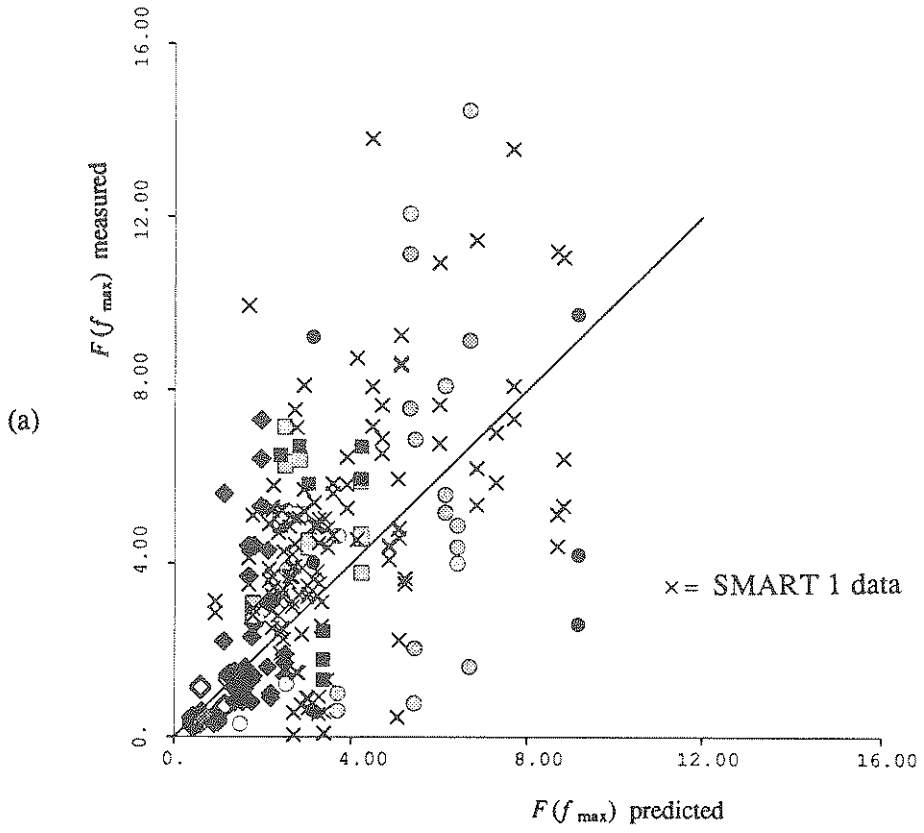
d = epicentral distance (km)

d_h = hypocentral distance (km)

M = magnitude

h = depth of epicenter (km)

$\alpha = (-42.0\pm1.26) + (6.94\pm1.26)\ln\left[\frac{10^M}{d_h^{1.9} k_3^3}\right]$		$r^2 = 0.69$
$\ln(\tau) = (0.574\pm0.095) + (0.011\pm0.0013)d$		$r^2 = 0.63$
$k_1 = (-10.6\pm1.19) + (5.39\pm0.36)\alpha$		$r^2 = 0.74$
$\ln(t_{max}) = (1.527\pm0.24) + (0.130\pm0.007)d$		$r^2 = 0.88$
$r_2 = 1.42\pm0.45$		
$b_2 = 1.67\pm0.82$		
$r_3 = (2.45\pm0.34) - (0.92\pm1.10)\tan^{-1}\frac{d}{h}$		$r^2 = 0.42$
$b_3 = 1.44\pm0.6$		
$k_3 = 0.338\pm0.078$		
$f_{max} = (17.2\pm1.21) - (2.08\pm0.19)M$	H	$r^2 = 0.54$
$f_{max} = (17.2\pm1.21) - (2.08\pm0.19)M$	V	$r^2 = 0.54$
$F(f_{max}) = (0.029\pm0.038) - (0.0260\pm0.0059)M$		$r^2 = 0.26$
$\phi_1 + \phi_2 + \phi_3 = 0.977\pm0.160$		



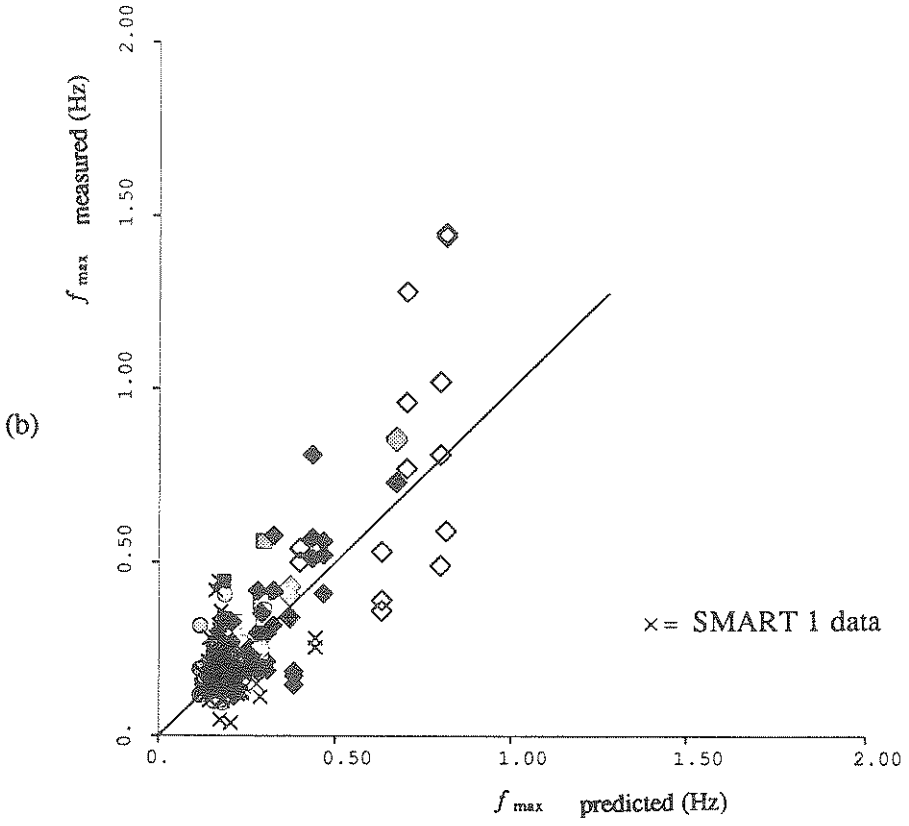
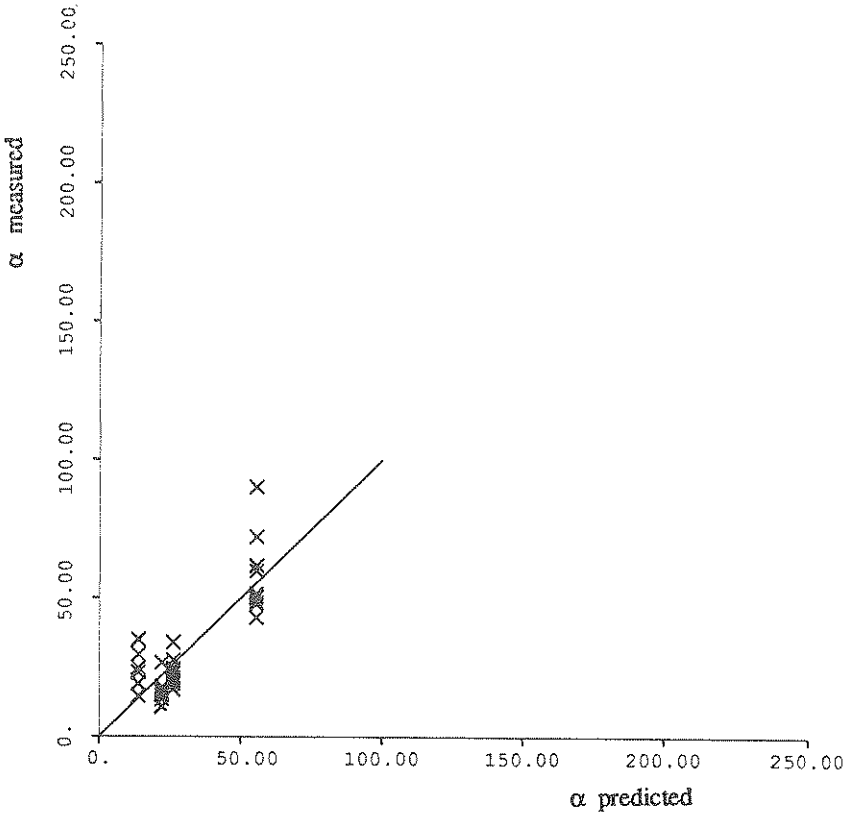


Fig. 6.4 Comparison of (a) $F(f_{max})$ measured and $F(f_{max})$ predicted and (b) f_{max} measured and f_{max} predicted using the Mexican model in Table 5.4.

(a)



(b)

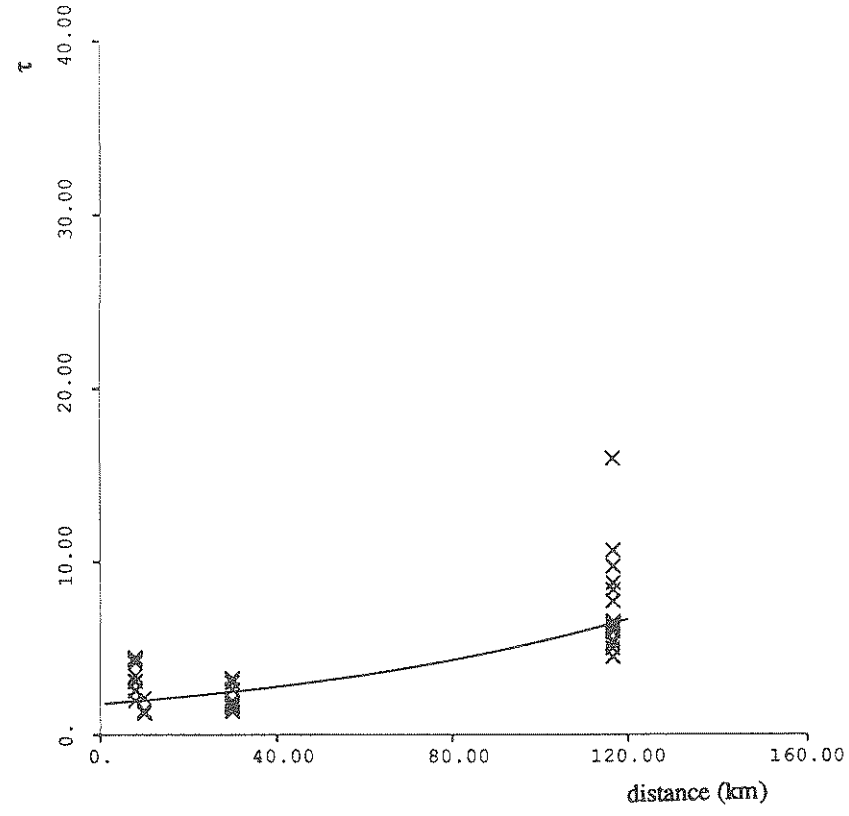
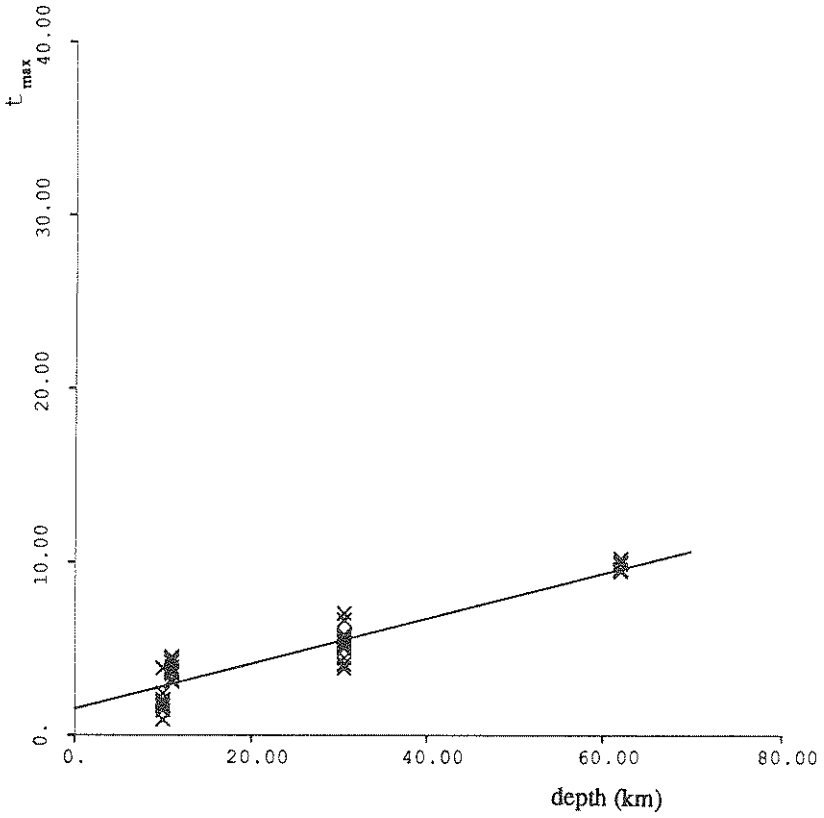


Fig. 6.5 (a) Comparison of α measured and α predicted and (b) relationship between τ measured and epicentral distance.

(a)



(b)

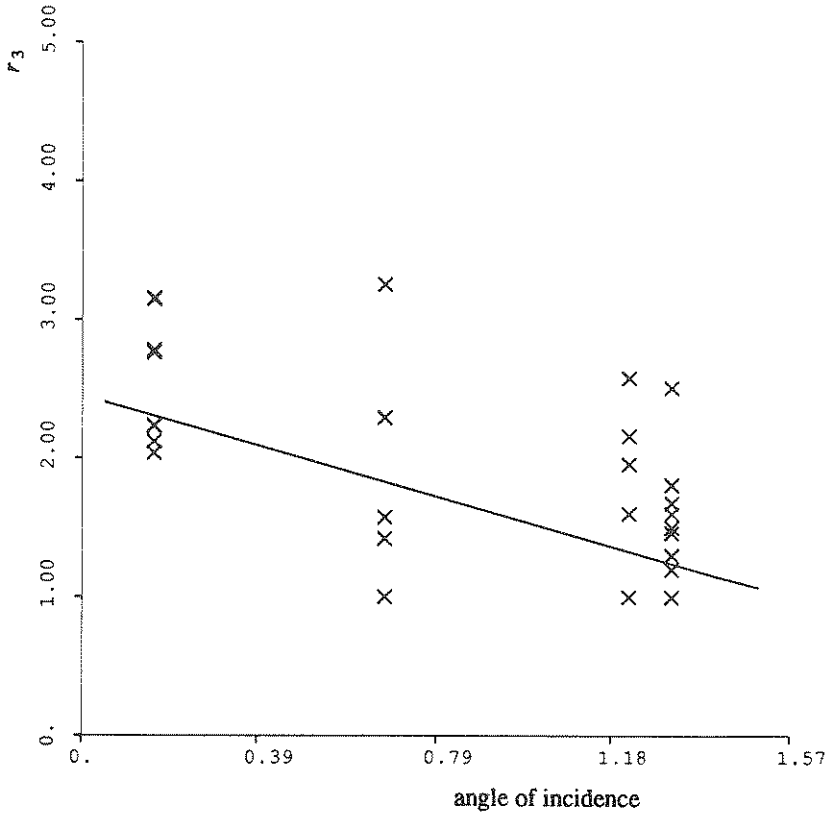


Fig. 6.6 Relationship between (a) t_{max} and depth and (b) r_3 and the angle of incidence between the source and the recording station.

The only parameter describing the vertical angle function that was found to vary with the physical variables was the ratio of the initial angle of shaking to the final angle of shaking, r_3. In Fig. 6.6b, r_3 is shown to decrease with the angle of incidence of the arriving waves. The angle of incidence is defined as the angle between the source and recording station and the vertical direction. Figure 6.7 presents an explanation of this behavior. In Fig 6.7a the expected behavior of the vertical angle envelope predicted for a low angle of incidence is shown. In the beginning of the record the ground motion is dominated by P-waves. The motion of the P-waves is longitudinal to the direction of propagation resulting in a high value of $\bar{\gamma}$ early in the record. Later in the record the SV-waves have a greater effect on the ground motion. Since the motion of the SV-waves is transverse to the propagation, the value of $\bar{\gamma}$ is low. For a high angle of incidence shown in Fig. 6.7b, the initial value of $\bar{\gamma}$ is lower and the final value is higher than in Fig. 6.7a. Thus the value of r_3 is lower for higher angles of incidence.

Due to the compactness of the SMART 1 array, the soil conditions were similar for each station. The value of k_3 could not be related to any of the physical variables and thus was estimated to be a constant with a mean of 0.338. This is consistent with the theory that k_3 varies with site conditions and not source parameters. Using an ANOVA analysis, no significant difference was found among the mean values of k_3 estimated for each event.

To estimate a relationship between k_3 and the soil properties, the values of k_3 estimated for Taiwan were combined with values previously estimated for Mexico and new values estimated for California for sites where shear wave velocity profiles were reported. The shear wave velocity profiles for Mexico City are from Herrera, et al (1965). The profiles for the California sites are from the U. S. Nuclear Regulatory Commission (1978). The shear wave velocity was calculated as the total thickness of the soil layer divided by the total time for a wave to propagate through the layer as calculated from the soil profile. The relationship between k_3 and the shear wave velocity is shown in Fig. 6.8. The values from Mexico consist of data recorded on the Mexico City lakebed and data recorded on rock from the September 19, 1985 Michoacan earthquake. The values from California were estimated from accelerograms recorded at Ferndale, Hollister, Taft, El Centro, Caltech (Millikan Library), Santa Barbara, and Cholame. The value of k_3 is lowest for sites where low values of

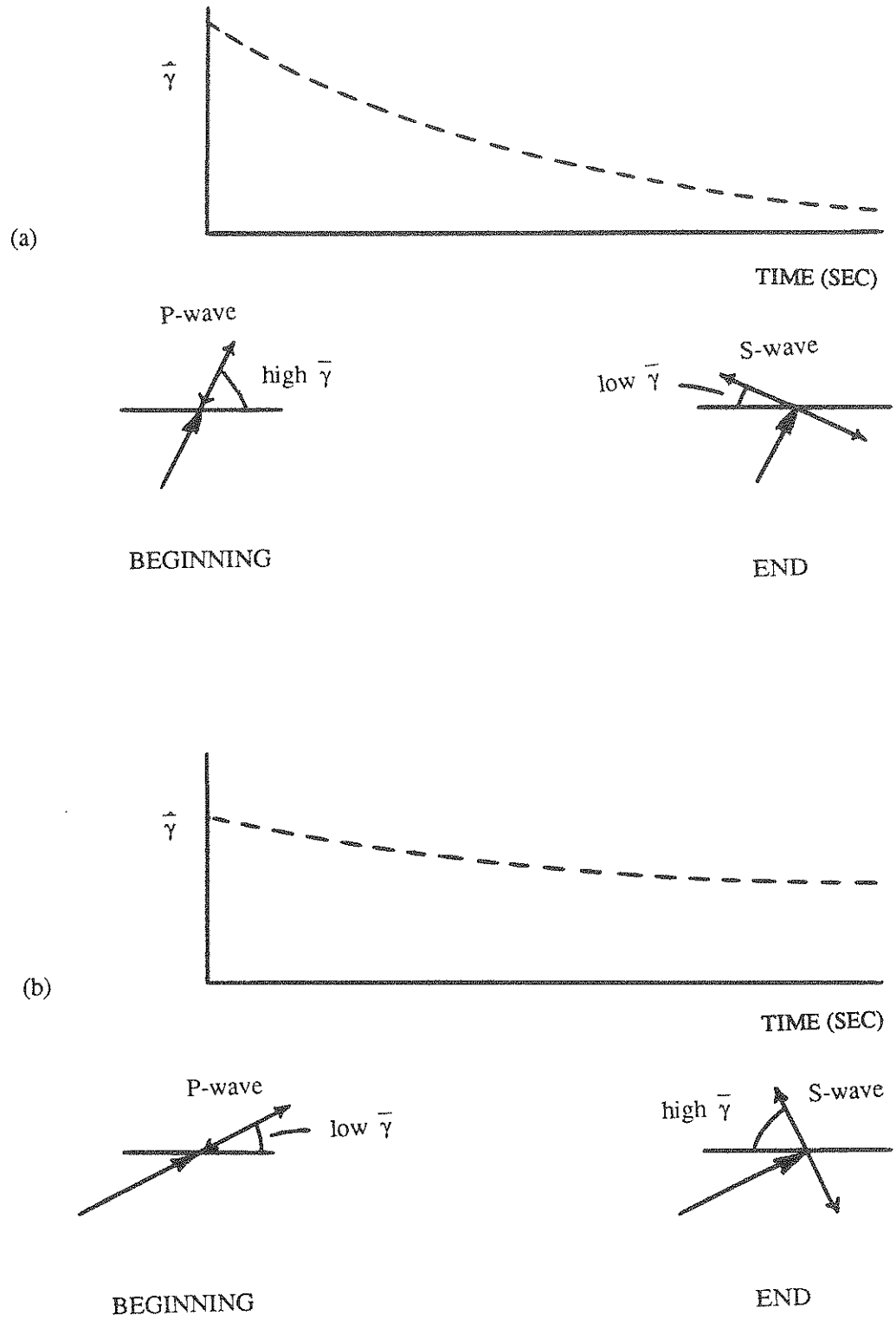


Fig. 6.7 Effect of the angle of incidence between the source and the recording station on the value of r_3.

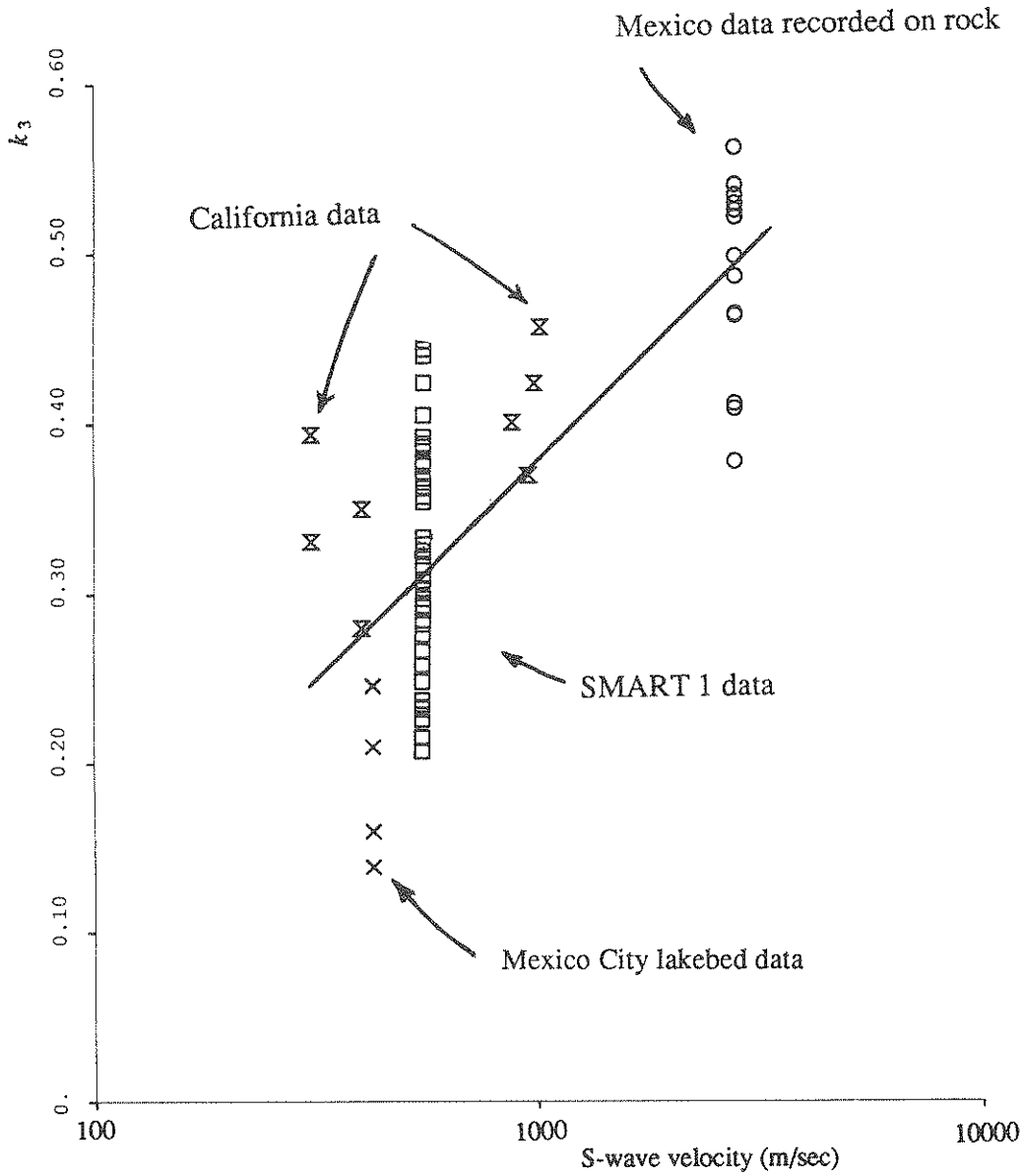


Fig. 6.8 Relationship between k_3 and S-wave velocity.

the shear wave velocity are reported. This is due to the increased refraction of the seismic waves at the bedrock-soil interface below the recording stations. When the recording station is located directly on rock, no refraction occurs and the value of k_3 is higher.

A comparison between the measured and predicted values of the parameters describing the frequency spectra is shown in Fig. 6.9. The value of the maximum of the ARMA Fourier spectrum was found to be weakly correlated to the magnitude of the earthquake. A much stronger correlation between the dominant frequency and the earthquake magnitude was found. As the magnitude of the earthquake increases, the dominant frequency of the ARMA Fourier spectrum decreases and the peak becomes more narrow. Also, the dominant frequency of the vertical components was found to be significantly higher at the 95% confidence level than the horizontal components by about 1.5 Hz.

6.4. Spatial Correlation of Modelling Parameters

One of the advantages of the Taiwan array is that correlations between modelling parameters at sites located close together can be measured. To measure the correlation of the modelling parameters between recording sites, each set of modelling parameters recorded at a site may be considered to be a series. Thus in this study the modelling parameters are a series of length four; one parameter calculated per each of the four events. Stations which recorded all four events and were modelled were stations C00, I06, M04, M05, M06, O03, and O09. The correlation of the modelling parameters between two stations is

$$\rho_\eta = \frac{\sum_i (\eta_{1i} - \overline{\eta_1})(\eta_{2i} - \overline{\eta_2})}{\sigma_1 \sigma_2} \tag{6.4.1}$$

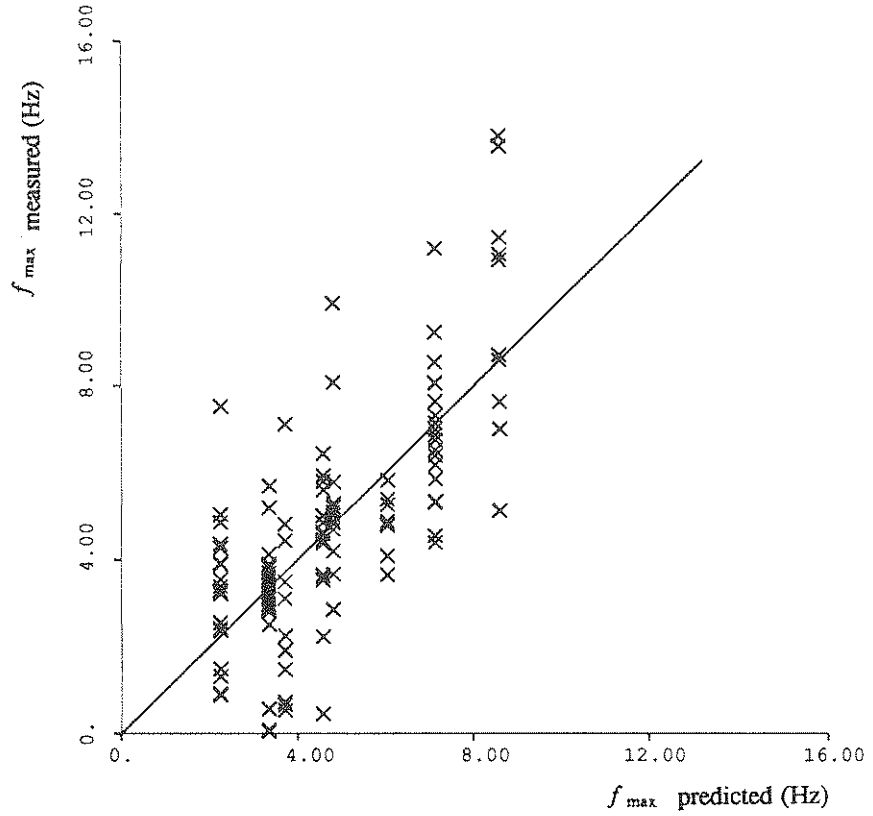
where

ρ_η = the correlation between the two series η_1 and η_2

η_{1i} = the parameter series estimated for the first station, i=1 to 4

σ_1 = the standard deviation of the first series

(a)



(b)

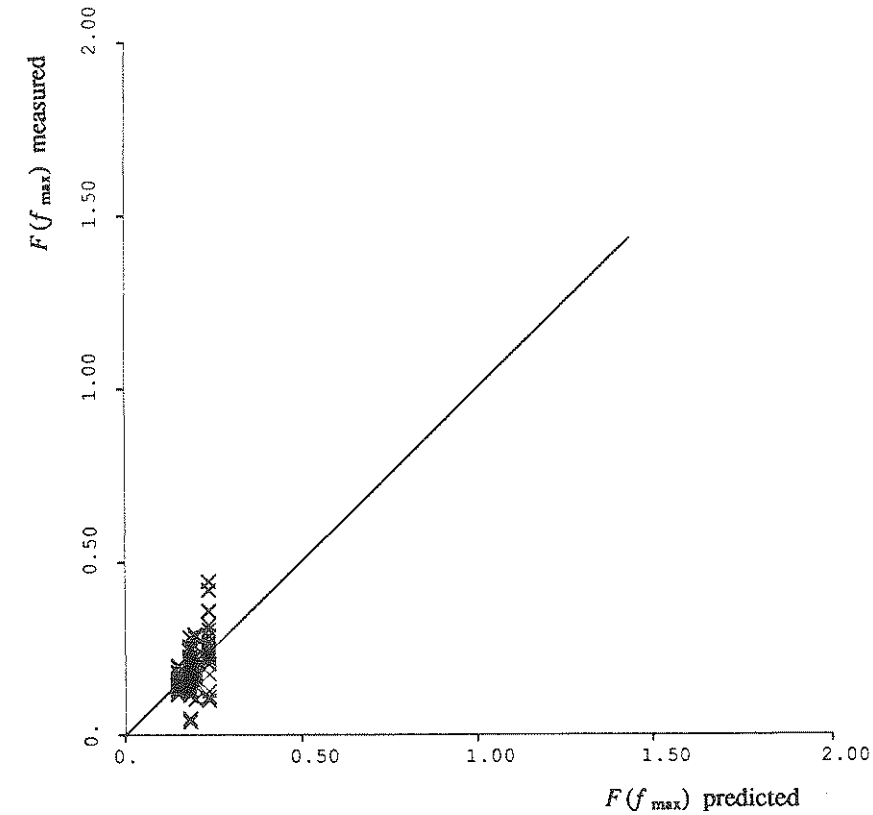


Fig. 6.9 Comparison of (a) f_{max} measured and f_{max} predicted and (b) $F(f_{max})$ measured and $F(f_{max})$ predicted.

$\overline{\eta_1}$ = the mean of the first series

η_{2i} = the parameter series estimated for the second station, i=1 to 4

σ_2 = the standard deviation of the second series

$\overline{\eta_2}$ = the mean of the second series .

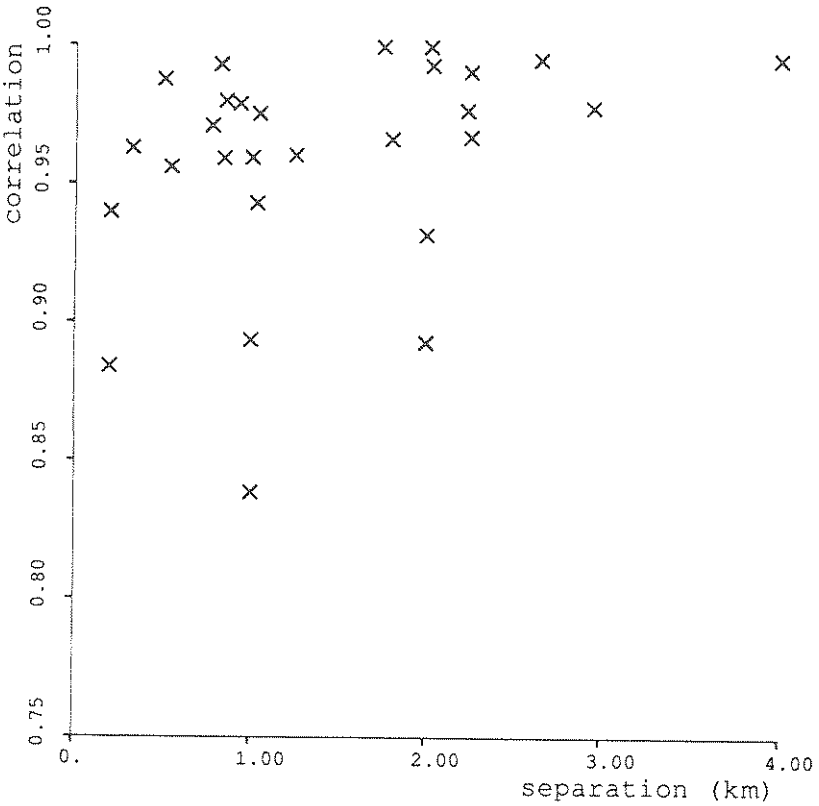
The correlations calculated are shown in Fig. 6.10a for α and 6.10b for k_3 versus separation distance. The values calculated for α are higher because of the correlation between α and the physical variables affecting its value. A regression analysis between the correlation of each parameter and the separation distance was performed. It was found that the slope of the regression line was not significantly different from zero and thus the correlation between the parameters was found to be unaffected by the separation distance. This result means that although the time series recorded by the SMART 1 array are correlated (Harichandran and Vanmarcke, 1984), the modelling parameters estimated for each station are independent of each other. Thus simulations for the entire array for a given set of physical variables may be generated from one set of modelling parameters. These simulations will possess realistic spatial distribution of statistical quantities, but will not be realistically correlated with each other. To model the cross-correlation among the accelerograms using the ARMA process, a multivariate ARMA model must be used to estimate the off-diagonal ARMA parameters. These parameters relate the correlation structure of the accelerograms recorded at different sites. This was not attempted in this study.

6.5. Simulation of the Records

To examine the quality of the model shown in Table 6.2, simulations based upon the physical variables listed in Table 6.1 were generated. Since the parametric relations shown in Table 6.2 were estimated using data generated from the first half of the study, the simulations were compared to records included in the second half of the study to measure the predictive value of the model.

Before any simulations were generated, a comparison of the modelling parameters calculated from Table 6.2 and the values estimated from the second half of the study was done. The results are presented in Figs. 6.11 to 6.14. The histogram shows the

(a)



(b)

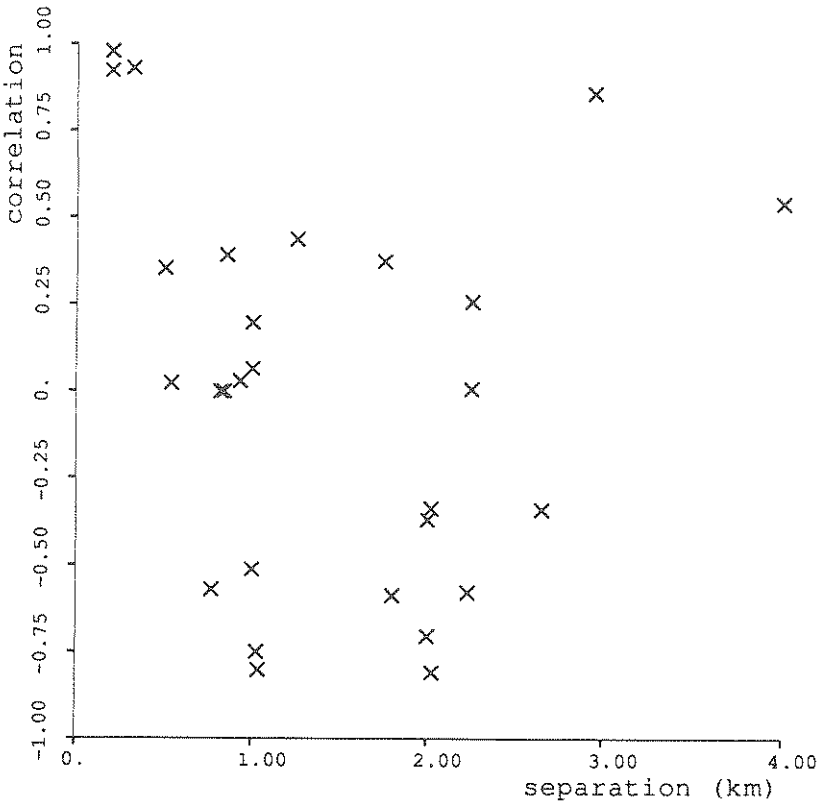


Fig. 6.10 Correlation between values of (a) α and (b) k_3 as a function of separation distance between stations.

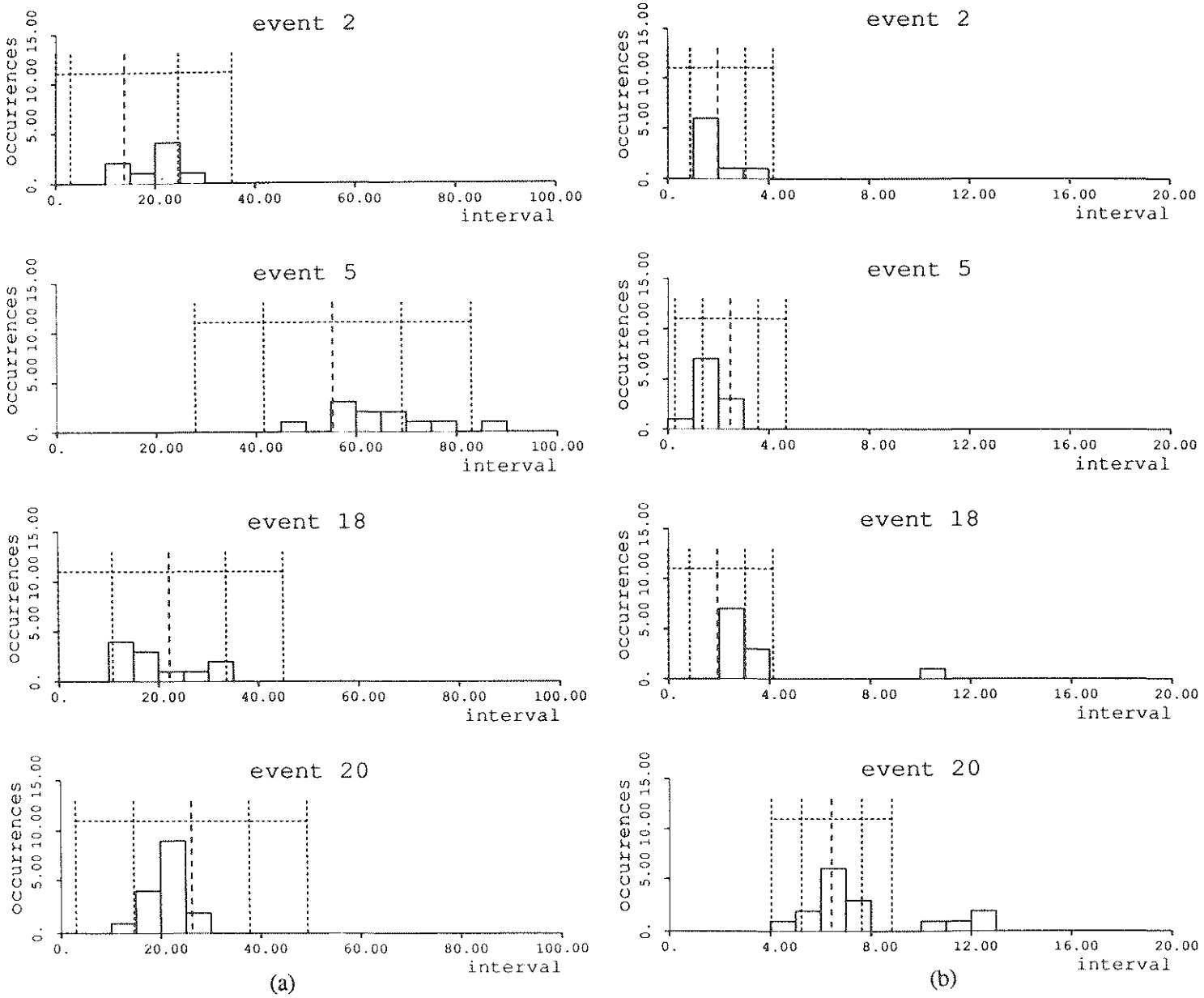


Fig. 6.11 Comparison of the value of (a) α predicted and α measured from the second half of the study and (b) τ predicted and τ measured from the second half of the study.

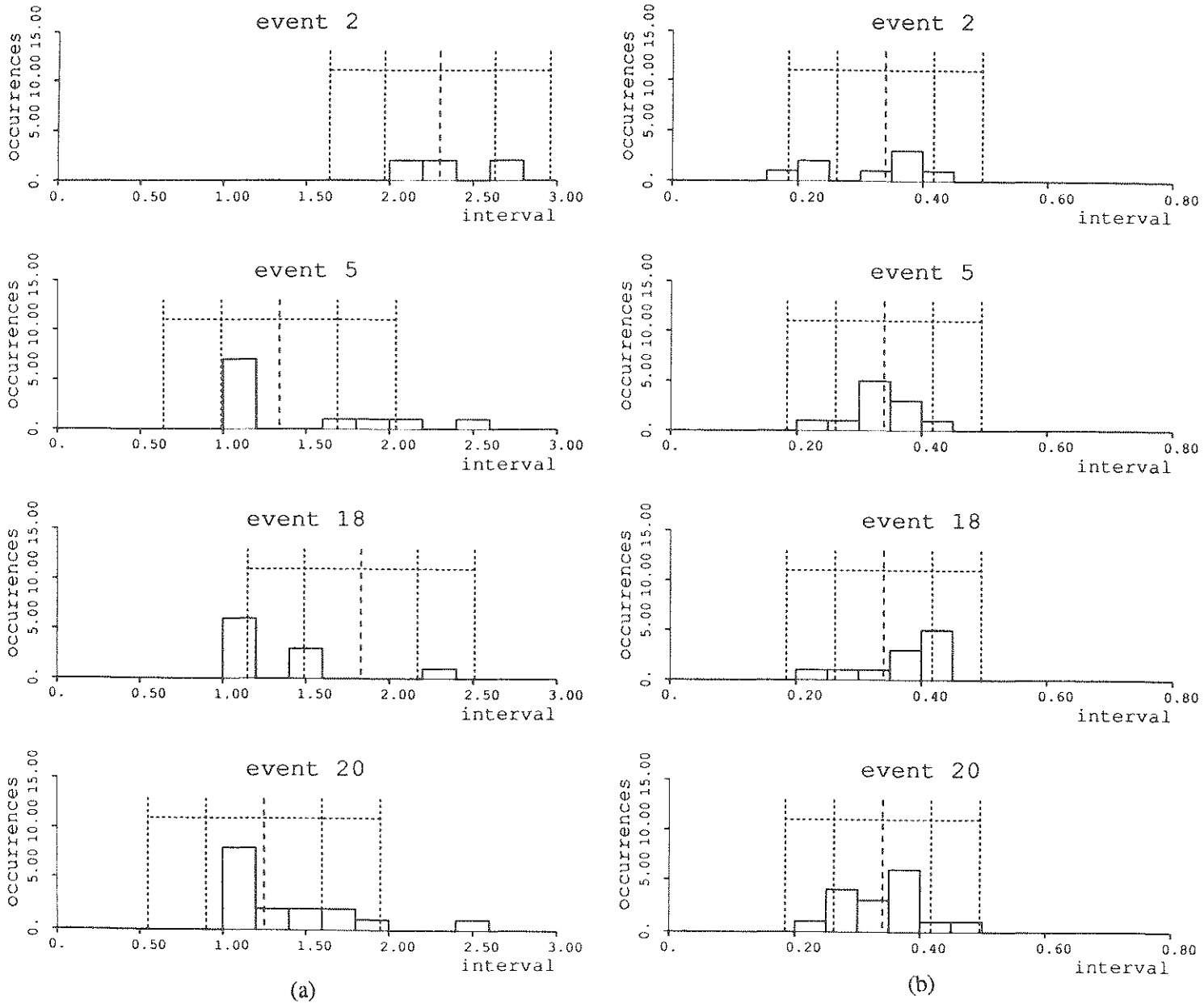


Fig. 6.12 Comparison of the value of (a) t_{max} predicted and t_{max} measured from the second half of the study and (b) k_3 predicted and k_3 measured from the second half of the study.

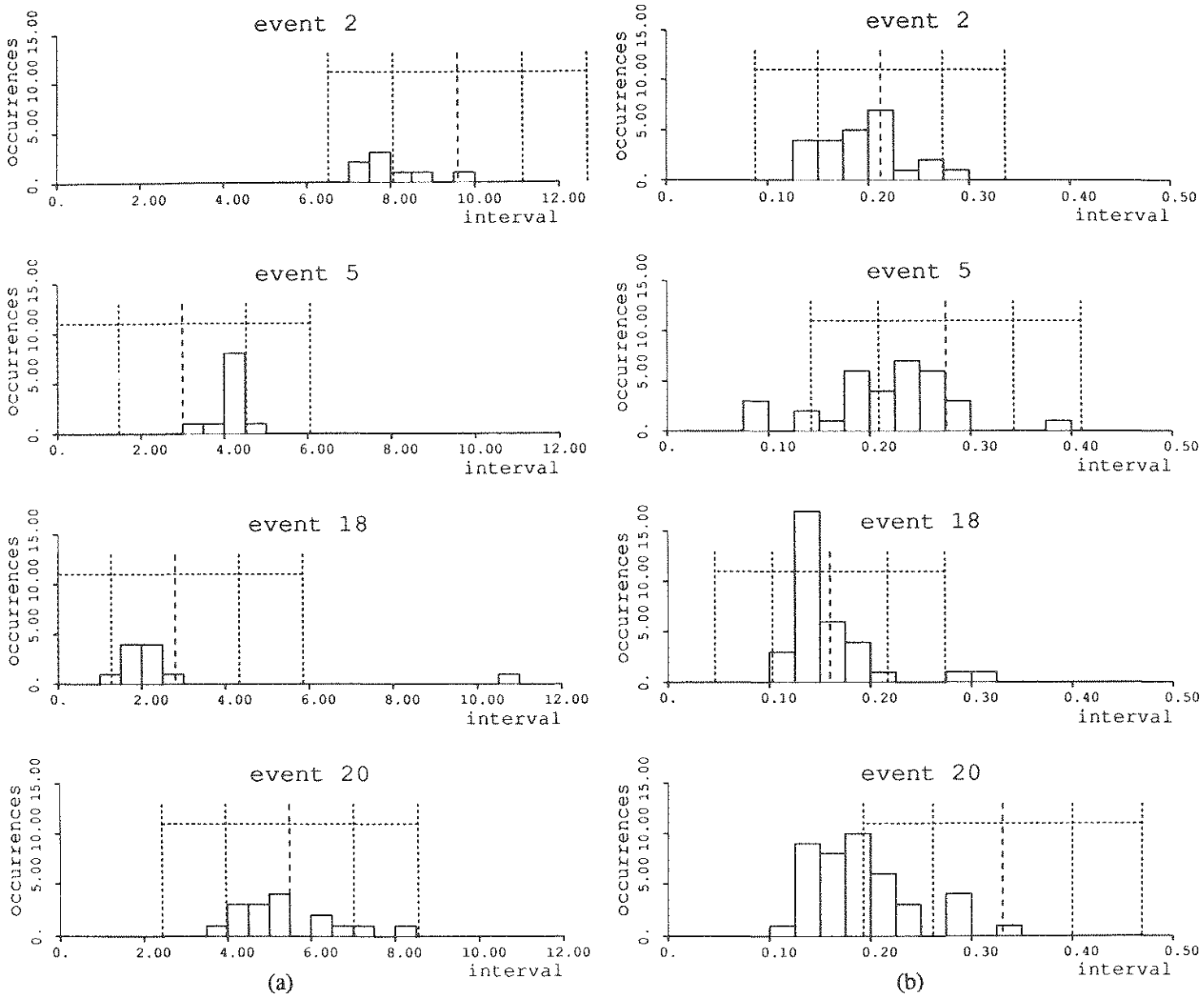


Fig. 6.13 Comparison of the value of (a) r_3 predicted and r_3 measured from the second half of the study and (b) $F(f_{max})$ predicted and $F(f_{max})$ measured from the second half of the study.

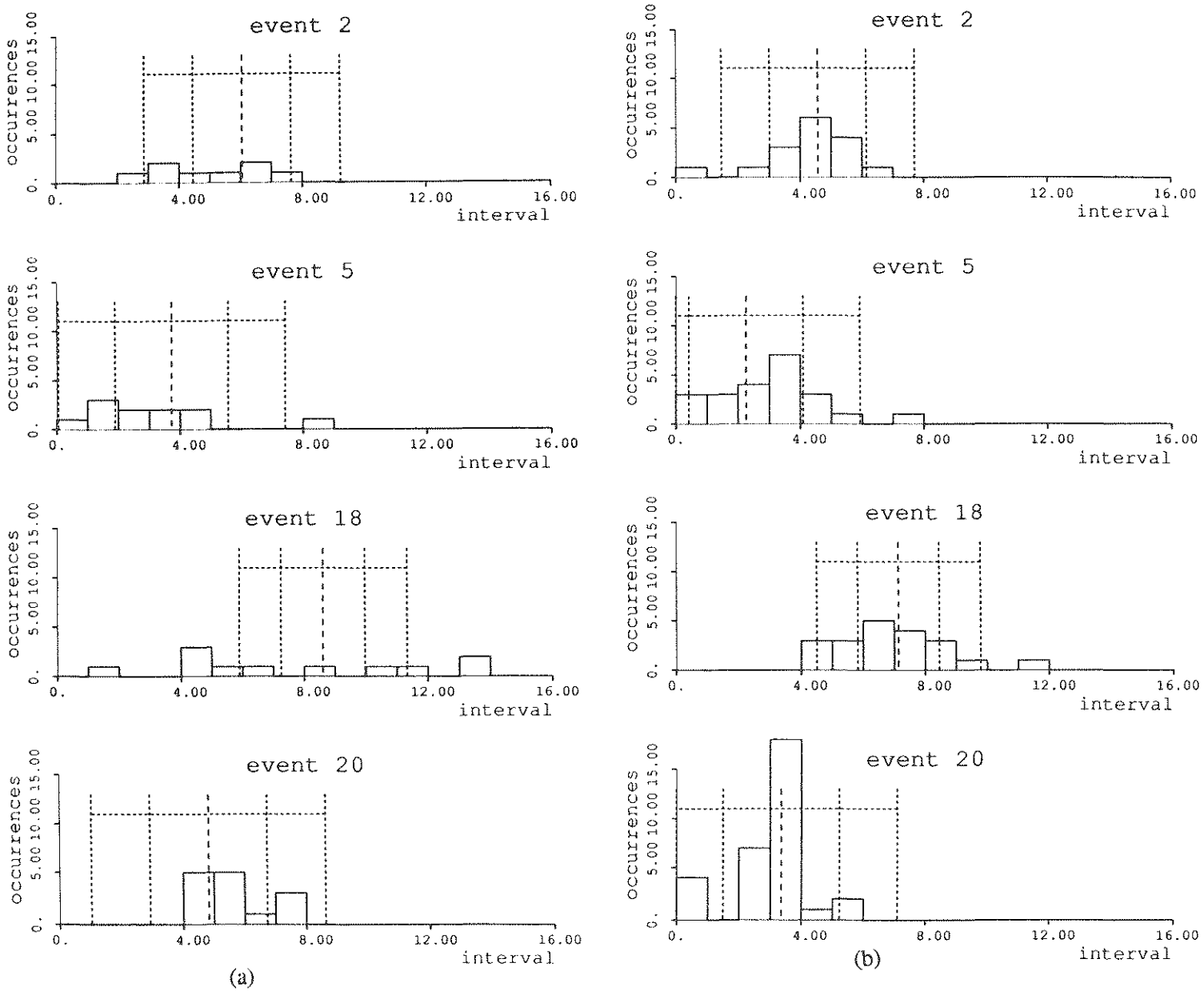


Fig. 6.14 Comparison of the value of (a) f_{max} (vertical) predicted and f_{max} (vertical) measured from the second half of the study and (b) f_{max} (horizontal) predicted and f_{max} (horizontal) measured from the second half of the study.

distribution of the values estimated from the original accelerograms. The mean value calculated from Table 6.2 and one and two standard error limits are indicated on each figure. It was found that the modelling parameters calculated from Table 6.2 in general predict the value of the modelling parameters for the second half of the study well. The one exception is $F(f_{max})$ predicted for event 20 for which half of the data falls outside the two standard error limit.

Following the procedure presented in Section 5.7, simulations were generated using modelling parameters calculated from Table 6.2 for each event. Simulations were also generated for α, f_{max}, and $F(f_{max})$ equal to the mean value calculated from Table 6.2 plus or minus two standard errors. For some cases this was not possible, however, because subtracting two standard errors from the mean value resulted in values of the parameters outside the possible range. As in Section 5.7, upper and lower envelopes of the frequency spectra were calculated based upon the simulations.

The simulated accelerograms and the Fourier and response spectra calculated from the simulations are shown in Figs. 6.15 to 6.26. A comparison of the simulated records with typical accelerograms recorded at station O03 is presented in Figs. 6.15, 6.18, 6.21, and 6.24. When comparing the time histories the two most important features of the simulations are the intensity of shaking and the duration of strong shaking. The duration of strong shaking is similar for all cases. The intensity of the strong shaking compares very well for event 20, is close for events 2 and 18, and is poorest for event 5.

Examing the differences between the frequency spectra is more useful in that all of the original records are shown and the envelopes measuring the possible extreme values are also presented. In general, the simulations match the original frequency spectra well. The only simulations which do not simulate the location of the dominant frequency and the shape of the frequency spectra well are the vertical components for events 5 and 20. The original data is generally bracketed well between the upper and lower envelopes. However, the difference between the two envelopes in some cases is quite large, indicating a great variability in possible ground motion.

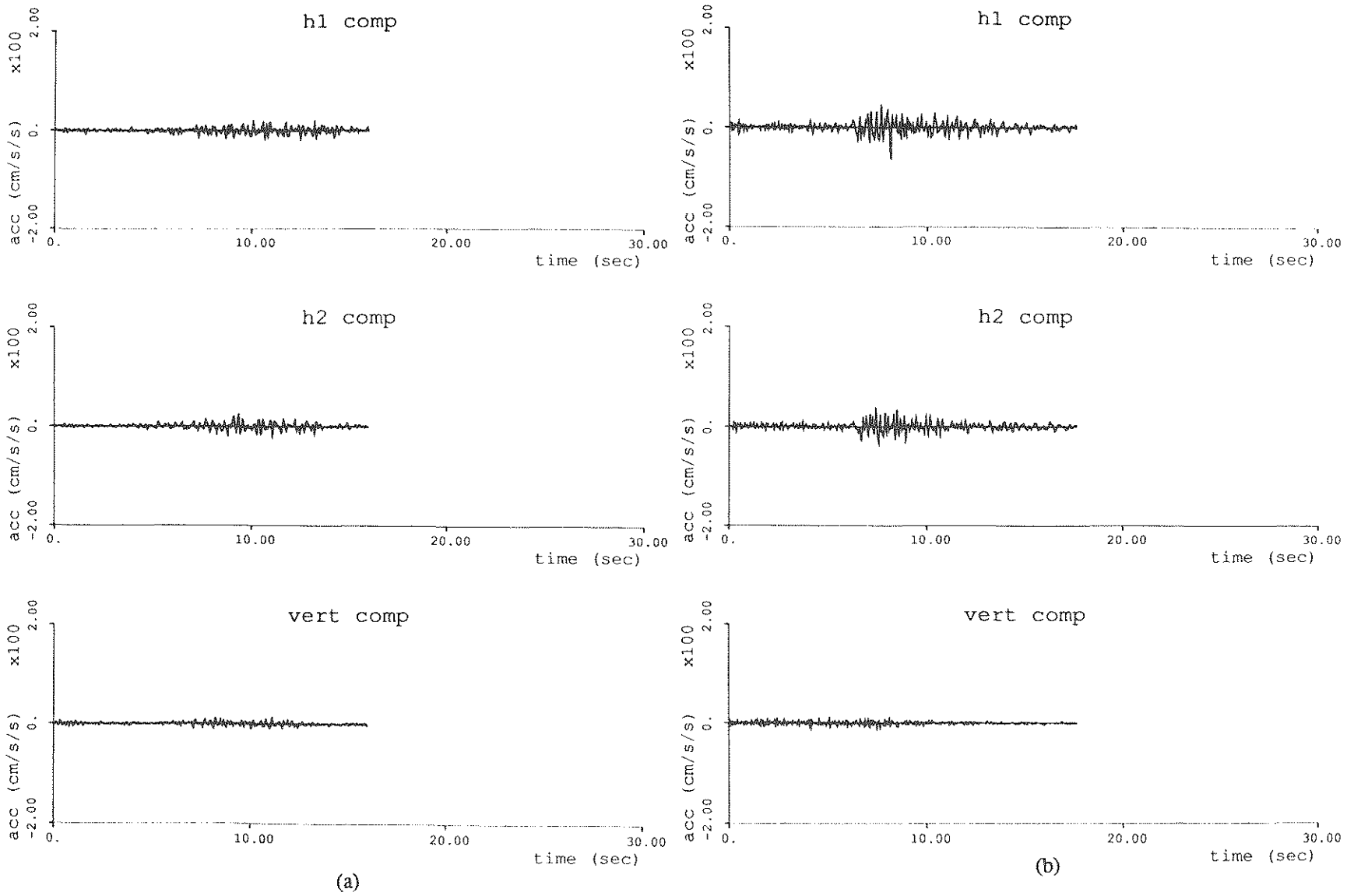


Fig. 6.15 Comparison of (a) simulation of event 2 accelerogram using parameters calculated from Table 6.2 with (b) an original accelerogram recorded at station O03.

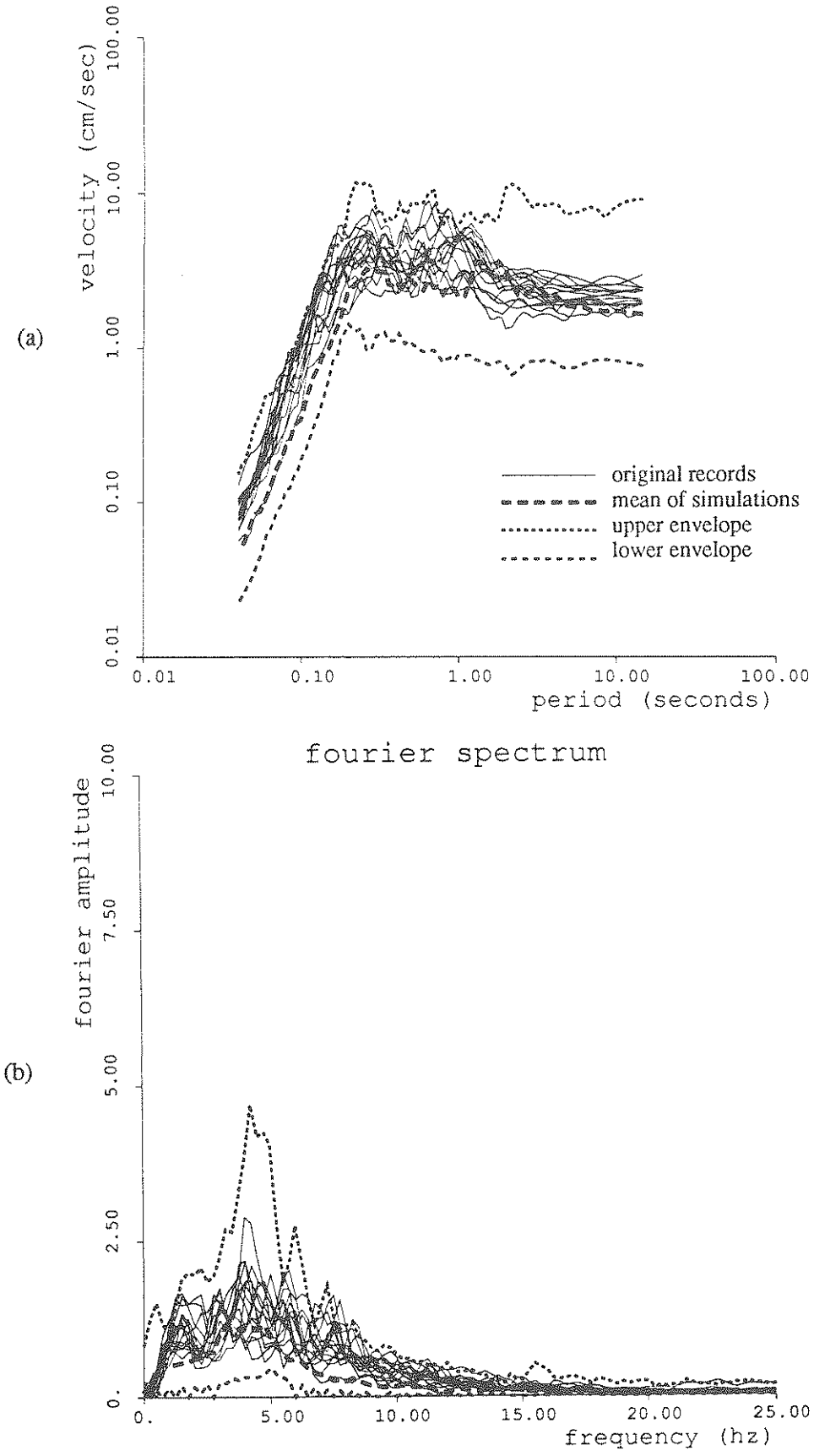


Fig. 6.16 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original recordings of event 2 and simulations using parameters calculated from Table 6.2.

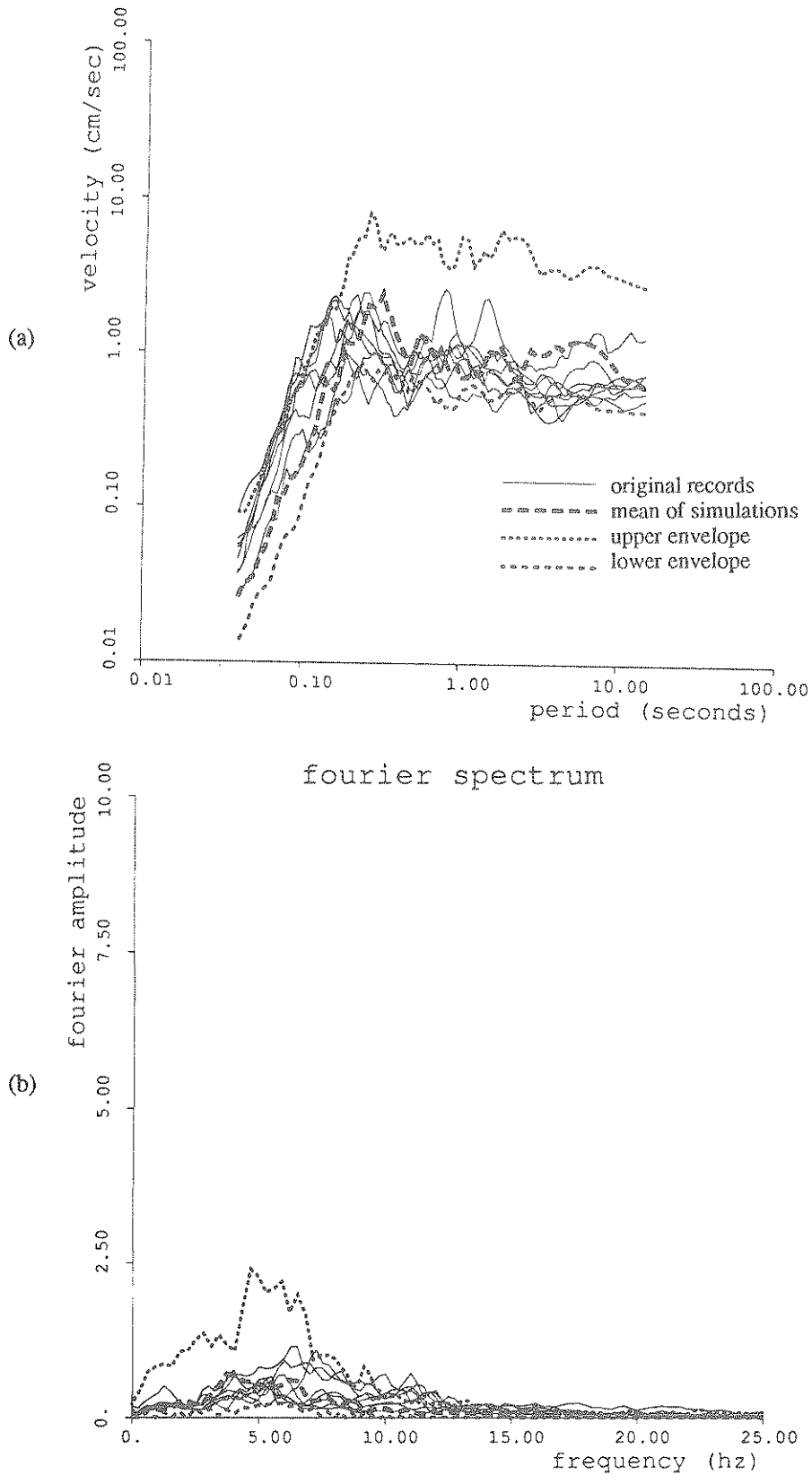


Fig. 6.17 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original recordings of event 2 and simulations using parameters calculated from Table 6.2.

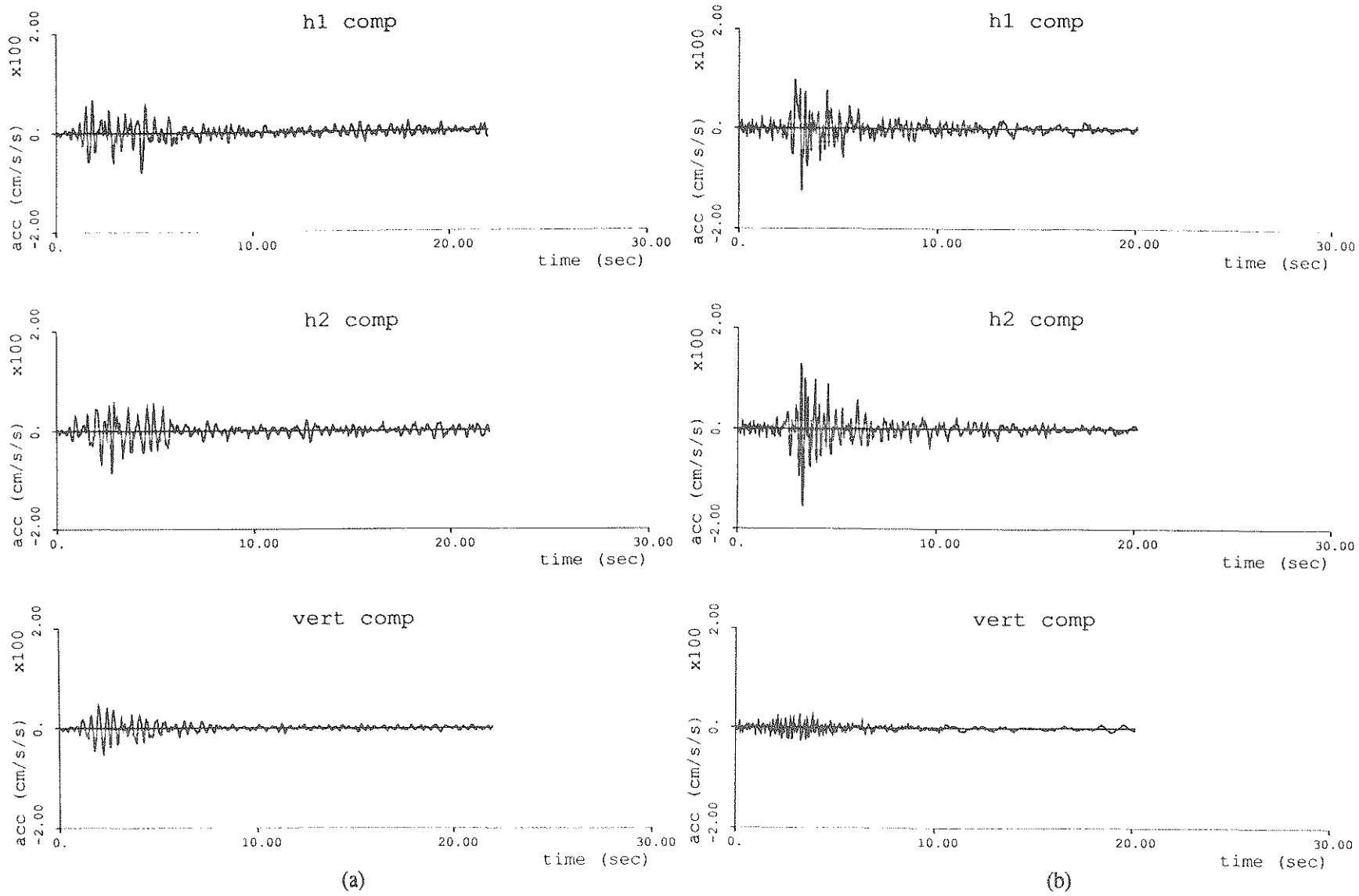


Fig. 6.18 Comparison of (a) simulation of event 5 accelerogram using parameters calculated from Table 6.2 with (b) an original accelerogram recorded at station O03.

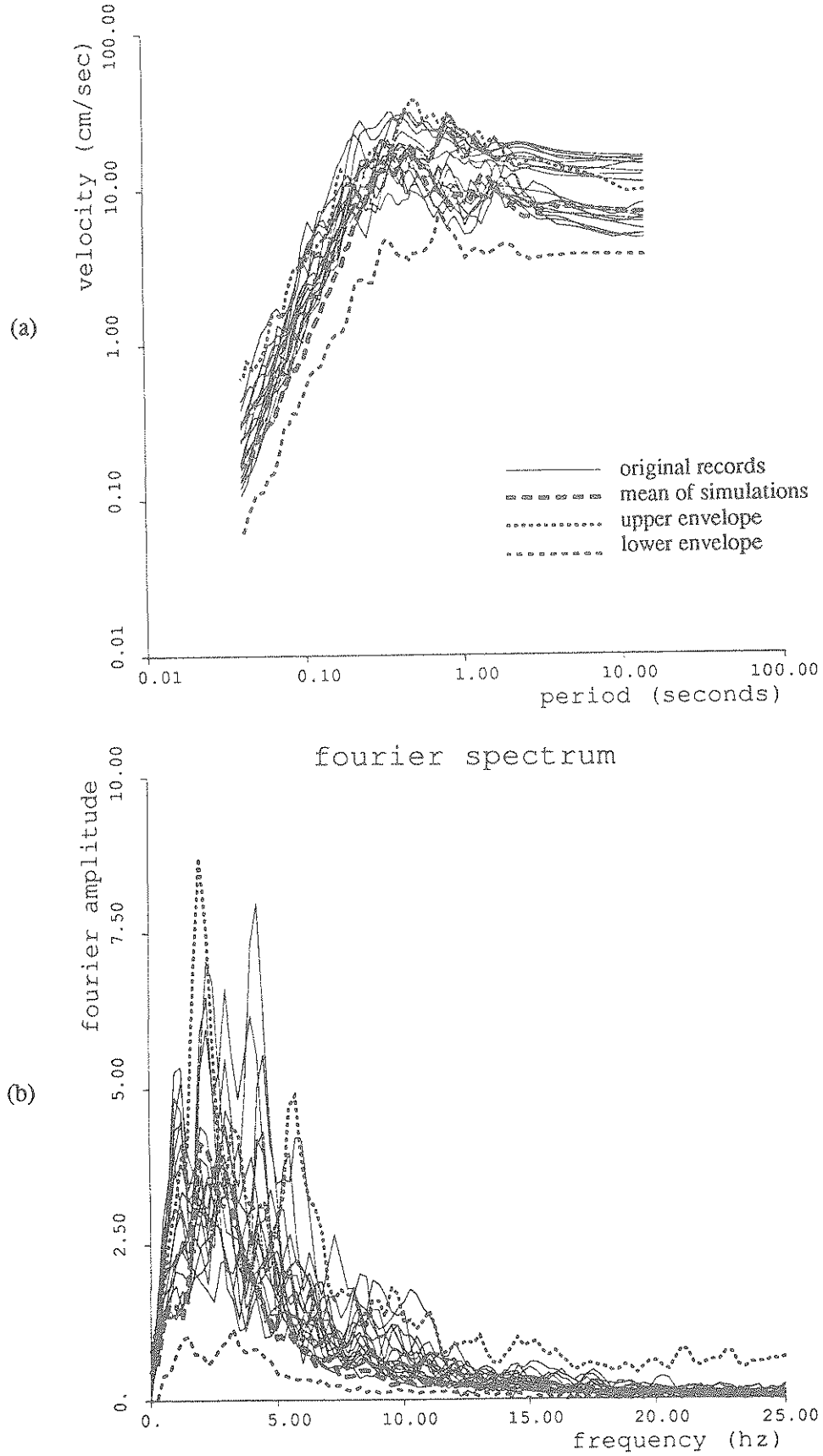


Fig. 6.19 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original recordings of event 5 and simulations using parameters calculated from Table 6.2.

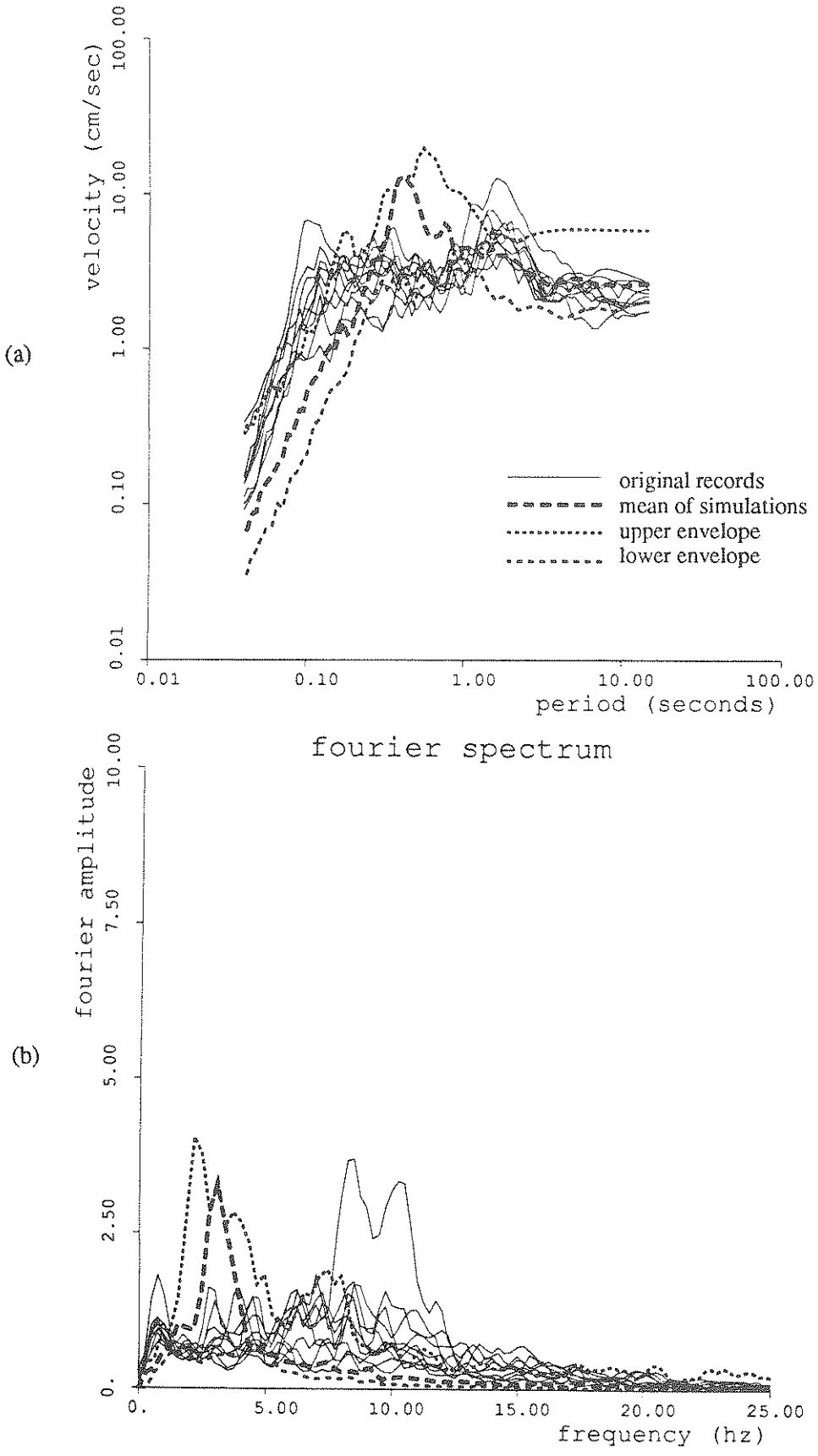


Fig. 6.20 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original recordings of event 5 and simulations using parameters calculated from Table 6.2.

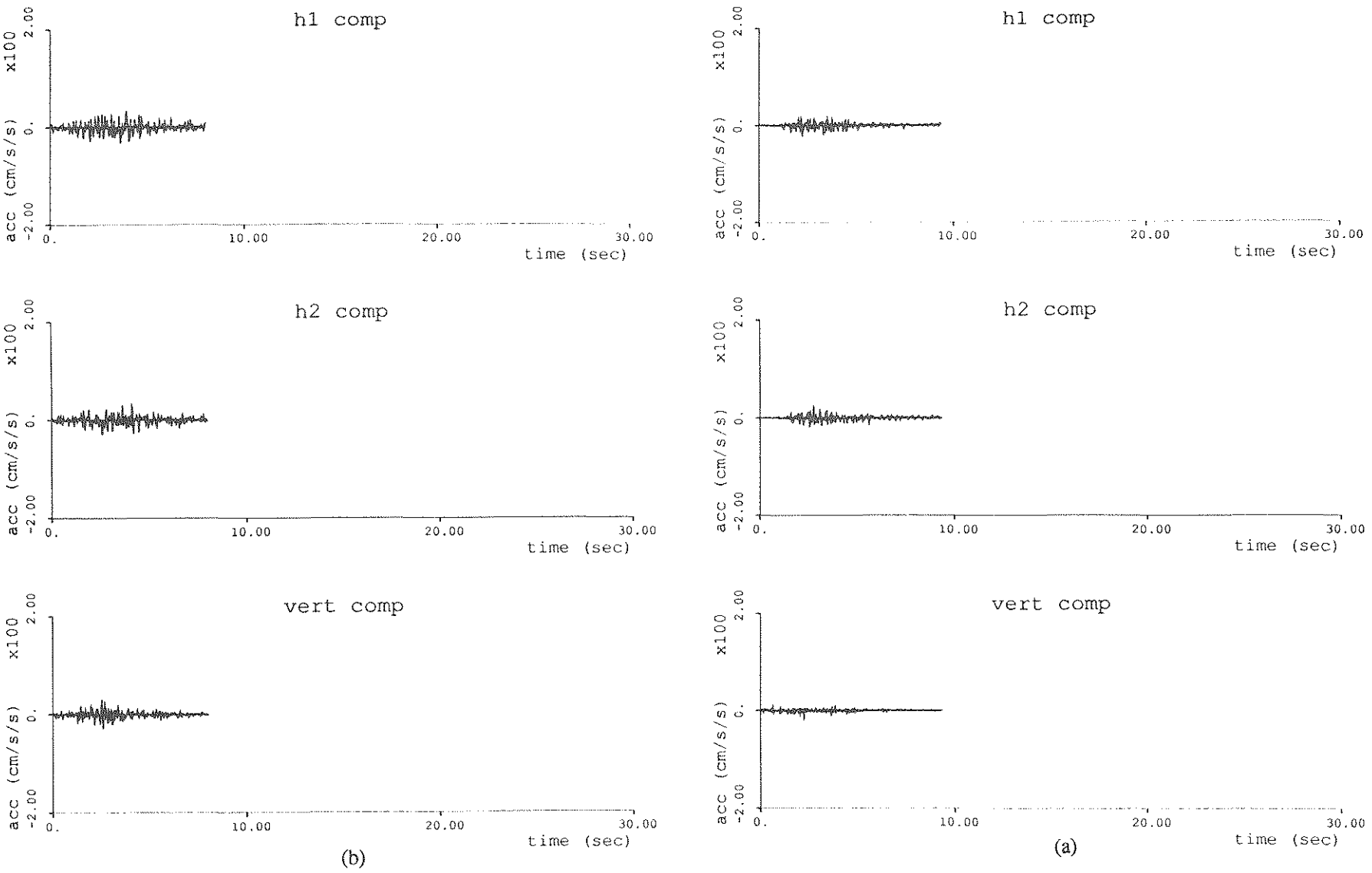


Fig. 6.21 Comparison of (a) simulation of event 18 accelerogram using parameters calculated from Table 6.2 with (b) an original accelerogram recorded at station O03.

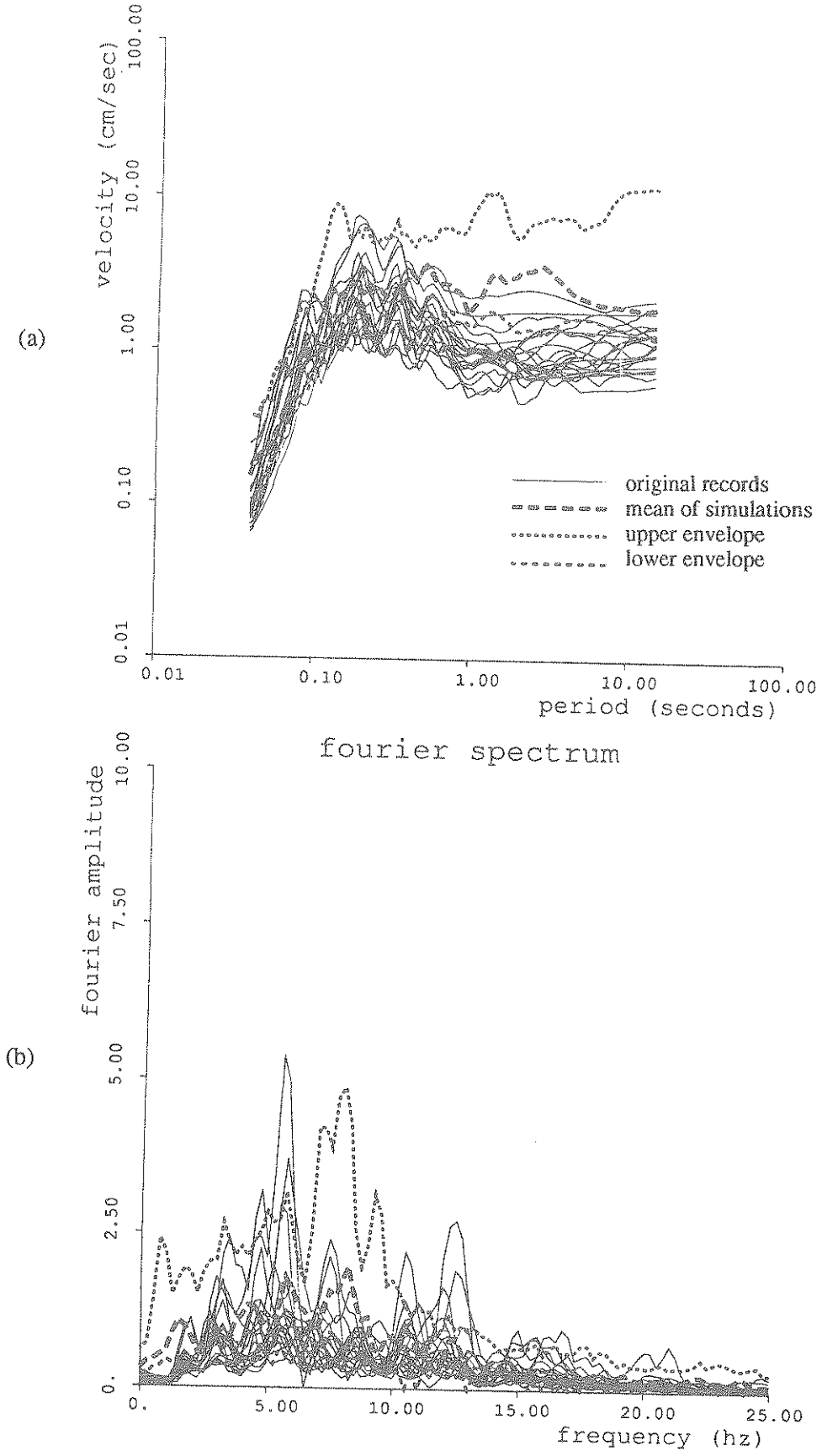


Fig. 6.22 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original recordings of event 18 and simulations using parameters calculated from Table 6.2.

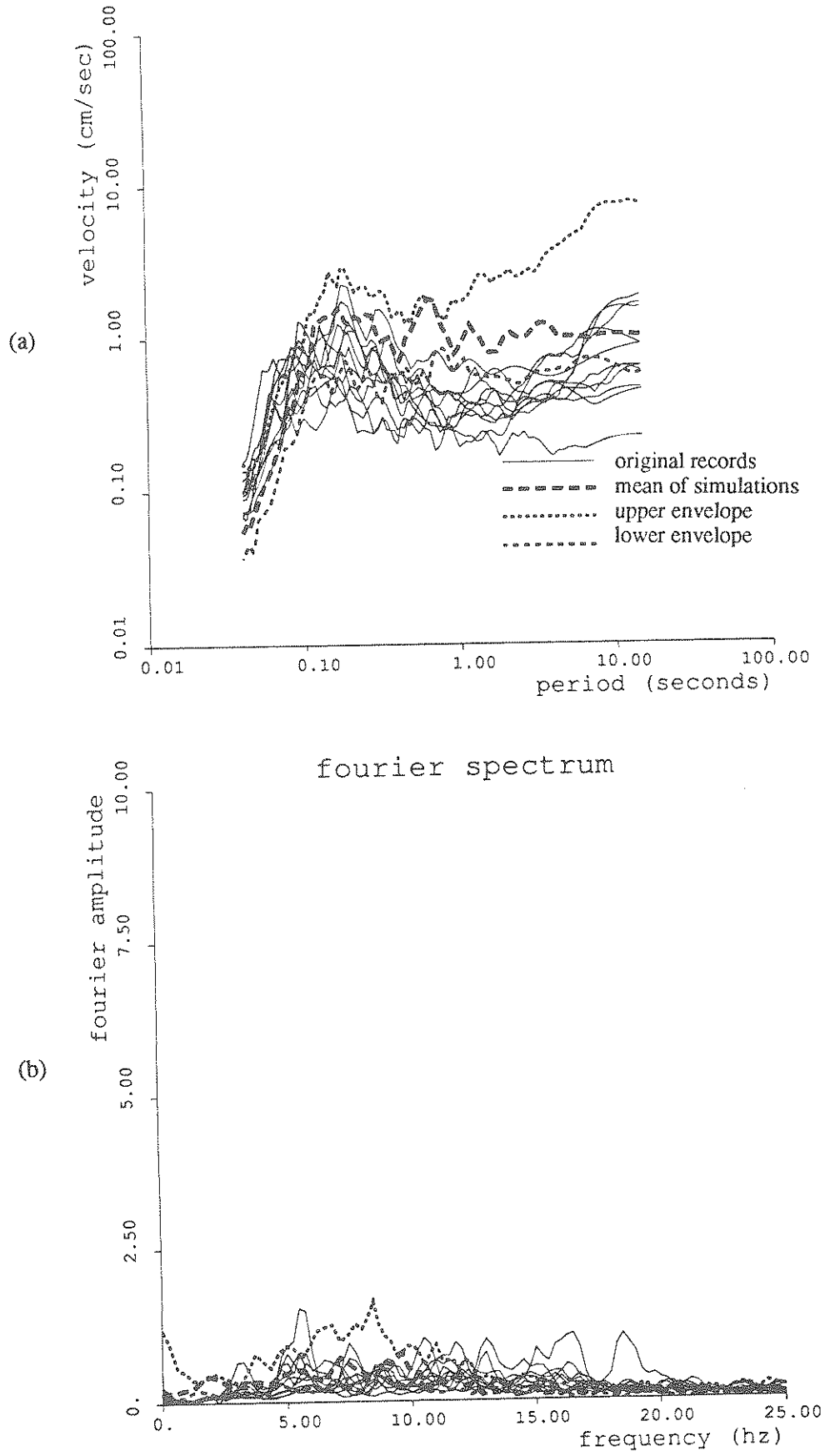


Fig. 6.23 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original recordings of event 18 and simulations using parameters calculated from Table 6.2.

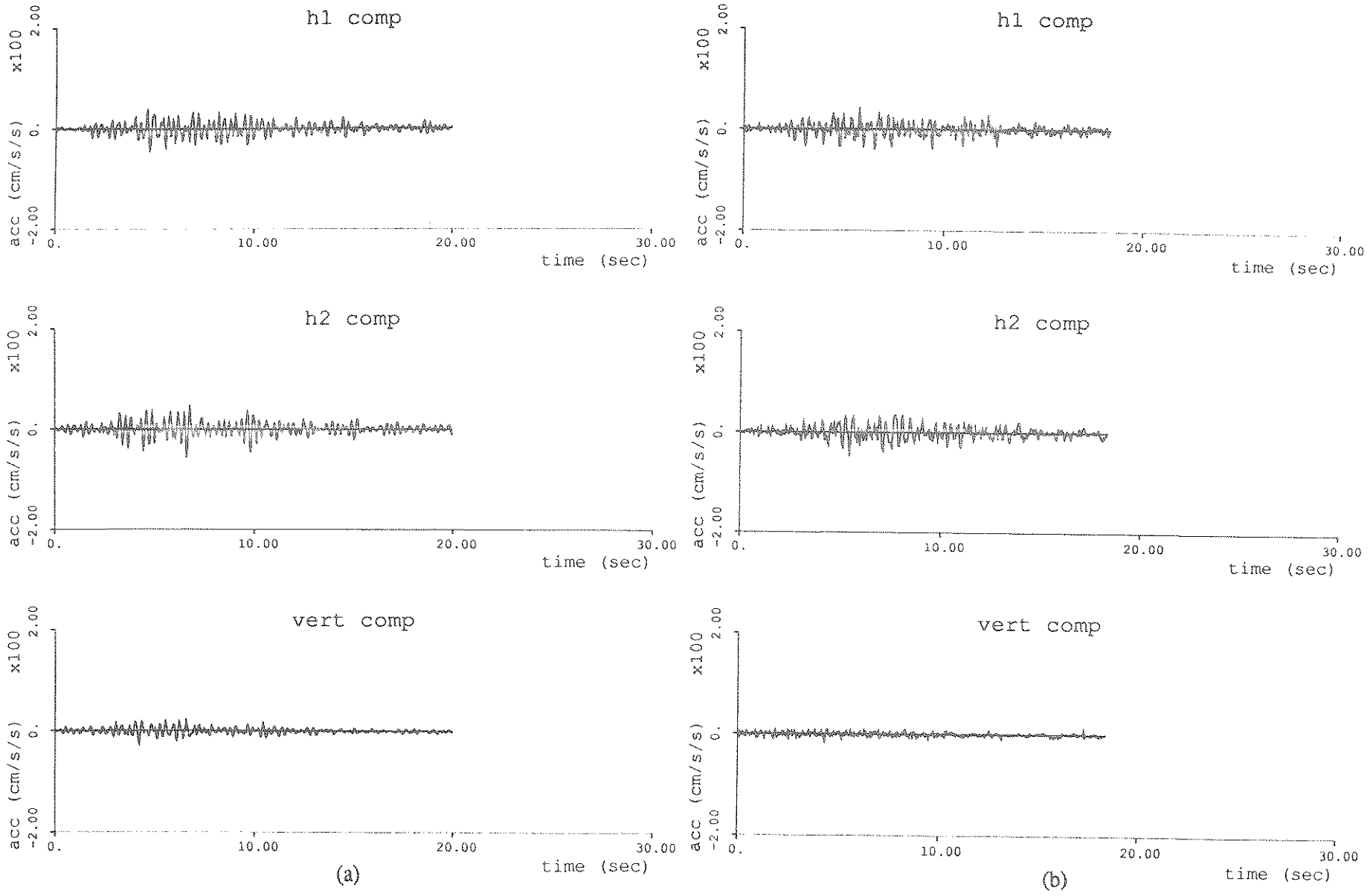


Fig. 6.24 Comparison of (a) simulation of event 20 accelerogram using parameters calculated from Table 6.2 with (b) an original accelerogram recorded at station O03.

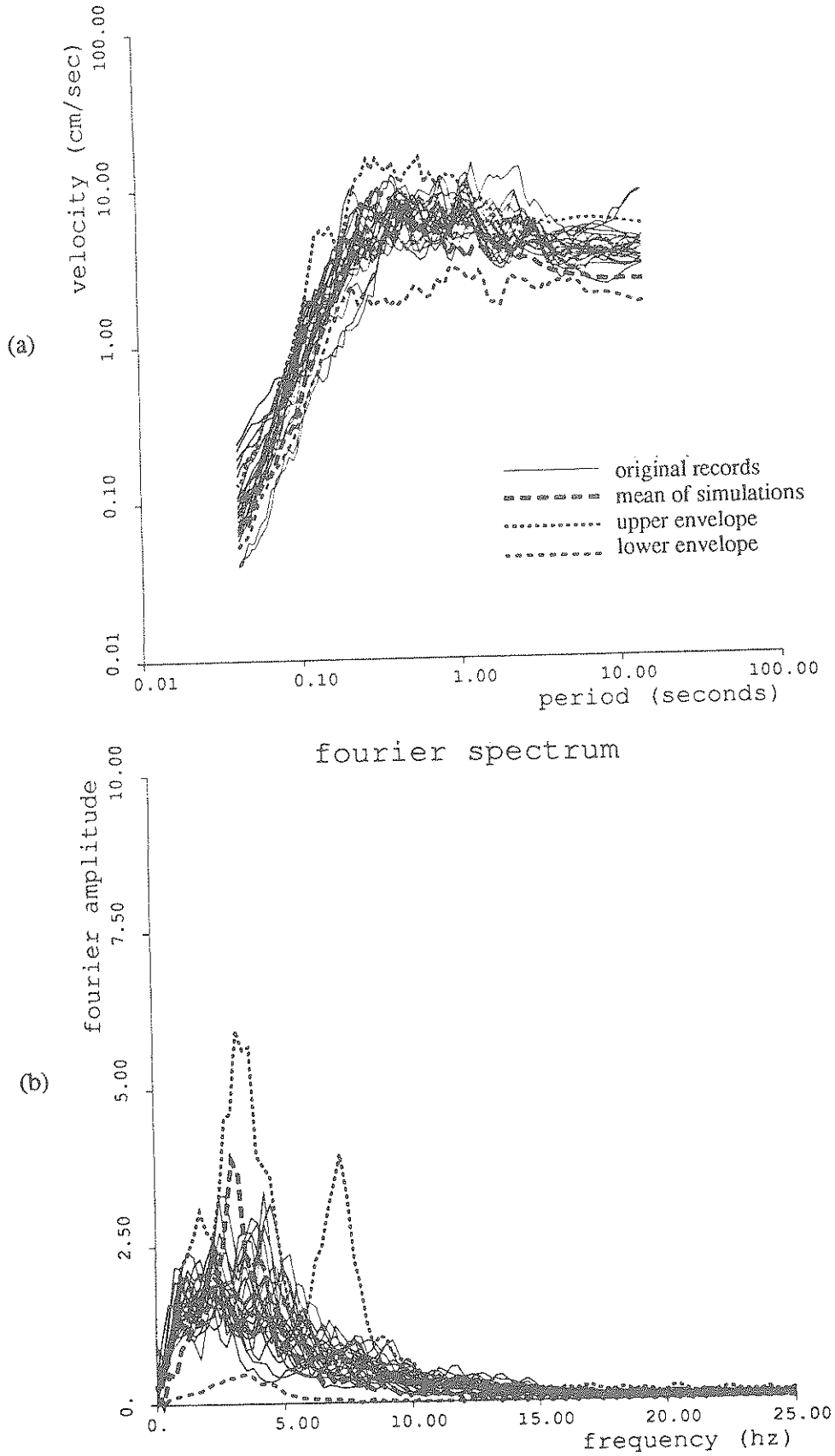


Fig. 6.25 Comparison of (a) response spectrum and (b) Fourier spectrum for the horizontal components of the original recordings of event 20 and simulations using parameters calculated from Table 6.2.

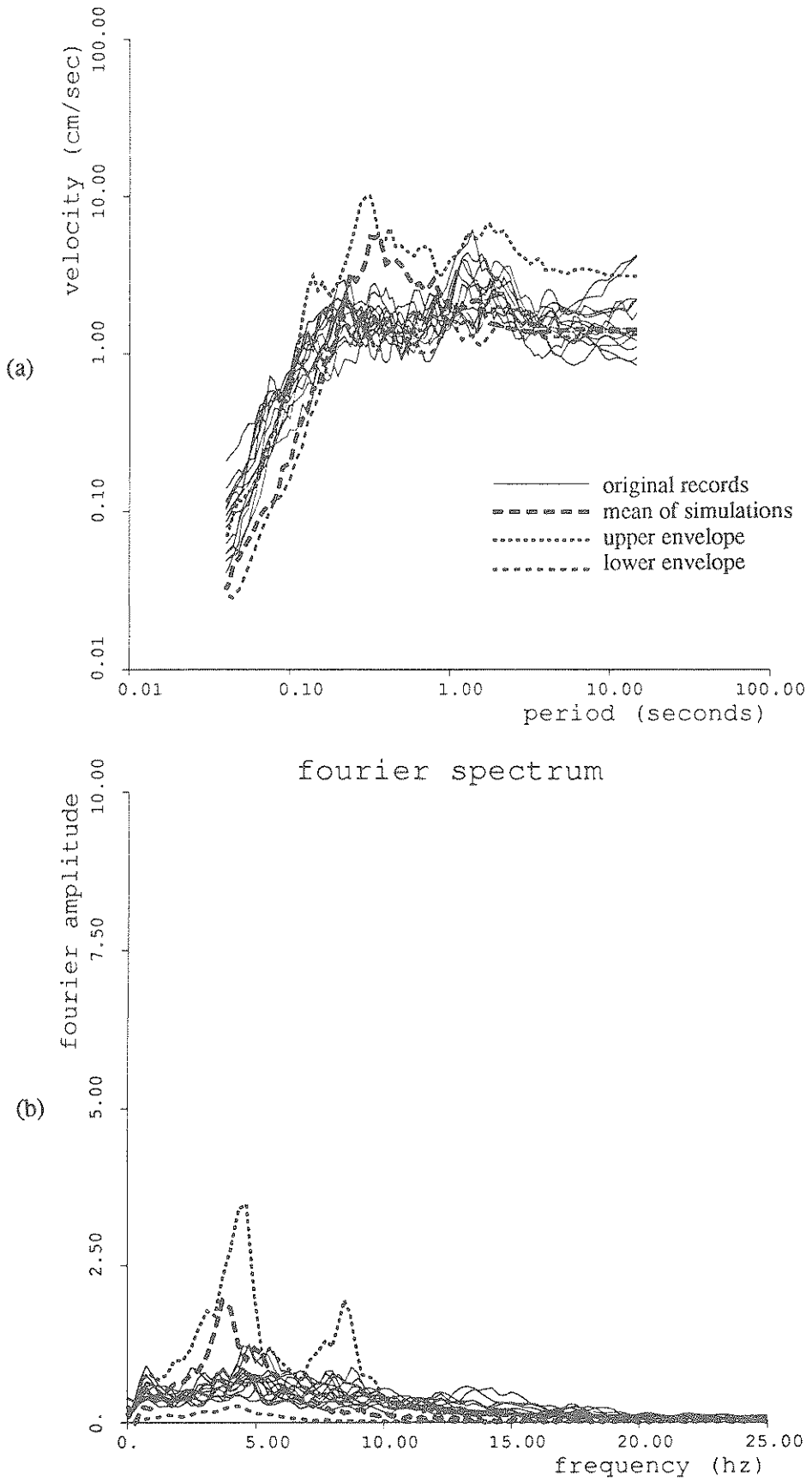


Fig. 6.26 Comparison of (a) response spectrum and (b) Fourier spectrum for the vertical components of the original recordings of event 20 and simulations using parameters calculated from Table 6.2.

References

Bolt, B. A., et al. (1982). *Preliminary Report on the SMART 1 Strong Motion Array in Taiwan,* Technical Report UCB/EERC-82/13, Earthquake Engineering Research Center, University of Calfornia, Berkely, California.

Harichandran, R. and E. Vanmarcke (1984). *Space-Time Variation of Earthquake Ground Motion,* Massachusetts Institute of Technology Technical Report, Cambridge, Massachusetts.

Herrera, I., E. Rosenblueth, and O. A. Rascon (1965). "Earthquake Spectrum Prediction for the Valley of Mexico," *Proceedings of the Third World Conference on Earthquake Engineering,* p. 61.

U. S. Nuclear Regulatory Commission (1978). *Geotechnical and Strong Motion Earthquake Data from U. S. Accelerograph Stations,* U. S. Nuclear Regulatory Commission, Technical Report NUREG-0029.

7. CONCLUSIONS

7.1. Conclusions

The Autoregressive-Moving Average (ARMA) process has been used to model strong-motion accelerograms stabilized by variance and frequency transformations. This was found to be an efficient method for describing the ground motions through a small number of parameters. By relating these parameters to physical variables, it is possible to reasonably predict the ground motion of a site in the examined regions where no strong motion data has been recorded.

Two procedures have been developed for modelling strong-motion accelerograms. The first method is a univariate procedure to model individual components of an accelerogram. The second method is a multivariate procedure to model the three accelerogram components simultaneously.

The univariate and multivariate procedures were used to model free-field, strong-motion accelerograms recorded in California, Mexico, and Taiwan resulting in the following models:

(1) A univariate model for modelling the ground motion in California based upon the earthquake magnitude, epicentral distance, and site location.

(2) A multivariate model for various soil conditions in Mexico City for ground motion expected from a magnitude 8+ earthquake along the Pacific Coast of Mexico.

(3) A multivariate model for sites located on rock at varying distances from a magnitude 8+ earthquake along the Pacific coast of Mexico.

(4) A multivariate model for any site condition, epicentral distance, or magnitude earthquake in Mexico.

(5) A multivariate model based on the SMART-1 data for northern Taiwan.

These models can be used to simulate the ground motion expected for an earthquake based upon the physical variables of the site and the earthquake. Because there is

variability in the ground motion for a given set of physical parameters, the standard errors of the regression coefficients relating the modelling parameters to physical variables were calculated. From the standard errors, confidence intervals can be calculated for the modelling parameters. By varying the values of the modelling parameters within these intervals, simulations possessing the possible range of frequency content, duration, and intensity may be computed.

7.2. Suggestions for future research

There are three main areas in which this work may be extended in the future: (1) further analysis of strong-motion accelerograms, (2) applying probability concepts to generate simulations for a certain level of risk, and (3) applying the modelling procedure to seismograms recorded in the eastern United States.

There exists a wealth of strong-ground motion data recorded in Japan which may be analyzed using the modelling procedures developed. Because of the subduction faulting in Japan, a comparison of the modelling parameters calculated in Japan with the results presented in this report will be useful in examining the effects of the faulting mechanism and the depth of the energy release on the modelling parameters.

A second region that may be studied is the eastern United States and Canada. A number of strong-motion records recorded from the 1982 New Brunswick earthquake exists. Additional east coast records are being collected by the National Center for Earthquake Engineering Research and will be made available for analysis.

Confidence intervals for the parametric relations calculated in the Mexico and Taiwan study have been presented in this report. By combining this information with the uncertainty in the earthquake magnitude and the location of the rupture zone on the fault, it will be possible to generate simulations for a given level of risk.

Finally, the application of the modelling procedure may be extended to seismograms recorded in the eastern United States. The analysis of seismograms will allow the modelling procedure to be applied to regions where no strong-motion data is available. A comparison between the modelling parameters estimated for east coast and

west coast seismograms may then be made.

APPENDIX

A.1 Derivation of Eq. (2.7.4)

By deriving a relationship between the variance of the white noise sequence, σ_a^2, and the variance of the time series, σ_z^2, it is possible to express the white noise variance as a function of the ARMA parameters for series with constant variance. From Box and Jenkins (1976), the variance of the time series, σ_z^2 or γ_0, and the covariance at lags 1 and 2, γ_1 and γ_2, of an ARMA (2,2) model may be expressed as

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma_a^2 - \sigma_a^2\theta_1(\phi_1 - \theta_1) - \sigma_a^2\theta_2[-\theta_2 + \phi_2 + \phi_1^2 - \theta_1\phi_1] \tag{A.1.1}$$

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 - \theta_1\sigma_a^2 - \theta_2\sigma_a^2(\phi_1 - \theta_1) \tag{A.1.2}$$

$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0 - \theta_2\sigma_a^2 \tag{A.1.3}$$

By letting

$$c_1 = 1 - \theta_1(\phi_1 - \theta_1) - \theta_2[-\theta_2 + \phi_2 + \phi_1^2 + \theta_1\phi_1] \tag{A.1.4}$$

$$c_2 = -\theta_1 - \theta_2(\phi_1 - \theta_1) \tag{A.1.5}$$

$$c_3 = -\theta_2 \tag{A.1.6}$$

Equations (A.1.1), (A.1.2), and (A.1.3) can be put in matrix form as

$$\begin{bmatrix} 1 & -\phi_1 & -\phi_2 \\ -\phi_1 & 1-\phi_2 & 0 \\ -\phi_2 & -\phi_1 & 1 \end{bmatrix} \begin{Bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{Bmatrix} = \begin{Bmatrix} c_1\sigma_a^2 \\ c_2\sigma_a^2 \\ c_3\sigma_a^2 \end{Bmatrix} \tag{A.1.7}$$

Using Kramer's rule to solve for γ_0, the relationship between the white noise variance and the time series variance expressed in Eq. (2.7.4) can be derived as

$$\sigma_a^2 = \frac{(1-\phi_2)(1-\phi_2^2) - \phi_1^2(1-\phi_2)}{(1-\phi_2)c_1 + (1+\phi_2)\phi_1 c_2 + (1-\phi_2)\phi_2 c_3}\sigma_z^2 \tag{A.1.8}$$

For time series stabilized to a variance of 1.0, this relation reduces to

$$\sigma_a^2 = \frac{(1-\phi_2)(1-\phi_2^2) - \phi_1^2(1-\phi_2)}{(1-\phi_2)c_1 + (1+\phi_2)\phi_1 c_2 + (1-\phi_2)\phi_2 c_3} \tag{A.1.9}$$

where the white noise variance is a function only of the ARMA parameters.

A.2 Derivation of Eq. (4.3.2.2)

As shown in the previous section, it is possible to derive a relationship between the white noise variance and the ARMA parameters fitted to a stabilized series. In this section the relationship will be derived for the ARMA (3,1) model used in the multivariate modelling procedure.

From Box and Jenkins (1976), the variance of the time series and the covariance at lags 1, 2, and 3 of an ARMA (3,1) model may be expressed as

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \phi_3\gamma_3 + \sigma_a^2[1 - \theta_1(\phi_1 - \theta_1)] \tag{A.2.1}$$

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 + \phi_3\gamma_2 - \theta_1\sigma_a^2 \tag{A.2.2}$$

$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0 + \phi_3\gamma_1 \tag{A.2.3}$$

$$\gamma_3 = \phi_1\gamma_2 + \phi_2\gamma_1 + \phi_3\gamma_0 \tag{A.2.4}$$

Equations (A.2.1), (A.2.2), (A.2.3), and (A.2.4) can be put in matrix form as

$$\begin{bmatrix} 1 & -\phi_1 & -\phi_2 & -\phi_3 \\ -\phi_1 & 1-\phi_2 & -\phi_3 & 0 \\ -\phi_2 & -\phi_1-\phi_3 & 1 & 0 \\ -\phi_3 & -\phi_2 & -\phi_1 & 1 \end{bmatrix} \begin{Bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{Bmatrix} = \begin{Bmatrix} \sigma_a^2[1 - \theta_1(\phi_1 - \theta_1)] \\ -\theta_1\sigma a^2 \\ 0 \\ 0 \end{Bmatrix} \tag{A.2.5}$$

By solving this set of equations for the variance, γ_0, the relationship between the white noise variance and the time series variance may be found as

$$\gamma_0 = \frac{c_6 - c_7}{c_1 - c_2 - c_3 - c_4 - c_5}\sigma_a^2 \tag{A.2.6}$$

$$c_1 = 1 - \phi_2 - \phi_3(\phi_1 + \phi_3) - \phi_1(\phi_1 + \phi_2\phi_3)$$

$$c_2 = \phi_2[\phi_1(\phi_1 + \phi_3) + \phi_2(1 - \phi_2)]$$

$$c_3 = \phi_1\phi_3[\phi_1(\phi_1 + \phi_3) + \phi_2]$$

$$c_4 = \phi_3(1 - \phi_2)(\phi_1\phi_2 + \phi_3)$$

$$c_5 = \phi_3^2[\phi_2^2 - \phi_3(\phi_1 + \phi_3)]$$

$$c_6 = [1 - \theta_1(\phi_1 - \theta_1)][1 - \phi_2 - \phi_1\phi_3 - \phi_3^2]$$

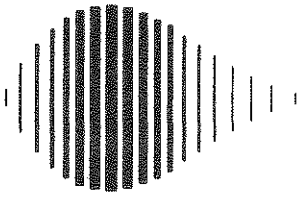
$$c_7 = \theta_1[\phi_1 + \phi_2(\phi_1 + \phi_3) + \phi_3(\phi_1^2 + \phi_1\phi_3 + \phi_2)]$$

By setting the time series variance equal to one, the white noise variance for any component i may be calculated as

$$\sigma_i^2 = \frac{c_1 - c_2 - c_3 - c_4 - c_5}{c_6 - c_7} \qquad \text{(A.2.7)}$$

References

Box, G. E. P. and G. M. Jenkins (1976). *Time Series Analysis Forecasting and Control,* Holden-Day, San Francisco.



National Center for Earthquake Engineering Research
State University of New York at Buffalo