

**DEPARTMENT OF THE INTERIOR
U. S. GEOLOGICAL SURVEY**

**Analysis of Recordings in Structural Engineering:
Adaptive Filtering, Prediction, and Control**

by

Erdal Şafak

**U. S. Geological Survey
Menlo Park, California**



Open-File Report 88-647

This report is preliminary and has not been reviewed for conformity with U.S. Geological Survey editorial standards and stratigraphic nomenclature. Any use of trade names is for descriptive purposes only and does not imply endorsement by the U.S. Geological Survey.

October 1988

CONTENTS

	Page
ABSTRACT	iii
1. INTRODUCTION	1
2. DISCRETE-TIME REPRESENTATION OF CONTINUOUS SYSTEMS	3
3. DISCRETE-TIME MODELS FOR SISO SYSTEMS	6
4. MODELS FOR UNKNOWN AND NOISY SYSTEMS	10
5. SYSTEM IDENTIFICATION	11
5.1. One-step-ahead prediction	11
5.2. Weighted least-squares method	12
5.3. Maximum likelihood method	14
5.4. Stochastic approximation	15
5.5. Recursive prediction error method (RPEM)	17
6. CONVERGENCE AND CONSISTENCY OF ESTIMATION	21
7. FISHER INFORMATION MATRIX AND CRAMER-RAO INEQUALITY	22
8. MODEL SELECTION	22
9. MODEL VALIDATION	24
10. SPECTRAL ESTIMATION	26
11. ADAPTIVE CONTROL	27
12. PREPROCESSING OF DATA	29
13. EXAMPLES	31
13.1. Identification of a time-invariant simulated system	31
13.2. Identification of a time-varying simulated system	32
13.3. Identification of a building with soil-structure interaction	34
13.4. Identification of a building with nonlinear behavior	36
13.5. Identification of a building from ambient vibration data	37
13.6. Spectral modeling of earthquake ground motions	37
13.7. Site amplification of earthquake ground motions	39
13.8. Source scaling of earthquake ground motions	41
13.9. Adaptive control of a simulated system	42
13.10. Adaptive control of ambient vibrations of a building	43
14. Discussion and further applications	44
15. Summary and conclusions	45
ACKNOWLEDGMENTS	46
REFERENCES	47
TABLES	52
FIGURES	59

ABSTRACT

Analyses of recordings include processing of data, determining analytical models that match the record, and identification of the system from which the record is obtained.

Current methods that are used to analyze recordings in structural engineering are based on the classical filtering and Fourier analysis approach. These methods assume that: (a) either the signal and noise spectra are nonoverlapping, or there is a frequency band where the signal to noise ratio is high (i.e., noise can be neglected), and (b) the properties of the signal within the selected time window are time-invariant.

Recently, new methods for record analysis have been developed based on the concepts of adaptive filtering and stochastic approximation. These methods are commonly known as the stochastic-adaptive methods. Stochastic-adaptive methods make use of the statistical properties of the record, and integrate filtering, modeling, and identification into a single algorithm. The advantages of stochastic-adaptive methods over the classical methods are: (a) they can remove the noise from the signal over the whole frequency band, (b) they can track time-varying characteristics of the signal, and (c) they make it possible to apply adaptive control on unknown systems.

In this report, a concise theory of stochastic-adaptive methods, and their applications in structural engineering are presented. The theoretical part includes the following topics: discrete models for dynamic systems, one-step-ahead prediction, stochastic approximation, recursive prediction error method, model selection, model validation, spectral estimation, and adaptive control. The application part presents ten examples by using earthquake, ambient vibration, and simulated data. The examples include identification of time-invariant and time varying simulated systems; identifications of buildings with soil-structure interaction, nonlinear behavior, and ambient vibrations; modeling of spectral shape, site amplification, and source scaling of earthquake ground motions; and adaptive minimum variance control of a simulated system and a building with ambient vibrations.

1. INTRODUCTION

Instrumentation of structural systems to investigate their dynamic behavior under various loads is becoming increasingly popular. Rapid developments in digital recording and computer technologies made instrumentation cheaper and more attractive today than they were before. In structural engineering, instrumentation has been used extensively to measure the quantities related to loads, such as earthquake ground accelerations, wind velocities, wave heights, and blast pressures. The development of models for structures has been mainly based on the theoretical approach, which makes use of the physical laws that govern the structural system (e.g., Newton's law), and the mechanical properties of the components (e.g., mass, stiffness, damping, etc.). When available, the recordings from structures were used to check the validity of the models. As a result of recent increases in the number of instrumented structures, the use of actual data along with the theoretical approach is becoming popular in structural analysis. The data from instrumented structures can be used to check the validity or to determine parameters of analytical models, and to develop empirical models. Instrumentation is also used for safety evaluation, where the load resistance characteristics of aging structures are determined by measuring their motions. A recent application of structural instrumentation is the fatigue detection of steel offshore platforms (e.g., Ibanez, 1987). The motion of the platform is continuously monitored. A sudden change in the characteristics of the recorded signal usually is a sign of crack initiation due to fatigue.

Increasing use of instrumentation has necessitated faster and more reliable methods for signal processing, modeling, and identification. It is well known that all recordings from dynamic systems contain noise due to mechanical imperfections in the recording instruments, and also due to ambient noise exists in the recording environment. Because of the random effects involved, it is not generally possible to determine the exact structure of this noise, so that it might be completely removed from the signal. The classical signal processing approach has been to remove the frequency components of the record that are dominated by noise by using band-pass filters. For the retained frequencies, although they still contain noise, it is assumed that the actual signal amplitudes are dominant over the noise, and therefore noise can be neglected. A large number of such filters are available in the literature (for detail, see Rabiner and Gold, 1975). Two types, the Butterworth and Ormsby filters, have commonly been used for earthquake recordings (Hudson, 1979).

During the last twenty years, new methods for signal processing based on the concepts of stochastic-adaptive filtering and prediction have been developed. These methods use statistical characteristics of data to filter the noise from the signal. The filter characteristics are initially unknown. They are estimated recursively in the time domain, and adjusted continuously by using the information extracted from the data. Such an approach inte-

grates filtering, modeling, and identification into a single algorithm. The filtering problem becomes equivalent to the estimation of the parameters of two recursive filters, one for the noise and one for the actual signal. Since the noise is generally unknown and random, a stochastic approach rather than a deterministic approach is used in the process. Various terms have been used in the literature to distinguish such signal processing methods, e.g., stochastic, adaptive, on-line, recursive, sequential, and real-time. The term that will be used in this report is stochastic-adaptive.

Stochastic-adaptive methods present four main advantages over the previous methods: (a) removal of the noise from the signal is done over the whole frequency band, which can not be accomplished by classical band-pass filters, (b) because of the recursive form of the algorithms, time-varying characteristics of the signal can be tracked, (c) only a small segment of the data is needed during the computations, and (d) the algorithm makes it possible to apply adaptive control on the system. The development of stochastic-adaptive methods are based on the pioneering works by Kolmogorov (1941) and Wiener (1949), and later by Kalman (1960) and Kalman and Bucy (1961). Today, these techniques are successfully being applied to various practical problems in guidance and navigation, automatic control, speech processing, and econometrics.

In this report basic principles of stochastic-adaptive filtering and prediction techniques are introduced, their use in modeling, identification, and control of discrete-time recordings is presented, and examples for applications in structural dynamics are given. Stochastic-adaptive techniques have been developed rather recently, mainly by researchers in statistics, electrical and control engineering, and econometrics. There are literally hundreds of papers on the subject, scattered in many journals in the above fields. The theory given in this report is concise and limited to that necessary to follow the examples. However, a large list of references is provided in the report for those interested in more detail.

The first part of the report presents the theoretical development, which includes sections 2 through 11. Section 2 gives the relationship between continuous and discrete-time representations of linear systems. Section 3 presents the time-domain and frequency-domain representation of discrete, single-input single-output (SISO) systems. Section 4 introduces a general discrete-time domain model and its special forms for unknown SISO systems with noise. Section 5 presents basic components of recursive identification algorithms, including one-step-ahead prediction, the least-squares and maximum likelihood methods, the concept of stochastic approximation, and the recursive prediction error method. Section 6 discusses the convergence and consistency of the identification, and section 7 gives the limits for the accuracy of the identification. Sections 8 and 9 present methods for model selection and model validity. Section 10 shows the use of the identification algorithm for spectral estimation. Section 11 introduces an adaptive control algorithm

as an extension of the identification algorithm.

The second part of the report, section 12, starts with guidelines for preprocessing the data. Section 13 presents ten examples for the application of the theory, using both simulated data and actual recordings. Examples presented are as follows:

1. identification of a time-invariant simulated system,
2. identification of a time-varying simulated system,
3. identification of a building with soil-structure interaction,
4. identification of a building with nonlinear behavior,
5. identification of a building using ambient vibration data,
6. spectral modeling of ground accelerations,
7. identification of earthquake site amplification,
8. identification of earthquake source scaling,
9. adaptive control of a simulated system, and
10. adaptive control of ambient vibrations of a building.

Section 14 discusses the other applications of the method, and section 15 is the summary and conclusions.

2. DISCRETE-TIME REPRESENTATION OF CONTINUOUS SYSTEMS

Linear dynamic systems are generally described by continuous-time domain ordinary or partial differential equations. Modern recording instruments, however, are all digital and give measurements in the discrete-time domain. Thus, it is appropriate first to show the relationship between continuous and discrete-time representations, and present methods for converting from one to another.

The most straightforward approach to convert from the continuous to discrete domain is to approximate the differentials by difference equations. There are three approximation rules commonly used; they are the forward rectangular rule, backward rectangular rule, and the trapezoid rule. The forward rectangular rule (also known as Euler's approximation) approximates an n th order derivative by the following equation:

$$y^{(n)}(t) = \left[\frac{y(t+1) - y(t)}{T} \right]^n = \left(\frac{q-1}{T} \right)^n y(t) \quad (2.1)$$

where T denotes the sampling interval, q is the shift operator, such that $q^{-k}y(t) = y(t-k)$, and the index $t-k$ enumerates the sampling instant. The forward rectangular rule uses present and future values of $y(t)$. A corresponding approximation using present and past values of $y(t)$ is the backward rectangular rule given by the equation

$$y^{(n)}(t) = \left[\frac{y(t) - y(t-1)}{T} \right]^n = \left(\frac{1 - q^{-1}}{T} \right)^n y(t) \quad (2.2)$$

If T is not sufficiently small (in comparison with the smallest period in the signal), rectangular approximations can give erroneous results because of the accumulation of errors. A better approximation is given by the trapezoid rule, or the so-called Tustin's method (Tustin, 1947), where the same derivative is approximated as

$$y^{(n)}(t) = \left(\frac{2}{T} \cdot \frac{1 - q^{-1}}{1 + q^{-1}} \right)^n y(t) \quad (2.3)$$

Inserting these discrete forms for continuous derivatives gives the equivalent difference equation for the system.

The second approach for converting continuous-time systems to discrete-time systems is based on covariance equivalence. The discrete-time system is determined by requiring that the output covariance function coincides at all the sampled points with that of the continuous system (Bartlett, 1946). For a simple damped oscillator with zero-mean white-noise excitation, for example, the continuous equation of motion is

$$\frac{d^2 y(t)}{dt^2} + 2\xi_0 \omega_0 \frac{dy(t)}{dt} + \omega_0^2 y(t) = \frac{1}{m} e(t) \quad (2.4)$$

where $e(t)$ is the white-noise input, $y(t)$ is the response, and m , ξ_0 , and ω_0 denote the mass, damping ratio, and the natural frequency, respectively, of the oscillator. The corresponding discrete system is given by the following equation

$$y(t) + \alpha_1 y(t-1) + \alpha_2 y(t-2) = \beta_1 x(t-1) + \beta_2 x(t-2) \quad (2.5)$$

The coefficients α_1 , α_2 , β_1 , and β_2 can be calculated in terms of m , ξ_0 , ω_0 , and T from the equivalence of the continuous and discrete output covariance functions. They are given by the equations (Gersch and Luo, 1972)

$$\alpha_1 = -2 \cos \omega_0 T \sqrt{1 - \xi_0^2} \exp(-\xi_0 \omega_0 T) \quad (2.6)$$

$$\alpha_2 = \exp(-2\xi_0 \omega_0 T) \quad (2.7)$$

$$\beta_1 = \frac{1}{2} (\sqrt{\delta_0 + 2\delta_1} + \sqrt{\delta_0 - 2\delta_1}) \quad (2.8)$$

$$\beta_2 = \frac{1}{2} (\sqrt{\delta_0 + 2\delta_1} - \sqrt{\delta_0 - 2\delta_1}) \quad (2.9)$$

where δ_1 and δ_0 are given as

$$\delta_1 = R_y(1) + \alpha_1 R_y(0) + \alpha_2 R_y(1) \quad (2.10)$$

$$\delta_2 = R_y(0) + \alpha_1 R_y(1) + \alpha_2 R_y(2) + \alpha_1 \delta_1 \quad (2.11)$$

$R_y(k)$ denotes the autocorrelation of $y(t)$ for lag k , calculated by the equation

$$R_y(k) = E[y(t)y(t-k)] = \frac{1}{N} \sum_{t=k}^N y(t)y(t-k) \quad (2.12)$$

where N is the number of sampling points.

Two other approaches for discretization of continuous systems are the pole-zero mapping and the hold equivalence. They both aim to match the continuous transfer function by a discrete equivalent. More on these two techniques can be found in Franklin and Powell (1980).

Regardless of the approach used for discretization, the discrete-time equivalent of a continuous SISO linear system can be represented by a linear difference equation of the following form

$$y(t) + a_1 y(t-1) + \dots + a_{n_a} y(t-n_a) = b_0 x(t) + b_1 x(t-1) + \dots + b_{n_b} x(t-n_b) \quad (2.13)$$

where $x(t)$ and $y(t)$ are the discrete input and output sequences, respectively, and a_j and b_j are called the parameters of the system.

The most important element of discrete-time representation is the sampling interval, T . The sampling interval determines the highest frequency, the so-called *Nyquist frequency*, that the discrete signal can contain. Nyquist frequency is given in hertz as $f_N = 1/2T$. No frequency information beyond f_N can be extracted from a signal sampled with time interval T . A continuous signal $f(t)$ with frequency content between $(-f_c, f_c)$ can be completely reconstructed from its sampled values by the equation

$$f(t) = \sum_{s=-\infty}^{\infty} f(sT) \frac{\sin 2\pi f_N(t-sT)}{2\pi f_N(t-sT)} \quad (2.14)$$

provided that $f_N > f_c$. This is known as *Shannon's sampling theorem* (Shannon, 1949). The techniques used in practice for reconstructing continuous signals from their sampled forms are much simpler. The most widely used one is the zero-order hold, where the signal amplitude is assumed constant (i.e., equal to the value at the first sampling point) between two sampling points. That is

$$f(t) = f(sT), \quad \text{for } sT \leq t < (s+1)T \quad (2.15)$$

The largest error, ϵ_0 , made by using the zero-order hold is

$$\epsilon_0 = \max_s |f(s+1) - f(s)| \leq T \max_t |f'(t)| \quad (2.16)$$

where $f'(t)$ is the derivative of $f(t)$. An improved version is the first-order hold, where the signal amplitude is assumed linear between two sampling points. First-order hold reconstruction is given by the equation

$$f(t) = f(sT) + \frac{t-sT}{T}[f(sT) - f(sT-T)], \quad \text{for } sT \leq t < (s+1)T \quad (2.17)$$

The largest reconstruction error, ϵ_1 , for the first-order hold is

$$\epsilon_1 = \max_s \max_t \left| f(t) - f(sT) - \frac{t-sT}{T}[f(sT) - f(sT-T)] \right| \leq T^2 \max_t |f''(t)| \quad (2.18)$$

3. DISCRETE-TIME MODELS FOR SISO SYSTEMS

The general form for discrete-time representation of a SISO system is given by Eq. 2.13. Equation 2.13 can be written in a more compact form by introducing the following polynomials in the backward-shift operator

$$A(q) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \quad (3.1)$$

$$B(q) = b_0 + b_1 q^{-1} + \dots + b_{n_b} q^{-n_b} \quad (3.2)$$

Equation 2.13 then becomes

$$y(t) = \frac{B(q)}{A(q)} x(t) \quad (3.3)$$

The coefficients a_j and b_j of the polynomials $A(q)$ and $B(q)$ can be constant (time-invariant systems), or functions of t (time-varying systems). The polynomial ratio $B(q)/A(q)$ is called the system transfer operator (the term operator is used since q is not a variable, but an operator). By actually dividing $B(q)$ by $A(q)$ an infinite power series $H(q)$ in the variable q^{-1} is obtained. Equation 3.3 then becomes

$$y(k) = H(q)x(k) \quad (3.4)$$

In terms of the more familiar impulse response functions, the input-output relationship can also be written as

$$y(t) = \sum_{k=1}^{\infty} h(k)x(t-k) \quad (3.5)$$

where $h(k)$ is the discrete impulse response function of the system. By simple manipulation, Eq. 3.5 becomes

$$y(t) = \sum_{k=1}^{\infty} h(k)[q^{-k}x(t)] = \left[\sum_{k=1}^{\infty} h(k)q^{-k} \right] x(t) \quad (3.6)$$

From a comparison of Eq. 3.6 with Eq. 3.4, we can write

$$H(q) = \sum_{k=1}^{\infty} h(k)q^{-k} \quad (3.7)$$

In terms of filtering, the transfer operator $H(q)$ (or $B(q)/A(q)$) represents a recursive linear filter, which converts the input signal $x(t)$ into the output signal $y(t)$.

System equations can also be expressed in the frequency domain by taking the Z -transforms of the time-domain Eqs. 2.13, 3.3, and 3.4. The Z -transform is the discrete equivalent of the continuous Laplace transform. The Z -transform of a discrete sequence $f(kT)$ is defined by the equation

$$Z[f(kT)] = \sum_{k=0}^{\infty} f(kT)z^{-k} \quad (3.8)$$

where $Z[\]$ denotes the Z transform, and z is any complex number. The theory of Z -transforms can be found in texts on discrete systems (e.g., Cadzow, 1973). By taking the Z -transform in Eqs. 3.3 and 3.4, we obtain the following frequency domain equation for the system

$$Z[y(t)] = \frac{B(z)}{A(z)} Z[x(t)] = H(z)Z[x(t)] \quad (3.9)$$

Polynomials $A(z)$ and $B(z)$ are the same as defined by Eqs. 3.1 and 3.2, respectively, with the q 's replaced by z 's. Because z is a variable rather than an operator, $H(z)$ is now called the transfer function. The roots of the numerator polynomial

$$b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b} = 0 \quad (3.10)$$

are called the zeros of the transfer function, whereas the roots of the denominator polynomial

$$1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{n_a} z^{-n_a} = 0 \quad (3.11)$$

are called the poles of the transfer function. The use of terms poles and zeros come from the observation that if $H(z)$ is plotted in three dimensions such that the horizontal axes are the real and complex parts of z and the vertical axis is $H(z)$, the resulting shape resembles a tent. The poles are where the tent is supported, and the zeros are where the tent is tied to the ground.

The transfer function can be represented in terms of more familiar harmonic functions by simply selecting $z = e^{i2\pi fT}$, where $i = \sqrt{-1}$, f denotes the cyclic frequency, and T is the sampling interval. $H(e^{i2\pi fT})$ is known as the frequency response function of the system. The physical meaning of $H(e^{i2\pi fT})$ is that the output $y(t)$ is obtained by multiplying the amplitude of each frequency component of the input $x(t)$ by $|H(e^{i2\pi fT})|$, and shifting its phase by $\arg H(e^{i2\pi fT})$.

In order to have a stable system, it is required that any bounded input gives a bounded output. This corresponds to the condition for the impulse response function $h(k)$ that

$$\sum_{k=1}^{\infty} |h(k)| < \infty, \quad \text{or} \quad \lim_{k \rightarrow \infty} h(k) = 0 \quad (3.12)$$

For the transfer function, the stability means that the poles should all be in complex-conjugate pairs with modulus less than one (i.e., located inside the unit circle in the complex plane).

For $n_b \leq n_a$ (Eqs. 3.1 and 3.2), the transfer function $H(z) = B(z)/A(z)$ can be put into the following form by using a partial fraction expansion (if $n_b > n_a$, first a polynomial division, then a partial fraction expansion should be made):

$$H(z) = \sum_{j=1}^{n_a} \frac{q_j}{1 - p_j z^{-1}} \quad (3.13)$$

where p_j is the j th complex root of the polynomial $A(z)$, and q_j is the corresponding residue of $H(z)$. The residue q_j can be calculated from the equation (Tretter, 1976)

$$q_j = \lim_{z \rightarrow p_j} (1 - p_j z^{-1}) H(z) = \frac{b_1 p_j^{-1} + \dots + b_{n_b} p_j^{-n_b}}{\prod_{\substack{k=1 \\ k \neq j}}^{n_a} (1 - p_k p_j^{-1})} \quad (3.14)$$

If the pairs of terms corresponding to pairs of complex-conjugate roots are combined, then

$$H(z) = \sum_{j=1}^{n_a/2} H_j(z) \quad (3.15)$$

with

$$H_j(z) = \frac{2\Re(q_j) - 2\Re(q_j \bar{p}_j) z^{-1}}{1 - 2\Re(p_j) z^{-1} + |p_j|^2 z^{-2}} \quad (3.16)$$

where $\bar{\cdot}$ and \Re denote the complex-conjugate and the real part, respectively. Each $H_j(z)$ is equivalent to a simple-damped oscillator. The form given by Eqs. 3.15 and 3.16 for $H(z)$ is known as the parallel form realization, where the filter output $y(t)$ is modeled as the linear combination of the outputs of second-order filters each subjected to input $x(t)$. A schematic of parallel form realization is given in Fig. 3.1. From the comparison of Eq. 3.16 with Eqs. 2.6 and 2.7, we can write

$$\alpha_{1j} = -2\Re(p_j) = -2 \cos 2\pi f_{0j} T \sqrt{1 - \xi_{0j}^2} \exp(-\xi_{0j} 2\pi f_{0j} T) \quad (3.17)$$

$$\alpha_{2j} = |p_j|^2 = \exp(-2\xi_{0j} 2\pi f_{0j} T) \quad (3.18)$$

where the frequency f_{0j} is in Hz. Solving for ξ_{0j} and f_{0j} , we obtain the frequency and damping of the corresponding oscillator in terms of pole locations as

$$\xi_{0j} = \frac{\ln(1/r_j)}{[\phi_j^2 + \ln^2(1/r_j)]^{1/2}} \quad (3.19)$$

$$f_{0j} = \frac{\ln(1/r_j)}{2\pi \xi_{0j} T} \quad (3.20)$$

where r_j and ϕ_j are the modulus and the arguments of the j th pole calculated from the equations

$$r_j = p_j \bar{p}_j \quad \text{and} \quad \phi_j = \tan^{-1} \left[\frac{\Im(p_j)}{\Re(p_j)} \right] \quad (3.21)$$

with $\Re(p_j)$ and $\Im(p_j)$ denoting the real and imaginary parts of p_j . Each $H_j(z)$ can be considered as a mode of the system. The numerator of $H_j(z)$ gives the weighting factor (i.e., effective participation factor; Beck, 1978) for that mode.

The above interpretation of the transfer function is based on the parallel form realization. Two other forms widely used in signal analysis are the ladder (cascaded) form and the state space form. They will not be given here, but can be found elsewhere (e.g., Tretter, 1976). A rigorous analysis of the relationship between time-series and linear systems is given in three sequential papers by Willems (1986*a*, 1986*b*, 1987).

4. MODELS FOR UNKNOWN AND NOISY SYSTEMS

As mentioned earlier, the recordings from dynamic systems are always contaminated by noise due to ambient noise existing in the recording environment, as well as the imperfections in the recording instrument. In a very general case, the dynamics that generate the noise is different from the dynamics of the actual system. Also, not only the output but also the input is contaminated with noise. Since the noise in the input goes through the system dynamics along with the actual input, the final noise in the output would include input noise modified by the system dynamics, as well as noise independent of the system dynamics. Since the system is linear, we will consider two linear filters to represent the system dynamics and the noise dynamics. We will also assume that the unknown noise is a stationary random process, which is the case in most real-life problems, so that according to the Wold decomposition theorem it can be written as the output of a linear, time-invariant system driven by a white-noise process (Wold, 1938).

Let $B(q)/[A(q)F(q)]$, and $C(q)/[A(q)D(q)]$ denote two linear filters representing the system dynamics and noise dynamics, respectively, with $A(q)$ and $B(q)$ defined as in Eqs. 3.1 and 3.2, and

$$C(q) = 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c} \quad (4.1)$$

$$D(q) = 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d} \quad (4.2)$$

$$F(q) = 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f} \quad (4.3)$$

We can write the following equations for the signal sequence $y_s(t)$ and the noise sequence $y_n(t)$:

$$y_s(t) = \frac{B(q)}{A(q)F(q)} x(t-k) \quad (4.4)$$

$$y_n(t) = \frac{C(q)}{A(q)D(q)} e(t) \quad (4.5)$$

where $e(t)$ denotes a white-noise sequence, and k is the time delay between input and output. The factor $1/A(q)$ in Eqs. 4.4 and 4.5, common to both filters, accounts for the effect of system dynamics on the output noise. The recorded output is the sum of signal and noise sequences. Therefore, by combining Eqs. 4.4 and 4.5, we obtain the input-output relationship (i.e., system equation) for a SISO linear system; that is

$$A(q)y(t) = q^{-k} \frac{B(q)}{F(q)} x(t) + \frac{C(q)}{D(q)} e(t) \quad (4.6)$$

Equation 4.6 represents a general family of model structures for SISO linear systems, and is known as the *black-box* model (Ljung and Söderström, 1983). A schematic of the black-box model is given in Fig. 4.1. In the majority of applications, not all the polynomials in Eq. 4.6 are needed (i.e., some of the polynomials can be taken equal to one). By eliminating various polynomials a number of special forms of the black-box model are obtained. Some of these forms are known by special names in the literature. They are summarized in Table 4.1.

5. SYSTEM IDENTIFICATION

System identification constitutes determining the coefficients of the polynomials in the black-box model, Eq. 4.6, for a given pair of input and output sequences. The steps for identification are outlined in the following subsections.

5.1. ONE-STEP-AHEAD PREDICTION

One-step-ahead prediction involves predicting the output at the next time step by using the information available at the present and past time steps. The main problem in doing this is the unknown noise term. To handle the noise, let us first rewrite Eq. 4.6, such that the white-noise term $e(t)$ is isolated. With some algebraic manipulations, we can write

$$y(t) = \left[1 - \frac{D(q)A(q)}{C(q)} \right] y(t) + q^{-k} \frac{D(q)B(q)}{C(q)F(q)} x(t) + e(t) \quad (5.1)$$

It can be shown by inserting the expressions for $A(q)$, $C(q)$, and $D(q)$ (from Eqs. 3.1, 4.1, and 4.2) inside the bracket on the right-hand side of Eq. 5.1, that the coefficient of $y(t)$ is a polynomial with order q^{-1} with no constant term. Therefore, the right-hand side of Eq. 5.1 includes $y(t)$ terms only up to time $t - 1$. Assume that all the input, output, and the polynomial coefficients are known up to time $t - 1$. Then, all the terms on right-hand-side of Eq. 5.1 are known, except the noise term $e(t)$. The best estimate we can make for $e(t)$ would be to use its expected value, i.e., zero. Therefore, by taking the expectation in Eq.

5.1, and also noting that $E[e(t)] = 0$, we can write for the best estimate of $y(t)$ in terms of past values of the input, output, and parameters, as

$$\hat{y}(t, \theta) = \left[1 - \frac{D(q)A(q)}{C(q)}\right]y(t) + q^{-k} \frac{D(q)B(q)}{C(q)F(q)}x(t) \quad (5.2)$$

$\hat{y}(t, \theta)$ is the expected value of $y(t)$ at time t for given θ , and θ is the vector composed of the model parameters (i.e., coefficients of the polynomials); that is

$$\theta = (a_1, \dots, a_{n_a}, b_1, \dots, b_{n_b}, c_1, \dots, c_{n_c}, d_1, \dots, d_{n_d}, f_1, \dots, f_{n_f})^T \quad (5.3)$$

where superscript T denotes the transpose. $\hat{y}(t, \theta)$ is known as the *one-step-ahead prediction* of $y(t)$. θ is included in the argument list of $\hat{y}(t, \theta)$ to emphasize that the predicted value depends on the past values of the model parameters.

The difference

$$\varepsilon(t, \theta) = y(t) - \hat{y}(t, \theta) \quad (5.4)$$

gives the error in the estimation at time t . Note by comparing Eqs. 5.1 and 5.2 that, if θ is estimated perfectly, then the error would be equal to a white-noise sequence, i.e., $\varepsilon(t, \theta) = e(t)$. Therefore, one of the criteria for accurate identification is to have the error sequence $\varepsilon(t, \theta)$ as close to a white-noise sequence as possible.

5.2. WEIGHTED LEAST-SQUARES METHOD

The goal of any identification algorithm is to minimize the total estimation error. First, however, we have to decide on how to measure the total estimation error. The most popular and algebraically the most convenient one is to use the least-squares approximation, which was first introduced by Gauss to calculate the six coefficients that determine the elliptical orbit of a planetary body (Gauss, 1809). An historical review and evolution of the least-squares method is given by Sorensen (1970).

The least-squares approximation uses quadratic criteria for measuring errors. We will also include a weighting factor in the criterion, so that we can have the flexibility of manipulating the effect of data on total error. With these, we will define the total error up to time t as

$$V(t, \theta) = \frac{1}{2} \gamma(t) \sum_{s=1}^t \beta(t, s) \varepsilon^2(s, \theta) \quad (5.5)$$

where $\beta(t, s)$ is the weighting factor, and $\gamma(t)$ is the normalization factor for $\beta(t, s)$, defined by

$$\gamma(t) = \frac{1}{\sum_{s=1}^t \beta(t, s)} \quad \text{or,} \quad \gamma(t) \sum_{s=1}^t \beta(t, s) = 1 \quad (5.6)$$

For time-invariant systems, weighting factors can all be taken equal to one. However, when the initial conditions are unknown, it is advantageous to use weighting factors that gradually decrease to much smaller values towards the beginning of data, so that the effect of unknown initial conditions on the total error is minimized. For time-varying systems, weighting factors are essential to track the time variation of system parameters. The weighting factors localize the identification by giving more weights to the current values, and by gradually discounting the past values.

A moving rectangular window, or exponential window can be used as the weighting factor for time-varying systems. For a rectangular window

$$\beta(t, s) = \begin{cases} 0, & \text{for } s < t - s_w \\ 1, & \text{for } s \geq t - s_w \end{cases} \quad (5.7)$$

where s_w is the length of the rectangular window in terms of number of data points (the actual window length is $s_w T$). s_w should be selected such that $s_w T$ is not smaller than the largest significant period of the system.

For an exponential window, first assume that $\beta(t, s)$ has the following recursive form

$$\beta(t, s) = \lambda(t) \beta(t-1, s) \quad \text{with} \quad 1 \leq s \leq t-1 \quad (5.8)$$

where

$$\lambda(t) \leq 1, \quad \text{and} \quad \beta(t, t) = 1 \quad (5.9)$$

This recursive form leads to the equations

$$\beta(t, s) = \prod_{j=s+1}^t \lambda(j) \quad \text{and} \quad \gamma(t) = \frac{\gamma(t-1)}{\lambda(t) + \gamma(t-1)} \quad (5.10)$$

If $\lambda(t)$ is chosen constant, i.e., $\lambda(t) = \lambda$, the weighting factors become exponential, that is

$$\beta(t, s) = \lambda^{t-s} \quad (5.11)$$

where the constant λ is called the *forgetting factor* (Ljung and Söderström, 1983). Also note from Eq. 5.10 that if $\lambda(t) = 1$ and $\gamma(0) = 1$, then $\gamma(t) = 1/t$.

Figure 5.1 shows three windows, a rectangular window with $s_w = 10$, and two exponential windows with $\lambda = 0.9$, and 0.7 . The smaller the λ value the sharper is the decay of the weighting factor. For constant $\lambda(t)$ values the weighting factors are moving windows of the same shape. To reduce the effect of initial conditions and speed up the convergence rates, a time-varying forgetting factor, which initially is small and grows towards the value λ_0 as more data points are processed, is more appropriate. Such time variation can be modeled by the following recursive equation:

$$\lambda(t) = \lambda_0 \lambda(t-1) + 1 - \lambda_0 \quad (5.12)$$

In order to start the recursion, λ_0 and $\lambda(0)$ values need to be specified. Suggested values are $\lambda_0 = 0.99$ and $\lambda(0) = 0.95$ (Ljung and Söderström, 1983; Young, 1984). However, these values should be used with care, since it is the sampling interval and the rate of change of the frequency content of the signal that actually determine the most appropriate values. As will be shown later by examples, identification results can be very sensitive to λ_0 and $\lambda(0)$ values. It is recommended that either the frequency content and time variations of the signal should be investigated by Fourier analysis techniques, or several combinations of λ_0 and $\lambda(0)$ should be tried, before deciding on the final values for identification.

In order to track fast variations in the system parameters, the window length should be small (i.e., the decay of the weighting factors should be steep). On the other hand, to remove the random noise from the signal, the number of sampled points used at any time (defined by the window length) should be large enough, so that averaging over the sampled points can actually remove the noise effects. Therefore, there is a trade-off between the time-tracking ability and the noise sensitivity of the identification.

Although the use of quadratic criteria makes the least-squares method analytically simple, it also makes the propagation of estimation errors very wide. An isolated error, such as a large instrument noise at one sampling point, can effect the estimation in all the other points. One way to prevent this is to check the estimation error (Eq. 35) at every sampling point for a sudden large jump, and if there is one, calculate the error for that point by using a power less than two, say 1.5.

5.3. MAXIMUM LIKELIHOOD METHOD

Another popular method for error minimization and parameter estimation is a statistical one, the maximum likelihood method, attributed to Fisher (1922). The basic principles of the maximum likelihood method can be summarized as follows. Assume that the output $y(t)$ is an observation (i.e., a sample) from a random process, whose probability density function depends on the unknown parameter vector θ . Let $P[y_N|\theta]$ denote the joint conditional probability density function of $y(t)$ for $t = 1 \cdots N$ (i.e.,

$P[y_N|\theta] = P[y(1), y(2), \dots, y(N)|\theta]$. For given θ , $P[y_N|\theta]$ shows the probability that the output will take the measured values $y(1), y(2), \dots, y(N)$. Once the observed values of $y(t)$ are inserted, $P[y_N|\theta]$ becomes a deterministic function of θ . This function is called the *likelihood function*. θ is then selected as the value which makes the likelihood function maximum (i.e., the observed output becomes as likely as possible). By using Bayes' rule, we can write

$$P[y_N|\theta] = P[y_N|y_{N-1}, \theta] P[y_{N-1}|y_{N-2}, \theta] \cdots P[y_2|y_1, \theta] P[y_1|\theta] \quad (5.13)$$

Now, assume that the conditional densities are Gaussian, such that

$$P[y_t|y_{t-1}, \theta] = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left[-\frac{(y - m_t)^2}{2\sigma_t^2}\right] \quad (5.14)$$

where m_t and σ_t are the conditional mean and standard deviation of $y(t)$. They both depend on unknown parameters θ , and the past data. Based on the arguments for Eq. 5.2, we can write

$$m_t = \hat{y}(t, \theta); \quad \text{therefore,} \quad y - m_t = y(t) - \hat{y}(t, \theta) = \varepsilon(t, \theta) \quad (5.15)$$

The equation for θ is obtained by maximizing $P[y_N|\theta]$. Instead of maximizing $P[y_N|\theta]$, we can maximize its logarithm. Using Eq. 5.15 in Eq. 5.14 and the resulting equation in Eq. 5.13, we can write for the negative logarithm of $P[y_N|\theta]$

$$-\ln P[y_N|\theta] = \frac{1}{2} \sum_{t=1}^N \left[\left(\frac{\varepsilon(t, \theta)}{\sigma_t} \right)^2 + \ln \sigma_t^2 \right] + \frac{N}{2} \ln 2\pi \quad (5.16)$$

Maximization of $P[y_N|\theta]$ is now equivalent to minimization of the right-hand side of Eq. 5.16. Note that if σ_t is constant or independent of θ , the maximum likelihood criterion becomes equal to the quadratic criterion of the least-squares method, given by Eq. 5.5, with unit weighting factor. The advantage of the least-squares method is its simplicity. The advantages of the maximum likelihood method are its independence from the model type, and better convergence properties. For Gaussian observations, however, these two methods are equivalent, and result in the same θ values. The method suggested by Friedlander (1982) combines the least-squares and the maximum likelihood methods. The formulation that will follow is based on the weighted least-squares method.

5.4. STOCHASTIC APPROXIMATION

For the least-squares method, to minimize the total error we require in Eq. 5.5 that

$$\frac{d}{d\theta} V(t, \theta) = V'(t, \theta) = 0 \quad (5.17)$$

where the superscript ' is used to denote derivative of V with respect to θ . Note in Eq. 5.1 that each $y(t)$ includes the unknown white-noise term $\epsilon(t)$. Therefore, $\epsilon(t, \theta)$ (Eq. 5.4), $V(t, \theta)$ (Eq. 5.5), and $V'(t, \theta)$ are all functions of white-noise terms $e(0), e(1), \dots, e(t)$. Since the noise is random, any new set of measurements would result in a different $V'(t, \theta)$ function, and consequently, different θ values. Thus, Eq. 5.17 should be considered by using the expected value of $V'(t, \theta)$ over the set of measurements, that is

$$E V'(t, \theta) = 0 \quad (5.18)$$

Solution of this equation can be accomplished by using *stochastic approximation techniques*. The general theory of stochastic approximation can be found in Albert and Gardner (1967), Nevelson and Khasminskii (1973), and Kushner and Clark (1978). A brief outline of the concept will be given below by summarizing the Robbins-Monro algorithm (Robbins and Monro, 1951), which is the prototype stochastic approximation scheme.

Let $Q(\eta, \nu)$ denote a function of unknown parameters η and the random measurement vector ν . The problem we want to solve is to determine the values of η which satisfy the equation

$$E Q(\eta, \nu) = 0 \quad (5.19)$$

Assume that the probabilistic distribution of ν is unknown (the exact form of $Q(\eta, \nu)$ may also be unknown). However, since measured values of ν are available, we can determine $Q(\eta, \nu)$ for any chosen η . Thus, if we have a large set of ν values, we can choose a value for η , calculate $Q(\eta, \nu)$ for each set of ν , take the average, and check whether it is zero; and, if not, continue the procedure by choosing new η values until it becomes zero. However, this is a very tedious and time consuming procedure. Moreover, in most cases we do not have a large set of ν . Robbins and Monro (1951) introduced a more efficient algorithm to solve the problem. They showed that η can be solved recursively by the following algorithm:

$$\eta_{j+1} = \eta_j + \alpha_j Q(\eta_j, \nu) \quad (5.20)$$

where subscript j denotes the iteration number, and α_j is a sequence of positive constants satisfying the conditions (Young, 1984)

$$\alpha_j > 0; \quad \sum_{j=1}^{\infty} \alpha_j = \infty; \quad \text{and} \quad \sum_{j=1}^{\infty} \alpha_j^2 < \infty \quad (5.21)$$

The harmonic sequence $1/j$ (i.e., $1, 1/2, 1/3, \dots$) is the best known example of the sequences that satisfy these conditions. For time-varying systems, the last condition in Eq. 5.21 is usually violated in order to track the time variations of the parameters; however, the algorithm still works.

Going back to our original problem, Eq. 5.18, we can write by analogy to Eq. 5.20 that

$$\theta_j(t) = \theta_{j-1}(t) + \alpha_j V'(t, \theta_{j-1}) \quad (5.22)$$

This equation is basically the stochastic equivalent of the gradient method that is widely used for minimization of deterministic functions. As is well known from deterministic theory, the gradient method becomes very slow when the iteration approaches the minimum (Luenberger, 1973). An improved version of the gradient method is obtained by modifying the search direction of the iteration by the second derivative (i.e., Hessian), of the function. This results in Newton's method. There are also methods that are the combination of gradient and Newton's methods, such as the Marquardt algorithm (Marquardt, 1963). Modification of Eq. 5.22 according to Newton's method gives

$$\theta_j(t) = \theta_{j-1}(t) + \alpha_j [V''(t, \theta_{j-1})]^{-1} V'(t, \theta_{j-1}) \quad (5.23)$$

If V is quadratic in θ this iteration would converge to minimum V in one step (Ljung and Soderstrom, 1983). To speed up the calculations, the usual practice is to add one more data point at each iteration. When this is done Eq. 5.23 becomes

$$\theta(t) = \theta(t-1) + \alpha_t [V''[t, \theta(t-1)]]^{-1} V'[t, \theta(t-1)] \quad (5.24)$$

5.5. RECURSIVE PREDICTION ERROR METHOD (RPEM)

The iterative algorithm given by Eq. 5.24 with Eqs. 5.2, 5.4 and 5.5 for system identification is known as the *Recursive Prediction Error Method*, RPEM for short. For ease of notation, denote that

$$\psi(t) = -\frac{d\varepsilon(t, \theta)}{d\theta} = \frac{d\hat{y}(t, \theta)}{d\theta} \quad \text{and} \quad R(t) = V''(t, \theta) \quad (5.25)$$

$\psi(t)$ is the gradient and $R(t)$ is the Hessian of the prediction. Note that $\psi(t)$ is a vector with dimension $d_\theta = n_a + n_b + n_c + n_d + n_f$, whereas $R(t)$ is a matrix with dimension $d_\theta \times d_\theta$. $V'(t, \theta)$ is calculated from Eq. 5.5 by taking the derivative with respect to θ , that is

$$V'(t, \theta) = -\gamma(t) \sum_{s=1}^t \beta(t, s) \psi(s) \varepsilon(s, \theta) \quad (5.26)$$

It can be shown by using Eq. 5.10 that Eq. 5.26 can be put into following recursive form:

$$V'(t, \theta) = \gamma(t) \left[\frac{\lambda(t)}{\gamma(t-1)} V'(t-1, \theta) - \psi(t) \varepsilon(t, \theta) \right] \quad (5.27)$$

or

$$V'(t, \theta) = V'(t-1, \theta) + \gamma(t) [-\psi(t) \varepsilon(t, \theta) - V'(t-1, \theta)] \quad (5.28)$$

If we assume that V' was actually minimized in the previous time step, we can then write $V'(t-1, \theta) = 0$. Therefore, Eq. 5.28 becomes

$$V'(t, \theta) = -\gamma(t) \psi(t) \varepsilon(t, \theta) \quad (5.29)$$

The components of $\psi(t)$ can be calculated by using Eq. 5.2 in Eq. 5.25. They are given by the following recursive equations:

$$\frac{\partial}{\partial a_j} \hat{y}(t, \theta) = -\frac{D(q^{-1})}{C(q^{-1})} y(t-j) \quad (5.30)$$

$$\frac{\partial}{\partial b_j} \hat{y}(t, \theta) = \frac{D(q^{-1})}{C(q^{-1})F(q^{-1})} y(t-j) \quad (5.31)$$

$$\frac{\partial}{\partial c_j} \hat{y}(t, \theta) = \frac{1}{C(q^{-1})} \varepsilon(t-j, \theta) \quad (5.32)$$

$$\frac{\partial}{\partial d_j} \hat{y}(t, \theta) = -\frac{1}{C(q^{-1})} \left[A(q^{-1}) y(t-j) - \frac{B(q^{-1})}{F(q^{-1})} x(t-j) \right] \quad (5.33)$$

$$\frac{\partial}{\partial f_j} \hat{y}(t, \theta) = -\frac{D(q^{-1})}{C(q^{-1})F(q^{-1})} \cdot \frac{B(q^{-1})}{F(q^{-1})} x(t-j) \quad (5.34)$$

The Hessian $R(t)$ is calculated by taking the second derivative of $V(t, \theta)$ with respect to θ . Thus, by taking one more derivative in Eq. 5.26 we obtain

$$R(t) = V''(t, \theta) = \gamma(t) \sum_{s=1}^t \beta(t, s) \left[\psi(t) \psi^T(t) + \psi'(t) \varepsilon(t, \theta) \right] \quad (5.35)$$

It can be shown that close to minimum $V(t, \theta)$, the term $\psi'(t) \varepsilon(t, \theta)$ becomes very small, such that it can be neglected (Ljung, 1987). Thus we can approximate

$$R(t) \approx \gamma(t) \sum_{s=1}^t \beta(t, s) \psi(t) \psi^T(t) \quad (5.36)$$

$R(t)$ can also be put in a recursive form:

$$R(t) = R(t-1) + \gamma(t)[\psi(t)\psi^T(t) - R(t-1)] \quad (5.37)$$

Using Eqs. 5.29 and 5.35 in Eq. 5.24, and also taking $\alpha_t = 1$ give the following final form for the RPEM algorithm:

$$\theta(t) = \theta(t-1) + \gamma(t)R^{-1}(t)\psi(t)\varepsilon(t) \quad (5.38)$$

The matrix inversion, $R^{-1}(t)$, in Eq. 54 makes it computationally inefficient. There are various algorithms developed to calculate $R^{-1}(t)$ without actually inverting the matrix. One is to use the matrix inversion lemma, by which the matrix inversion can be put into a form of recursive multiplications (Householder, 1964). A straightforward application of the matrix inversion lemma, however, leads to equations that are susceptible to accumulation of round-off errors. Numerically more sound forms are obtained by using the so-called factorization techniques, such as the square root algorithm (Potter, 1963), or the U - D factorization algorithm (Bierman, 1977). Application of these algorithms to obtain $R^{-1}(t)$ is given in Ljung and Söderström (1983).

In order to start the recursion in Eq. 5.38, we need the initial values for the vectors $\theta(t)$, $\psi(t)$, and the matrix $R^{-1}(t)$. For stable systems, it can be assumed that

$$\theta(0) = 0, \quad \psi(0) = 0, \quad \text{and} \quad \gamma(0)R^{-1}(0) = K \cdot I \quad (5.39)$$

where K is a large constant and I is the identity matrix. A suggested value for K is (Franklin and Powell, 1980)

$$K = \frac{10}{N+1} \sum_{t=0}^N y^2(t) \quad (5.40)$$

If there is a large amplitude difference between the input and output, either two K values proportional to amplitudes should be used (so that the elements of R^{-1} will have the same magnitude during the recursion), or the input and output should be scaled to have similar magnitudes prior to recursion.

Application of the RPEM algorithm to an input-output set can be summarized by the following steps:

1. Read and preprocess the input-output series (as will be discussed in section 12).
2. Select λ_0 and $\lambda(0)$ for the weighting factor (Eq. 5.12).
3. Select initial values $\theta(0)$, $\psi(0)$, and $\gamma(0)R^{-1}(0)$ (Eq. 5.39); set $V(0) = 0$ (Eq. 5.5).
4. Calculate for the next time step $\hat{y}(t, \theta)$ (Eq. 5.2), $\varepsilon(t, \theta)$ (Eq. 5.4), $V(t, \theta)$ (Eq. 5.5), $\psi(t)$ (Eqs. 5.30-5.34), $\gamma(t)$ (Eq. 5.10), and $R(t)$ (Eq. 5.37).
5. Calculate $\theta(t)$ (Eq. 5.38).
6. Move to next time step, and repeat steps 4 and 5.

For strongly time-varying systems, step 5 may require iteration over θ as shown by Eq. 5.23. To select the best model order and the weighting factor, the algorithm is repeated for different model orders and weighting factors, and the variation of total estimation error $V(N, \theta)$ (Eq. 5.5) is observed.

The RPEM algorithm presented above is for the general black-box model structure (Eq. 4.6), where all the polynomials exist. For the special model structures given in Table 4.1, the equations would be simplified significantly. Although the general procedure for the black-box models would also apply to any special model, it is possible to develop more efficient, and model-specific algorithms for special model structures. To give an example, consider the ARX model given by the equation

$$A(q)y(t) = B(q)x(t) + e(t) \quad (5.41)$$

It can be shown for the ARX model that the one-step-ahead prediction of the output (Eq. 5.2) is

$$\hat{y}(t, \theta) = \phi^T(t)\theta \quad (5.42)$$

where

$$\phi(t) = [-y(t-1), \dots, y(t-n_a), x(t-1), \dots, x(t-n_b)]^T \quad (5.43)$$

Equation 5.42 is simply a linear regression equation, with $\phi(t)$ being the regression vector. θ can be calculated by taking $\hat{y}(t, \theta) = y(t)$ and using standard regression analysis. If the system is time-invariant, and all the data is available beforehand, then the ARX model can also be identified in off-line fashion using correlation methods. If N denotes the number of data points, the total error can be written as

$$V(N, \theta) = \sum_{t=1}^N \beta(N, t) [y(t) - \phi^T(t)\theta]^2 \quad (5.44)$$

Minimization of $V(N, \theta)$ with respect to θ gives

$$\theta = \frac{\sum_{t=1}^N \beta(N, t) \phi(t) y(t)}{\sum_{t=1}^N \beta(N, t) \phi(t) \phi^T(t)} \quad (5.45)$$

6. CONVERGENCE AND CONSISTENCY OF ESTIMATION

As for any iteration method, the convergence and consistency of the above algorithm should be investigated before it is applied to real-life problems. This involves showing that the algorithm converges to a global minimum not a local minimum, and that the estimated parameters are unique. These problems have been investigated extensively (Åström and Bohlin, 1965; Hannan, 1973; Ljung, 1976, 1978, 1981; Anderson et.al., 1978; Anderson and Taylor, 1979; Dugard et. al., 1980; Solo, 1981; Fogel, 1981; Tyskin et.al., 1981; Lai and Wei, 1982; Rootzen and Sternby, 1984; and Stoica and Nehorai, 1987). Basic results are summarized below.

For convergence and consistency, first the system and the model should satisfy the following three conditions: (a) the system is stable, i.e., bounded input gives bounded output (system does not have negative damping), (b) the selected model includes the true system, meaning that the number of parameters used in the model is greater or equal to that of the true system (although we do not know much about the true system, we can assume that this condition is eventually satisfied by increasing the number of parameters), and (c) the input is *persistently exciting*, which means that the input should have a non-zero spectral amplitudes at frequencies corresponding to frequencies of the system (generally satisfied because of the existence of noise).

For the black-box, ARARMAX, and ARMAX models (see Table 4.1), the convergence to a local minimum is theoretically possible. Various model validity checks that will be discussed later will show clearly whether a local or global minimum is reached. If a local minimum is encountered, the calculations should be restarted by using different initial values for θ . The process is repeated until the global minimum is obtained. In practical applications, however, local minima are not encountered frequently. In none of the applications that will be presented later has a local minimum been encountered. For some special forms of the black-box model, it is possible to prove analytically that no local minima exist. They can be summarized as follows:

- a. For FIR models there are no local minima (since the criterion function is quadratic in θ).
- b. For ARX models there are no local minima (since the criterion function is quadratic in θ).
- c. For ARMA models there are no local minima (Åström and Söderström, 1974).
- d. For ARARX models there are no local minima if the signal-to-noise ratio is large (Söderström, 1974).

- e. For Box-Jenkins models there are no local minima only if $F = 1$ (Söderström, 1975a).
- f. For Output-Error models there are no local minima if the input is white noise (Söderström, 1975a).

The uniqueness of the identification of a black-box model is satisfied if the polynomials for the true system (which will be distinguished by the superscript * from those of the estimated system) satisfy all of the following conditions (Åström and Bohlin, 1965):

- a. there is no common factor to all three polynomials $z^{n_a} \cdot A^*(z)$, $z^{n_b} \cdot B^*(z)$, and $z^{n_c} \cdot C^*(z)$;
- b. there is no common factor to polynomials $z^{n_b} \cdot B^*(z)$ and $z^{n_b} \cdot F^*(z)$;
- c. there is no common factor to polynomials $z^{n_c} \cdot C^*(z)$ and $z^{n_d} \cdot D^*(z)$; and
- d. if $n_a > 1$, there is no common factor to polynomials $z^{n_d} \cdot D^*(z)$, and $z^{n_f} \cdot F^*(z)$.

7. FISHER INFORMATION MATRIX AND CRAMER-RAO INEQUALITY

No matter what the model structure and the estimation algorithm are, there is a limit to the achievable accuracy of the estimated parameters. This limit is known as the *Cramer-Rao lower bound* (Cramer, 1946; Rao, 1965). Let θ_0 denote the true value of the parameter vector θ , and $\hat{\theta}(y^N)$ its estimate based on N sampling points. Then the Cramer-Rao lower bound is given by the inequality

$$E[\hat{\theta}(y^N) - \theta_0][\hat{\theta}(y^N) - \theta_0]^T \geq M_I^{-1} \quad (7.1)$$

The matrix M_I is known as the *Fisher information matrix* (Fisher, 1922), defined by the equation

$$M_I = -E \left[\frac{d^2}{d\theta^2} \ln P(y_N|\theta) \right]_{\theta=\theta_0} \quad (7.2)$$

where $P(y_N|\theta)$ denotes, as before, the conditional joint probability density of $y(t)$, $t = 1, \dots, N$, for given θ . Since θ_0 is unknown, M_I cannot be evaluated; therefore, the Cramer-Rao lower bound does not have any significance, as far as the validation of estimated θ values is concerned. It is important, however, to know that there is such a lower bound for accuracy. Any unbiased estimator θ (i.e., θ which makes the expected value of residuals zero) that attains the Cramer-Rao lower bound is said to be *efficient*. It can be shown that in the linear Gaussian case, the weighted least-squares and maximum likelihood methods result in efficient estimators (Brogan, 1987).

8. MODEL SELECTION

Model selection involves the selection of the form and the order of the model, and constitutes the most important part of system identification. Any prior information about the

system behavior and noise sources can be useful. The physical laws governing the system, if known, give the dynamic equilibrium equations, which can be a good starting point for model selection. For example, equilibrium equations for the uni-directional vibration of a six-story building suggest that the building would have six distinct frequencies, provided that the input has sufficient frequency content to excite all the modes (i.e., the input is persistently exciting). We therefore conclude that the denominator polynomial for the system transfer function should be at least twelfth order. Another useful tool is the Fourier analysis of the data. The number of dominant peaks in the amplitude spectra of the input and output give a good idea of the model order. It is also important to check whether the data is linear or nonlinear by using various tests available in the literature (e.g., Haber, 1985; Varlaki, et.al., 1985). Young (1978) and Priestly (1980) have suggested that almost any nonlinear system can be considered as a linear system with time-varying parameters. However, from the identification stand point, we prefer to start with a model that is closest to the actual system. In order to be able to describe all the possible forms of system dynamics the model should have a sufficient number of parameters, i.e., the model should be *flexible*. On the other hand, overparametrization causes a singular or ill-conditioned $R(t)$ matrix in Eq. 5.38, and pole-zero cancellations in the transfer function. This requires that the model should contain the smallest number of free parameters to represent the system adequately, i.e., the model should be *parsimonious*. A general recommendation for selecting the model type is to start with the simplest model, and continue with the next simplest model until the model validity tests (which will be given in the next section) are satisfied, or until a pole-zero cancellation occurs (see, Söderstöm, 1975b for tests for pole-zero cancellation).

For selecting the model order, a straightforward approach is to investigate the variation of total estimation error, $V(N, \theta)$ (Eq. 5.5), with model order. Normally, the total estimation error decreases with increasing model order. However, as schematically shown in Fig. 8.1, the decrease is very sharp at the beginning, and gradually flattens as the order increases. The end of the steep decline usually indicates the optimal model order. The beginning of the flat region suggests that any additional parameter is not significantly improving the model for the system, although it may improve the model for noise. The model order can be taken as the number of parameters in the denominator polynomial of the system transfer function.

Another test available for selecting model order is to use *Akaike's information theoretic criterion* (AIC) (Akaike, 1981). AIC seeks a model order that minimizes the *information distance* between the conditional probability density functions of the measurements and the true system. The definition of the information distance is due to Kullback and Leibler (1951). The minimization of information distance corresponds to maximizing the entropy

of the measurement probability density with respect to the true probability density. AIC gives the following equation for the best model order

$$\theta_{AIC} = \min_{\theta} \left[\frac{1}{N} \sum_{t=1}^N -\ln P(t|\theta) + \frac{d_{\theta}}{N} \right] \quad (8.1)$$

where, as previously defined, $P(t, \theta)$ is the conditional (i.e., for given θ) probability density function of the measurements, and d_{θ} denotes the dimension of the vector θ (i.e., number of parameters). For Gaussian prediction errors $\varepsilon(t)$ and the maximum likelihood criterion function, Eq. 8.1 becomes

$$\theta_{AIC} = \min_{\theta} \left\{ \ln \left[\frac{1}{N} \left(1 + \frac{2d_{\theta}}{N} \right) \sum_{t=1}^N \varepsilon^2(t) \right] \right\} \quad (8.2)$$

AIC is based on the maximum likelihood estimation. For the least-squares estimation, the criterion corresponding to AIC is the *Akaike's final prediction-error criterion* (FPE). It is given by the equation

$$\theta_{FPE} = \min_{\theta} \left[\frac{N + d_{\theta}}{N - d_{\theta}} \frac{1}{N} \sum_{t=1}^N \frac{1}{2} \varepsilon^2(t) \right] \quad (8.3)$$

Note that for $N \gg d_{\theta}$ Eqs. 8.2 and 8.3 result in the same criterion. Also note that both equations penalize using too many parameters (i.e., the criterion function increases with increasing d_{θ}).

Various other methods have also been suggested for model order selection (Hsia, 1977; Schwarz, 1978; Hannan and Quinn, 1979; Fine and Hwang, 1979; Inagaki, 1981; Unton, 1981; Rissanen 1983; Fuchs, 1987). The method by Hsia (1977) allows iteration on model order as well as on parameter estimates. A review of existing methods is given by Stoica et. al. (1986).

9. MODEL VALIDATION

The final stage of an identification process is to confirm that the estimated model is a realistic approximation of the actual system. This is known as model validation. There are several tools available for model validation. The first and simplest test is to compare the estimated system transfer function with that obtained from the standard Fourier analysis. A second test may be to compare the output of the estimated system with the actual output. Although a perfect match is not expected, these two tests should give a fairly good match. Any gross mismatch is a clear indication of an incorrect model.

Another set of tests can be made over the residuals $\varepsilon(t)$ (Eq. 5.4), of the model (Anscombe and Tukey, 1963). As Eqs. 5.1 and 5.2 indicate, the whole estimation algorithm

is based on the condition that the difference between the model output and the actual output (i.e., the residual) is a white-noise series. Therefore, the closer $\varepsilon(t)$ is to white-noise the better is the model. A first step would then be to plot $\varepsilon(t)$ and its Fourier amplitude spectrum, and see whether they look similar to those of a white-noise series. More precise checks can be made by using various statistical tests for whiteness. By using the autocovariance of the residuals, for example, we can calculate the parameter ζ_M for a sufficiently long correlation lag M (e.g., $M = 20$), as

$$\zeta_M = \frac{N}{[R_\varepsilon(0)]^2} \sum_{\tau=1}^M [R_\varepsilon(\tau)]^2 \quad (9.1)$$

where $R_\varepsilon(\tau)$, the autocovariance of $\varepsilon(t)$, is

$$R_\varepsilon(\tau) = \frac{1}{N} \sum_{t=1}^{N-\tau} \varepsilon(t)\varepsilon(t+\tau) \quad (9.2)$$

If $\varepsilon(t)$ is indeed a white-noise series, then ζ_M would be asymptotically $\chi^2(M)$ (Chi-square) distributed (Ljung, 1987). To give an example, assume that for a valid identification we require $\varepsilon(t)$ be white-noise with 90-percent probability. We then go to the χ^2 distribution table and find the value $\chi_{90}^2(M)$, corresponding to the 90-percentile and correlation lag M , and require for model validity that $\chi_{90}^2 > \zeta_M$.

In most applications, we are more interested in determining the system transfer function than determining the noise transfer function. This means that it is permissible to have an inaccurate model for the noise. In this case, there is no need to force the residuals to be white noise, since they would not be with inaccurate noise models. Instead, we can investigate the correlation of the residuals with input. This would show us whether there are any components common to both residuals and input. If there are, it means that more information can be extracted from the residuals as regards to the relationship between input and output than what is given by the model. The correlation of residuals with input is characterized by the cross-covariance function, $R_{\varepsilon x}(\tau)$, which is given by

$$R_{\varepsilon x}(\tau) = \frac{1}{N} \sum_{t=\tau}^N \varepsilon(t)x(t-\tau) \quad (9.3)$$

If $\varepsilon(t)$ is independent of $x(t)$ then $\sqrt{N}R_{\varepsilon x}(\tau)$ would be asymptotically normal distributed with zero mean, and the standard deviation σ

$$\sigma = \sum_{\tau=-N}^{\tau=N} R_\varepsilon(\tau)R_x(\tau) \quad (9.4)$$

where $R_\varepsilon(\tau)$ and $R_x(\tau)$ are the autocovariance functions of $\varepsilon(t)$ and $x(t)$, calculated as in Eq. 9.2. If we want the input and residuals to be independent with α -percent probability, we require that (Ljung, 1987)

$$|R_{\varepsilon x}(\tau)| \leq \sqrt{\frac{\sigma}{N}} N_\alpha \quad (9.5)$$

N_α is the α -percentage level of $N(0, 1)$ normal distribution. Sometimes, Eq. 9.5 is satisfied for $\tau > 0$, but not for $\tau < 0$. This does not necessarily mean that the model is wrong. It does mean however that there is a feedback between the output and input (i.e., the system is not an open-loop, but a closed-loop system). This was observed in the analysis of building vibrations with soil-structure interaction effects, which will be presented in section 13.

The AIC and FPE criteria described in the previous section for model selection can also be used to validate a model among several models. More on model validation can be found in Ljung(1987), Bohlin(1987), Söderström (1987), and Brogan (1987).

10. SPECTRAL ESTIMATION

The common approach in modeling the spectral shape of a signal is to use the Fourier amplitude spectrum calculated by using fast Fourier transforms. For a discrete signal of N sampling points, the Fourier amplitude spectrum is given by $N/2$ amplitude-frequency pairs. In other words, description of the Fourier amplitude spectrum requires $N/2$ parameters. Also implicit in this approach are the assumptions that the signal is stationary, and the signal to noise ratio is high (or the signal and noise spectra do not overlap). For time-varying signals, such as ground motion records, the first assumption is not valid. The second assumption is valid only for a certain frequency band. To overcome these, the data is windowed and filtered. The windowing aims to incorporate the nonstationary characteristics of the motion, whereas the filtering removes the segments of the data believed to be dominated by noise. Various smoothing and curve-fitting techniques are applied to the Fourier amplitude spectrum for model development. In this section, we will show how the adaptive identification algorithm, the RPEM, presented above can be used for spectral estimation.

Assume that the record at hand is the output of a linear, time-varying filter due to an unknown input. Since we have no access to the input, we will assume that the input is a zero-mean, white-noise random process. As shown earlier in section 3, when evaluated at $z = e^{i2\pi fT}$ the modulus of the system transfer function gives the spectral ratio, the ratio of the Fourier amplitude spectrum of the output to that of the input. For a white-noise input, the input spectrum is constant; therefore, the spectral ratio is simply the scaled

form of the output spectrum. The scaling constant is equal to the RMS value of the white-noise input. Thus, for a black-box model structure with white-noise input, the spectral estimation of $y(t)$, $S_y(f)$, can be calculated from the equation

$$S_y(f) = \frac{|B(e^{-2\pi fT})|}{|A(e^{-2\pi fT})F(e^{-2\pi fT})|} \sigma_x \quad (10.1)$$

where σ_x is the RMS value of the zero-mean, white-noise input $x(t)$.

The adaptive filtering approach for spectral estimation is closely related to the maximum entropy spectrum introduced by Burg (1975). It can be shown that the maximum entropy spectrum corresponds to the model that $B(q) = F(q) = 1$ in Eq. 10.1 (van den Bos, 1971), which is the form that an AR (autoregressive) model would give. The use of more complicated models than the AR model allows the modeling of sharp dips as well as sharp peaks with fewer parameters.

11. ADAPTIVE CONTROL

An important extension of the stochastic-adaptive identification algorithm is adaptive control. Adaptive control involves simultaneous estimation of the system model and the control input as the output of the system is recorded. The difference from classical control is that the system is unknown. At every sampling point, the control mechanism first learns the system by using adaptive identification, then determines the necessary control input based on the control criterion. Such control mechanisms are known as *self-tuning regulators* (Åström and Wittenmark, 1973). The basic principles of the adaptive control will be summarized below by developing the adaptive control law for ARMAX models.

Consider an ARMAX model with a k -step time delay between input and output:

$$A(q)y(t) = q^{-k}B(q)u(t) + C(q)e(t) \quad (11.1)$$

where the input is denoted by $u(t)$ rather than $x(t)$ to emphasize the difference that $u(t)$ is the yet to be determined control input. The problem we want to solve is this: knowing $y(t)$ how can we determine $A(q)$, $B(q)$, $C(q)$, and $u(t)$ so that $y(t)$ becomes as close to a specified value $y^*(t)$ as possible?

Since there is a k -step time delay in the system, an input at time t can only affect the outputs at times $t+k$ and beyond. We determine the predicted output at time $t+k$ based on the information up to time t . By multiplying both sides of Eq. 11.1 with q^k we obtain

$$y(t+k) = \frac{B(q)}{A(q)}u(t) + \frac{C(q)}{A(q)}e(t+k) \quad (11.2)$$

Based on the data up to time t we can compute $e(t), e(t-1), \dots, e(0)$ as the difference between recorded outputs and one-step-ahead predicted outputs, i.e., $e(t) = y(t) - \hat{y}(t, \theta)$, $e(t-1) = y(t-1) - \hat{y}(t-1, \theta)$, \dots etc. (see section 5.1 on one-step-ahead prediction). We cannot calculate the values $e(t+1), e(t+2), \dots, e(t+k)$; however, we know that they are independent of the data up to time t . To separate the predictable and unpredictable portions of $e(t)$, let us write the second term on the right-hand side of Eq. 11.2 as follows

$$\frac{C(q)}{A(q)}e(t+k) = S(q)e(t+k) + \frac{G(q)}{A(q)}e(t) \quad (11.3)$$

where it can be shown that the order of polynomial $S(q)$ is $k-1$, whereas the order of polynomial $G(q)$ is the larger of $n_c - k$ and $n_a - 1$. From Eq. 11.3, by writing $e(t)$ as $q^{-k}e(t+k)$ and multiplying both sides by $A(q)$, we can obtain the following identity:

$$C(q) = A(q)S(q) + q^{-k}G(q) \quad (11.4)$$

Also from Eq. 11.1 we can write

$$e(t) = \frac{A(q)}{C(q)}y(t) - q^{-k}\frac{B(q)}{C(q)}u(t) \quad (11.5)$$

Using Eq. 11.5 in Eq. 11.3, and the resulting equation in Eq. 11.2, we obtain

$$y(t+k) = S(q)e(t+k) + \left[\frac{B(q)}{A(q)} - q^{-k}\frac{B(q)G(q)}{C(q)A(q)} \right] u(t) + \frac{G(q)}{C(q)}y(t) \quad (11.6)$$

Finding the expression for $q^{-k}G(q)$ from Eq. 11.4, and using it inside the bracket on the right hand side of Eq. 11.6 result in:

$$y(t+k) = S(q)e(t+k) + \frac{G(q)}{C(q)}y(t) + \frac{B(q)S(q)}{C(q)}u(t) \quad (11.7)$$

Since the term $S(q)e(t+k)$ is independent of data (i.e., it is a linear combination of future noise terms) it cannot be predicted with the available information. The best we can assume is that its expected value is zero. Therefore, by taking the expectation in Eq. 11.7, we can write for the k -step-ahead prediction $\hat{y}(t+k)$

$$\hat{y}(t+k) = E[y(t+k)] = \frac{G(q)}{C(q)}y(t) + \frac{B(q)S(q)}{C(q)}u(t) \quad (11.8)$$

For control, assume that we require $y(t) = y^*(t)$, where $y^*(t)$ is a specified sequence. Thus, for time $t+k$, we simply set the predicted output $\hat{y}(t+k)$ equal to the desired output $y^*(t+k)$. We can then write

$$\frac{G(q)}{C(q)}y(t) + \frac{B(q)S(q)}{C(q)}u(t) = y^*(t+k) \quad (11.9)$$

Equation 11.9 represents the k -step-ahead adaptive control law for the ARMAX model. The control input is calculated from the control law as

$$u(t) = -\frac{G(q)}{B(q)S(q)}y(t) + \frac{C(q)}{B(q)S(q)}y^*(t+k) \quad (11.10)$$

For $y^*(t) = 0$ the control law is known as the *minimum variance control*. For systems with unit delay (i.e., $k = 1$), where the control input at time t influences the output at time $t+1$, the equations are simplified significantly. For $k = 1$, it can be shown in Eq. 11.3 that $S(q) = 1$ and $G(q) = C(q) - A(q)$. Therefore, for unit delay and minimum variance control, the equation for control input can be written

$$u(t) = \frac{A(q) - C(q)}{B(q)}y(t) \quad (11.11)$$

As the above equations show, the adaptive control algorithm involves two steps, identification and control. At every time step the system is first identified by using an adaptive identification algorithm, such as the RPEM, then the control input is calculated based on the identified system. This procedure is repeated at every time step.

The algorithm given above is one of several algorithms used for adaptive control. It is chosen for its direct connection to the identification algorithm presented earlier. Although the theory of adaptive control is still being developed, there is a large number of references already available in the literature, where more rigorous analysis of the existing methods, stability, and convergence properties can be found (e.g., Åström, 1970, 1983; Wittenmark, 1975; Landau, 1979; Goodwin and Sin, 1983; Åström and Wittenmark, 1984; Aloneftis, 1987).

12. PREPROCESSING OF DATA

In general, recordings in their initial (i.e., raw) form do not provide a good identification of the system. Much better identification can be obtained if data is preprocessed prior to identification. Preprocessing involves removal of mean and outliers, filtering, decimation, and synchronization of input and output.

In identification we are interested in the dynamic characteristics of the system, since any static relationship can easily be determined by other means. The static part in the system is characterized by non-zero mean values in the input and output signals. The

identification algorithm RPEM can handle the static part along with the dynamic part by adding unknown constants in the parameter vector; however, this would increase the size of the problem, and adversely affect the convergence properties. A better alternative is to remove the mean values of the input and output prior to identification. There may also be isolated outliers (i.e., erroneous large peaks) at several points in the data due to various reasons, such as temporary sensor failure, an accidental shock to the instrument, etc. These outliers can disturb the identification, especially if the least-squares criterion is used for the error measurement. They should be removed prior to identification, unless special precautions are included in the algorithm as discussed in section 5.2.

Another important parameter in the identification is the selection of the sampling interval. Sampling interval is related to the frequency resolution of the signal through the Nyquist frequency, f_N . No frequency information beyond f_N can be obtained from the signal. Frequency components higher than f_N are folded back and superimposed over the smaller frequencies in the spectrum. This distortion is called *aliasing*. To prevent aliasing, signals should be filtered by using an anti-aliasing filter, which is a low-pass filter with cut-off frequency equal to f_N . In most situations, due to the characteristics of the system, we are only interested in up to a certain frequency, which may be much smaller than the Nyquist frequency. If this is the case, there is no need for a high sampling rate. High sampling rates force the algorithm to identify only the high frequency part of the system (Ljung and Söderström, 1983). We first low-pass filter the signal, then decimate it (i.e., increase the sampling interval) such that the Nyquist frequency is near the highest frequency in which we are interested. This would not only help to reduce the noise effects, since in general the noise spectrum has a broader band-width, but also reduce the size of the computations. We also need to high-pass filter the signal to eliminate the very low-frequency drifts in the signal. These filterings can be done off-line if we have access to the total record before the identification, or on-line simultaneous with the identification.

A final step in the preprocessing of data is the synchronization of the input and output, if they are not recorded in synchronous fashion. A lack of synchronization can be handled by properly selecting the time delay between input and output during the identification. In fact, the time delay may be due to physical characteristics of the system so that the input and output should not be synchronized. However, estimating the time delay is often not easy, and requires a trial and error procedure. It is best to use synchronized input and output, and assume that there is no time delay. The synchronization of input and output can be accomplished by matching some characteristic points in the signals, or by shifting them until the cross-correlation between input and output is maximum.

13. EXAMPLES

13.1. TIME-INVARIANT SIMULATED SYSTEM

The first example is a simulated time-invariant system, given by a sixth-order ARMAX model (see Table 4.1) with polynomials

$$A(q) = 1 + 0.4q^{-1} - 0.2q^{-2} - 0.4q^{-3} + 0.3q^{-4} + 0.5q^{-5} - 0.1q^{-6} \quad (13.1)$$

$$B(q) = 0.7q^{-1} - 0.7q^{-2} + 0.8q^{-3} + 0.4q^{-4} - 0.3q^{-5} \quad (13.2)$$

$$C(q) = 1 - 0.2q^{-1} + 0.7q^{-2} \quad (13.3)$$

The input sequence $x(t)$ is assumed to be a pseudorandom binary sequence (i.e., a random sequence of zeros and ones) with RMS value equal to one. The amplitude of the white-noise series $e(t)$ is determined for a given signal-to-noise ratio (SNR). The SNR is calculated as the ratio of the RMS value of the signal sequence $y_s(t)$ to that of the noise sequence $y_n(t)$. For an ARMAX model the sequences $y_s(t)$ and $y_n(t)$ are

$$y_s(t) = \frac{B(q)}{A(q)} x(t) \quad (13.4)$$

$$y_n(t) = \frac{C(q)}{A(q)} e(t) \quad (13.5)$$

The model is simulated for $SNR = 5$ assuming $N = 500$ points with $T = 0.05$ sec. for the time interval (although the actual value of the time interval is not needed in the estimation). The plots of output, signal, and noise series are given in Fig. 13.1.1. The output is the sum of the signal and noise sequences.

Assuming that the output and input are recorded and the noise is unknown, the parameters of the model are determined by using the RPEM algorithm. Initial values of the parameters and the gradient were assumed zero. The initial value for the matrix $\gamma(t)R^{-1}(t)$ in Eq. 5.38 is assumed to be diagonal with elements all equal to 1000. In order to remove the initial transient part created by the unknown initial conditions, three passes are made over the data. The first pass uses $\theta = 0$ as the initial conditions for the parameters. The other two passes use the θ values obtained at the end of the previous pass. The parameters were first estimated by using the values $\lambda_0 = 1$, $\lambda(0) = 1$ (Eq. 5.12) for the forgetting factors. This pair corresponds to a uniform weighting factor of one for all the points. As explained in section 5.2, uniform weighting factors give results that are least sensitive to the effects of noise, which is a desirable property, but they also make the algorithm worst for time-tracking ability. However, since the system is time-invariant,

$\lambda_0 = \lambda(0) = 1$ is appropriate. Fig. 13.1.2 gives the time variation of estimated parameters for this case. Exact values of parameters are also shown in Fig. 13.1.2 by arrows on the right hand side of the plots. The estimates are very good when compared to exact values. Recall that the parameters were estimated assuming that the noise is unknown.

To investigate the effect of weighting factors on the algorithm, another estimate of the parameters were made by using the forgetting factors $\lambda_0 = 0.99$ and $\lambda(0) = 0.95$, values suggested by Ljung and Söderström (1983) and Young (1984). These values correspond to a time-varying exponential weighting profile. The weighting factors have a sharp decay at the beginning of the data to eliminate rapidly the effects of initial conditions, then become flatter as more data points are fed. Figure 13.1.3 gives the profiles of weighting factors at two intermediate and the last sampling points (100, 300, and 500) during the recursion. The time variation of the estimated parameters are given in Fig. 13.1.4, again with the exact values marked by arrows on the right side of the plots. As the figure shows, the estimated values for the parameters are not constant. Instead, they fluctuate around the exact values such that their time averaged (by eye) values give a good estimate of the exact values. The reason for this is the exponentially decaying weighting factors used in the algorithm. With weighting factors, the algorithm has now the ability to track the time varying characteristics of the system, but at the same time the estimates are more sensitive to noise effects. Since the system is not time-varying, the use of weighting factors has adverse effects on the estimation.

To see the effect of noise level in the estimation, the system was simulated for two other signal-to-noise ratios, $SNR = 1$, and $SNR = 10$, and the parameters were estimated for uniform weighting factors ($\lambda_0 = \lambda(0) = 1$). The results are presented in Fig. 13.1.5 for $SNR = 1$, and in Fig. 13.1.6 for $SNR = 10$. As expected, the figures indicate that the higher the SNR value, the more accurate the estimates are. It is important to note in Fig. 13.1.5, however, that even for $SNR = 1$, where the RMS value of the unknown noise is equal to that of the signal (i.e., very noisy data), the algorithm still gives fairly good estimate of the parameters.

13.2. TIME-VARYING SIMULATED SYSTEM

Next we will consider a time-varying simulated system. Again, the system considered is an ARMAX system with orders $n_a = 2$, $n_b = 1$, and $n_c = 2$ for output, input, and noise, respectively. The system is defined by the polynomials

$$A(q) = 1 + a_1(t)q^{-1} + a_2(t)q^{-2} \quad (13.6)$$

$$B(q) = b_1(t)q^{-1} \quad (13.7)$$

$$C(q) = 1 - 0.75q^{-1} + 0.50q^{-2} \quad (13.8)$$

where it is assumed that the output and input parameters (a_1 , a_2 , and b_1) are time-varying, whereas the noise parameters ($c_1 = -0.75$ and $c_2 = 0.50$) are constant. For the time-varying parameters, the selected forms of time variation are plotted in Fig. 13.2.1. Since this is a second order system (i.e., $A(q)$ is second order), it corresponds to a simple-damped oscillator, whose frequency and damping can be calculated in terms of parameters a_1 and a_2 (see section 3). The time variation of corresponding oscillator frequency and damping values are plotted in Fig. 13.2.2.

The system is simulated by using a pseudorandom binary sequence for $x(t)$, and a white-noise sequence for $e(t)$. The length of the series is $N = 1000$, with time interval $T = 0.1$. The input is scaled such that its averaged Fourier amplitude is equal to one (it can be shown that a pseudorandom binary sequence has a flat spectrum, same as a white-noise sequence). The signal-to-noise ratio is assumed to be $SNR = 2$, and the RMS value of $e(t)$ was chosen accordingly. The plots for the signal and noise series, and corresponding Fourier amplitude spectra are given in Fig. 13.2.3. Note in Fig. 13.2.3 that the noise amplitudes are as high as half of the signal amplitudes. The output series (signal plus noise) and its Fourier amplitude spectrum are given in Fig. 13.2.4. Since the input is selected to have unit amplitude spectrum, the output spectrum also represents the spectral ratio. Also plotted in Fig. 13.2.4 superimposed over the Fourier spectrum are the exact transfer functions at three instances, $t = 20, 50$ and 80 sec . As the figure shows, the transfer function changes with time, whereas the spectral ratio is time-invariant. This example clearly shows that for time-varying systems use of spectral ratios to estimate transfer functions can give misleading results.

The system is identified by using the RPEM. It was first assumed that $\lambda_0 = \lambda(0) = 1$ (i.e., uniform weighting factors). As explained in the previous example, with these λ values the algorithm would not track the time variations in the system. Figure 13.2.5 shows the comparison of estimated values of the parameters (dashed lines) with the exact ones (solid lines). The estimated values are nearly constant, approximately equal to the average of the time-varying exact values, and cannot track the time variations in the system. The values estimated are probably the same as would have been obtained from a non-recursive (i.e., off-line) estimation algorithm.

Next, the parameters were estimated by using exponentially decaying weighting factors defined by the forgetting factor parameters $\lambda_0 = 0.99$, and $\lambda(0) = 0.99$. Having a value close to one for $\lambda(0)$ means that the weighting factor profile does change significantly with time. In other words, the time tracking ability of the algorithm is similar throughout the data. The estimated values and the comparison with the exact ones are given in Fig. 13.2.6. As seen in the figure, the estimation is much better this time. The time variation of the parameters is tracked fairly well. For the constant noise parameters, however, the

estimates are not very smooth for the reasons explained in the previous example.

13.3. IDENTIFICATION OF A BUILDING WITH SOIL-STRUCTURE INTERACTION

This example will show identification of a building with soil-structure interaction by using earthquake induced vibration recordings. The building considered is a 22-story, triangular in plan, reinforced concrete building located in Viña del Mar, Chile. The building has a mat foundation without piles resting on sand. Following the 3 March 1985, Central Chile earthquake of magnitude 7.8, the building was temporarily instrumented to obtain aftershock data (Çelebi, 1985). A set of records was obtained from an aftershock with a magnitude 4.5. Detailed analyses of the building and the data set can be found elsewhere (Çelebi, 1985; Bongiovanni et al., 1987). The analysis showed that there was a significant amount of soil-structure interaction in the form of rocking during the earthquake. Here, we will consider only two records from the data, just to give an example of the application of the RPEM.

The records considered are the velocities at the basement and the top floor, both recorded in the same direction. The original records had a sampling rate of 200 per second, which gives a frequency resolution of 100 Hz. From the size of the building it was decided that any frequency component beyond 10 Hz would be insignificant for the building. Thus, the records were low-pass filtered beyond 10 Hz, and decimated. They were also high-pass filtered beyond 0.1 Hz to remove any low-frequency drifts in the data. The processed velocity records of the basement and top floor are given in Fig. 13.3.1 along with their Fourier amplitude spectra. An ARMAX model is considered for the building. By using the criteria for model order selection given in section 8, the orders were selected as 16, 15, and 3 for the polynomials $A(q)$, $B(q)$, and $C(q)$ in the model. The records were synchronized by using the maximum cross-correlation criterion. The basement record is the input, and the top-floor record is the output of the model. Using the RPEM algorithm the parameters of the model were determined. Trying various forgetting factors indicated that the parameters were time invariant (i.e., building response is linear). Thus, the final results were obtained using $\lambda_0 = \lambda(0) = 1$ for the forgetting factor. The values for the parameters are given in Table 13.3.1. Also included in the table are the complex poles (modulus r_j , and argument ϕ_j), and corresponding natural frequencies (f_{0j}), damping ratios (ξ_{0j}), and weighting factors (last two columns). Note that all the poles have modulus less than one (i.e., they are inside the unit circle in the complex plane), two of the poles, poles one and ten, are real (i.e., pole's argument is zero, or corresponding damping ratio is one), and the remaining poles are in complex-conjugate pairs, each resulting in the same frequency, damping ratio, and weighting factor. The most dominant mode of the system is the one corresponding to the pole closest to the unit circle (i.e., pole with the highest r value).

For this building, it is the 13th (or the 14th) pole, which corresponds to a frequency of 0.917 Hz., and damping ratio 0.014. The estimated transfer function for the building and its match with the spectral ratio is shown in Fig. 13.3.2. The spectral ratio is obtained as the ratio of the smoothed Fourier amplitude spectrum of the top-floor record to that of the basement record. The figure shows a good match. The dominant mode of the building is clearly seen in the figure. Recall that, unlike the spectral ratio, the estimated transfer function is completely known analytically in the form of a recursive filter. The filter is simply the weighted sum of ten damped oscillators (as shown in Table 13.3.1) all with known frequency, damping, and weighting factors.

As discussed in section 9, after the model is selected and the parameters are determined, the next step is to check the validity of the model. The match of the transfer function with spectral ratio has already been given in Fig. 13.3.2. Next we will check the residuals of the identification following the procedure given in section 9. Residuals are defined as the difference between the model output and the recorded output. In order to have a valid identification, the residuals should be a white-noise sequence. The plot of the residual time series and its Fourier amplitude spectrum are given in Fig. 13.3.3. Visual inspection of the spectrum suggests that the residuals are close to a white-noise sequence, since the peaks are distributed in all frequencies. A more accurate check is to test the autocovariance of the residuals. This is also given in Fig. 13.3.3. Two straight lines in the figure show the 90-percent confidence levels for whiteness based on the χ^2 test. For model validity (with 90-percent confidence) the autocovariance should not exceed these levels, except at zero lag. Fig 13.3.3 shows that this criterion is also satisfied. Another check is the cross-covariance of the residuals with input. For positive correlation lags, the cross-covariance shows whether more information can be extracted from residuals by the input. For negative correlation lags, the cross-covariance shows if there is a feedback in the system (i.e., if there are any components common to both input and the residuals). The cross-covariance is also given in Fig. 13.3.3 with 90-percent confidence levels (straight lines). Figure 13.3.3 shows that the confidence criterion is satisfied for positive lags but not for negative lags. This does not necessarily mean the model is wrong, but an indication that the system has a feedback. In other words there is a frequency component in the motion common to both basement and top-floor records. The physical explanation of this is that the building is subjected to rocking which has already been shown by previous studies on the building (Bongiovanni et al. 1987).

As a final test for modal validity, a comparison of model output with the recorded output is given in Fig. 13.3.4. This is a much more strict test than the previous ones. However, Fig 13.3.4 shows that the match is fairly good. Based on all these, we can conclude that the estimated model for the building is satisfactory.

13.4. IDENTIFICATION OF A BUILDING WITH NONLINEAR BEHAVIOR

The second building investigated is the Imperial County Services Building in El Centro, California, which is a six-story, frame and shear-wall, office building. This building suffered significant damage during the October 15, 1979, magnitude 6.5, Imperial Valley, California earthquake. At the time of the earthquake the building was instrumented by a 13-channel accelerometer, central recorder system, installed by the California Division of Mines and Geology. The data set obtained during the earthquake contributed significantly to the understanding of inelastic behavior of structures during earthquakes. A large number of studies were done by using the data set. The detailed analysis of the recordings can be found elsewhere (e.g., Rojahn and Mork, 1982). Here, again by using only two records, the dynamic characteristics of the building and their time variations will be investigated. The records used are the east-west accelerations at the ground level and at roof (records 13 and 4 in Rojahn and Mork, 1982). The records were preprocessed as in the previous building. Time histories of the records are given in Fig. 13.4.1. Corresponding Fourier amplitude spectra and the spectral ratio (roof/ground level) are given in Fig. 13.4.2. Following the guidelines for model selection, an ARX model with orders 12 and 11 for the polynomials $A(q)$ and $B(q)$ was selected. Since it was known that the building was damaged extensively during the earthquake (i.e., the response was nonlinear), a forgetting factor with parameters $\lambda_0 = 0.99$ and $\lambda(0) = 0.95$ was used in the algorithm, so that the time variations in the parameters can be tracked properly. Calculated time variations of the parameters are plotted in Fig. 13.4.3. The parameters clearly show three distinct regions. The initial part between 0 to 3 seconds corresponds to linear behaviour of the original structure. The second part between 3 to 9 seconds is where the damage takes place. The final part from 9 seconds and beyond corresponds to vibration of the damaged structure.

The transfer functions from each region calculated at times 1, 5, and 25 second are shown in Fig. 13.4.4. Their comparison with the spectral ratio is given in Fig. 13.4.5. As the figures show the transfer function changes with time, which can not be detected by the the spectral ratio approach. The values of the parameters, including the poles and the corresponding frequencies, damping ratios, and weighting factors are given in Table 13.4.1 for the damaged structure at time $t = 25$ seconds.

The residuals of the identification, also at $t = 25$ sec. are plotted in Fig. 13.4.6, along with its Fourier amplitude spectrum, autocovariance function, and the cross-covariance with input. As seen in the figure, the autocovariance of the residuals exceeds the 90-percent confidence bounds for small lags; however, the cross-covariance satisfies the confidence criterion. This observation suggests that the model gives a good estimate of the transfer

function for the actual system, but not for the noise.

13.5. IDENTIFICATION OF A BUILDING FROM AMBIENT VIBRATION DATA

The identification algorithm RPEM can also be used to identify structures based on ambient vibration data. The difference in the ambient vibration case is that there is no record of the input. However, the ambient excitations, such as wind and traffic noise, generally have a broad band width. Thus, a logical assumption for the input would be a white-noise sequence. Since white noise has a flat spectrum, the estimated transfer function becomes equivalent to the output spectrum scaled by a constant (see section 10). The unknown scaling constant affects only the numerator of the transfer function. Since the denominator is unchanged, the modal frequencies and damping ratios can still be determined accurately. The modal weighting factors, however, cannot be determined because of the unknown input.

As an example for ambient vibration analysis, ambient acceleration data from the top-floor of a 9-story reinforced-concrete building is considered. The data and its Fourier amplitude spectrum are given in Fig. 13.5.1. Going through the model selection procedure, an ARX model with orders 24 and 23 for the polynomials $A(q)$ and $B(q)$ were found to be appropriate. Since the data is stationary, forgetting factor is assumed to be one. The parameters were estimated using a simulated white-noise sequence with unit spectral amplitude for input. The values for parameters and the corresponding modal properties are presented in Table 13.5.1. The estimated transfer function and its match with the output Fourier amplitude spectrum are given in Fig. 13.5.2. Note in the figure that the closely spaced two large peaks between 1 and 2 Hz are matched very well. The closeness of these two dominant peaks required the use of such a high order model. The model validity tests on the residuals are summarized in Fig. 13.5.3. The tests show that the model is appropriate.

13.6. SPECTRAL MODELING OF EARTHQUAKE GROUND MOTIONS

A common approach for spectral modeling of earthquake ground motions is to use the Fourier amplitude spectrum. This approach assumes that the signal is stationary, and the signal to noise ratio is high (or the signal and noise spectra do not overlap). For ground motion records, the first assumption is usually not valid. The second assumption is valid only for a certain frequency band. In this example, we will estimate the spectrum by using the RPEM as explained in section 10. The frequency domain plots for this and the following two examples are all given on log-log scales in order to be consistent with the usual practice in earthquake engineering and seismology.

Assume that the motion recorded on the surface is the output of a linear, time-varying filter due to an unknown input. Since we have no access to the input, it will be assumed that the input is a random process of specified characteristics. One such process, suggested by Kormylo (1979) and Mendel (1983), is the Bernoulli-Gaussian sequence, which is the product of a Bernoulli sequence (i.e., a random sequence of zeros and ones) and a zero-mean, Gaussian white-noise sequence. The Bernoulli-Gaussian sequence can be considered as a random spike sequence. A sample function of a Bernoulli-Gaussian sequence and its Fourier amplitude spectrum are given in Fig. 13.6.1. It can be shown that a Bernoulli-Gaussian sequence is white; thus, the sample values are statistically independent, and its spectrum is constant (Mendel, 1983). We will model the system by an ARX model, and simulate the Bernoulli-Gaussian input such that it has unit variance. With these, the spectrum of the output can be written by putting $\sigma_x = 1$ and $F(q) = 1$ in Eq. 10.1 as

$$S_y(f) = \frac{|B(e^{-2\pi fT})|}{|A(e^{-2\pi fT})|} \quad (13.9)$$

It should be noted here that $x(t)$ can also be assumed to be an ordinary white-noise sequence similar to $e(t)$, and the parameters of $x(t)$ and $e(t)$ can be combined into a single set of parameters. This results in ARMA models (see Table 4.1). Time-invariant forms of ARMA models have been used by a number of researchers for ground motion modeling (Polhemus and Çakmak, 1981; Chang, et al. 1982; Çakmak, et al. 1985). It is suggested that the Bernoulli-Gaussian sequence is a more realistic representation of the earthquake source mechanism (Mendel, 1983).

An example for spectral modeling will be given by using a record from the 1971 San Fernando earthquake (record F88S70E obtained at the basement of Glendale municipal services building). The time history and the Fourier amplitude spectrum of the record are given in Fig. 13.6.2. The model order is selected by investigating the variation of total prediction error with parameter n_a . This variation is given in Fig. 13.6.3. The figure suggests that a model with $n_a > 4$ is appropriate. As a first choice, $n_a = 4$, $n_b = 3$, and $n_c = 0$ (4-3 model) were used. The values for the parameters a_j and b_j are calculated recursively by using the RPEM algorithm. Their time variations are given in Fig. 13.6.4. The a_j values are fairly constant beyond 4 seconds. Recall that the a_j 's determine the location of the peaks (i.e., the dominant frequencies) of the transfer function. The b_j values exhibit more fluctuations, especially within the first 10 seconds. b_j 's do not affect the peak locations, but the peak amplitudes. The initial part of the curves, about 10 steps or so (i.e., up to the 0.2 second mark), represents the transient part where the results are not reliable due to the effects of unknown initial values. The values of a_j and b_j , pole locations, corresponding modal frequencies and damping ratios, and the weighting factors for each mode are presented for $t = 18$ seconds in Table 13.6.1. The dominant mode is the one that

corresponds to the pole closest to the unit circle (i.e., the mode corresponding the largest r_j value in Table 13.6.1). The corresponding amplitude plot of the transfer function, and the match with the Fourier amplitude spectrum are given in Fig. 13.6.5. The match seems satisfactory, except for very low frequencies.

A closer look at the model fit can be made by investigating the residuals of the model. They are presented in Fig. 13.6.6. According to Fig. 13.6.6, although the amplitudes of the autocovariance function are fairly small, they exceed the 90-percentile value at significantly high lags. The test for cross-covariance of the residuals with input also fails at one point, although overall it is better than the autocovariance test.

For comparison, consider a higher order model with $n_a = 8$, $n_b = 7$, and $n_c = 0$ (8-7 model). The amplitude of the frequency response function, again at $t=18$ seconds, and the match with the Fourier amplitude spectrum is given in Fig. 13.6.7, with the corresponding values of the parameters, pole locations, frequencies and damping ratios, and the modal weighting factors in Table 13.6.1. Since $n_a = 8$ there are four peaks in the transfer function. The increase of the model order improves the match at higher frequencies. The residuals and the related tests are given in Fig. 13.6.8. The autocovariance and cross-covariance functions are also better in this model. The autocovariance function exceeds the confidence level for small lags, but it passes the test beyond 0.1 second. The cross-covariance function passes the confidence test for all lags. Further studies showed that higher order models did not improve the match at low frequencies, nor the test for autocorrelation. This is believed to be because of low-frequency errors in the Fourier spectra caused by windowing.

Two more figures, Figs. 13.6.9 and 13.6.10, are presented to show the variation of the transfer function with time and with model order. Fig. 13.6.9 shows the superimposed transfer functions (i.e., estimated spectra) at 1, 5, and 18 seconds for (8-7) model. The figure confirms that the transfer function varies with time. Fig. 13.6.10 shows the superimposed transfer functions for (4-3), (8-7), and (12-11) models, at 18 seconds. The higher-order models basically improve the match at higher frequencies (i.e., a two order increase in n_a results in the determination of the next higher mode). The low frequencies do not seem to be affected by model increase.

13.7. SITE AMPLIFICATION OF EARTHQUAKE GROUND MOTIONS

It is well known that local geology and topography can significantly alter the characteristics of earthquake ground motions. The alteration is in the form of amplification of the components at certain frequencies, and attenuation of the components at other frequencies. This observation suggests that site amplification can be approximated by a finite order filter.

The commonly used technique for quantifying site amplification for a given site has been to calculate the spectral ratio, that is, the ratio of the Fourier amplitude spectrum

of the record at that site to the Fourier amplitude spectrum of the record at a nearby reference site (usually a rock site). The accuracy of the spectral ratio approach is very much dependent on the signal-to-noise ratio. The standard deviation of the calculated spectral ratio from the exact ratio at frequency f is directly proportional to $A_n(f)/A_x(f)$, where $A_n(f)$ and $A_x(f)$ are the Fourier amplitude spectra of the noise and input, respectively (Rake, 1980; Ljung, 1985). To reduce the effect of noise, various smoothing techniques are applied to the input and output spectra before taking the ratio (Robinson, 1967). Irrespective of how accurate they are, the spectral ratios that are calculated as above are difficult to describe analytically in terms of simple parameters, especially when there are several dominant peaks.

Here, it will be assumed that the site amplification can be approximated by a discrete linear filter with input $x(t)$ and output $y(t)$, representing the recorded ground accelerations from a rock site and from a soil site, respectively. Assume that the sites are close to each other, so that the differences in the signals are due to site effects only (i.e., no distance effects). The transfer function of this system represents a linear model for the site amplification. When evaluated at $z = e^{i2\pi fT}$, the amplitude of the transfer function gives the spectral ratio in the usual sense; and when expanded into partial fractions as discussed in section 3, the transfer function gives the equivalent parallel system for the site in terms of modal frequencies, dampings, and weighting factors.

As an example, the approach presented above was applied to a site amplification problem observed during the March 3, 1985, magnitude 7.8, central Chile earthquake. The investigations after the earthquake showed that the local geological and topographical site conditions caused large variations in ground motion amplitudes and structural damage (Çelebi, 1987). The records are from a rock site, and an alluvial site. The distance between the stations is five km. The strong-motion accelerograms are shown in Fig. 13.7.1. Corresponding Fourier amplitude spectra and the smoothed spectral ratio are given in Fig. 13.7.2. It is clear from Figs. 13.7.1 and 13.7.2 that the amplitudes and the frequency contents of the motions are significantly different.

After investigating the variation of total estimation error with model order, an ARX model with $n_a = 12$ and $n_b = 11$ is chosen. The parameters were calculated recursively by using the RPEM algorithm. Their time variations are plotted in Fig. 13.7.3. The amplitude of the transfer function, and its match with the spectral ratio at $t = 18$ sec. are given in Fig. 13.7.4. The corresponding numerical values for the parameters, pole locations, modal frequencies and damping ratios, and the weighting factors are given in Table 13.7.1. Since $n_a = 12$, the model gives the first six modes of the site. The results of model validity checks on the residuals are given in Fig. 13.7.5., which shows the time series of the residuals, and its amplitude spectrum, autocovariance, and cross-covariance

with the input. The 90-percentile confidence levels are also given in the figure. The autocovariance curve crosses the 90-percentile boundaries at several points, although the overall correlation is fairly small. The cross-covariance curve crosses the boundaries at one location.

Next, a higher order model with $n_a = 16$ and $n_b = 15$ is tried. The transfer function and its match with the spectral ratio are plotted in Fig. 13.7.6, again at $t = 18$ sec. The corresponding numerical values for the parameters and modal characteristics are given in Table 13.7.1. The residual tests are presented in Fig. 13.7.7, which shows slight improvement over the previous model. The cross-covariance values now are all inside the 90-percentile confidence interval. The auto-covariance curve still exceeds the boundaries at some points, but the amplitudes are smaller than those of the previous model. Further studies on this site done using different models and records have basically shown that models beyond 16th order do not provide any improvement.

An important comment that should be made here is that the procedure used for site amplification in this example is applicable only to sites with a non-layered medium. For a layered medium, there is a strong correlation between the input and the output (i.e., part of the output becomes input), because of the reflection of waves at the layer boundaries. This makes the system a closed-loop system, whereas the RPEM algorithm presented is for open-loop systems.

13.8. SOURCE SCALING OF EARTHQUAKE GROUND MOTIONS

The term source scaling is used to characterize the relationship between two records obtained at the same site from two different earthquakes. In seismology, this is known as the source deconvolution, or the empirical Green's function deconvolution. It is assumed that the site effects and the path effects in the records are the same, so that the differences between the records can be attributed solely to the source effects. The knowledge of source scaling is important for predicting motions for large earthquakes by extrapolating from small earthquakes. As for the site amplification, the common approach has been to use the spectral ratio of single station recordings of co-located events.

Here, again, the discrete linear filtering approach will be used to model the source scaling. Assume, at a given site, that the motion from a large earthquake is a linear function of motions from small earthquakes, represented by an integral (e.g., a convolution) or a differential equation. It can be shown that this assumption leads to a discrete time domain equation for the relationship between the large earthquake and the small earthquake in the form given by an ARX model (Table 4.1). The output and input of the model are the records from the large earthquake, and the small earthquake, respectively.

To give an example, two records from aftershocks of the 1983 Borah Peak, Idaho, earthquake are considered (Boatwright, 1985). The records are a day apart, and have

the RMS values of 2.5 and 14.2 cm/sec^2 . The plots of the time series are given in Fig. 13.8.1. Figure 13.8.2 gives the corresponding Fourier amplitude spectra and the smoothed spectral ratio (large earthquake/small earthquake). After studying the variation of error with model types and orders an ARX model with $n_a = 6$ and $n_b = 5$ is selected. The time variation of the estimated parameters are plotted in Fig. 13.8.3. The numerical values for the parameters, and the modal characteristics are given in Table 13.8.1 at $t = 10$ sec. The corresponding amplitude of the transfer function and the match with the spectral ratio are shown in Fig. 13.8.4. The model validity tests on the residuals are presented in Fig. 13.8.5. The tests for autocorrelation and cross-correlation in Fig. 13.8.5 show that the residuals fail the 90-percentile confidence criteria, suggesting a higher order model. Thus, a higher order model with $n_a = 14$, $n_b = 13$, and $n_c = 0$ was selected next. The resulting transfer function and the match with the spectral ratio is given in Fig. 13.8.6; the numerical values for parameters are given in Table 13.8.1; and the residual checks are given in Fig. 13.8.7. The residual tests show the model is appropriate for the transfer function.

13.9. ADAPTIVE CONTROL OF A SIMULATED SYSTEM

This example will show adaptive control of a simulated system by using the theory presented in section 11. Consider a system described by an ARMAX (Table 4.1) model with polynomials

$$A(q) = 1 - 1.5q^{-1} + 0.6q^{-2} \quad (13.10)$$

$$B(q) = q^{-1} + 0.5q^{-2} \quad (13.11)$$

$$C(q) = 1 + 0.2q^{-1} - 0.5q^{-2} \quad (13.12)$$

Assume that we want to control the output by using the minimum variance control strategy. If the system is identified previously by using some other inputs, so that the polynomials $A(q)$, $B(q)$, and $C(q)$ are all known prior to control, then the calculation of the control signal $u(t)$ is straightforward from Eq. 11.11. Using Eq. 11.11, we can write for the minimum variance control input

$$u(t) = \frac{-1.7 + 1.1q^{-1}}{1 + 0.5q^{-1}} y(t) \quad (13.13)$$

To see how the control works, the system is simulated assuming $\sigma_e = 2.5$ for the standard deviation of $e(t)$. The results are given in Fig. 13.9.1, which presents the output without control (i.e., $u(t) = 0$), output with minimum variance control, and the input required for the control. As seen from the figure, the control signal significantly reduces the output amplitudes.

If the system characteristics are unknown, the control should be performed simultaneously with system identification. Let us consider the same system, assuming that we do not know the coefficients of the polynomials. We have access only to the uncontrolled output signal (signal obtained assuming $u(t) = 0$ in the model, which is plotted at the top of Fig. 13.9.1). Assume that we chose a control model the same as the original system model, i.e., an ARMAX model with second order polynomials for $A(q)$, $B(q)$, and $C(q)$. Using the two-step algorithm the coefficients of the polynomials are determined along with the control signal for a minimum variance control. The results are presented in Figs. 13.9.2 and 13.9.3. Figure 13.9.2 gives uncontrolled output, controlled output, and control signal, while Fig. 13.9.3 gives the time-variation of parameters. In plotting the control input it is assumed that the amplitudes remain constant between the time intervals. As seen in Fig 13.9.2, control is accomplished in a few seconds. The larger amplitudes at the beginning of the controlled signal are due to initial transient phase in the parameter estimation algorithm. Figure 13.9.3 shows that estimated values for the parameters are not exactly equal to those of the original system, although the model type and polynomial orders were all assumed the same as the original system. This is due to the closed-loop nature of the system, where the input is determined from the output. Thus, the algorithm should not be expected to give the same parameters as the open-loop original system, since it is designed for control rather than identification.

To show that the control can still be achieved without exactly matching the system, we next consider an ARX model for the same system, assuming a fourth order polynomial for $A(q)$ and a third order polynomial for $B(q)$. The results for this case are given in Figs. 13.9.4 (input and output sequences), and in Fig 13.9.5 (time variation of parameters). Although a different model and different number of parameters are used, the system is still controlled successfully.

13.10. ADAPTIVE CONTROL OF AMBIENT VIBRATIONS OF A BUILDING

The last example will show how the ambient vibration amplitudes of the building in section 13.5 can be reduced by applying adaptive control techniques. For control purposes a much smaller order model was used. An ARX model with orders 6 and 2 for the polynomials $A(q)$ and $B(q)$ is considered for control. Forgetting factor parameters used are $\lambda_0 = 0.99$ and $\lambda(0) = 0.95$ (using $\lambda_0 = \lambda(0) = 1$ gave the same results). Using the adaptive control algorithm of section 11, parameters of the model and the corresponding control input were determined at every sampling point. Figure 13.10.1 shows the uncontrolled output, controlled output, and the required control input. As the figure shows, the control is achieved very quickly, in about two seconds. The time variation of control parameters are shown in Fig. 13.10.2.

It is important to note that the example presented shows only the application of the adaptive control algorithm. In other words, it is an example for the algorithm, not for the control mechanism. The actual implementation of the control involves designing the mechanism which would simultaneously generate and apply the calculated control input to the structure. Such a mechanism may be easy to develop for equipment-type structures, but not for buildings. However, the technique has successfully been applied to fairly large aerospace structures.

14. DISCUSSION AND FURTHER APPLICATIONS

Stochastic-adaptive filtering, prediction, and control techniques are powerful tools that can be applied to a wide range of problems dealing with analysis of discrete-time signals. One application is the adaptive signal processing (Widrow and Stearns, 1984), where the prediction, filtering, and smoothing of time series can be performed by using the same formulation. We have already given the equations for one-step-ahead and k -step-ahead predictions (Eqs. 5.2 and 11.8). In Eq. 11.8, if $k = 0$ the resulting equation gives the filtered value of the output at t , whereas if $k < 0$ the resulting equation gives the smoothed value of the output at $t - k$. Another application is the adaptive noise cancelling (Widrow, et. al., 1975), where the noise in a recording is removed in the time domain by using a reference noise recording.

The examples presented in the report are related to building vibrations and earthquake ground motions. The main goal in the examples is to present the methodology, rather than presenting the best application or the best example. Similar techniques can be applied to data from other branches of structural engineering, such as wind velocity vs. building motion in wind engineering, or wave height vs. platform motion in offshore dynamics.

Stochastic-adaptive techniques are based on solid theoretical foundations, and they have been used successfully in various other engineering fields during the last ten years or so. However, straightforward applications of theoretical equations to raw data can give very misleading results. Examples presented in this paper, and numerous other examples which will not be given here, strongly suggest that the preprocessing of data and the selection of model type, model order, and initial values are crucial for accurate identification. There are no set rules for these choices. The familiarity with the behavior of the system, and the experience with analysis of similar data are the most useful assets in applications. One more point regarding the application of the algorithm is the computational errors for systems with very low damping. Due to fixed-point, finite word length arithmetic in computers, the roundoff errors can accumulate to a significant level. It is recommended that the highest precision available in the computer be used during the computations.

The RPEM presented in the report incorporates the time-varying characteristics of the

model by using an exponentially decaying running window. A more complicated approach would be to introduce an additional set of equations that actually models the time variation of parameters as an auto-regressive process (Bohlin, 1977). The model used in this study is based on the transfer function representation and parallel form realization of system equations. There are other forms which can also be used to model systems, such as state-space forms and ladder forms (state-space form with diagonal covariance matrix for the state vector). The first form leads to the use of Kalman-Bucy filters (Kalman and Bucy, 1961), whereas the second form leads to lattice filters (Morf, 1977) for identification. Application of these filters for modeling and identification in structural dynamics will be presented in a separate study.

15. SUMMARY AND CONCLUSIONS

This report gives a concise review of the theory of stochastic-adaptive filtering, prediction, and control techniques for discrete-time recordings, and presents ten examples for applications in structural dynamics. Included in the theoretical section of the report are the discrete-time representation of single-input, single-output (SISO) systems, models for SISO systems with noise, the concept of stochastic approximation, the recursive prediction error method (RPEM) for system identification, spectral estimation, and adaptive control. The practical aspects of the RPEM as regards the convergence and consistency of the results, model selection, and model validation are also given in the report. The applications include identification and control of building vibrations and simulated systems, as well as analysis of earthquake ground motions.

It is clear that the stochastic-adaptive methods have several advantages over the currently used techniques. The major ones are: (a) the filtering of the noise is done over the whole frequency band, (b) time-varying characteristics of the signal can be tracked, (c) the models obtained are simple recursive filters; (d) due to the recursive form of the algorithm, only a small segment of the data is needed during the computations, and (e) the method provides a basis for applications of adaptive control techniques to unknown systems.

More on stochastic-adaptive filtering, prediction, and control can be found in Åström and Eykhoff (1971), Söderström and Stoica (1983), Ljung and Söderström (1983), Young (1984), Goodwin and Sin (1984), Graupe (1984), Ljung (1987), Brogan (1987), Caines (1988), and Söderström and Stoica (1988). The special issues of the journals *IEEE Transactions on Automatic Control* (1974), and *Automatica* (1980) give reviews of the methods. Papers by Saridis (1974), Isermann et. al. (1974), and Söderström and Stoica (1981) give comparison of different adaptive identification algorithms, whereas papers by Åström (1980), Isermann (1980), and Wittenmark and Åström (1984) give practical aspects for applications.

ACKNOWLEDGMENT

I would like to thank to Dr. M. Çelebi and P. Mork of the U. S. Geological Survey in Menlo Park for providing some of the data used in the examples.

REFERENCES

- Akaike, H. (1981). Modern development of statistical methods, in *Trends and Progress in System Identification*, (Ed. by P. Eykhoff), Pergamon Press, Elmsford, N.Y.
- Albert, A. E., and L. A. Gardner (1967). *Stochastic Approximation and Nonlinear Regression*, MIT Press, Cambridge, Mass.
- Aloneftis, A. (1987). *Stochastic Adaptive Control Results and Simulations*, Lecture Notes in Control and Information Sciences, 98, Springer-Verlag, Heidelberg.
- Anderson, B.A.O., J.B. Moore, and R.M. Hawkes (1978). Model approximation via prediction error identification, *Automatica*, Vol. 14, pp 615-622.
- Anderson, T.W. and J.B. Taylor (1979) Strong consistency of least- squares estimates in dynamic models, *Ann. Stat.*, Vol. 7, pp. 484- 489.
- Anscombe, F.J. and J.W. Tukey (1963). The examination and analysis of residuals, *Technometrics*, Vol. 5, pp. 141-160.
- Åström, K.J. (1970). *Introduction to Stochastic Control Theory*, Academic Press, New York.
- Åström, K.J. (1980). Maximum Likelihood and Prediction Error Methods, *Automatica*, Vol. 16, pp. 551-574.
- Åström, K.J. (1983). Theory and applications of adaptive control- A survey, *Automatica*, Vol. 19, pp. 471-486.
- Åström, K.J. and T. Bohlin (1965). Numerical identification of linear dynamic systems from normal operating records, *Theory of Self-Adaptive Control Systems* (Ed. by H. Hammond), Plenum Press, New York.
- Åström, K.J. and B. Wittenmark (1973). On self-tuning regulators, *Automatica*, Vol. 9, pp. 195-199.
- Åström, K.J. and T. Söderström (1974). Uniqueness of the maximum likelihood estimates of the parameters of an ARMA model. *IEEE Trans. Auto. Contl*, Vol. AC-19, pp. 769-773.
- Åström, K.J. and B. Wittenmark (1984). *Computer Controlled Systems*, Prentice-Hall, Englewood Cliffs, N.J.
- Åström, K.J. and P. Eykhoff (1971). System Identification-A Survey, *Automatica*, Vol. 7, pp. 123-162.
- Automatica* (1980). Special Section-System Identification Tutorials, Vol. 16, No. 5.
- Bartlet, M.S. (1946). The theoretical specification and sampling properties of autocorrelated time series, *J. Roy. Stat. Soc., Ser. B*, 8, pp. 27-41.
- Beck, J.L. (1978). Determining models of structures from earthquake records, *Earthquake Engineering Research Laboratory*, EERL 78-01, California Institute of Technology, Pasadena, California.
- Bierman, G. J. (1977). *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York.
- Boatwright, J. (1985). Characteristics of the aftershock sequence of the Borah Peak, Idaho earthquake determined from digital recordings of the events, *Bull. Seism. Soc. Am.* 75, 1265-1284.
- Bohlin, T. (1977). Four cases of identification of changing systems, *System Identification: Advances and Case Studies* (R.K. Mehra and D.G. Lainiotis, editors), Academic Press, New York, NY.
- Bohlin, T. (1987). Validation of identified models, in *Encyclopedia of Systems and Control* (Ed. by M. Singh), Pergamon Press, Oxford.
- Bongiovanni, G., M. Çelebi, and E. Şafak (1988). Seismic rocking response of a triangular building founded on sand, *Earthquake Spectra*, Vol. 3, pp. 793-810.
- van den Bos, A. (1971). Alternative interpretation of maximum entropy spectral analysis, *IEEE Transactions on Information Theory*, vol. IT-17, pp. 493-494.
- Box, G.E.P. and G.M. Jenkins (1970). *Time series analysis, forecasting and control*, Holden-Day, San Francisco, CA.
- Brogan, W.L. (1987). An introduction to system identification, in *Short Course Notes on Kalman Filtering II*, UCLA Extension, Oct. 25-30.

- Burg, J.P. (1975). Maximum entropy spectral analysis, *Ph.D. Thesis*, Department of Electrical Engineering, Stanford University, Stanford, CA.
- Cadzow, J.A. (1973). *Discrete-Time Systems*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Çakmak, A.S., R.I. Sherif, and G.W. Ellis (1985). Modeling earthquake ground motions in California using parametric time series method, *Soil Dynamics and Earthquake Engineering*, Vol. 4, No. 3, pp.124-131.
- CDMG (1987). *Significant building records obtained by the California Strong Motion Instrumentation Program*, California Division of Mines and Geology (CDMG), Sacramento, California.
- Çelebi, M. (1985). Preliminary evaluation of performance of structures (*in Preliminary report of investigations of the central Chile earthquake of March 3, 1985, edited by S. T. Algermissen*), *U. S. Geol. Surv. Open-File Rep.* 85-542.
- Çelebi, M. (1987). Topographical and geological amplification determined from strong-motion and aftershock records of the 3 March 1985 Chile Earthquake, *Bull. Seismol. Soc. Am.*, Vol. 77, pp. 1147-1167.
- 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, pp.651-662.
- Cramer, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N.J.
- Dugard, L., I.D. Landau, and H.M. Silveria (1980). Adaptive state estimation using MRAS technique - convergence analysis and evaluation, *IEEE Trans. Auto. Contl.*, Vol. AC-25, pp. 1169-1182.
- Fine, T. and W.E. Hwang (1979). Consistent estimation of system order, *IEEE Trans. Auto. Contl.*, AC-24, pp. 387-402.
- Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics, *Philos. Trans. Roy. Soc. London Ser. A*, 222, pp. 309-368.
- Fogel, E. (1981). A fundamental approach to convergence analysis of least squares algorithms, *IEEE Trans. Auto. Contl.*, Vol. AC-26, No. 3, pp. 646-655.
- Friedlander, B. (1982). A Modified Prefilter for Some Recursive Parameter Estimation Algorithms, *IEEE Trans. on Auto. Contl.*, Vol. AC-27, No. 1, pp. 232-235.
- Fuchs, J.-J. (1987). On estimating the order of an ARMA process, *Automatica*, Vol. 23, pp. 779-782.
- Gauss, K. F. (1809). *Theoria motus corporum caelestium*. English translation: *Theory of the Motion of the Heavenly Bodies*, Dover, New York (1963).
- Gersch, W. and S. Luo (1972). Discrete time series synthesis of randomly excited structural system response, *J. Acoust. Soc. Amer.*, 51, pp. 402-408.
- Goodwin, G. C. and K. S. Sin (1984). *Adaptive Filtering, Prediction and Control*, Prentice-Hall, Englewood Cliffs, N.J.
- Graupe, D. (1984). *Time Series Analysis, Identification, and Adaptive Filtering*, Robert E. Krieger Pub., Malabar, Florida.
- Haber, R. (1985). Nonlinearity tests for dynamic processes. *proc.*, 7th IFAC Symp on System Identification, York, U.K., pp. 409-413.
- Hannan, E.J. (1973). The asymptotic theory of linear time series models. *J. App. Prob.*, Vol. 10, pp. 135-145.
- Hannan, E.J. and B.G. Quinn (1979). The determination of the order of an autoregression, *J. R. Stat. Soc. Set-B*, 41, pp. 190-195.
- Householder, A. S. (1964). *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York.
- Hsia, T. C. (1977). *System Identification*, Lexington Books, Lexington, Mass.
- Hudson, D.E. (1979). *Reading and interpreting strong motion accelerograms*, Earthquake Engineering Research Institute, El Cerrito, California.
- Ibanez, P. (1987). Experimental evaluation of offshore tower dynamics, *proc. Soc. for Exper. Mech.*, Spring Conf., June 1987, Houston, Texas, pp.587-589.
- IEEE Trans. on Auto. Contl.* (1974). Special Issue on System Identification and Time-Series Analysis, Vol 19, No. 6.

- Inagaki, M. (1981). On Estimating the Orders of an Auto regressive Process, *IEEE Trans. on Auto. Contl.*, Vol. AC-26, No. 2, pp. 570-571.
- Isermann, R. (1980). Practical Aspects of Process Identification, *Automatica*, Vol. 16, No. 5.
- Isermann, R., U. Bauer, W. Bamberger, P. Knepo, and H. Sieberg (1974). Comparison of six online identification and parameter estimation methods, *Automatica*, Vol. 10, pp. 81-103.
- Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*, Academic Press, New York.
- Kailath, T. (1976). *Lectures on linear least-squares estimation*, Springer-Verlag, Wien, West Germany.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *J. Basic Eng. Trans. ASME, Ser. D*, Vol. 82, No. 1, pp. 35-45.
- Kalman, R. E., and R. S. Bucy (1961). New results in linear filtering and prediction theory, *J. Basic Eng., Trans. ASME, Ser. D*, Vol. 83, No. 3, pp. 95-108.
- Kolmogorov, A. N. (1941). Interpolation and extrapolation of stationary random sequences, *Bull. Acad. Sci. USSR Ser. Math.*, Vol. 5; Translation: Rand Corp., Santa Monica, Calif., Memo RM-3090 PR.
- Kormylo, J. (1979). Maximum-likelihood seismic deconvolution, *Ph.D. Thesis*, Dep. of Elect. Eng., University of Southern California, Los Angeles, CA.
- Kozin, F. and H.G. Natke (1986). System identification techniques, *Structural Safety* (special issue), Vol. 3, No. 3&4, pp. 269-316.
- Kullback, S. and R. A. Leibler (1951). On information and sufficiency, *Ann. Math. Statist.*, Vol. 22, pp. 79-86.
- Kushner, H. J. and D. S. Clark (1978). *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, New York.
- Lai, T. L. and C. Z Wei (1982). Asymptotic properties of projections with applications to stochastic regression problems, *J. Multivariate Analysis*, Vol. 12, pp. 346-370.
- Landau, I. D. (1979). *Adaptive Control—The Model Reference Approach*, Marcel Dekker, New York.
- Lee, D.T.L. (1980). Canonical ladder form realizations and fast estimation algorithms, *Ph.D. Thesis*, Dep. of Elect. Eng., Stanford University, Stanford, CA.
- Mendel, J.M. (1983). *Optimal Seismic Deconvolution*, Academic Press, New York, NY.
- Morf, M. (1974). Fast algorithms for multivariable systems, *Ph.D. Thesis*, Dept. of Elect. Eng., Stanford University, Stanford, Ca.
- Ljung, L. (1976). Consistency of the least-squares identification method. *IEEE Trans. Auto. Contl.*, Vol. AC-21, pp. 779-781.
- Ljung, L. (1978). Convergence analysis of parametric identification methods. *IEEE Trans. Auto. Contl.*, Vol. AC-23, pp. 770-783.
- Ljung, L. (1981). Analysis of a general recursive prediction error identification algorithm. *Automatica*, Vol. 17, pp. 89-100.
- Ljung, L. (1985). On the estimation of transfer functions, *Automatica*, vol. 21, pp. 677-696.
- Ljung, L. (1987). *System Identification: Theory for the User*, Prentice Hall, Englewood Cliffs, New Jersey.
- Ljung, L. and Soderstrom, T. (1983). *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, Mass.
- Luenberger, D. G. (1973). *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass.
- Marquardt, D.W. (1963). An Algorithm for Least Squares Estimation of Nonlinear Parameters, *SIAM Jour. Applied Math*, 11, pp. 431-441.
- Nevelson, M. B., and R. E. Z Khasminskii (1973). Stochastic approximation and recurrence estimation, in *Translations of Mathematical Monographs*, Vol. 74, American Mathematical Society, Providence, R.I., Chap. 7.
- Pandit, S.M. and S.M. Wu (1983). *Time Series and System Analysis with Applications*, Wiley, New York.

- Polhemus, N.W. and A.S. Çakmak (1981). Simulation of earthquake ground motions using autoregressive moving average (ARMA) models, *Earthquake Engineering and Structural Dynamics*, Vol. 9, pp.343-354.
- Potter, J.E. (1963). New statistical formulas, *Memo 40*, Instrumentation Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.
- Priestly, M.B. (1980). State dependent models: a general approach to nonlinear time series analysis, *Time Ser. Anal.*, Vol. 1, No. 1, pp. 47-73.
- Rabiner, L.E. and B. Gold (1975). *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ.
- Rake, H. (1980). Step response and frequency response methods, *Automatica*, vol. 16, pp. 519-526.
- Rao, C.R. (1965). *Linear Statistical Inference and Its Applications*, 2nd Ed., Wiley, New York.
- Rissanen, J. (1983). A universal prior for integers and estimation by minimum description length, *Ann. Statist.*, 11, pp. 416-431.
- Robbins, H., and S. Monro (1951). A stochastic approximation method, *Ann. Math. Stat.*, Vol. 22, pp. 400-407.
- Robinson, E.A. (1967). *Multichannel Time Series Analysis with Digital Computer Programs*, Holden-Day, San Francisco, CA.
- Rojahn, C. and P.N. Mork (1982). An analysis of strong-motion data from a severely damaged structure: The Imperial County Services Building, El Centro, California, *U. S. Geological Survey Professional Paper*, 1254, pp. 357-375.
- Rootzen, H. and J. Sternby (1984). Consistency in least-squares estimation—A Bayesian approach. *Automatica*, Vol. 20, pp. 471-777.
- Safak, E. and M. Çelebi (1987). On identification of site amplification from earthquake records, *proceedings, Soc. for Exp. Mech., Spring Conf., June 1987, Houston, Texas*, pp.737-746.
- Safak, E. (1988a). Stochastic system identification in structural dynamics, in *Probabilistic Methods in Civil Engineering* (Ed. by P.D. Spanos), Amer. Soc. Civ. Eng., New York, N.Y.
- Safak, E. (1988c). Optimal-adaptive filters for modeling spectral shape, site amplification, and source scaling, *Soil Dynamics and Earthquake Engineering*, in print.
- Saridis, G N. (1974). Comparison of six on-line identification algorithms, *Automatica*, Vol. 10, pp. 69-79.
- Schwarz, G. (1978). Estimating the dimension of a model, *Ann. Statist.*, Vol. 6, pp. 461-464.
- Shannon, C.E. (1949). Communications in the presence of noise, *proc., IRE* 37, pp. 10-21.
- Söderström, T. (1974). Convergence properties of the generalized least squares identification method. *Automatica*, Vol. 10, pp. 617-626.
- Söderström, T. (1975a). On the uniqueness of maximum likelihood identification, *Automatica*, Vol. 11, pp. 193-197.
- Söderström, T. (1975b). Test of pole-zero cancellation in estimated models. *Automatica*, Vol. 11, pp. 537-541.
- Söderström, T. (1987). Identification: model structure determination, In *Encyclopedia of Systems and Control* (Ed. by M. Singh), Pergamon Press, Oxford, England, pp. 2287-2293.
- Söderström, T. and P. Stoica (1981). Comparison of some instrumental variable methods. Consistency and accuracy aspects. *Automatica*, Vol. 17, pp. 101-115.
- Söderström, T. and P. Stoica (1983). *Instrumental Variable Methods for System Identification*. Lecture Notes in Control and Information Sciences, Springer-Verlag, New York.
- Solo, V. (1981). The second order properties of a time series recursion. *Ann. Statist.*, Vol. 9, pp. 307-317.
- Sorensen, H.W. (1970). Least squares estimation from Gauss to Kalman, *IEEE Spectrum*, 7, pp. 63-68.
- Sorensen, H.W. (1981). *Parameter Estimation: Principle and Problems*, Marcel Dekker, New York.
- Stoica, P., P. Eykhoff, P. Janssen, and T. Söderström (1986). Model-structure selection by cross-validation, *Int. J. Control*, 43, pp. 1841-1878.

- Stoica, P. and A. Nehorai (1987). On the uniqueness of prediction error models for systems with noisy input-output data, *Automatica*, Vol. 23, pp. 541-543.
- Tretter, S.A. (1976). *Introduction to Discrete-Time Signal Processing*, Wiley, New York.
- Tustin, A. (1947). A method of analyzing the behavior of linear systems in terms of time series, *JIEE* (London), 94, pt. IIA, pp. 130-142.
- Tsyppkin, T.Z., E.D. Avedyan, and O.V. Gulinskiy (1981). On convergence of the recursive identification algorithms, *IEEE Trans. Auto. Contl.*, Vol. AC-26, No. 5, pp. 1009-1017.
- Unton, F. Z. (1981). On a Procedure for Testing the Order of Time Series, *IEEE Trans. on Auto. Contl.*, Vol. AC-26, No. 2, pp. 573-575.
- Varlaki, P., G. Terdik, and V. A. Lototsky (1985). Tests for linearity and bilinearity of dynamic systems. *proc. 7th IFAC Symp. on System identification*, York, U.K., Pergamon Press, New York, pp. 731-736.
- van den Boom, A.J.W., and van den Enden, A.W.M. (1974). The determination of the orders of process and noise dynamics, *Automatica*, 10, pp. 244-256.
- Widrow, B., J.R. Glover, J.M. McCool, J. Kaunitz, C.S. Williams, R.H. Hearn, J.R. Zeidler, E. Dong, and R.C. Goodlin (1975). Adaptive noise cancelling: principles and applications, *Proc. IEEE*, Vol. 63, No. 12, pp. 1692-1716.
- Widrow, B. and S. Stearns (1984). *Adaptive Signal Processing*, Prentice- Hall, Englewood Cliffs, N.J.
- Wiener, N. (1949). *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, Wiley, New York.
- Willems, J.C. (1986a). From time series to linear system- Part I. Finite dimensional linear time invariant systems, *Automatica*, Vol. 22, No. 5, pp. 561-580.
- Willems, J.C. (1986b). From time series to linear system- Part II. Exact modelling, *Automatica*, Vol. 22, No. 6, pp. 675-694.
- Willems, J.C. (1987). From time series to linear system- Part III. Approximate modelling, *Automatica*, Vol. 23, No. 1, pp. 87-115.
- Wittenmark, B. (1975). Stochastic Adaptive Control Methods: A Survey, *Int. J. Control*, Vol. 21, No. 5, pp. 705-730.
- Wittenmark, B. and K.J. Åström (1984). Practical Issues in the Implementation of Self-Tuning Control, *Automatica*, Vol. 20, No. 5, pp. 595- 605.
- Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*, Almquist and Wicksell, Uppsala, Sweden.
- Young, P.C. (1978). A general theory of modeling for badly defined systems, in *Modeling, Identification, and Control in Environmental Systems* (Ed. by G.S. Vansteenkiste), North Holland, Amsterdam.
- Young, P.C. (1984). *Recursive Estimation and Time-Series Analysis*, Springer-Verlag, New York.

TABLE 4.1

Discrete-time models for SISO systems

Abbreviations:

AR : autoregressive,

MA : moving average,

X : auxiliary input,

FIR : finite impulse response.

1. Black – Box : $A \cdot y = \frac{B}{F} \cdot x + \frac{C}{D} \cdot e$

2. Box – Jenkins : $y = \frac{B}{F} \cdot x + \frac{C}{D} \cdot e$

3. ARARMAX : $A \cdot y = B \cdot x + \frac{C}{D} \cdot e$

4. ARARX : $A \cdot y = B \cdot x + \frac{1}{D} \cdot e$

5. ARMAX : $A \cdot y = B \cdot x + C \cdot e$

6. ARMA : $A \cdot y = C \cdot e$

7. ARX : $A \cdot y = B \cdot x + e$

8. Output – Error : $y = \frac{B}{F} \cdot x + e$

9. FIR : $y = B \cdot x + e$

TABLE 13.3.1

Filter parameters, pole locations, and the corresponding modal frequency, damping, and the weighting factors for the building with soil-structure interaction in Viña del Mar, Chile.

ARMAX Model with $n_a = 16$, $n_b = 15$, $n_c = 3$

j	a_j	b_j	c_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j\bar{p}_j)$
0	1.000	0.	1.000						
1	-1.458	-0.393	-0.156	0.496	0.000	2.233	1.000	0.000	0.000
2	1.606	0.014	-0.559	0.907	-0.571	1.844	0.168	46.331	41.811
3	-0.363	-0.127	0.597	0.907	0.571	1.844	0.168	46.331	41.811
4	-0.373	-0.387	0.000	0.871	1.486	4.751	0.093	19.856	-17.171
5	0.699	0.263	0.000	0.871	-1.486	4.751	0.093	19.856	-17.171
6	-0.415	-0.370	0.000	0.927	0.739	2.365	0.102	-76.777	118.343
7	0.275	-0.021	0.000	0.927	-0.739	2.365	0.102	-76.777	118.343
8	-0.061	0.096	0.000	0.980	1.216	3.870	0.016	-56.042	38.989
9	0.128	-0.043	0.000	0.980	-1.216	3.870	0.016	-56.042	38.989
10	-0.119	-0.068	0.000	0.986	0.000	0.046	1.000	-644.330	-635.051
11	0.647	0.123	0.000	0.909	-1.174	3.748	0.081	-12.378	-24.829
12	-0.591	0.107	0.000	0.909	1.174	3.748	0.081	-12.378	-24.829
13	0.701	-0.250	0.000	0.996	0.288	0.917	0.014	18.788	7.330
14	-0.120	0.313	0.000	0.996	-0.288	0.917	0.014	18.788	7.330
15	0.028	-0.182	0.000	0.954	0.982	3.129	0.048	24.946	22.124
16	0.188	0.000	0.000	0.954	-0.982	3.129	0.048	24.946	22.124

TABLE 13.4.1

Filter parameters, pole locations, and the corresponding modal frequency, damping, and the weighting factors at $t = 25$ sec. for the building with nonlinear response (Imperial County Services Building).

ARX Model with $n_a = 12$, $n_b = 11$; $t = 25$ sec.

j	a_j	b_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j\bar{p}_j)$
0	1.000	0.000						
1	-5.276	0.008	0.882	0.780	5.027	0.159	23.697	-142.289
2	14.504	-0.038	0.882	-0.780	5.027	0.159	23.697	-142.289
3	-27.766	0.097	0.898	-1.306	8.340	0.082	571.951	-157.141
4	41.108	-0.176	0.898	1.306	8.340	0.082	571.951	-157.141
5	-49.373	0.244	0.894	1.158	7.410	0.096	-457.427	510.994
6	49.237	-0.271	0.894	-1.158	7.410	0.096	-457.427	510.994
7	-41.101	0.243	0.940	0.435	2.796	0.141	130.695	79.285
8	28.560	-0.171	0.940	-0.435	2.796	0.141	130.695	79.285
9	-16.168	0.090	0.930	1.498	9.550	0.048	-236.982	825.928
10	7.110	-0.033	0.930	-1.498	9.550	0.048	-236.982	825.928
11	-2.195	0.007	0.973	0.103	0.677	0.256	-1072.202	1092.045
12	0.362	0.000	0.973	-0.103	0.677	0.256	-1072.202	1092.045

TABLE 13.5.1

Filter parameters, pole locations, and the corresponding modal frequency, damping, and the weighting factors for the building with ambient vibrations.

ARX Model with $n_a = 24$, $n_b = 23$

j	a_j	b_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j\bar{p}_j)$
0	1.000	0.000						
1	0.158	-0.794	0.381	0.000	3.070	1.000	0.000	0.000
2	-0.177	-0.400	0.402	-1.047	4.420	0.657	0.000	0.000
3	-0.435	-0.362	0.402	1.047	4.420	0.657	0.000	0.000
4	0.303	0.125	0.831	0.000	0.590	1.000	5.101	-4.239
5	0.363	0.090	0.756	-1.458	4.726	0.188	0.090	-0.027
6	0.499	-0.017	0.756	1.458	4.726	0.188	0.090	-0.027
7	0.346	-0.443	0.883	-0.471	1.549	0.255	20.812	11.373
8	-0.016	-0.608	0.883	0.471	1.549	0.255	20.812	11.373
9	-0.038	-0.243	0.951	-1.502	4.784	0.033	-108.340	-196.705
10	0.292	-0.227	0.951	1.502	4.784	0.033	-108.340	-196.705
11	0.007	-0.203	0.857	0.906	2.926	0.168	-60.796	41.920
12	-0.043	0.179	0.857	-0.906	2.926	0.168	-60.796	41.920
13	0.112	0.317	0.853	1.230	3.949	0.128	-12.431	4.238
14	0.218	0.082	0.853	-1.230	3.949	0.128	-12.431	4.238
15	-0.061	-0.334	0.917	-0.205	0.709	0.391	63.157	47.888
16	0.028	0.012	0.917	0.205	0.709	0.391	63.157	47.888
17	0.030	-0.190	0.974	-0.969	3.085	0.027	271.213	-154.572
18	-0.021	0.297	0.974	0.969	3.085	0.027	271.213	-154.572
19	-0.097	-0.304	0.991	0.389	1.240	0.023	-571.705	754.051
20	-0.036	-0.728	0.991	-0.389	1.240	0.023	-571.705	754.051
21	-0.116	-0.405	0.929	-0.773	2.473	0.095	24.449	65.973
22	0.003	0.626	0.929	0.773	2.473	0.095	24.449	65.973
23	0.001	-0.547	0.982	0.550	1.752	0.033	773.135	-769.838
24	0.007	0.000	0.982	-0.550	1.752	0.033	773.135	-769.838

TABLE 13.6.1

Filter parameters, pole locations, and the corresponding modal frequency, damping, and the weighting factors for the San Fernando record (record F88S70E).

$n_a = 4, n_b = 3; t = 18 \text{ sec.}$

j	a_j	b_j	c_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j\bar{p}_j)$
0	1.000	0.000	1.000						
1	-2.885	1.209	0.000	0.928	0.997	7.955	0.075	0.830	-0.043
2	3.672	1.079	0.000	0.928	-0.997	7.955	0.075	0.830	-0.043
3	-2.543	1.423	0.000	0.959	0.207	1.684	0.198	0.142	-0.109
4	0.792	0.000	0.000	0.959	-0.207	1.684	0.198	0.142	-0.109

$n_a = 8, n_b = 7; t = 18 \text{ sec.}$

j	a_j	b_j	c_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j\bar{p}_j)$
0	1.000	0.000	1.000						
1	-4.511	0.118	0.000	0.896	0.625	5.052	0.173	0.115	-0.381
2	10.328	0.288	0.000	0.896	-0.625	5.052	0.173	0.115	-0.381
3	-15.661	0.395	0.000	0.943	1.460	11.624	0.040	0.823	5.660
4	17.089	0.218	0.000	0.943	-1.460	11.624	0.040	0.823	5.660
5	-13.711	0.048	0.000	0.926	1.046	8.346	0.073	-1.006	1.493
6	7.907	0.188	0.000	0.926	-1.046	8.346	0.073	-1.006	1.493
7	-3.007	0.591	0.000	0.977	0.184	1.473	0.127	0.271	-0.266
8	0.585	0.000	0.000	0.977	-0.184	1.473	0.127	0.271	-0.266

TABLE 13.7.1

Filter parameters, pole locations, and the corresponding modal frequency, damping, and the weighting factors for the site amplification example (records are from the March 3, 1985 Chilean earthquake).

$n_a = 12, n_b = 11; t = 18 \text{ sec.}$

j	a_j	b_j	c_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j\bar{p}_j)$
0	1.000	0.000	1.000						
1	-4.121	-0.008	0.000	0.928	0.912	7.284	0.082	229.574	80.451
2	9.649	0.030	0.000	0.928	-0.912	7.284	0.082	229.574	80.451
3	-16.660	-0.067	0.000	0.920	1.311	10.453	0.064	1191.369	1052.694
4	23.273	0.115	0.000	0.920	-1.311	10.453	0.064	1191.369	1052.694
5	-27.382	-0.160	0.000	0.944	-1.135	9.044	0.050	436.193	1222.136
6	27.639	0.184	0.000	0.944	1.135	9.044	0.050	436.193	1222.136
7	-24.082	-0.180	0.000	0.948	0.583	4.662	0.091	-262.090	106.185
8	17.982	0.148	0.000	0.948	-0.583	4.662	0.091	-262.090	106.185
9	-11.280	-0.097	0.000	0.927	-1.484	11.826	0.051	-161.601	-557.945
10	5.708	0.046	0.000	0.927	1.484	11.826	0.051	-161.601	-557.945
11	-2.130	-0.015	0.000	0.960	0.183	1.492	0.218	-77.697	129.498
12	0.462	0.000	0.000	0.960	-0.183	1.492	0.218	-77.697	129.498

$n_a = 16, n_b = 15; t = 18 \text{ sec.}$

j	a_j	b_j	c_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j\bar{p}_j)$
0	1.000	0.000	1.000						
1	-4.523	-0.001	0.000	0.933	0.731	5.842	0.094	-171.195	232.077
2	11.793	0.000	0.000	0.933	-0.731	5.842	0.094	-171.195	232.077
3	-22.980	0.006	0.000	0.943	-1.513	12.045	0.039	-2345.374	-2339.097
4	36.880	-0.024	0.000	0.943	1.513	12.045	0.039	-2345.374	-2339.097
5	-51.014	0.056	0.000	0.950	-1.233	9.822	0.042	-1912.515	-1970.833
6	62.412	-0.101	0.000	0.950	1.233	9.822	0.042	-1912.515	-1970.833
7	-68.565	0.147	0.000	0.944	0.502	4.024	0.113	-162.119	44.324
8	68.156	-0.181	0.000	0.944	-0.502	4.024	0.113	-162.119	44.324
9	-61.456	0.194	0.000	0.943	1.346	10.719	0.044	763.469	-100.972
10	50.148	-0.183	0.000	0.943	-1.346	10.719	0.044	763.469	-100.972
11	-36.746	0.147	0.000	0.947	-0.934	7.446	0.058	1321.034	-388.209
12	23.817	-0.102	0.000	0.947	0.934	7.446	0.058	1321.034	-388.209
13	-13.295	0.057	0.000	0.939	1.016	8.099	0.062	-45.058	-712.433
14	6.113	-0.026	0.000	0.939	-1.016	8.099	0.062	-45.058	-712.433
15	-2.096	0.007	0.000	0.980	0.168	1.347	0.119	-6.407	107.599
16	0.421	0.000	0.000	0.980	-0.168	1.347	0.119	-6.407	107.599

TABLE 13.8.1

Filter parameters, pole locations, and the corresponding modal frequency, damping, and the weighting factors for the source amplification example (records are from the 1983 Borah Peak, Idaho earthquake).

$n_a = 6, n_b = 5; t = 10 \text{ sec.}$

j	a_j	b_j	c_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j \bar{p}_j)$
0	1.000	0.000	1.000						
1	-2.695	-0.309	0.000	0.925	-1.499	11.948	0.052	8.161	12.166
2	3.971	0.664	0.000	0.925	1.499	11.948	0.052	8.161	12.166
3	-4.311	-0.660	0.000	0.885	1.010	8.100	0.120	-8.031	1.894
4	3.403	0.520	0.000	0.885	-1.010	8.100	0.120	-8.031	1.894
5	-1.916	0.037	0.000	0.966	0.215	1.734	0.161	5.597	-5.485
6	0.624	0.000	0.000	0.966	-0.215	1.734	0.161	5.597	-5.485

$n_a = 14, n_b = 13; t = 10 \text{ sec.}$

j	a_j	b_j	c_j	r_j	ϕ_j	f_{0j}	ξ_{0j}	$2\Re(q_j)$	$2\Re(q_j \bar{p}_j)$
0	1.000	0.000	1.000						
1	-3.996	-0.060	0.000	0.923	1.212	9.666	0.066	27.828	74.174
2	9.322	0.339	0.000	0.923	-1.212	9.666	0.066	27.828	74.174
3	-16.628	-0.474	0.000	0.941	-1.001	7.984	0.060	23.005	-37.370
4	24.445	0.420	0.000	0.941	1.001	7.984	0.060	23.005	-37.370
5	-30.897	0.113	0.000	0.952	-1.336	10.636	0.037	-35.395	-18.397
6	34.337	-0.823	0.000	0.952	1.336	10.636	0.037	-35.395	-18.397
7	-33.883	1.646	0.000	0.883	0.400	3.337	0.297	-1.697	1.850
8	29.846	-2.223	0.000	0.883	-0.400	3.337	0.297	-1.697	1.850
9	-23.343	2.382	0.000	0.909	0.851	6.815	0.111	-8.318	1.027
10	16.023	-2.100	0.000	0.909	-0.851	6.815	0.111	-8.318	1.027
11	-9.398	1.531	0.000	0.907	1.525	12.160	0.064	32.199	10.071
12	4.507	-0.767	0.000	0.907	-1.525	12.160	0.064	32.199	10.071
13	-1.623	0.312	0.000	0.967	0.193	1.556	0.173	-1.820	0.596
14	0.339	0.000	0.000	0.967	-0.193	1.556	0.173	-1.820	0.596

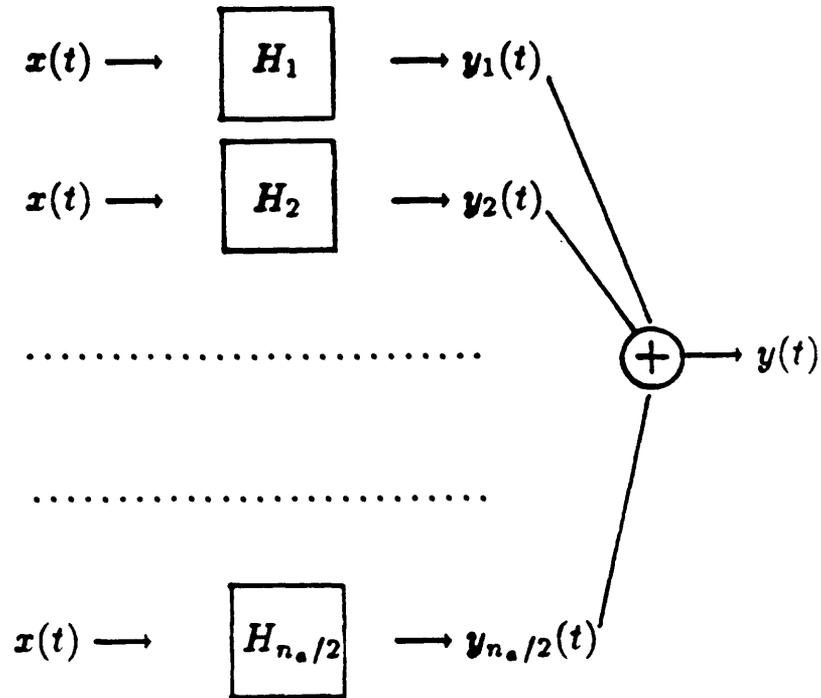


FIG. 3.1- Schematic of parallel form realization of SISO systems.

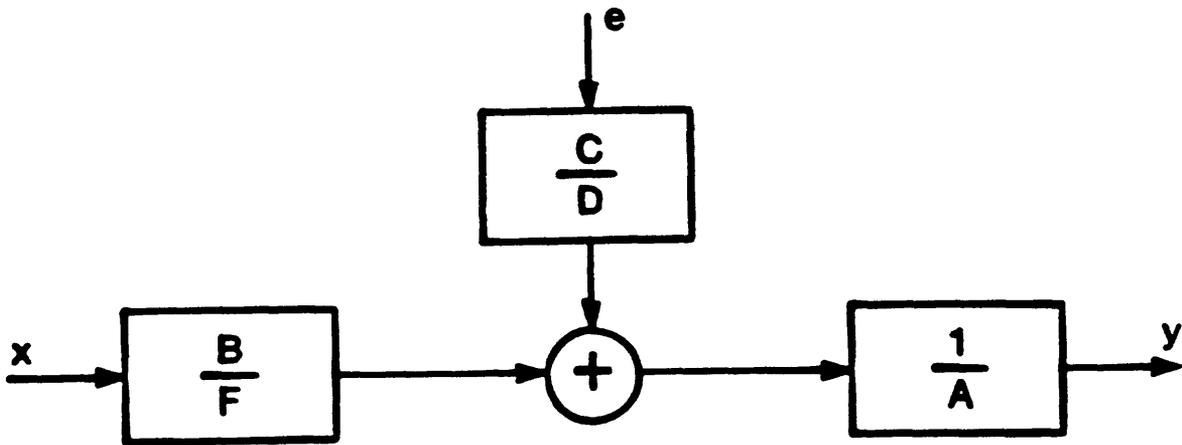


FIG. 4.1- Schematic of black-box models for SISO systems.

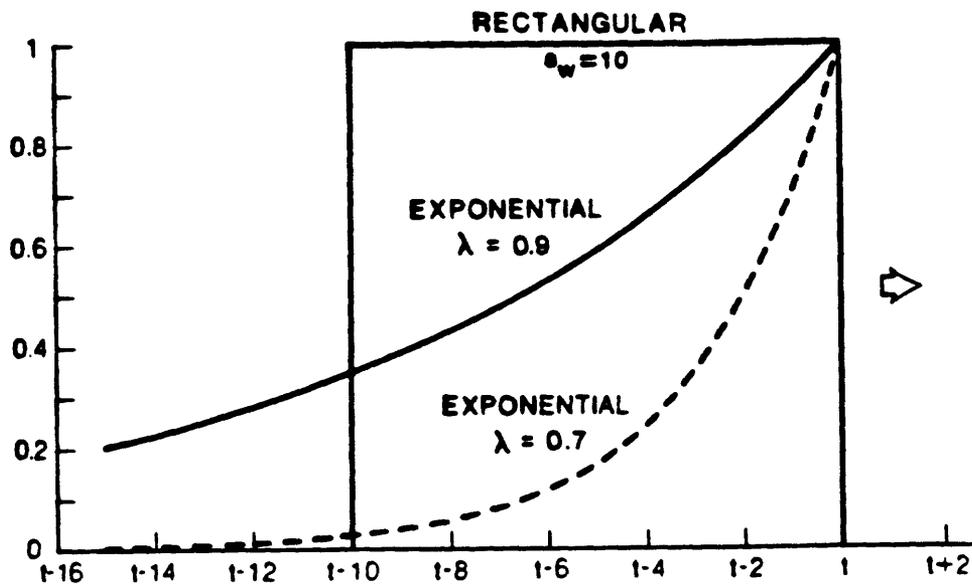


FIG. 5.1- Three sample windows for weighting factors.

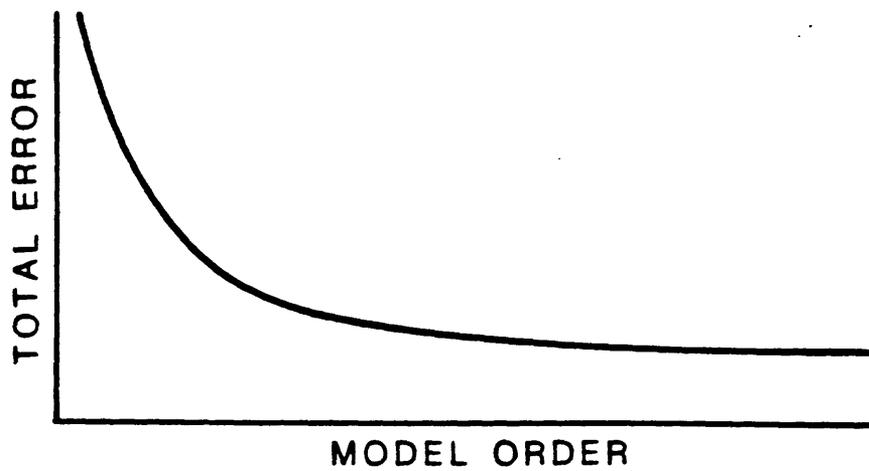


FIG. 8.1- Variation of total estimation error with model order.

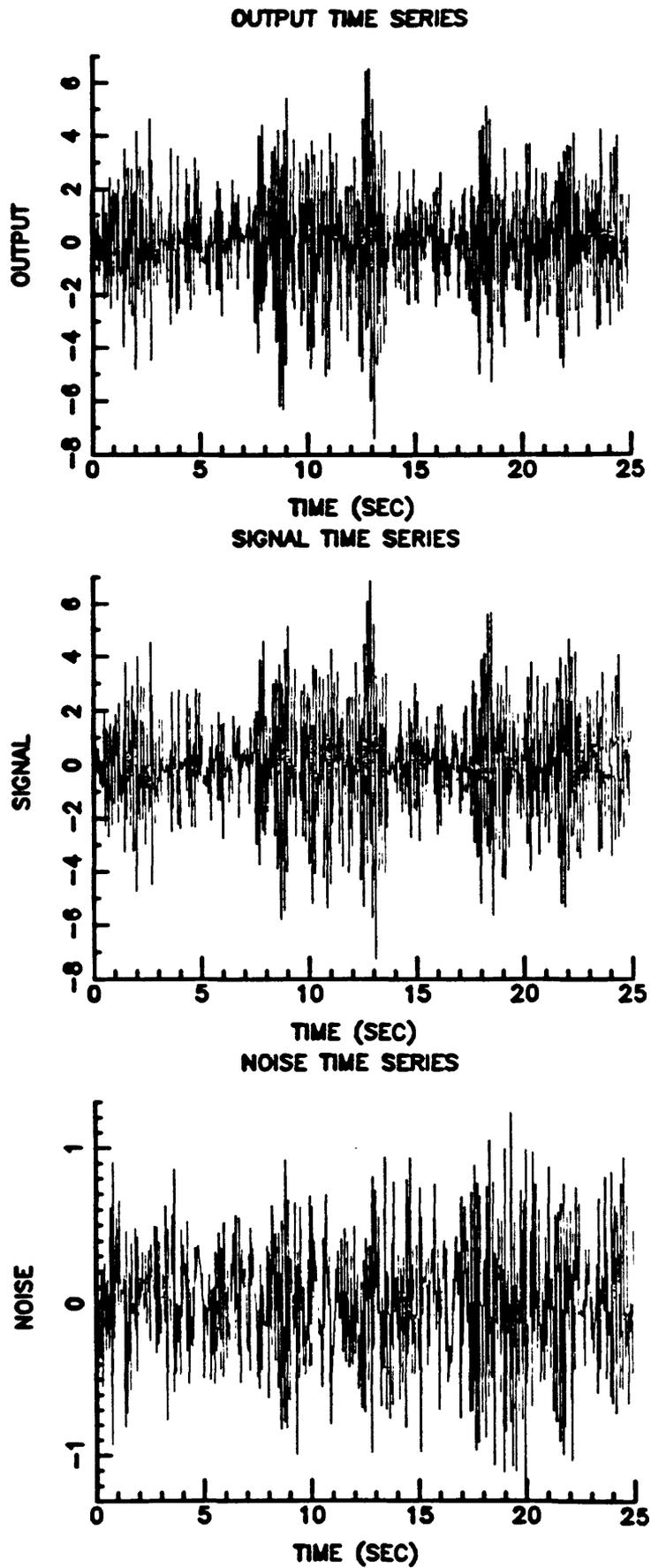


FIG. 13.1.1- Simulated output, signal, and noise time series (output is equal to signal plus noise).

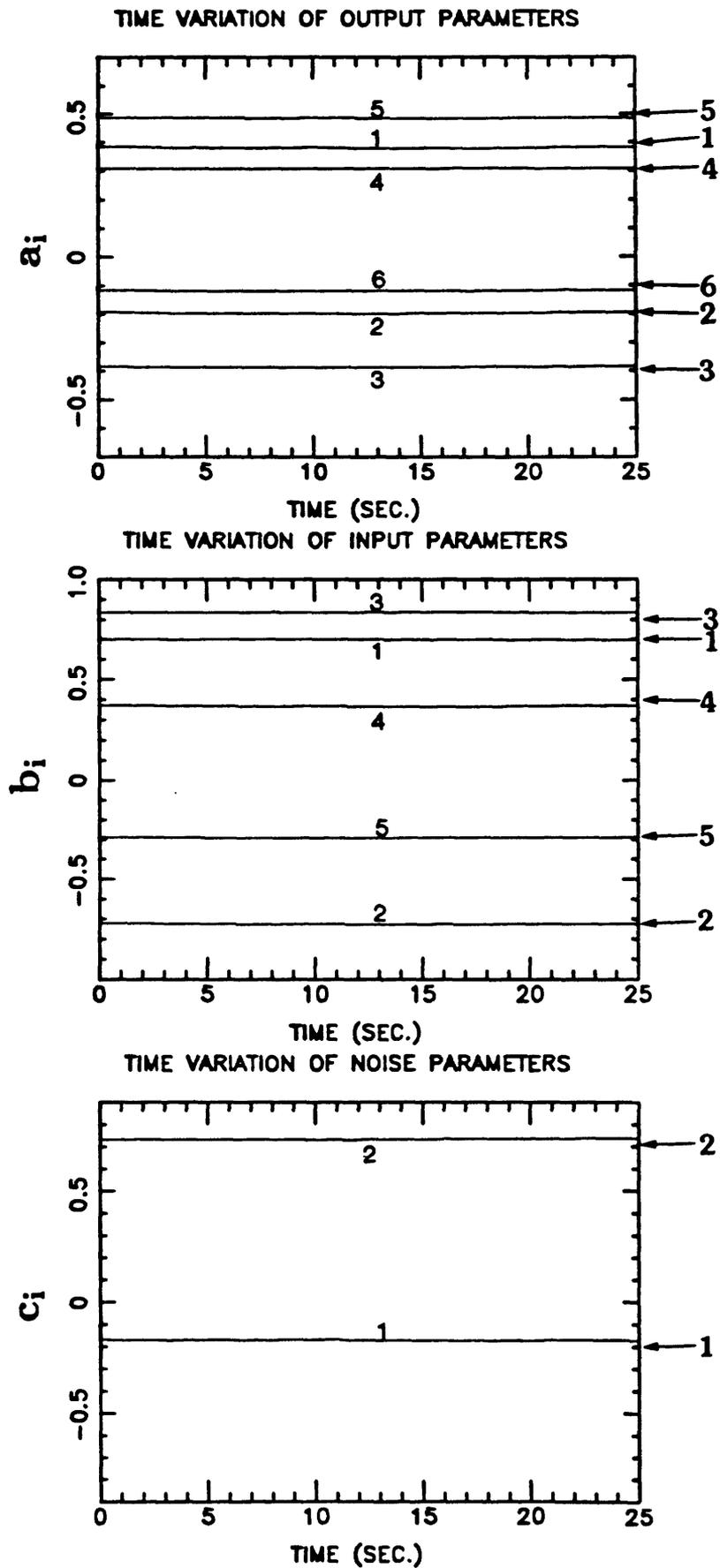


FIG. 13.1.2- Estimated model parameters for $\lambda_0 = \lambda(0) = 1$ and $SNR = 5$, and comparison with exact values (marked by arrows on the right margin).

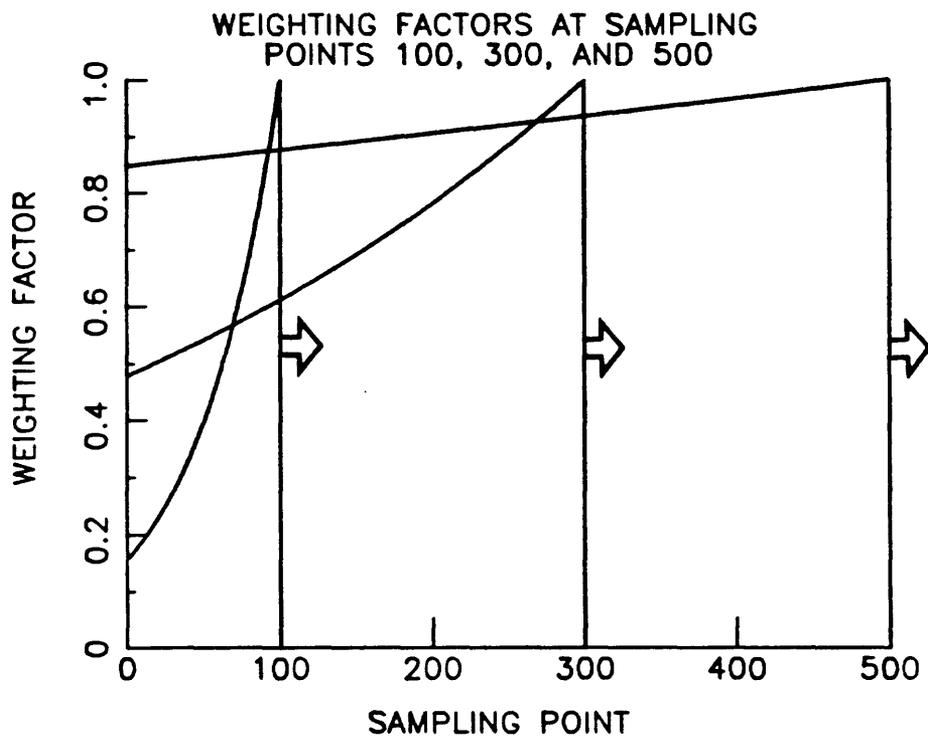


FIG. 13.1.3- Profile of weighting factors for $\lambda_0 = 0.99$ and $\lambda(0) = 0.95$ at sampling points 100, 300, and 500.

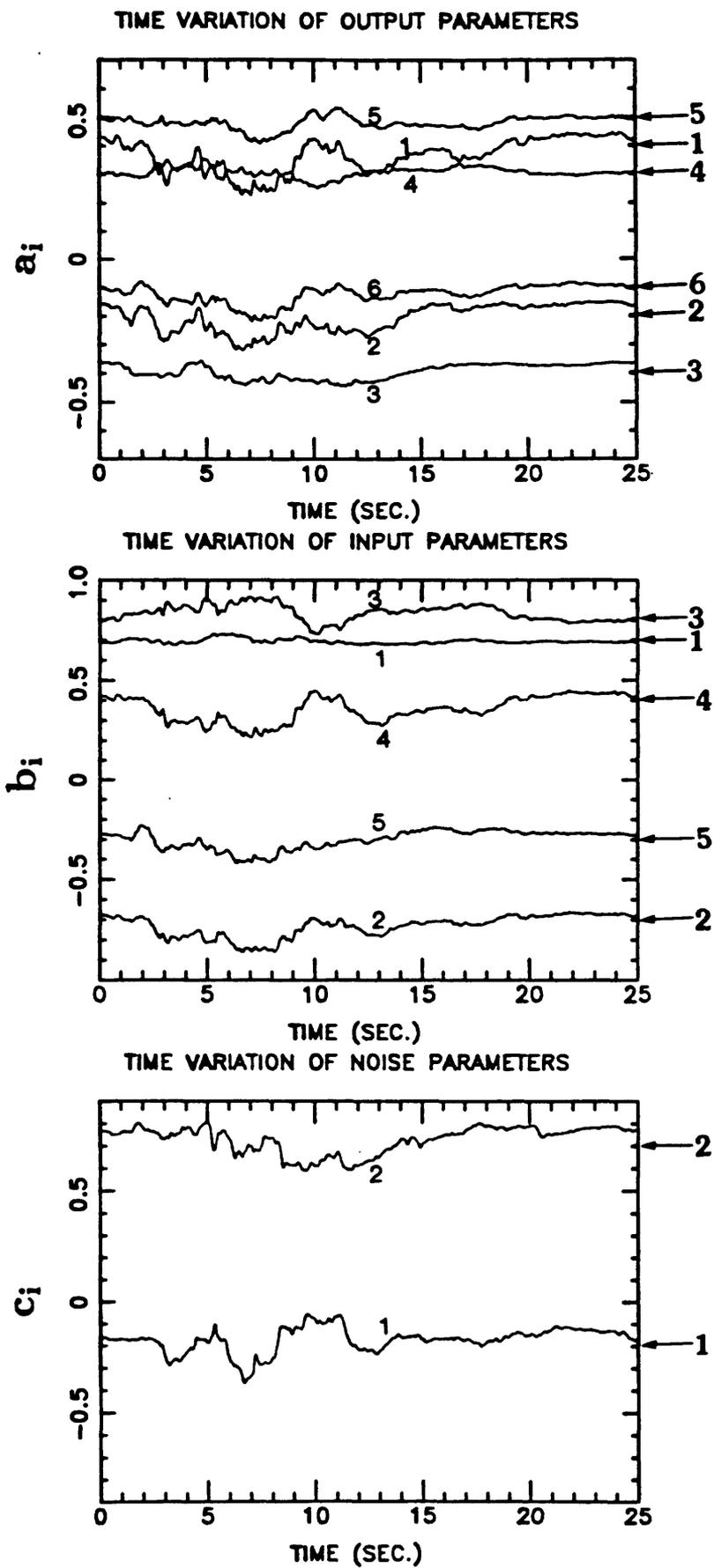


FIG. 13.1.4- Estimated model parameters for $\lambda_0 = 0.99$, $\lambda(0) = 0.95$ and $SNR = 5$, and comparison with exact values (marked by arrows on the right margin).

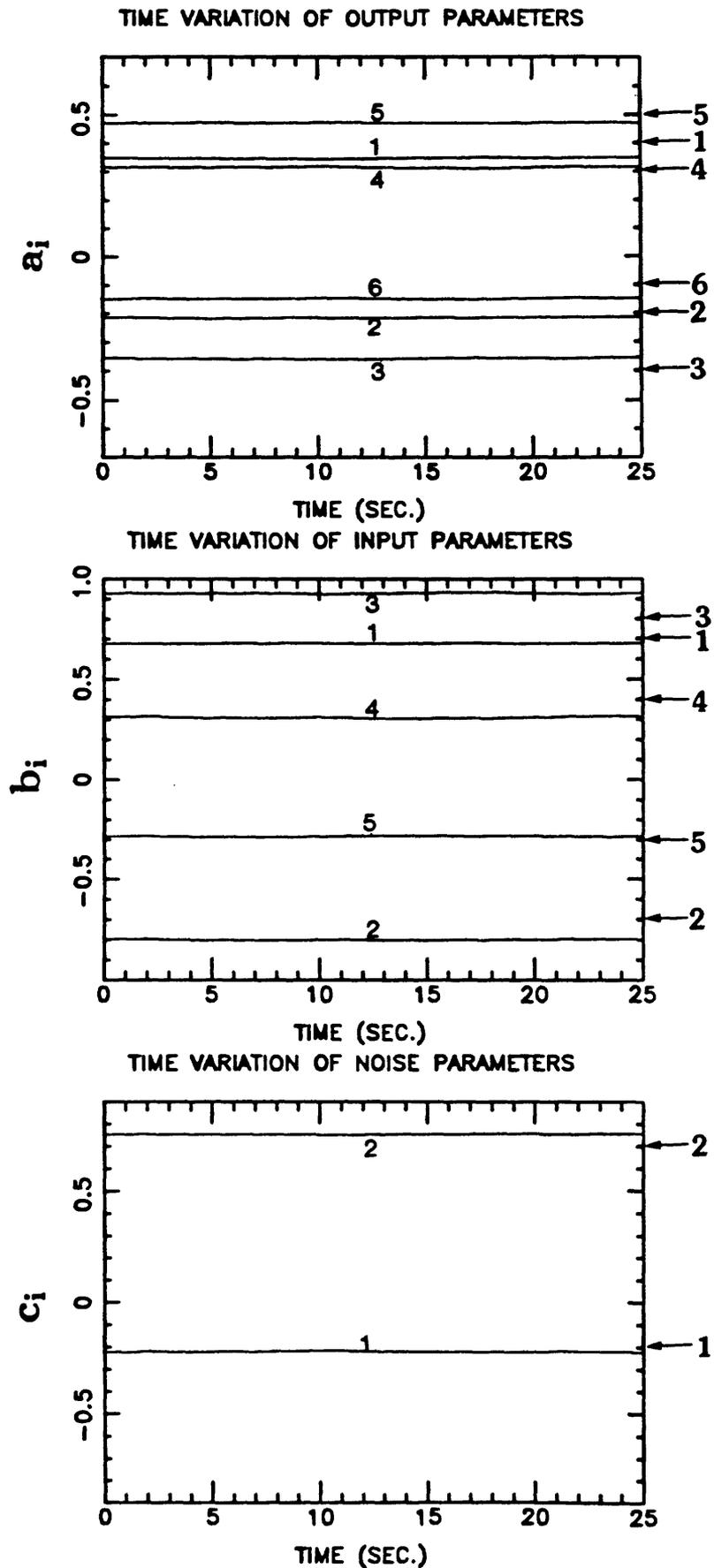


FIG. 13.1.5- Estimated model parameters for $\lambda_0 = \lambda(0) = 1$ and $SNR = 1$, and comparison with exact values (marked by arrows on the right margin).

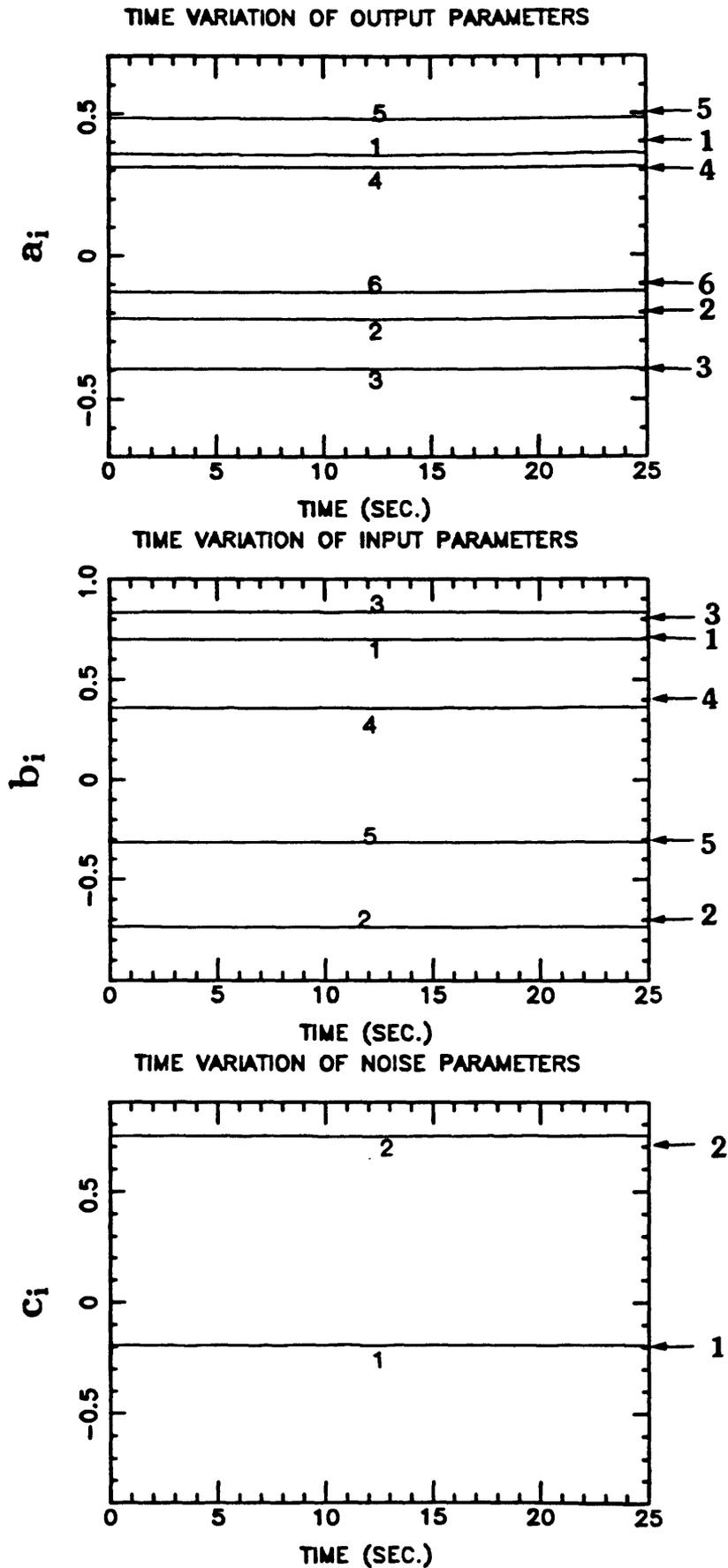


FIG. 13.1.6- Estimated model parameters for $\lambda_0 = \lambda(0) = 1$ and $SNR = 10$, and comparison with exact values (marked by arrows on the right margin).

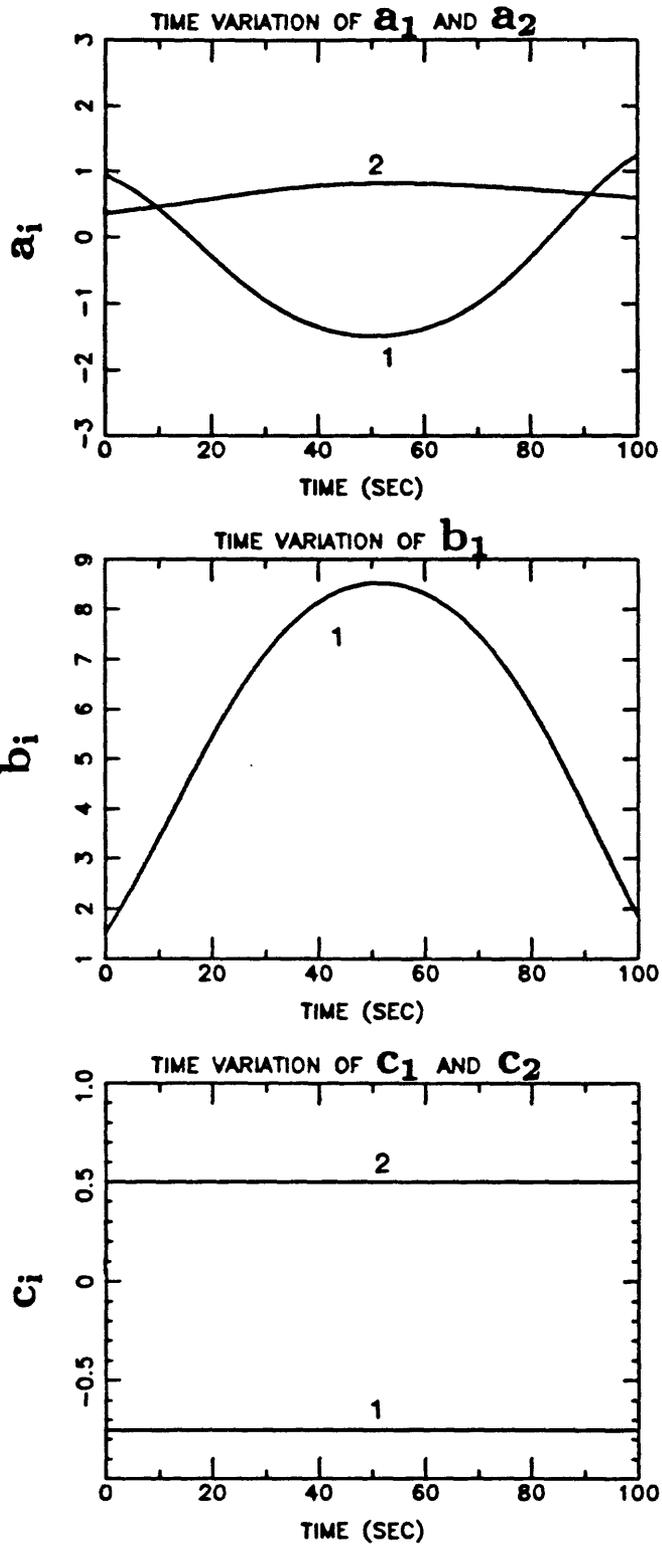


FIG. 13.2.1- Selected time variation of model parameters.

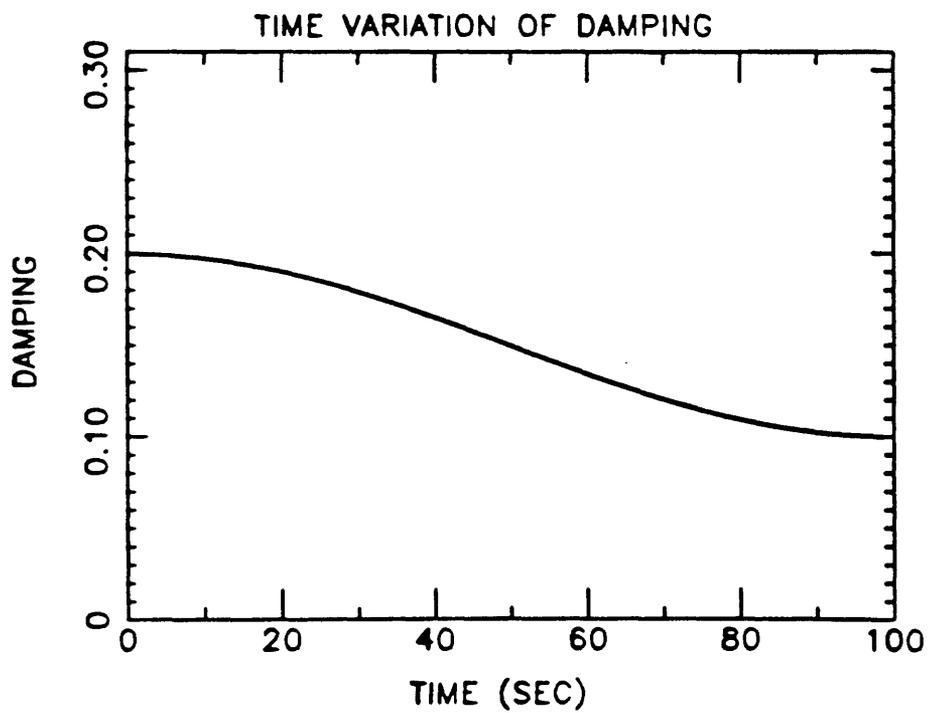
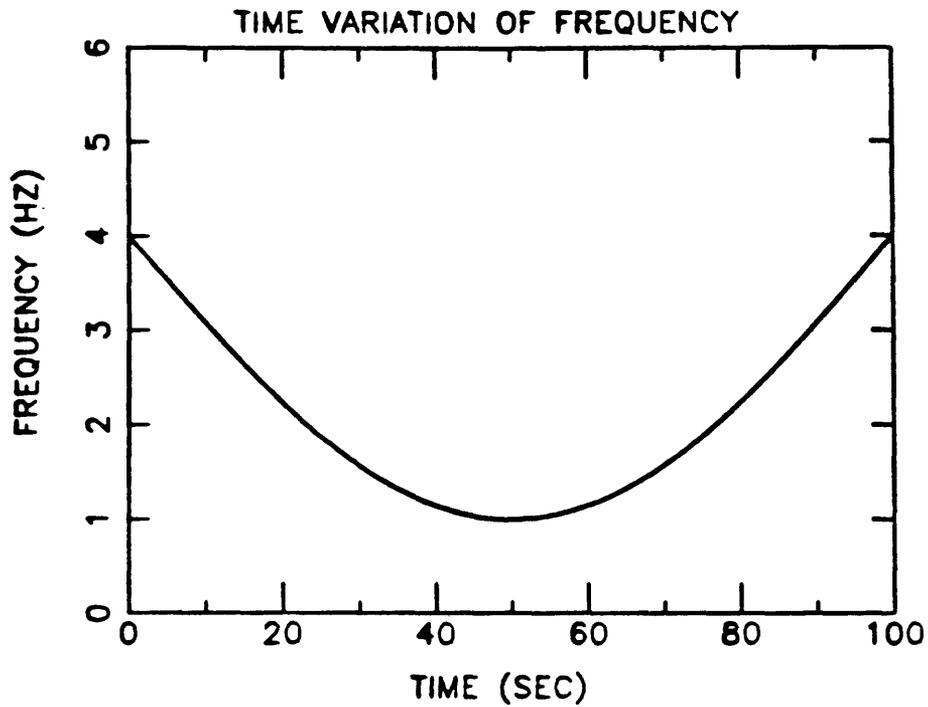


FIG. 13.2.2- Time variations of frequency and damping ratio of the corresponding oscillator.

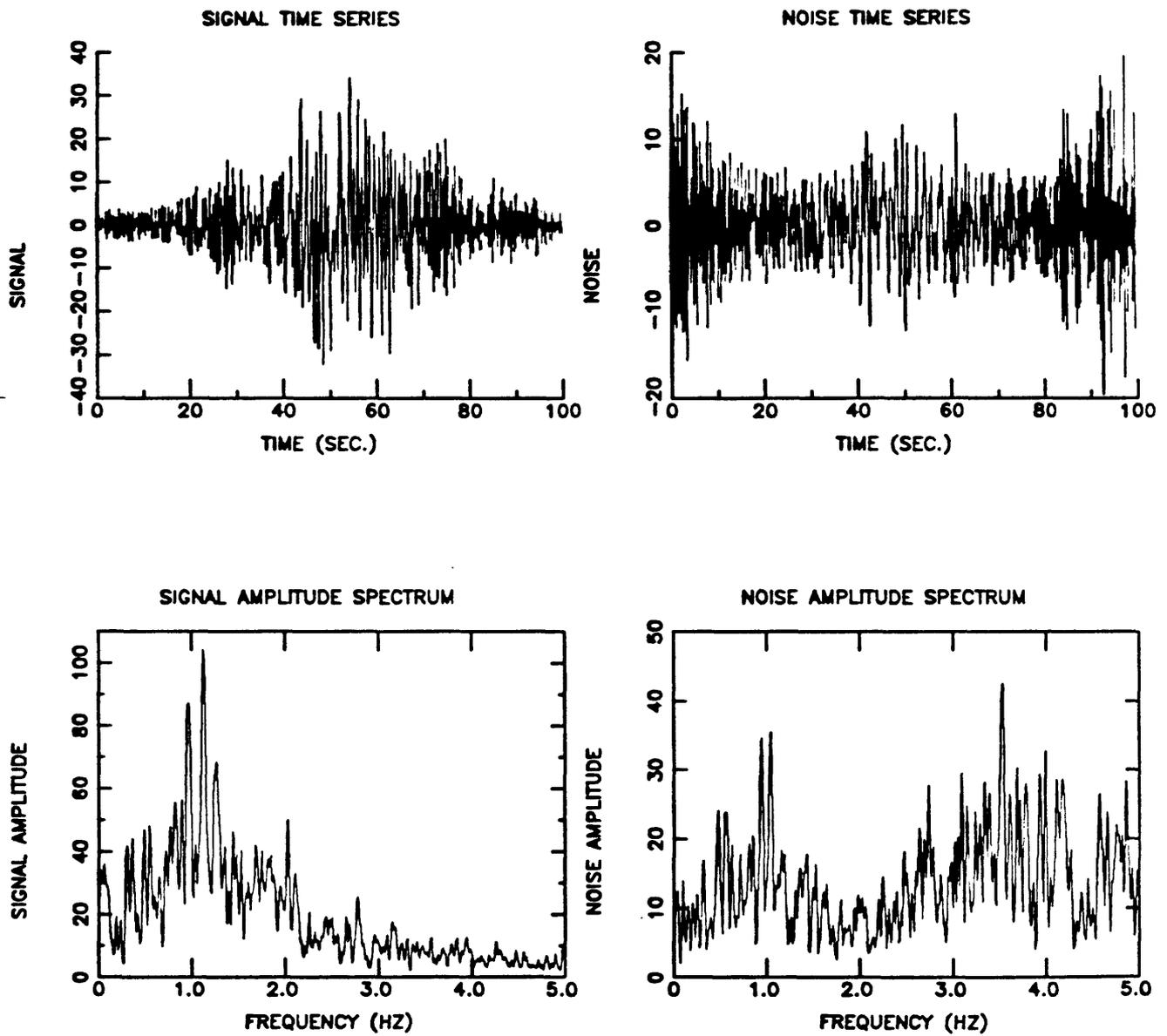


FIG. 13.2.3- Simulated signal and noise series and their Fourier amplitude spectra.

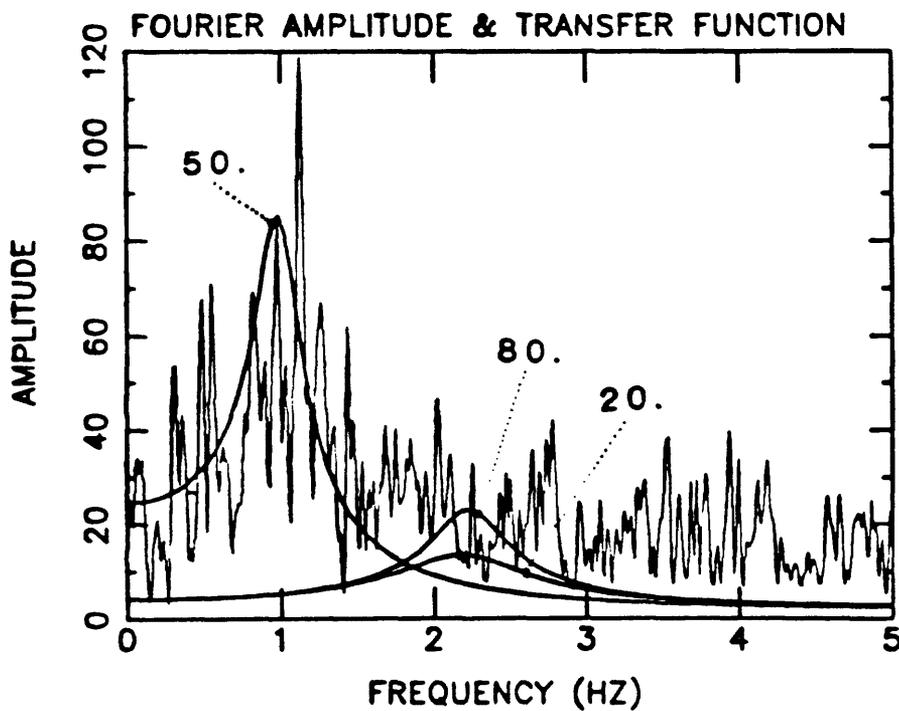
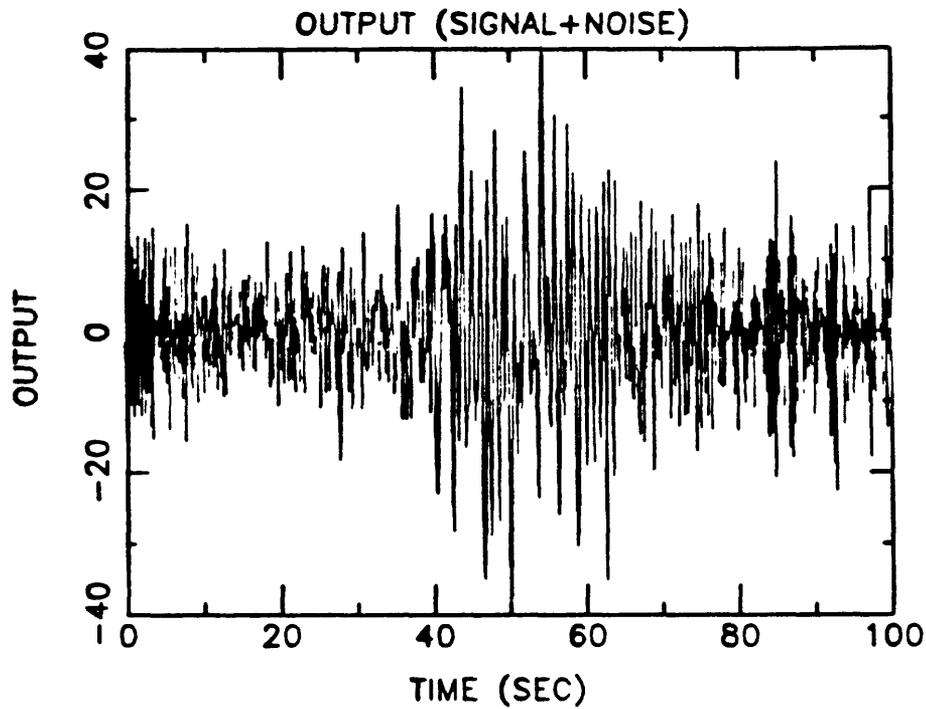


FIG. 13.2.4- Simulated output (signal plus noise) series, output Fourier amplitude spectrum, and comparison with exact transfer function calculated at $t = 20, 50,$ and 80 seconds.

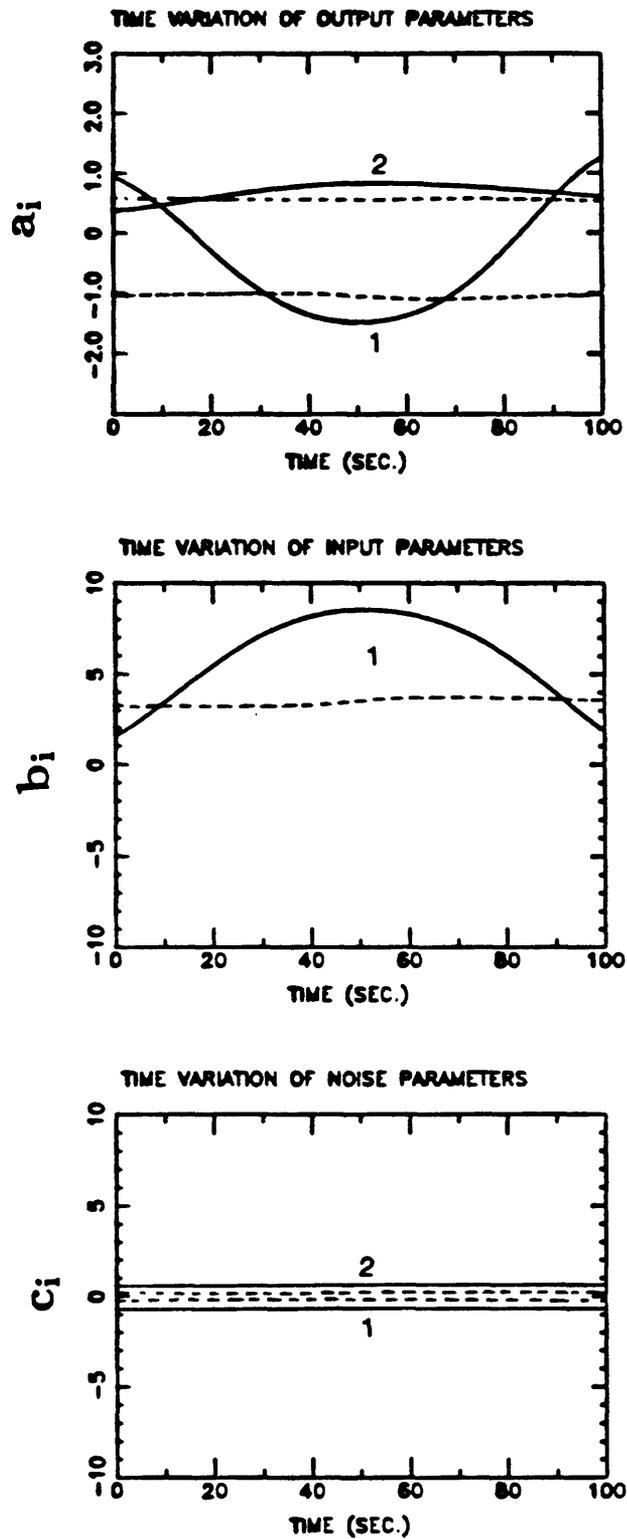


FIG. 13.2.5- Estimated model parameters for $\lambda_0 = \lambda(0) = 1$ (dashed lines), and comparison with exact values (solid lines).

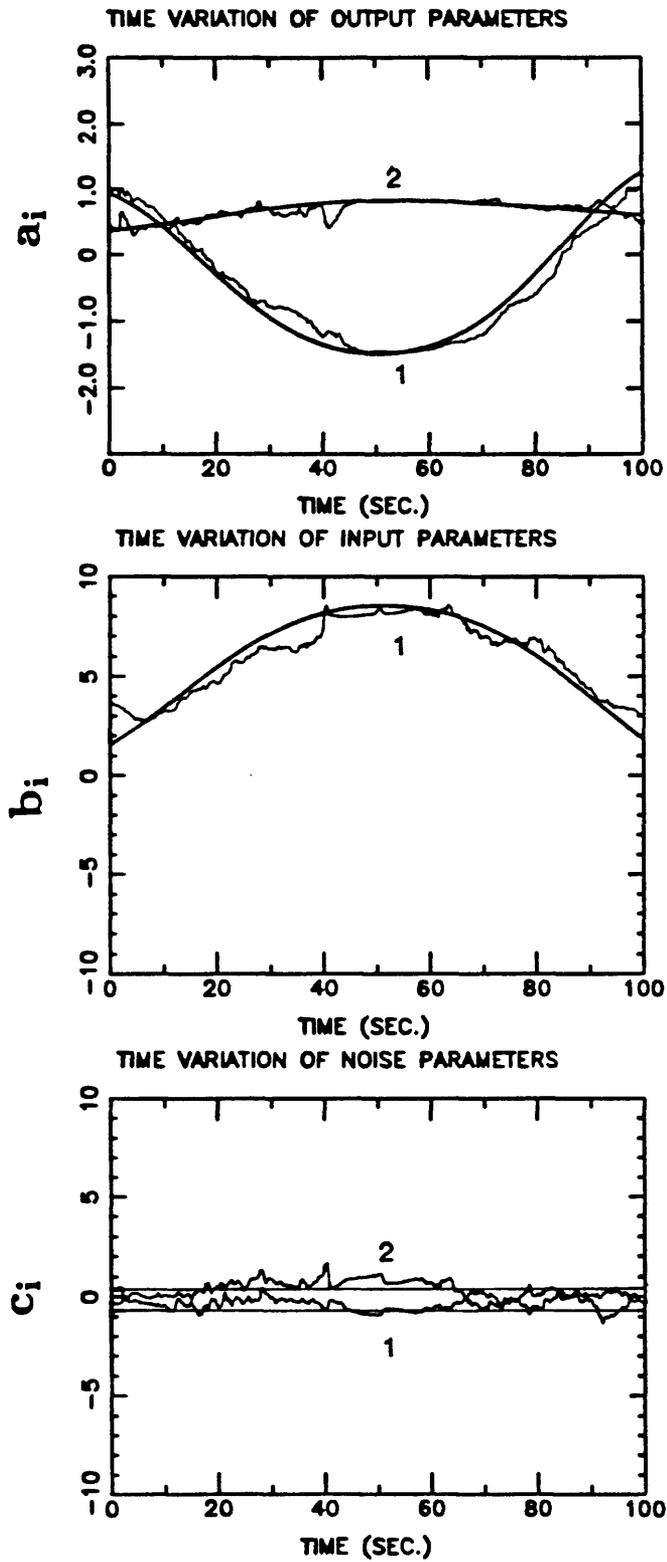


FIG. 13.2.6- Estimated model parameters for $\lambda_0 = 0.99$, $\lambda(0) = 0.99$ (jagged lines), and comparison with exact values (smooth lines).

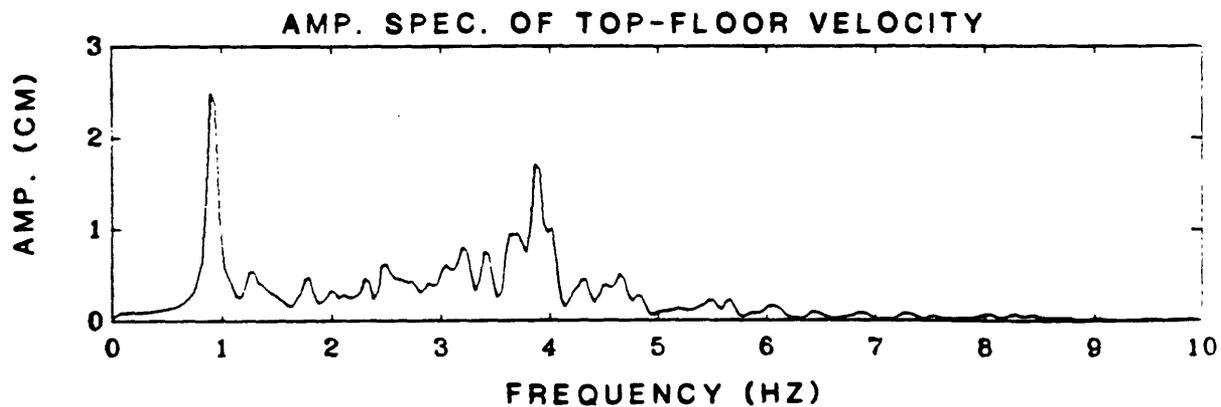
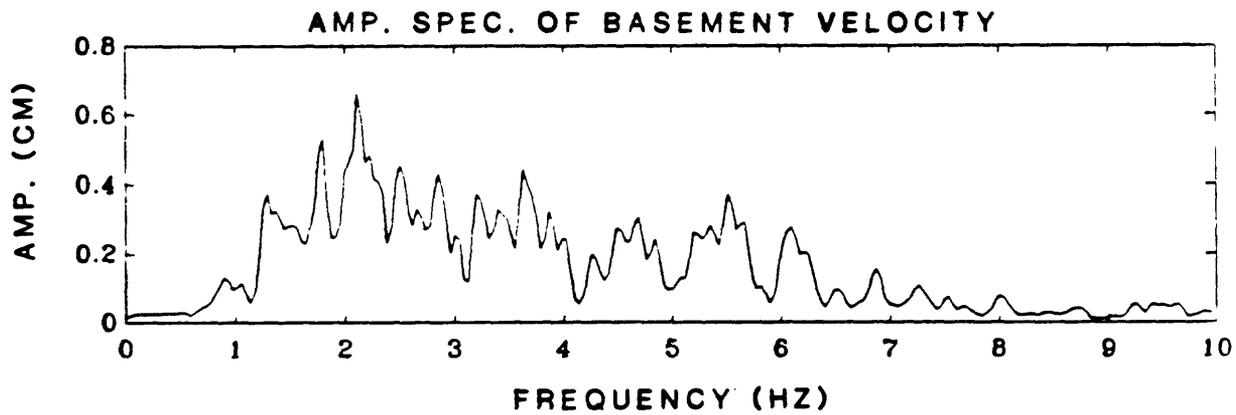
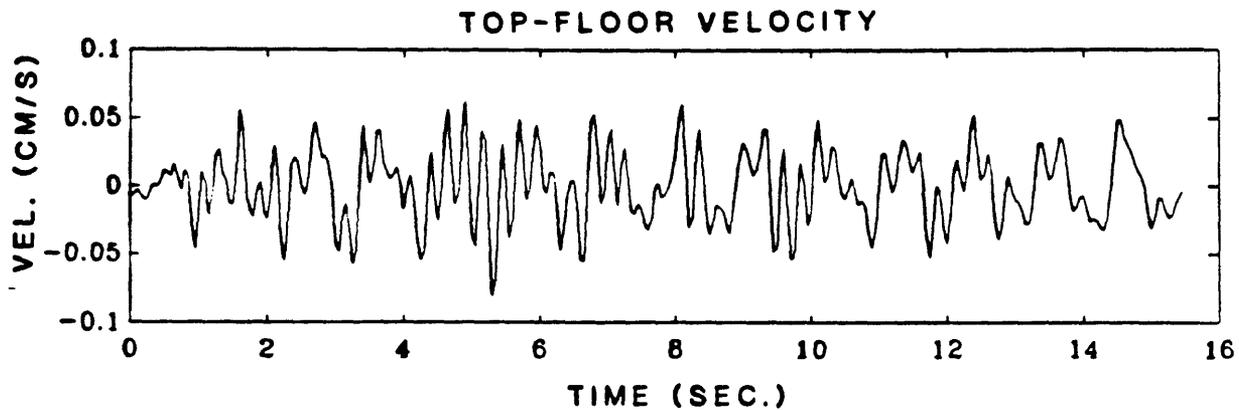
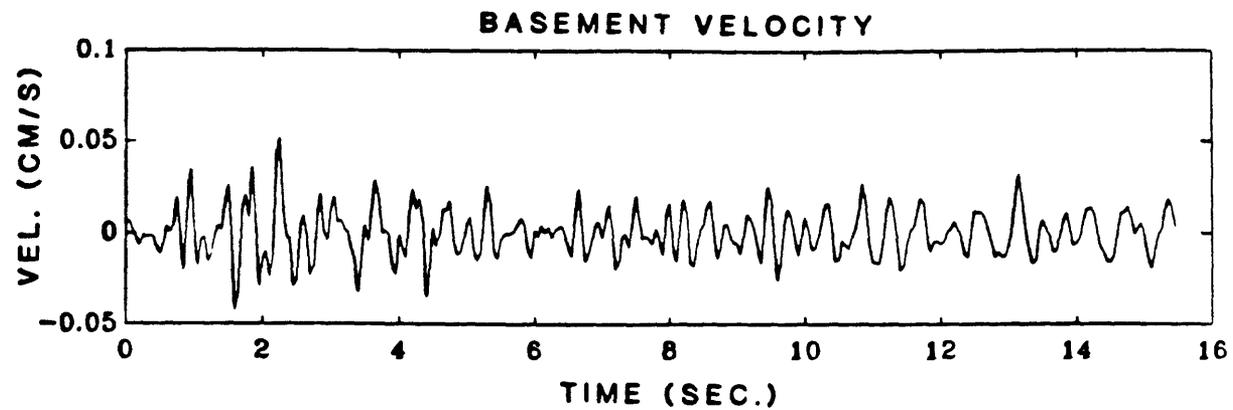


FIG. 13.3.1- Recorded basement and top-floor velocities, and their Fourier amplitude spectra.

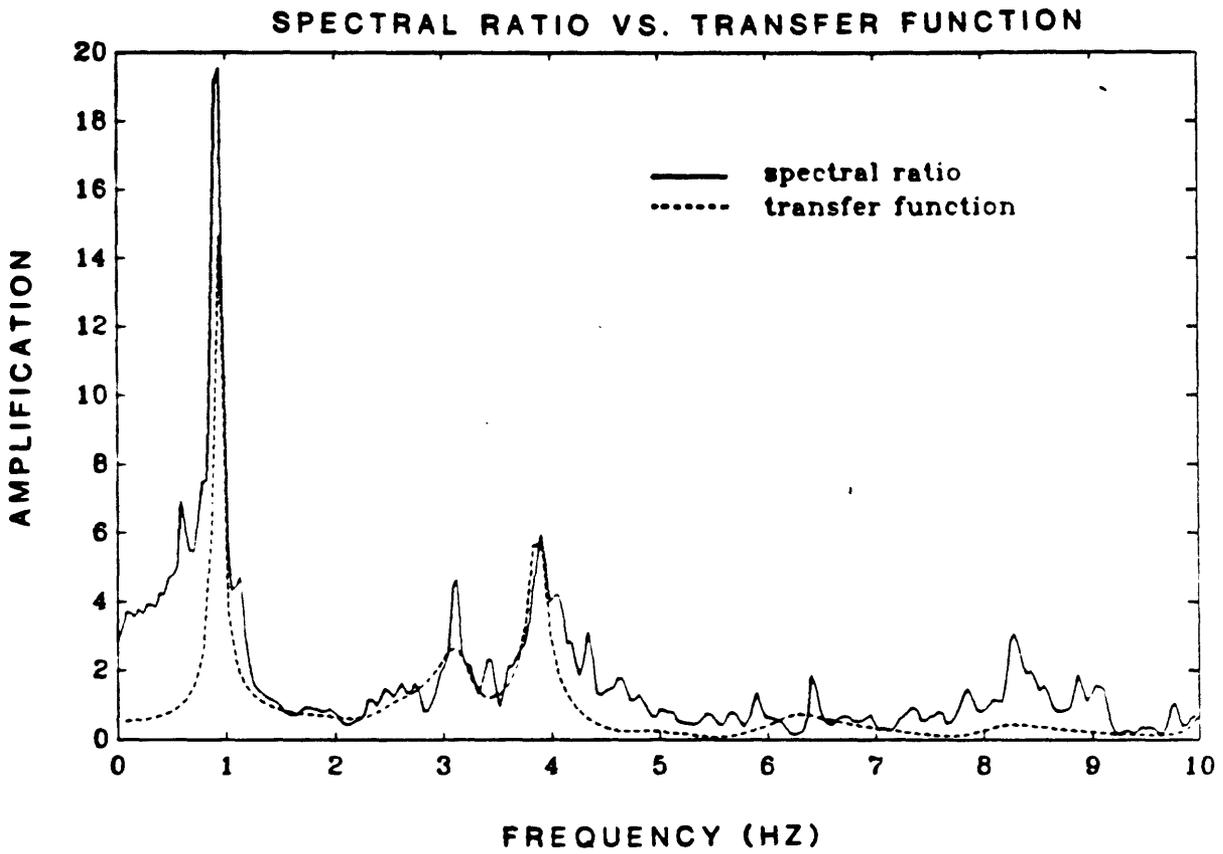


FIG. 13.3.2- Comparison of estimated transfer function with spectral ratio.

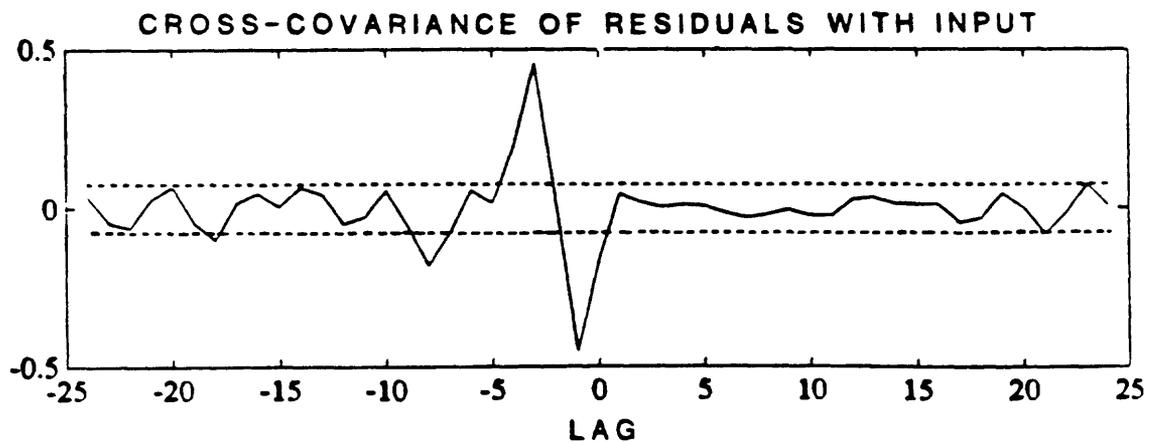
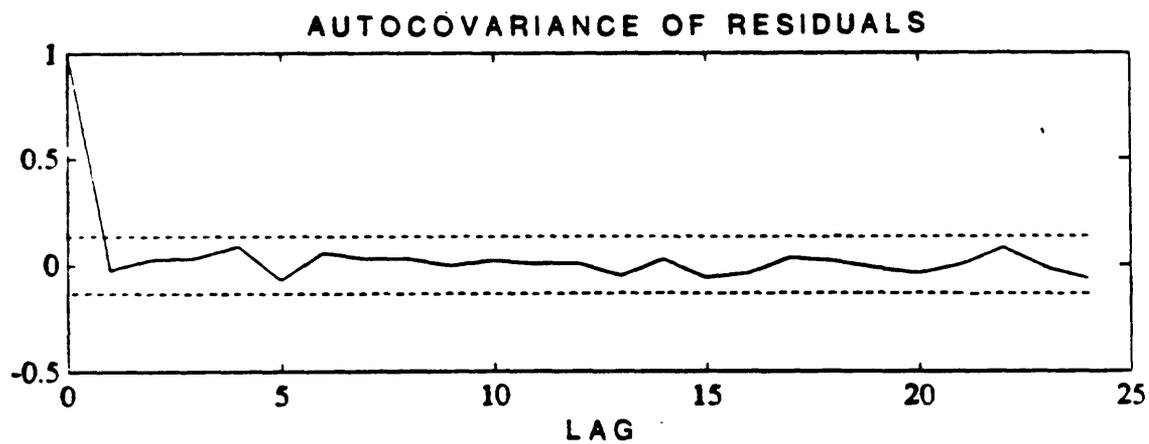
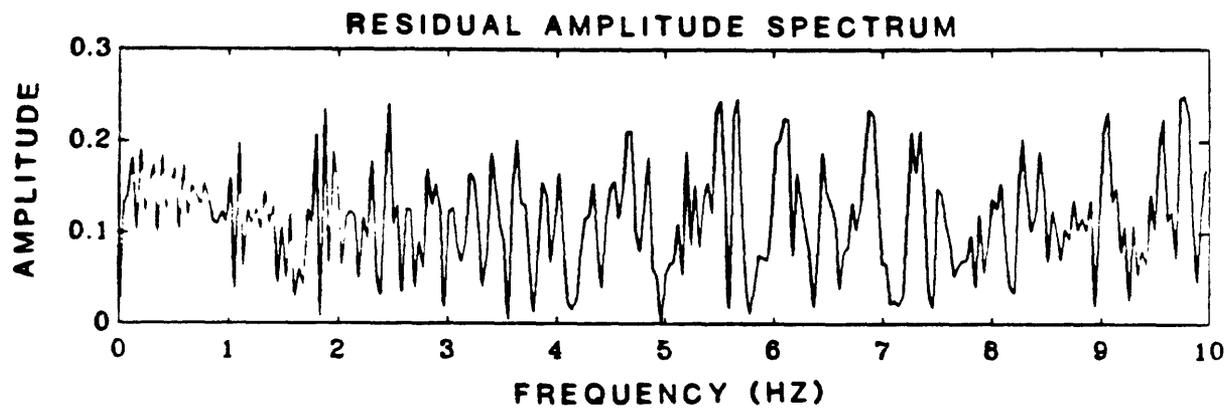
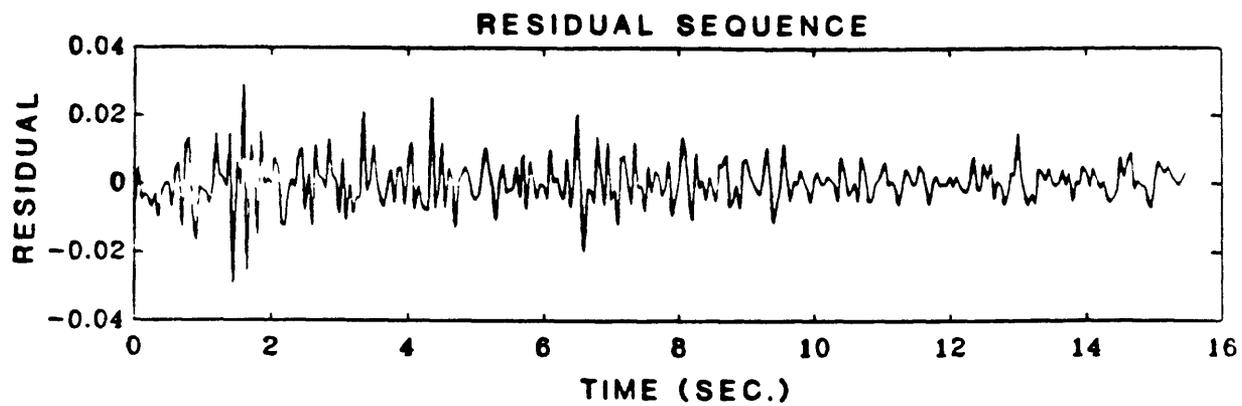


FIG. 13.3.3- Residual time series, Fourier amplitude spectrum, autocovariance, and cross-covariance with input (straight lines in autocovariance and cross-covariance plots show 90-percent confidence levels for whiteness and independence, respectively).

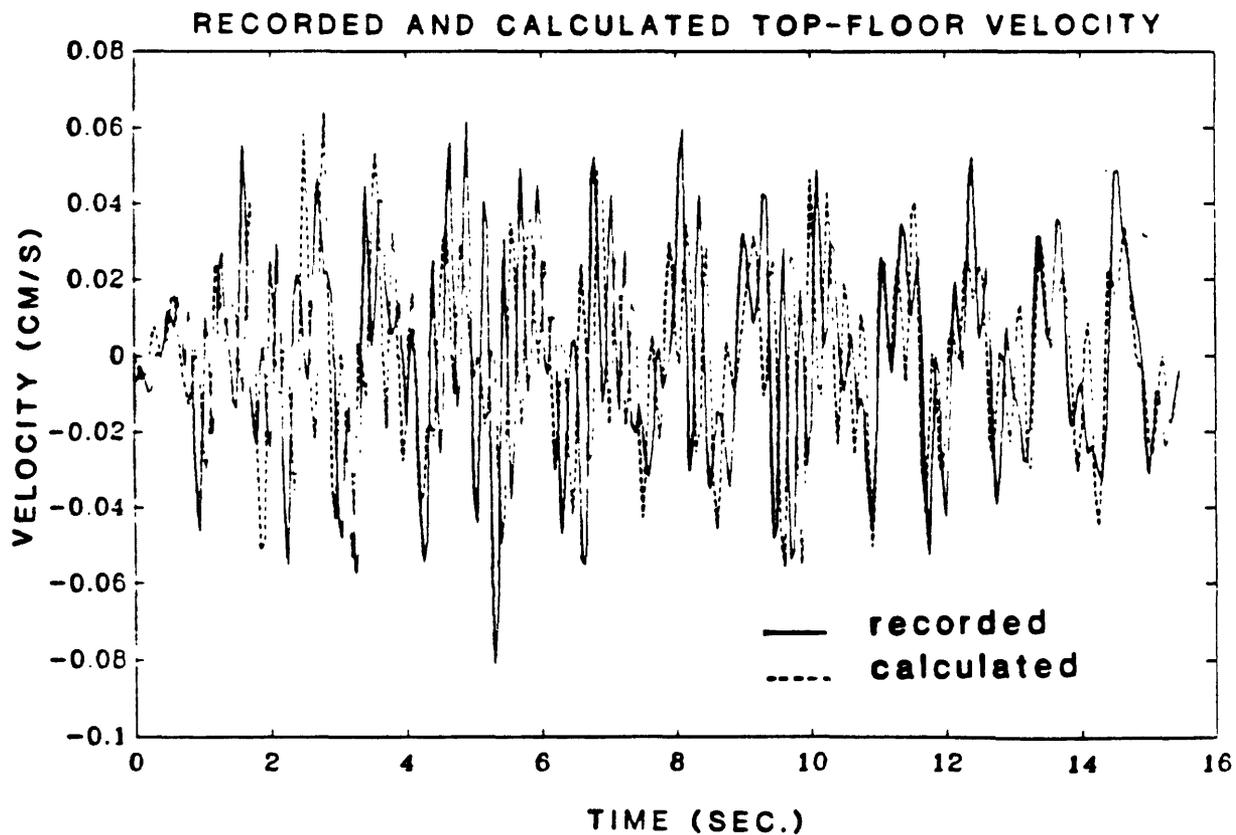
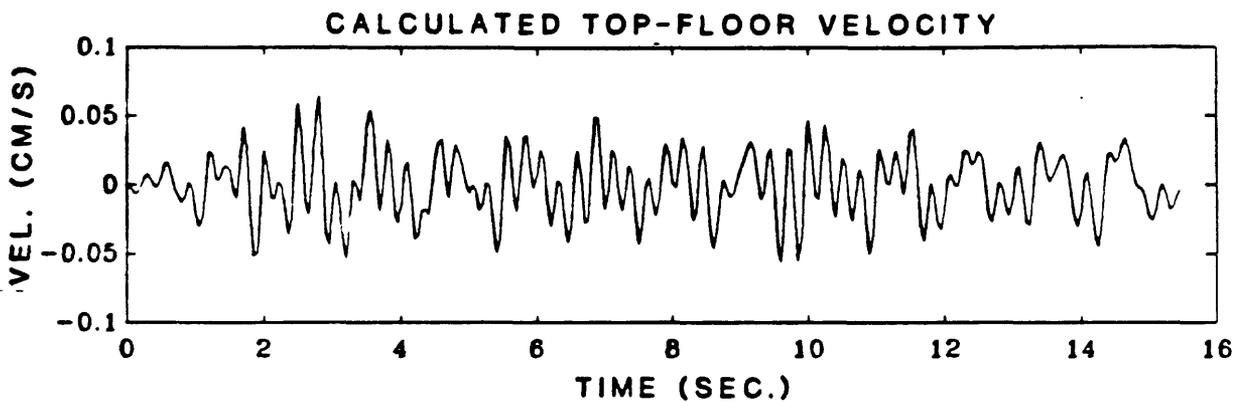
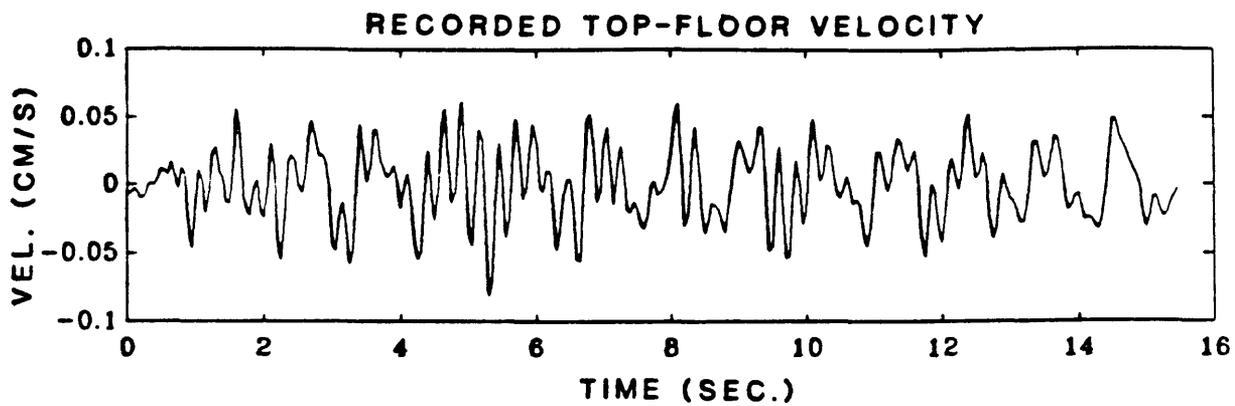


FIG. 13.3.4- Comparison of calculated top-floor velocity with recorded velocity (plotted separately and together).

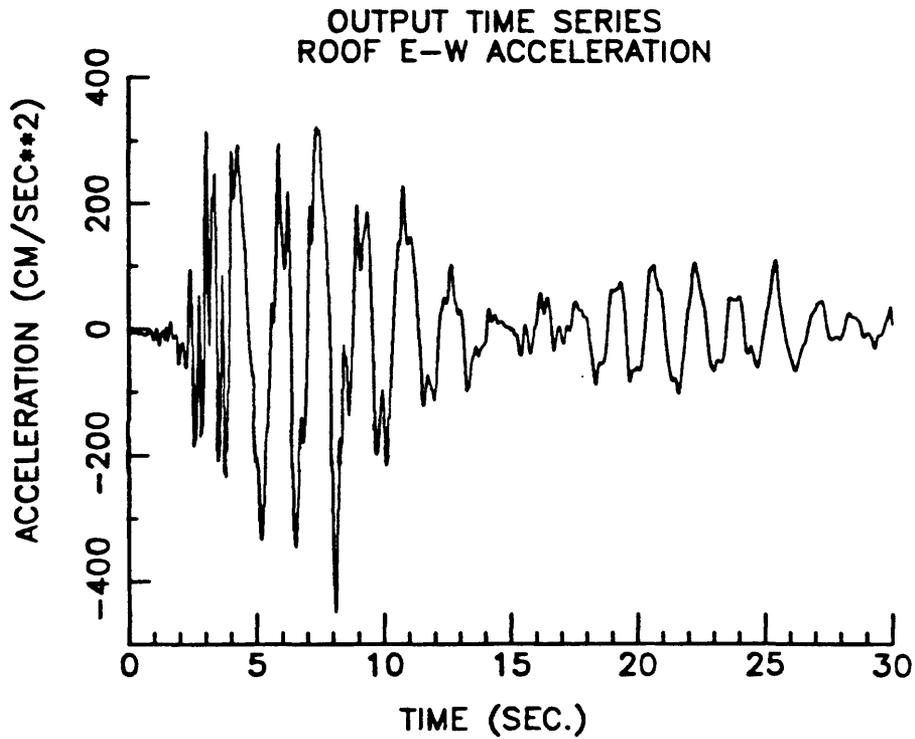
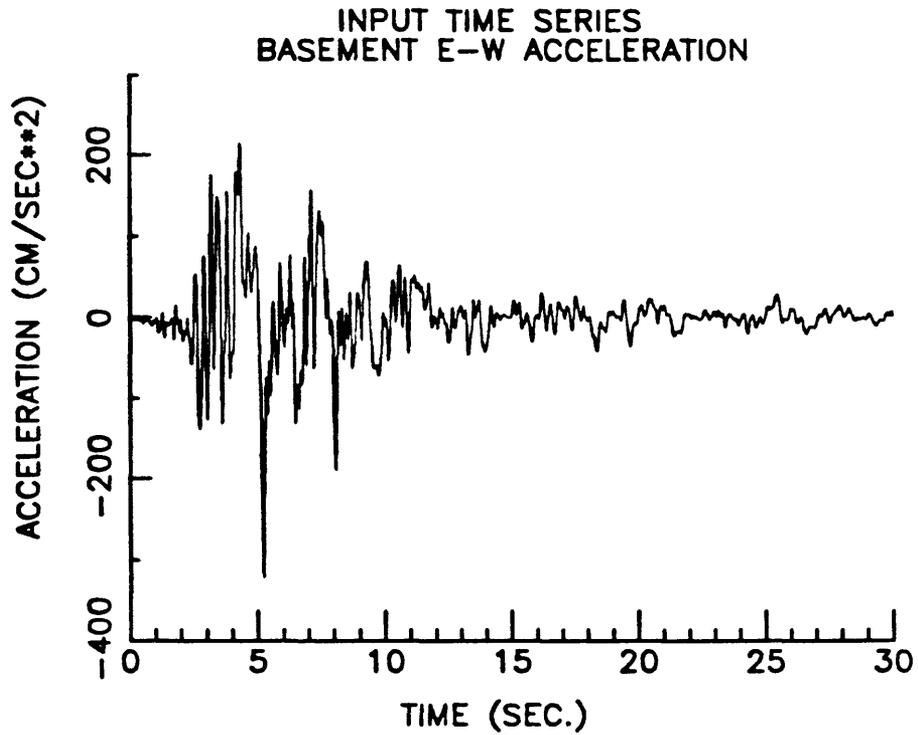


FIG. 13.4.1- Recorded basement and roof East-West accelerations.

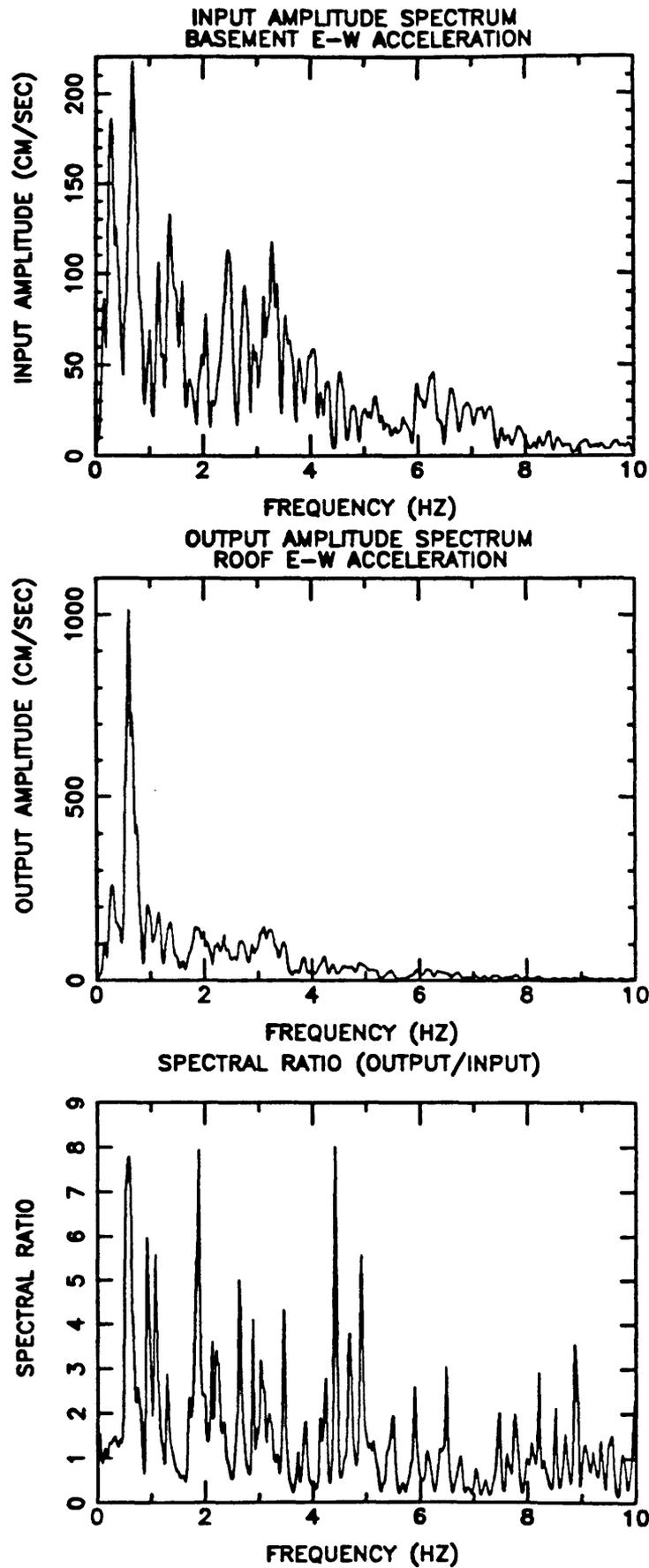
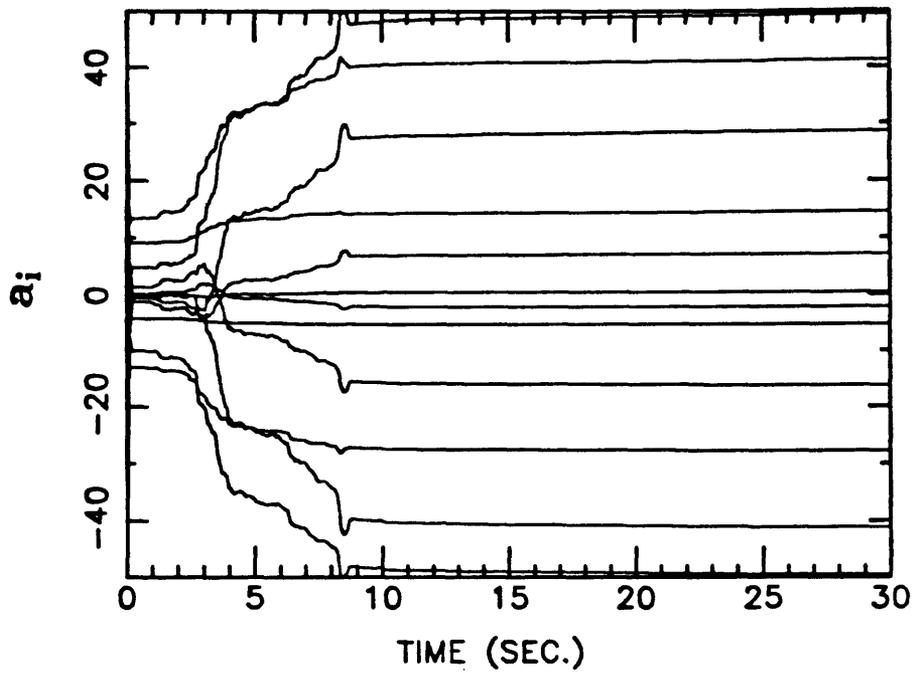


FIG. 13.4.2- Fourier amplitude spectra of basement and roof accelerations, and the spectral ratio (roof/basement).

TIME VARIATION OF OUTPUT PARAMETERS



TIME VARIATION OF INPUT PARAMETERS

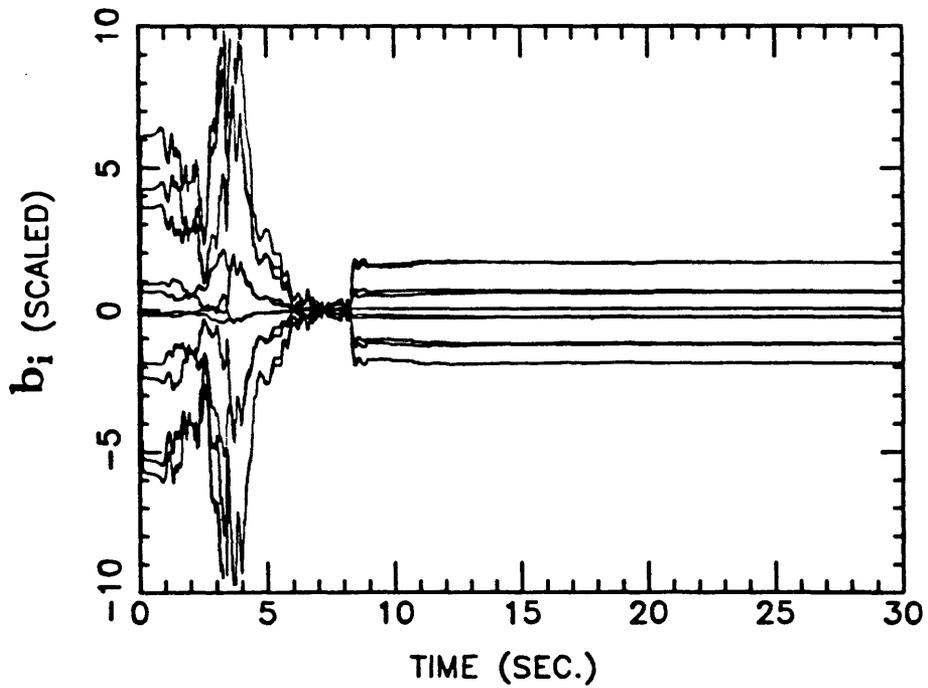


FIG. 13.4.3- Time variation of model parameters.

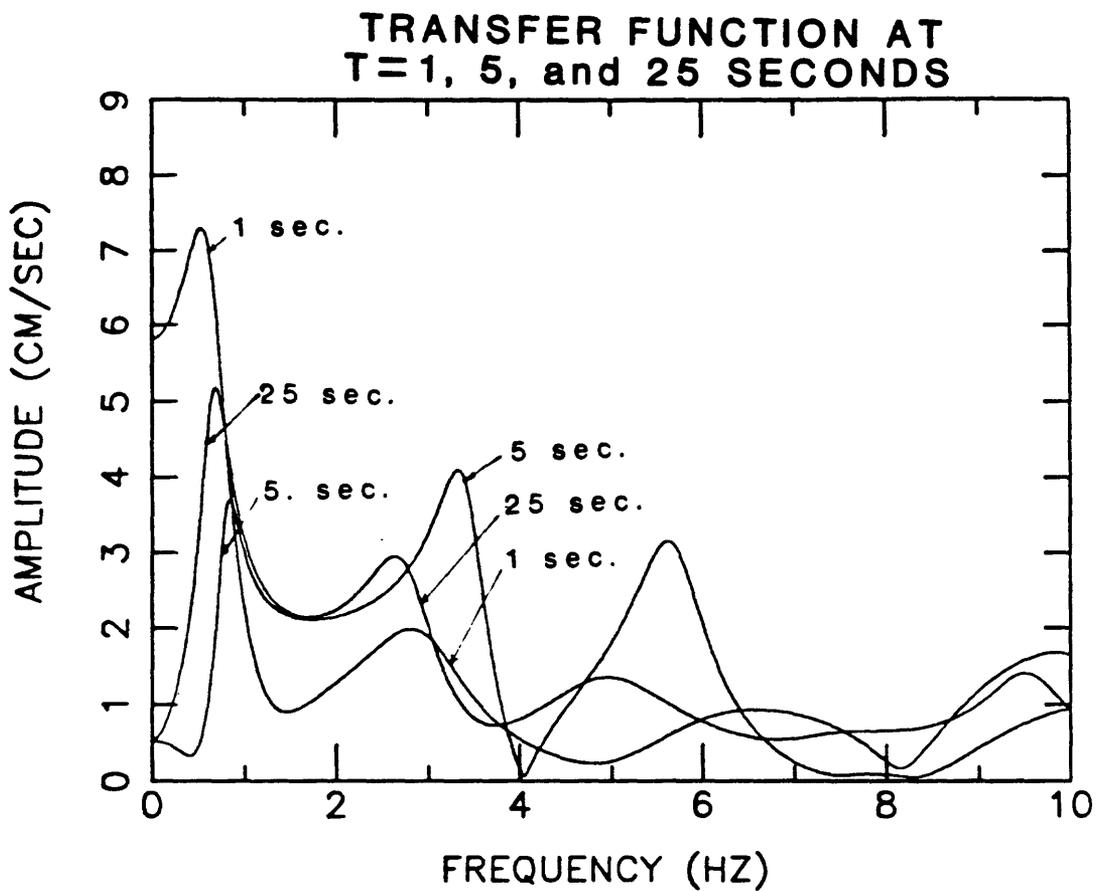


FIG. 13.4.4- Estimated transfer function at $t = 1, 5,$ and 25 seconds.

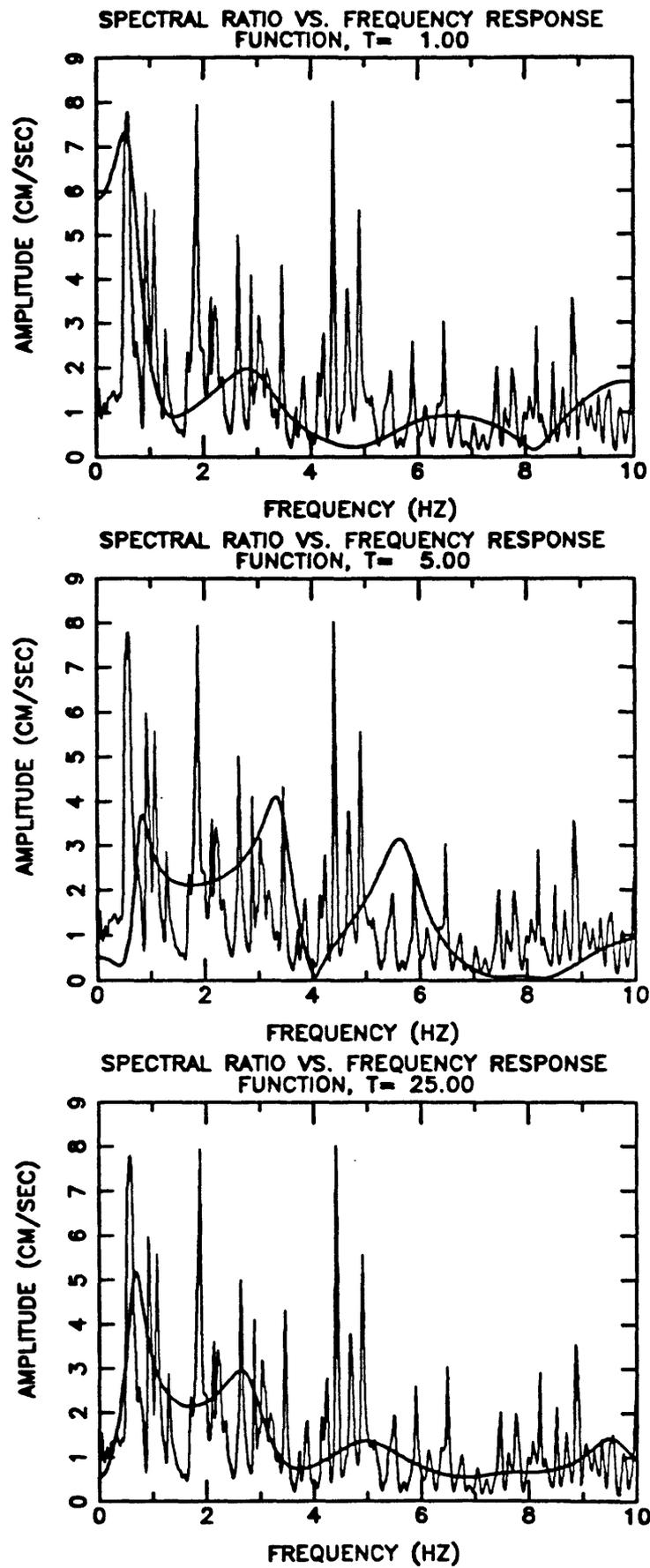


FIG. 13.4.5- Comparison of estimated transfer function with spectral ratio at $t = 1, 5,$ and 25 seconds.

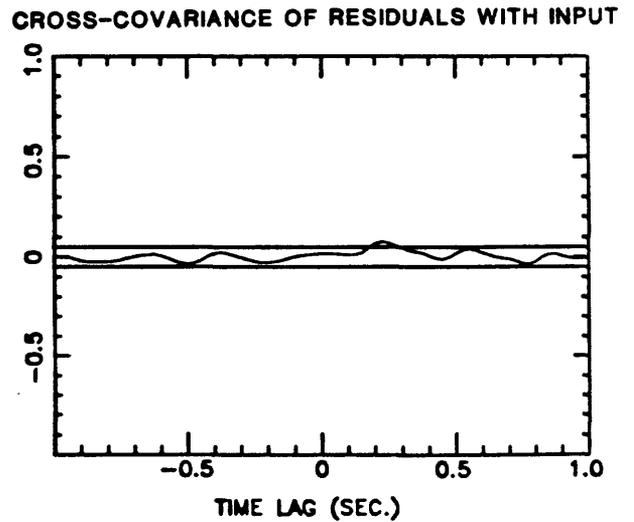
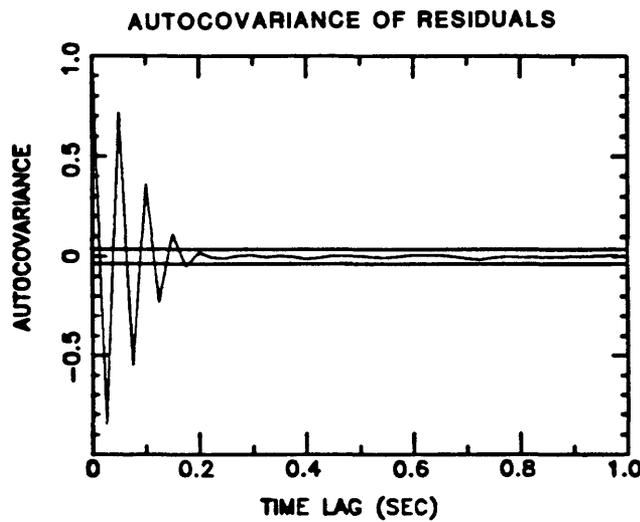
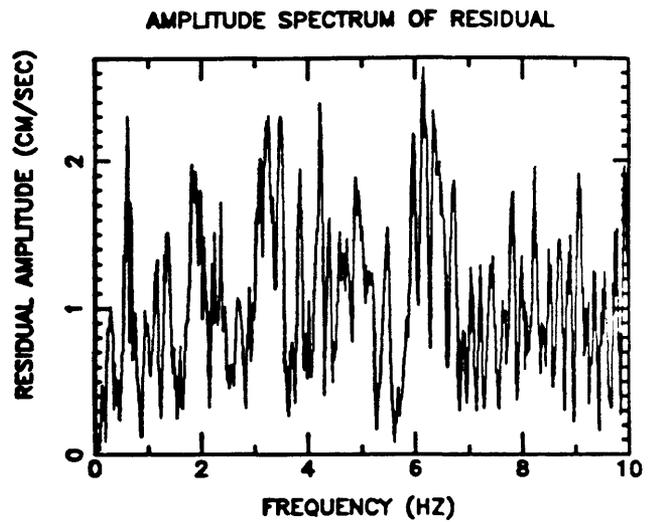
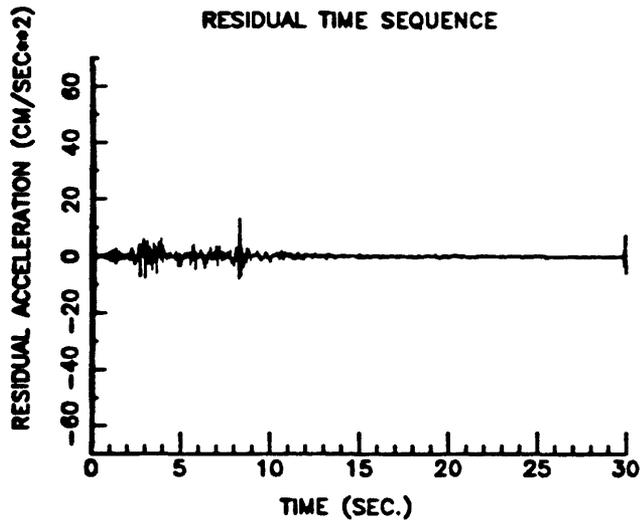


FIG. 13.4.6- Residual time series at $t = 25$ seconds, its Fourier amplitude spectrum, autocovariance, and cross-covariance with input (straight lines in autocovariance and cross-covariance plots show 90-percent confidence levels for whiteness and independence, respectively).

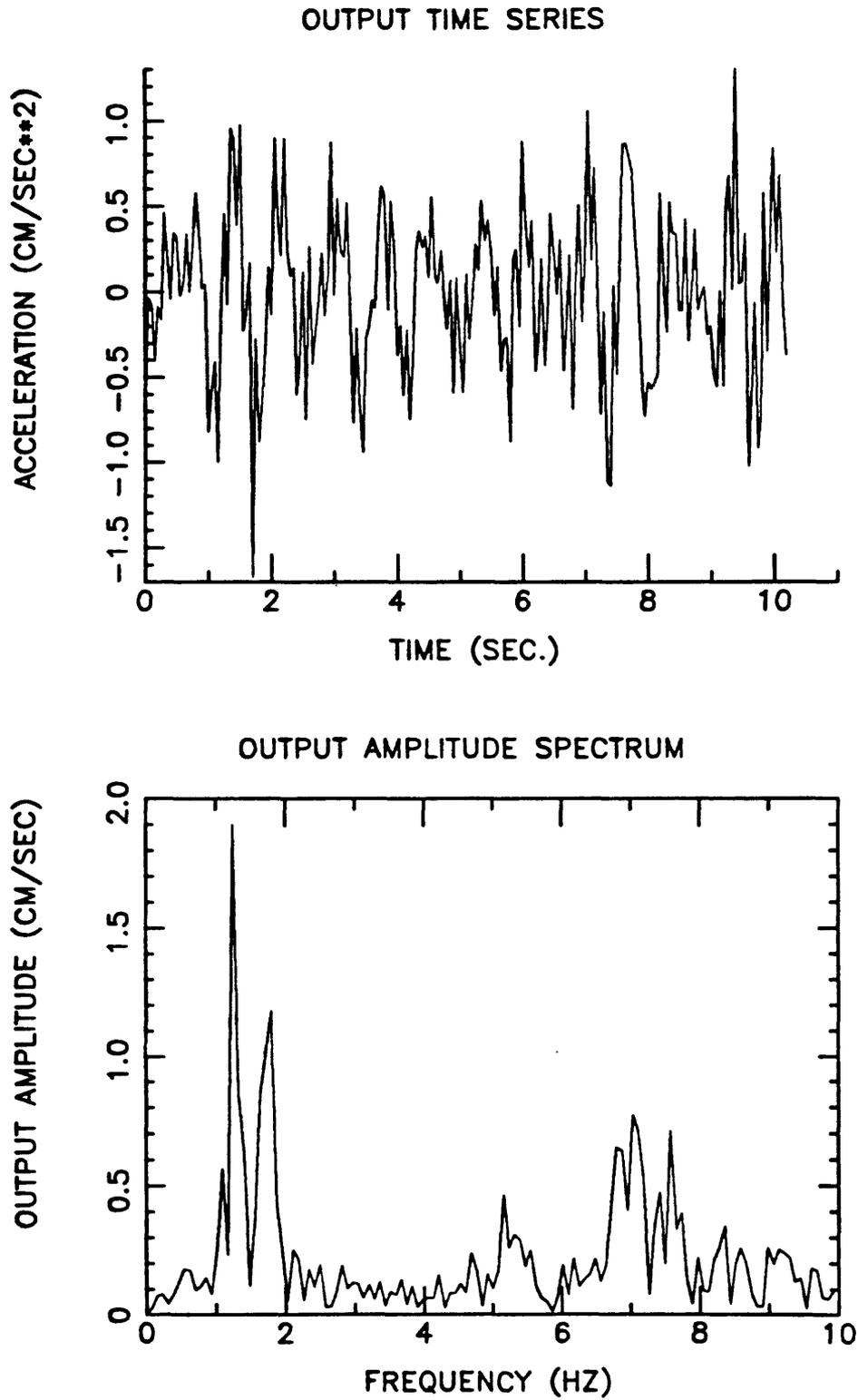


FIG. 13.5.1- Recorded top-floor ambient acceleration, and its Fourier amplitude spectrum.

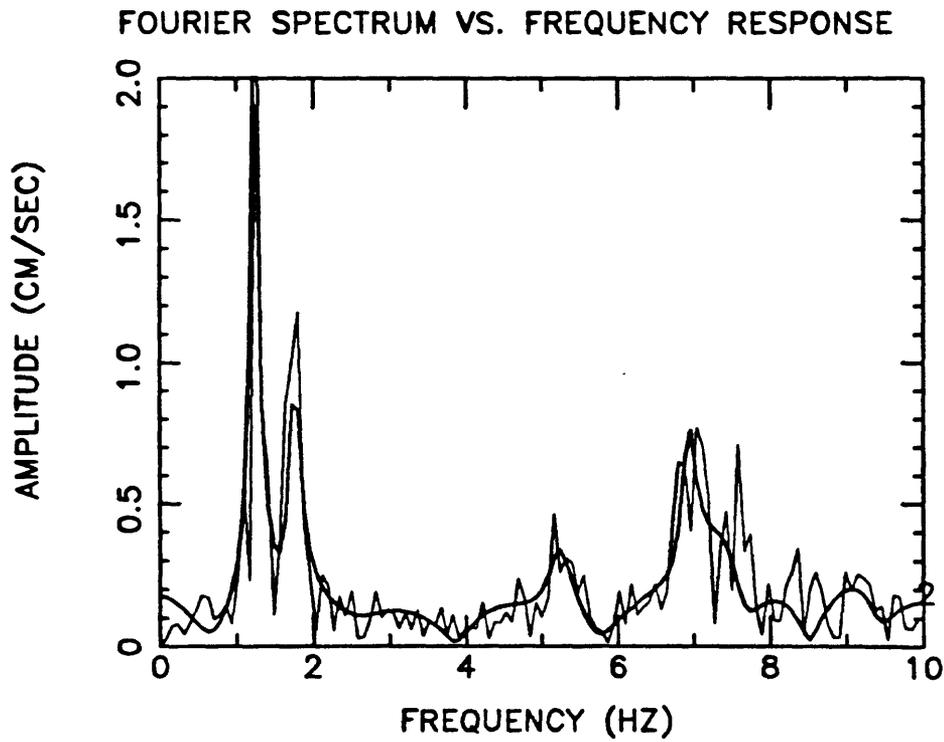
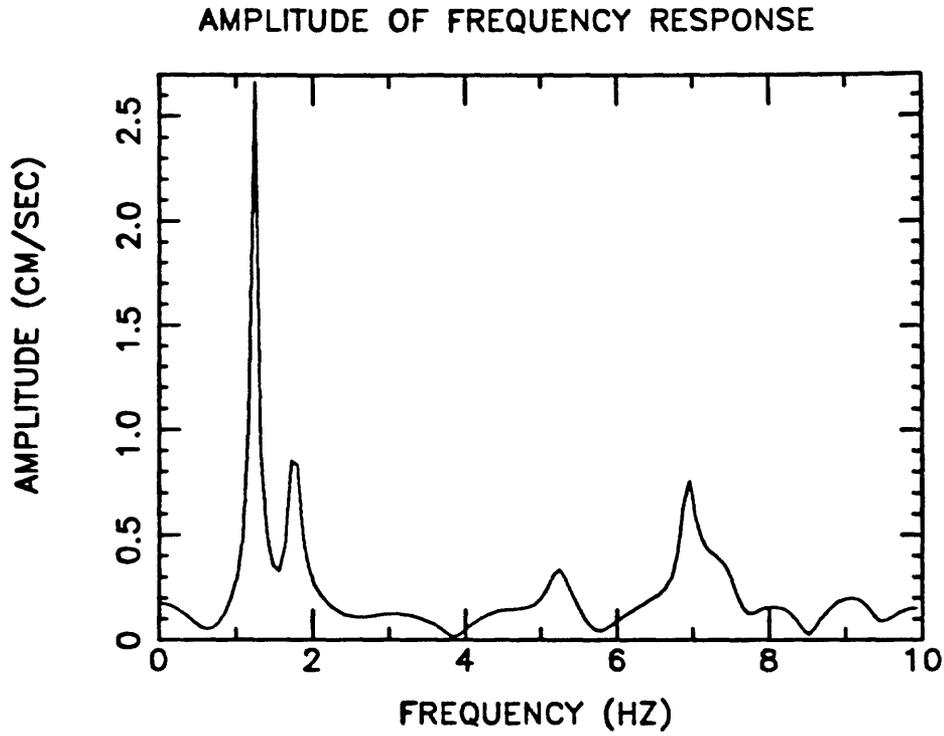


FIG. 13.5.2- Estimated transfer function, and comparison with Fourier spectrum.

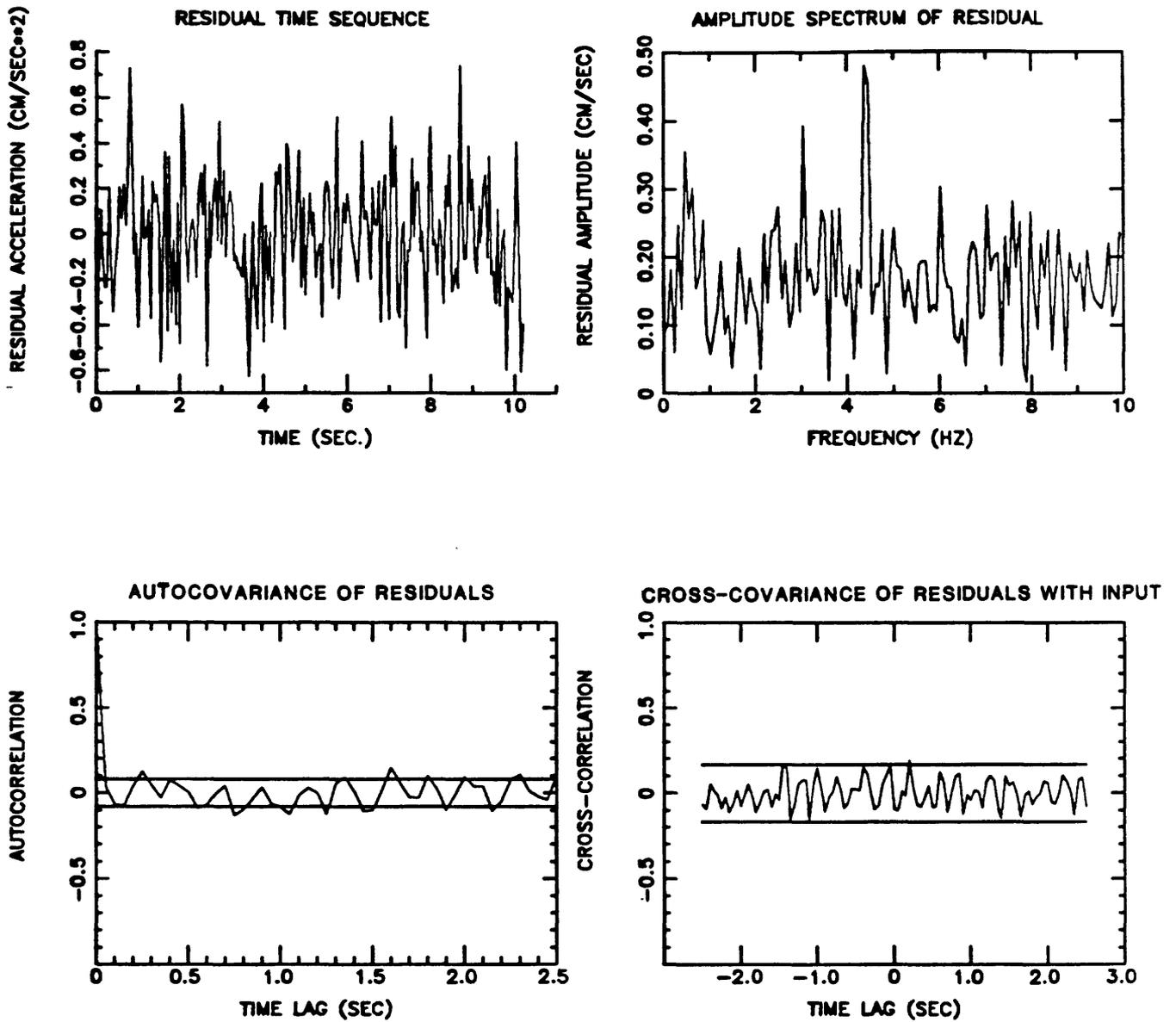


FIG. 13.5.3- Residual time series, its Fourier amplitude spectrum, autocovariance, and cross-covariance with input (straight lines in autocovariance and cross-covariance plots show 90-percent confidence levels for whiteness and independence, respectively).

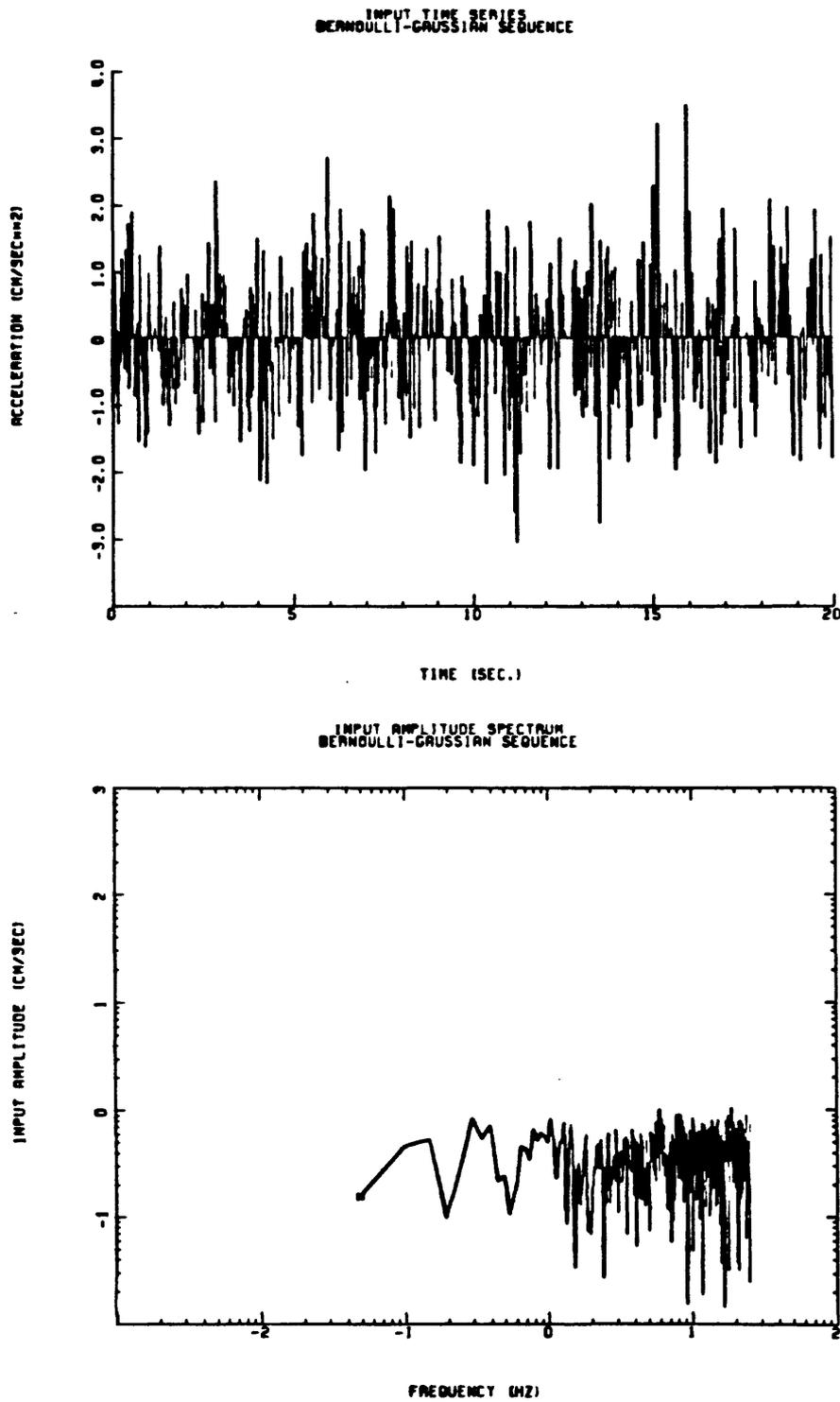


FIG. 13.6.1- A sample function of the Bernoulli-Gaussian sequence, and its Fourier amplitude spectrum.

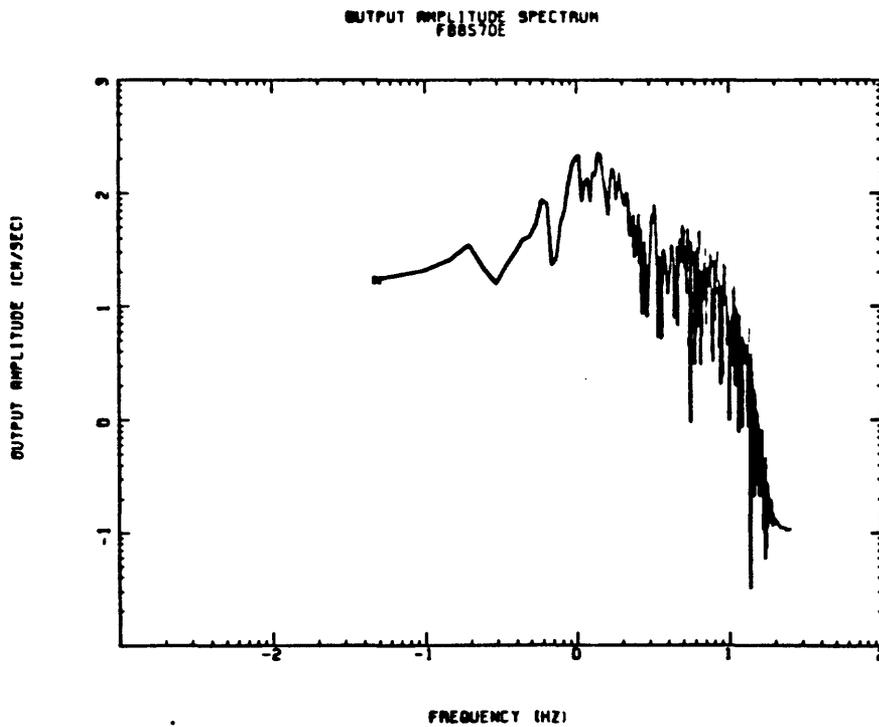
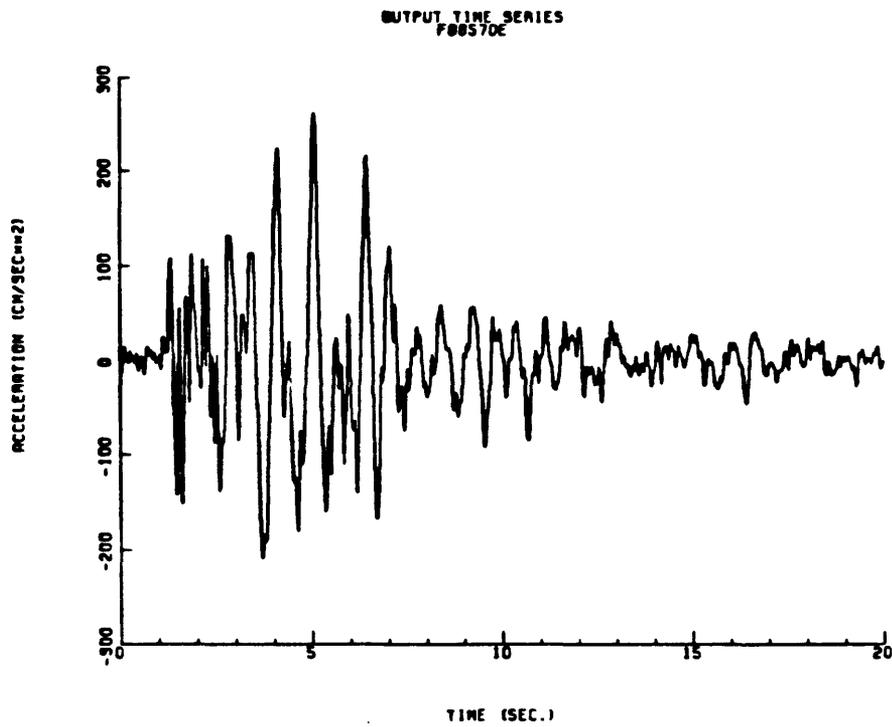


FIG. 13.6.2- Time history and the Fourier amplitude spectrum of the record F88S70E from the 1971 San Fernando earthquake.

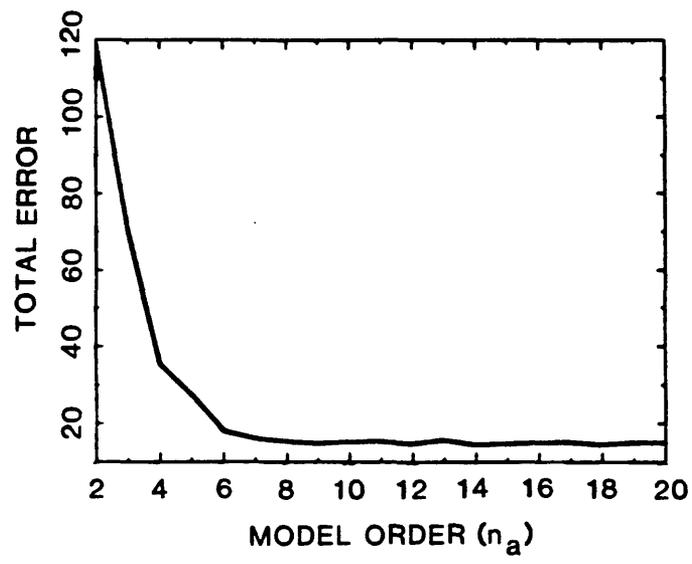
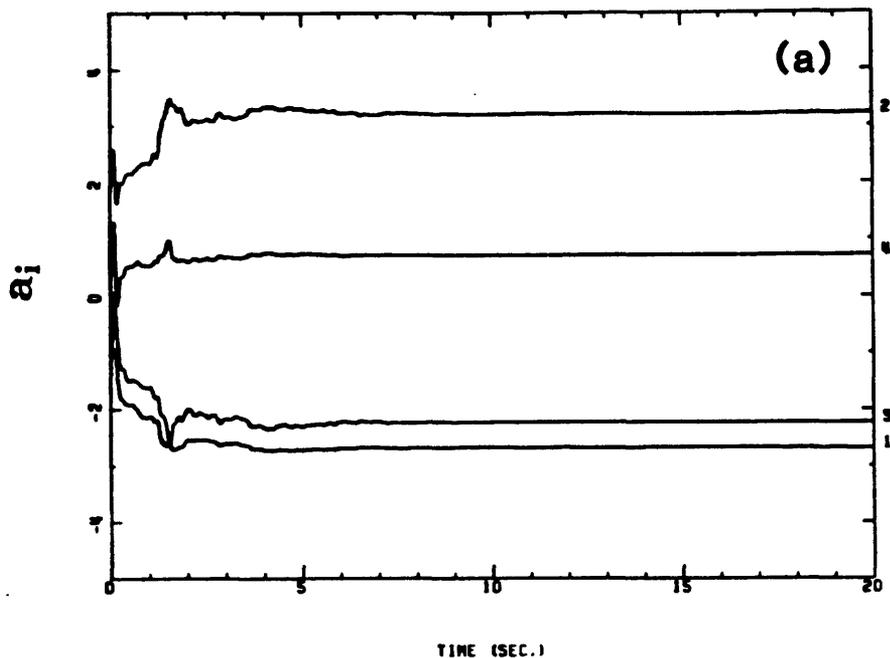


FIG. 13.6.3- Model order versus prediction error for the San Fernando record: n_a vs. $V(N)$ with $n_b = n_a - 1$ and $n_c = 0$.

TIME VARIATION OF OUTPUT PARAMETERS
SAN FERNANDO, FEB 27 1971



TIME VARIATION OF INPUT PARAMETERS
SAN FERNANDO, FEB 27 1971

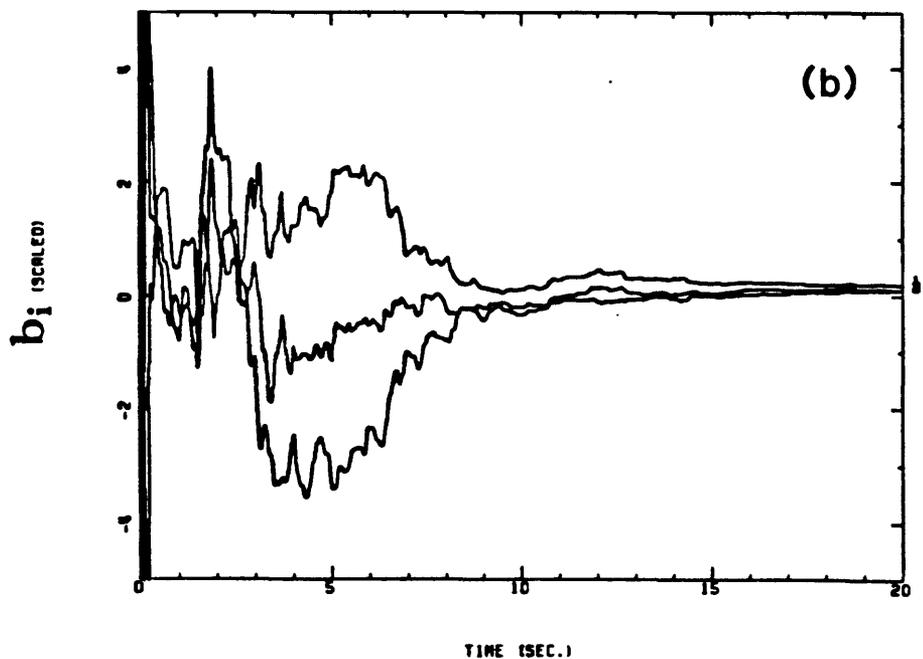
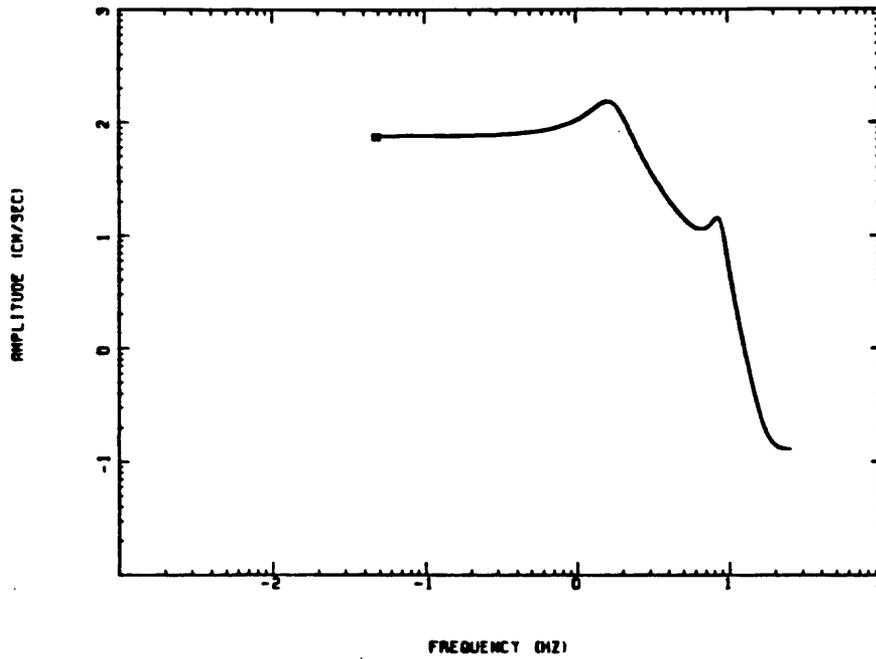


FIG. 13.6.4- Time variation of model parameters for the San Fernando record: (a) a_j , $j = 1 \sim 4$; (b) b_j , $j = 1 \sim 3$.

AMPLITUDE OF FREQUENCY RESPONSE FUNCTION, T= 18.00
SAN FERNANDO, F88570E



FOURIER SPECTRUM VS. FREQUENCY RESPONSE FUNCTION, T= 18.00
SAN FERNANDO, F88570E

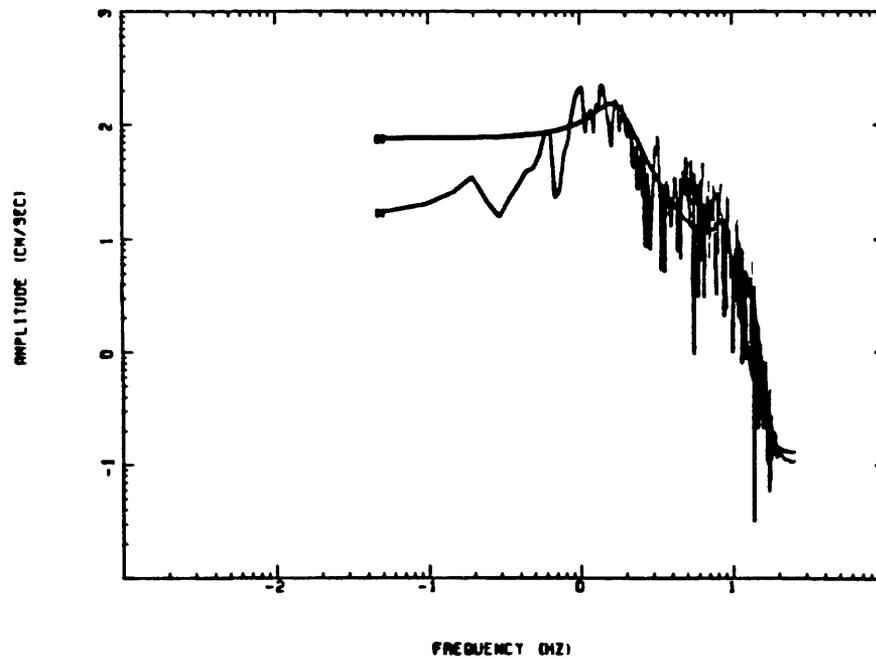


FIG. 13.6.5- The amplitude of the transfer function for (4-3) model of the San Fernando record evaluated at $t = 18$ sec., and its match with the Fourier amplitude spectrum.

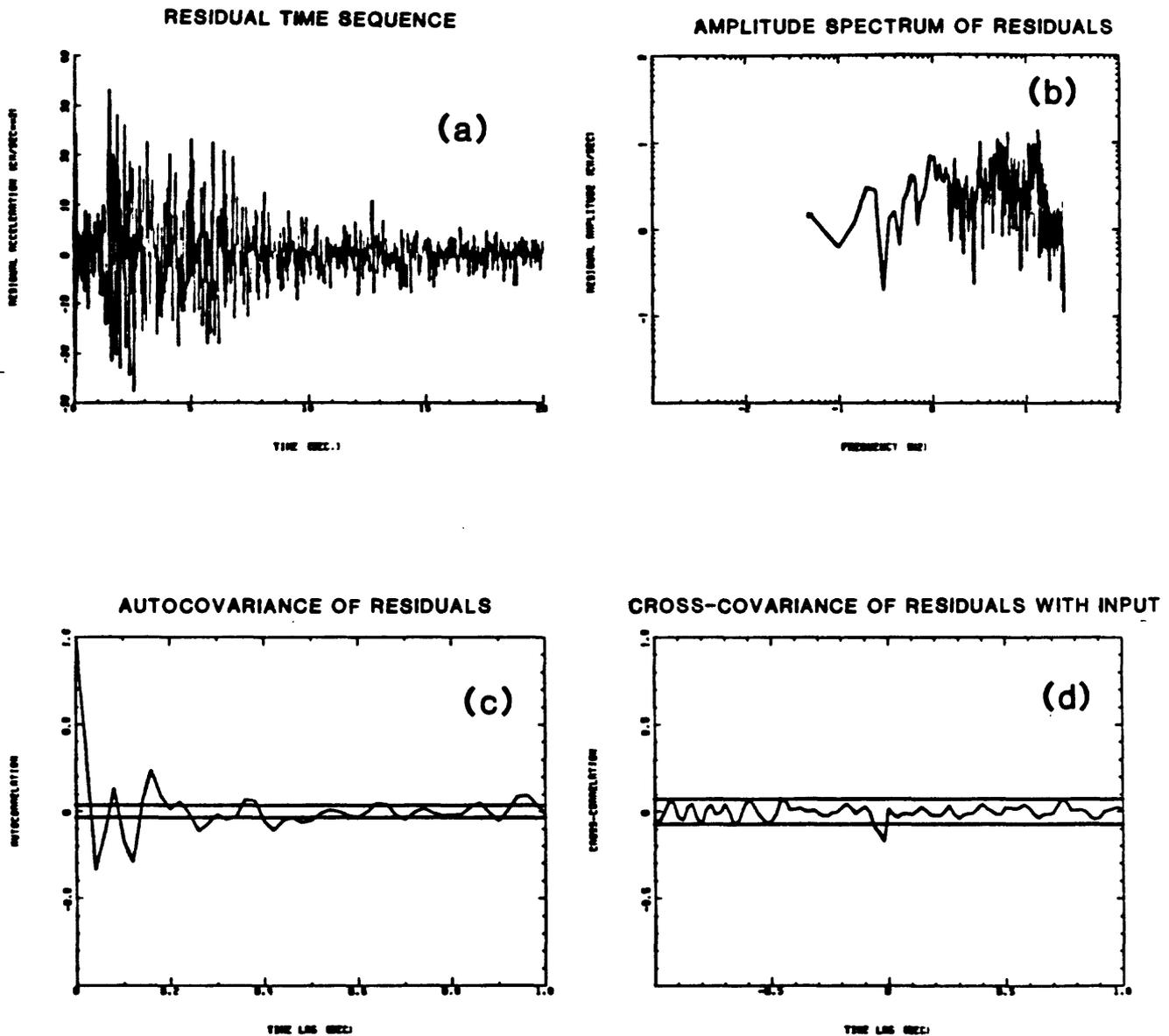
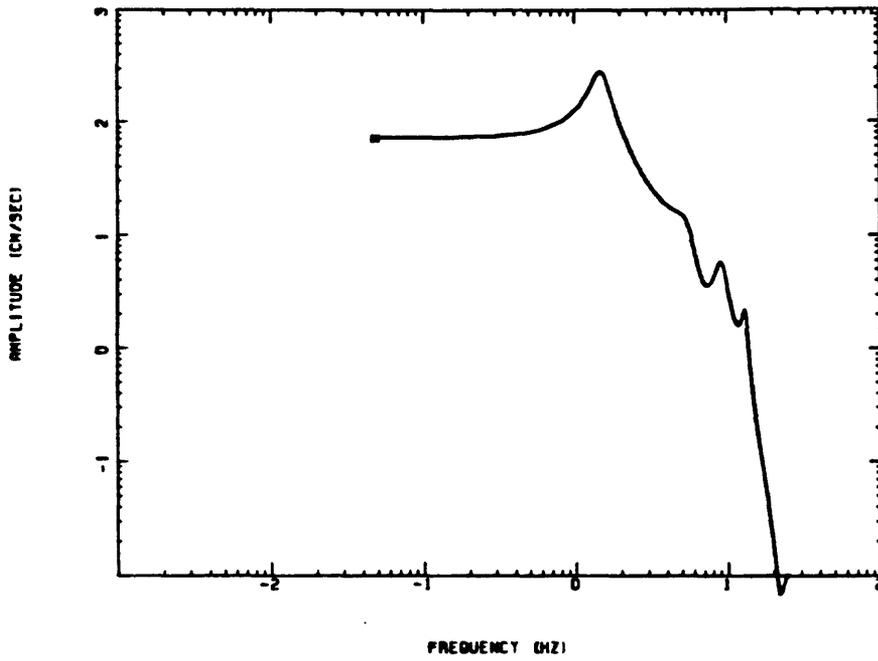


FIG. 13.6.6- The residual tests for (4-3) model of the San Fernando record evaluated at $t = 18$ sec.: (a) time history of the residual; (b) Fourier amplitude spectrum of the residuals; (c) autocovariance function of the residuals, and the 90-percentile confidence limits for whiteness according to the Chi-square χ^2 test; (d) cross-covariance function of the residuals with input, and the 90-percentile confidence limits according to (0,1) normal distribution law.

AMPLITUDE OF FREQUENCY RESPONSE FUNCTION, T= 18.00
SAN FERNANDO. F88570E



FOURIER SPECTRUM VS. FREQUENCY RESPONSE FUNCTION, T= 18.00
SAN FERNANDO. F88570E

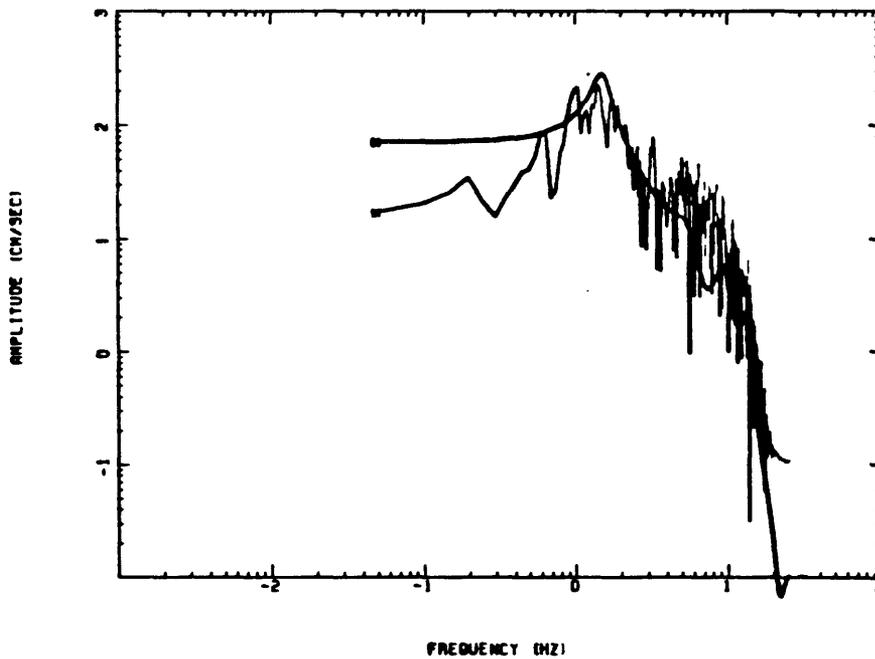


FIG. 13.6.7- The amplitude of the transfer function for (8-7) model of the San Fernando record evaluated at $t = 18$ sec., and its match with the Fourier amplitude spectrum.

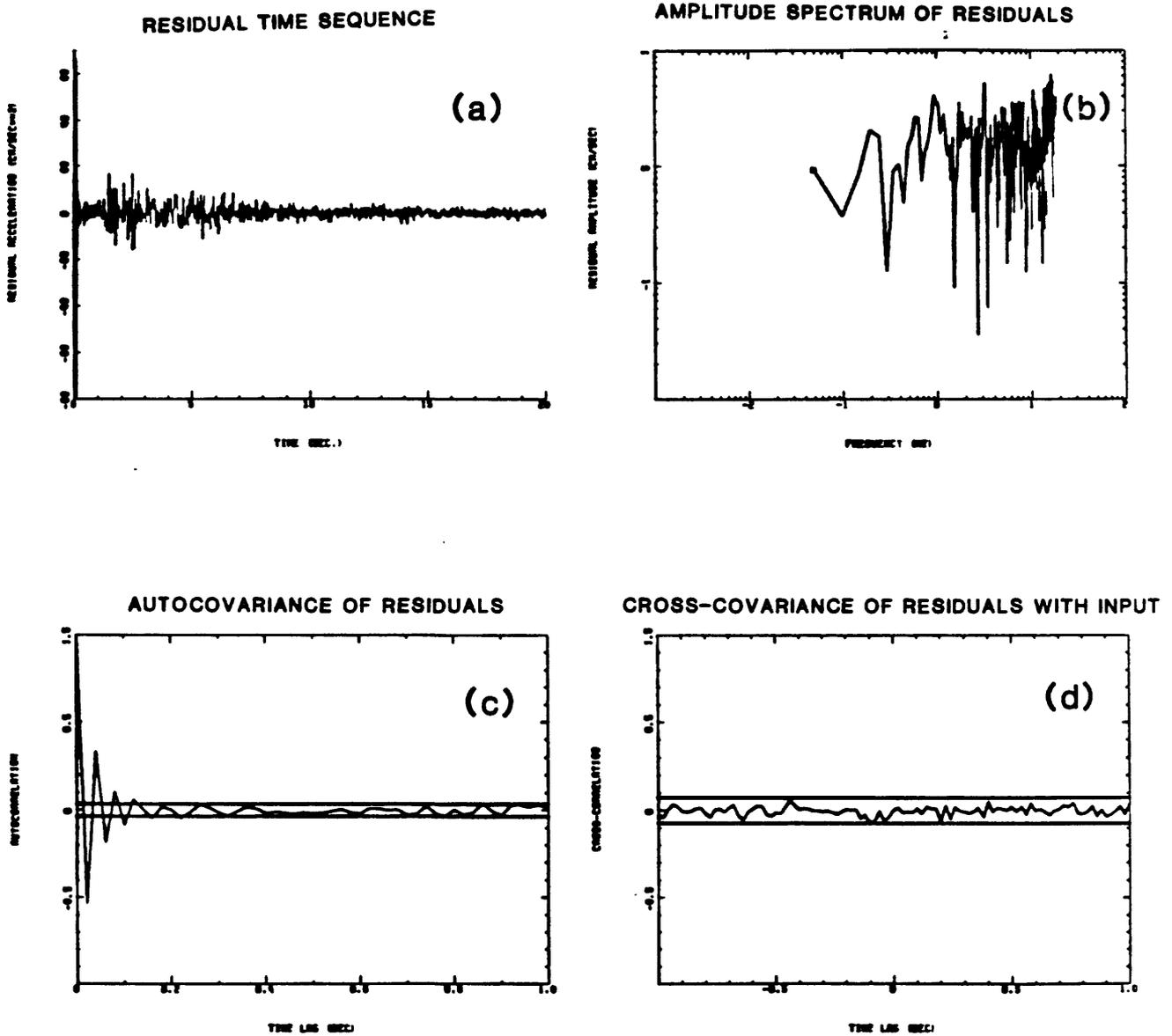


FIG. 13.6.8- The residual tests for (8-7) model of the San Fernando record evaluated at $t = 18$ sec.: (a) time history of the residual; (b) Fourier amplitude spectrum of the residuals; (c) autocovariance function of the residuals, and the 90-percentile confidence limits for whiteness according to the Chi-square χ^2 test; (d) cross-covariance function of the residuals with input, and the 90-percentile confidence limits according to (0,1) normal distribution law.

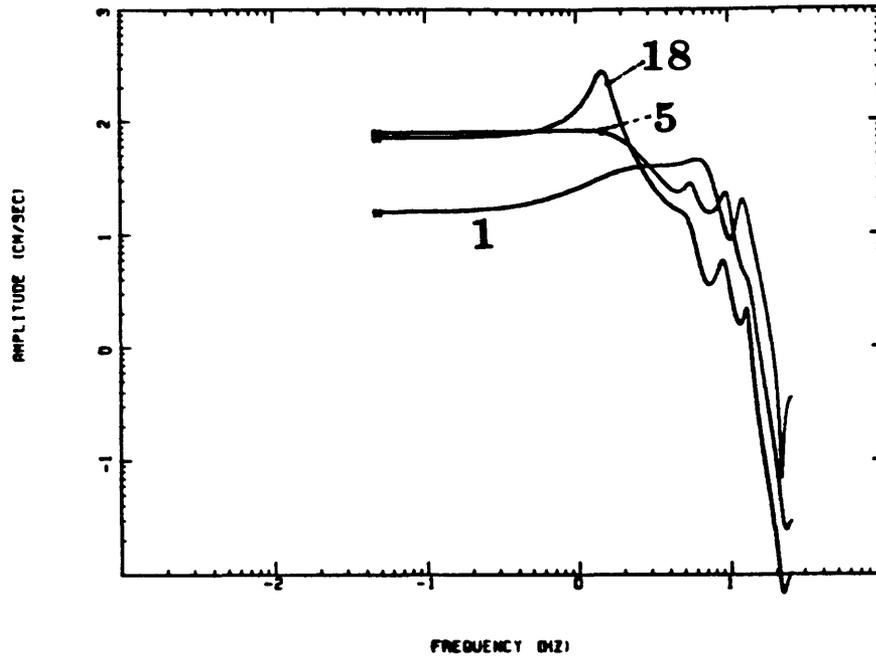


FIG. 13.6.9- The amplitude of the transfer function for (8-7) model of the San Fernando record evaluated at times $t = 1, 5$ and 18 sec.

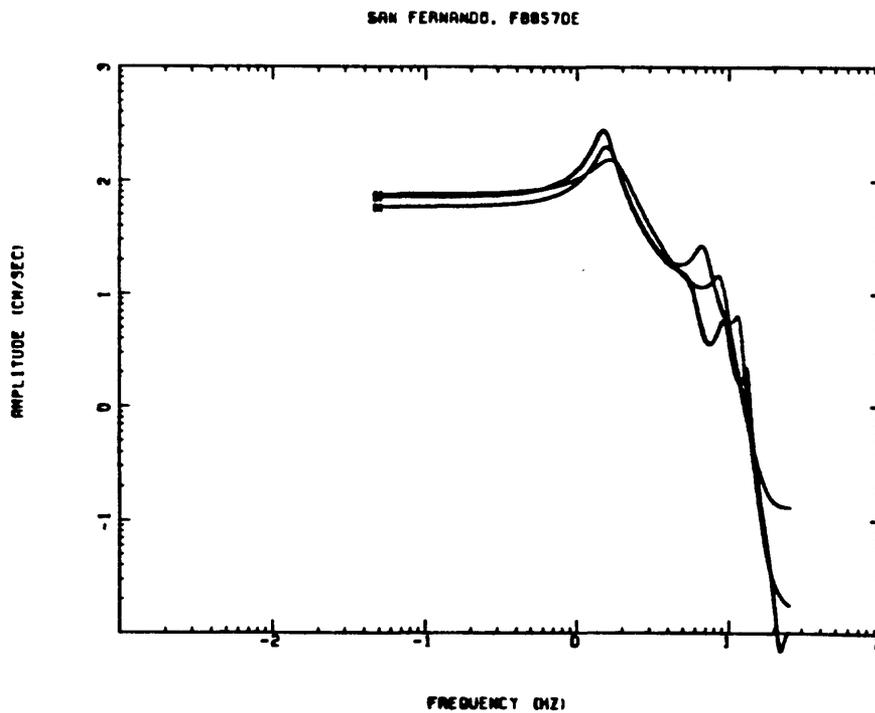


FIG. 13.6.10- The amplitude of the transfer function of the San Fernando record evaluated at $t = 18$ sec. for models (4-3), (8-7) and (12-11).

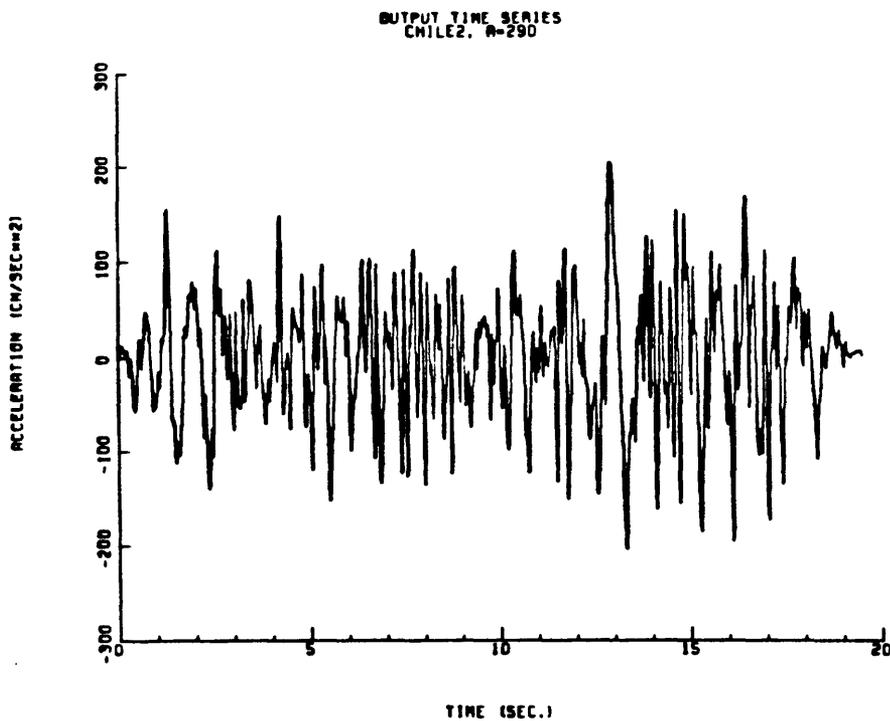
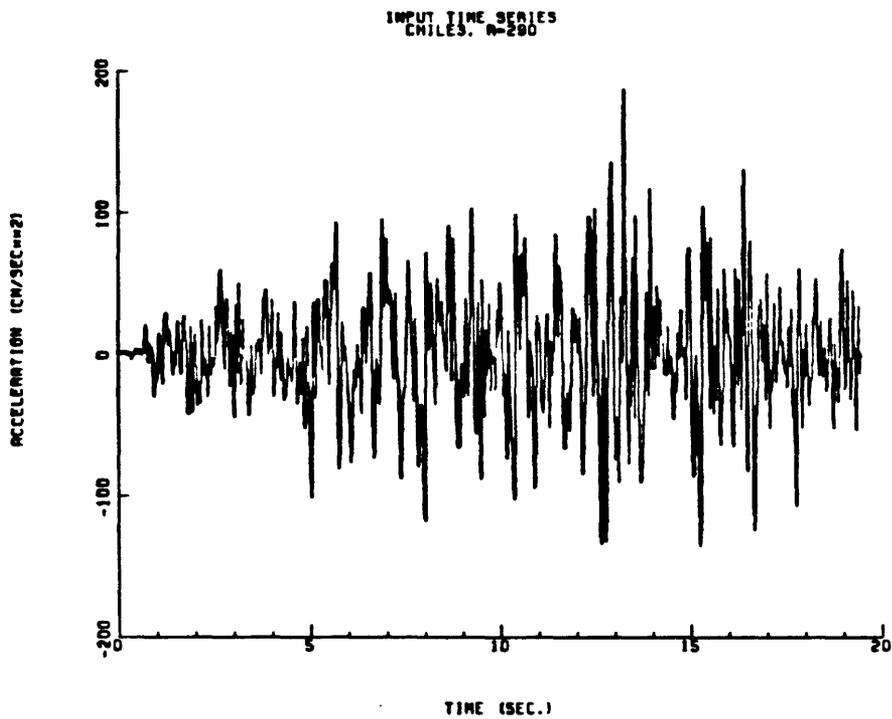


FIG. 13.7.1- Input (rock site) and output (soil site) time series for the site amplification example. The records are from the 3 March 1985 Chilean earthquake.

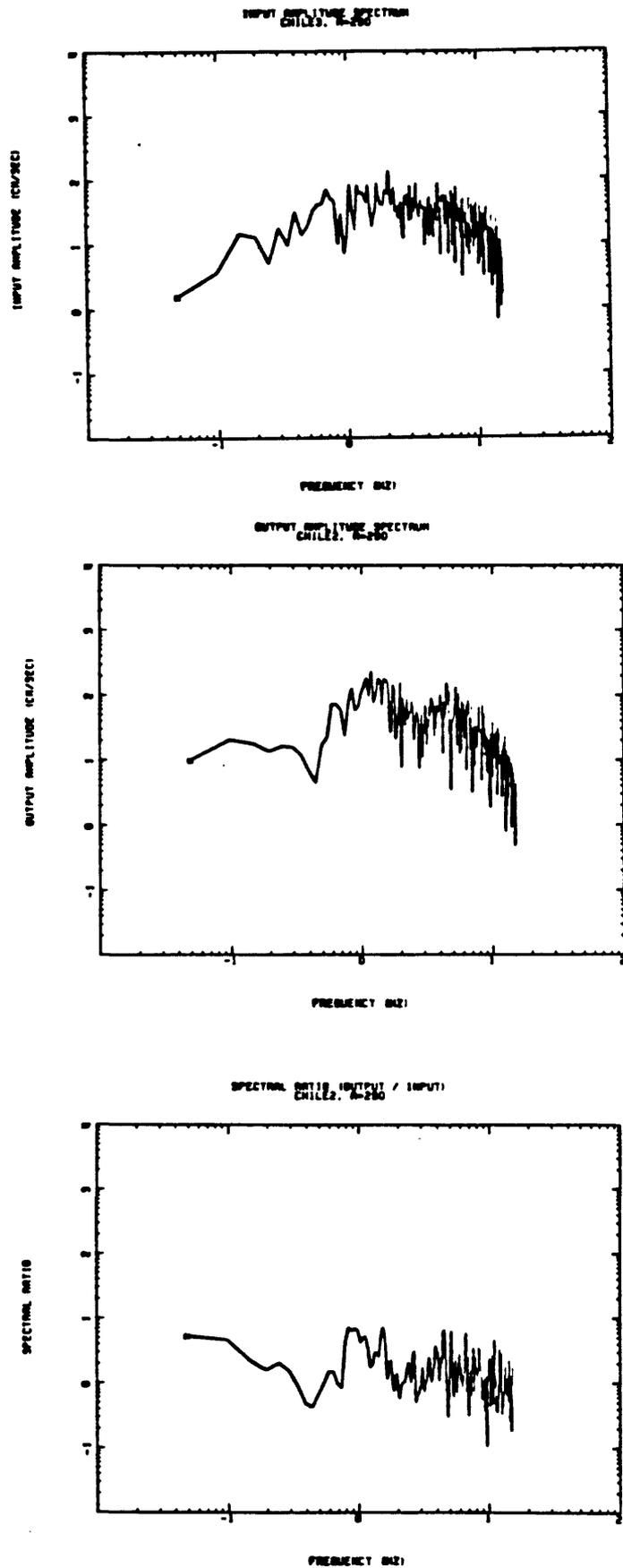


FIG. 13.7.2- Input (rock site) and output (soil site) Fourier amplitude spectra and the smoothed spectral ratio (soil site/rock site) for the site amplification example. The records are from the 3 March 1985 Chilean earthquake.

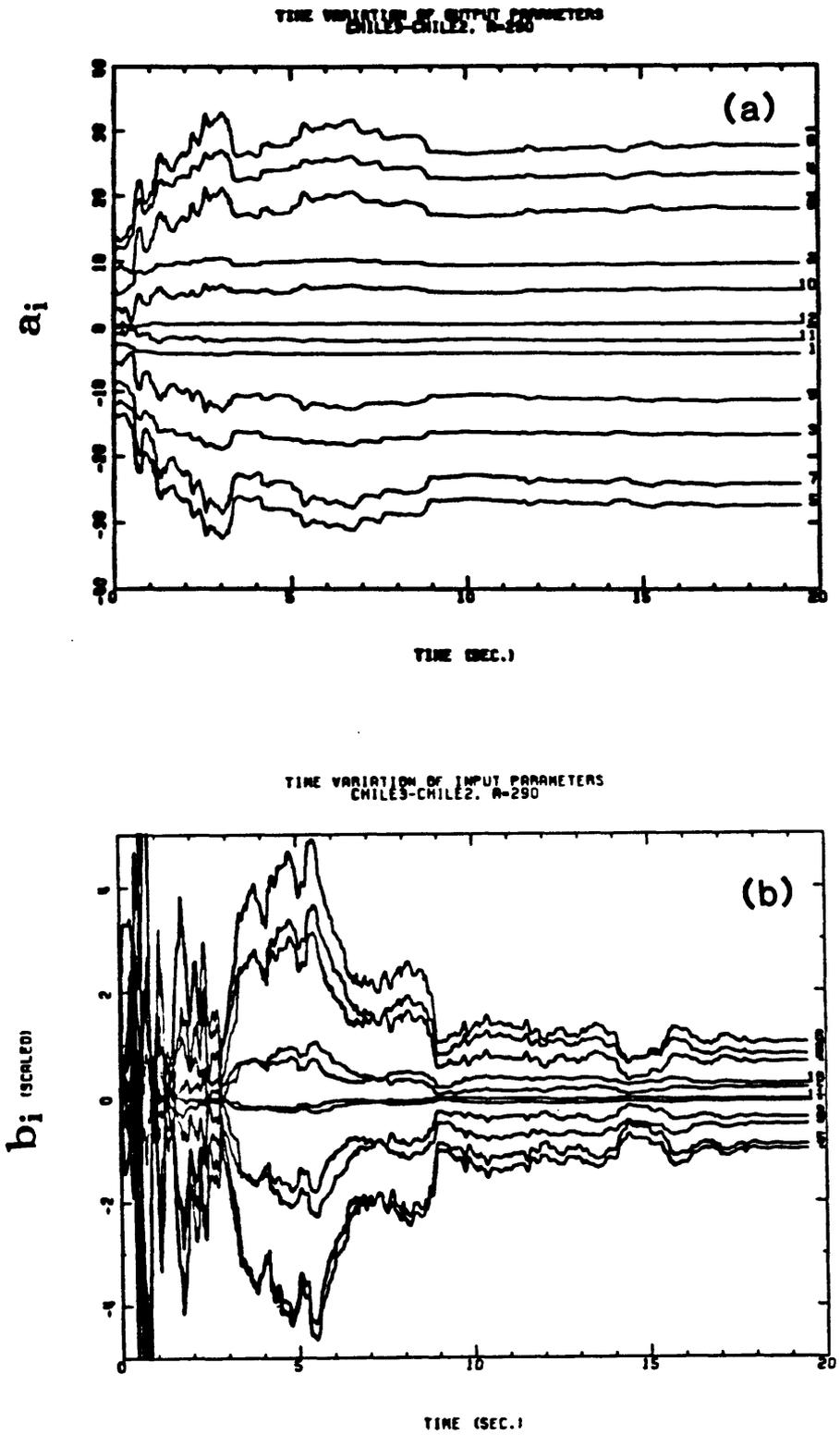


FIG. 13.7.3- Time variation of model parameters for the site amplification example (1985 Chilean earthquake): (a) a_j , $j = 1 \sim 12$; (b) b_j , $j = 1 \sim 11$.

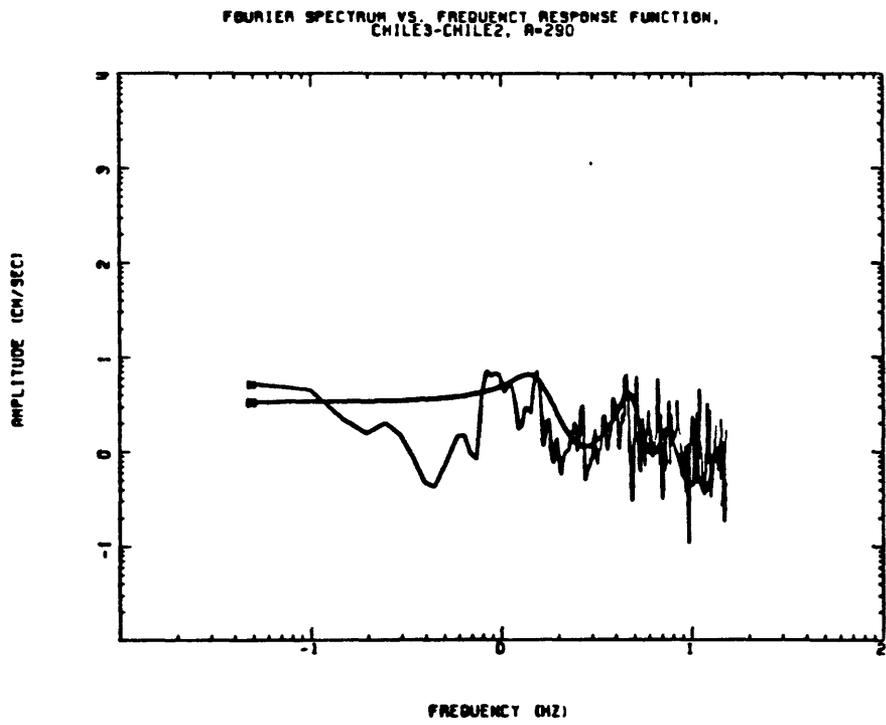
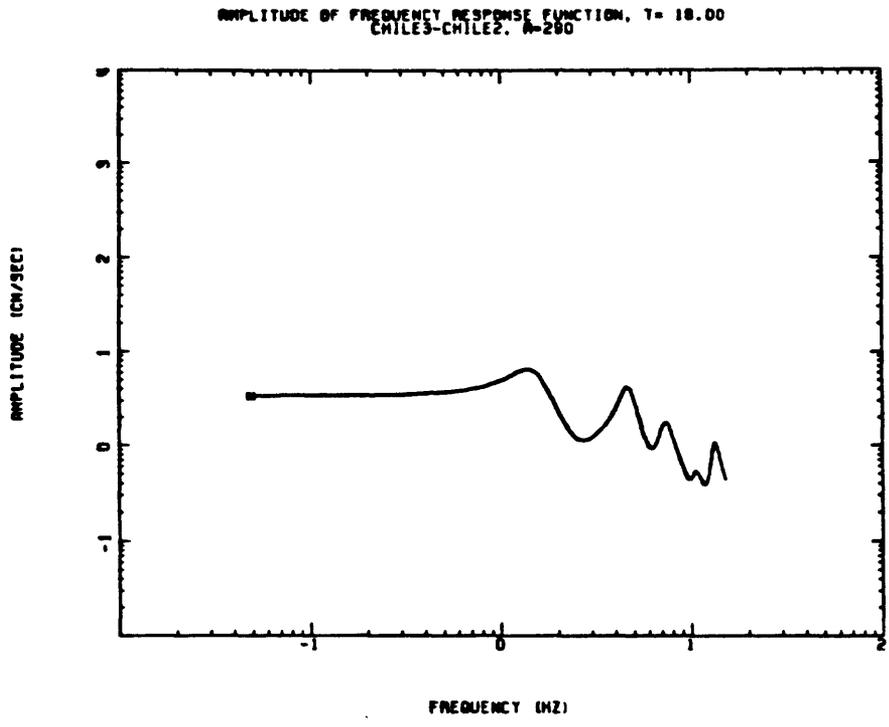


FIG. 13.7.4- The amplitude of the transfer function for (12-11) model for the site amplification example (1985 Chilean earthquake) evaluated at $t = 18$ sec., and its match with the Fourier amplitude spectrum.

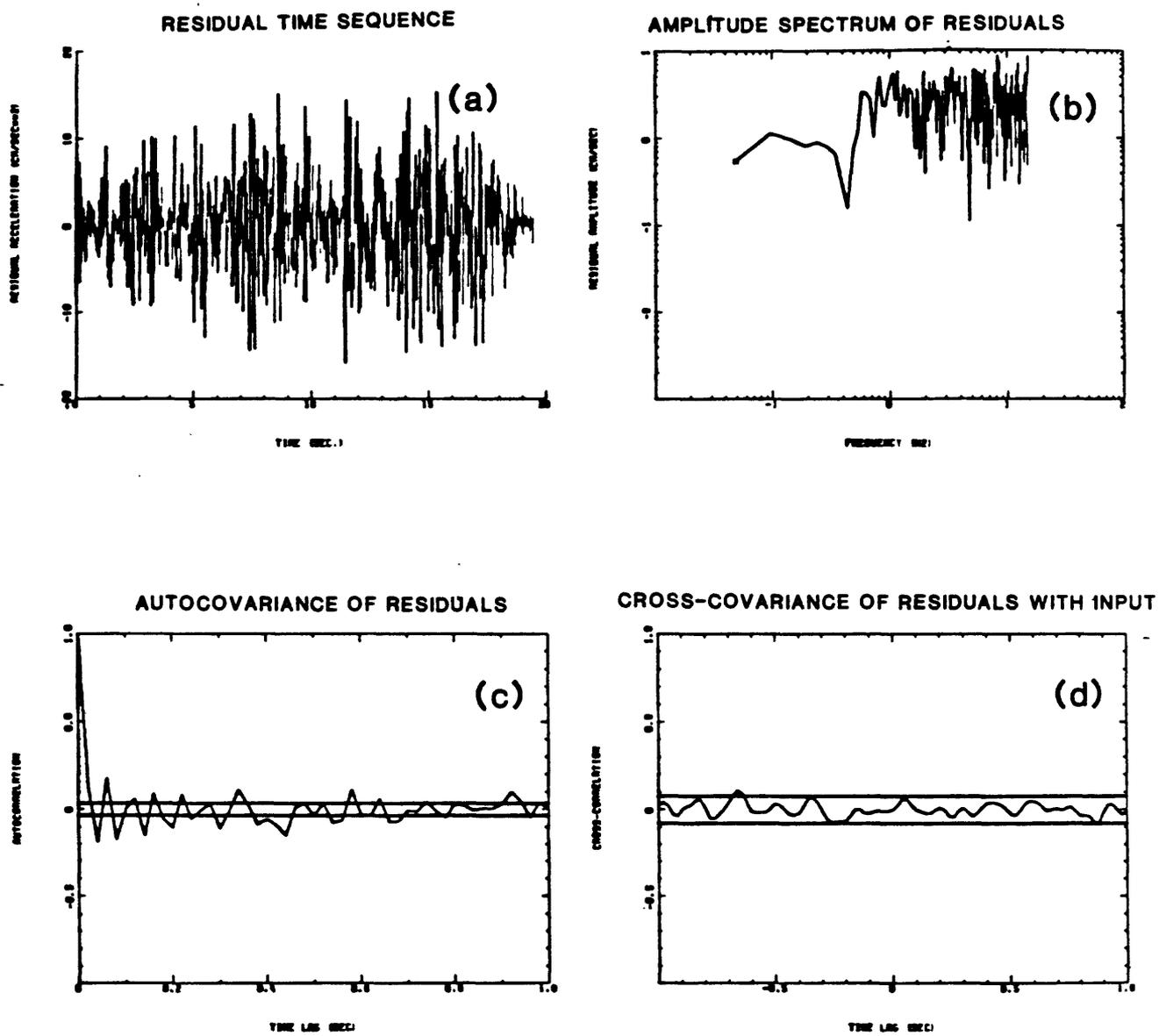


FIG. 13.7.5- The residual tests for (12-11) model for the site amplification example (1985 Chilean earthquake) evaluated at $t = 18$ sec.: (a) time history of the residual; (b) Fourier amplitude spectrum of the residuals; (c) autocovariance function of the residuals, and the 90-percentile confidence limits for whiteness according to the Chi-square χ^2 test; (d) cross-covariance function of the residuals with input, and the 90-percentile confidence limits according to (0,1) normal distribution law.

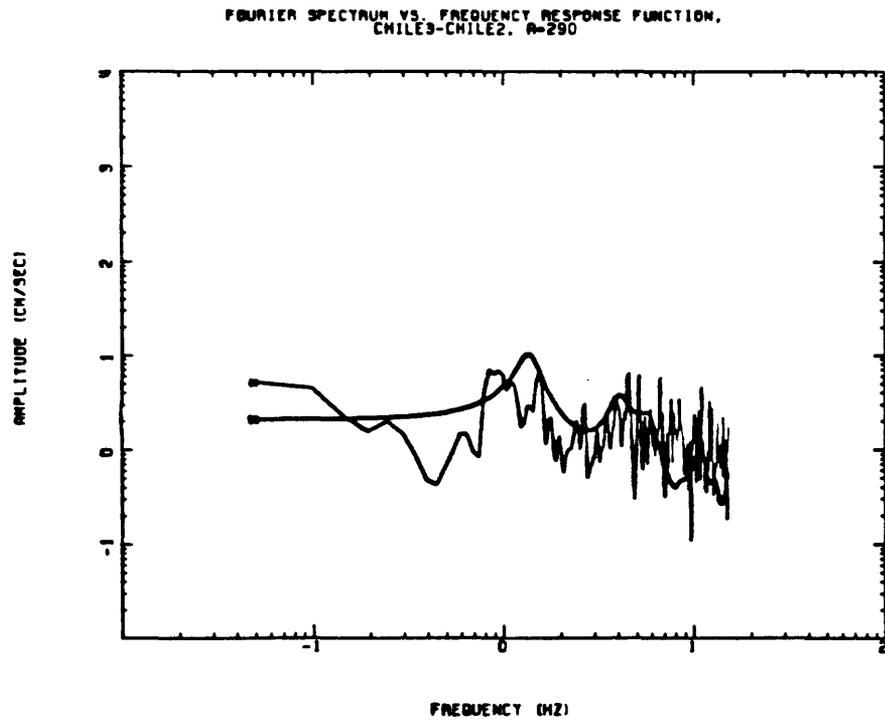
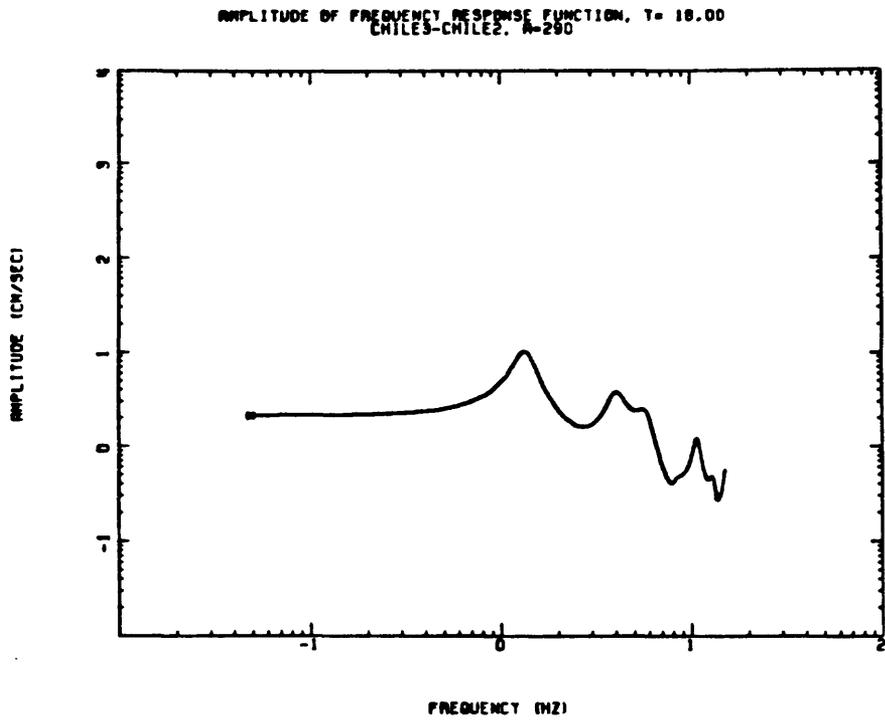


FIG. 13.7.6- The amplitude of the transfer function for (16-15) model for the site amplification example (1985 Chilean earthquake) evaluated at $t = 18$ sec., and its match with the Fourier amplitude spectrum.

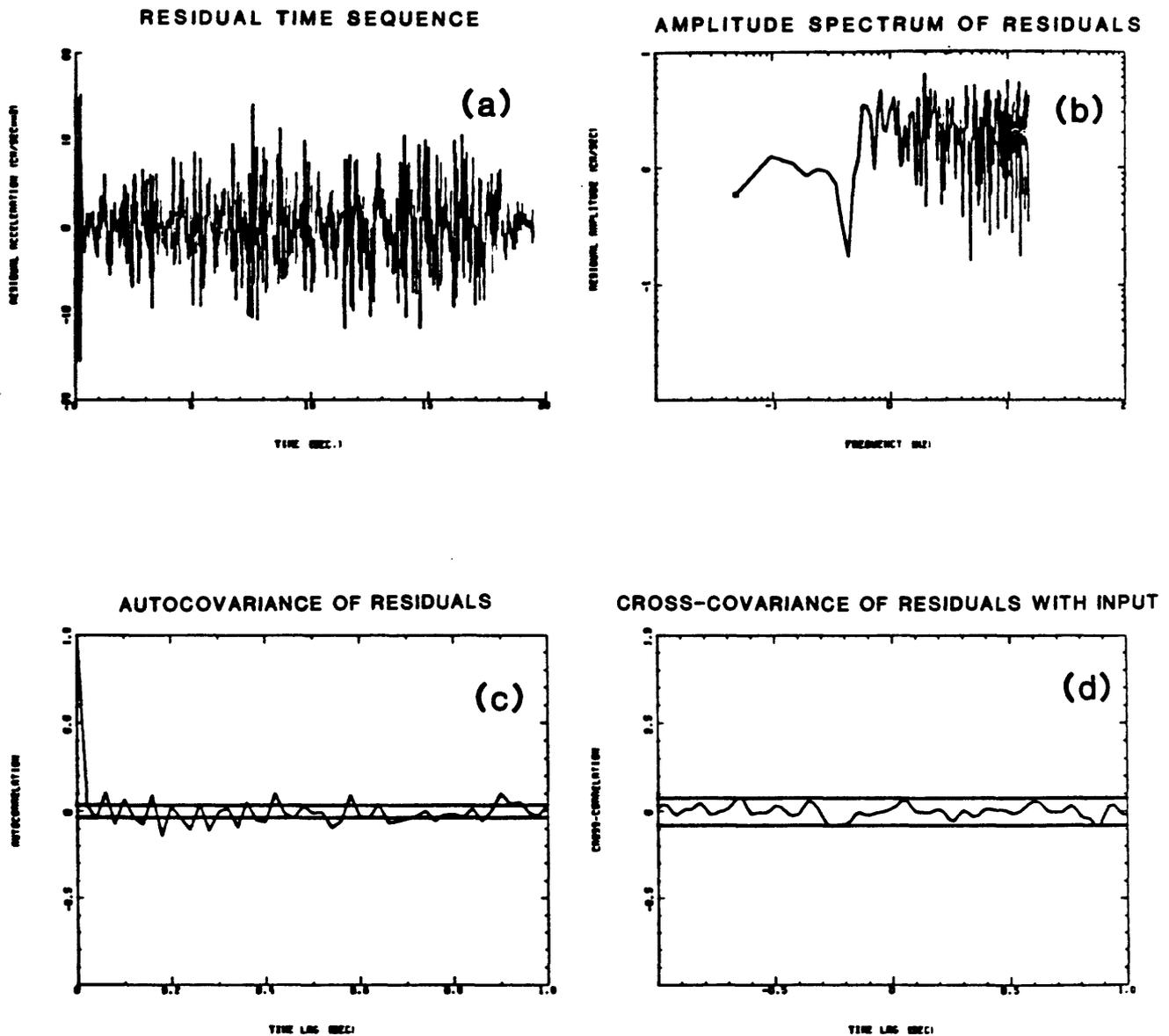


FIG. 13.7.7- The residual tests for (16-15) model for the site amplification example (1985 Chilean earthquake) evaluated at $t = 18$ sec.: (a) time history of the residual; (b) Fourier amplitude spectrum of the residuals; (c) autocovariance function of the residuals, and the 90-percentile confidence limits for whiteness according to the Chi-square χ^2 test; (d) cross-covariance function of the residuals with input, and the 90-percentile confidence limits according to (0,1) normal distribution law.

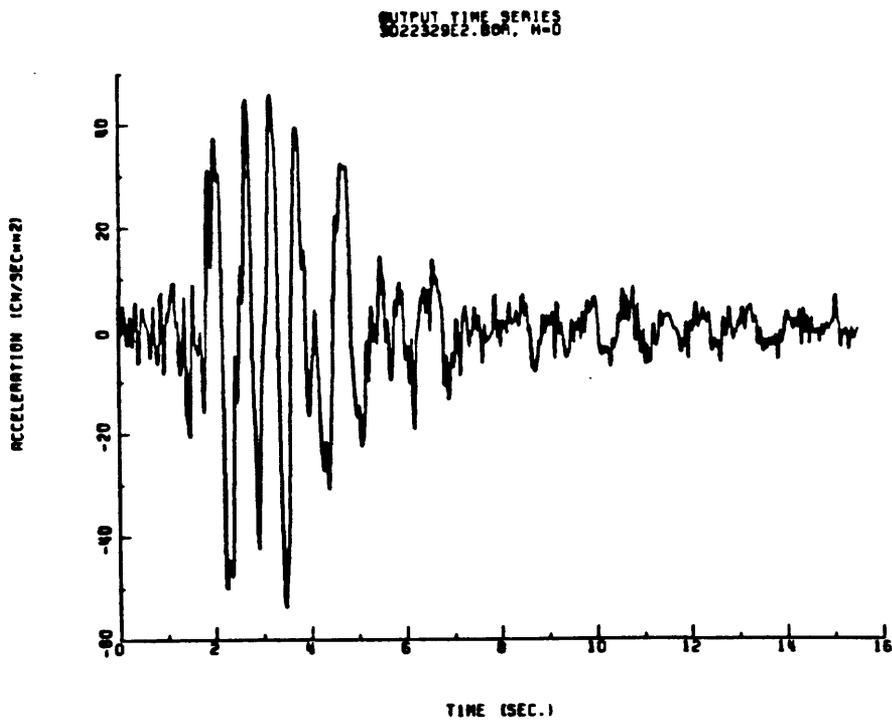
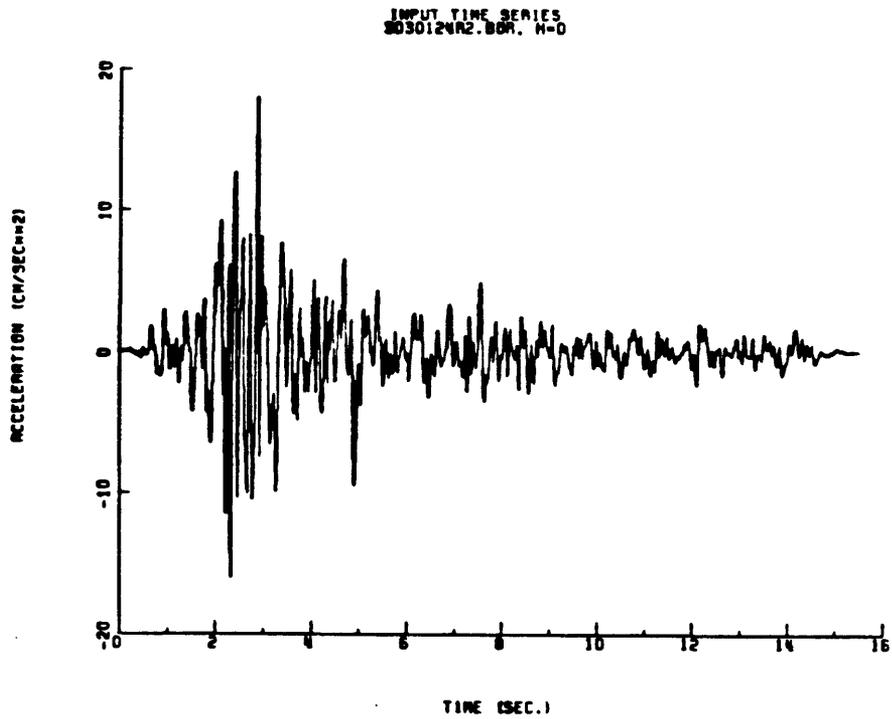


FIG. 13.8.1- Input (small earthquake) and output (large earthquake) time series for the source amplification example. The records are from the aftershocks of 1983 Borah Peak, Idaho earthquake.

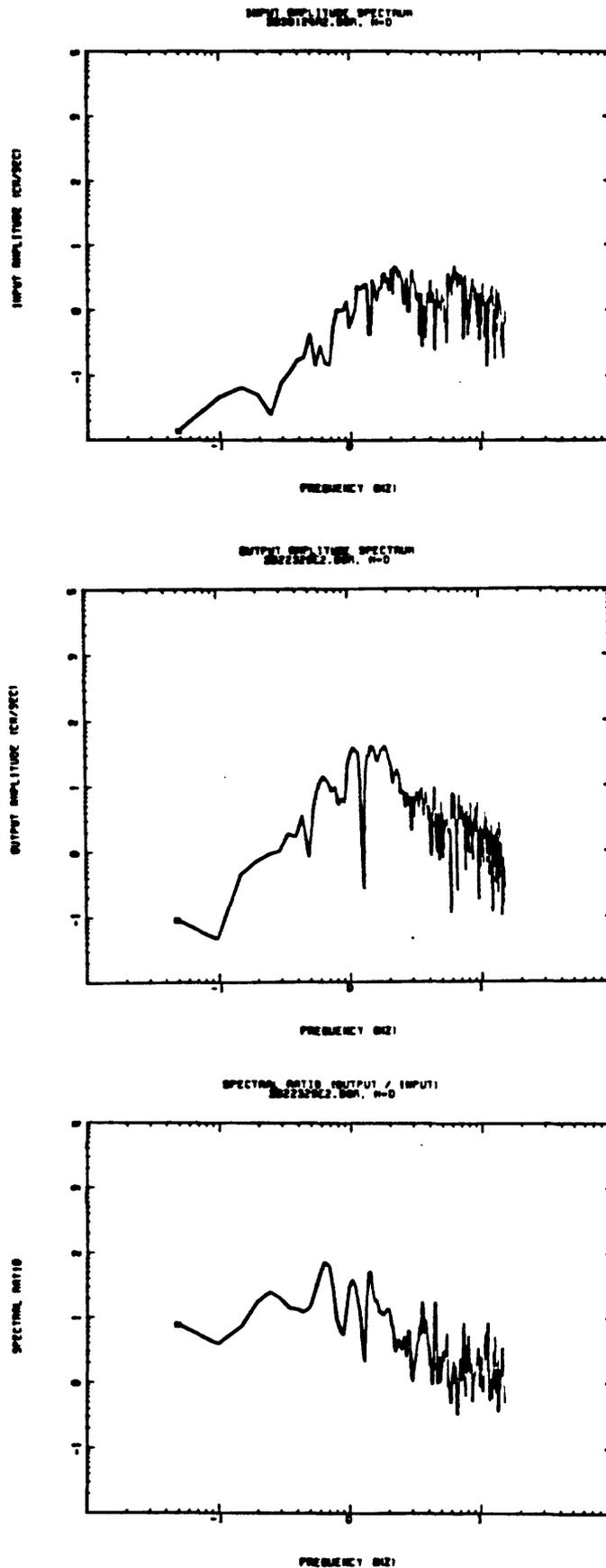


FIG. 13.8.2- Input (small earthquake) and output (large earthquake) Fourier amplitude spectra and the smoothed spectral ratio (large earthquake/small earthquake) for the source scaling example. The records are from the aftershocks of 1983 Borah Peak, Idaho earthquake.

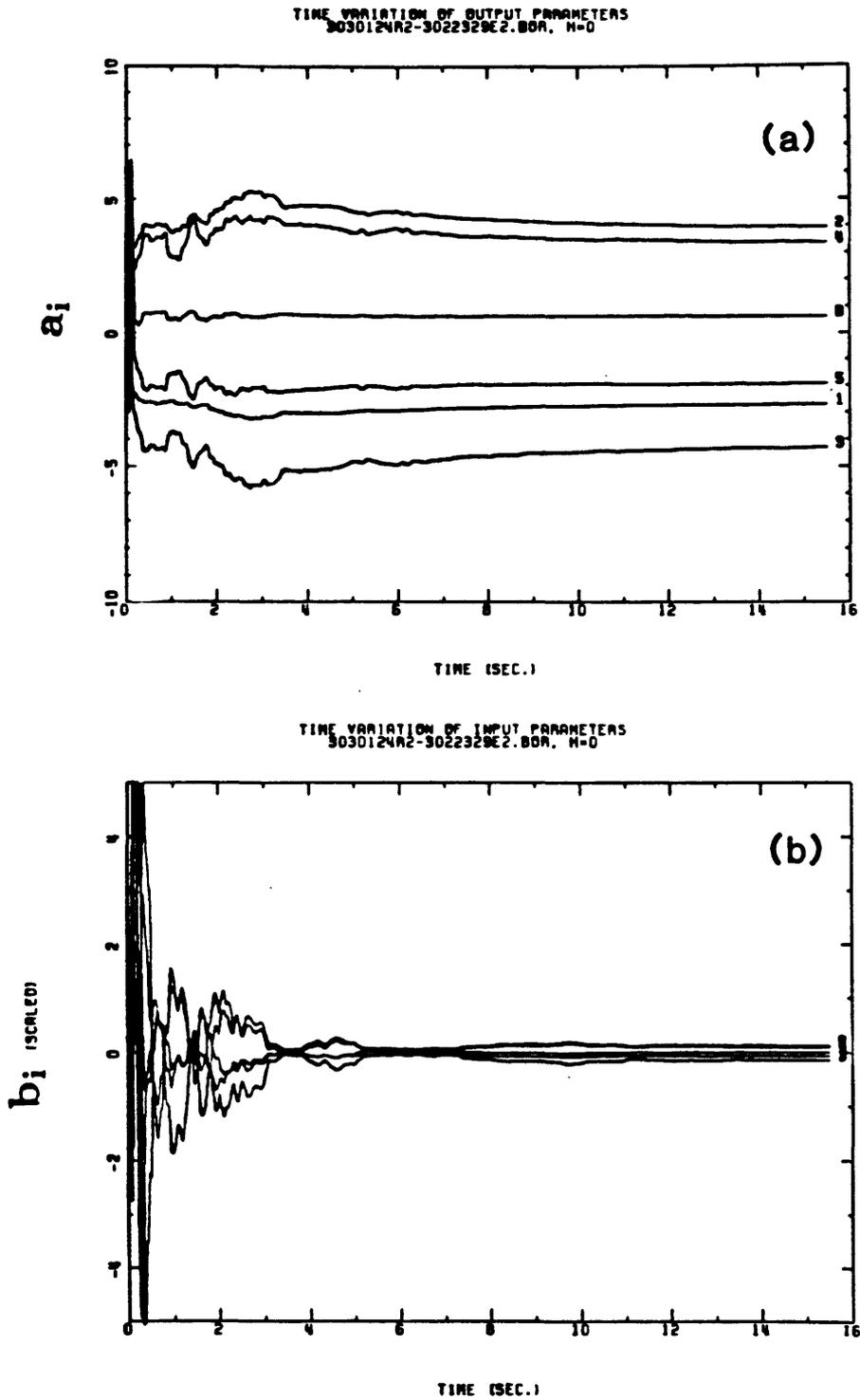
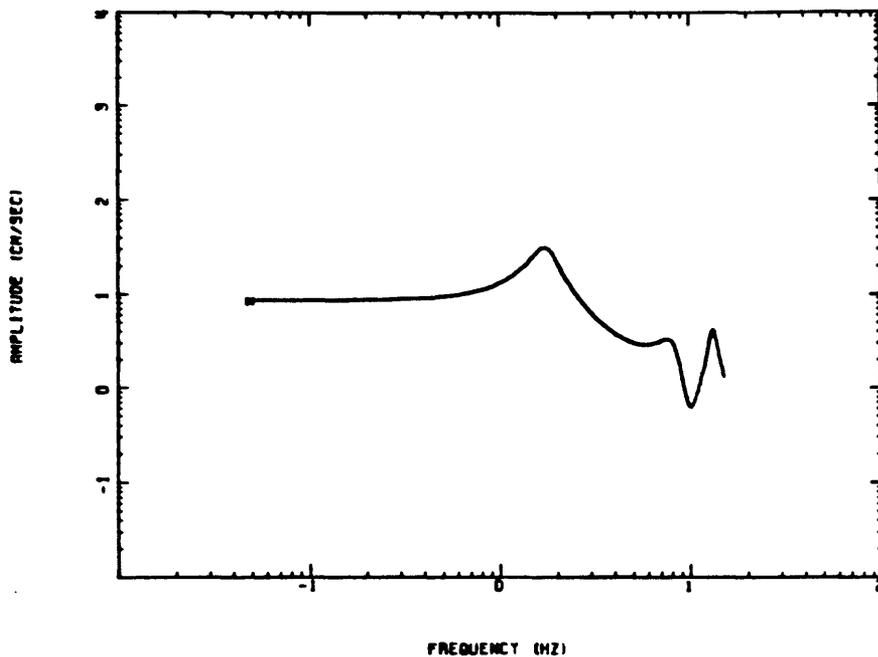


FIG. 13.8.3- Time variation of model parameters for source scaling example (1983 Borah Peak earthquake): (a) $a_j, j = 1 \sim 6$; (b) $b_j, j = 1 \sim 5$.

AMPLITUDE OF FREQUENCY RESPONSE FUNCTION, T= 10.00
3030124A2-3022329E2.00A, H=0



FOURIER SPECTRUM VS. FREQUENCY RESPONSE FUNCTION,
3030124A2-3022329E2.00A, H=0

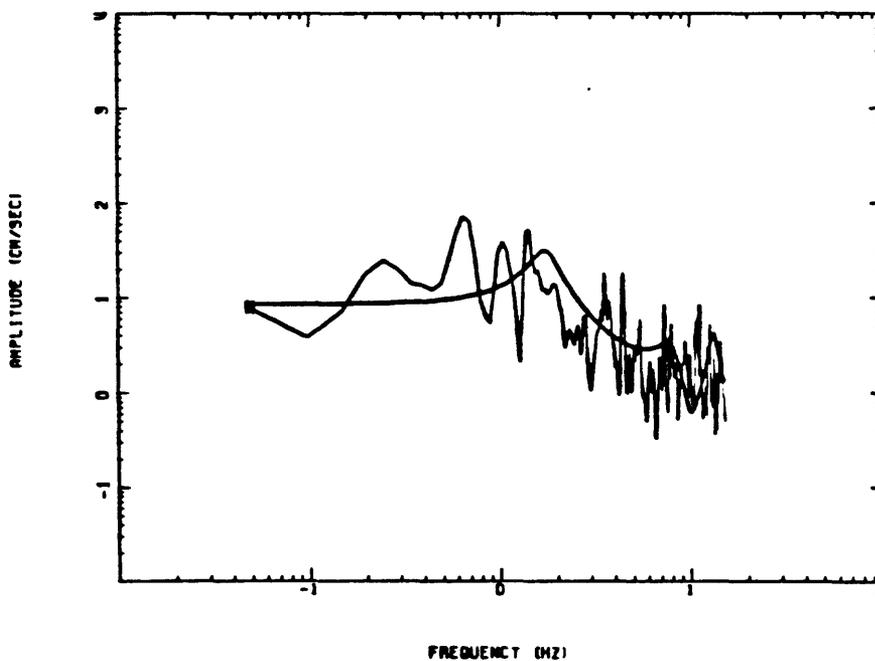


FIG. 13.8.4- The amplitude of the transfer function for (6-5) model for the source scaling example (1983 Borah Peak earthquake) evaluated at $t = 10$ sec., and its match with the Fourier amplitude spectrum.

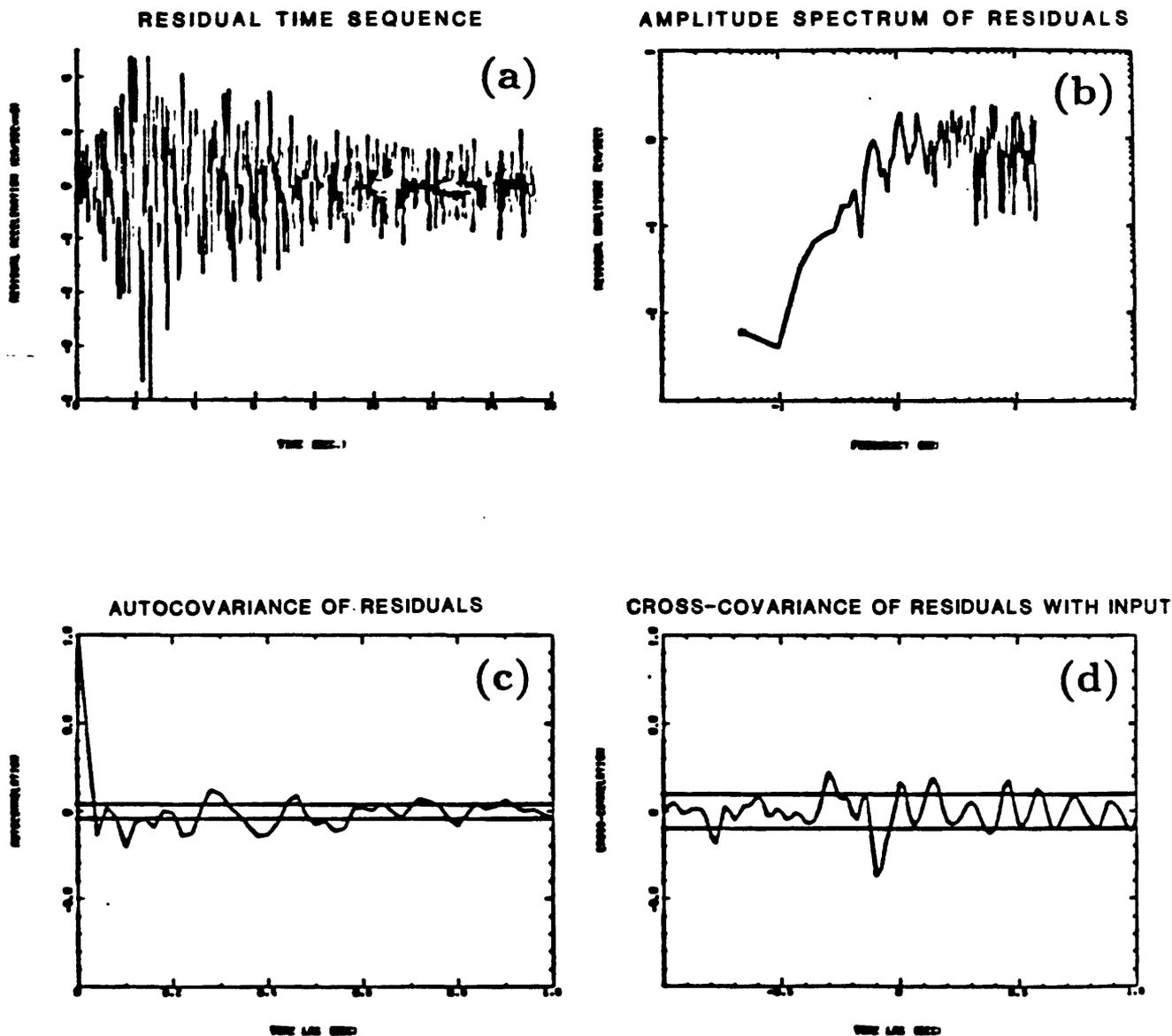
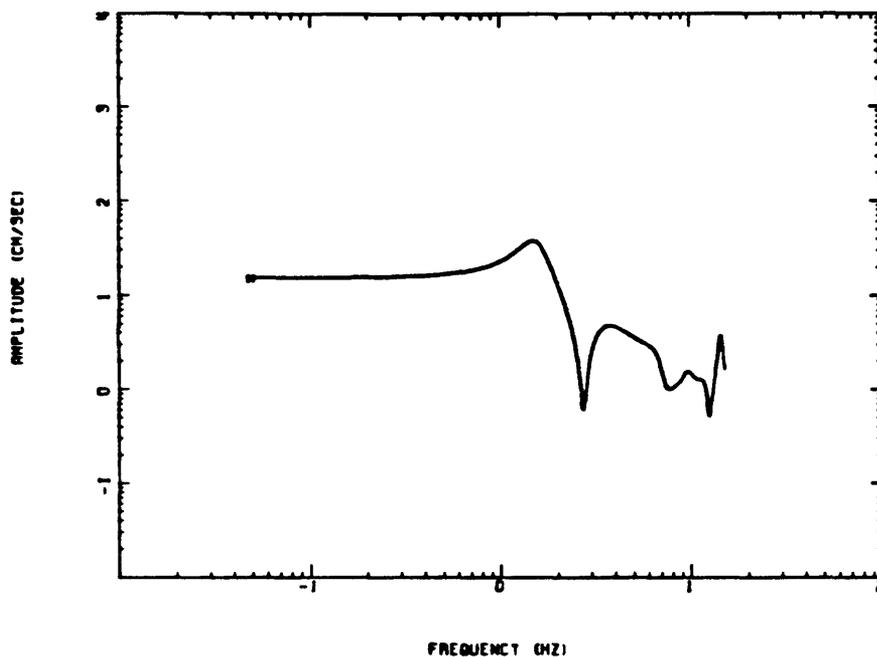


FIG. 13.8.5- The residual tests for (6-5) model for the source scaling example (1983 Borah Peak earthquake) evaluated at $t = 10$ sec.: (a) time history of the residual; (b) Fourier amplitude spectrum of the residuals; (c) autocovariance function of the residuals, and the 90-percentile confidence limits for whiteness according to the Chi-square χ^2 test; (d) cross-covariance function of the residuals with input, and the 90-percentile confidence limits according to (0,1) normal distribution law.

AMPLITUDE OF FREQUENCY RESPONSE FUNCTION, T= 14.00
3030124A2-3022329E2.BOR. M=0



FOURIER SPECTRUM VS. FREQUENCY RESPONSE FUNCTION,
3030124A2-3022329E2.BOR. M=0

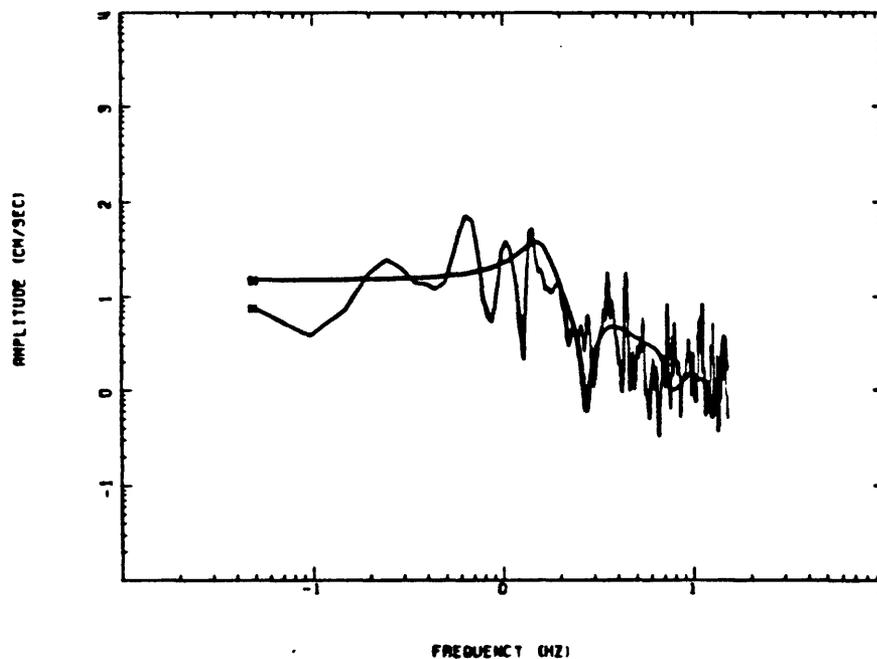


FIG. 13.8.6- The amplitude of the transfer function for (14-13) model for the source scaling example (1983 Borah Peak earthquake) evaluated at $t = 10$ sec., and its match with the Fourier amplitude spectrum.

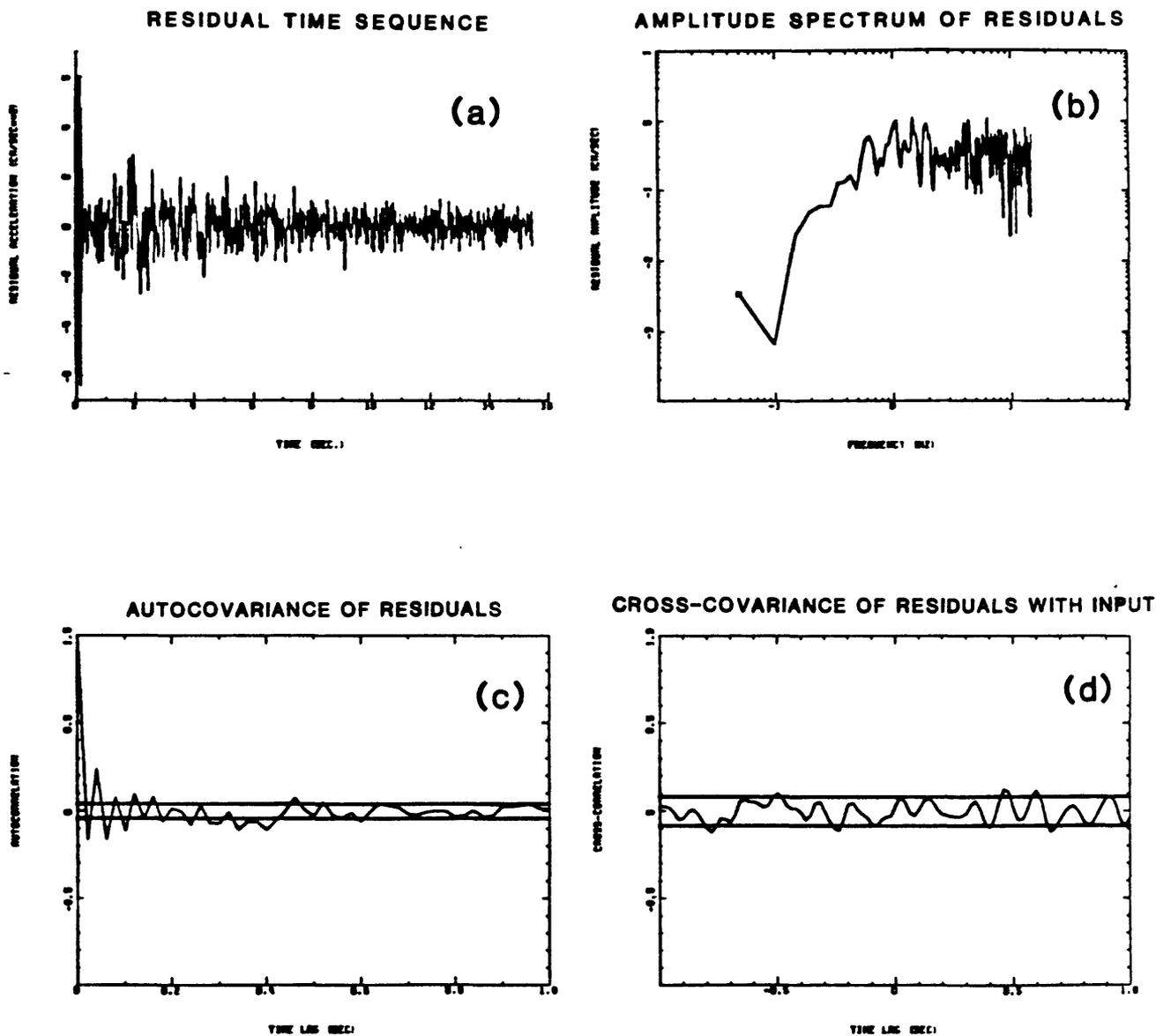


FIG. 13.8.7- The residual tests for (14-13) model for the source scaling example (1983 Borah Peak earthquake) evaluated at $t = 10$ sec.: (a) time history of the residual; (b) Fourier amplitude spectrum of the residuals; (c) autocovariance function of the residuals, and the 90-percentile confidence limits for whiteness according to the Chi-square χ^2 test; (d) cross-covariance function of the residuals with input, and the 90-percentile confidence limits according to (0,1) normal distribution law.

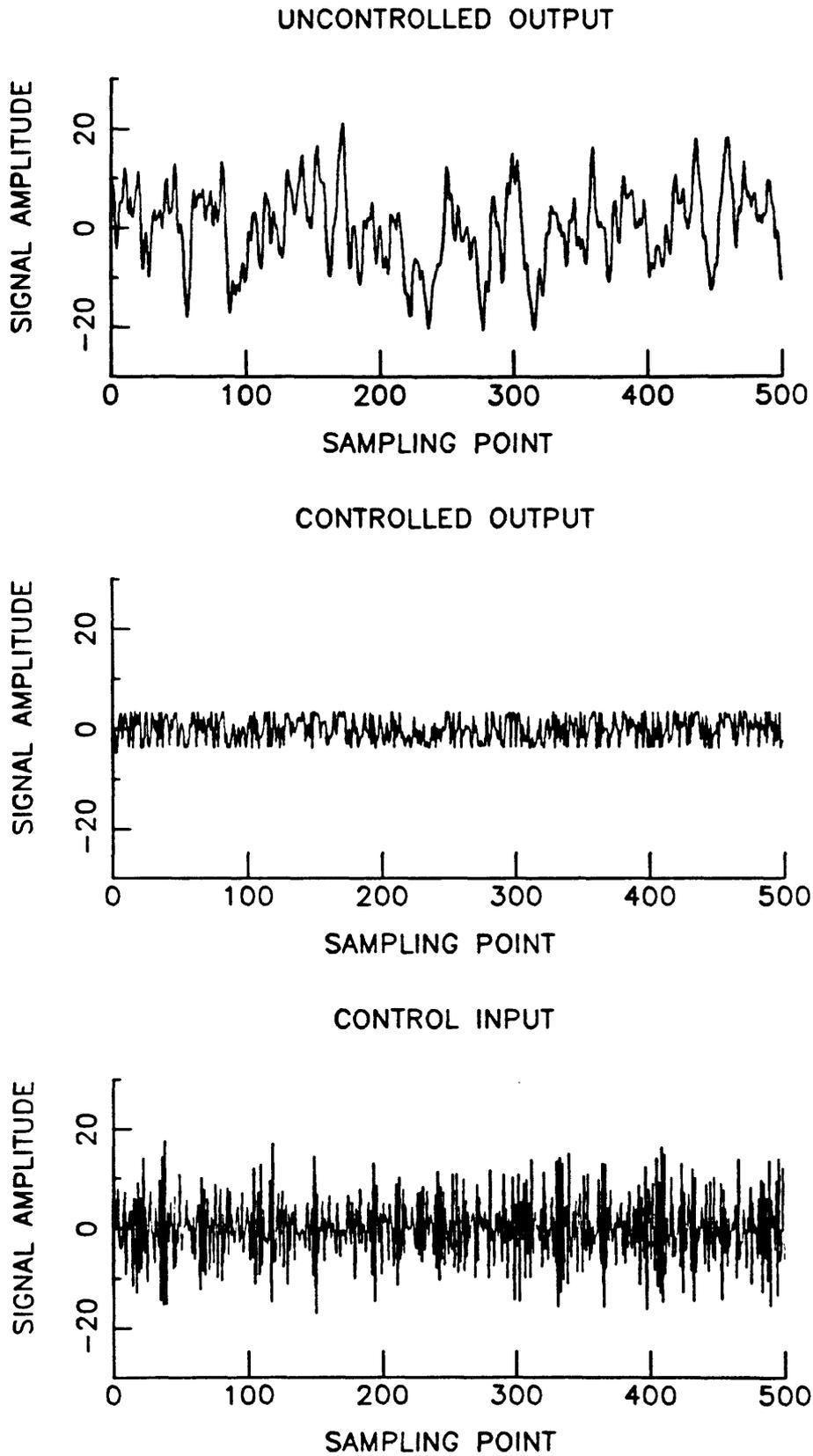


FIG. 13.9.1- Simulated uncontrolled output, controlled output, and control input (known system).

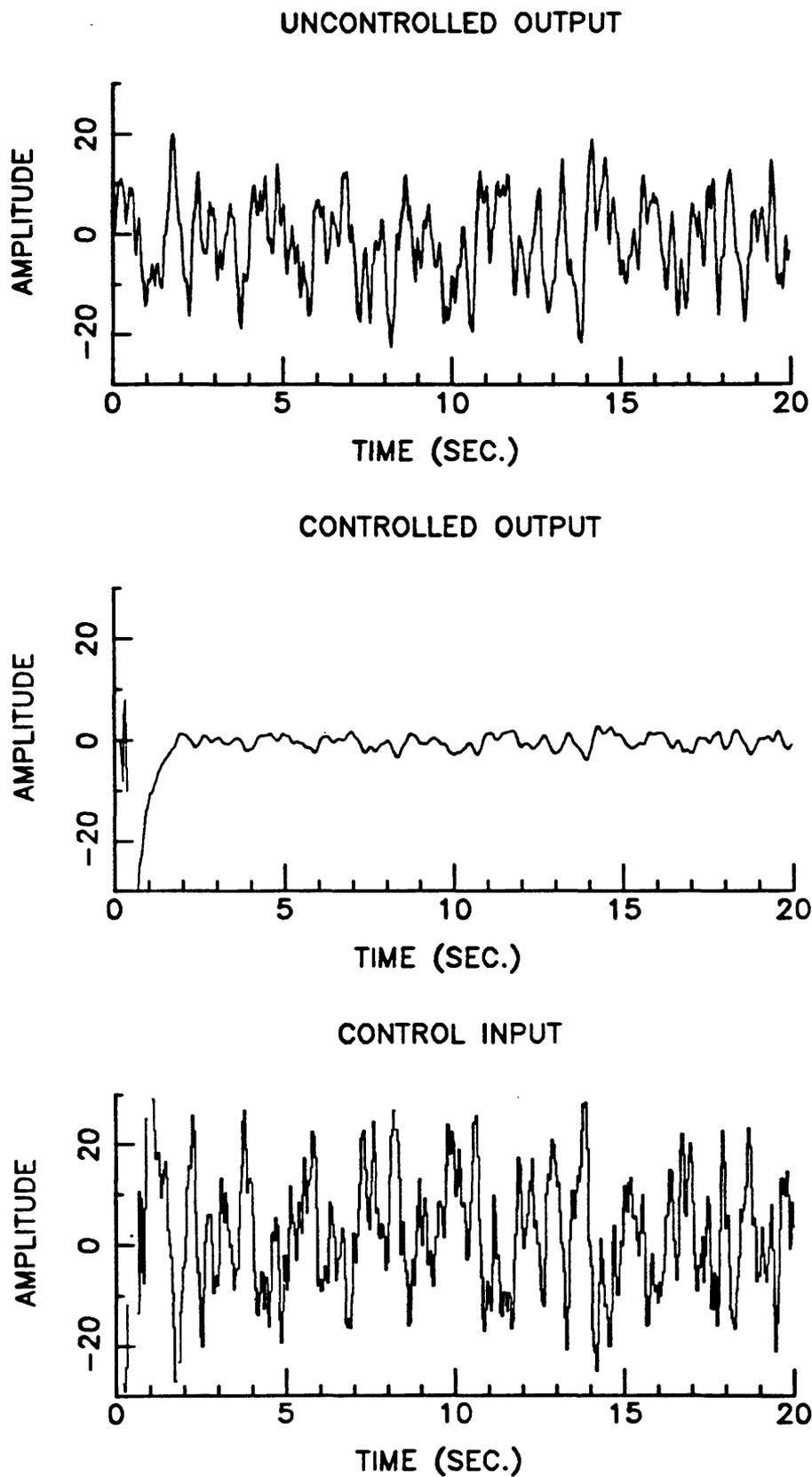


FIG. 13.9.2- Simulated uncontrolled output, controlled output, and control input (unknown system, ARMAX model).

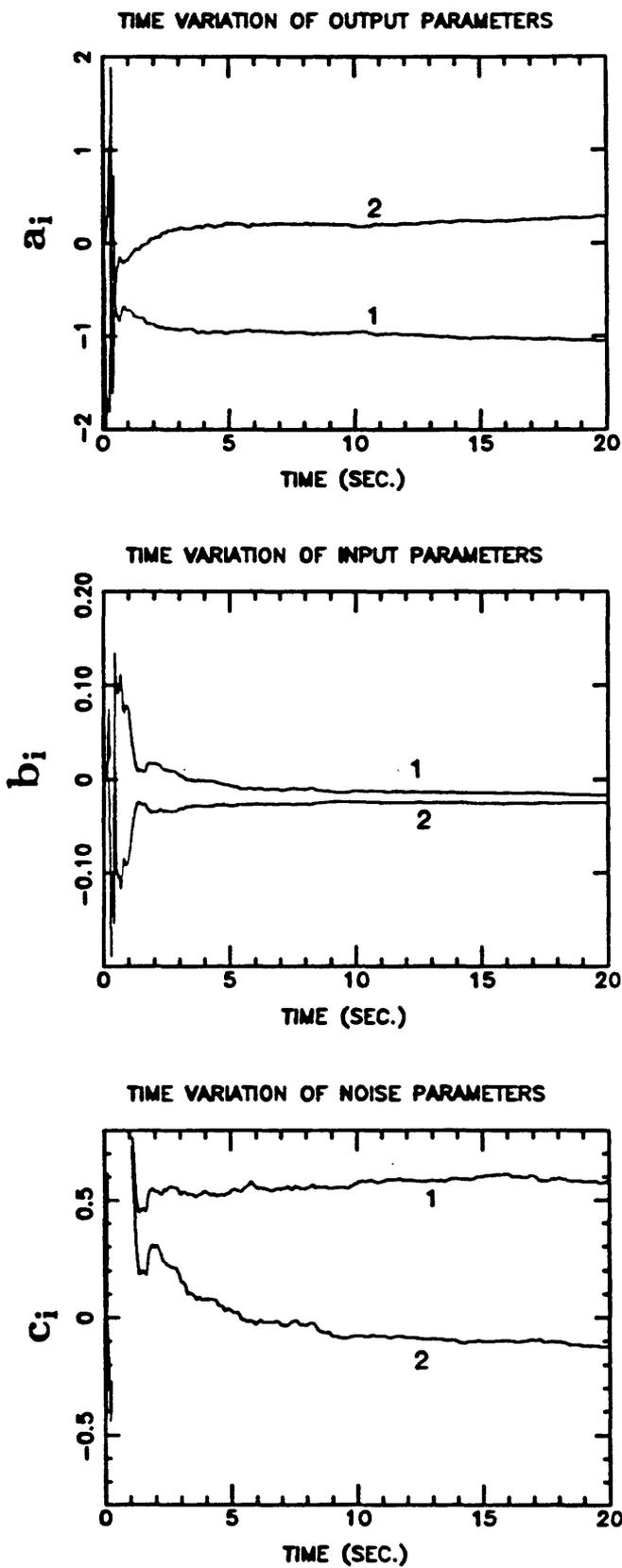


FIG. 13.9.3- Time variation of model parameters (ARMAX model).

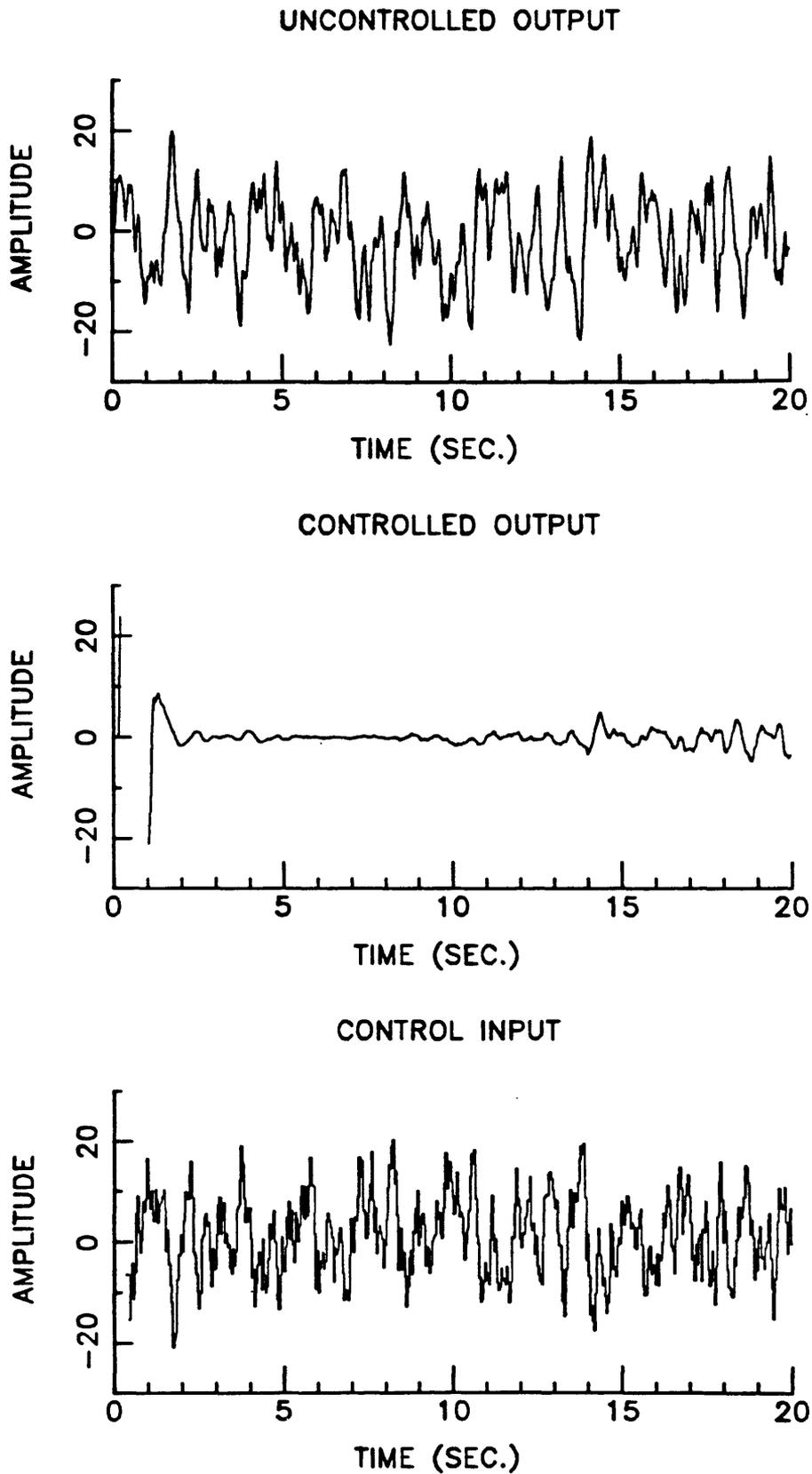


FIG. 13.9.4- Simulated uncontrolled output, controlled output, and control input (unknown system, ARX model).

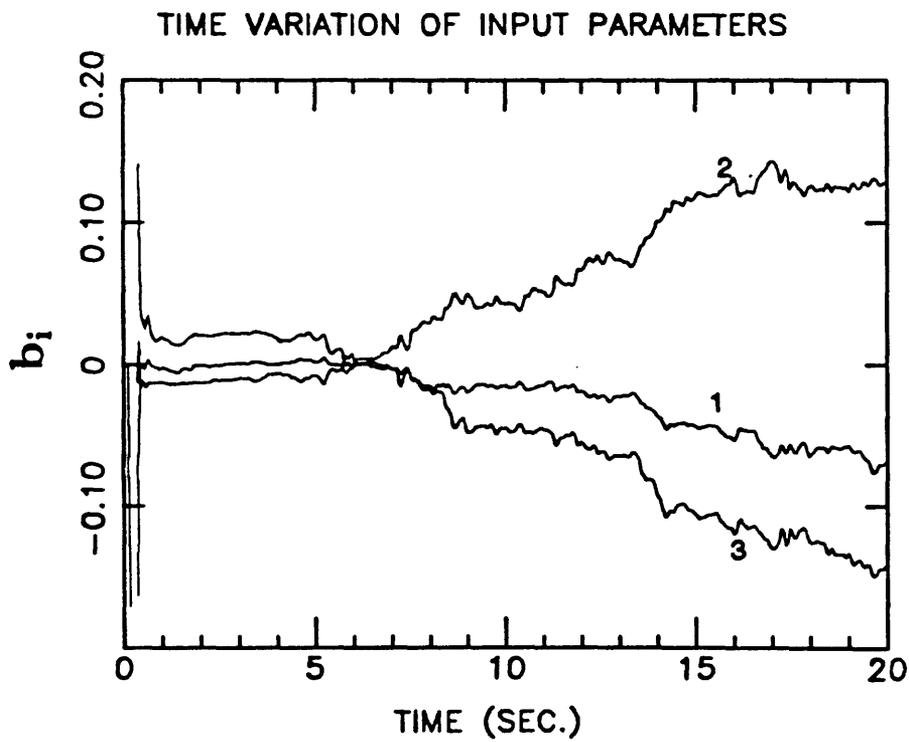
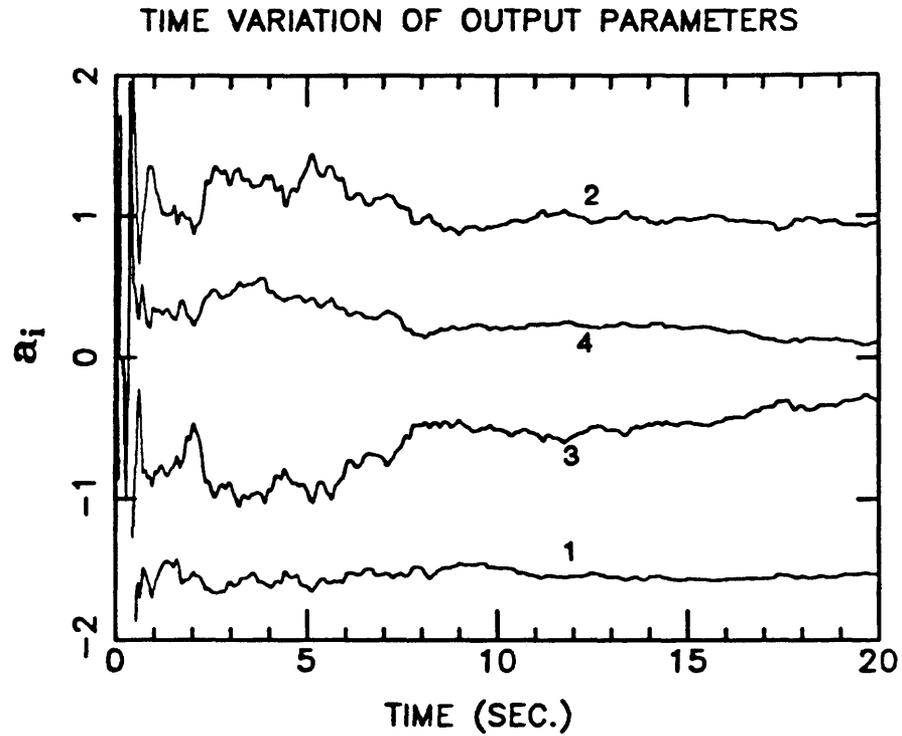


FIG. 13.9.5- Time variation of model parameters (ARX model).

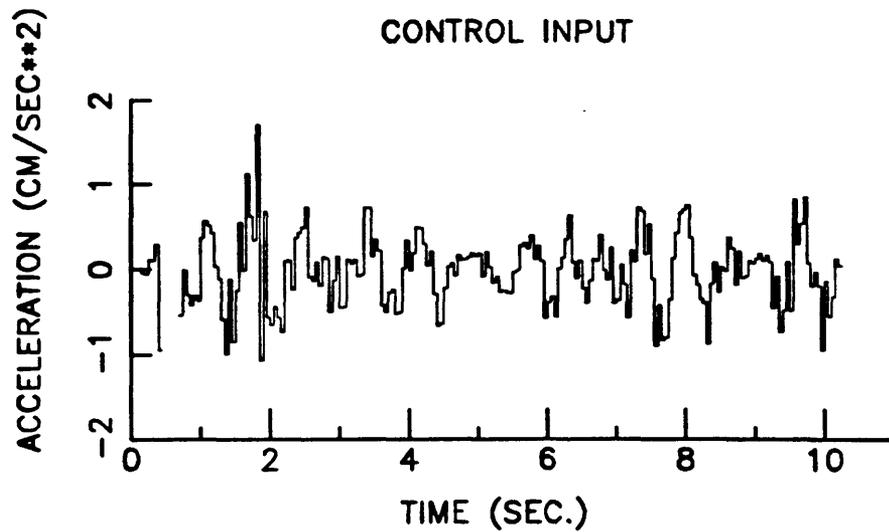
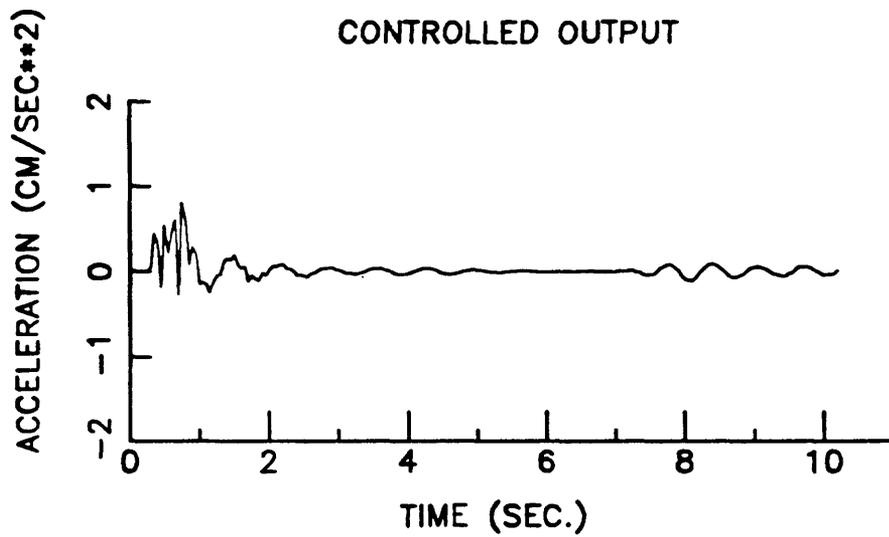
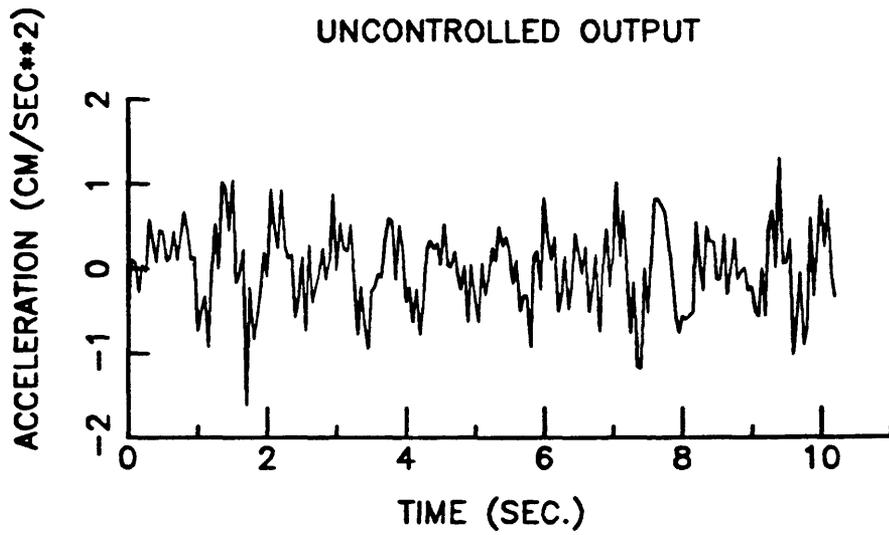


FIG. 13.10.1- Recorded (uncontrolled) top-floor acceleration, controlled top-floor acceleration, and control input.

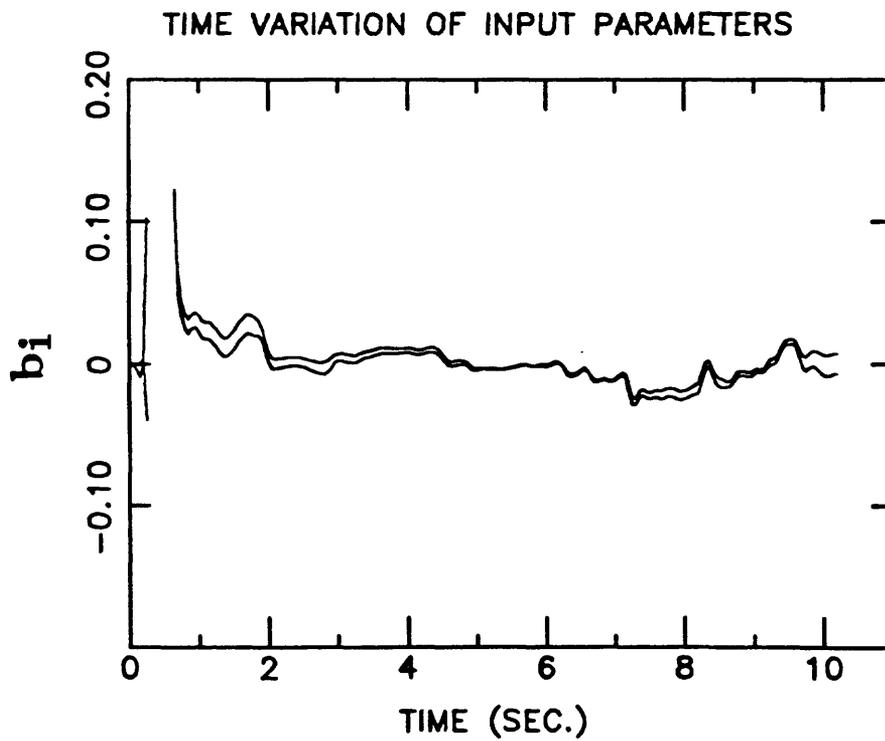
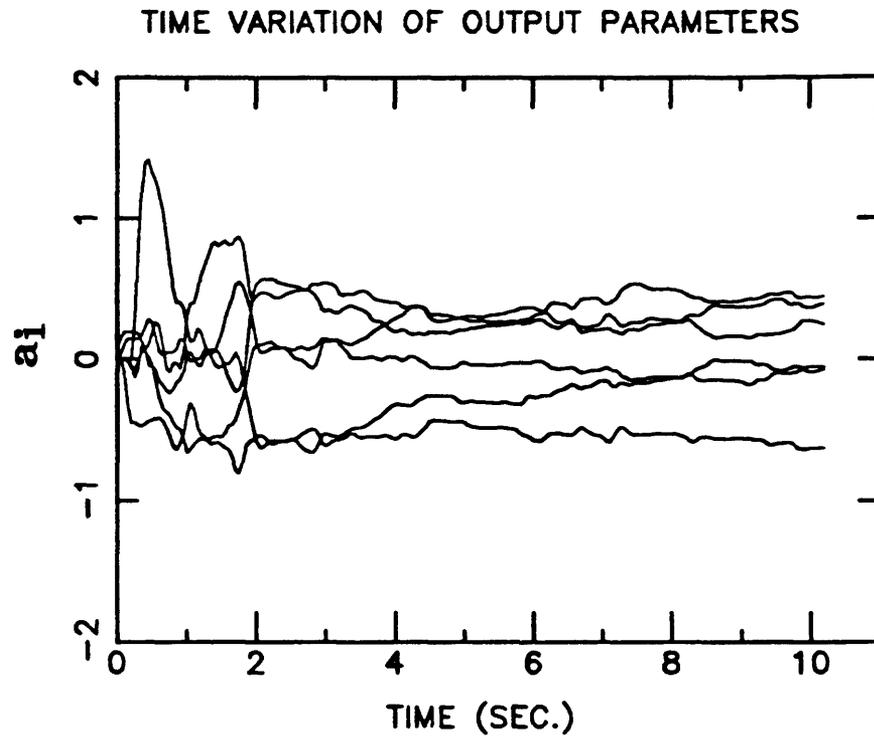


FIG. 13.10.2- Time variation of model parameters (ARX model).