HOKKAIDO UNIVERSITY

Title	A Study of Detection and Extraction Methods for Microearthquake Waves by Autoregressive Models
Author(s)	TAKANAMI, Tetsuo
Citation	Journal of the Faculty of Science, Hokkaido University. Series 7, Geophysics, 9(1), 67-196
Issue Date	1991-03-25
Doc URL	http://hdl.handle.net/2115/8782
Type	bulletin (article)
File Information	9(1)_p67-196.pdf

Instructions for use

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Jour. Fac. Sci., Hokkaido Univ., Ser. VII (Geophysics), Vol. 9, No. 1, 67-196, 1991.

A Study of Detection and Extraction Methods for Microearthquake Waves by Autoregressive Models

Tetsuo Takanami

Urakawa Seismological Observatory, Faculty of Science, Hokkaido University, Sapporo, 060, Japan

(Received October 5, 1990)

Abstract

In this paper, we study the methodology for the objective determination of P and S-arrival times and for the extraction of necessary information from the observed microearthquake waves. Computationary efficient and reliable computer programs have been developed for the seismic signal analysis of data obtained by Hokkaido University Earthquake Recording System. The principal subjects in this paper are as follows.

(1) We first review the conventional statistical methods (Orbit spectral analysis, Moving window analysis, and REMODE filter analysis) for the discriminations of seismic waves and test them by applying to the microearthquake waves.

(2) We secondly apply the univariate autoregressive (AR) model to determine the onset of P and S-waves of microearthquakes by using the minimum AIC (Akaike Information Criterion) estimations and develop an efficient new algorithm based on the adoption of Householder transformation for least squares computation in the minimum AIC calculation. This algorithm provides a very simple and fast procedure for handling additional observations, viz. it is useful for the on-line fitting of locally stationary AR models. The onset time of P-wave can be obtained by applying the univariate AR model to only vertical component seismogram.

(3) It is usually difficult to determine the onset time of S-wave by single component seismogram. We thirdly develop a three dimensional AR model for determining the onset time of S-wave by using one vertical and two horizontal components seismograms. It is based on the extension of the algorithm to three variate AR modeling. Here we also propose a simple and easy method to determine the onset time of S-wave by using the minimum AIC value which is obtained by summing up the three components AIC data sets. This algorithm is useful for on-line processing.

(4) Finally, we develop a model for decomposition of time series into several components. In the model each component is expressed by the AR model. The crucial problem of estimating changing variance of the model has been solved by the techniques of piece-wise modeling and local modeling.

The extraction of microearthquake signal from noisy data, and the decomposition of P and S-wave signals have been shown to exemplify the power of the proposed procedure. The stability of the procedure and a possible simplification of the procedure are also considered by using the same data set.

Table of contents

1. Introduction

The microearthquake data collected by a network of seismographs such as the Hokkaido array operated by RCEP (The Research Center for Earthquake Prediction of Hokkaido University) may contain a wealth of information about the tectonic process and structure of the earth's crust under the array.

Microearthquake data are currently being used in earthquake prediction, locating earthquakes and monitoring seismically active areas. The use of microearthquakes in earthquake prediction is based on the idea that large earthquakes share the same tectonic causes as the numerous small ones occurring in the same general area (Aki, 1968; Takagi, 1982).

Precursory variations in a large number of geophysical parameters are currently being studied. Variations in V_p/V_s or P wave travel time residuals (Semenov, 1969; Aggarrwal et al., 1973; Ohtake, 1973; Utsu, 1973; Robinson et al., 1974; Iizuka, 1979; Yoshii, 1982), epicenter recurrence patterns, anomalous seismicity (Kerr, 1978; Motoya, 1984; Matu'ura, 1986), and temporal changes in fault plane solutions (Nakajima, 1974) may be particularly effective in an earthquake prediction program (Rikitake, 1976; Mogi, 1977; Lindh, et al., 1978; Kerr, 1978; Rikitake, 1979; Asada, 1982). Detailed fault zone structure can be determined from the location of microearthquakes (Eaton et al., 1970). Tectonic stress indicators such as fault plane solutions and stress drops can be determined from first motion and spectra studies (McNally and McEvilly, 1977; Aki, 1969, 1968; Brune, 1970; Nakajima, 1974).

Microearthquakes can be also used in three dimensional seismology (study on three-dimensional seismic-velocity structure) to determine structural details around an active zone and to map magma chambers and geothermal areas. In three dimensional seismology, the arrival times from a large number of stations are used to determine the fine scale structure of a small area. The ultimate resolution of three dimensional seismology deqends on the wave-length of the first arrival, which may be about a few hundred meters for microearthquakes. In many experiments (e.g. studies in Hokkaido, Sakata et al., 1981; Takanami, 1982; Miyamachi and Moriya, 1987; Nakanishi, 1989), their resolutions are limited by the number of stations. A higher density of stations and the accurate arrival times from the stations would increase the resolution.

The full exploitation of these data for the purpose of earthquake prediction has been hampered by a lack of a strict, uniform procedure for analyzing microearthquake data. For example, the bias and the error in elementary measurements such as picking the times of first arrivals by eye may vary from

time to time because of a change in personnel engaged in the work. Workers in statistical seismology are often dismayed by the non-uniformity of the data set in time, when they are studying epicenter migrations and the changes in b-value. Unfortunately there is an evidence that one anomalous P-delay attributed to dilatancy may be due in part to subjective bias of the personnel who read the records (Lindh et al., 1978 ; Okada, 1980).

It is impossible to avoid subjective bias in reading seismograms, some people tend to pick out arrival time earlier than the real one, others tend to pick later ones (Freeman, 1966a, 1966b, 1968). It is also impossible to get any objective estimate of the errors involved. This is important in first moiton studies. It is shown that the wrong sense of first motion is often picked if the signal to noise ratio is lower than a critical value (Aki, 1976 ; Pearce and Barley, 1977).

Earthquakes occurring within a few hundred kilometers from a seismic network are called local events. Local earthquakes are often characterized by

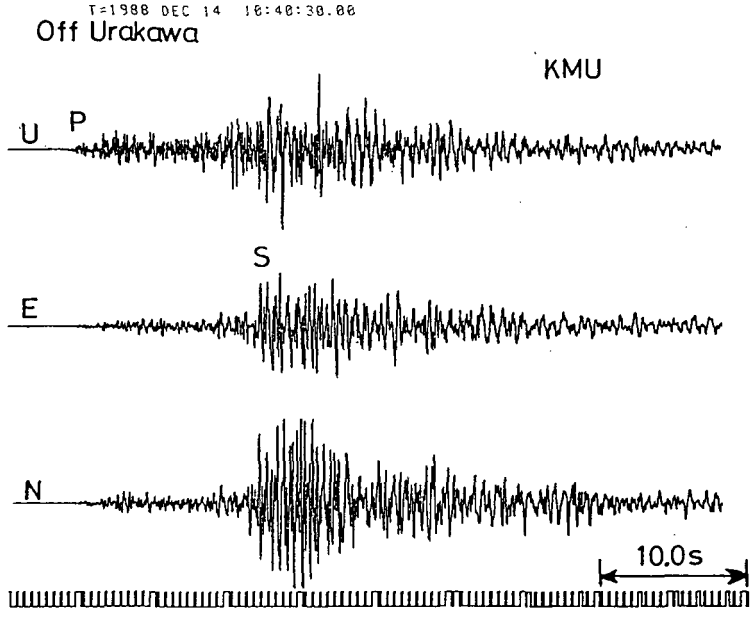


Fig. 1.1. Example of seismic signals recorded by the short-period, three-components seismographs of station KMU of the RCEP, Hokkaido University Microearthquake Network. The event occurred at 41.43°N, 142.34°E, at 10 : 40, on Dec. 14, 1988. The depth is 41 km, and magnitude is 4.1. The time scale is shown by the last trace.

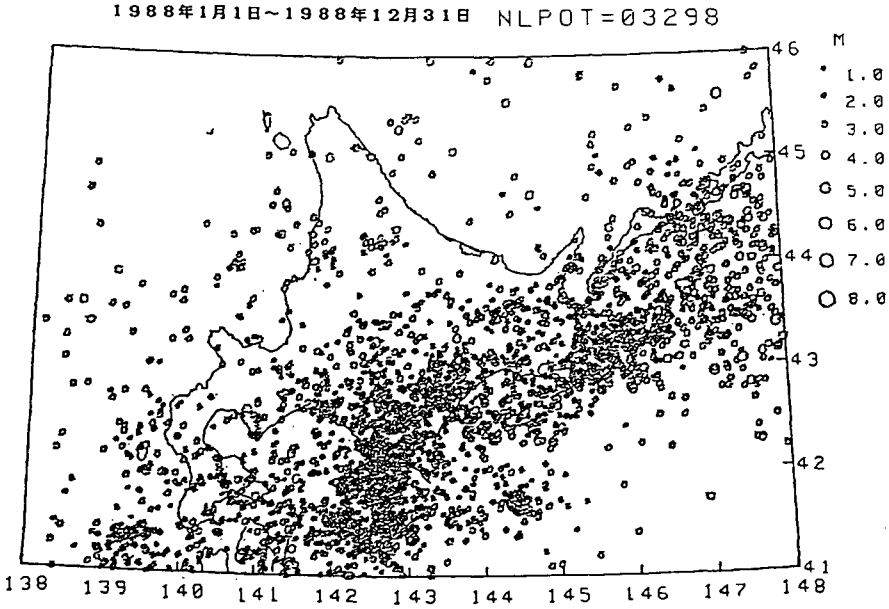


Fig. 1.2. Map showing earthquake epicenters (for 1988) as determined by the RCEP Hokkaido University Microearthquake Network. The events of 3298 are located on the map.

high frequency waves as shown in Fig. 1.1. On the other hand, the transient signal of instrumental or cultural origin often has the same predominant period throughout its duration. Therefore, for a weak microearthquake it is difficult to determine the onset times of P-waves and/or S-waves.

Microearthquake networks are designed to record as many earthquakes as nature and instruments permit, and they produce an immense volume of data. To fully utilize the information obtained from these microearthquake networks, automatic processing of the massive data can be a serious problem to be solved. For instance, the RCEP network detects thousands of microearthquakes per year (see Fig. 1.2). Therefore, the number of P and S-phases should reach several ten thousands. Specifically, during one year it records several ten thousands of seismograms from thousands of events. Even if seismograms can be processed manually at this rate, earthquake swarms and aftershock sequences can increase the seismicity level of an area by an order of magnitude. Rapid, objective analysis of such a volume of data could be crucial for earthquake prediction. Further, since the microearthquake network is required to determine earthquake hypocenters precisely, it is crucial to develop an automat-

ed system that can determine the first P-arrival times accurately and efficiently.

The waveforms of microearthquakes recorded at neighboring stations have been observed to be remarkably dissimilar in general. This has led to computer algorithms that process each incoming digital seismogram as an independent time series.

In this paper, the methodology of the objective determination of arrival times and the extraction of necessary information from the observed seismograms is studied. Computationary efficient and reliable computer programs are developed for the analysis of seismic data obtained by Hokkaido University earthquake recording system.

The main points of the paper are summarized as follows. There is always ambiguity associated with measuring first arrival time from seismograms whether it is done by a seismologist or by a machine since the seismic signals are of unknown shape and are contaminated with noise. This ambiguity increased with distance from the epicenter because of structure and attenuation of the earth.

In order to develop the earthquake processing for reducing the ambiguity, some statistical techniques are studied using the seismograms of microearthquakes. The process of detecting arrivals should be efficient in real time. Since the seismogram is a time series, we can introduce a statistical method, autoregressive (AR) modeling, which has been mainly developed in the field of signal processing. It is an efficient procedure for determining the onset times of P and S-arrivals in on-line system. We will estimate the arrival time accuracy by applying the method to many seismograms recorded by the RCEP network. The AR modeling is based on the parametrizing the spectral contents of seismic signals and/or noises. We shall later propose an extended AR modeling for the decomposition of noisy seismograms into seismic signals and unwanted noises or into P, S-waves, and noises.

Before going on to the developments of methodologies described above,we will review some well known techniques to process intricately seismic waves, that is, a visual method for particle motion trajectories of P and S-waves and a statistical analysis of three components seismograms. As for the statistical analysis, we will review and check two kinds of techniques, i.e. moving window analysis and rectilinear motion detector (REMODE filter analysis). We will consider their usefulness in the area of automatic signal processing for determining onset times of P and S arrivals. We will also review the automatic signal processings, which are currently being used in various seismological observations.

This peper is organized in twelve major chapters. Chapter 1 provides an introduction to automatic processing for microearthquakes. It begins with describing the important properties inherent in microearthquake data. It also deals with the problems in actual microearthquake data processing and with the intrinsic value in this data for rapid earthquake prediction.

Chapter 2 explains the properties of microearthquake data and shows the highly sensitive seismographic stations of Hokkaido University Seismic Network.

Chapter 3 reviews the automatic processings, which are currently used, to determine the onset times of P-arrivals. In this chapter, we also take up an autoregressive modeling for non-stationary seismic waves.

Chapter 4 describes the most general mathematical statement of the properties of seismic wave propagation in an isotropic perfectly elastic homogeneous medium, and shows how the polarization techniques based on statistical analysis can really manipulate the properties of seismic wave propagation, that is, this chapter is also concerned with the usefulness of the polarization techniques applying to the seismograms recorded at Hokkaido University Seismic Network. This chapter is further divided into sections covering different polarization techniques for particle motion analysis, moving window analysis, and REMODE filtering analysis.

Chapter 5 is concerned with the mathematical aspect of probability model for general time series. In this chapter, we review the stochastic process, the stationary process, the autoregressive process, the non-stationary time series, the on-line type model fitting procedure, the efficient least squares computation by Householder transformation, and the minimum AIC estimation of locally stationary models.

Chapter 6 extends the idea described in chapter 5 to the automatic deter mination of P-arrival times. In this chapter we develop the algorithm for the least squares computation by the Householder transformation to reduce the CPU-time by one to several hundreds. And it describes the successful adaptation of locally stationary modeling to the determinations of the onset times of P-aves with several kinds of noises.

In the chapter 7, we turn our attention to the 3-dimensional time series model. In particular, we develop the method for 3-dimensional locally stationary AR model fitting to determine onset times of S-waves.

Chapter 8 deals with the application of methods developed in the previous chapters to the seismogams recorded at Hokkaido University Seismic Network. In this chapter, we also test the distributions for residuals between onset times

read by a seismologist and those determined by the on-line processing developed in the previous chapters.

Chapters 9 and 10 are addressed to the problem of decomposition of a time series into several components. In chapter 9, we propose to use autoregressive model for each of the background noise and earthquake signal. Chapter 10 deals with the develoment of decomposition into the background noise, P-wave, and S-wave. Chapter 9 and 10 provie the state space model, the Kalman filter, and smoother algorithms for their decompositions.

Finally, chapter 11 summarizes the principal results obtained throughout this study.

2. Microearthquakes and microearthquake network

This chapter reviews some principles and applications of microearthquake network. This reviews heavily draw on our experience with the Microearthquake Network of RCEP, Hokkaido University.

Earthquakes of magnitudes less than 3 are generally referred to as microearthquakes. In order to extend the range of application of seismological studies to microearthquakes, it is necessary to have a network of closely spaced and highly sensitive seismographic stations. Such a network is usually called a microearthquake network. From this network we can obtain large amounts of information about the seismicity because of a high occurrence rate of microearthquakes and the large number of recording stations. One of the important applications is the monitoring the seismicity for the purpose of the earthquake prediction.

Historically, microearthquake networks evolved from temporamy expedition that studied aftershocks of large earthquakes to reconnaissance surveys, and finally to permanent telemetered networks. Methodologies and techniques have been developed to study microearthquakes (more or less independently) by various groups.

Gutenberg and Richter (1941) found that there is a relation between the magnitude M of earthquakes and their frequancy of occurrence N,

$$\log N = a - bM, \tag{2.1}$$

where a and b are constants. Many studies of earthquake statistics have shown that Eq. 2.1 holds and that the value of b is approximately 1. This means the number of earthquakes increases by tenfold as the magnitude decreases by one. Therefore, the smaller the magnitude of earthquakes one can record, the more

the information about the seismicity one can collect. After the introduction of Richter's magnitude scale, it has been conveniently used to classify earthquakes more definitely to avoid ambiguity. A common classification of earthquakes according to magnitude is as follow (Hagiwara, 1964),

Magnitude	Classification
$M \geqq 7$	Major earthquake
$5 \leqq M < 7$	Moderate earthquake
$3 \leqq M < 5$	Small earthquake
$1 \leqq M < 3$	Microearthquake
$M < 1$	Ultra-microearthquake

However, all earthquakes of magnitude less than 3 are frequently called micro earthquakes, and the term Ultra-microearthquake is seldom used.

2.1 Hokkaido University Seismic Network

The regional seismic network deployed in Hokkaido by the Research Center for Earthquake Prediction of Hokkaido University (RCEP) is designed to study small and microearthquakes. The dynamic range of the seismic recording system is only 54 dB. Therefore, if an earthquake of magnitude larger than the designed magnitude level of the network occurs within the network area, the saturation causes loss of information on the precise behavior of seismograms. To attain a broad dynamic range of seismic recording, two kinds of observation systems which consist of one vertical component geophone of a natural frequency of 2 Hz and three components (vertical, north-south and east-west) of seismographs of a natural frequency of 1 Hz, are set at the same station and their output signals are amplified with different amplification factors. Recordings by 2 Hz and 1 Hz seismographs are suited for the microearthquakes and for the regional ones, respectively. In the recording system, ground velocity signals from 2 Hz geophone at each station is first digitized by an 8 bit nonlinear AD converter with 92.3 samples/sec (2400 bps/26 bit), and digitized with 42.15 samples/sec for 1 Hz geophone, then waves with frequency above 30 Hz are eliminated by an anti-aliasing filter (Butterworth filter of order 6), and transformed to pulse code modulation (PCM) data. These PCM data were transmitted to the central recording statiom of Hokkaido University (e.g. Kasahara, 1976; Maeda, 1978) and recorded on a magnetic tape by a high density data recorder (HDDR of Yamatake Honeywell Co.) which drives continuously for two days.

Details of the Hokkaido University Seismic Network are given in the following section.

2.2 Network recording system

We will review some factors that should be considered in laying out a microearthquake network. In all cases, we assume that the primary objective of the network is to locate earthquakes that occur within the network area and to estimate their magnitudes. At RCEP, we have been developing the network in the whole of Hokkaido for the past two decades (Fig. 2.1)

In 1967, the Urakawa Seismological Observatory installed a tripartite network with a span of several hundreds meters for recording microearthquakes at Kamikineusu, Urakawa, southern Hokkaido (Motoya, 1969). In 1976, the RCEP network was organized to telemeter seismic signals from 9 field stations to a recording and processing center of Hokkaido University, Sapporo (Kasahara, 1976; Maeda, 1978).

Telemetry transmission and centralized timing and recording systems are

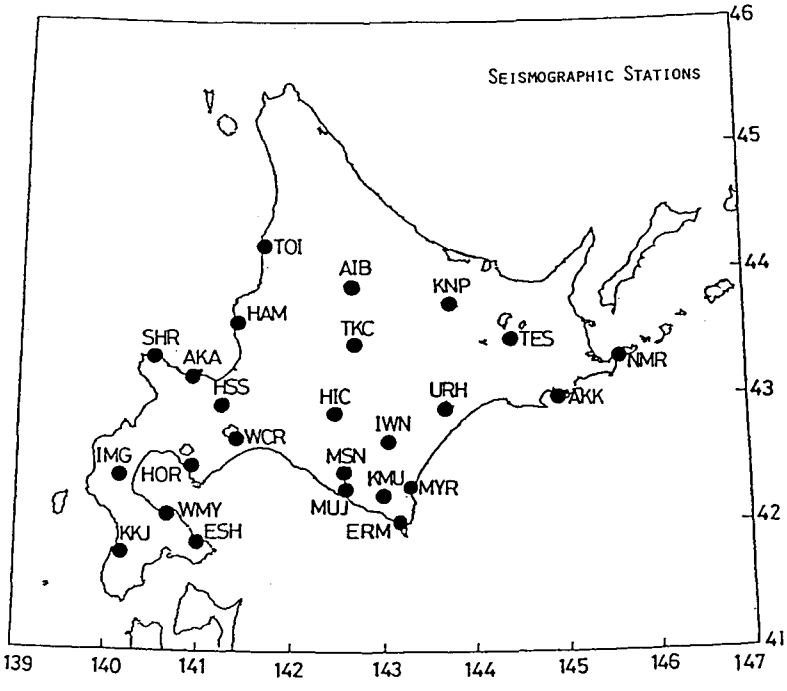


Fig. 2.1. Map showing seismographic stations in Hokkaido. These were the stations as of Dec. 31, 1988.

commonly used for operating microearthquake networks. In this method, the output signal from a seismometer is conditioned and amplified at the field station, and properly converted for the telemetry transmission to the central location. At this center, signals from many field stations and a common time base are recorded simultaneously. The recording is made on magnetic tape and paper chart. The common time base eliminates the need for individual clock correction for each station.

By early 1989, the network had grown to 23 stations. Although the basic idea of telemetry transmission and recording has not changed, the instrumentation has been modified since 1983. For instance sampling rate of 1Hz geophoneh was changed from 92.3 to 42.15 samples/sec. In this paper, we use the seismograms digitized with 92.3 samples/sec.

3. Review of data processing for determining onset times of seismic arrivals

This chapter is concerned with reviewing the data processing to determine onset times of seismic wave arrivals being currently used. Before processing time series, most investigators apply bandpass filtering to reduce noises that are outside the frequency range of interest and to eliminate the dc bias level.

The concept of characteristic function is useful in designing computer algorithms for determining first P-arrival times. The incoming signal in digital form is seldom used directly in the algorithms. It is usually transformed into one or more time series that are more suitable for determining first P-arrival times. One of the simplest characteristic functions is obtained by performing a difference operation on the incoming digital signal. The transformed time series is referred to here as the characteristic function $f(k)$, where k is a time index.

Stewart (1977) defined the following characteristic function. If d_k denotes the difference between adjacent data points, i.e., $d_k=X_k-X_{k-1}$, then

$$f(k)\begin{cases} =d_k, & \text{if } g(k)\neq g(k-1) \\ =d_k, & \text{if } g(k)=g(k-1) \text{ and } h(k)=8 \\ =d_k+d_{k-1}, & \text{if } g(k)=g(k-1) \text{ and } h(k)\neq 8 \end{cases} \tag{3.1}$$

where

$$g(k)\begin{cases} =+1, & \text{if } d_k \text{ is positive or zero} \\ =-1, & \text{if } d_k \text{ is negative} \end{cases}$$

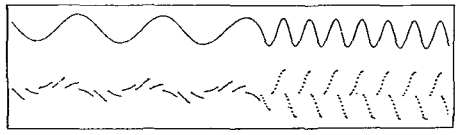


Fig. 3.1. Effect of the characteristic function $f(k)$ of Eq. 3.1. on a 1-Hz and 3-Hz sinusoidal signal. The upper signal is the input function, and the lower signal is the computed characteristic function. Digitizing rate is 50 samples/sec. (after Stewart, 1977).

and

$$h(k)=|\sum_{k-7}^{k} g(k)|. \tag{3.2}$$

This characteristic function $f(k)$ is suitable to detect less impulsive P-wave arrivals. Figure 3.1 shows the behavior of this characteristic function when it is applied to a sinusoidal signal. Using $f(k)$, a moving time average $\overline{f(k)}=|\sum f(k)|$ is computed. If $|f(k)/\overline{f(k)}|$ exceeds 2.9 at a time index $k=j$, then a tentative detection is declared at the time t_j. Within the next 0.5 sec, four additional tests are performed to see if the tentative detection can be confirmed as an event. If so, the first P-arrival time is computed according to $t_j-h(j)\cdot \Delta t$, where Δt is the digitization time interval.

Allen (1978) devised a characteristic function to make it sensitive to variations in signal amplitude and its first derivative. If the amplitude of incoming signal is X_k at time t_k, then the modified signal amplitude is defined by

$$R_k=C_1R_{k-1}+(X_k-X_{k-1}), \tag{3.3}$$

where C_1 is a weighting constant. Allen's characteristic function is defined by

$$f(k)=R_k^2+C_2(X_k-X_{k-1})^2, \tag{3.4}$$

where C_2 is also a weighting constant. Using this characteristic function, a short-term and long-term average are computed by

$$\alpha(k)=\alpha(k-1)+C_3[f(k)-\alpha(k-1)],$$

and

$$\beta(k)=\beta(k-1)+C_4[f(k)+\beta(k-1), \tag{3.5}$$

where C_3 and C_4 are constants with values ranging from 0.2 to 0.8 and from 0.005 to 0.05, respectively. If $\alpha(k)/\beta(k)$ exceeds 5 at a time index $k=j$, then an exent is declared to have occurred at time the t_j. The first P-arrival time is determined by the intersection of the slope of the modified signal amplitude at the time t_j with the zero level of the modified signal. So far, all the characteristic

functions discussed depend on the first difference of the digitized data of incoming signal.

Anderson (1978) proposed a different method of defining a characteristic function. If the amplitude of the incoming signals is X_k at time t_k, then the modified signal amplitude is defined by

$$Y(k)=X_k-\overline{X}_k, \tag{3.6}$$

where $\overline{X}_k$ is a moving time average of X_k and is used to estimate the offset in the incoming signal. A moving time average of the absolute values of the modified signal amplitude is defined by

$$z(k)=(n+1)^{-1}\sum_{i=k-n}^{k}|Y(i)|, \tag{3.7}$$

where the length of the time window is $n\cdot\Delta t$, and Δt is the sampling interval. Anderson (1978) introduced a characteristic function that is obtained from the series $\{t_k, Y_k\}$ by saving only the zero crossing points and those points corresponding to the maximum amplitude (either positive or negative) between two successive zero crossing. This characteristic function removes the higher frequency components of the seismic signal, but preserves the oscillatory character of the signal. It consists of a series of serration, each of which is referred to as a blip. To detect the first P-arrivals, the characteristic function $f(k)$ is examined blip. A blip at a time τ_i is defined by three points $\{\tau_j, b_j\}$, $\{\tau_{j+1}, b_{j+1}\}$, $\{\tau_{j+2}, b_{j+2}\}$. If the duration of a blip, i.e., $\tau_{j+2}-\tau_j$, exceeds 0.06 sec and $b_{j+1}/z(\tau_{j+1})$ exceeds 6, then a P-arrival is declared to have occurred at the time τ_j.

So far, the characteristic functions of the incoming signals were calculated in order to detect the first P-arrivals. They are aimed principally at focusing the time variation of amplitude of signal and have no clear theoretical basis for discrimination between seismic signal and background noise. To deal successfully with the subjects, a different method, which adopts parametric models of autoregressive type, can be used efficiently.

Tjøstheim (1975a) referred to Box and Jenkins (1970) which contains general method for identification, fitting, estimating and diagnostic checking of general ARMA (Autoregressive-moving average) time series models. ARMA models represent the wide sense stationary time series $X(t)$ as

$$\begin{aligned}&X(t)-a_1X(t-1)-\cdots-a_pX(t-p)\\&=z(t)-b_1z(t-1)-\cdots-b_qz(t-q),\end{aligned} \tag{3.8}$$

where $X(t)$ is the observed time series, and $z(t)$ is a white noise time series such

that expectation $E\{X(t), z(s)\}=0$ for $s>t$. Although, the ARMA model is a general expression for the stationary time series, it is demonstrated that the digitized short-period noise at Norwegian Seismic Array, NORSAR can be described with a satisfactory statistical fit by an autoregressive model (b_1, b_2, ..., $b_q=0$ in Eq. 3.8).

It is natural to consider next the possibility of fitting autoregressive model to the discrete time series defined by a digitized P-wave signal. At NORSAR, the autoregressive analysis was undertaken for 40 underground nuclear explosions and 45 earthquakes from Eurasian (Tjøstheim, 1975b). It was found that in most cases a reasonable statistical fit is obtained using a low order autoressive model. The method utilized both amplitude and frequency information. Shirai and Tokuhiro (1979) used a log-likelihood ratio function to determine onset times of P and S-waves of local earthquakes. The log-likelihood function, $\Lambda_{P/N}$ for P-wave to background noise is

$$\Lambda_{P/N}(k)=[\log(\bar{\sigma}_N^2/\bar{\sigma}_P^2)+\{x(k)-\bar{\boldsymbol{a}}_N^T\boldsymbol{u}_N(k)\}^2/\bar{\sigma}_N^2 - \{x(k)-\bar{\boldsymbol{a}}_P^T\boldsymbol{u}_P(k)\}^2/\bar{\sigma}_P^2]/2, \tag{3.9}$$

where $\bar{\boldsymbol{a}}$ and $\bar{\sigma}$ are least squares estimators of AR coefficient $\boldsymbol{a}$ and variance σ^2, and $\boldsymbol{u}^T(k)=\{x(k-1), ..., x(k-q)\}$. To detect the first P-arrivals, log-likelihood ratio function $\Lambda_{P/N}(k)$ is examined. If $\Lambda_{P/N}(k)=0$, a P-arrival is declared to have occurred at the time k. In the same way, the determination of onset time of S-wave is carried out as if $\Lambda_{S/N}(k)=0$. Their method allows backtracking over the data to improve the reliability of the timing to a few hundredths of a second. It will be reasonable to use an autoregressive model to describe the behavior of seismic wave. Using the concept of log-likelihood ratio, Hamaguchi and Suzuki (1979) have developed the method for determining onset times of P-waves. Morita and Hamaguchi (1981) have also implemented this method to determine onset times of S waves. As for non-stationary time series such as seismic signals, the log-likelihood ratio is not always satisfied with declaration of first P or S arrivals. Further, AR models used in the log-likelihood ratio function should be established a priori. However, AR models of both noise and seismic signal are generally time dependent. Therefore, there is a problem of objectivity to judge whether it can determine the proper onset time of P or S-wave.

Ozaki and Tong (1975) first proposed adaptive procedure of modeling nonstationary process by a locally stationary AR model. They extended Akaike's minimum final prediction error (FPE) procedure for AR model fitting (Akaike, 1970) to locally stationary situations using the AIC developed by

Akaike (1973). The FPE and AIC are defined as

$$\mathrm{FPE}=S(M)\{1+(M+1)/N\}\{1-(M+1)/N\}^{-1}, \tag{3.10}$$

and

$$\mathrm{AIC}=-2\log\{\text{maximum likelihood}\}+2\{\text{number of parameters}\}, \tag{3.11}$$

where M is the order of the AR model. N is the data length and $S(M)$ is the sample variance of residuals of the AR model of order M.

Kitagawa and Akaike (1978) modified the procedure based on the adoption of Householder transformation for least squares computations. Yokota et al. (1981) applied the AR model to the procedure for the declaration of the arrival times of P-waves recorded by the telemetering network of microearthquake observation in Kanto, central Japan. The usefulness of the procedure based on the locally stationary AR model was successfully substantiated by the implementation to the microearthquake observation network.

The practice is being taken up in automatic processing system (e.g. Hasegawa et al., 1986). In any automatic processing of seismic waves, it is important not only to make efforts to read the arrival times of P waves and/or S waves with a high accuracy for each station but also to check the reliability of the results of reading in order to reject the inaccurately read items automatically. Morita and Hamaguchi (1984) and Maeda (1985) have tested the procedures using the seismic wave data observed by their microearthquake networks. According to their studies, the method for automatic processing of the seismic wave data is proved to be useful for the practical purpose. In general such a model was obtained mainly by using classical least squares method for log-likelihood estimators of the AR model. Therefore, after Kitagawa and Akaike (1978), we introduce a new procedure based on the adoption of Householder transformation for such least squares computations. This algorithm provides a very simple and fast procedure for handling additional observations which is useful for the on-line fitting of locally stationary AR models.

4. Methods of representing trajectories of particle motions

Polarization is quantitatively characterized by parameters that determine the trajectories of particle motion. In general, moving particles describe intricate trajectories in space, and for this reason methods of representing such trajectories are of great importance in their analysis and understanding.

We shall begin this chapter by theoretically reviewing of particle motions

for P and S-waves in an isotropic perfectly elastic homogeneous media. In the following sections, we will review and check the statistical methods to show the main properties of their trajectories. In general, the trajectories may be represented in various methods. Here, we will focus on the orbit spectrum analysis as the most direct method of event detection by visual, and on the principal components analysis as a statistical analysis. The principal components analysis is the three components analysis, which is here limited to the two kinds of polarization methods, that is a moving window analysis and a rectilinear motion detector (REMODE filter) analysis.

4.1 Basic information on P- and S-wave polarizations in seismic wave theory

The purpose of this section is to provide certain fundamental behaviors of P and S-waves by the seismic wave theory.

The propagation of seismic waves in a complex, laterally varying, 2 or 3-dimensional structure is a considerably complicated process. Analytical solutions of the elastodynamic equatiens for such types of media are not known. The most general mathematical statement of the properties of wave propagation in an isotropic perfectly elastic homogeneous medium described by the Lamē elastic parameters λ and μ and by the density ρ is included in the standard partial differential equation

$$\delta^2\mathbf{A}/\delta x^2+\delta^2\boldsymbol{A}/\delta y^2+\delta^2\boldsymbol{A}/\delta z^2=1/\boldsymbol{C}^2\cdot\delta^2\boldsymbol{A}/\delta t^2, \tag{4.1}$$

or

$$\nabla^2\boldsymbol{A}=1/\boldsymbol{C}^2\cdot\delta^2\boldsymbol{A}/\delta t^2. \tag{4.2}$$

This may be a single equation, in which case $\boldsymbol{A}$ is a scalar quantity, or there may be a set of three equations in each of which $\boldsymbol{A}$ is one component of a vector. $\boldsymbol{C}$ is the phase velocity of propagation of wave. The wave velocity for P-wave is given by $C_P^2=(\lambda+2\mu)/\rho$, and the one for S-wave is given by $C_S^2=\mu/\rho$. It can be shown that any arbitrary initial displacement of components (u, v, w) will separate into two parts, propagating with two characteristic velocities C_P and C_S. If recording of seismic signal begins with a large sharp P-wave, the direction of the first impulse can be seen in all three components. The ray lies in the vertical plane through epicenter and station. If the initial impulse is not clear, it may be possible to compare directions of displacement in three components at some later instant during the group of P-waves. Direction for S-wave cannot be derived so simply because it may arrive having polarized in any plane. In general P-waves are most clearly recorded in the vertical component and S-

waves in the horizontal components. P and S-waves are polarized on their own directions.

The next sections will deal with the problems of detecting and discriminating polarized wave-forms in three component time series.

4.2 Brief review of studies of seismic wave polarization

4.2.1 Introduction

In this section, we will provide background material from the areas of illustration technique of trajectory, principal components analysis, and the method of enhancement of highly polarized waves.

The polarization properties of electromagnetic fields in optics and radio transmission have been the subject of considerable study and much of the theory is derived from the work (Born and Wolf, 1965). Early examples which include the necessary background mathematical exposition are the studies of Paulson et al. (1965) and Fowler et al. (1967).

In seismolgy, the types of elastic waves which are produced by an earthquake or an explosion are well known from theoretical and laboratory model studies. However, the seismic recordings are always contaminated by noise which makes the detection and interpretation of small seismic events difficult.

Therefore polarization analysis has been used to devise filters which will separate elastic body waves into compressional (P) and shear (S) phases and also enhance or attenuate surface Rayleigh and Love waves as desired (Flinn, 1965 ; Archambeau and Flinn, 1965 ; Montalbetti and Kanasewich, 1970).

Dziewonski et al. (1969), Kaneko and Watanabe (1982) have studied variation of amplitude of signal as a function of time and frequency to obtain accurate identification of wave type and direction of the principal axis of trajectory at specific discrete frequency. Kaneko and Watanabe (1982) have developed a unique orbit analysis for visually identifying the waves that are nearly linearly polarized at specific discrete frequencies. Therefore, they called it "Orbit Spectral Analysis". This analysis can resolve complex transient signals composed of several dominant periods that arrive at the recording station simultaneously.

Vidale (1986) has shown that polarization analysis of strong motion data from the 1971 San Fernando earthquake aids in the discrimination between wave types, which is important for the understanding of the complicated earthquake induced shaking observed in basins.

4.2.2 Determination of parameters of elliptically- and multiply-polarized waves by statistical analysis

4.2.2.1 Brief review of mathematical representation for principal components analysis

In this section, we will review the mathematical representation for principal components in three-dimensional space.

Generally speaking, principal components analysis is a multivariate technique for examining relationship among several quantitative variables. It is used for summarizing data and detecting linear relationships. Given a data set with p numeric variables, p principal components can be computed. Each principal component is a linear combination of the original variables, with coefficients equal to the eigenvectors of the correlation or covariance matrix. We use the principal components analysis for understanding statistically seismic waves of microearthquakes in the next section.

Consider a plane quasi-monochromatic wave propagation in the z direction. A quasi-monochromatic wave is one on which most of the energy is confined over a small band width, df, about the mean frequency, f,

$$df/f \ll 1. \tag{4.3}$$

To observe the polarization properties at any one point, it is required that the complex amplitude and phase are relatively constant over a time, T, which is called the coherence interval and is determined by

$$1/f < T < 1/df. \tag{4.4}$$

The recorded components are real functions of time but may be written as

$$\begin{aligned} E_x(t) &= A_x(t) \exp\{i(\phi_x(t) - 2\pi ft)\}, \\ E_y(t) &= A_y(t) \exp\{i(\phi_y(t) - 2\pi ft)\}, \end{aligned} \tag{4.5}$$

and

$$E_z(t) = 0.$$

where the amplitude, A, and phase, ϕ, are functions of time, but vary slowly relative to the period of the sinusoidal oscillation. It will be assumed that the time average of both components is zero. If the wave is monochromatic and invariant with time, the field is perfectly polarized.

Letting $-2\pi ft = \tau$ and taking the real part of Eq. 4.5, we have

$$E_x = A_x \cos(\tau + \phi_x),$$

and

$$E_y = A_y \cos(\tau + \phi_y). \tag{4.6}$$

To eliminate τ we make use of the following two trigonometric identities,

$$\cos(\tau +/- \phi) = \cos\tau\cos\phi -/+ \sin\tau\sin\phi,$$

and

$$\sin(\tau - \phi) = \sin\tau\cos\phi - \cos\tau\sin\phi. \tag{4.7}$$

Using the first of these on Eq. 4.6 gives

$$E_x/A_x = \cos\tau\cos\phi_x - \sin\tau\sin\phi_x.$$

and

$$E_y/A_y = \cos\tau\cos\phi_y - \sin\tau\sin\phi_y. \tag{4.8}$$

Hence

$$(E_x/A_x)\sin\phi_y - (E_y/A_y)\sin\phi_x = \cos\tau\sin(\phi_y - \phi_x),$$

and

$$(E_x/A_x)\cos\phi_y - (E_y/A_y)\cos\phi_x = \sin\tau\sin(\phi_y - \phi_x). \tag{4.9}$$

Squaring and adding give

$$(E_x/A_x)^2 + (E_y/A_y)^2 - 2(E_x/A_x)(E_x/A_y)\cos\delta = \sin^2\delta, \tag{4.10}$$

where

$$\delta = \phi_y - \phi_x. \tag{4.11}$$

Equation 4.10 is the equation of a conic section. It is an ellipse, since the associated determinant is not negative,

$$\begin{aligned} &\begin{bmatrix} 1/A_x^2 & -\cos\delta/(A_xA_y) \\ -\cos\delta/(A_xA_y) & 1/A_y^2 \end{bmatrix} \\ &= 1/(A_x^2A_y^2)[1-\cos^2\delta] = \sin^2\delta/(A_x^2A_y^2) \\ &\geqq 0. \end{aligned} \tag{4.12}$$

The ellipse is inscribed to a rectangle whose sides are parallel to the coordinate axes and whose lengths are $2A_x$ and $2A_y$ (Fig. 4.1). The ellipse touches the sides at the points $(+/-A_x,\ +/-A_y\cos\delta)$ and $(+/-A_x\cos\delta,\ +/-A_y)$. The wave represented by Eq. 4.6 is then said to be elliptically polarized. In general, the axes of the ellipse are not in the OX and OY directions. Eq. 4.10 may be written in the matrix form as

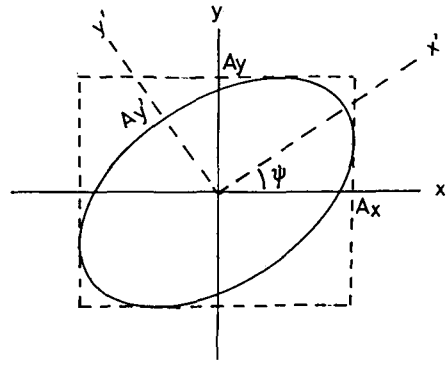


Fig. 4.1. Polarization ellipse describing the locus of the end point of vector.

$$[E_y E_x]\begin{bmatrix} A_x^2 & -A_x A_y \cos\delta \\ -A_x A_y \cos\delta & A_y^2 \end{bmatrix}\begin{bmatrix} E_y \\ E_x \end{bmatrix} = A_x^2 A_y^2 \sin^2\delta, \tag{4.13}$$

or

$$\boldsymbol{E}\boldsymbol{S}\boldsymbol{E}^T = A_x^2 A_y^2 \sin^2\delta, \tag{4.14}$$

where $\boldsymbol{E}^T$ is the transpose of $\boldsymbol{E}$.

An orthogonal transformation may be made to rotate by an angle ϕ to a new co-ordinate axis (x', y'), in which the major axis of the ellipse coincides with the x' axis. The rotation matrix, $\boldsymbol{T}$,

$$\boldsymbol{T} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}, \tag{4.15}$$

will convert $\boldsymbol{S}$ to the diagonal form, $\mathbf{S}'$,

$$\boldsymbol{T}^T\boldsymbol{S}\boldsymbol{T} = \boldsymbol{S}'. \tag{4.16}$$

The equation for the ellipse is then

$$[E_{x'}E_{y'}]\begin{bmatrix} A_{x'}^2 & 0 \\ 0 & A_{y'}^2 \end{bmatrix}\begin{bmatrix} E_{y'} \\ E_{x'} \end{bmatrix} = A_{x'}^2 A_{y'}^2 \tag{4.17}$$

In the rotated coordinate system, the field components are

$$E_{x'} = A_{x'}\cos(\tau+\phi),$$

and

$$E_{y'} = A_{y'}\sin(\tau+\phi). \tag{4.18}$$

Born and Wolf (1965) derive the polarization parameters. The angle ψ is given by

$$\tan 2\psi = 2A_{x'}A_{y'}\cos(\delta)/(A_{x'}^2 - A_{y'}^2). \tag{4.19}$$

The ellipticity is given by

$$\tan\beta = A_{x'}/A_{y'} \quad (-\pi/4 < \beta \leqq \pi/4), \tag{4.20}$$

where β is given by

$$\sin(2\beta) = 2A_{x'}A_{y'}\sin(\delta)/(A_{x'}^2 + A_{y'}^2). \tag{4.21}$$

When looking, face-on, into the propagating wave the polarization is said to be right handed when the rotation is clockwise ($\beta \geqq 0$) and left handed when the rotation is counter clockwise ($\beta < 0$). This convention is opposite to the usual terminology for a right or left handed screw.

Waves which have time varying amplitudes and phases as in Eq. 4.5 are not analyzed easily by their field vectors.

4.2.2.2 Brief review of applications of polarization filters in seismology

The object of this section is to present the method to determine direction and rectilineality in polarized particle motion, the moving window analysis, and the method for rectilinear motion detector by the statistical analysis of orthogonal components.

Various workers have applied polarization filtering techniques to recorded seismic data for improvement of the signal to noise ratio. Lewis and Meyer (1968) apply a phase filter of the REMODE type as described by Archambeau and Flinn (1965) and originally developed by Mims and Sax (1965) with subsequent work by Griffin (1966a, 1966b) to data recorded during the Early Rise experiment in the summer of 1966. Archambeau, Flinn and Lambert (1969) used the same type of filter to study multiple P phases from NTS explosions. In another application, Basham and Ellis (1969) used a REMODE filter designated as a P detection filter to process P-wave codas of numerous seismic events recorded in western Alberta. The polarization filter applied to 25 seconds of record following the P onset aids in identification of numerous compressional wave arrivals including P, pP and sP. PcP and PKP phases for events at appropriate epicentral distances are also detected. The design of the polarization filter used by Montalbetti and Kanasewich (1970) is a variation on one described by Flinn (1965) in which both rectilinearity and direction of particle motion are considered. In order to obtain measures of these two quantities, the covariance matrix for a set of N points taken over each of the three orthogonal components of ground motion, R (radial), T (transverse) and Z (vertical) is computed. For

a three-component digital seismogram, a specified time window of length $N\Delta t$, where Δt is the sampling interval, is considered. To determine the covariance matrix for this set of observations, the means, variances and covariances must be calculated for the three variables R, T and Z.

Next we will review the statistical analysis of orthogonal components (a principal components analysis) of motion in the following sections 4.2.2.2.1 ~4.2.2.2.3 and then will implement polarization filter, which is induced from the principal components analysis, to the seismic waves of microearthquakes recorded by the Hidaka network in section 4.3.

4.2.2.2.1 Direction of polarization and rectilinearity

We define the mean or expected value of N observations of the random variable $X_{1i}(i=1, 2, ..., N)$ as

$$\mu_1=1/N\sum_{i=1}^{N}X_{1i}=E(X_1). \tag{4.22}$$

The covariance between N observations of two variables X_1 and X_2 is given by

$$\boldsymbol{Cov}[X_1, X_2]=1/N\sum_{i=1}^{N}(X_{1i}-\mu_1)(X_{2i}-\mu_2), \tag{4.23}$$

where μ_1 and μ_2 are computed as in Eq. 4.22.
It is evident that

$$\boldsymbol{Cov}[X_1, X_2]=\boldsymbol{Cov}[X_2, X_1]. \tag{4.24}$$

The matrix with $\boldsymbol{Cov}[X_r, X_s]$ in its rth row and sth column $(r, s=1, ..., n)$ is the covariance matrix for the set of n random variables $X_j(j=1, ..., n)$. If $\overline{X}$ is vector of the random variables and $\bar{\mu}$ the vector of means for each of these variables, the covariance matrix $\boldsymbol{V}$ is defined by

$$\boldsymbol{V}=E[(\overline{X}-\bar{\mu})(\overline{X}-\bar{\mu})^t]. \tag{4.25}$$

The superscript t indicates the column transpose of vector.
For the case of three variables $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ considered over the time window $N\Delta t$, Eq. 4.25 is represented by

$$\boldsymbol{V}=\begin{bmatrix} \boldsymbol{Var}[X] & \boldsymbol{Cov}[X, Y] & \boldsymbol{Cov}[X, Z] \\ \boldsymbol{Cov}[X, Y] & \boldsymbol{Var}[Y] & \boldsymbol{Cov}[Y, Z] \\ \boldsymbol{Cov}[X, Z] & \boldsymbol{Cov}[Y, Z] & \boldsymbol{Var}[Z] \end{bmatrix}, \tag{4.26}$$

where the covariances and variances are defined in Eq. 4.23.

If the covariance matrix given by Eq. 4.26 is diagonalized, an estimate of the rectilinearity of particle motion trajectory over the specified time window can be obtained from the ratio of the principal axes of this matrix. The direction

of polarization may be measured by considering the eigenvector of the largest principal axis. If λ_1 is largest eigenvalue and λ_2 the next largest eigenvalue of the covariance matrix, then a function of the form

$$F(\lambda_1, \lambda_2)=1(\lambda_2/\lambda_1)^n \tag{4.27}$$

would be close to unity when rectilinearity is high ($\lambda_1 \gg \lambda_2$) and close to zero when the two principal axes approach one another in magnitude (low rectilinearity). The direction of polarization can be determined by considering the components of the eigenvector associated with the largest eigenvalue with respect to the coordinate directions X, Y and Z. Suppose now that in the matrix of Eq. 4.32 the variance terms are small with respect to those on the main diagonal, on which, $\boldsymbol{Var}(\boldsymbol{X}) > \boldsymbol{Var}(\boldsymbol{Y}) > \boldsymbol{Var}(\boldsymbol{Z})$. The lengths of the two principal axes would thus correspond very closely to $\boldsymbol{Var}(\boldsymbol{X})$ and $\boldsymbol{Var}(\boldsymbol{Y})$, and the eigenvector associated with the major axis, $\boldsymbol{Var}(\boldsymbol{X})$, $\boldsymbol{Var}(\boldsymbol{X}) \gg \boldsymbol{Var}(\boldsymbol{Y})$, then $F(\lambda_1, \lambda_2)$ in Eq. 4.27 would be close to unity and we would have high rectilinearity in the X direction. If $\boldsymbol{Var}(\boldsymbol{X}) \sim \boldsymbol{Var}(\boldsymbol{Y})$, then the direction of polarization would still predominantly along the X coordinate axis, but the rectilinearity would be low. If the off-diagonal terms of the matrix were significant, the diagonalization would introduce a rotation, and the orientation in space of the major axis would be given by the components of its eigenvector relative to the original X, Y, Z coordinate system.

4.2.2.2.2 Moving window analysis

In a previous section 4.2.2.2.1, variances and covariances of recorded ground motions were obtained for successive time intervals using the relation

$$\mu_{ij}=\langle[a_i(t)-\overline{a_i}][a_j(t)-\overline{a_j}]\rangle,\ t=t_1,\ t_2 \qquad i,j=x,y,z, \tag{4.28}$$

in which the the triangular brackets denotes time averaging and they are taken over the interval $t_1 \leqq t < t_2$ but where the mean values $\overline{a_i}$ and $\overline{a_j}$ are found by averaging $a_i(t)$, $a_j(t)$ over the entire duration of motion. In the present investigations, variances and covariances are obtained as continuous functions of time t_0 using the so-called "moving-window" technique, i.e. using the relation

$$\mu_{ij}(t_0, \Delta t)=\langle[a_i(t)-\bar{a}_i][a_j(t)-\bar{a}_j]\rangle,$$
$$t=t_0-\Delta T/2,\ t_0+\Delta T/2,\ i,j=w,y,z, \tag{4.29}$$

where the time averages are taken over the interval ΔT centered at time t_0.

Having obtained all nine covariance functions for the recorded components of motion in accordance with Eq. 4.29, the corresponding time dependent directions of principal components obtained, i.e. giving the principal transformation matrix P as a function of time t_0 and time window length ΔT, i.e. $P=P(t_0,$

ΔT). This time dependent principal transformation matrix then allows one to obtain the time dependent directions of principal components of motion $a_x(t)$, $a_y(t)$ and $a_z(t)$ and their corresponding principal variances $\sigma_1(t_0)^2$, $\sigma_2(t_0)^2$ and $\sigma_3(t_0)^2$. ΔT should be taken sufficiently long so that the higher frequency fluctuations are essentially removed but the slower time dependent characteristics are retained, i.e. the time average over duration ΔT will be essentially equal to the average taken across the ensemble.

The direction of each principal component axis is given by angles ϕ and θ as shown in Fig. 4.2. Angle ϕ is the declination of the principal axis from the vertical axis through point "0"; thus, its value falls in the range $0° \leqq \phi \leqq 90°$. Angle θ is measured from the north axis to the projection of the northerly extension of the principal axis on a horizontal plane containing point "0". By this definition, θ lies in the range $-90° < \theta \leqq 90°$. The angle θ_E in Fig. 4.2 represents the horizontal direction of the axis passing through the seismographic site location (point "0") and the epicenter. Since this angle is measured in a similar manner to that of angle θ, it also lies in the range $-90° < \theta_E \leqq 90°$. Length OA in Fig. 4.2 represents the magnitude of the variance of principal ground motion. The square root of this quantity (sigma) can be used to represent the intensity functions of the corresponding nonstationary process (Iyengar and Iyengar, 1969; Kubo and Penzen, 1975).

4.2.2.2.3 Enhancement of weak microseismic data by REMODE filter

A particular form of a polarization filter which considers only the rectilinearity in the radial and vertical directions instead of in three dimensional space has been called "REMODE" for rectilinear motion detector. It was originated by Mims and Sax (1965) in the time domain and by Archambeau and Flinn (1966 a, b) in the frequency domain. Additional subsequent development was made by Griffin (1966a, b).

The vertibal (Z) and redial (R) components of seismic signals are rotated so

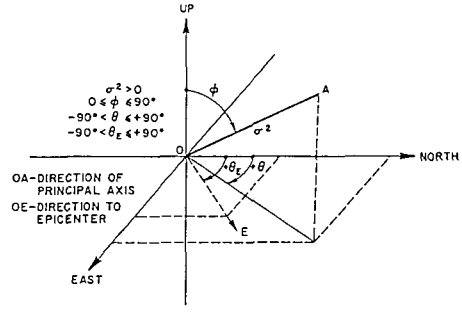


Fig. 4.2. Directions of principal axes in three dimensional space.

Polarization filter

Rotation of Components

$$Z' = Z\cos\left(\frac{\pi}{4} - i\right) + N\cos\alpha\sin\left(\frac{\pi}{4} - i\right) + E\sin\alpha\sin\left(\frac{\pi}{4} - i\right)$$

and

$$R' = Z\sin\left(\frac{\pi}{4} - i\right) - N\cos\alpha\cos\left(\frac{\pi}{4} - i\right) - E\sin\alpha\cos\left(\frac{\pi}{4} - i\right)$$

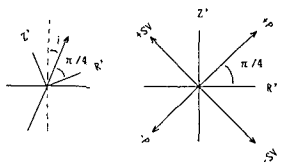


Fig. 4.3. Rotation to obtain the Z and R components for a REMODE filter. On the right: particle motion for either an incident compressional (P) or a vertical polarized shear (SV) wave.

that the expected direction of the incident body waves bisect the angle between the two orthogonal components ($\bar{Z}$, $\bar{R}$ in Fig. 4.3). That is

$$\bar{Z} = Z\cos(\pi/4-\theta)+N\cos\alpha\sin(\pi/4-\theta)+E\sin\alpha\sin(\pi/4-\theta),$$

and

$$\bar{R} = Z\sin(\pi/4-\theta)-N\cos\alpha\cos(\pi/4-\theta)+E\sin\alpha\cos(\pi/4-\theta), \quad (4.30)$$

where θ is the angle of incidence, α is the azimuth of a great circle path, Z, N, E are the vertical, north-south and east-west components of motion. All the P and SV motion is represented by these two rotated components so that the similar shape of $\bar{Z}$ and $\bar{R}$ signal wave forms contrasts with dissimilar shape for any noise.

The filter operator is obtained from a cross correlation function, $C(T)$ of $\bar{Z}(t)$ and $\bar{R}(t)$ over a window, W, centered at some time, t, on the record.

$$C(+T)=\sum_{t-W/2}^{t+W/2}\bar{Z}(t)\bar{R}(t+T). \quad (4.31)$$

To ensure that the operator is an even function and it introduces no phase distortion, the negative lags are generated from the positive lags.

$$C(-T)=C(+T). \quad (4.32)$$

If the polarization in the $\bar{R}-\bar{Z}$ plane in predominantly rectilinear, $C(T)$ will be large. If the motion is elliptical or random, it will be small. By convolving $C(T)$ with the original time series, motion of high rectilinearity is enhanced and

elliptically polarized motion is attenuated. The output from the filter is given by

$$Y_R(t)=P_PK\sum_{T=-L}^{L}\bar{R}(t-T)C(T),$$

and

$$Y_Z(t)=P_PK\sum_{T=-L}^{L}\bar{Z}(t-T)C(T), \tag{4.33}$$

where P_P is a polarization operator and K is a normalizing factor.

The operator, $C(T)$ is different for every data point, t, so the processing is computationally expensive. It is recommended that the maximum lag, $T_{\max}$ be half the interval, L. A normalizing factor, K, determined from autocorrelation of the input may be used to enhance weak signals with high rectilinearity.

$$K(t)=\left[\sum_{t-W/2}^{t-W/2}\bar{R}^2(t)\sum_{t-W/2}^{t+W/2}\bar{Z}^2(t)\right]^{-1/2} \tag{4.34}$$

The REMODE operator, as described above, will enhance both linearly polarized P and SV phases. Note that for P or compressional motion in Fig. 4.3,

$$\bar{R}(t)\bar{Z}>0, \tag{4.35}$$

while for vertically polarized shear waves

$$\bar{R}(t)\bar{Z}<0. \tag{4.36}$$

The polarization operator, P_P is defined as

$$P_P=\begin{vmatrix}1 & \bar{R}(t)\bar{Z}(t)>0,\\ 0 & \bar{R}(t)\bar{Z}(t)<0,\end{vmatrix} \tag{4.37}$$

for P phases and

$$P_S=\begin{vmatrix}1 & \bar{R}(t)\bar{Z}(t)<0,\\ 0 & \bar{R}(t)\bar{Z}(t)>0,\end{vmatrix} \tag{4.38}$$

for SV phases.

4.3 Application of polarization analyses to Hidaka array data

In the section 4.2, we reviewed the fundamental methods to present particle motion properties by the statistical analysis. In this section, we are concerned with the real applications of them to the microseismic waves recorded by the Hidaka array.

4.3.1 Plotting trajectory of particle motion by orbit spectral analysis

4.3.1.1 Introduction of orbit spectral analysis

Here, we provide fundamental procedure of the orbit spectral analysis for a visual trajectory of particle motion.

The variation of amplitude of a signal as a function of time and frequency is studied by the multiple filter technique (Dziewonski, 1969; Kaneko and Watanabe, 1982; Loh, 1985) to obtain an accurate identification of wave types and directions of principal trajectory axes at specific discrete frequencies. This method can resolve complex transient signals composed of several dominant periods that arrive at the recording station simultaneously. When the orthogonal components of the motion are analyzed, the results can be used to study the three dirmensional vibration of the recording site.

The multifilter technique is based on a bandpass filter, $H_n(\omega)$, which operates on the Fourier transform of the signal; a Gaussian filter has been used for the orbit spectral analysis. We use the multifilter technique based on the use of a Butterworth band-pass filter, $H_n(\omega)$, the corresponding digital filter of which is $H(z)$. Its window function can be expressed as

$$|H_n(\omega)|^2=[1-(\omega/\omega_{01})^{2n}\cdot(\omega/\omega_{02})^{2n}]/[1+(\omega/\omega_{01})^{2n}]\cdot[1+\omega/\omega_{02})^{2n}], \quad (4.39)$$

where ω_{01} and ω_{02} are the cut-off frequencies, and

$$\omega_{01}<\omega_{02}. \quad (4.40)$$

Here n, ω_{01} and ω_{02} are determined so as to satisfy the following conditions,

$$\begin{aligned}
&1/\{1+(\omega_L/\omega_{01})^{2n}\}=1/\{1+(\omega_H/\omega_{02})^{2n}\}\geqq RP,\\
&1/\{1+(\omega_{c1}/\omega_{01})^{2n}\}=1/\{1+\omega_{c2}/\omega_{02})^{2n}\}\geqq RC,\\
&n\geqq\{\log(\sqrt{(1/RP-1)}/\sqrt{(1/RC)-1})/\log(\omega_L/\omega_{c1})\},\\
&\omega_{c2}(1/RC-1)^{-1/2n}>\omega_{02}>\omega_L(1/RC-1)^{-1/2n},
\end{aligned} \quad (4.41)$$

and,

$$\omega_{c1}(1/RC-1)^{-1/2n}<\omega_{01}<\omega_H(1/RC-1)^{-1/2n},$$

where $\omega_{c1}=2\pi FC_1$, $\omega_{c2}=2\pi FC_2$, $\omega_L=2\pi FL$, and $\omega_H=2\pi FH$. The parameters FC_1, FL, FH, FC_2, RP, and RC are explained in Fig. 4.4. In terms of Laplace transform, i.e. after introducing the complex frequency as the independent variable, the squared absolute value of the frequency response becomes,

$$|H(p)|^2=H(p)^*H(p)=1/(1+(-1)^n\cdot p^{2n}), \quad (4.42)$$

where $p=\sigma+j\omega$ and $H(p)^*$ is complex conjugate of $H(p)$. In order to determine the digital representation $H(z)$, we replace p in $H(p)$ by $(1-z^{-1})/(1$

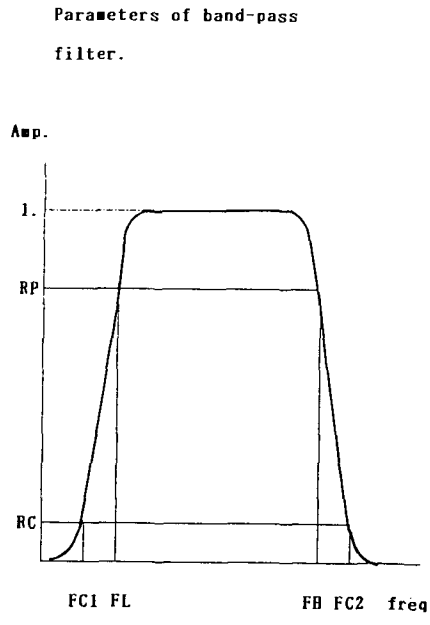


Fig. 4.4. Typical response of Butterworth digital filter. Parameters of FH and FL are cut off frequencies of the filter. FC_1 and FC_2 are the parameters to define the truncation response. RP and RC are the parameters to define the amplitude of the filter. $FH=1.1\cdot FO$, $FL=0.9\cdot FO$. $FC_1=0.75\cdot FO$, $FC_2=1.25\cdot FO$, $RC=0.05$, $RP=0.9$, and FO is a centered frequency of the band-pass filter.

$+z^{-1}$), which gives coefficients a_n, b_n of

$$H(z)=\left(\sum_{n=0}^{M} a_n z^{-n}\right)/\left(1+\sum_{n=1}^{L} b_n z^{-n}\right). \tag{4.43}$$

The nth output value becomes

$$y_n=\sum_{i=0}^{M} a_i x_{n-i}-\sum_{i=1}^{L} b_i y_{n-i}. \tag{4.44}$$

where y_n is assumed to be zero for $n<0$. This Eq. 4.44 describes the recursive filtering procedure. With the use of y_n, the orbit spectra can be constructed at specified frequency bands as shown in Fig. 4.5. Figure 4.6 shows the flow diagram of this analysis.

4.3.1.2 Some results of orbit spectral analysis

In this section, we show the results obtained by the orbit spectral analysis.

One set of the Hidaka array data, recorded from the earthquake of March 21, 1982 is used to illustrate the orbit spectral analysis. Figures 4.7, 4.8, and 4.9 show the orbits obtained from the seismograms of microearthquake at station MYR. The direction of the maximum variance of P wave correlates reasonably well with the direction to the epicenter which is located 45 km apart from the station in direction of S85° W, and the hypocentral depth for this earthquake

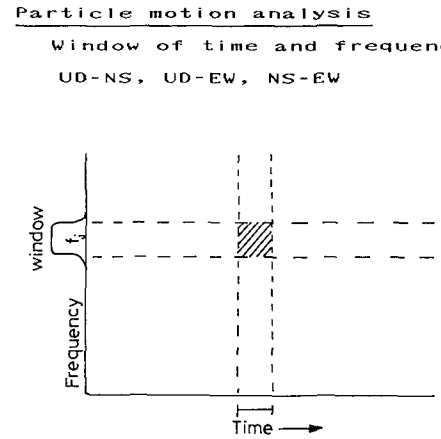


Fig. 4.5. A unit window cell is shown by a hatched square. A particle motion by two component seismograms is obtained through each window. Length of time for the unit cell is 1.083 sec. Frequency bandwidth of the cell is the width of Butterworth filter. Centered frequency of the filter is determined by the relation, $F(j)=FMA/(10.0 ** \log(FMA/FMI)/(IF-1)) ** -(j-1)$ $(j=1, IF)$. Here, $F(j)$, FMA, FMI, and IF are centered, maximum, minimum frequencies, and cell number of windows. $(FMA=15, FM=2, IF=18)$.

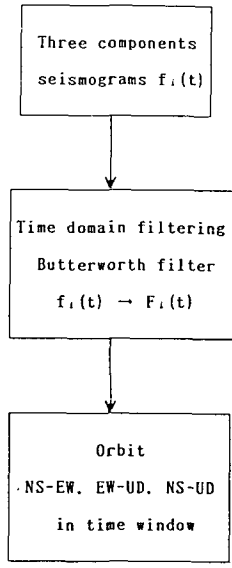


Fig. 4.6. Flow diagram for the orbit spectral analysis. A set of band-pass filters ($FO=2.0$ to 15.0) is used. Each centered frequency is given by $F(j)$ in Fig. 4.5.

is 26 km. Since the wave speed of the P-wave is different from one of the S-wave, we can identify the orbits for their own waves in the different time windows where they have the predominant orbits in their specific frequency bands. For example, for time window centered at $t_0=6.48$ sec, the first dominant wave is composed of several orbits, of which maximum trajectory is found at $f_0=11.63$ Hz. We can identify P-wave by the trajectories of orbits. The next predominant orbital group can be seen in the time window centered at $t_0=16.20$ sec. They contain surely the orbits of S-wave from the point of view of the arrival time. However we cannot easily identify S-wave by the trajectories since they vary both in size and in shape momently. This fact means that the trajectory of the S-wave is likely to change more abruptly than that of P wave does. S-wave is likely to be contaminated by the several later phases generated after P-wave. The direction of maximum variance of horizontal compo-

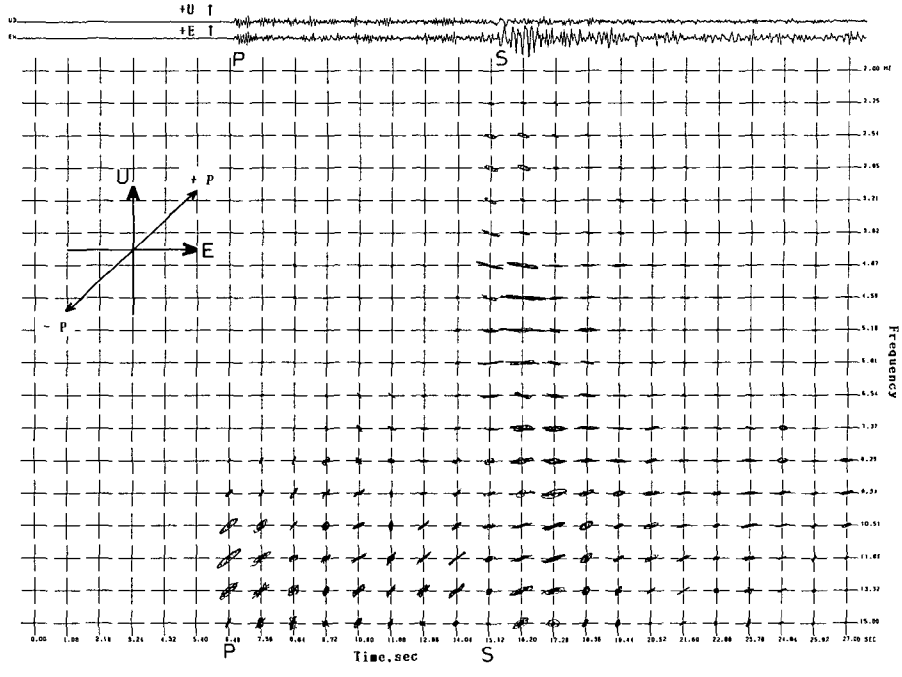


Fig. 4.7. Example of orbit spectral analysis at station MYR is shown. East-west and vertical components seismograms are used. The event occurred off Urakawa (42.13°N, 142.13°E). Depth, magnitude, and origin time are 26 km, 2.0, and 08 : 42, Mar. 21, 1982.

nents of S wave can not be determined because the trajectories over S-wave are likely circular or elliptical. Besides, the direction of maximum variance for S-wave does not seem to fit in with the direction of epicenter. From such visual orbit spectral analysis, we can understand that the change of wave amplitude of S-wave does not provide us with the good decision of S-arrival. When the orthogonal components of the motion are analyzed, the results can be used to study the three dirmensional motion of the recording site. Accordingly, we will consider to develop a statistical method for quantitative analysis of orthogonal components of the ground motion.

4.3.2 Applications of polarization filters

In this section, we show the results obtained by the polarization filters based on the statistical analysis of orthogonal components.

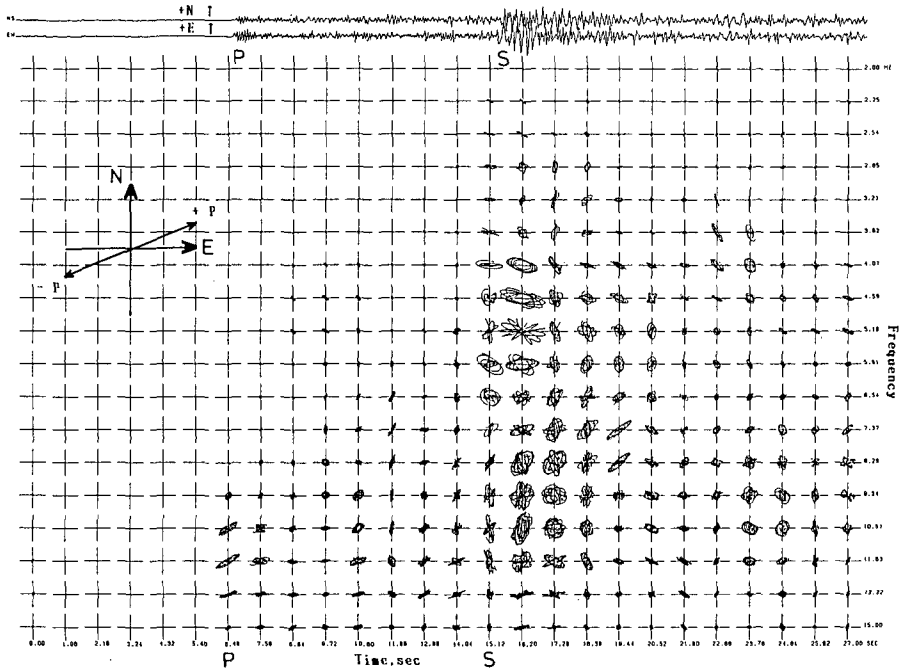


Fig. 4.8. Example of orbit spectral analysis at station MYR is shown. The components are north-south and east-west. The hypocentral parameters of the event is the same as those described in Fig. 4.7.

4.3.2.1 Some results of moving window analysis

Here we discuss the results obtained by the moving window analysis, which is one of the polarization filters.

Direction angle ϕ and θ, and the square root of the principal variance (σ) have been obtained as functions of time t_0 for the major, intermediate and minor principal axes of the ground motion at station MYR. The results are shown in Fig. 4.10.1 through 4.10.2. These results of Figs. 4.10.1 and 4.10.2 are respectively obtained by using time window length ΔT equal to two and five seconds at discrete values of one-half second apart. The solid, short dashed and intermediate-dashed curves in these figures represent respectively the major, minor and intermediate principal axes and the horizontal long-dashed straight line represents the direction θ_E to the reported epicenter. It should be noted from the definition of θ that as the horizontal direction of a principal axis rotates in a continuous manner through the east-west direction, the value of θ changes instantaneously by 180°, i.e. changes from $+90°$ to $-90°$ or from $-90°$ to 90°

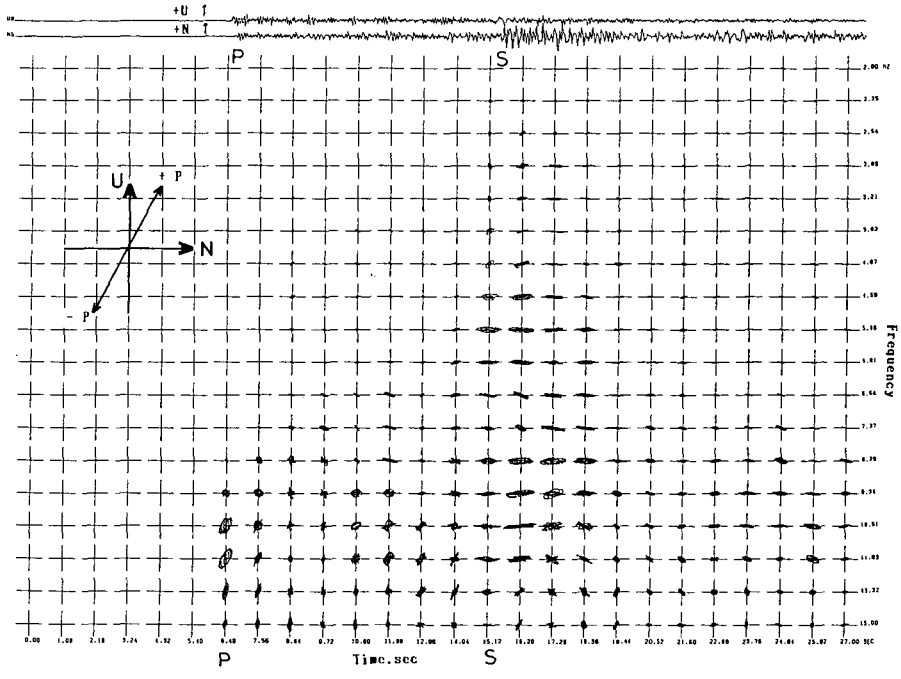


Fig. 4.9. Example of orbit spectral analysis at station MYR is shown. The components are north-south and vertical. The hypocentral parameters of the event is the same as those described in Fig. 4.7.

depending upon whether the horizontal projection of the principal axis is rotating clockwise or counterclockwise. This explains the sudden jumps which appear in the functions of θ which take place over single spacings of the prescribed discrete values of t_0, namely over one-half second spacings.

It should be noticed that if ground processes are represented by the product of stationary random processes and deterministic intensity functions and if the intensity functions for three components vary with time, any two of the principal variance functions differ from each other by fixed constant only in which case the directions of principal axes are fixed, i.e. they are invariant over the entire duration.

The time domain moving-window analysis described above has been applied to the ground motions recorded at Hidaka stations of the RCEP network just before the Urakawa-Oki earthquake of March 21, 1982. Velocity records at 6 stations have been evaluated. Stationa MYR and IWN are located in the east side of the Hidaka mountains and stations MSN, KMU, ERM, and HIC are

located on the west side of the mountains. Their locations are illustrated together with the epicenters of the foreshocks in Fig. 4.11. Table 4.1 summarizes the station locations, epicentral coordinates. One finds increased fluctuations in $\mu_{ij}(t_0, \Delta t)$ and the corresponding directions of principal axes as the values of time window length Δt in Eq. 4.29 is shorter and shorter. In fact, when $\Delta T = 2.0$ seconds, the major principal axis of ground motion coincides with the instantaneous resultant velocity vector which changes its direction rapidly in a random fashion over the entire sphere of space.

The certain correlations should be noted as follows:

(1) The directions of principal axes are sometimes suddenly interchanged. This interchange which occurs mainly after the period of high intensity motion is due to a corresponding change in direction along which the seismic waves such as P and S-waves have maximum energy.

(2) The directions of principal axes after P-waves do not quite correspond to the azimuths of epicenters.

(3) The relation among the directions of major principal axes after P-waves and of those after S-waves is not simple and so can not provide

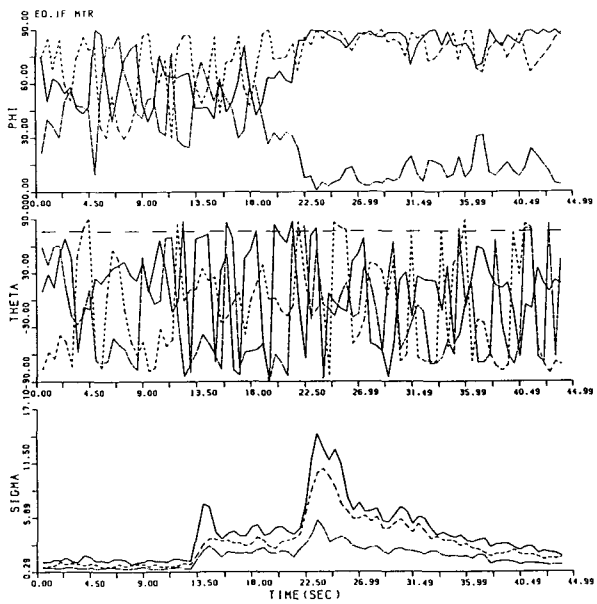


Fig. 4.10.1. Time dependent directions of principal axes and square roots of principal variances at station MYR. Time window length, ΔT is 2.0 sec. The parameters of hypocenter of event 1F are shown in Table 4.1.

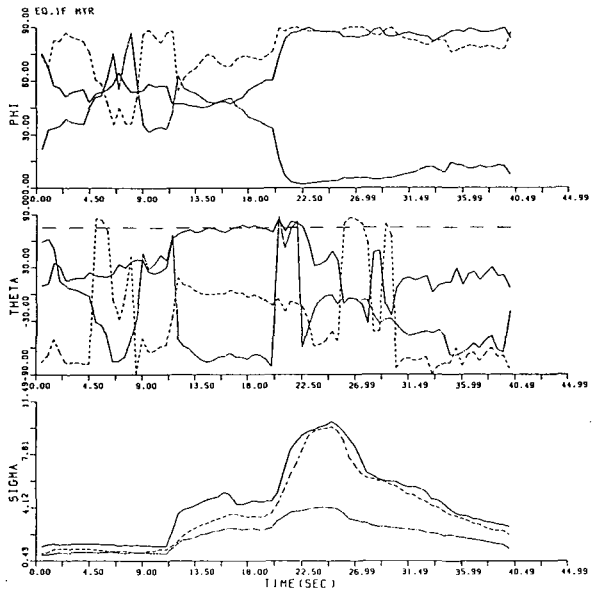


Fig. 4.10.2. Time dependent directions of principal axes and square roots of principal variances at station MYR. Time window length, ΔT is 5.0 sec. The parameters of hypocenter are the same as those of Fig. 4.10.1.

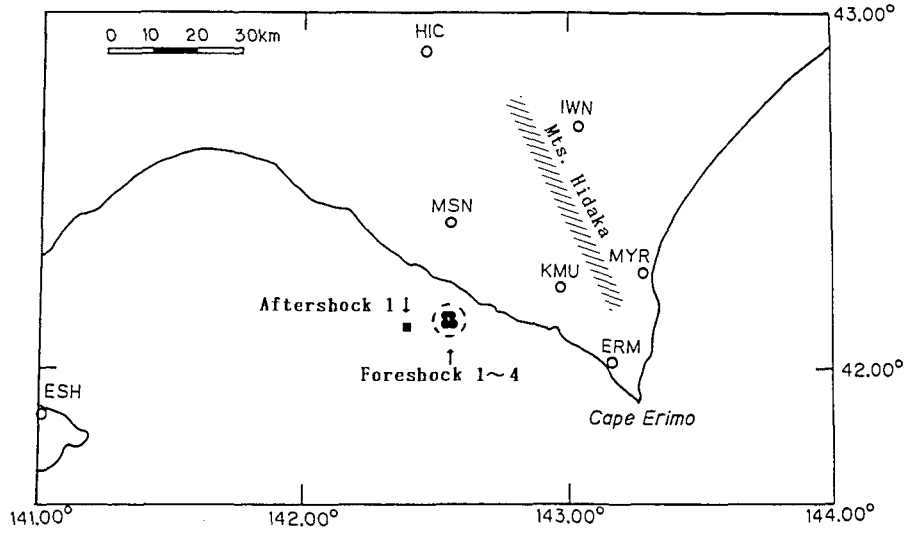


Fig. 4.11. Map showing the fore shock epicenters of the 1982 off Urakawa earthquake. The parameters of hypocenters shown in Table 4.1.

an identification of P or S-waves for the automatic processing.

(4) The principal axes of the ground motions show the averaging characteristics over the time windows ΔT, therefore the resolutions of onset times of incoming phases cannot be expected to be less than ΔT.

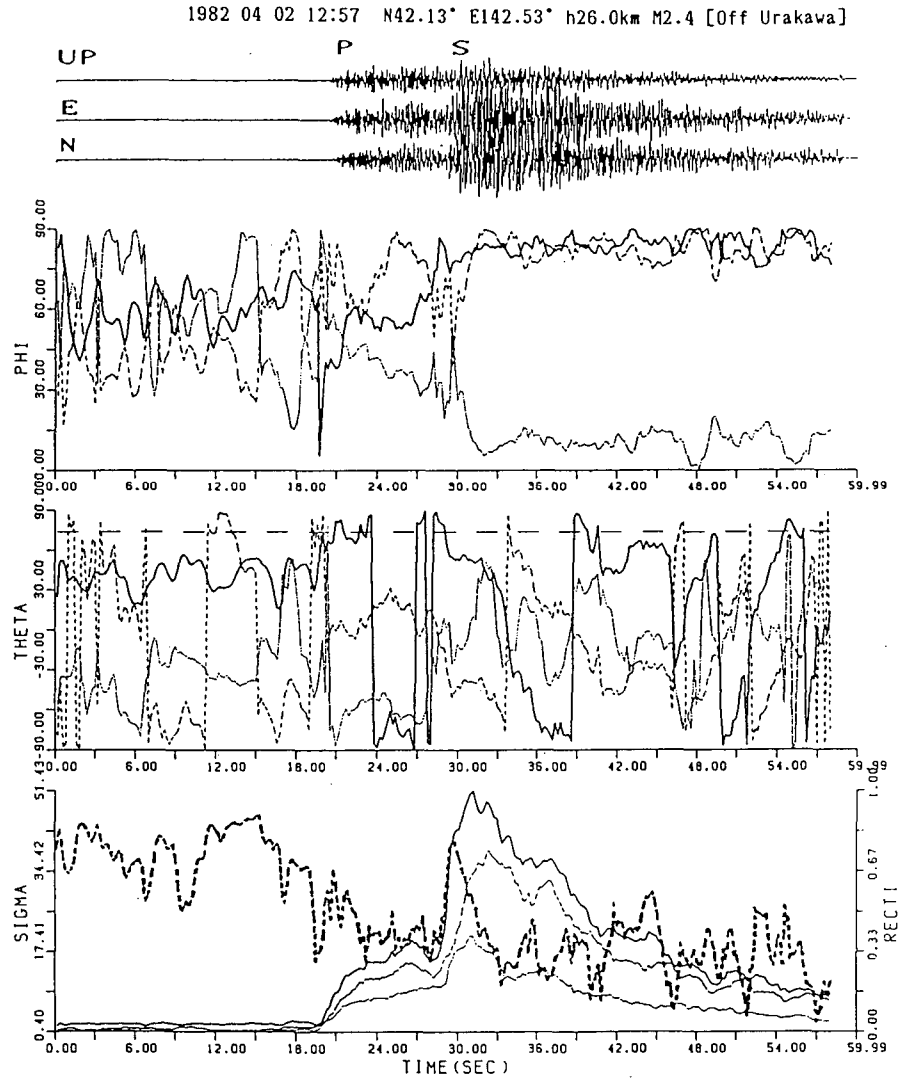


Fig. 4.12. Time dependent directions of principal axes and square roots of principal variances at station MYR. Time window length, ΔT is 2.0 sec. The parameters of hypocenter shown at the top of this figure.

Table 4.1. Source parameters of the events used in several analyses.

Date	Time	Long.	Lat.	Depth	M	Event No.
Mar. 21, 1982	07 : 45 53.0	142.557°E	42.158°N	31.0 km	1.9	Foreshock-1
Mar. 21, 1982	08 : 42 52.0	142.555°E	42.131°N	26.0 km	2.0	Foreshock-2
Mar. 21, 1982	08 : 49 20.7	142.561°E	42.158°N	33.7 km	2.1	Foreshock-3
Mar. 21, 1982	09 : 33 15.0	142.574°E	42.133°N	31.1 km	2.3	Foreshock-4
Dec. 12, 1982	12 : 04 24.0	142.380°E	42.090°N	18.8 km	2.6	Aftershock-1

For an example of shorter time window, the directions of principal axes, intensity function, and three components seismograms are shown together with a rectilinear curve in Fig. 4.12. ΔT used in Fig. 4.12 is 2.0 seconds and so the parameters of them should change rapidly rather than the case of $\Delta T = 5.0$ seconds. This microearthquake is an aftershock of the Urakawa-Oki earthquake of March 21, 1982 and the hypocentral parameters are quite the same as those of the above foreshocks, and recorded at station MYR.

Rectilinear curves shown in Fig. 4.12 does not reinforce the identification of P-wave or S-wave as the changes of the directions of principal axes, and the intensity functions cannot do so even if ΔT becomes shorter. So we are really in trouble. Now we will move on to the next application of REMODE filter.

4.3.2.2 Some results of REMODE filter

Here we show the results obtained by the REMODE filter, which was already explained in section 4.2.2.2.3.

Polarization filter known as REMODE is designed to enhance P and SV-wave motions which are rectilinears. We apply this filter to the same microearthquake data used in the previous applications. One example for station MSN is shown in Fig. 4.13. Attenuation of the SV-trace shown as in the bottom of Fig. 4.13 illustrates how the filter enhances the motion which exhibits a preferred direction of polarization. Previous applications of moving window analysis to the same microearthquake data indicated that this processor was difficult to detect and identify S-wave. Much of noise following the P phase has frequency content and polarization so similar to signal that this processing does not isolate weak signals effectively. On the other hand, MEMODE shows a good separation of phases P and S.

In general, REMODE is superior to the moving window analysis in enhancing P and S-waves of microearthquakes. However, it sometimes fails in enhancing P and S-waves of low S/N ratios, as shown in the case of the third event at MSN (Fig. 6.1(c)). And also the results at KMU are not well due to the

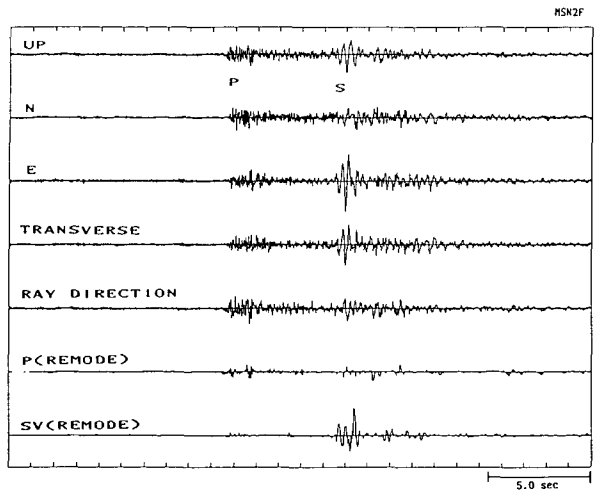


Fig. 4.13. Example of polarization filter, REMODE at station MSN. The parameters of hypocenter of event-2 shown in Table 4.1.

stationary hum noises. The effects of the noises contaminated in the seismic signals are serious for REMODE processor as well as the previous moving window analysis.

It is therefore suggested that such statistical analysis as the moving window analysis or REMODE, becomes unstable easily when the ratio of signal to noise decreases. And what is even worse, the moving window analysis or REMODE is conducted in the appropriate window of which length needs to be longer as the characteristics of the signals is kept, that is, the time resolution depends strongly on the window length. They are not suitable for determining onset time of instantaneous P and S-waves

Generally speaking, the spectral characteristics of seismic waves are not the same as those of the background noises and their features always very in time. So we will then review an autoregressive modeling, which is one of powerful time series analyses methods and provides us with information about spectral features of time series. Finally, we will apply the model to the microseismic signals.

5. Review of mathematical aspect of probability model for time series

5.1 Stochastic processes

This section describes the probability models for time series, which are

called stochastic processes. Most physical processes in the real world involve a random or stochastic element in their structure, and a stochastic process can be described as 'a statistical phenomenon that evolves in time according to probabilistic laws'. Many authors use the term 'stochastic process' to describe both the real physical process and mathematical model of it. The word 'stochastic', which is of Greek origin, is used to mean 'pertaining to chance', and many writers use 'random process' as a synonym for stochastic process. Mathematically, a stochastic process may be defined as a collection of random variables $\{X(t),\ t \in T\}$, where T denotes the set of time-points at which the process is defined. We will denote the random variable at time t by $X(t)$ if T is continuous (usually $-\infty < t < \infty$), and by X_t if T is discrete (usually $t=0, \pm 1, \pm 2, \cdots$). Thus a stochastic process is a collection of random variables which are ordered in time. For a single outcome of the process we have only one observation on each random variable and these values evolve in time according to probabilistic laws.

In most statistical problems we are concerned with estimating the properties of a population from a sample. In time series analysis, however, it is often impossible to make more than one observation at a given time so that we only have one observation on the random variable at time t. Nevertheless we may regard the observed time series as just one example of the infinite set of time series which might have been observed. This infinite set of time series is sometimes called the ensemble. Every member of the ensemble is a possible realization of the stochastic process. The observed time series can be thought of as one particular realization, and will be denoted by $x(t)$ for $(0 \leqq t \leqq T)$ if the observations are continuous, and by x_t for $t=1, \cdots, N$, if the observations are discrete.

Time series analysis is essentially concerned with evaluating the properties of the probability model which generated the observed time series. One way of describing a stochastic process is to specify the joint probability distribution of $X_{t1}, \cdots, X_{tn}$ for any set of time $t_1, \cdots, t_n$ and any value of n. But this is rather complicated and is not usually attempted in practice. A simpler way of describing a stochastic process is to give the moments of the process, particularly the first and second moments which are called the mean, variance and covariance functions.

Higher moments of a stochastic process may be defined in an obvious way, but are rarely used in practice, since a knowlege of two functions of mean and covariance in sufficient for the Gaussian processes and even for non-Gaussian processes the estimated higher moments are usually unreliable.

5.2 *Stationary processes*

An important class of stochastic processes for describing time series is the so called stationary processes. A mathematical definition of a stationary time series will be given later on. However it is now convenient to introduce the idea of stationarity from an intuitive point of view. Roughly speaking, a time series is said to be stationary if there is no systematic change in mean (no trend), no systematic change in variance, and no periodic variations.

Most of the theories developed in time series analysis are based on stationarity, and for this reason time series analysis often requires one to transform a nonstationary series into a stationary one so as to use these theories. For example one may remove the trend and seasonal variation from a set of data and then try to model the variation in the residuals by means of a stationary stochastic process.

A time series is said to be strictly stationary if the joint distribution of $X(t_1), \cdots, X(t_n)$ is the same as the joint distribution of $X(t_1+\tau), \cdots, X(t_n+\tau)$ for all $t_1, \cdots, t_n, \tau$. In other words shifting the time origin by an anount τ has no effect on the joint distributions which must therefore depend only on the intervals between $t_1, \cdots, t_n$. The above definition holds for any value of n. In particular, if $n=1$, it implies that the distribution of $X(t)$ must be same for all t so that

$$\mu(t)=\mu, \tag{5.1}$$

$$\sigma^2(t)=\sigma^2 \tag{5.2}$$

are both constants which do not depend on the value of t. Here $\mu(t)$ and $\sigma^2(t)$ are defined by

$$\mu(t)=E(X_t), \tag{5.3}$$

and

$$\sigma^2(t)=Var(X_t), \tag{5.4}$$

where $E(\cdot)$ and $Var(\cdot)$ denote expectation and variance.

5.3 *Autoregressive processes*

Suppose that $\{Z_t\}$ is a purely random process with mean zero and variance σ_z^2. Then a process $\{X_t\}$ is said to be an autoregressive process of order p if

$$X_t=\alpha_1 X_{t-1}+\cdots+\alpha_p X_{t-p}+Z_t, \tag{5.5}$$

It is a kind of a multiple regression model, but X_t is regressed not on indepen-

dent variables but on part values of X_t itself. Here the term autoregressive. An autoregressive process of order p will be abbleviated to an AR(p) process.

5.3.1 First-order case

For simplicity, we begin by examining the first-order case, where $p=1$. Then

$$X_t=\alpha X_{t-1}+Z_t. \tag{5.6}$$

The AR(1) process is sometimes called a Markov process. By successive substitution in Eq. 5.6 we may write

$$\begin{aligned} X_t&=\alpha[\alpha X_{t-2}+Z_{t-1}]+Z_t \\ &=\alpha^2[\alpha X_{t-3}+Zt_{-2}]+\alpha Z_{t-1}+Z_t \end{aligned} \tag{5.7}$$

and eventually we find that X_t may be expressed as an infinite-order MA process in the form (provided $-1<\alpha<1$)

$$X_t=Z_t+\alpha Z_{t-1}+\alpha^2 Z_{t-2}+\cdots. \tag{5.8}$$

It can be seen that, a first order MA process can be expressed by an AR model of infinite order. This duality between AR and MA processes is useful for a variety of purposes.

Rather than use successive substitution, it is simpler to use the backward shift operator B. Then Eq. 5.6 may be written

$$(1-\alpha B)X_t=Z_t \tag{5.9}$$

so that

$$\begin{aligned} X_t&=Z_t/(1-\alpha B) \\ &=(1+\alpha B+\alpha^2 B^2+\cdots)Z_t \\ &=Z_t+\alpha Z_{t-1}+\alpha^2 Z_{t-2}+\cdots. \end{aligned} \tag{5.10}$$

Then we find

$$E(X_t)=0 \tag{5.11}$$

and

$$Var(X_t)=\sigma_2{}^2\cdot(1+\alpha^2+\alpha^4+\cdots). \tag{5.12}$$

Thus the variance is finite provided that $|\alpha|<1$, in which case

$$Var(X_t)=\sigma_2{}^2/(1-\alpha^2)=(\sigma_2)^2 \tag{5.13}$$

5.3.2 General-order case

As in the first-order case, we can express an AR process of finite order as a MA process of infinite order. This may be done by successive substitution, or

by using the backward shift operator. Then Eq. 5.5 may be written an

$$(1-\alpha_1 B-\cdots-\alpha_p B^p)X_t=Z_t, \tag{5.14}$$

or

$$\begin{aligned} X_t &= Z_t/(1-\alpha_1 B-\cdots-\alpha_p B^p) \\ &= f(B)Z_t, \end{aligned} \tag{5.15}$$

where

$$\begin{aligned} f(B) &= (1-\alpha_1 B-\cdots-\alpha_p B^p)^{-1} \\ &= (1+\beta_1 B+\beta_2 B^2+\cdots). \end{aligned} \tag{5.16}$$

The relationship between the α's and the β's may then be found. Having expressed X_t as a MA process of infinite order, it follows that $E(X_t)=0$. The variance is finite provided that $\sum(\beta_i)^2$ converges, and this is a necessary condition for stationarity.

An equivalent way to express the stationarity condition is to say that the roots of the equation

$$\phi(B)=1-\alpha_1 B-\cdots-\alpha_p B^p=0 \tag{5.17}$$

must lie outside the unit circle (Box and Jenkins, 1970)

AR models have been applied to many situations in which it is resonable to assume that the present value of a time series can be reasonably approximated by a weighted average of the past values.

In practice most time series are non-stationary. In order to apply a stationary model discussed in the above section, it is necessary to remove nonstationarity.

Ozaki and Tong (1975) extended the AR model fitting procedure developed by Akaike (1969, 1970, 1976) to non-stationry situation. They considered a locally stationary process and fitted a stationary AR model to each stationary block of the data. The goodness of fit of the global model composed of these local stationary models is measured by the corresponding AIC and the partition of the data into blocks which minimizes the AIC specifies the best model. The homogeneity of the data is checked each time as a block of prescribed number of new data is added and the additional one is pooled to the original one if these two blocks of data are considered to be homogeneous. Otherwise a new process of modeling starts with the new block.

Next we will review the minimum AIC procedure for the fitting of the locally stationary AR model proposed by Kitagawa and Akaike (1978). Ozaki and Tong (1975) used the conventional technique of fitting of AR model de-

scribed in Akaike (1970), which is developed for the analysis of long stationary data and is not quite suitable for the application to a non-stationary situation where the analysis of a short span of data is necessary.

5.4 A procedure for the modeling of non-stationry time series

Kitagawa and Akaike (1978) proposed a new procedure based on the Householder transformation, which is a powerful tool for the solution of the least squares problem. This transformation allows the necessary modification of a fitted model due to the addition of observations or the addition and deletion of regressors quite easily. Especially with this procedure even the change of the mean value of the process can be handled very easily.

Given a set of observation $\{x_1, \cdots, x_N\}$, they considered the situation where the time interval $[1, N]$ is divided into k blocks, each of length $n_i (n_1+n_2+\cdots+n_k=N$; k and n_i are unkonwn), and the following locally stationary AR model is being fitted to the data :

$$x_n = a_0^i + \sum_{m=1}^{M(i)} a_m^i x_{n-m} + \varepsilon_n^i, \quad \left(\sum_{j=1}^{i-1} n_j + 1 \leqq n \leqq \sum_{j=1}^{i} n_j ;\ 1, \cdots, k\right) \tag{5.18}$$

where ε_n^i is a Gaussian white noise with $E(\varepsilon_n^i)=0$, $E(\varepsilon_n^i)^2=\sigma_i^2$ and $E(\varepsilon_n^i x_{n-m})=(m>0)$.

Ignoring the inital $M(1)$ observations, the likelihood of the locally stationary AR model is defined by

$$\prod_{i=1}^{k} (2\pi\ \sigma_i^2)^{-n_i/2} \exp\left\{-(2\sigma_i^2) \sum_{n=p_i}^{q_i} \left(x_n - a_0^i - \sum_{m=1}^{M(i)} a_m^j x_{n-m}\right)^2\right\}, \tag{5.19}$$

where $p_1=M(1)+1$, $p_i=1+\sum_{j=1}^{i-1} n_j$, $q_i=\sum_{j=1}^{i} n_j$.

For simplicity of notation N and n_1 are replaced by $N-M(1)$ and $n_1-M(1)$, respectively. Thus by denting $a^i=(a_0^i,\ a_1^i, \cdots, a_{M(i)}^i)$, the logarithm of the approximate likelihood is given by

$$L(x\ ;\ k,\ n_i,\ M(i),\ a^i,\ \sigma_i^2,\quad (i=1, \cdots, k)) = (-1/2) \sum_{i=1}^{k} \left\{ n_i \log 2\pi\ \sigma_i{}^2 + 1/\sigma_i{}^2 \sum_{n=p_i}^{q_i} \left(x_n - a_0^i - \sum_{m=1}^{M(i)} a_m^i x_{n-m}\right)^2 \right\}. \tag{5.20}$$

For any given a_m^i's the maximum of the function L is attained at

$$\sigma_i^2 = 1/n_i \sum_{n=p_i}^{q_i} \left(x_n - a_0^i - \sum_{m=1}^{M(i)} a_m^i x_{n-m}\right)^2. \tag{5.21}$$

Therefore the overall maximum of the log-likelihood L is given by

$$L^*(x\,;\ k,\ n_i,\ M(i),\quad (=1,\cdots,k))$$
$$=(-1/2)\sum_{i=1}^{k}\{n_i \log 2\pi\ \bar{\sigma}_i{}^2+n_i\}$$
$$=-N/2(1+\log 2\pi)-1/2\sum_{i=1}^{k} n_i \log\ \bar{\sigma}_i{}^2, \tag{5.22}$$

where $\bar{\sigma}_i{}^2$ is the approximate maximum likelihood estimate of $\sigma_i{}^2$ and is obtained by minimizing $\sigma_i{}^2$ with respect to a_m^i's.

AIC (Akaike Information Criterion) was proposed for the selection of the best statistical model (Akaike, 1973). Since AIC is obtained by estimating the Kullback-Leibler information number (Kullback and Leibler, 1951) of the true distribution with respect to the assumed model, it can be considered as a reasonable measure of the badness of the estimated model. In the case of locally stationnary AR model, AIC is given by

$$\text{AIC}=-2(\text{maximum log likelihood})+2(\text{number of parameters})$$
$$\text{AIC}=\sum_{i=1}^{k} n_i \log\ \bar{\sigma}_i{}^2+2\left(\sum_{i=1}^{k} M(i)+2\right). \tag{5.23}$$

The minimum AIC estimates (MAICE) of the number of stationary blocks, the size of each block and the order of the AR model fitted to each block are defined as those values of k, n_i $(i=1,\cdots,k)$ and $\mathrm{M}(i)$ $(i=1,\cdots,k)$ which minimize the AIC.

5.5 *On-line type fitting procedure*

Consider the situation where an AR model, AR_0, has been fitted to the set of data $(x_1,\cdots,x_n\}$ and an additional set of m observations $(x_{n+1},\cdots,x_{n+m}\}$ is newly obtained where m is a prescribed number. We consider two competing models. The first one is defined by connecting two different AR models, the model AR_0 and the model AR_1 which is fitted to the newly obtained data $\{x_{n+1},\cdots,x_{n+m}\}$. The AIC of this jointed model is given by

$$\text{AIC}_1=\log \sigma_0{}^2+m \log \sigma_1{}^2+2(M_0+M_1+4), \tag{5.24}$$

where $\sigma_0{}^2$ and M_0 are the innovation variance and the order of the AR model AR_0, $\sigma_1{}^2$ and M_1 are those of AR_1, respectively. The second model is an AR_2, fitted to the whole span of the pooled data $\{x_1,\cdots,x_{n+m}\}$. The AIC of this model is given by

$$\text{AIC}_2=(n+m)\log \sigma_2{}^2+2(M_2+2), \tag{5.25}$$

where $\sigma_2{}^2$ is the innovation variance and M_2 is the order of the fitted model.

If AIC_1 is less than AIC_2, we switch to the new model AR_1. Otherwise, the

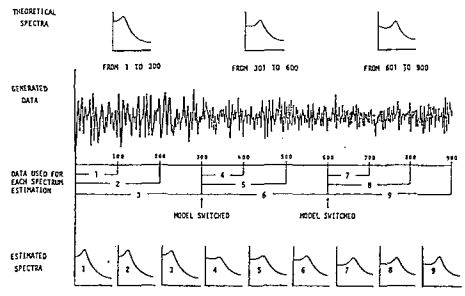


Fig. 5.1. Example of power spectral density function estimation. Minimum AIC Estimation (MAICE) with the locally stationary autoregressive models at each block is compared with the theoretical spectrum. Data length $N=900$, basic span $m=100$, highest order $k=5$. The axes of abscissa of the spectra: normalized frequencies by sampling rate ($f=0.0$ to 0.5). (after Fig. 2 of Kitagawa and Akaike, 1978).

two sets of data are considered to be homogeneous and the model AR_2 is accepted. The procedure repeats these steps whenever a set of m new observations is given (see Fig. 5.1). We will call m as the basic span. The procedure is so designed as to follow the change of the structure of the time series, while if the structure remains unchanged, it will improve the model by using the additional observations.

5.6 Least squares computation by Householder transformation

In this section, we will present a computationary efficient method for the fitting of AR models. The efficiency of the algorithm is especially important in the fitting of locally AR model.

First we assume that the mean value of the process $\{x_n\}$ is zero. The least squares estimates of the k th order AR coefficients are obtained by minimizing the sum of squares

$$1/(N-K)\sum_{n=k+1}^{N}(x_n-a_ix_{n-1}-\cdots-a_kx_{n-k})^2 \tag{5.26}$$

To solve this least squares problem, defined the matrix $\boldsymbol{X}$ and the vectors $\boldsymbol{y}$ and $\boldsymbol{a}$ by

$$\boldsymbol{X}=\begin{bmatrix} x_k & x_{k-1} & \cdots & x_1 \\ x_{k+1} & x_k & \cdots & x_2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{N-1} & x_{N-2} & \cdots & x_{N-K} \end{bmatrix}, \tag{5.27}$$

$$\boldsymbol{y}=\begin{bmatrix} x_{k+1} \\ x_{k+2} \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix}, \tag{5.28}$$

and

$$\boldsymbol{a}=\begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_k \end{bmatrix} \tag{5.29}$$

The least squares method minimizes $\| \boldsymbol{y}-\boldsymbol{X}\boldsymbol{a} \|$, where $\| \cdot \|$ denotes the Euclidean norm, and the solution is given by the normal equation

$$\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{a}=\boldsymbol{X}^T\boldsymbol{y}. \tag{5.30}$$

As a numerical procedure, the direct solution of the normal equation is not quite efficient and the procedure realized by first orthogonalizing the column vectors of $\boldsymbol{X}$ and then solving the resultant equation supersedes the normal equation approach in both manipulability and numerical accuracy (Golub, 1965).

A Householder transformation is an orthogonal transformation defined by a matrix $\boldsymbol{P}=\boldsymbol{I}-2\boldsymbol{u}\boldsymbol{u}^T$, where $\boldsymbol{u}$ is an arbitrary vector with $\| \boldsymbol{u} \|=1$. Let $\boldsymbol{X}_{(1)}=\boldsymbol{X}$ and $\boldsymbol{X}_{(2)}, \cdots, \boldsymbol{X}_{(k+1)}$ be defined by $\boldsymbol{X}_{(k+1)}=\boldsymbol{P}_{(k)}\boldsymbol{X}_{(k)}$ $(k=1, \cdots, K)$, where $\boldsymbol{P}_{(k)}$ is a Householder transformation and is chosen so that $x_{i,k}^{(k+1)}=0$ $(i=k+1, k+2, \cdots, N-K)$, where $x_{i,k}^{(k+1)}$ denotes (i, k)th element of $\boldsymbol{X}^{(k+1)}$. Such a $\boldsymbol{P}^{(k)}$ is obtained by $\boldsymbol{P}^{(k)}=\boldsymbol{I}-\boldsymbol{v}_k\boldsymbol{v}_k^T/h_k$ with

$$\boldsymbol{v}_k{}^t=(0, \cdots, 0, x_{k,k}{}^{(k)}\pm\tau, x_{k+1,k}{}^{(k)}, \cdots, x_{N-K,k}{}^{(k)}), \tag{5.31}$$

where $\tau^2=\sum_{i=1}^{N-K}\{x_{i,k}{}^{(k)}\}^2$ and $h_k=\tau^2\mp\tau x_{k,k}{}^{(k)}$ (Kitagawa and Akike, 1978). After the application of the Householder transforations for k times

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{Q}\boldsymbol{X}=\boldsymbol{R}=\begin{bmatrix} \boldsymbol{S} \\ - \\ \boldsymbol{0} \end{bmatrix}, \tag{5.32}$$

where $\boldsymbol{Q}=\boldsymbol{P}^{(k)}\boldsymbol{P}^{(k-1)}\boldsymbol{P}\cdots\boldsymbol{P}^{(1)}$ is an orthogonal matrix and $\boldsymbol{S}$ is an upper triangular matrix. The matrix $\boldsymbol{S}$ and the vector $\boldsymbol{Q}\boldsymbol{y}$ keep the complete information

for the least squares fitting of AR models up the K-th order.
Denote $\boldsymbol{S}$ and $\boldsymbol{z}=\boldsymbol{Qy}$ by

$$\boldsymbol{S}=\begin{bmatrix} S_{11} & \cdots & S_{1k} \\ & \ddots & \vdots \\ \mathbf{0} & & S_{kk} \end{bmatrix}, \tag{5.33}$$

and

$$\boldsymbol{z}^T=(z_1, \cdots, z_{n-k}). \tag{5.34}$$

For each $k \leqq K$, the least squares estimates of the coefficients of the k-th order AR model

$$x_n=\sum_{i=1}^{k} a_i^k x_{n-1}+\varepsilon_n^k \tag{5.35}$$

are obtained by solving the linear equation

$$\begin{bmatrix} S_{11} & \cdots & S_{1k} \\ & \ddots & \vdots \\ \mathbf{0} & & S_{kk} \end{bmatrix} \begin{bmatrix} a_1^k \\ \vdots \\ a_k^k \end{bmatrix} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_k \end{bmatrix}. \tag{5.36}$$

The corresponding estimate $d(k)$ of $\sigma_k{}^2=E(\varepsilon_n^k)^2$ is given by Golub (1965)

$$d(k)=1/(N-K)\cdot\sum_{i=k+1}^{N-K} z_i{}^2. \tag{5.37}$$

We note that for the fitting of the generalized model with a constant term a_0

$$x_n=a_0+\sum_{m=1}^{k} a_m x_{n-m}+\varepsilon_n, \quad (k=1, \cdots, K). \tag{5.38}$$

We have only to define $\boldsymbol{X}$ by

$$\boldsymbol{X}=\begin{bmatrix} 1 & x_k & \cdots & x_1 \\ 1 & x_{k+1} & \cdots & x_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N-1} & \cdots & x_{N-K} \end{bmatrix}. \tag{5.39}$$

Under the Gaussian assumption, since the least squares estimates are reasonable approximations to the maximum likelihood estimates, the minimum AIC procedure for the fitting of an AR model can be defined as follows.

1) Assume that a set of data $\{x_n : n=1, \cdots, N\}$ is given.

2) Replace x_n by $x_n-\bar{x}$, where $\bar{x}=1/N \sum_{n=1}^{N} x_n$

3) Determine the upper limit K of the order of AR models to be fitted to the data.

4) Define the $(N-K)\times(K+1)$ matrix

$$\boldsymbol{X}=\begin{bmatrix} x_K & x_{K-1} & \cdots & x_1 & x_{K+1} \\ x_{K+1} & x_K & \cdots & x_2 & x_{K+2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ k_{N-1} & x_{N-2} & \cdots & x_{N-K} & x_N \end{bmatrix}. \tag{5.40}$$

5) Reduce the matrix $\boldsymbol{X}$ to an upper triangular matrix

$$\boldsymbol{S}=\begin{bmatrix} S_{11} & \cdots & S_{1K} & S_{1,K+1} \\ & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \\ & & & \cdot \\ & & S_{KK} & S_{K,K+1} \\ \mathbf{0} & & & S_{K+1,K+1} \end{bmatrix} \tag{5.41}$$

by the successive application of the Householder transformation described in the preceding section.

6) Compute the AIC of the AR model of order m by

$$\mathrm{AIC}(m)=(N-K)\log(d(m))+2(m+2), \tag{5.42}$$

where $d(m)$, the least squares estimate of the innovation variance, is in this case given by

$$d(m)=1/(N-K)\cdot \sum_{i=m+1}^{K+1} s_{i,K+1}^2. \tag{5.43}$$

7) Adopt the m which gives the minimum of $\mathrm{AIC}(m)$ $(m=0, 1, \cdots, K)$ as the order L of the model.

8) The minimum AIC estimates of the AR coefficients a_i^L $(i=1, \cdots, L)$ are obtained by

$$a_L^L=s_{L,L}^{-1}s_{L,K+1}, \tag{5.44}$$

and

$$a_i^L=s_{1,i}^{-1}\left(s_{i,K+1}-\sum_{j=i+1}^{L} a_j^L s_{i,j}\right) \quad (i=L-1, \cdots, 1). \tag{5.45}$$

5.7 *Minimum AIC estimation of locally stationary models*

In this section we will review an implementation of the minimum AIC procedure for the fitting of locally stationary AR model. Let the matrix $\boldsymbol{X}$ be defined by

$$\boldsymbol{X}=\begin{bmatrix} 1 & x_K & \cdots & x_1 & x_{K+1} \\ 1 & x_{K+1} & \cdots & x_2 & x_{K+2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{N0-1} & \cdots & x_{N0-K} & x_{N0} \end{bmatrix} \tag{5.46}$$

and is reduced to the upper triangular matrix $\boldsymbol{S}$ by the Householder transformation. When an additional set of observation $\{x_{N0+1}, \cdots, x_{N0+M}\}$ is obtained the matrix $\boldsymbol{Z}$ is constructed as

$$\boldsymbol{Z}=\begin{bmatrix} 1 & x_{N0} & \cdots & x_{N0-K+1} & x_{N0+1} \\ 1 & x_{N0+1} & \cdots & x_{N0-K+2} & x_{N0+2} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 1 & x_{N0+M-1} & \cdots & x_{N0+M-K} & x_{N0+M} \end{bmatrix}. \tag{5.47}$$

By repeated applications of the Householder transformation, the matrix $\boldsymbol{Z}$ is reduced to an upper triangular matrix $\boldsymbol{S}_1$ such as

$$\boldsymbol{Q}_1\boldsymbol{Z}=\begin{bmatrix} \boldsymbol{S}_1 \\ - \\ \boldsymbol{0} \end{bmatrix} \tag{5.48}$$

Again by applying the Householder transformations to the matrix

$$\begin{bmatrix} \boldsymbol{S} \\ - \\ \boldsymbol{S}_1 \end{bmatrix}, \tag{5.49}$$

we obtain an upper triangular matrix $\boldsymbol{S}_2$ by

$$\boldsymbol{P}\begin{bmatrix} \boldsymbol{S} \\ - \\ \boldsymbol{S}_1 \end{bmatrix}=\begin{bmatrix} \boldsymbol{S}_2 \\ - \\ \boldsymbol{0} \end{bmatrix}. \tag{5.50}$$

From the property of the orthogonal transformation, it can be seen that the

triangular matrix S_2 is one and the same as the one obtained by reducing the augmented matrix

$$\begin{bmatrix} X \\ - \\ Y \end{bmatrix}. \tag{5.51}$$

This means that the least squares estimates of the coefficients of an AR model obtained by pooling two consecutive time series can be obtained with little additional computations. The procedure by Ozaki and Tong (1975) assumes the zero initial and end conditions for each block. It is one of advantages of the present procedure that the fitting of an AR model is realized with the initial condition given by its preceding block. Another advantage is that it can be applied to the situation where the mean value of the process varies between blocks. The minimum AIC procedure for the fitting of a locally stationary AR model is summarized as follows.

1) Assume that a set of data $\{x_n;\ n=1, \cdots, N\}$ is given.

2) Let AIC (S) and AR (S) denote the minimum of AIC (m) among $m=0, 1, \cdots, K$ and the minimum AIC estimate of AR model obtained through the matrix S, respectively.

3) Set the upper limit K of the order of AR models and choose M, the length of the basic span of data.

4) Construct of $M\times(K+2)$ matrix

$$X=\begin{bmatrix} 1 & x_K & \cdots & k_1 & x_{K+1} \\ 1 & x_{K+1} & \cdots & k_2 & x_{K+2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & k_{K+M-1} & \cdots & x_M & x_{K+M} \end{bmatrix}. \tag{5.52}$$

5) Reduce the matrix X to an upper triangular matrix S_0.

6) Determine the minimum AIC autoregressive model AR(S_0) and put AIC$_0$ $\equiv$AIC(S_0).

7) Set $N_0=M$

8) Construct the $M\times(K+2)$ matrix

$$\boldsymbol{Y}=\begin{bmatrix} 1 & x_{N0+K} & \cdots & x_{N0+1} & x_{N0+K+1} \\ 1 & x_{N0+K+1} & \cdots & x_{N0+2} & x_{N0+K+2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{N0+K+M-1} & \cdots & x_{N0+M} & x_{N0+K+M} \end{bmatrix}. \tag{5.53}$$

9) Reduce the matrix $\boldsymbol{Y}$ to an upper triangular matrix $\boldsymbol{S}_1$.

10) Determine the minimum AIC autoregressive model AR($\boldsymbol{S}_1$) and put $\mathrm{AIC}_1 \equiv \mathrm{AIC}(\boldsymbol{S}_1)+\mathrm{AIC}_0$,

11) Construct of $2(K+2)\times(K+2)$ matrix

$$\boldsymbol{Z}=\begin{bmatrix} \boldsymbol{S}_0 \\ - \\ \boldsymbol{S}_1 \end{bmatrix}. \tag{5.54}$$

12) Reduce the matrix $\boldsymbol{Z}$ to an upper triangular matrix $\boldsymbol{S}_2$

13) Determine the minimum AIC autoregressive model AR($\boldsymbol{S}_2$) and put $\mathrm{AIC}_2 \equiv \mathrm{AIC}(\boldsymbol{S}_2)$.

14) If AIC_1 is less than AIC_2, replace the current model AR($\boldsymbol{S}_2$) by AR ($\boldsymbol{S}_1$), and set $\boldsymbol{S}_0=\boldsymbol{S}_1$ and $\mathrm{AIC}_0=\mathrm{AIC}_1$.

If AIC_1 is greater than or equal to AIC_2, replace the current model AR($\boldsymbol{S}_0$) by AR($\boldsymbol{S}_2$), and set $\boldsymbol{S}_0=\boldsymbol{S}_2$ and $\mathrm{AIC}_0=\mathrm{AIC}_2$.

15) If N_0+M becomes equal to N, stop the procedure, otherwise replace N_0 by N_0+M and go back to step 8).

The present proceure is useful for the automatic detection or monitoring of the change of spectral characteristics of industrial and natural processing. In the next chapter, we will develop the procedure to facilitate automatic determination of arrival time by an on-line system. The performance of the proposed procedure will be checked by applying it to weak seismic signals of microearthquakes contaminated with background noises.

6. A new efficient procedure for the estimation of onset times of P-waves

In this chapter a computationally efficient procdure is developed for the fitting of locally stationary autoregressive model. The amount of computations is bounded by a function of the data length and the model order only and does not depend on the number of possible arrival time. This facilitates automatic determination of arrival time by on-line system. On-line system FUNIMAR

(fast univariate case of minimum AIC method of AR model fitting) is developed to implement the procedure. The method is checked by applying to weak seismic signals from earthquakes, which are superimposed with the background noises, such as traffic noise, electronic hum noise, and heavy microtremor.

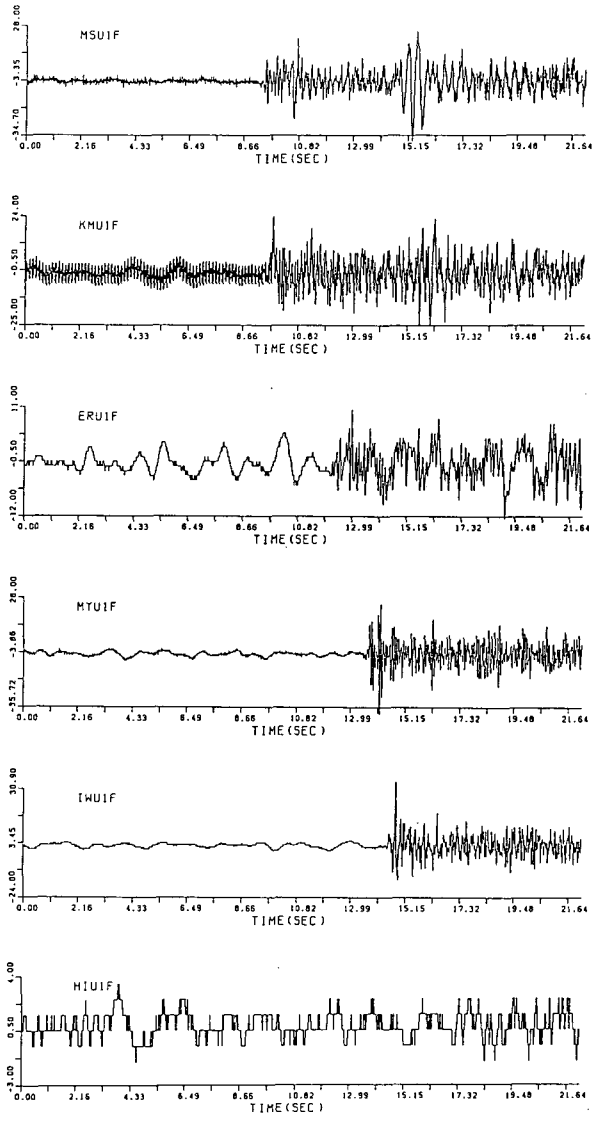
6.1 Introduction

In recent years, the automatic processing of seismic signals for the detection of seismic activity becomes realistic due to the establishment of well equipped nation-wide seismological network (Hamaguchi and Suzuki, 1979). In an earthquake prediction it is necessary to distinguish rapidly a foreshock sequence from background seismicity of a region which is facilitated by rapid location of earthquakes and processing large quantities of data. Neeldess to say, the seismic signals observed by seismometers are contaminated by various kinds of signals, such as microtremors, microseisms, and artificial vibrations.

These observed data sets have been traditionally handled by empirical methods based on the expertise of human operator to single out real seismic signals from the various noises. For the automatic processing of seismic signals, it thus becomes necessary to develop a method that can automatically detect seismic wave from noisy data. Some attempts have been made based on the autoregressive modeling of the observed seismic signals (Tjøsteim, 1975; Shirai and Tokuhiro, 1979; Hamaguchi and Suzuki, 1979; Yokota et al., 1981; Hamaguchi and Morita, 1980; Morita and Hamaguchi, 1981; Maeda, 1985, 1986; Hasegawa et al., 1986). An AR model is very useful for analysis of a stationary time series.

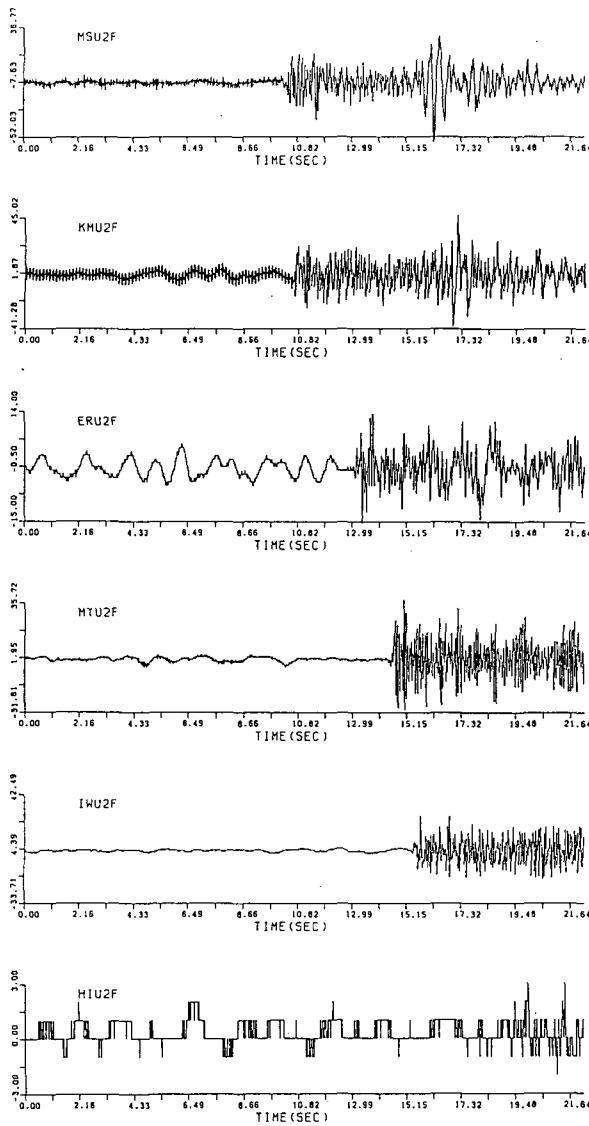
However from the statistical point of view, the main feature of the signals observed by seismometers is the non-stationarity. Although seismic waves are non-stationary, it may be reasonable to consider that it can be approximated by an AR model on each properly divided time interval (Ozaki and Tong, 1975; Kitagawa and Akaike, 1978). The use of the locally stationary AR model was thus motivated and it was shown that it is actually usful for detection of arrival of P-waves in the noisy data (Yokota et al., 1981). A significant merit of the time series method is that we can automatically determine the arrival time of the P-waves by just looking for the time point that attains the minimum value of the AIC (Akaike information criterion) of a locally stationary AR model, namely by finding the model that best fits to the observed data (Ozaki and Tong, 1975; Kitagawa and Akaike, 1978). The CPU-time of this method depends on the number of data (data lenght) and on the order of AR model.

The objective of this chapter is to develop a computationlly efficient

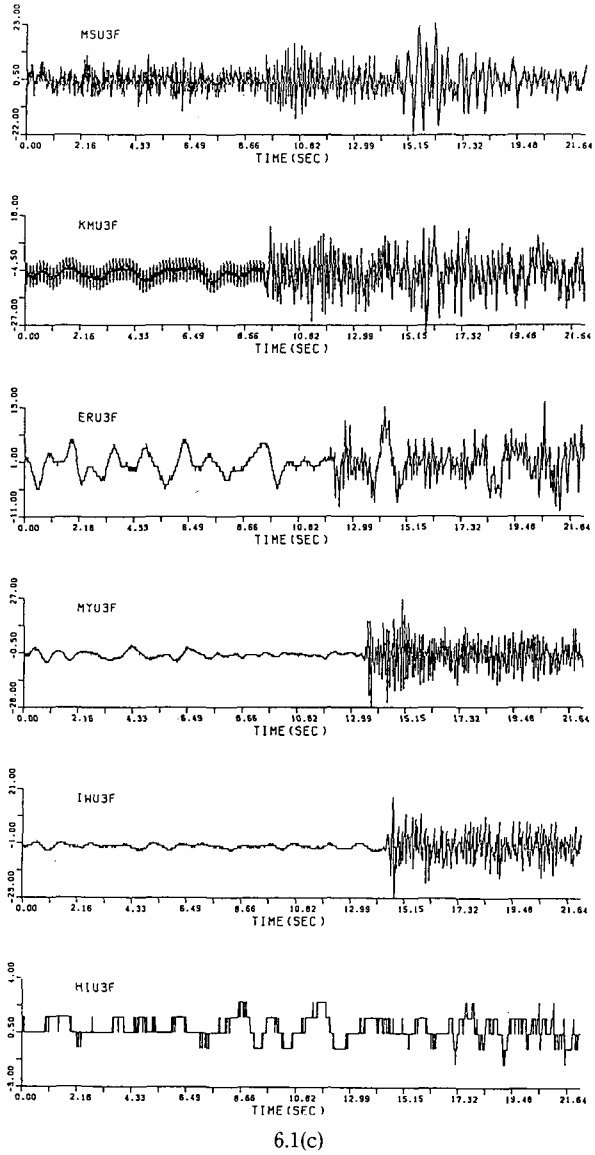


6.1(a)

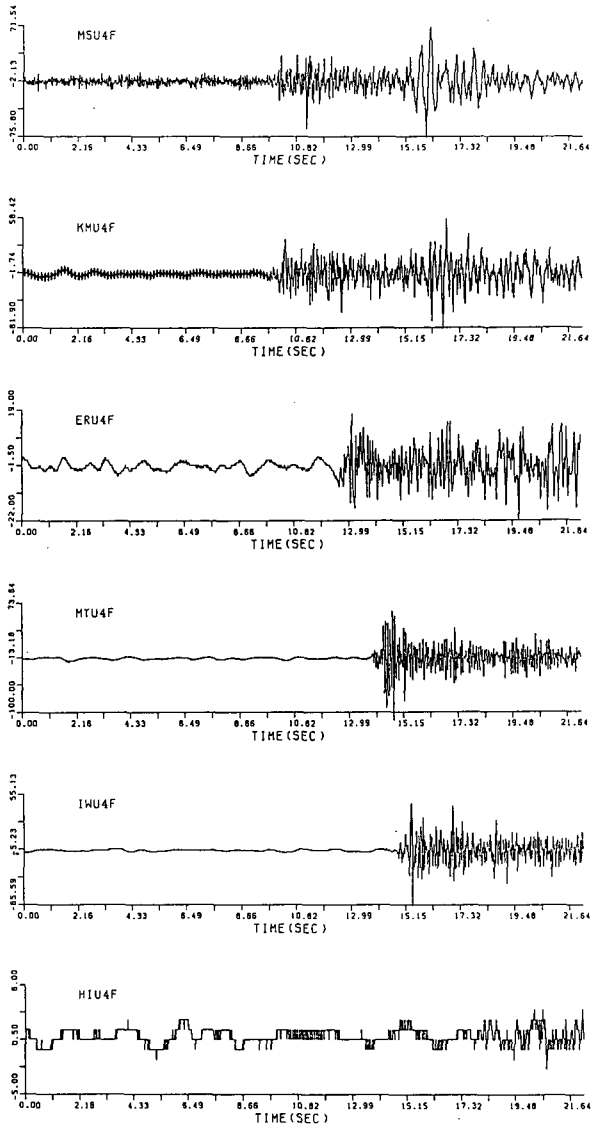
Fig. 6.1. Example of seismograms used in chapter 6. The parameters of earthquakes shown in Table 4.1. (a) seismograms of foreshock No. 1. (b) seismograms of foreshock No. 2. (c) seismograms of foreshock No. 3. (d) seismograms of foreshock No. 4. Label in the upper left-hand corner to seismic records is named after station, component (up-down), and number of foreshock.



6.1(b)



6.1(c)



6.1(d)

algorithm for the fitting of univariate locally stationary AR model so that it can apply to the on-line processing of seismic wave. The procedure is particularly useful for the automatic determination of the arrival time of P and S-waves. We apply the procedure to some microearthquakes occurred of Urakawa, Hokkaido, Japan (Fig. 6.1). At present, this procedure is successfully used in the on-line system of the Research Center for Earthquake Prediction (RCEP) of Hokkaido University (Suzuki et al., 1986).

6.2 Estimation of arrival time by a locally stationary AR model ; A review

In this section, we will review the use of time series model for the estimation of arrival time of a seismic wave. The seismometers are under the influence of various kinds of noises such as traffic noise, electronic hum noise, and heavy microtremor. As a result, the observed seismogram shows a random behavior, which is not in general a white noise sequence. This sometimes makes the decision of the arrival time difficult. However, it will be reasonable to assume that the spectral characteristics before and after the arrival of the seismic wave are quite different. From the point of view of time series modeling, this means that the models for the time series before and after the arrival of seismic wave are quite different. Since the spectrum of the time series can be well expressed by an appropriate AR model, it will be also reasonable to use an AR model for each time series. The use of locally stationary AR model is thus motivated. In the general definition of locally stationary AR model (Ozaki and Tong, 1975 ; Kitagawa and Akaike, 1978), the time interval is divided into k sub-intervals. However for the estimation of the arrival time, it seems sufficient to use the following simple model for which time interval is divided into two parts.

We assume that we have a time series $\{x_n\ ;\ n=1, \cdots, N\}$. This series can be divided into two subseries and each of them can be expressed by an AR model,

$$x_n = \sum_{m=1}^{M(1)} a_m^1 x_{n-m} + \varepsilon_n^1, \quad (1 \leqq n \leqq p_2),$$

and

$$x_n = \sum_{m=1}^{M(2)} a_m^2 x_{n-m} + \varepsilon_n^2, \quad (p_2 \leqq n \leqq N) \tag{6.1}$$

where ε_n^i is a Gaussian white noise with mean zero and variance ${\sigma_i}^2$. a_m^i is the autoregressive coefficient and $M(i)$ is the order of the i-th model. It should be noted that p_2 corresponds to the unknown arrival time. The above two AR models, background motion model and earthquake model constitute our locally

stationary AR model for the estimation of arrival time of P-wave.

Ignoring the initial distribution, the approximate likelihood of the locally stationary AR model is given by

$$\prod_{i=1}^{3}(2\pi\ \sigma_i^2)^{-N_i/2}\exp\left\{-(2\sigma_i^2)-\sum_{n=p_i}^{q_i}\left(x_n-\sum_{m=1}^{M(i)}a_m^i x_{n-m}\right)^2\right\}, \tag{6.2}$$

where $q_0=M(1)$, $q_1=p^2-1>$ $q_2=N$, $p_1=M(1)+1$, $N_i=q_i-q_{i-1}$.
Thus by denoting $a^i=(a_1^i, \cdots, a_{M(i)}^i)$, the approximate log-likelihood L is given by

$$\begin{aligned}&L(p_2,\ M(i),\ a^i,\ \sigma_i^2\quad (i=1,2))\\&\quad=-1/2\sum_{i=1}^{2}\left\{N_i\log 2\pi\ \sigma_i^2+\sigma_i^{-2}\sum_{n=p_i}^{q_i}\left(x_n-\sum_{m=1}^{M(i)}a_m^i x_{n-m}\right)^2\right\}.\end{aligned} \tag{6.3}$$

For arbitrarily given a_m^i's the maximum of the log-likelihood L is attained at

$$\sigma_i^2=1/N_i\sum_{n=p_i}^{q_i}\left(x_n-\sum_{m=1}^{M(i)}a_m^i x_{n-m}\right)^2. \tag{6.4}$$

Therefore, by substituting (6.4) into (6.3), the log−likelihood function for the estimation of a_m^i is reduced to the following form

$$\begin{aligned}&L^*(p_2,\ M(i),\ a_m^i,\quad (i=1,2))\\&\quad=-1/2\sum_{i=1}^{2}\{N_i\log 2\pi\sigma_i^2+N_i\}\\&\quad=-(N-q_0)/2(1+\log 2\pi)-1/2\sum_{i=1}^{2}N_i\log\sigma_i^2.\end{aligned} \tag{6.5}$$

The maximum likelihood estimate of a_m^i is obtained by maximizing (6.5) or equivalently minimizing σ_i^2 with respect to a_m^i's. It should be noted that the maximum likelihood estimate of two AR models is obtained by fitting two AR models independetly to $\{x_{p1}, \cdots, x_{p2-1}\}$ and $\{x_{p2}, \cdots, x_N\}$ by the least squares method. A computationary efficient Householder method used for AR model fitting was shown in chapter 6.6. So far we have assumed that the arrival time p_i and the model order $M(i)$ are given. But in practice, they are unknown and have to be determined from the data. We will determine these unknown constants so that the corresponding model best fits to the data. This can be realized by minimizing the AIC criterion. In the case of the locally stationary AR model, AIC is given by

$$\begin{aligned}&\mathrm{AIC}=-2\ (\text{maximum likelihood})+2(\text{number of parameters})\\&\quad=(N-q_0)(1+\log 2\pi)+\sum_{i=1}^{2}\{N_i\log\bar{\sigma}_i^2+2M(i)+2\}.\end{aligned} \tag{6.6}$$

In the following argument, it is indispensable to use one and the same data

set. We will thus set $q_0=\max M(1)$ instead of $M(1)$. However, we can ignore the first term of Eq. 6.6 since it is common to all possible models. It should be noted that $\bar{\sigma}_i^2$ depends on the selection of p_2 and $M(i)$. For fixed p_2, AIC is a function of the model orders, $M(1)$ and $M(2)$. Therefore by minimizing AIC, we can determined the best selection of orders. By substituting these into Eq. 6.6, the AIC becomes a function of only p_2. We will denote this AIC_{p2}. Then, by finding the p_2 which attains the minimum of AIC, we can get the best (in the sense of AIC) estimate of the arrival time of the seismic wave. We will call this, the minimum AIC estimate of the arival time.

Incidentally, $M(1)$, a_m^1, σ_1^2 and $M(2)$, a_m^2, σ_2^2 which correspond to this arrival time are the minimum AIC estimates of background and seismic wave models, respectively. By substituting these estimates into the theoretical relation between AR model and power spectrum

$$p_j(f)=\sigma_j^2/\left|1-\sum_{m=1}^{M(j)}a_m^j\exp(-2\pi imf)\right|^2,\quad (j=1,2), \tag{6.7}$$

we can get the power spectra of background noise and seismic wave (Kitagawa and Takanami, 1985). Here f is the frequency and i is the square root of (-1).

6.3 *A computationary efficient implementation of least squares fitting of locally stationary AR model*

In order to find the best estimate of the arrival time, we have to find the best noise model and the best earthquake model for each possible division of time interval, namely for all n_{p2} $(n_0\leqq n_{p2}\leqq n_1)$.

In this section, we will present a computationally efficient procedure based on Householder transformation, which is described in Chapter 5.6, that can yield these models successively.

To be specific, we assume that we have a time series $\{x_1, \cdots, x_N\}$ and prior to the analysis we know that the arrival time is in the interval $\{n_0, n_1\}$. It is also assumed that the required resolution is p points, namely we have a fit model for each possible division, n_0, n_0+p, $\cdots$, $n_0+lp\equiv n_1$. We first consider fitting AR model to the data set $\{x_1, \cdots, x_{n0}\}$. For this, we define the following $(n_0-k)\times(k+1)$ matrix X_0,

$$\boldsymbol{X}_0=\begin{bmatrix} x_k & x_{k-1} & \cdots & x_1 & \cdots & x_{k+1} \\ x_{k+1} & x_k & \cdots & x_2 & \cdots & x_{k+2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{n0-1} & x_{n0-2} & \cdots & x_{n0-k} & \cdots & x_{n0} \end{bmatrix} \tag{6.8}$$

The application of the the Householder transformation to the matrix $\boldsymbol{X}_0$ yields the following upper triangular matrix $\boldsymbol{R}_0$.

$$\boldsymbol{U}\boldsymbol{X}_0=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} & r_{1,k+1} \\ & r_{22} & \cdots & r_{2k} & r_{2,k+1} \\ & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & & \cdot \\ & \mathbf{0} & & & r_{k+1,k+1} \end{bmatrix}=\boldsymbol{R}_0. \tag{6.9}$$

Using this upper triangular matrix $\boldsymbol{R}_0$, the variance of the noises and AIC's of the AR models with order $j(j=0, \cdots, k)$ are obtained by

$$\sigma_0^2(j)=(n_0-k)^{-1}\sum_{i=j+1}^{k+1} r^2{}_{i,\,k+1} \qquad (j=0, 1, \cdots, k), \tag{6.10}$$

$$\mathrm{AIC}_0(j)=(n_0-k)\log\sigma_0^2(j)+(2j+1) \qquad (j=0, 1, \cdots, k). \tag{6.11}$$

For fixed n_0, $\mathrm{AIC}_0(j)$ is the criterion for the selection of best AR order $M(j)$ for the data set $\{x_1, \cdots, x_{n0}\}$. We then define,

$$\mathrm{AIC}_0^N=\min_j \{\mathrm{AIC}_0(j)\} \qquad (j=0, \cdots, k). \tag{6.12}$$

AIC_0^N is the AIC of the best noise model obtained under the assumption that the arrival time is n_0+1 and will be used later for the estimation of the arrival time.

Next we will fit an AR model to the augmented data set $\{x_1, \cdots, x_{n0+p}\}$. Obviously this can be done by defining $(n_0+p-k)\times(k+1)$ matrix $\boldsymbol{X}_1$ and reducing to an upper triangular form, $\boldsymbol{S}_1$, by a Householder transformation as follows,

$$\boldsymbol{X}_1=\begin{bmatrix} x_k & x_{k-1} & \cdots & x_1 & x_{k+1} \\ x_{k+1} & x_k & \cdots & x_2 & x_{k+2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{n0+p-1} & x_{n0+p-2} & \cdots & x_{n0+p-k} & x_{n0+p} \end{bmatrix}, \tag{6.13}$$

and

$$\boldsymbol{U}_1\boldsymbol{X}_1=\begin{bmatrix} s_{11}s_{12} & \cdots & s_{1k} & s_{1,k+1} \\ & s_{22} & \cdots & s_{2k} & s_{2,k+1} \\ & & \cdots & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & \cdot & \cdot \\ \mathbf{0} & & & & s_{k+1,k+1} \end{bmatrix}=\boldsymbol{S}_1. \tag{6.14}$$

However, by the property of the orthogonal transformation, it can be seen that one and the same upper triangular matrix is obtained by first organizing $(k+1+p)\times(k+1)$ matrix $\boldsymbol{R}_1$ and then reducing to an upper triangular form $\boldsymbol{S}_1'$ by a proper Householder transformation as follows,

$$\boldsymbol{R}_1=\begin{bmatrix} r_{11}r_{12} & \cdots & r_{1k} & r_{1,k+1} \\ & r_{22} & \cdots & r_{2k} & r_{2,k+1} \\ & & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot \\ & 0 & \cdots & \cdot & \cdot \\ & & & & r_{k+1,k+1} \\ x_{n0} & & \cdots & x_{n0+k-1} & x_{n0+1} \\ \cdot & & \cdots & \cdot & \cdot \\ \cdot & & \cdots & \cdot & \cdot \\ \cdot & & \cdots & \cdot & \cdot \\ x_{n0+p-1} & x_{n0+p-2} & & x_{n0+p-k} & x_{n0+p} \end{bmatrix}, \tag{6.15}$$

and

$$\boldsymbol{U}_1'\boldsymbol{R}_1=\begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1k} & s_{1,k+1} \\ & s_{22} & \cdots & s_{2k} & s_{2,k+1} \\ & & \cdots & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & & s_{k+1,k+1} \\ & & 0 & & \end{bmatrix}=\boldsymbol{S}_1'. \tag{6.16}$$

This is a very important modification. Note that the number of rows of the matrix $\boldsymbol{X}_1$ and $\boldsymbol{S}_1$ are n_0+p-k whereas those of $\boldsymbol{R}_1$ and $\boldsymbol{S}_1'$ are $k+p+1$. The variance of the noise and AIC of the AR models for the data set $\{x_1, \cdots, x_{n0+p}\}$ is obtained by

$$\sigma_0^2(j)(n_o-k+p)^{-1}\sum_{i=j+1}^{k+1}s_{i,k+1}^2 \qquad (j=0, 1, \cdots, k), \tag{6.17}$$

$$\mathrm{AIC}_1(j)=(n_0-k+p)\log\sigma_{n0+p}^2(j)+2(j+1) \qquad (j=0, 1, \cdots, k), \tag{6.18}$$

and

$$\mathrm{AIC}_1^N=\min_j\{\mathrm{AIC}_1(j)\} \qquad (j=0, \cdots, k). \tag{6.19}$$

We can repeat this method until we get entire AR models and corresponding AIC's for the data set $\{x_1, \cdots, x_{n0}\}$, $\{x_1, \cdots, x_{n0+p}\}$, $\cdots$, $\{x_1, \cdots, x_{n1}\}$. As the result, we obtained $\mathrm{AIC}_0{}^N$, $\mathrm{AIC}_1{}^N$, $\cdots$, $\mathrm{AIC}_W{}^N$.

In the same way, we can also get entire earthquake models. Namely, we first fit an AR model to the data set $\{x_{n1-p+1}, \cdots, x_N\}$, and obtain AIC_W^P. AIC_{W-1}^P is then obtained by augmenting $x_{n1-2p+1}, \cdots, x_{n1-p}$ in the same way as (6.16). Repeating this method W times, we obtain $\mathrm{AIC}_0{}^P, \cdots, \mathrm{AIC}_W^P$. Then the sum of two AIC's is as follows,

$$\mathrm{AIC}_i=\mathrm{AIC}_i{}^N+\mathrm{AIC}_i{}^P \qquad (\mathrm{i}=1, W), \tag{6.20}$$

which expresses the AIC of the locally stationary AR model for which the time interval is assumed to be divided into $\{1, n_o+ip\}$ and $\{n_0+ip+1, N\}$. This means that the arrival time is n_0+ip+1. Thus the best estimate of arrival time can be found by selecting the minimum among $\mathrm{AIC}_0, \cdots, \mathrm{AIC}_W$.

For the Householder transformation, the amount of necessary multiplication (and the addition) can be evaluated as

$$\begin{aligned}\sum_{i=1}^{k+1}(n+1-i)(k+2-i)&=\sum_{i=1}^{k+1}\{i^2-(n+k+3)i+(n+1)(k+2)\}\\&=(k+1)(k+2)(2k+3)/6-(n+k+3)(k+1)(k+2)/2+(n+1)(k+2)(k+1)\\&=(k+1)(k+2)(3n-k)/6\\&\sim nk^2/2,\end{aligned} \tag{6.21}$$

where n is the number of rows of the matrix $\boldsymbol{X}_1$ depends on the assumed arrival time. Therefore, the total amount of multiplication (and addition) for the comparison of all possible locally stationary AR models, is roughly of the order of

$$\begin{aligned}&\sum_{i=0}^{W}(n_0+ip)k^2/2+\sum_{i=0}^{W}(N-n_0-ip)k^2/2\\&=\{n_0k^2(W+1)+pk^2W(W+1)/2\}/2+\{(N-n_0)k^2(W+1)\\&\quad-pk^2W(W+1)/2\}/2\\&=Nk^2(W+1)/2\end{aligned}$$

$$\sim Nk^2(n_1-n_0)/2p. \tag{6.22}$$

On the other hand, for the Householder transformation (6.16), the amount of multiplication is

$$\begin{aligned}&\sum_{i=1}^{k+1}(p+1)(k+2-i)=\sum_{i=1}^{k+1}\{(p+1)(k+2)-(p+1)i\}\\&=(p+1)(k-2)(k+1)-(p+1)(k+1)(k+2)/2\\&=(p+1)(k+1)(k+2)/2\\&\sim pk^2/2\end{aligned} \tag{6.2}$$

Therefore the total amount of multiplication by the present method is

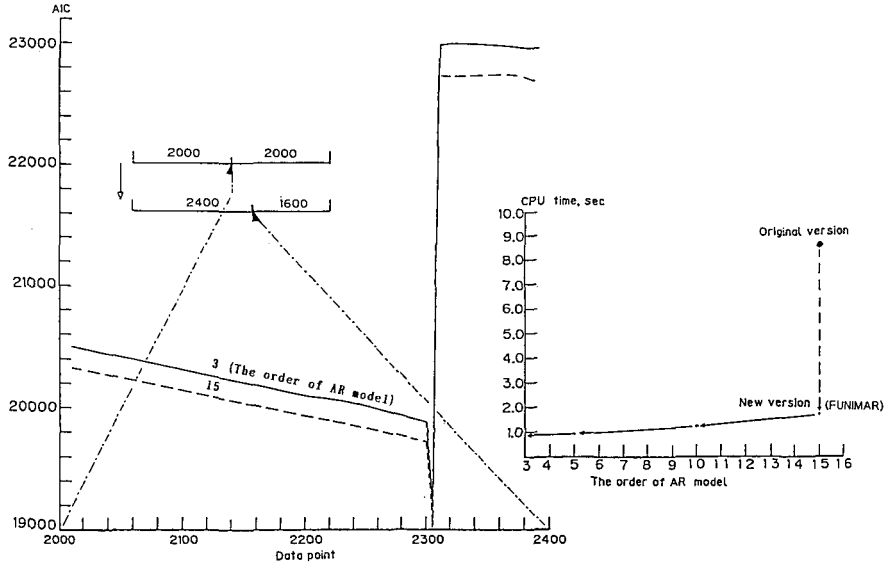


Fig. 6.2. Comparison of CPU times by conventional procedure and by FUNIMAR. Right: Comparison of the CPU times by traditional least squares computation by Eq. (6. 14) and by the newly developed least squares computation by Eq. (6. 16), FUNIMAR. Further comparison of CPU times according to the order of AR model is shown by using the procedure FUNIMAR. Left: Examples of AIC curves of the AR model of order 3 and of the AR model of order 15 (dashed line). The test data is an updown component seismogram recorded at station ESH. The parameters of hypocenter is shown in Table 4.1.

$$\begin{aligned} & n_0k^2/2+\sum_{i=1}^{W}pk^2/2+(N-n_1)k^2/2+\sum_{i=1}^{W}pk^2/2 \\ & =(N+n_0-n_1)k^2/2+pk^2W \\ & =(N+n_0-n_1)k^2/2+(n_1-n_0)k^2 \\ & =Nk^2/2+k^2(n_1-n_0)/2 \quad (<Nk^2). \end{aligned} \tag{6.24}$$

It should be noted that the amount of operations for the proposed method does not depend on the number of models, W. For example, if $N=3000$, $n_0=1000$, $n_1=2000$, $k=10$ and $p=2$, the necessary amount of multiplications are respectively

$$(3000)(10^2)(1000)/4=7.5\times10^7, \tag{6.25}$$

and

$$1/2(3000)(10^2)+1/2(10^2)(1000)=2\times10^5. \tag{6.26}$$

Therefore, by the present procedure, the amount of computation for Householder transformation is reduced to about 1/400 of the original procedure. The FORTRAN program FUNIMAR is developed to implement this procedure. The CPU-time for FUNIMAR is approximately one tenth of that for the time required by the original procedure as shown in Fig. 6.2.

6.4 Test of the proposed procedure on real seismograms

The procedure developed in the previous section has been applied to the seismograms observed at the stations of RCEP. As shown in section 2.2, ground velocity signal from geophone at each station is first digitized by an 8 bit nonlinear AD converter with 92.3 samples/sec (2400 bps/26 bit), then waves with frequency above 30 Hz were eliminated by an anti-aliasing filter (Butterworth filter of order 6).

For the illustration of the new procedure, we use seismograms of four foreshocks of the 1982 Off Urakawa Earthquake of M 7.1 (Sakai, 1984) recorded at six stations in the Hidaka region and those of an aftershock recorded at the station ESH. In particular, much attention is paid to noisy seismograms, which are contaminated with several kinds of noises, such as sudden traffic noise, electronic hum, or relatively strong microseisms (see Fig. 6.1). Table 4.1 summarizes the information about the earthquakes used in the present study. The magnitudes of these earthquakes are about 2. Locations of the epicenters and the seismic stations are shown in Fig. 4.11. They are located very close to the main shock. We also check the feasibility of estimating the arrival time of S-wave. Usually the first portion of S-wave train is superimposed on the strong

coda wave induced by P-wave. Thus it is not easy to decide the onset time of S-wave by eyes. The seismogram recorded at Esan (ESH) is used to examine whether the current precedure has the ability to pick out the correct onset time of S-wave disturbed by such P-coda wave.

6.4.1 The wave buried in traffic noise recorded at station Misono

The distances from the epicenters to the nearest station Misono (MSN) is about 30 km. The station MSN is located near a road and occasionally suffers from traffic noises (see Event 3 of Fig. 6.1). We first apply a band-pass filter of 0.08 Hz-7.0 Hz to enhance seismic signals (Fig. 6.3). Although the traffic noises are suppressed considerably by this filter, the onset times of P-waves in any of the three components are still ambiguous. We apply the present procedure to the unfiltered data recorded by the up-down component seismograph. AIC_n and the seismogram are shown in Fig. 6.4. AIC_n indicates the badness of the locally stationary AR model which assumes that the P-waves arrives at n_0+n_p-th time point. The arrow on this record indicates the arrival time determined by the AIC criterion. In this case the AIC_n has a clear minimum. It indicates

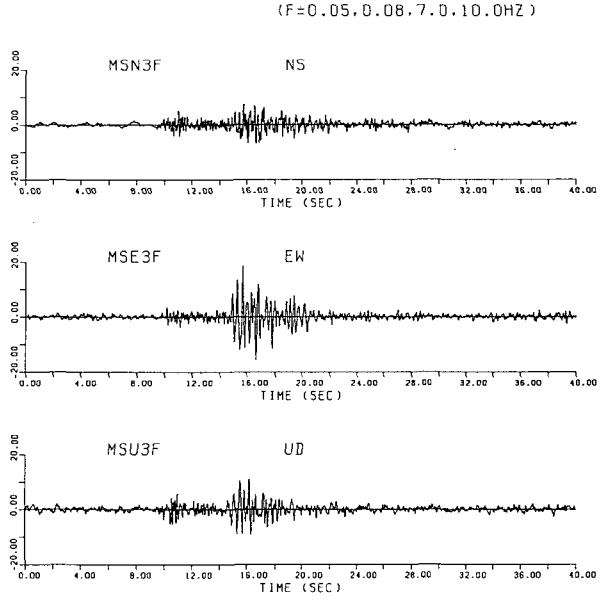


Fig. 6.3. Filtered three-component seismograms of foreshock No. 3 at station MSN. The frequency of the bandpass filter ranges between 0.08 and 7.0 Hz. The outside of the frequency range of spectra is tapered by half-cosine function. Note its unfiltered up-down component seismogram shown at the top of Fig. 6.1(c).

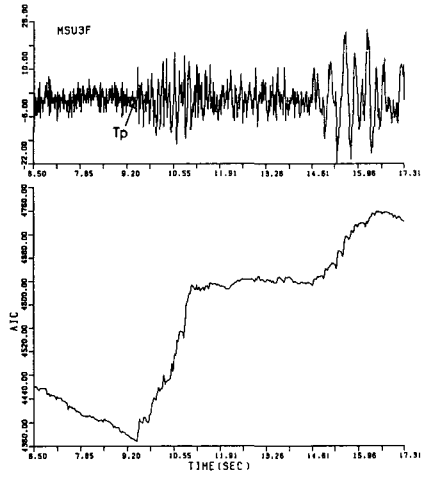


Fig. 6.4. The determined onset time of P-arrival of the event No. 3 recorded at MSN and AIC_n obtained by the FUNIMAR. An arrow indicates the onset time inferred from Minimum AIC Estimation (MAICE).

that we can get a good estimate of the onset time of P-wave. An implication of this result is that at least in this case, the AR model of the seismic signal is considerably different from that of the traffic noise. It should be noted that the procedure can determine the onset time of seismic arrival without prior noise reduction filter.

6.4.2 The seismogram superimposed on the strong microtremor noises recorded at station Erimo

The station Erimo (ERM) is located at a distance of 0.5 km from the shore line of the Pacific and has epicentral distance of about 50 km from the locations of the foreshocks. The ratios of seismic signal to the background noise are about 0.5 or less. Besides the geometrical spreading, they are strongly affected by attenuation in the crustal structure along the ray paths (e.g. Takanami, 1982). In the ordinary routine work, it is very hard to identify the onset times of such weak P-waves recorded on the paper chart. The AIC_n values together with the original up-down component seismogram of the event No. 3 are shown in Fig. 6.5. The arrow indicates the estimated onset time of P-wave, i.e. the time that corresponds to the minimum of AIC_n. The trace of time versus AIC_n shows a sharp wedge in the neighborhood of the minimum of AIC_n. This indicates that if the noise is the microtremor type we may get a good estimate of onset time of P-wave even though the signal to noise ratio is as low as 0.5 or less.

6.4.3 The weak seismogram recorded at station Hidaka

The station Hidaka (HIC) is located at a distance of 85 km from the epicenters of the foreshocks and is apart from towns and roads. The ampli-

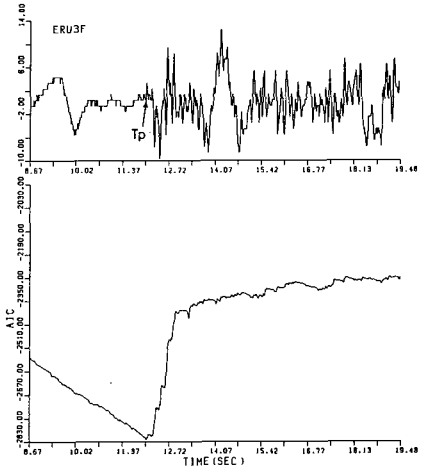


Fig. 6.5. The determined onset time of P-arrival of the event No. 3 recorded at ERM. Other conventions are the same as in Fig. 6.4.

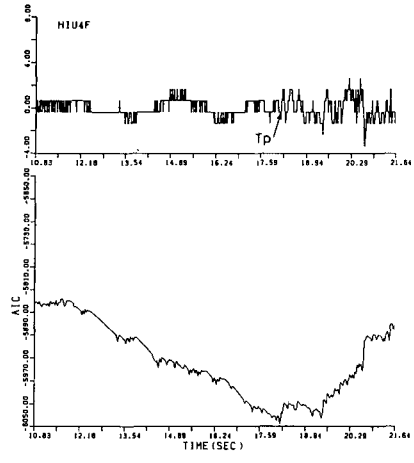


Fig. 6.6. The determined onset time of P-arrival of the event No. 4 recorded at HIC. Other conventions are the same as in Fig. 6.4.

tudes of the seismogram recorded at station HIC is very weak and almost equal to one LSB (least significant bit) of the digital data and is comparable to those of the background noises due to traffic, machinery, or oceanic sources. The typical seismogram and the corresponding AIC_n are shown in Fig. 6.6 The behavior of AIC_n is not so monotonous as those of previous examples. Two local minima of AIC_n are found at 18.18 sec and 19.40 sec as candidates of the P-arrival time. The arrow assigned on the seismogram corresponds to the smallest one of them.

6.4.4 The seismogram with strong hum noise recorded at station Kamikineusu

The station Kamikineusu (KMU) is located a distance of about 25 km from the epicenters of the foreshocks. On this occasion, the seismogram obtained at station KMU is contaminated by a strong electronic noise, namely, hum with a frequency of 50 Hz, which remains even after the application of the Butterworth low pass filters of order 6. As shown in Fig. 6.7 the seismic signals of the small foreshocks are unclear due to the presence of hum, and it is hard to decide the onset times of P-waves. The arrow in the figure indicates the estimate of the arrival time determined by the minimum AIC procedure. Considering the information from other stations, it is confirmed that the position of arrow actually indicates the onset time of the first P-arrival. This example is a corroboration that the present procedure works even under the presence of

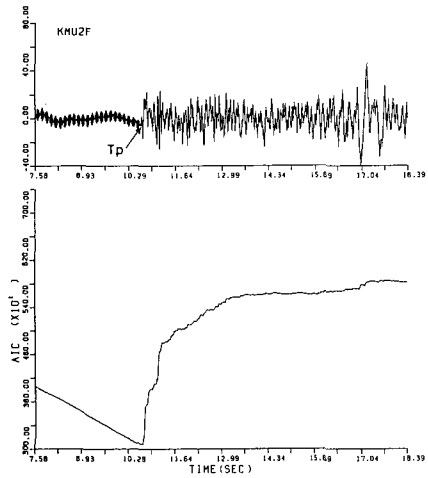


Fig. 6.7. The determined onset time of P-arrival of the event No. 2 recorded at KMU. Other conventions are the same as in Fig. 6.4.

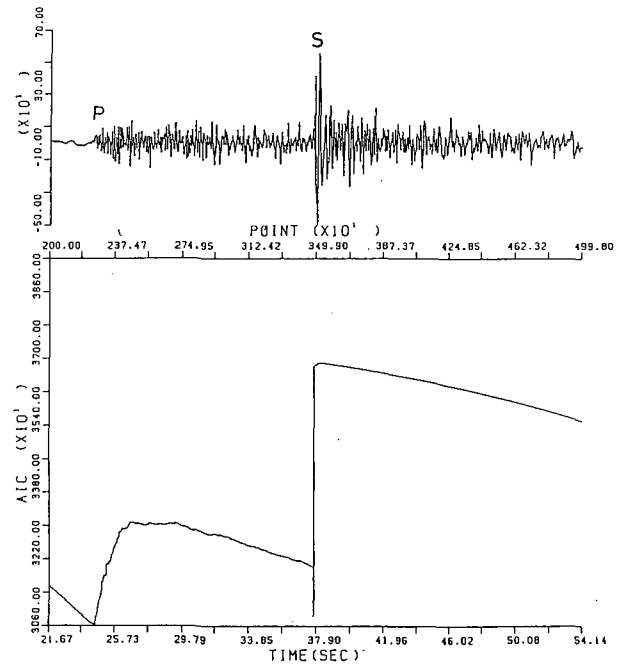
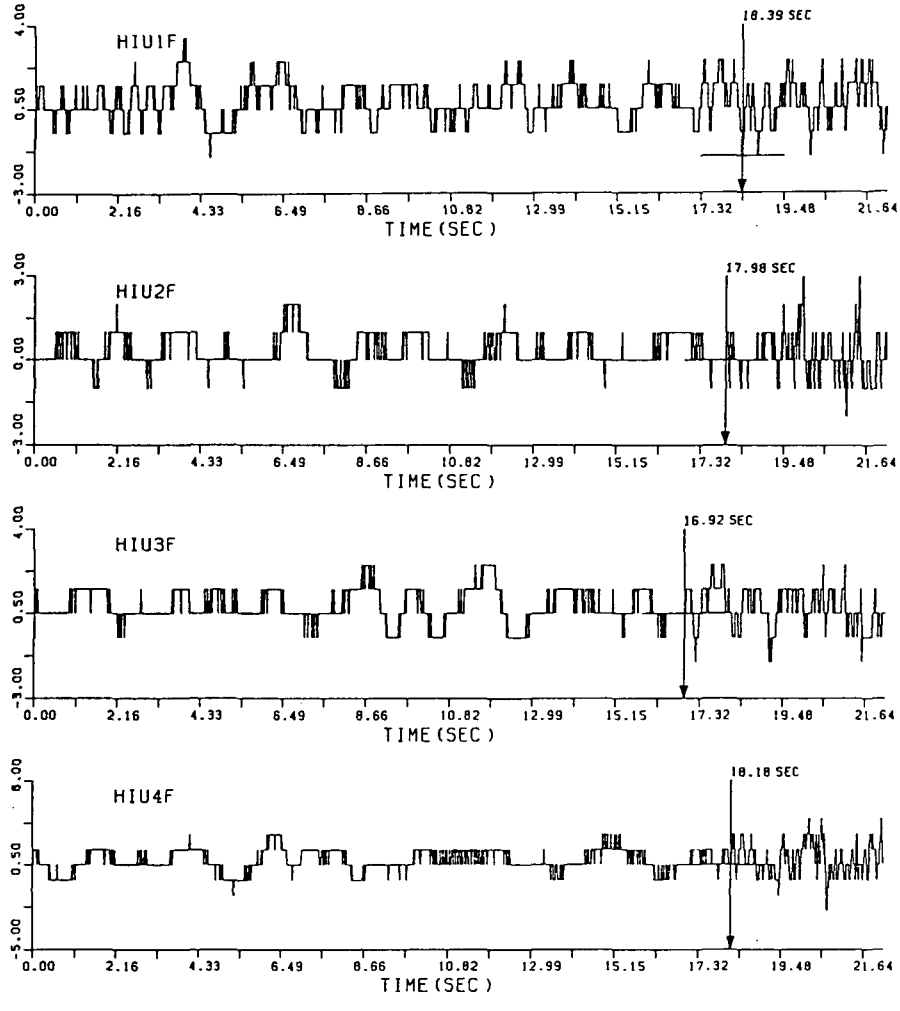


Fig. 6.8. The seismic record of ESH (aftershock No. 1 in Table 4.1) and AIC_n curve. The two clear minima in the AIC_n curve indicate the onset times of P-wave and S-wave, respectively.

strong hum noise.

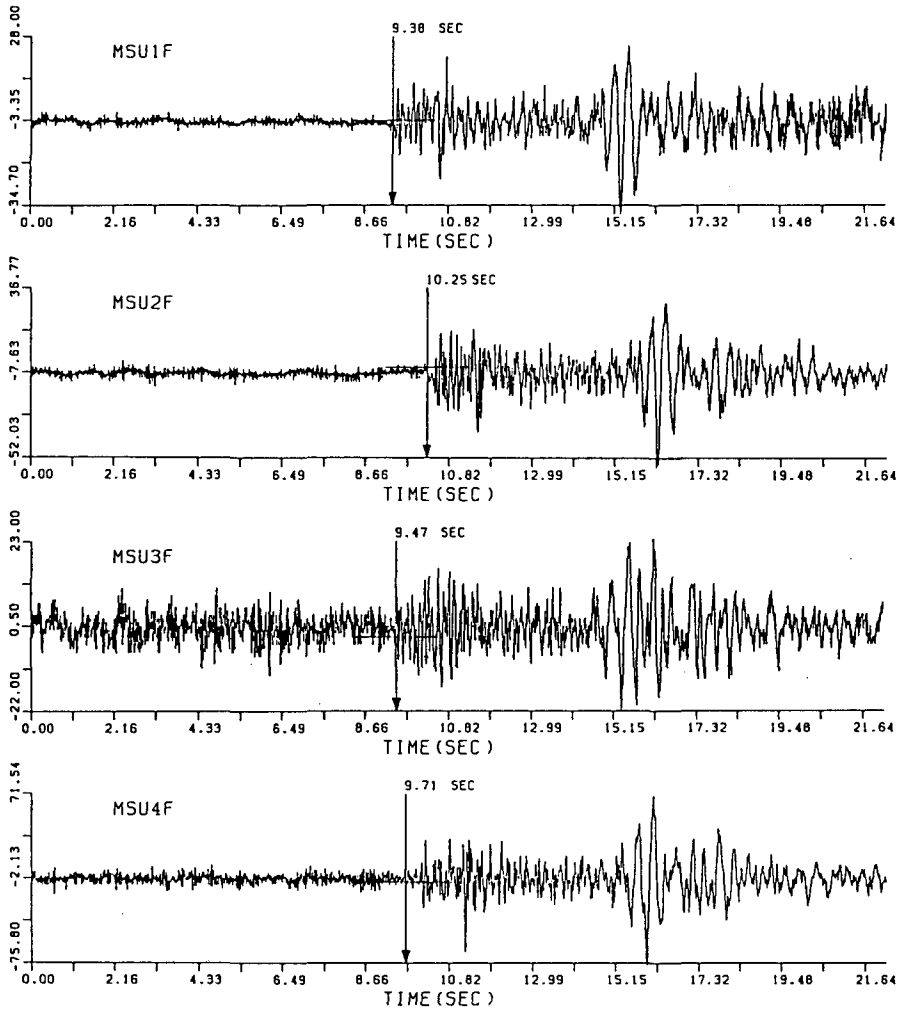
6.4.5 Availability for the estimation of the onset time of S-wave

The S-waves of local earthquakes, whose S-P times are not so large, are contaminated by the P-wave trains. However, it is expected that P and S-waves have different spectra and thus the locally stationary AR model can



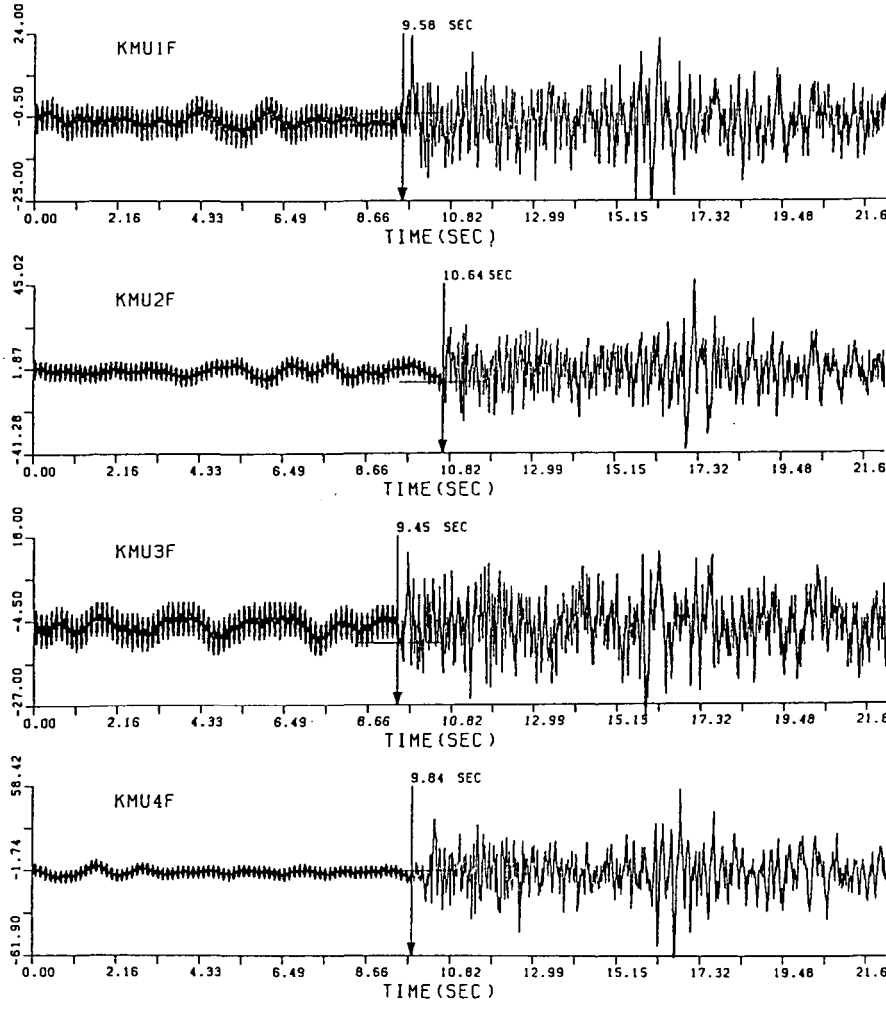
6.9(a)

Fig. 6.9. The onset times determined by using MAICE (Minimum AIC Estimation). The determined onset times of P-waves arrowed with crosses on seismic records. (a) station HIC ; (b) station MSN ; (c) station KMU ; (d) station ERM ; (e) station MYR ; (f) station IWN.



6.9(b)

distinguish these two waves. Figure 6.8 represents the seismogram recorded at the station ESH. This station is located west-south-west of the epicenter. The event used here is one of the aftershocks of the Off Urakawa earthquake and occusred on December 12, 1982. Figure 6.8 represents the two local minima of the AIC. The first local minimum of the AIC corresponds to the onset time of P-wave determined by the AR model of P-wave. The second one exactly corresponds to the onset time of S-wave determined by the AR model of S-wave. It is therefore suggested that the procedure FUNIMAR, which has been

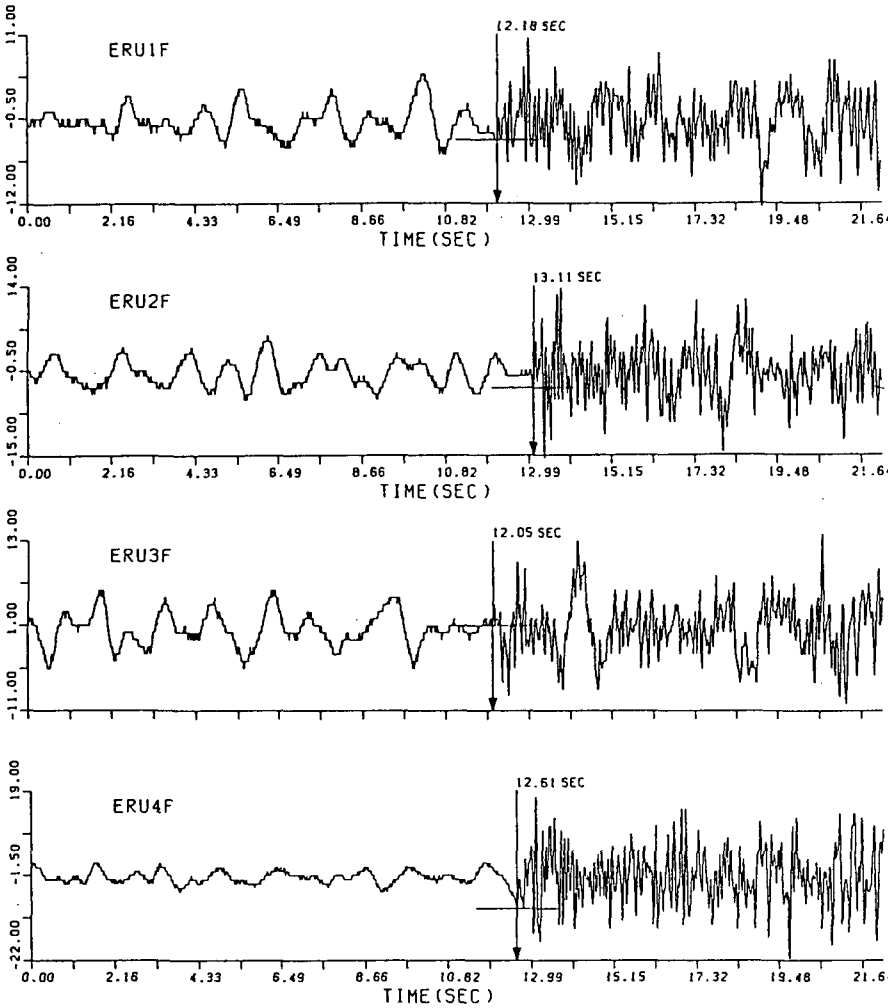


6.9(c)

focused mainly on the decision of the onset time of first P-arrival, can be also useful for the determination of the onset time of S-arrival.

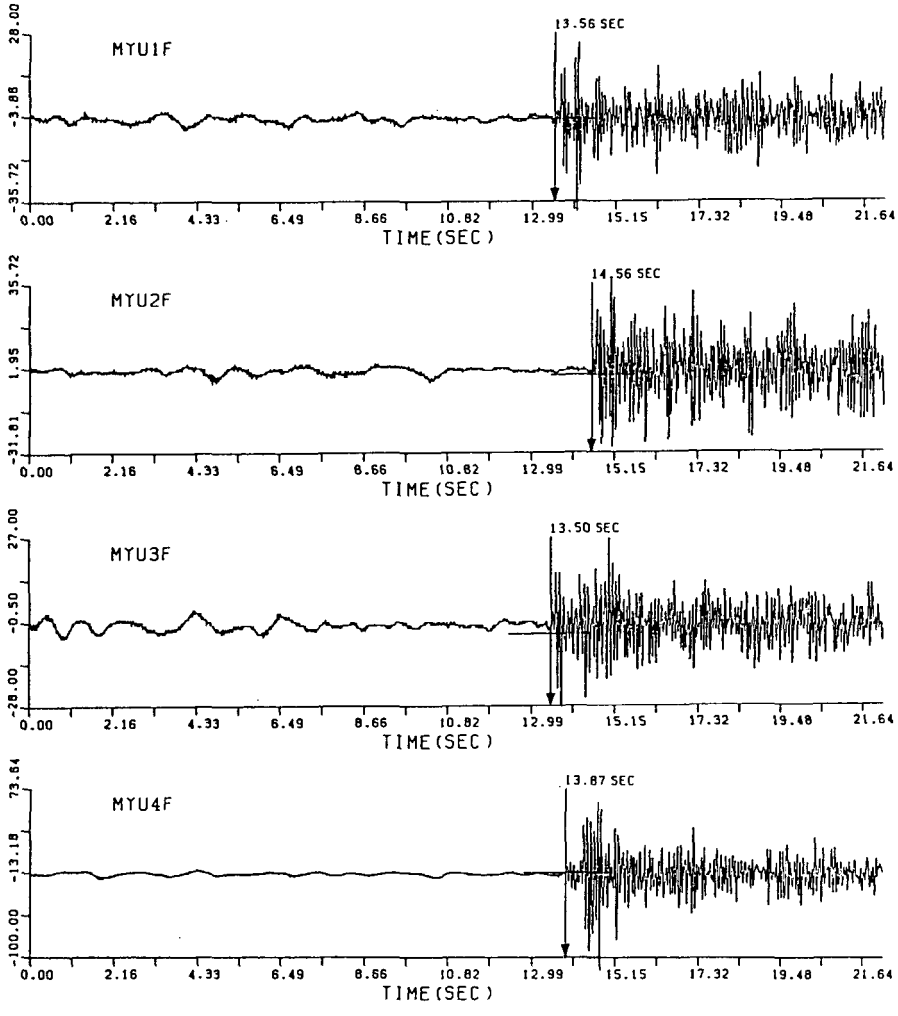
6.5 Discussion and conclusion

In the present chapter, we have shown a new, efficient, and numerically stable method for the fitting of locally stationary AR model which is suited to on-line precessing. In our method, we use a Householder transformation for the least squares computation. The proposed method has been applied to the



6.9(d)

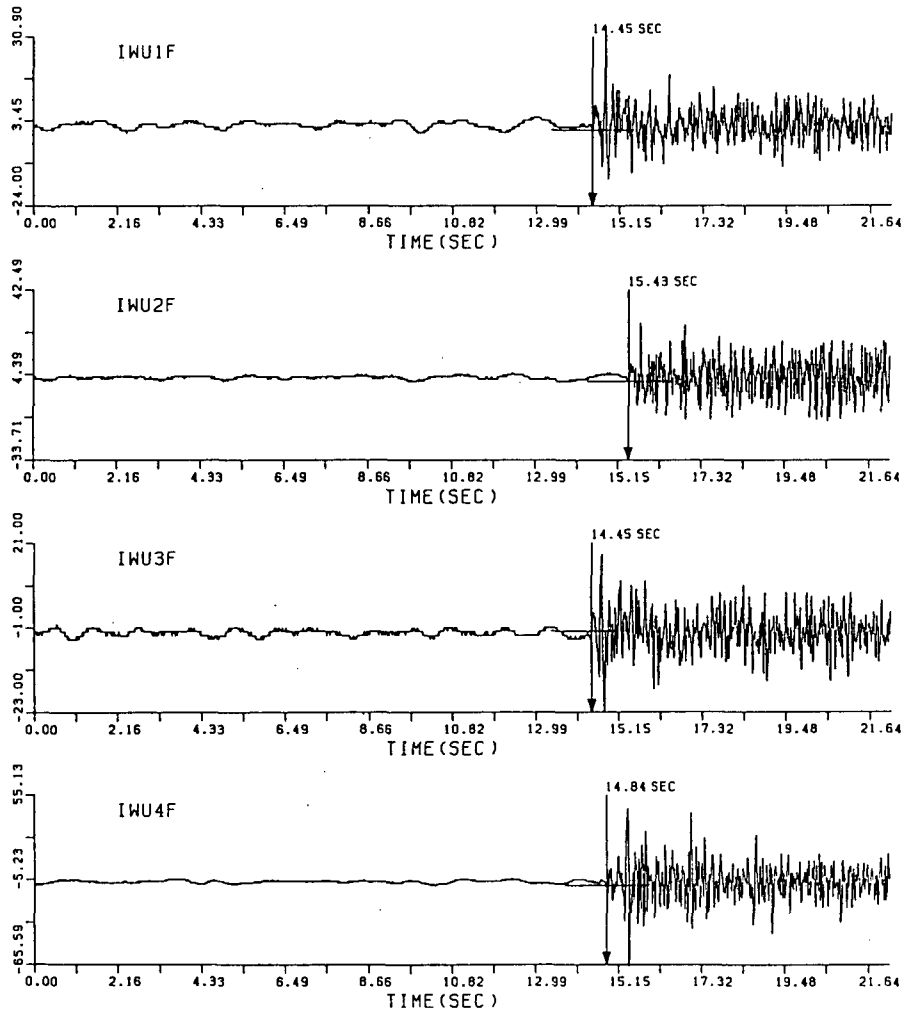
typical up-down component records of microearthquakes. It is shown that the proposed procedure can reasonably identify the arrival time of seismic wave even when it is contaminated by various kind of noises, such as traffic noise, hum, and when the signal to noise ratio is very small (Fig. 6.9). If the amplitude of seismic signal is greater than the level of LSB, the present procedure can reasonably determine the onset time of the P-arrival. Even for the case when we have difficulty in determining the onset time of seismic arrival, it might be able to alleviate it by the use of multi-channel version of present procedure.



6.9(e)

Further, by the proper modification of orthogonal transformation, CPU-time is significantly reduced. Namely, $O(Nk^2W)$ operations are reduced to only $O(Nk^2)$, where N, W, and k are the length of data set, the number of models checked, and the upper limit of the order of AR model, respectively.

By the implementation of the modified procedure for the AR model fitting, it is shown that the computing times are reduced to about one tenth of the time required by the original one and the AR model with order 5 will be enough to determine the onset time of seismic phase in the background noise without



6.9(f)

decrease of accuracy. In addition, the onset times of S-waves, which are mixed with the scattered coda waves induced by P-waves, can be also determined successfully by the present method. Morita and Hamaguchi (1984) discussed a mathematical aspect for the confidence interval of the onset time estimate. They concluded that the 90% confidence interval time was estimated to be 0.2 and 0.8 seconds for P and S onsets when the signal to noise ratios were about 10 dB. Judging from the present study, the confidence interval of the onset time does depend not only on such ratios but also the difference between two AR

models i.e., the spectral content of seismic signal and that of background noise.

7. Estimation of arrival times of seismic waves by 3-dimensional time series model

7.1 Introduction

In the estimation of arrival time of P-wave, the use of univariate time series is considered to be reasonable, since the P-wave is compression wave and a dominant part of the motion appears in the vertical component. However, since S-wave is a shear type wave, for the estimation of arrival time of S-wave the analysis on the horizontal plane, namely on the east-west and north-south components, seems to be necessary.

In view of the fact that even after the arrival of S-wave the coda of P-wave remains and that S-wave also includes the vertical motion, the use of all of three components is desirable. We also have an anticipation that even for the detection of P-wave, the analysis of three-variate time series will give more precise information about the onset time. We are thus motivated to use a three-variate locally stationary autoregressive (3D-LSAR) model which consists of the following two local models.

In section 7.2, a procedure for the estimation of arrival time of a seismic wave is developed based on a three-variate locally stationary autoregressive model (3D-LSAR) fitting. In section 7.3, a computationally efficient procedure for 3D locally stationary AR model fitting is developed. In section 7.4, the posterior probability of the arrival time of seismic wave is derived by using the likelihood of the LSAR models. Section 7.5 is devoted to several numerical examples where the proposed procedure is compared with the procedure based on the 1D LSAR models. The detection of the S-wave is also considered in this section.

7.2 Estimation of the arrival time and 3-dimensional locally stationary AR modeling

Let $y_n=(y_{nU}, y_{nN}, y_{nE})^t$, $(n=1, \cdots, N)$ be a three variate time series, where y_{nU}, y_{nN}, and y_{nE} express the up-down, north-south, and east-west components of the seismogram, respectively. Obviously, the characteristics of the series, e. g., the variances and spectra, change over time due to the arrival of a certain seismic wave such as P-wave or S-wave. However, it may be reasonable to assume that each of the seismogram before and after the arrival of the seismic wave is stationary and can be expressed by a single time series model. In this modeling the arrival time of the seismic wave, n_A, corresponds to the change

point of the time series model.

Background noise model,

$$y_n=\sum_{i=1}^{m_1}\boldsymbol{A}_{i1}\boldsymbol{y}_{n-i}+\boldsymbol{w}_{n1} \qquad (n=1, \cdots, n_{A-1}) \tag{7.1}$$

Here m_1 is the autoregressive order, $\boldsymbol{A}_{i1}$ is the 3×3 autoregressive coefficient matrix for i-lag component, and $\boldsymbol{w}_{n1}$ is the innovation sequence with mean 0 and covariance matrix $\boldsymbol{\Sigma}_1$. This model expresses the dynamics of the background motion. It should be noted that in the detection of S-wave, the coda of P-wave as well as the background motion are expressed by the "background noise" model.

Signal model,

$$y_n=\sum_{i=1}^{m_2}\boldsymbol{A}_{i2}\boldsymbol{y}_{n-1}+\boldsymbol{w}_{n2} \qquad (n=n_A, \cdots, N) \tag{7.2}$$

Here m_2, $\boldsymbol{A}_{i2}$, and $\boldsymbol{w}_{n2}$ are autoregressive order, autoregressive coefficient matrix and the innovation of the signal model, respectively. The variance matrix of the innovation $\boldsymbol{w}_{n2}$ is denoted by $\boldsymbol{\Sigma}_2$. This model dexpresses the dynamics of the seismic wave.

Assuming the arrival time n_A and orders of autoregressions m_1 and m_2, the distribution of the time series is given by

$$\boldsymbol{y}_n\sim N(\sum_{i=1}^{m_1}\boldsymbol{A}_{i1}\boldsymbol{y}_{n-1},\ \boldsymbol{\Sigma}_1) \qquad (n=1, \cdots, n_{A-1}),$$

and

$$\boldsymbol{y}_n\sim N(\sum_{i=1}^{m_2}\boldsymbol{A}_{i2}\boldsymbol{y}_{n-1},\ \boldsymbol{\Sigma}_2) \qquad (n=n_A, \cdots, N). \tag{7.3}$$

Therefore, given the observations $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_N$, the log-likelihood of the LSAR model is given approximately as follows

$$\begin{aligned}L(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)=&-(1/2)\{3N-m_1)\log 2\pi+(n_A-m_1)\log|\boldsymbol{\Sigma}_1|\\&+(N-n_A)\log|\boldsymbol{\Sigma}_2|+\sum_{n=m_1+1}^{n_A-1}\boldsymbol{w}_{n1}^t\boldsymbol{\Sigma}_1^{-1}\boldsymbol{w}_{n_1}+\sum_{n_A}^{N}\boldsymbol{w}_{n_2}^t\boldsymbol{\Sigma}_2^{-1}\boldsymbol{w}_{n_2}\},\end{aligned} \tag{7.4}$$

where $\boldsymbol{A}_1=(A_{1_1}, \cdots, A_{m_{1,1}})$, $\boldsymbol{A}_2=(A_{12}, \cdots, A_{m_2,2})$ and from Eqs. 7.1 and 7.2, $\boldsymbol{w}_{nj}$ are

$$\boldsymbol{w}_{nj}=\boldsymbol{y}_n-\sum_{i=1}^{m_j}\boldsymbol{A}_{ij}\boldsymbol{y}_{n-i} \qquad (j=1, 2). \tag{7.5}$$

The maximum likelihood estimates of $\boldsymbol{A}_{ij}$ and $\boldsymbol{\Sigma}_j(i=1, \cdots, m_j: 1, 2)$ are approximately given by maximizing (7.3).

However, from the form of the log-likelihood function, it can be easily seen that the parameters of the background noise model and the signal model can be independently obtained by minimizing

$$(n_{A-1}-m_1)\log|\boldsymbol{\Sigma}_1|+\sum_{n=m_1+1}^{n_A-1}\boldsymbol{w}_{n1}^t\boldsymbol{\Sigma}_1^{-1}\boldsymbol{w}_{n_1},$$

and

$$(N-n_A)\log|\Sigma_2|+\sum_{n=n_A}^{N}\boldsymbol{w}_{n2}^t\boldsymbol{\Sigma}_2^{-1}\boldsymbol{w}_{n_2}. \tag{7.6}$$

The computationally efficient procedure for the fitting of these models will be shown in the next section. The fitted model can be evaluated by the AIC criterion defined by

$$\text{AIC}=-2L(\bar{\boldsymbol{A}}_1, \bar{\boldsymbol{A}}_2, \bar{\boldsymbol{\Sigma}}_1, \bar{\boldsymbol{\Sigma}}_2)+2 \text{ (the number of estimated parameters)},$$

where $\bar{\boldsymbol{A}}_1$, $\bar{\boldsymbol{A}}_2$, $\bar{\boldsymbol{\Sigma}}_1$, and $\bar{\boldsymbol{\Sigma}}_2$ are the maxjmum likelihood estimates of $\boldsymbol{A}_j$ and $\boldsymbol{\Sigma}_j(j=1, 2)$, respectively. In the estimation of the arrival time, the crucial problem is the estimation of the dividing point n_A. This point can be determined by finding the minimum of the AIC.

7.3 *Computationally efficient procedure for 3-dimensional locally stationary AR model fitting*

7.3.1 *Householder method for multivariate AR model fitting*

We will first briefly review the procedure for the fitting of multivariate AR model developed for the program **MULMAR** (Multivariate case of minimum AIC method of AR model fitting) in **TIMSAC-78** (A time series analysis and control program package) by Akaike et al. (1979). This program can be used easily in the Oogata Computer Center of Hokkaido University (Takanami, 1984).

Assume that three-variate time series $\{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_N\}$ is given and we are fit to multivariate AR model

$$\boldsymbol{y}_n=\sum_{i=1}^{m}\boldsymbol{A}_i\boldsymbol{y}_{n-i}+\boldsymbol{w}_n, \quad \boldsymbol{w}_n\sim N(0, \boldsymbol{\Sigma}).$$

The main idea of **MULMAR** is to use an autoregressive model with instantaneous response

$$\boldsymbol{y}_n=\boldsymbol{B}_o\boldsymbol{y}_n+\sum_{i=1}^{m}\boldsymbol{B}_i\boldsymbol{y}_{n-i}+\boldsymbol{v}_n, \quad \boldsymbol{v}_n\sim N(0, \boldsymbol{V}). \tag{7.7}$$

It is, here, assumed that the coefficient of the instantaneous response is of the form

$$\boldsymbol{B}_0=\begin{bmatrix} 0 & 0 & 0 \\ b_{21} & 0 & 0 \\ b_{31} & b_{32} & 0 \end{bmatrix}$$

and the covariance matrix $\boldsymbol{V}$ is of diagonal form,

$$\boldsymbol{V}=\begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}.$$

Since

$$\boldsymbol{y}_n=(\boldsymbol{I}-\boldsymbol{B}_0)^{-1}\sum_{i=1}^{m}\boldsymbol{B}_i\boldsymbol{y}_{n-i}+(\boldsymbol{I}-\boldsymbol{B}_0)^{-1}\boldsymbol{y}_n,$$

this AR model with the instantaneous response is equivalent to the ordinary AR model with the relation

$$\begin{aligned}\boldsymbol{A}_i&=(\boldsymbol{I}-\boldsymbol{B}_0)^{-1}\boldsymbol{B}_i,\\ \boldsymbol{\Sigma}&=(\boldsymbol{I}-\boldsymbol{B}_0)^{-1}\boldsymbol{V}(\boldsymbol{I}-\boldsymbol{B}_o)^{-t}.\end{aligned}\tag{7.8}$$

It should be noted that these two models have the same number of parameters.

The significant merit of the use of the AR model with the instantaneous response is that it can be obtained by independently fitting the univariate models for each of the three components. This can be justified as follows. Since the covariance matrix is of diagonal form,

$$N\log|\boldsymbol{V}|+\sum_{n=1}^{N}\boldsymbol{v}_n^t\boldsymbol{V}^{-1}\boldsymbol{v}_n=\sum_{i=1}^{3}\{N\log\sigma_i^2+\sum_{n=1}^{N}\boldsymbol{v}_{ni}^2/\sigma_i^2\}.\tag{7.9}$$

Therefore, if the matrix $\boldsymbol{B}_i$ is divided as

$$\boldsymbol{B}_i=\begin{bmatrix} b_{i1} \\ b_{i2} \\ b_{i3} \end{bmatrix},$$

each of the parameter set $\{b_{i1},(i=1,\cdots,m_1)\}$, σ_1^2, $\{b_{i2},(i=1,\cdots,m_2)$, $\sigma_2^2\}$ and $\{b_{i3},(i=1,\cdots,m_3)$, $\sigma_3^2\}$ are independently estimated by minimizing

$$N\log\sigma_j^2+\sum_{n=1}^{N}\boldsymbol{v}_{nj}^2/\sigma_j^2.\tag{7.10}$$

For any given b_{ij}, (7.10) is minimized when

$$\sigma_j^2=1/N\sum_{n=1}^{N}\boldsymbol{v}_{nj}^2,\tag{7.11}$$

and by substituting this estimate into Eq. 7.10, it can be seen that b_{ij} are obtained by minimizing

$$N \log \sigma_j^2 + N,$$

or equivalently by minimizing σ_j^2. This means that by using the special expression for the multivariate AR model, the maximum likelihood estimates of the multivariate AR model are obtained by solving the least squares problem for each of the three components. Further more, the log-likelihood and AIC of the three-variate AR model are obtained as the sum of those of three components models.

We will next show an algorithm which can solve these three least squares problems efficiently. The least squares estimation of the multivariate AR model can be realized by first making $(N-m)\times(3m+3)$ matrix

$$\boldsymbol{X}=\begin{bmatrix} y^t_m & \cdots & y^t_1 & y^t_{m+1} \\ y^t_{m+1} & \cdots & y^t_2 & y^t_{m+2} \\ \cdot & \cdot\ \cdot & & \cdot \\ \cdot & \cdot\ \cdot & & \cdot \\ \cdot & \cdot\ \cdot & & \cdot \\ y^t_{N-1} & \cdots & y^t_{N-m} & y^t_N \end{bmatrix}, \tag{7.12}$$

and reduce this matrix to an upper triangular form by an orthogonal transformation (i.e. Householder transformation)

$$\boldsymbol{S}_1=\begin{bmatrix} s_{11} & \cdots & s_{1,3m+3} \\ & \cdot & \cdot \\ & \cdot & \cdot \\ & & \cdot \\ & & s_{3m+3,3m+3} \\ \multicolumn{3}{c}{\mathbf{0}} & \end{bmatrix}. \tag{7.13}$$

The $(3m+1)\times(3m+1)$ upper left triangular matrix contains sufficient information for the fitting of the model for the first component (UD-component). In particular the innovation variance $\sigma_1^2(j)$ and AIC(j) of the j-th order model

$$\boldsymbol{y}_{nU}=\sum_{i=1}^{j} \boldsymbol{b}_{iU}\boldsymbol{y}_n+\boldsymbol{w}_{nU}, \tag{7.14}$$

where $\boldsymbol{b}_{iU}=(b_i(1,1),\ b_i(1,2),\ b_i(1,3))$, $\boldsymbol{y}_n=(y_{nU}, y_{nN}, y_{nE})^t$, are obtained by

$$\sigma_1^2(j)=1/(N-m)\sum_{i=3j+1}^{3m+1} s^2_{i,3m+1} \qquad (j=0, 1, \cdots, m),$$

and

$$\mathrm{AIC}(j)=(N-m)\log \sigma_1^2(j)+2(3j+1). \tag{7.15}$$

Incidentally, the regression coefficients of the UD-component model with

order j are obtained by solving the linear equation

$$\begin{bmatrix} s_{11} & \cdots & s_{1,3j} \\ & \ddots & \\ 0 & & s_{3j,3j} \end{bmatrix} \begin{bmatrix} b_1(1,1) \\ \vdots \\ b_j(1,3) \end{bmatrix} = \begin{bmatrix} s_{1,3m+1} \\ \vdots \\ s_{3j,3m+1} \end{bmatrix}. \tag{7.16}$$

However, it should be noted that for the present purpose of the estimation of the arrival time, only the AIC values of the best model are necessary and we do not actually solve this linear equation.

For the computation of the AIC of the second (the North-South) component model,

$$y_{nN} = b_0(2,1)y_{nU} + \sum_{i=1}^{j} b_{i2}y_n + w_{nN}, \tag{7.17}$$

we first transform the matrix (7.12) to the following form

$$\boldsymbol{S}_2 = \begin{bmatrix} s_{11} & \cdots & s_{1,3m} & s_{1,3m+1} & s_{1,3m+1} & s_{1,3m+3} \\ s_{21} & \cdots & s_{2,3m} & & s_{2,3m+2} & s_{2,3m+3} \\ \vdots & & \vdots & & \vdots & \vdots \\ & & s_{3m+1,3m} & & s_{3m+1,3m+2} & s_{3m+1,3m+3} \\ & & & & s_{3m+2,3m+2} & s_{3m+2,3m+3} \\ & & 0 & & & s_{3m+3,3m+3} \end{bmatrix}. \tag{7.18}$$

This can be done by using as appropriate Householder transformation with only a little addtional computations. Then the upper left $(3m+2)\times(3m+2)$ sub-diagonal matrix contains sufficient information for the fitting of the regression model for the second component which has an instant response from the first component. The residual variance and AIC of the j-th order model is given by

$$\sigma_2^2(j) = 1/(N-m)\sum_{i=3j+2}^{3m+2} s_{i,3m+2}^2 \qquad (j=0,1,\cdots,m),$$

and

$$\mathrm{AIC}_2(j) = (N-m)\log\sigma_2^2(j) + 2(3j+2). \tag{7.19}$$

It can be seen that the $(3m+1)$-th column of the matrix $\boldsymbol{S}_1$ which is used as the vector of objective variable in fitting the model for the first component, is now used as the vector of a regressor corresponding to the instantaneous response from the first variable. Similarly the model for the third variable (East-West

component) can be obtained from

$$\boldsymbol{S}_3=\begin{bmatrix} s_{11}\cdots & s_{1,3m} & s_{1,3m+1} & s_{1,3m+2} & s_{1,3m+3} \\ s_{21}\cdots & s_{2,3m} & & s_{2,3m+2} & s_{2,3m+3} \\ s_{31}\cdots & s_{3,3m} & & & s_{3,3m+3} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ & s_{3m+2,3m} & & & s_{3m+2,3m+3} \\ & & & & s_{3m+3,3m+3} \\ & \mathbf{0} & & & \end{bmatrix}, \tag{7.20}$$

$$\sigma_3^2(j)=1/(N-m)\sum_{i=3j+3}^{3m+3} s_{i,3m+3}^2 \qquad (j=0, 1, \cdots, m),$$

and

$$\mathrm{AIC}_3(j)=(N-m)\log\sigma_3^2(j)+2(3j+3). \tag{7.21}$$

The AIC of the original 3D autoregressive model is given by

$$\mathrm{AIC}(j)=\min_j \mathrm{AIC}_1(j)+\min_j \mathrm{AIC}_2(j)+\min_j \mathrm{AIC}_3(j). \tag{7.21}$$

7.3.2 Augmentation of the data

In the previous section, we showed the algorithm of the fitting of multivariate AR model. We will now show a method of fitting an AR model for an augmented data set $\{y_1, \cdots, y_N, y_{N+1}, \cdots, y_{N+p}\}$. Here p is the number of the new data which might be 1. This can be performed by first organizing the following $(3m+3+p)\times(3m+3)$ matrix $\boldsymbol{R}_1$,

$$\boldsymbol{R}_1=\begin{bmatrix} & & \boldsymbol{S}_1 & \\ y_N{}^t & \cdots & y^t{}_{N-m+1} & y^t{}_{N+1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ y^t{}_{N+p-1} & \cdots & y^t{}_{N+p-m} & y^t{}_{N+p} \end{bmatrix} \tag{7.22}$$

and reduce to an upper triangular form. Then applying the same method as presented in the previous section to this matrix, we can get the AIC values of the best AR model fitted to this data set.

7.3.3. The number of necessary operations

For the Householder transformation of $N\times k$ matrix, the amount of multiplication (and additions) is approximately evaluated as $nk^2/2$ in the proceding chapter 6. On the other hand, the necessary operations for the transforma-

tion (7.22) amount to

$$\sum_{i=1}^{p+1}(p+1)(k+1-i)\sim pk^2/2.$$

Therefore the total amount of multiplications (and additions) for the comparison of all possible 3D-LSAR models by the present method is

$$\begin{aligned}&1/2n_0(3m+3)+\sum_{i=1}^{W}p(3m+3)^2/2+1/2(N-n_1)(3m+3)^2\\&\qquad+\sum_{i=1}^{W}p(3m+3)^2/2\\&\fallingdotseq 9/2Nm^2+9/2m^2(n_1-n_0)\leqq 9Nm^2.\end{aligned}\tag{7.24}$$

where W is the repeating number described in section 6. This means that the total number of computation for the fitting of 3D-LSAR model is less than twice that for fitting single three variate AR model.

On the other hand, the number of necessary operations for fitting a ordinary 3D-AR model to entire data set is approximately $(9/2)Nm^2$, and fitting 3D-LSAR model without recurive formula shown in section (7.3.7) requires

$$\sum_{i=0}^{W}1/2(n_0+ip)(3m+3)+\sum_{i=0}^{W}1/2(N-n_0-ip)k^2\sim(9/2p)Nm^2(n_1-n_0).\tag{7.25}$$

Therefore by using the present procedure, the number of necessary computations is reduced by

$$9Nm^2/\{(9/2p)Nm^2(n_1-n_0)\}=2p/(n_1-n_0)=2/W\ll 1.\tag{7.26}$$

Incidentally, fitting a 1D-LSAR model by the same recursion requires $(1/2)\cdot Nm^2+(1/2)\cdot m^3(n_1-n_0)$. Summarizing, the necessary computing time by the present method is only twice that for ordinary 3D-AR model, and is 9 times as much as that for 1D-LSAR model and is $2p/(n_1-n_0)$ times of the conventional method.

7.4 Test of autoregressive model fitting

Fitting the 3-dimesional autoregressive models to nonstationary time series will be applied to the microearthquake data recorded at stations of RCEP. To check the validity of the 3-dimensional autoregressive model, we first applied FUNIMAR to every component of seismograms, and then apply the 3-dimensional autoregressive model to the same data set. Figure 7.1 shows the results of FUNIMAR, which is applied to the three components seismograms observed at MYR (see Foreshock 2 in Table 4.1 and seismograms in Fig. 4.7 to 4.9). The original record is sampled at each 0.01084 second with minimum resolution of

1.0. The minimum AIC estimates of the arrival time obtained by the procedure FUNIMAR are $n=1351$, 1361, and 1342 for the east-west, the north-south, and the vertical components as shown in Fig. 7.1, respectively. The point 1342 is surely the onset time of P-wave judging from the original seismograms. The difference in the minimum AIC point means the discrepancy between the start times of ground motions for their own components. Therefore the motions of P waves in the east-west and north-south component seismograms start 0.098 and 0.206 seconds later than that of the vertical, respectively. This result might be due to their own S/N of the seismograms.

By way of experiment, the three kinds of AIC values obtained above are summed up at every point, and the result is shown in Fig. 7.2 with a label '**AIC(EW)+AIC(NS)+AIC(UD)**'. This trace of AIC versus time shows the minimum at $n=1342$ which is exactly the same minimum point found in the trace of AIC versus time for the vertical component as shown in Fig. 7.1. This means that the vertical component of the seismogram has more precise informa-

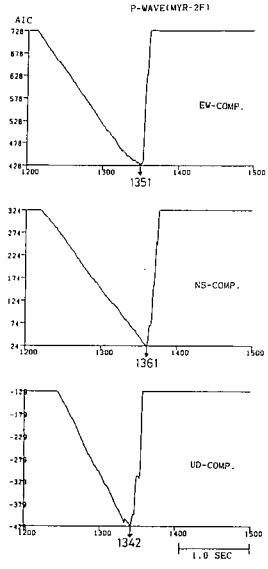


Fig. 7.1. The AIC curves of three component P waves recorded at station MYR. FUNIMAR is applied to each component seismogram. The earthquake is foreshock No. 2 in Table 4.1. Time scale shown at the bottom. Display of 300 data points (about 3.3 sec.).

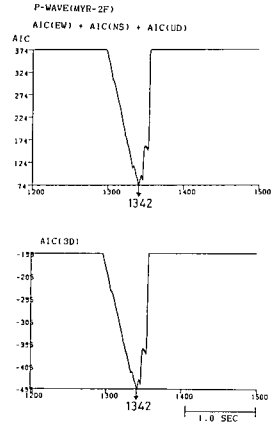


Fig. 7.2. The three-dimensional AIC curves of the P wave used in Fig. 7.1. Top: AIC curve by the summation of the three AIC's shown in Fig. 7.1. Bottom: AIC curve by three dimensional (3D) procedure. The parameters of the event is the same as those in Fig. 7.1.

tion about the onset time of P-wave than the other components. Since the slope of the AIC around the minimum is steeper, we can get more reliable estimate of the onset time by using the sum of three AIC sequences.

Next, we will take up another procedure, the 3-variate autoregressive model fitting. Its mathematical aspect is described in preceding section. The data set used is completely the same as the one tested above. At the bottom of Fig. 7.2 the AIC curve is shown with labeling '**AIC(3D)**'. It resembles the result obtained by the sum of AIC's of three univariate models; and the minimum point of the AIC completely coincides with that in the case of summing separate AIC procedures. This is a very important evidence that the onset time of P-wave can be determined by either model fitting of the two different procedures. The procedure by summing up AIC's of separate modelings at every point can save more CPU-time than that by the 3-dimensional autoregressive model owing to the smaller amount of necessary calculation. The CPU-time for the summing procedure is three times as long as that for the single trace, whereas the 3-dimensional procedure does spend CPU-time about 9 times as long as the single trace. Therefore the summing procedure '**AIC(EW)+AIC(NS)+AIC(UD)**' is suitable for preliminary on-line system.

Next, we will examine the advantage of the use of the 3-variate autoregressive model for determing onset time of S-wave. Those two procedures examined above are applied also to the three components seismograms in which S-waves are surely found. Figure 7.3 shows the results of the FUNIMAR procedure, applied to each component seismogram. It can be seen that the minimum AIC estimates of the onset time from the east-west, the north-south and the vertjcal component seismograms are 2065, 2085, and 1913, respectively. The difference of the minimum AIC points demonstrates that it is difficult to determine the onset time of S-wave from only one component of the seismogram. Moreover the shapes of AIC traces suggest the difficulty of the determination of the onset time of S-wave. The many local minima are found on each AIC curves as shown in Fig. 7.3, which has been already inferred from the behaviors of the particle motions of S-waves as shown in the preceding chapters. In other words, it is usually hard to determine precisely the onset time from using a single trace, especially only from the vertical component of the seismogram as seen in Fig. 7.3.

Now we will consider other techniques which are described in the previous sections. We first applied the procedure '**AIC(EW)+AIC(NS)+AIC(UD)**' to the three components seismograms. The result is illustrated in the upper half of Fig. 7.4. The AIC curve with label '**AIC(EW)+AIC(NS)+AIC(UD)**' has

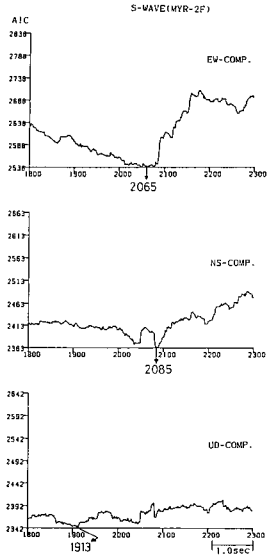


Fig. 7.3. The AIC curves of three component S-waves recorded at station MYR. Other conventions are the same as in Fig. 7.1. Time scale is shown at the bottom. Display for 500 data points (about 5.4 sec.) The parameters of the event is the same as in Fig. 7.1.

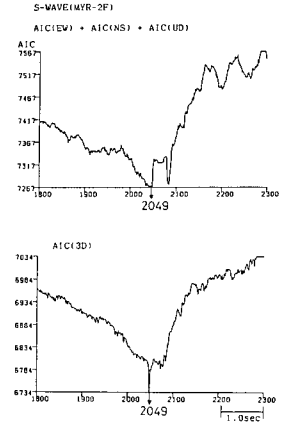


Fig. 7.4. The three-dimensional AIC curves of the S-wave used in Fig. 7.3. Top: AIC curve by the summation of the three AIC's shown in Fig. 7.3. Bottom: AIC curve by three dimensional (3D) procedure. The parameters of the event is the same as those in Fig. 7.1.

two significant local minima. The first minimum of the AIC is located at point 2049, and is smaller than the second local minimum. The second local minimum of the AIC probably corresponds to the point where some different seismic phase arrives.

For the time we consider that this minimum of the AIC indicates the arrival of the main S-wave. We also apply the 3-variate autoregressive model and compute the '**AIC(3D)**' for these three components seismograms. The resultant AIC curve labeled with '**AIC(3D)**' is shown in the bottom half of Fig. 7.4. The slope around the minimum AIC is steeper. Moreover this point accords with that of the above minimum AIC, which was temporary designated as the onset time of S-wave. This is a noticeable coincidence, that is, the onset time to S-wave can be determined if the three components seismograms are used. To make sure, we also examine other seismograms, and some of the results are shown in Fig. 7.5 to Fig. 7.7, respectively.

The results shown in Fig. 7.5 are obtained from the three components

seismograms observed at Iwanai (IWN) at 9:49, March 21, 1982. The hypocenter parameters of the earthquake are shown in Table 4.1. Figure 7.5 shows the minimum AIC point by the summing procedure places at location 2184 as a possible onset point (time) of S-wave, and that this location coincides with the one obtained by the other 3-dimensional procedure. This example exhibits that the application of 1-variate autoregressive model can yield the same results at the 3-variate AR model. About 0.2 seconds after the global minimum, a comparable local minimum can be seen in both the AIC curves. Such a phenomenon is also observed in the first test of S-wave onset time determination by using three components seismograms (Fig. 7.4). It is probably caused by the complicated crustal structure along the ray path. Another supplementary example shown in Fig. 7.6 is obtained for the seismograms at Misono (MSN). The earthquake occurred at 7:45, March 21, 1982. The source parameters of the event are shown in Table 4.1. The two AIC curves labeled with '**AIC(EW)+AIC(NS)+AIC(UD)**' and '**AIC(3D)**' have their minimum AIC at 1362 and 1363, respectively. Although the estimated onset time (point) of S-wave is almost identical, the features of the curves after the arrivals of S-wave are quite different from each other. It is a subject of the future study why such a discrepancy suddenly appears after the S-wave. There are some comments on the use of the sum of independently computed three AIC's. In the multivariate AR model, if we assume that the three components are independent, both of $\boldsymbol{A}_i$ and $\boldsymbol{\Sigma}$ become of diagonal form and the AIC of the 3-variate time series reduces

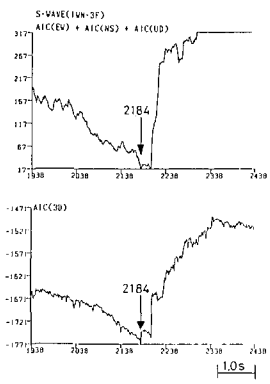


Fig. 7.5. The three-dimensional AIC curves of the S-wave at station IWN. The earthquake is fore-shock No. 3 in Table 4.1. Other conventions are the same as in Fig. 7.4.

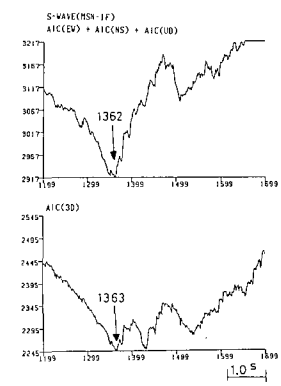


Fig. 7.6. The three-dimensional AIC curves of the S-wave at station MSN. The earthquake is fore-shock No. 1 in Table 4.1. Other conventions are the same as in Fig. 7.4.

the sum of three AIC's. Therefore, this AIC value expresses the goodness of the assumption of the independence of three components. The above mentioned actual behavior of AIC indicates that this assumption is reasonable before the arrival of S-wave but is not after that.

The other supplementary examination is carried out by using the three components seismograms, which are observed at station Iwanai (IWN) at 9 : 33, March, 21, 1982. The hypocentral paramentes of this event are also listed in Table 4.1. The resultant AIC curves are shown in Fig. 7.7. The two minima of their curves which are obtained through the AIC summing procedure and the 3-dimensional one are found at 2229 and 2222, respectively. In this case, the difference of the estimate of arrival time is seven in points, which is equivalent to only 0.076 seconds. Considering the theoretical motions of S-waves, 3-dimensional procedure provides an accurate arrival times.

In the present chapter, we have examined the techniques to determine the onset times of S-waves by using three components seismograms. The reason why we use the three components seismograms is that the S-waves of earthquakes are usually complicated by the coda waves excited by P-waves. We have tried to check the validity for determining them by using three different procedures, the FUNIMAR, the AIC summing, and the 3-dimensional processes. The process FUNIMAR is applied to each component seismogram. The three minima of the AIC obtained are quite different from each another. Thus the FUNIMAR does not seem suited for the application to S-wave. Another procedure is the summing process, **AIC(EW)+AIC(NS)+AIC(UD)** method. This procedure mostly can provide the same estimate as the other procedure, **AIC(3D)** method, and thus can replace the 3-dimensional one. Considering the necessary computation, we might suggest the summing procedure in actual on-

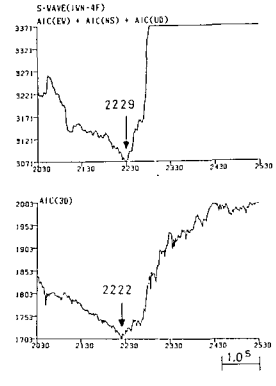


Fig. 7.7. The three-dimensional AIC curves of the S-wave at station IWN. The earthquake is fore-shock No. 4 in Table 4.1. Other conventions are the same as in Fig. 7.4.

line system.

8. Application and statistical evaluation

8.1 An application of AR model fitting for real automatic data processing

The procedures developed in the previous chapters have been applied to the seismograms recorded on the data recorder (high density data recorder, HD-DR) of RCEP. The HDDR can continuously record the signals telemetered from all seismic stations of RCEP for two days. Every second day, we change the tape and the recorded one is transferred to the retrieval and editing of seismic signals based on the lists of person arrival time readings made in advance.

We have incorporated the softwares of autoregressive model fittings into the automatic playback operation. In the course of retrieval, we have copied the seismograms displayed on the cathode ray tube (CRT) as often as possible we can. We have used these hard copies for re-reading the onset times of P- and S-phases and also for the investigations of the reliability of the procedures by the statistical analyses. The procedure FUNIMAR has been in operation for determing the onset times of P-waves, while the summing procedre has been used for determing the onset times of S-waves. The feasibility of the two procedures for the automatic processing has been already confirmed in the preceding chapters.

Table 8.1 shows the diagram indicating the correlation between the number of seismograms read by persons and the number of events captured by the preliminary processing for earthquake detection. This preliminary earthquake detection is made by judging whether the discrepancy between the initial model, which was made at the initial part of the time interval, and the current model becomes greater than a specified level that is set by a trial and error method. It can be seen a good coincidence between the number of stations which detect the small earthquakes and the number of the stations which are reported in advance. The greater the event becomes, the less the number of stations becomes. The maximum number of the stations is 24, which is restricted by the number of channels of paper chart recorder. Fig. 8.1 shows the comparisons of the rate of the number of detected evets to those of the events picked by persons in advance. The figure shows two examples observed for station numbers of 6 and 15, respectively. Table 8.1 and Fig. 8.1 exhibit that the events, which are designated as earthquakes by persons, are captured all by the preliminary processing.

Table 8.1. Comparison of the number of detected P-waves out of those of hand picked P-waves. NS=number of hand picked P-waves per one event.

		N_P																									
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
	1	1	55																								56
	2	0	1	34																							35
	3	0	2	6	20																						28
	4	0	3	5	5	43																					56
	5	0	1	2	6	3	22																				34
	6	0	2	2	3	6	6	29																			48
	7	0	1	0	1	3	5	4	20																		34
	8	0	0	0	2	1	2	7	0	37																	49
	9	0	0	0	0	3	1	4	8	6	24																46
	10	0	1	0	0	0	2	2	3	5	3	15															31
	11	0	0	1	0	0	1	1	4	3	7	0	14														31
NS	12	0	0	0	1	1	0	1	1	0	1	7	1	13													26
	13	0	0	1	0	0	0	0	0	0	0	2	5	1	6												15
	14	0	0	0	0	0	0	0	0	0	1	2	1	6	0	9											19
	15	0	0	0	0	0	0	0	0	0	0	0	0	1	6	3	2										12
	16	0	0	0	0	0	0	0	0	0	1	1	0	0	0	4	1	6									13
	17	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	11	0	1								16
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6	0	2							10
	19	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	8	0	2						12
	20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	4	2	0					8
	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	3	1				8
	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5	2	0			8
	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	4	1	0		7
	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	5	2	1	9

N_P=number of detected P-waves by FUNIMAR. Number of cases for N_P (P-waves)-events out of NS (P-waves)-events shown in the corresponding mesh.

8.1.1. An application of the FUNIMAR for the vertical components seismograms

We will first show the examples obtained by the efficient procedure, FUNIMAR in Fig. 8.3 (a)-(f). The source parameters and the epicenters of these earthquakes are shown in Table 8.2 and Fig. 8.2, respectively. Figure 8. 3(a) shows the examples of clear beginnings of seismic signals recorded by the vertical component. The AIC curves, which are shown together, show the simple features. Nobody can be misled in the judgment of the onset times of P-

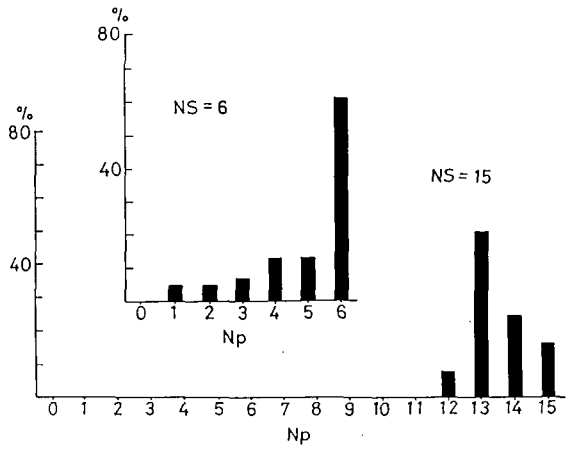


Fig. 8.1. Ratio of the number of P-wave arrivals picked by FUNIMAR to that specified by persons. Two examples for NS=6 and NS=15 are shown. Explanation of parameters: NS=number of P-times read by persons in advance and NP=number of P-times detected by FUNIMAR applied to the NS seismograms. Other cases shown in Table 8.1.

waves, which is also understood by the behaviors of the AIC curves. Namely, the AIC curves show monotonous decrease until they attain their own minima. In such a case, the points of minima could be obtained precisely with the accuracy of the sampling rate, so that we can anticipate that the accuracy of the onset time is 0.0108 seconds.

The next examples shown in Fig. 8.3(b) have the strong later-phases about 0.5-1.0 seconds after the fore running small P-waves. Such the distinct later phases with the weak first P-waves are sometimes observed at the stations of

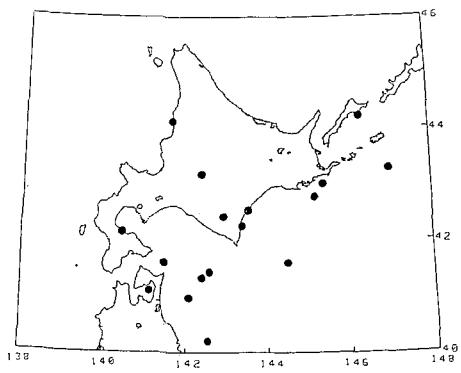


Fig. 8.2. Map showing the epicenters of the earthquakes used in FUNIMAR incorporated into the automatic playback operation.

RCEP, and have been investigated by some researchers, e.g. Motoya (1969), Moriya (1972), and Sugiyama (1989). Here we consider the determination of the onet times of the first P-waves in such a situation. It is a very important subject to investigate how large later phases influence the determination of the onset times. Figure 8.3(b) suggests that the AIC curves can designate the onset times of the first P-waves without any effects from the later phases. The arrival time can be detected easily by finding the minimum of the AIC curve.

The third examples, shown in Fig. 8.3(c), have the blurred rising of P-waves. Such the blurred risings are often found in the seismograms of deep or distant earthquakes. It is usually difficult to judge whether we could correctly identify the onset times of P-waves. At any rate, the judgment is supported by the change of spectral content of the seismogram by P-wave. The AIC curves clearly show their minima as shown in this figure, and so we may determined the onset times even for blurred rising P-waves. Even though a seismogram itself does not clearly exhibit the onset time of the P-wave, the AIC curve can substitute for seismogram. The above examples are the seismograms with relatively high S/N ratios.

In Fig. 8.3(d) the seismic signals are readily apparent compared with the background noises. In this case, the prominent frequencies of seismic signals are equal to those of their own background noises. If P-wave becomes in phase with the background noise, the onset time of P-wave is likely to be misled. Figure 8. (d) shows that the AIC curves are useful for the determinations of arrival times. It implies further that the AIC curves can reasonably designate the onset times of P-waves. The last two example groups, which are shown in Fig. 8.3(e) and (f), are composed of the seismic waves with ambiguously rising P-waves. Some of them show the flat AIC curves around its minimum. This suggests a limitation of the accuracy in determining the onset times of P-waves by using the FUNIMAR. The FUNIMAR, however, is a flexible procedure, so we may develop easily the parameters used in the FUNIMAR by experience. In the following section, we will investigate statistically the usefulness of the FUNIMAR, which has been working in the automatic playback system.

8.1.2 Statistical evaluation of the onset times of P-waves

Whenever the onset times of P-waves are determined through the playback processing, they are saved up in the disk file. So, we can easily retrieve data in order to evaluate the validity of the FUNIMAR for the automatic system.

We check the validity of the procedure by the bias and variance of the differences of the onset times, that is, we evaluate the mean and variance of the differences between the onset times obtained by the automatic processing and

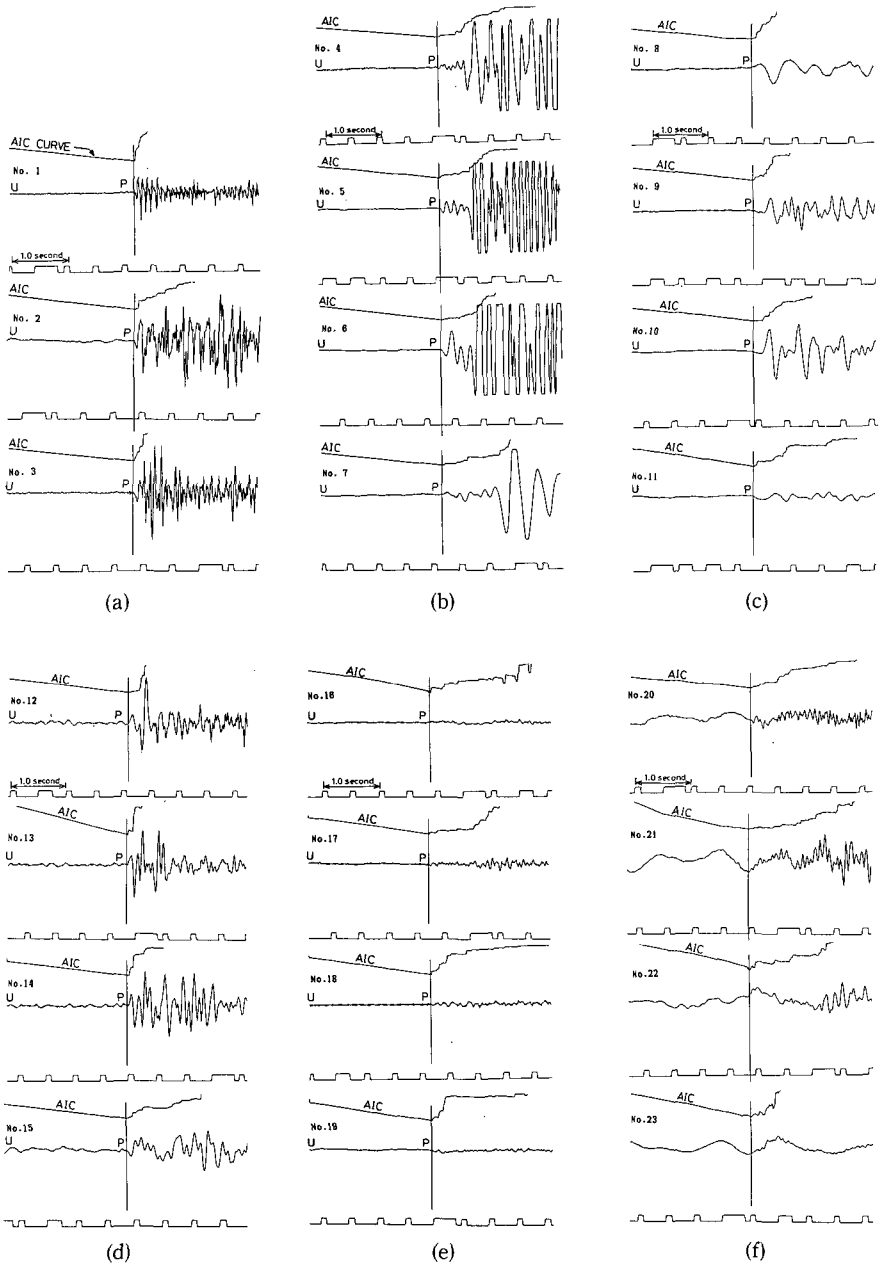

AIC CURVE
No. 1
U
P
1.0 second
AIC
No. 2
No. 3
No. 4
No. 5
No. 6
No. 7
No. 8
No. 9
No.10
No.11
No.12
No.13
No.14
No.15
No.16
No.17
No.18
No.19
No.20
No.21
No.22
No.23
(a)
(b)
(c)
(d)
(e)
(f)

Table 8.2. List of event parameters used for determining the onset time of P-waves.

ORIGIN TIME	ST. CODE	LOCATION		DEP.	MAG.
		LAT.	LONG.		
No. 1 1988 Sep. 23 (267)02h24m39s	IWN	N42.57°	E143.62°	70.8km	M1.8
No. 2 1988 Nov. 28 (333)06 33 31	MYR	N42.51°	E142.83°	51.8km	M2.8
No. 3 1988 Sep. 27 (271)05 46 00	IWN	N43.22°	E142.51°	173.4km	M3.2
No. 4 1988 Sep. 27 (271)19 12 07	KMU	N41.64°	E144.55°	45.5km	M3.6
No. 5 1988 Nov. 29 (334)18 15 02	NMR	N43.26°	E147.16°	33.0km	M3.5
No. 6 1988 Nov. 29 (334)19 56 13	KNP	N42.80°	E145.26°	44.1km	M3.5
No. 7 1988 Sep. 24 (268)23 05 36	KNP	N44.11°	E141.64°	15.8km	M3.5
No. 8 1988 Sep. 26 (270)05 06 32	TKC	N35.82°	E139.85°	419.7km	M4.4
No. 9 1988 Oct. 01 (275)05 57	TKC				
No. 10 1988 Dec. 01 (336)00 10 16	KNP	N40.25°	E142.85°	54.3km	M3.9
No. 11 1988 Sep. 24 (268)23 05 36	IMG	N44.11°	E141.64°	15.8km	M3.5
No. 12 1988 Sep. 25 (269)09 12 06	IWN	N41.56°	E142.55°	37.2km	M3.5
No. 13 1988 Sep. 27 (271)15 44	TKC				M0.5
No. 14 1988 Sep. 27 (271)13 44	TKC				M0.8
No. 15 1988 Jul. 20 (202)11 25 03	WCR	N42.26°	E140.20°	13.1km	M3.6
No. 16 1988 Jul. 24 (206)17 14 48	KMU	N41.09°	E141.12°	96.2km	M2.6
No. 17 1988 Sep. 25 (269)22 54 16	KMU	N42.35°	E143.39°	36.8km	M1.0
No. 18 1988 Sep. 25 (269)09 12 06	MSN	N41.56°	E142.55°	37.2km	M3.5
No. 19 1988 Oct. 01 (275)02 39 43	URH	N43.02°	E145.41°	57.5km	M2.2
No. 20 1988 Oct. 10 (284)09 46 06	HSK	N41.42°	E142.36°	27.2km	M3.2
No. 21 1988 Nov. 30 (335)14 45 09	MMA	N41.66°	E141.62°	14.8km	M2.8
No. 22 1988 Nov. 29 (334)19 22 13	TMR	N44.19°	E146.27°	187.5km	M3.6
No. 23 1988 Dec. 01 (275)02 12 47	MMA	N41.03°	E142.17°	64.3km	M1.9

Fig. 8.3. Examples of the results through the automatic playback operation. The vertical solid line marks the picked P-wave arrival time by FUNIMAR. (a): examples of clear P-waves. (b): examples of weak P-waves with strong later phases. (c): examples of emergent P-waves. (d): examples of P-waves contaminated with relatively strong back-ground noises. (e): examples of very weak P-waves. (f): examples of very weak P-waves comtaminated with relatively strong back-ground noises. Explanation about events and stations shown in Table 8.2.

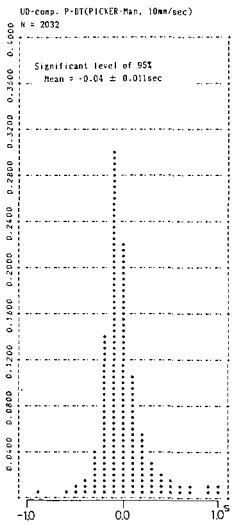


Fig. 8.4. Histogram of difference between the machine picked arrival time and the hand picked arrival time (machine-person). The relative frequencies are obtained through the optimized classification (after Sakamoto et al., 1983).

the P-times reported by persons who read the seismograms on paper charts. The paper chart recorder whose speed in 1.0 cm/sec has 24 channels. The recorded chart is put in the operation for the phase readings by a person using a digitizer one day later. Figure 8.4 represents the histogram obtained from these data. In making the histogram, we omit the outliers greater than 1.0 seconds or less than -1.0 seconds in time difference in view of the distribution. The histogram shows a symmetric distribution with a mean of -0.04 ± 0.01 seconds of 95%. From these observations, it now becomes certain that the FUNIMAR is useful as an automatic processor for determining onset times of P-waves.

8.1.3 An application of the summing procedure for the three components seismograms

We set the summing procedure in our computer system. After the theoretical S-arrival times are assumed using a list of hypocentral locations, the procedure is applied in this neighborhood. Several hundreds earthquakes are used to determine S-onset times by the summing procedure.

We will first show the examples obtained by the summing procedure in Fig. 8.6(a)-(e). The source parameters and the epicenters of these earthquakes are shown in Table 8.3 and Fig. 8.5, respectively. Figure 8.6(a) shows the examples of clear beginnings of seismic signals recorded by the horizontal components. The vertical long line labeled "S" means the minimum point in the curve of AIC, which is obtained by the summing procedure, '**AIC(EW)+AIC(NS)**

+**AIC(UD)**'. The short line with a dot shows the minimum point in the AIC curve for each component seismogram which is obtained by the FUNIMAR. The short line drawn on the vertical component seismogram suggests that the vertical component seismogram makes only a small contribution in determining the onset time of S-wave. On the other hand, both horizontal component

Table 8.3. List of event parameters used for determining the onset times of S-waves.

ORIGIN TIME	ST. CODE	LOCATION		DEP.	MAG.
		LAT.	LONG.		
No. 1 1988 Nov. 18 (323)04h58m59s	NMR	N43.12°	E147.42°	43.4km	M2.7
No. 2 1988 Nov. 18 (323)08 08 20	ESH	N42.53°	E141.86°	141.9km	M2.9
No. 3 1988 Nov. 19 (324)00 20 21	NMR	N42.78°	E147.44°	33.0km	M3.5
No. 4 1988 Nov. 19 (324)06 15 24	NMR	N42.86°	E147.26°	20.4km	M3.4
No. 5 1988 Nov. 22 (327)06 21 43	HSS	N42.25°	E142.57°	24.7km	M3.5
No. 6 1988 Nov. 22 (327)06 21 43	IMG	N42.25°	E142.57°	24.7km	M3.5
No. 7 1988 Nov. 22 (327)09 32 06	NMR	N42.67°	E142.55°	143.7km	M3.5
No. 8 1988 Nov. 24 (329)01 11 45	IWN	N42.00°	E144.33°	52.2km	M3.7
No. 9 1988 Nov. 24 (329)01 11 45	KMU	N42.00°	E144.33°	52.2km	M3.7
No. 10 1988 Nov. 24 (329)12 45 22	NMR	N43.18°	E145.34°	69.7km	M2.8
No. 11 1988 Nov. 24 (329)12 45 22	IWN	N43.18°	E145.34°	69.7km	M2.8
No. 12 1988 Nov. 24 (329)21 34 15	URH	N42.73°	E143.62°	112.0km	M2.0
No. 13 1988 Nov. 24 (329)21 34 15	IWN	N42.73°	E143.62°	112.0km	M2.0
No. 14 1988 Nov. 24 (329)23 11 24	URH	N44.17°	E142.48°	236.4km	M3.9
No. 15 1988 Nov. 24 (330)16 07 04	MYR	N42.60°	E145.25°	23.5km	M3.2

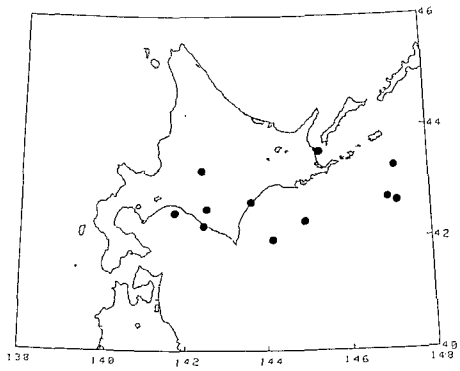
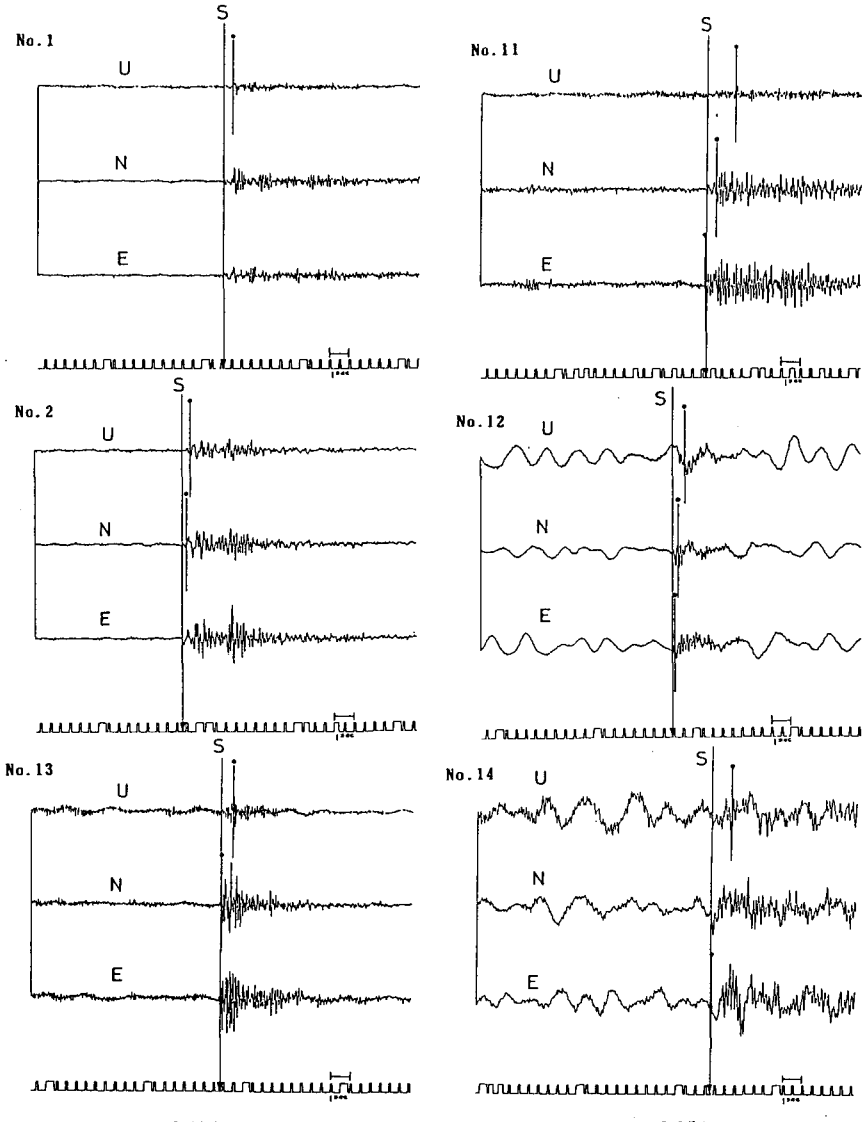


Fig. 8.5. Map showing the epicenters of the earthquakes used in the summation procedure incorporated into the automatic playback operation.

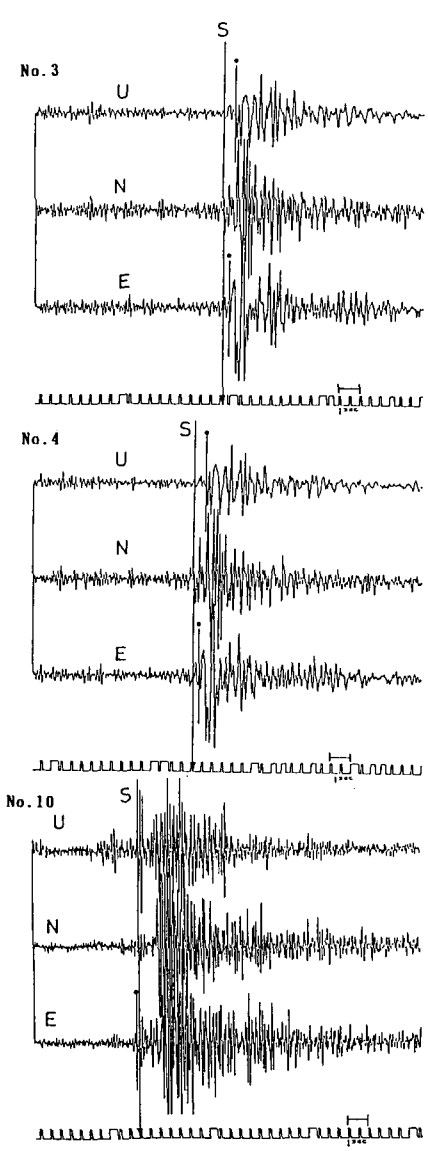
seismograms make comparable contributions to the AIC value in the summing procedure. The events of No. 12 and No. 14 are recorded in the strong background noises as shown in Fig. 8.6(b). However, the conspicuous difference of spectral contents between the background noises and seismic signals led to a great success in this processing. The vertical component seismograms at a S/N ratio lower than 1.0 make less contributions in determining them as before. Figure 8.6(c) exhibits the case when relatively strong S-waves are mixed with the complicated background noises. Even though the predominant frequencies of the background noises are quite close to those of the S-waves as shown in Fig. 8.6(c), the onset times of S-wave can properly be detected by minimizing the AIC's. Moreover, Fig. 8.6(d) represents a more complicated examples of S-waves, which are contaminated with stronger noises, in particular, the event of No. 8 has the strong coda waves, which interfere with the judgment of the onset time of S-wave. However, the results by the automatic processing demonstrate that the minima of AIC's correspond to the onset times of S-waves. The last examples, which are shown in Fig. 8.6(e) are the weak seismograms mixed with the noises having similar frequency contents. They are appropriate examples to investigate how the arrival times of ambiguous signals are determine, or to judge whether our procedure properly detects the onset times of S-waves. Figure 8.6(e) shows that the minima of the AIC reasonably indicate the arrival times. The examples presented above suggest that the automatic processing will succeed in the determination of the onset times of S-waves even if we are not assured where the S-waves begin in visual seismograms.

In the following section, we will evaluate statistically the validity of the procedure for S-waves, which has been working in the automatic playback system.

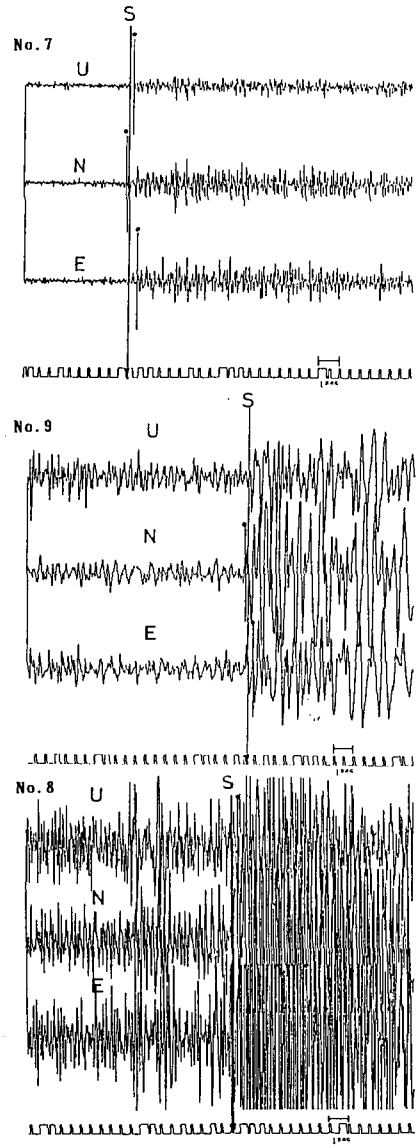


8.6(a) 8.6(b)

Fig. 8.6. Examples of the results through the automatic playback operation. The long solid line marks the picked S-wave arrival time by the summation procedure. The short solid line marks the picked time by applying FUNIMAR to each component. Symbols of U, N, E are the seismograms of up-down, north-south, and east-west components, respectively. Several kinds of examples are shown in (a)~(e). Explanation about events and stations shown in Table 8.3.



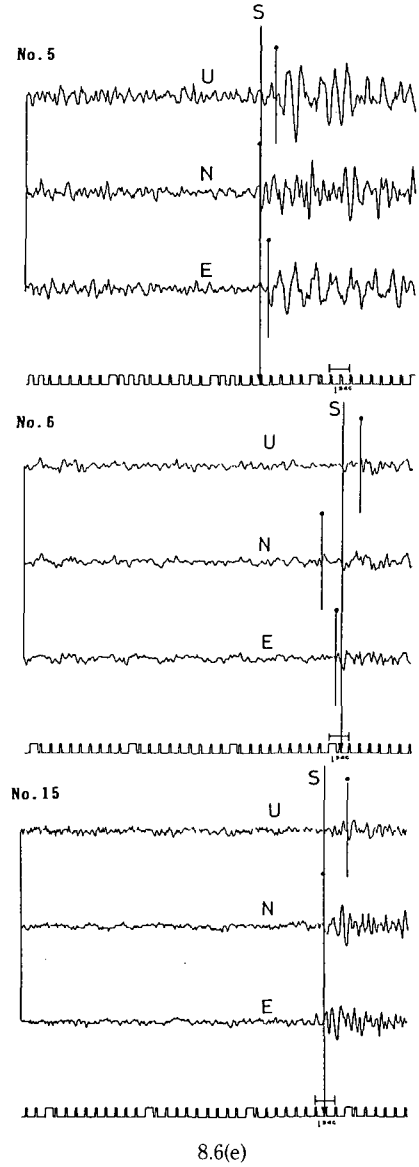
8.6(c)



8.6(d)

8.1.4 Statistical evaluation of the onset times of S-waves

Since we have few arrival time data for S-waves, which have been read by persons in RCEP, we use the hard copy of the seismograms obtained in the course of the automatic processing for S-waves. The time scale of the seismo-



8.6(e)

grams is exactly 4.7 cm/sec, which is larger than the paper speed of 1.0 cm/sec of the 24 channels chart recorder. A seismologist of RCEP reads carefully the three components seismograms of approximately two hundreds earthquakes. It has already been known that an onset time obtained by the summing proce-

dure is nearly in dentical to that by the procedure FUNIMAR applied to each horizontal component seismogram. However we can not insist on the accuracy of the onset times of S-waves since nobody knows which component seismogram indicates the onset time correctly. Therefore, we will adopt temporary measures as to the statistical estimation for the results, namely the time differences between the FUNIMAR's and the person's readings have been investigated for each component seismogram.

The three kinds of histogmams for up-down, north-south, and east-west component seismograms, which are made by the differences of onset times of S-waves between the above two methods, are shown in Fig. 8.7. As in the case of P-waves, the statistics, the mean and the variance, are obtained for the estimation error of S-time for each of three components. The means of them are -0.20 ± 0.094 sec, $+0.06\pm0.076$ sec, and $+0.01\pm0.06$ of 95%, respectively.

The sign '$-$' indicates that the automatic processing determines the onset time earlier than the person. The data sets of sample size 198, 227, and 227 are used in the investigations for vertical, north-south, and east-west component S-arrivals, respectively. As we have seen before, the mean of time differences for vertical component is several times as large as those for the horizontal components. Namely, for the determination of the onset times of S-waves we should

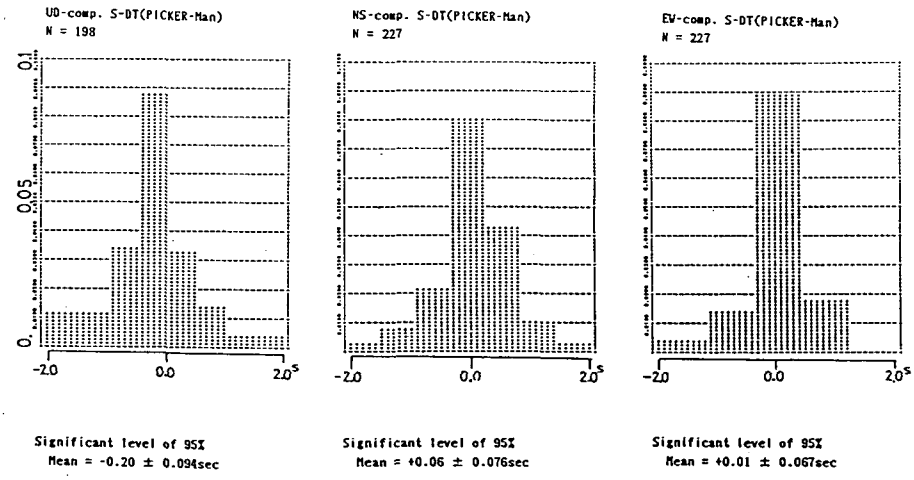


Fig. 8.7. Histogram of difference between the machine picked S-arrival time and the hand picked arrival time (machine-person), for each hand picked arrival quality. The relative frequencies are obtained through the optimized classification (after Sakamoto et al., 1983). In the left : histogram for up-down component, in the middle : histogram for north-south component, in the right : histogram for east-west component.

use the horizontal components of the seismograms.

We have investigated the validity of the automatic detection of S-waves by comparison with the S-time read by a person. In conclusion, our procedure based on the autoregressive model fitting, especially the summing procedure is an efficient method to determine onset times of S-waves. Although it has been well recognized that the procedure based on locally stationary AR model is quite efficient for the detection of P-waves, our investigation demonstrates that it becomes also useful for the detection of S-wave by the proper extension of the model.

9. Extraction of signal by a time series model and screening out microearthquake signals

9.1 Introduction

This chapter is addressed to the problem of decomposition of an observed time series into several components. A specific problem we are considering here is the extraction of microeathquake signal from noisy data. The earthquake is recorded by measuring the ground motion. The ground is always disturbed by various natural phenomena like ocean, wind and tide and by various human activities. As a result, even in normal situation without any earthquakes, the ground fluctuates with significant spectrum on its characteristic frequency band. The motion is called the stationary background noise. Therefore if the amplitude of the earthquake signal is very small, it is quite difficult to distinguish the earthquake signal from the background noise. We will consider here the problem of extracting small earthquake signal from such noisy data.

In a series of papers (Akaike, 1973 ; Kitagawa, 1981 ; Kitagawa and Gersch, 1984), models for nonstationary time series with drifting mean value function are developed and are applied to the seasonal adjustment of economic time series. In the paper (Kitagawa, 1983 ; Kitagawa and Gersch, 1985), a model for nonstationary time series with changing covariance structure is developed and is applied to the estimation of changing spectrum. These works are motivated by Akaike (1979) that solved the following smoothing problem originally posed by Whittaker (1923): Given the observation $y(n)$, $n=1, \cdots, N$, that are obtained from

$$y(n)=f(n)+e(n)$$

where $f(n)$ is an unknown smooth function and $e(n)$ is an ***i.i.d.*** (independent

identically distributed) from $N(0, \sigma^2)$. Whittaker (1923) suggested that the solution for $f(n)$ should balance a tradeoff between infidelity to the data and infidelity to a k-th order difference equation constraint on $f(n)$. The choice of a tradeoff parameter has been left unsolved to investigators. Akaike (1979) showed that the tradeoff parameter can be considered as the hyper-parameter of a Bayesian model and that it can be determined by maximizing the likelihood of the Bayesian model. It can be shown that Akaike's Bayesian model can be expressed in a simple state space form and that the tradeoff parameter is the ratio of the system noise variance and the observational noise variance that can now be estimated by the maximum likelihood method (Kitagawa, 1981). The state space representation facilitates an efficient Kalman filter algorithm for the computation of likelihood. It has been also found that a more complicated component like autoregressive process and an effect from another variable can be incorporated into the model by using the state space form. In the paper (Gersch and Kitagawa, 1983; Kitagawa and Gersch, 1984), the smooth trend, seasonal and stationry stochastic components each is modeled by stochastically perturbed linear difference equation of a certain order. In the papers (Kitagawa, 1983; Kitagawa and Gersch, 1985), the time series is modeled by autoregressive model with time varying coefficients. The smooth change of the autoregressive coefficients is modeled by a stochastically perturbed linear difference equation. The time series we are considering in this paper also has the second type of nonstationarity. Here a signal with different stochastic character is superimposed on a stationary time series. As the result, the observed time series has time varying covariance structure and thus can be modeled by an autoregressive model with time varying coefficients. Since the time series has a specific structure in our case, it will be advantageous to use a model that explicitly models this structure. For example, the record of the seismograph is the superposition of the background noise, earthquake signal and observational noise. In this case, a proper model for the seismic data will be

$$\begin{aligned}(\text{observation}) = &(\text{background noise})\\ &+(\text{earthquake signal})\\ &+(\text{observational noise}).\end{aligned}$$

In this chapter, we will propose to use autoregressive models for each of the background noise and earthquake signal. These models can be effectively estimated by the minimum AIC procedure that was developed as a natural extension of the maximum likelihood method to the situation where the type of the model or the number of parameters is unknown (Akaike, 1973). Since the

amplitude of the earthquake signal varies significantly with time, the innovation variance of the earthquake model is inevitably a function of time. Thus the estimation of changing variance is the crucial problem to get a good result. We will show two procedures for estimating this time varying variance.

In the first procedure, we divide the time interval into several segments and estimate the variance at each segment. The segment length should be set so that it is small enough to follow the change of variance but is large enough to get stable estimate. In the second procedure a new concept of local likelihood is exploited to alleviate this problem of choosing segment length.

In section 9.2, our basic state space model and the Kalman filter and smoother algorithms for the decomposition are shown. In section 9.3, the procedure for the estimation of component models and changing variance are presented. In section 9.4, actual records of a seismograph are analyzed to exemplify the procedure. Some properties and possible simplification of the procedure are discussed in section 9.5.

9.2 The basic model for decomposition

For the decomposition of a time series, we will consider here the model

$$y(n)=r(n)+s(n)+w(n), \tag{9.1}$$

where $w(s)$ is a white noise sequence such that $w(n)\sim N(0,\ \sigma^2)$ and $r(n)$ and $s(n)$ are both autoregressive processes,

$$r(n)=\sum_{i=1}^{m} a(i)r(n-i)+u_r(n),$$

and

$$s(n)=\sum_{i=1}^{j} b(i)s(n-i)+u_s(n). \tag{9.2}$$

Here $u_r(n)$ and $u_s(n)$ are white noise sequences such that $u_r(n)\sim N(0,\ \tau_1^2)$ and $u_s(n)\sim N(0,\ \tau_2^2)$. For the moment, we assume that the parameters of the models σ^2, τ_1^2, τ_2^2, m, j, $b(i)$ and $d(i)$ are known.

The model (9.1) and (9.2) can be combined into a state space model form

$$\boldsymbol{x}(n)=\boldsymbol{F}\boldsymbol{x}(n-1)+\boldsymbol{G}\boldsymbol{v}(n),$$

and

$$\boldsymbol{y}(n)=\boldsymbol{H}\boldsymbol{x}(n)+\boldsymbol{w}(n), \tag{9.3}$$

where $\boldsymbol{x}(n)$ is a $(m+j)$-dimensional state vector, $\boldsymbol{y}(n)$ is an observation at time n. $\boldsymbol{v}(n)$ is a 2-dimensional Gaussian system noise and $\boldsymbol{w}(n)$ is a Gaussian

observational noise with $E[\boldsymbol{v}(n)]=0$, $E[\boldsymbol{w}(n)]=0$, $E[\boldsymbol{v}(n)\boldsymbol{v}(n-i)^t]=\delta(i)\boldsymbol{Q}(n)$, $E[\boldsymbol{w}(n)\boldsymbol{w}(n-i)^t]=\delta(i)\boldsymbol{R}$, and $E[\boldsymbol{v}(n)\boldsymbol{w}(n-i)^t]=0$. The matrices $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{H}$, $\boldsymbol{R}$ and $\boldsymbol{Q}(n)$ are defined by

$$\boldsymbol{F}=\begin{bmatrix} a(1) & \cdots & a(m) & & & & \\ 1 & & & & & & \\ & \ddots & 0 & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ \hline & & & & b(1) & \cdots & b(j) \\ & & & & 1 & & \\ & & & 0 & & \ddots & \\ & & & & & \ddots & \\ & & & & & 1 & 0 \end{bmatrix}, \tag{9.4a}$$

$$\boldsymbol{G}=\begin{bmatrix} 1 & 0 \\ \cdots & \cdots \\ \cdots & \cdots \\ \cdots & \cdots \\ 0 & 0 \\ & \\ 0 & 1 \\ 0 & 0 \\ \cdots & \cdots \\ \cdots & \cdots \\ \cdots & \cdots \\ 0 & 0 \end{bmatrix}, \tag{9.4b}$$

$$\boldsymbol{H}=(1\,0\cdots0\,|\,1\,0\cdots0),$$

$$\boldsymbol{Q}=\begin{bmatrix} \tau_1^{\,2} & 0 \\ 0 & \tau_2^{\,2} \end{bmatrix}, \tag{9.4c}$$

and

$$\boldsymbol{R}=(\sigma^2).$$

Given the observations, $y(1), \cdots, y(N)$, and the initial conditions $\boldsymbol{x}(0\,|\,0)$ and $\boldsymbol{V}(0\,|\,0)$, the one step ahead predictor and the filter together with their estimation error covariance matrices are obtained from the Kalman filtering algo-

rithm,

Prediction (Time Update)

$$\begin{aligned}&\boldsymbol{x}(n\mid n-1)=\boldsymbol{F}\boldsymbol{x}(n-1\mid n-1),\\&\boldsymbol{V}(n\mid n-1)=\boldsymbol{F}\boldsymbol{V}(n-1\mid n-1)\boldsymbol{F}^t+\boldsymbol{G}\boldsymbol{Q}(n)\boldsymbol{G}^t.\end{aligned}\tag{9.5}$$

Filtering (Observation Update)

$$\begin{aligned}&\boldsymbol{K}(n)=\boldsymbol{V}(n\mid n-1)\boldsymbol{H}^t\{\boldsymbol{H}\boldsymbol{V}(n\mid n-1)\boldsymbol{H}^t+\boldsymbol{R}\}^{-1},\\&\boldsymbol{x}(n\mid n)=\boldsymbol{x}(n\mid n-1)+\boldsymbol{K}(n)\times\{\boldsymbol{y}(n)-\boldsymbol{H}\boldsymbol{x}(n\mid n-1\},\\&\boldsymbol{V}(n\mid n)=\{\boldsymbol{I}-\boldsymbol{K}(n)\boldsymbol{H}\}\boldsymbol{V}(n\mid n-1).\end{aligned}\tag{9.6}$$

Using these estimates, the smoothed value of the state $\boldsymbol{x}(n)$ which is the estimate given by the entire observations $y(1), \cdots, y(n)$, is obtained by the fixed interval smoothing algorithm.

$$\begin{aligned}&\boldsymbol{A}(n)=\boldsymbol{V}(n\mid n)\boldsymbol{F}^t\boldsymbol{V}(n+1\mid n)^{-1},\\&\boldsymbol{x}(n\mid N)=\boldsymbol{x}(n\mid n)+\boldsymbol{A}(n)\times\{\boldsymbol{x}(n+1\mid N)-\boldsymbol{x}(n+1\mid n)\},\\&\boldsymbol{V}(n\mid N)=\boldsymbol{V}(n\mid n)+\boldsymbol{A}(n)\times\{\boldsymbol{V}(n+1\mid N)-\boldsymbol{V}(n+1\mid n)\}\boldsymbol{A}(n)^t.\end{aligned}\tag{9.7}$$

Once the smoothing of the state is performed, the smoothed values of $r(n)$ and $s(n)$ are obtained as the first and $(m+2)$-st elements of $\boldsymbol{x}(n\mid N)$, respectively.

9.3 Estimation of parameters of the model

So far, we have assumed that the state space model is completely known. In actual situations, however, the model has various unknown parameters. For the determination of these parameters, we will use the log-likelihood of the state space model,

$$\begin{aligned}L(\theta)&=\log f(y(1), \cdots, y(N)),\\&=\prod_{n=1}^{N} f(y(n)\mid y(1), \cdots, y(n-1))\\&=-1/2\{N\log 2\pi+\sum_{n=1}^{N}\log v(n)+\sum_{n=1}^{N} e(n)^2/v(n)\},\end{aligned}\tag{9.8}$$

where

$$\begin{aligned}&v(n)=\boldsymbol{H}\boldsymbol{V}(n\mid n-1)\boldsymbol{H}^t+\sigma^2,\\&e(n)=y(n)-\boldsymbol{H}\boldsymbol{x}(n\mid n-1).\end{aligned}\tag{9.9}$$

The order of the autoregressive model, m and l, can be determined by minimizing

$$\mathrm{AIC}=-2\max l(\theta)+2\ (\text{number of parameters}).\tag{9.10}$$

The procedure for model fitting realized by minimizing AIC criterion is called the minimum AIC procedure (Akaike, 1973). Hereafter we will show the minimum AIC procedure for the determination of component models and parameters of model.

9.3.1 Background noise model

The background noise model, AR(m)+white noise, can be determined by fitting the model

$$y(n)=r(n)+w(n),$$

and

$$r(n)=\sum_{i=1}^{m} a(i)r(n-i)+u(n) \tag{9.11}$$

to a stationary part of the data. Since the model has a state space representation (9.3) with

$$\boldsymbol{F}=\begin{bmatrix} a(1) & a(2) & \cdots & a(m) \\ 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & 1 & 0 \end{bmatrix},$$

$$\boldsymbol{G}=\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{9.12}$$

$$\boldsymbol{H}=(1\,0\cdots0), \quad \boldsymbol{Q}(\tau_1^2), \quad \boldsymbol{R}(\sigma^2),$$

It is easy to compute the likelihood of the parameters by using (9.8) and (9.9). The parameters of the model, $a(1), \cdots, a(m)$, τ_1^2 and σ^2, are obtained by maximizing the log-likelihood. The order of the autoregression, m, is determined so that the corresponding AIC value becomes smallest.

9.3.2 Earthquake model

The autoregressive model for the earthquake signal can be obtained by fitting the model (9.4) to a part of data where the earthquake signal apparently exists. Here if we assume the stationarity of the background noise and observational noise, it follows that the parameters m, $a(1), \cdots, a(m)$, τ_1^2 and σ^2 are known and only l, $b(1), \cdots, b(l)$ and τ_2^2 are the unknown quantities. They can be estimated by the minimum AIC procedure.

9.3.3 Time varying variance

The variance τ_2^2 of the second autoregressive model is related to the magnitude of the earthquake signal and hence depends on time. When there is no earthquake signal, τ_2^2 should be zero. When earthquake signal arrives, τ_2^2 becomes positive and decreases to zero as the earthquake signal (coda) dies out. In extracting the earthquake signal, this parameter plays the roll of the signal to noise ratio. Thus the estimation of the time varying variance, $\tau_2^2=\tau_2^2(n)$, is an inevitable and crucial problem in our procedure. If τ_2^2 takes a constant value, we can estimate it by maximizing the likelihood. In our case it is actually a function of time. Here we will show two procedures for the estimation of time varying variance.

In the first procedure, we assume that the variance is piece-wisely constant. This assumption was successfully applied to the fitting of autoregressive model with time varying coefficinents for modeling of nonstationary time series (Kitagawa and Akaike, 1978; Ozaki and Tong, 1975). Specifically, we divide the time interval $[1, N]$ to several segments $[1, N_1]$, $(N_1+1, N_2]$, $\cdots$, $[N_{p-1}+1, N_p]$. For simplicity we assume that each segment has the same number of observations, M, possibly except for the last one. On each segment, we can determine the best value of τ_2^2 by maximizing the likelihood. As far as we have experienced it is enough to search for the best value over a coarse grid $c_0 2^{-k}(k=1, \cdots, k_{\max})$. In determining the segment length M, we should take into account the following two facts: if M is too small we can not get reliable estimate, on the other hand, if it is too large, the estimated piece-wise constant function cannot be a good replica of true function.

The second procedure has been developed to alleviate this problem of the choice of segment length. Here we use a new concept of local likelihood. As shown by Akaike (1973), the log-likelihood of a model can be derived as a natural estimate of the expected log likelihood, or of Kullback-Leibler information number except for a common constant, which are natural measures of similarity of true and estimated models. In the case of nonstationary time series, the true model itself veries with time. Therefore, if we measure the goodness of a fixed model by the expected log-likelihood, its value will become a function of time. In other words, the goodness of a model also depends on time. In such a situation, how should we evaluate the model? Our definition of local likelihood follows; The log-likelihood (9.8) substituted the data length N by one

$$L(n)=1/2\{\log 2\pi\, v(n)+e(n)^2/v(n)\} \tag{9.13}$$

is an unbiased estimate of the expected log-likelihood of the model at time n, although the variance is very large and hence is of no use in its form. But if we assume that the true model changes quite slowly and smoothly, the corresponding expected log-likelihood also should change quite smoothly. A natural idea then is to smooth the log-likelihood, $L(n)$, by a proper method to get stabilized estimate. By the smoothing, the variance of the estimate is reduced significanly at the expense of loss of unbiasness. The bias caused by this smoothing will be small if the true model changes slowly and smoothly. For the smoothing of the log-likelihood, we can use the procedure for smoothing chi-square variate proposed by Kitagawa and Gersch (1985). Define $t(m)=\log z(m)+\gamma$ $(m=1, \cdots, n/2)$ where $z(m)=2^{-1}(L(2m-1)+L(2m))$ and γ the Euler constant. We can see $t(m)$ thus defined is distributed nearly normal and with moments

$$E(t(m)]=\log c(m), \qquad Var(t(m))=\pi^2/6,$$

where $c(m)$ is the average of true expected log-likelihood at time $2m-1$ and $2m$. Upon approximating this distribution by a normal distribution with the above two moments and using the smoothess prior model

$$t(m)=2t(m-1)-t(m-2)+v(m),$$
$$v(m)\sim N(0, \rho^2),$$

we can get the smoothed value $\bar{t}(m)$ of $t(m)$. The local likelihood of the model is then obtained by $\bar{L}(2m)=\bar{L}(2m-1)=\exp\{\bar{t}(m)\}$.

The procedure for the estimation of changing variance based on the local likelihood is summarized as follows,

(1) For each value of $\tau_2^2=c_0 2^{-k}(k=1, \cdots, k_{max})$.
  (1-1) Compute $L_k(n)$ $(n=1, \cdots, N)$.
  (1-2) Obtain smoothed value $L_k(n)$ $(n=1, \cdots, N)$.
(2) For each time $n(n=1, \cdots, N)$.
  (2-1) Find k^* that attains the maximum value of $L_k(n)$ $(k=1, \cdots, k_{max})$.
  (2-2) Set $\tau_2^2(n)=c_0 2^{k^*}$.

9.4 *Example : Screening out microearthquake signal from noisy data*

Figure 9.1 shows the record of north-south component of earthquake observed at Hidaka station (HIC) at 9 : 33, March 21, 1982. The magnitude was $M=2.3$. The epicenter is located about 90 km away from the station location. The observed signal is minuscule relative to the background noise. The original record is also sampled at each 0.01084 second with minimum resolution of

1.0. Since from the preliminary investigation, it is found that the sampling interval is too small for time series modeling, four consecutive observations are taken averaged to get one observation for each four original observations. In Fig. 9.1 the time series thus obtained is named "original data". Due to this sampling scheme the sampling error variance of this record is 1/48. We first fit an AR plus white noise model to the stationary part of the data, $n=1$ through 500. Table 9.1 summarizes the fitted models ($m=1, \cdots, 10$). From the table we can see that the AR (7) plus white noise model best fit the data. The estimated parameters of the model are shown in Table 9.2. This is the MAICE model for the stationary background noise. We then fit AR (7)+AR (j)+white noise model ($j=1, \cdots, 10$) to the data from $n=701$ through 1200 where earthquake signal apparently exists. In fitting the models, we have assumed that the AR (7)+white noise model is the same as the one given Table 9.2. The second AR model, the one for the earthquake signal, is estimated by the MAICE procedure. From Table 9.3, we can see that AR (7) is the best model for the earthquake signal. The coefficients of the estimated model are shown in Table 9.4.

Figure 9.2 exhibits the decomposition by the piece-wise modeling of the changing variance. Here the segment length M, is set to 50. The result looks fairly reasonable. The arrival of P wave ($n=550$ through 850) that was not so clear in the original record has become eminent by this decomposition. The only problem here is that we have assumed that the τ_2^2 is a piece-wise constant function with segment length of 50. This assumption may cause false earthquake signal just before the true one. But the decomposition with too small segment length will produce less stable results.

Figure 9.3 looks much nicer than Fig. 9.2. This result is obtained regardless the segment length. Besides, the estimated background noise looks much

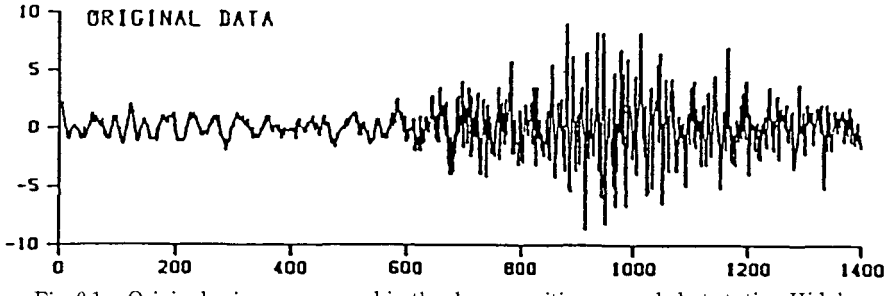


Fig. 9.1. Original seismogram used in the decomposition, recorded at station Hidaka (HIC). Parameters of the hypocenter shown as event of No. 4 fore-shock in Table 4.1.

Table 9.1. Comparison of models for background noise; log-likelihood, number of estimated parameters and AIC's.

Order	L	k	AIC
1	−239.88	3	585.75
2	−207.44	4	422.88
3	−197.46	5	404.92
4	−196.26	6	404.52
5	−194.21	7	402.41
6	−193.31	8	402.62
7	−191.90	9	401.80*
8	−191.80	10	403.64
9	−190.65	11	403.37
10	−190.61	12	405.22

* shows the minimum.

Table 9.2. Minimum AIC model for background noise: Autoregressive coefficients, system noise variance and observational noise variance.

k	$a(k)$
1	0.36543
2	0.78480
3	0.20257
4	−0.16981
5	−0.20106
6	−0.10616
7	−0.03171
$\sigma^2 = 0.04980$	
$\tau_1^2 = 0.04284$	

more uniform over the entire interval of time than the one in Fig. 9.2. This indicates that the estimation of the time varying variance by the local likelihood is appropriate. The figure of the time varying variance also shows the arrivals of P and S waves and gradual decline of earthquake signal.

9.5 Discussion

In the decomposition of the seismic record shown in the previous section, we used two autoregressive models shown in Table 9.2 and 9.4. The theoretical

Table 9.3. Comparison of models for earthquake signal: log-likelihood, numbers of estimated parameters and AIC's.

Order	L	k	AIC
1	−1104.20	2	2212.40
2	−1088.47	3	2182.93
3	−1086.03	4	2179.06
4	−1075.88	5	2161.75
5	−1074.39	6	2160.78
6	−1068.05	7	2150.10
7	−1061.89	8	2139.78*
8	−1061.45	9	2140.90
9	−1061.23	10	2142.45
10	−1060.61	11	2143.22

* shows the minimum.

Table 9.4. Minimum AIC model for earthquake signal: Autoregressive coefficients.

k	a (k)
1	0.72991
2	−0.38229
3	0.16226
4	−0.07983
5	−0.27143
6	0.27017
7	−0.13443

spectra of the corresponding stochastic processes are shown in Fig. 9.4. Inferring from the frequency contents, what we have to do here is to decompose, at each time, the spectrum of a time series as

$$p(f, n)=\tau_1^2 p_1(f)+\tau_2^2 p_2(f)+\sigma^2.$$

Here again, the variance τ_2^2 is an unknown function of time. It should be emphasized that we could estimate this function by the modeling in time domain.

It will be reasonable to ask how reliable or stable the background and earthquake models are. To examine this problem, we apply the models to the records observed by the same seismograph but at different times. The data shown in Fig. 9.5 was observed at 8 : 49, March 21, 1982. The magnitude of the

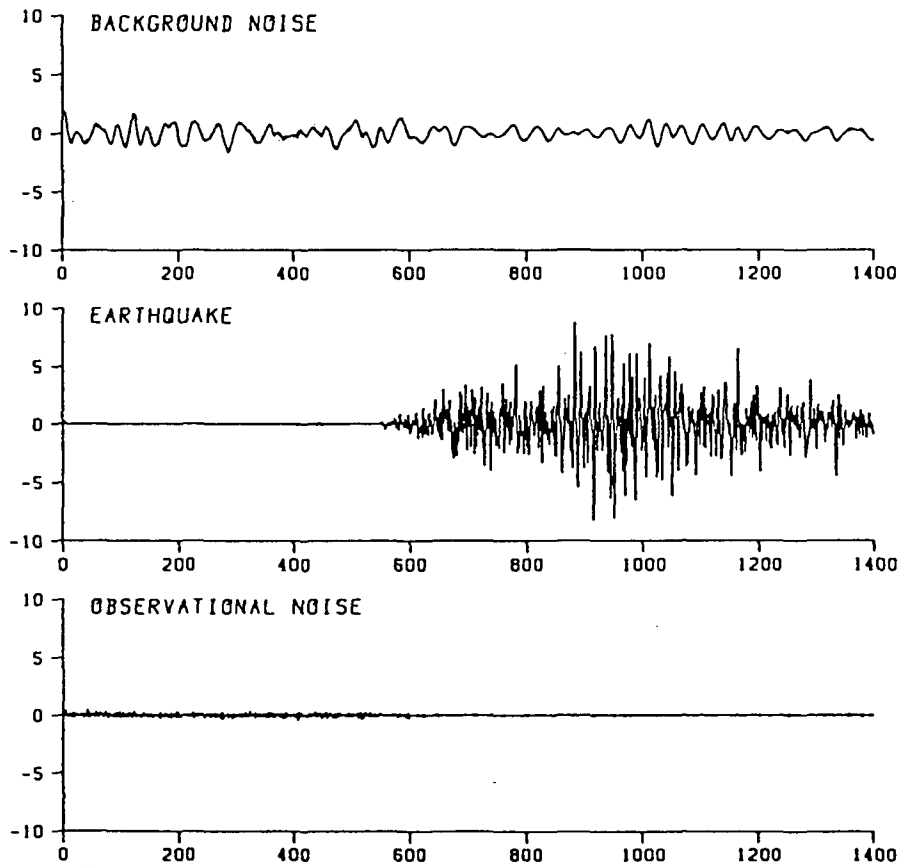


Fig. 9.2. Decomposition by piece-wise constant model : background noise, earthquake signal, observational noise.

earthquake is estimated to $M=2.1$ and the estimated epicenter is close to the one in Fig. 9.1. We emphasize that in this case we do not estimate the AR models from the data but used the ones fitted to the data in Fig. 9.1. In spite of the smaller signal and incorrect models (in the above sense), we can get quite good decomposition. The existence of earthquake signal that is not so clear in the original record in highlighted. From the figure, we can see that P-wave arrives around at $n=400$, S-wave around at $n=680$ and then coda gradually decreases and ends around at $n=1300$. By this decomposition, the background noise and observational noise are obtained to be stationary over time and the decay of coda becomes clear that was not so apparent in the original record.

Fitting AR+white noise and AR+AR+white noise models shown in subsec-

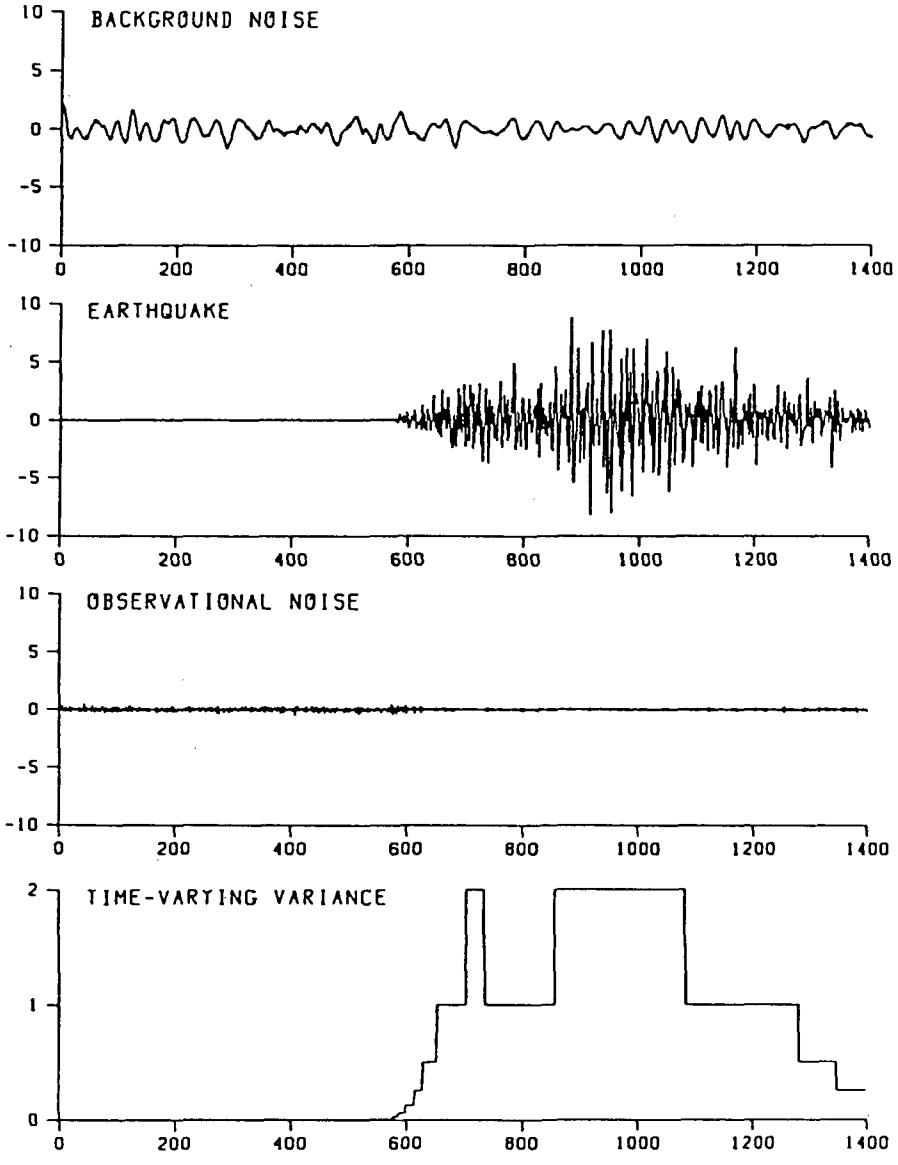


Fig. 9.3. Decomposition by local likelihood: Background noise, earthquake signal, observational noise and the variance of the earthquake model.

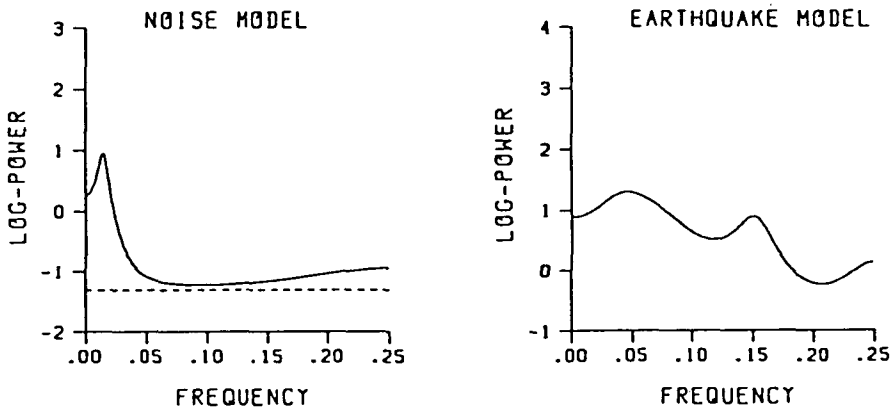


Fig. 9.4. Theoretical spectra of component models: background noise plus observational noise (····) and earthquake signal model. Frequency normalized by sampling rate.

tions 9.3.1 and 9.3.2 are laborious, since we have to estimate the parameters of each model by using numerical optimization procedure. A possible simplification of the procedure is to estimate these parameters by fitting simple AR models to the data, rather than fitting AR+white noise model. Table 9.5 shows the two AR models fitted to the same data as that of Table 9.2 and 9.4. Figure 9.6 shows the decomposition by this simplified model. Since in this case we cannot estimate the observational noise variance, it is set equal to 1/48, the theoretical sampling error variance. Due to the smaller value of observational noise variance, the fluctuation of the background noise is slightly increased in compensation for the reduction of observational noise. But the earthquake signal itself is reasonably extracted from the noisy data. Taking into account the significant reduction of computing cost, this simplified method is recommended for first attempt, although there is no guarantee that it does always work well. From Table 9.1, it is seen that it may also be possible to use lower order model, say $m=3$, for background noise since the difference of AIC value is not so significant.

It may be possible to use this procedure for the decomposition of earthquake signal into several components, e.g. P and S waves and other components. We will develop the procedure for the decomposition of earthquake signal into P and S waves and background noise components in the later chapter 10.

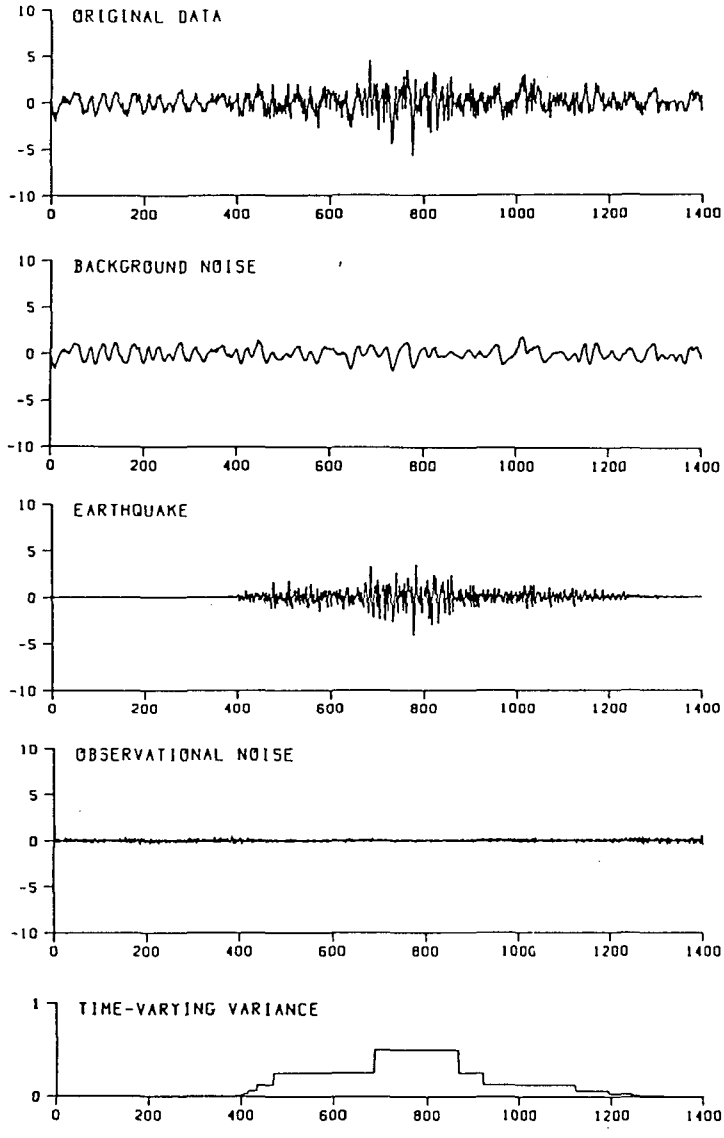


Fig. 9.5. Seismogram of No. 3 fore-shock recorded at station Hidaka (HIC), and decomposition by local likelihood: background noise, earthquake signal, observational noise and the variance of the earthquake model.

Table 9.5. Autoregressive coefficients for background noise and earthquake signal estimated by simplified method.

k	a (k)	b (k)
1	0.36703	0.72810
2	0.52956	−0.37715
3	0.20859	0.14170
4	−0.01162	−0.06454
5	−0.09087	0.32463
6	−0.07556	0.32463
7	−0.12467	−0.19012

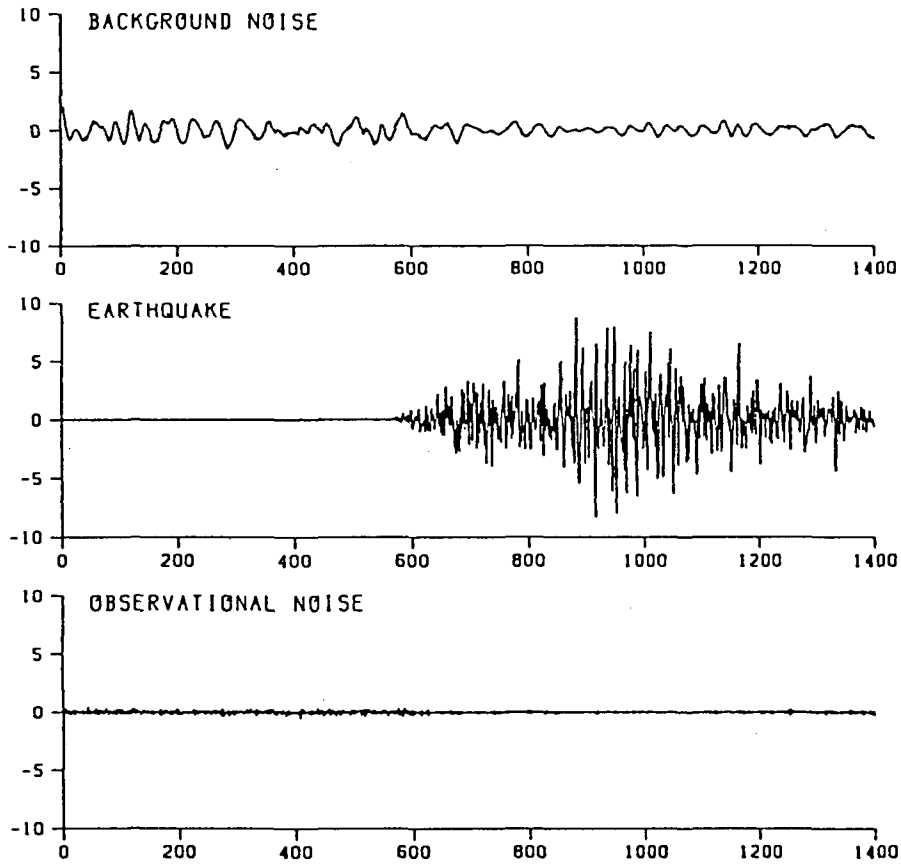


Fig. 9.6. Decomposition of the seismic data shown in Fig. 9.1 by the simplified method : background noise, earthquake signal, observational noise.

9.6 Conclusion

A model for the decomposition of time series into several components has been shown. The estimation of the parameters of the model is crucial problem. The component AR models are determined by the minimum AIC procedure. Two methods for the estimation of changing variance were proposed here. As an example, microearthquake signals are screened out from noisy data. It is found that the fitted model is applicable to another earthquake signal that came from the same epicenter and was observed at the same location as the original one. It is also shown that at least in our case, the simplified procedure proposed in section 9.5 works well and that necessary computation is reduced significantly.

10. Extension of the procedure for the decomposition to four components case

10.1 Extension of the model

As an extension of the model used in the previous sections, we will consider here the model

(observation) = (background noise)
+ (P-wave signal)
+ (S-wave signal)
+ (observational noise)

namely,

$$y(n)=r(n)+P(n)+S(n)+w(n), \tag{10.1}$$

where $w(n)$ is a white noise sequence such that $w(n)\sim N(0, \sigma^2)$ and $r(n)$, $P(n)$ and $S(n)$ are autoregressive processes:

$$r(n)=\sum_{i=1}^{m} a(i)r(n-i)+u_r(n),$$

$$P(n)=\sum_{i=1}^{j} b(i)P(n-i)+u_p(n),$$

and

$$S(n)=\sum_{i=1}^{k} c(i)S(n-i)+u_s(n). \tag{10.2}$$

Here, $u_r(n)$, $u_p(n)$ and $u_s(n)$ are white noise sequences such that $u_r(n)\sim N(0, \tau_1^2)$, $u_p(n)\sim N(0, \tau_2^2)$ and $u_s(n)\sim N(0, \tau_3^2)$. For the moment, we assume that

the parameters of the models σ^2, τ_1^2, τ_2^2, τ_3^2, m, j, k, $a(i)$, $b(i)$ and $c(i)$ are known.

The model (10.1) and (10.2) can be combined into a state space model form

$$\boldsymbol{x}(n)=\boldsymbol{F}\boldsymbol{x}(n-1)+\boldsymbol{G}\boldsymbol{v}(n),$$

and

$$y(n)=\boldsymbol{H}\boldsymbol{x}(n)+\boldsymbol{w}(n), \tag{10.3}$$

where $\boldsymbol{x}(n)$ is a $(m+j+k)$-dimensional state vector, $y(n)$ is an observation at time n. $\boldsymbol{v}(n)$ is a 3-dimensional Gaussian system noise, $\boldsymbol{w}(n)$ is a Gaussian observational noise with $E[\boldsymbol{v}(n)]=0$, $E[\boldsymbol{w}(n)]=0$, $E[\boldsymbol{v}(n)\boldsymbol{v}(n-i)^t]=\delta(i)\boldsymbol{Q}(n)$, $E[\boldsymbol{w}(n)\boldsymbol{w}(n-i)^t]=\delta(i)\boldsymbol{R}$, and $E[\boldsymbol{v}(n)\boldsymbol{w}(n-i)^t]=0$. The matrices $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{H}$, $\boldsymbol{R}$ and $\boldsymbol{Q}(n)$ are defined by

$$\mathbf{F}=\begin{bmatrix}
a(1) & \cdots & a(m) & & & & & & & \\
1 & & & & & & & & & \\
 & \ddots & & & & & & & & \\
 & & \ddots & & 0 & & & & & \\
 & & & \ddots & & & & & & \\
 & & & & 1 & 0 & & & & \\
\multicolumn{10}{c}{\dotfill} \\
 & & & & & b(1) & \cdots & b(j) & & \\
 & & & & & 1 & & & & \\
 & & & & & & \ddots & & & \\
 & & & & 0 & & \ddots & & & \\
 & & & & & & & 1 & 0 & \\
\multicolumn{10}{c}{\dotfill} \\
 & & & & & & & c(1) & \cdots & c(k) \\
 & & & & & & & 1 & & \\
 & & & & & & & & \ddots & \\
 & & & & 0 & & & & \ddots & \\
 & & & & & & & & & 1 \quad 0
\end{bmatrix}, \tag{10.4a}$$

$$\mathbf{G}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ & \cdots\cdots & \\ & \cdots\cdots & \\ & \cdots\cdots & \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ & \cdots\cdots & \\ & \cdots\cdots & \\ & \cdots\cdots & \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ & \cdots\cdots & \\ & \cdots\cdots & \\ & \cdots\cdots & \\ 0 & 0 & 0 \end{bmatrix}, \tag{10.4b}$$

$$\mathbf{H}=(10\cdots 0\,|\,10\cdots 0\,|\,10\cdots 0),$$

$$\mathbf{Q}=\begin{bmatrix} \tau_1^2 & 0 & 0 \\ 0 & \tau_2^2 & 0 \\ 0 & 0 & \tau_3^2 \end{bmatrix}, \text{ and}$$

$$\mathbf{R}=(\sigma^2).$$

The autoregressive model for the P-wave signal can be obtained by fitting the model, $y(n)=r(n)+P(n)+w(n)$ to a part of data where P-wave signal apparently exists. Here if we assume the stationarity of the background and observational noise, it follows that the parameters m, $a(1), \cdots, a(m)$, τ_1^2 and σ^2 are known and only j, $b(1), \cdots, b(j)$ and τ_2^2 are the unknown quantities. They can be estimated by the minimum AIC procedure. Then the autoregressive model for S-wave signal can be obtained by fitting the model (10.1), $y(n)=r(n)+P(n)+S(n)+w(n)$ to a part of data where S-wave signal exists. Parameters m, $a(1), \cdots, a(m)$, τ_1^2, τ_2^2 and σ^2 are known and only k, $c(1), \cdots, c(k)$ and τ_3^2 are the unknown quantities. They can also be estimated by the minimum AIC procedure. The variance τ_2^2 and τ_3^2 of the second and the third autoregressive models are relate to the magnitude of the earthquake signal and hence depend on time. Here we use the second procedure for estimation of time varying variance shown in previous section 9.3.

10.2 Example : Decomposition of microearthquake signal into P-wave, S-wave, and background noise

Figure 10.1 shows the record of east-west component of microearthquake signal observed at Hiroo, MYR, Japan, at 07 : 45, March 25, 1982. The magnitude is $M=1.9$. The epicenter is about 40 km to the west of the station MYR. The original record is also sampled at each 0.01084 seconds with minimum resolution of 1.0. We will call it original data again.

The problem here is the estimation of the parameters of autoregressive models of background noise, P-wave and S-wave with their innovation variances. Following the extended model described in section 10.1, we obtain the models of background noise, P-wave and S-wave by the minimum AIC estimation procedure.

We first fit an AR+white noise model to the stationary part of the data, $n=200$ through 400. From the model fitting we can see that AR (4)+white noise model best fits the data. This is the MAICE model for the stationary background noise. The estimated parameters of the model are shown in Table 10.1. We then fit AR (4)+AR (j)+white noise model to the data from $n=650$ through 1000 where P-wave signal apparently exists. The second AR model, the one for the P-wave signal, is estimated by the MAICE procedure. We also find that AR (4) is the best model for the P-wave signal. The estimated parameters of the model are shown in Table 10.2. We finally fit AR (4)+AR (4)+AR (k)+white noise model to the data from $n=1050$ through 1450 where S-wave signal apparently exists. The third AR model, the one for the S-wave signal, is also estimated by the MAICE procedure. We can see that AR (6) is the best model for the S-wave signal. The coefficients of the estimated model are shown in Table 10.3.

Figure 10.1 exhibits the decomposition by the piece-wise modeling of changing variances for model P and S-waves. The result is free from the segment length as the second method for the estimation of changing variance described in the chapter 9. The theoretical spectra of the corresponding stochastic processes are shown in Fig. 10.1 Incidentally, a small wave train apparently exists prior to the main S-wave as found in the middle of Fig. 10.1. This small train, however, has gone unnoticed before. According to the above decomposition, it is apart of S-wave model. In the particle motion analysis in the preceding chapter, we observed certainly that the polarizations of S-waves frequently showed abrupt changes in direction. Since crustal microearthquakes are caused by the release of accumulated shear stress, and since the S-waves

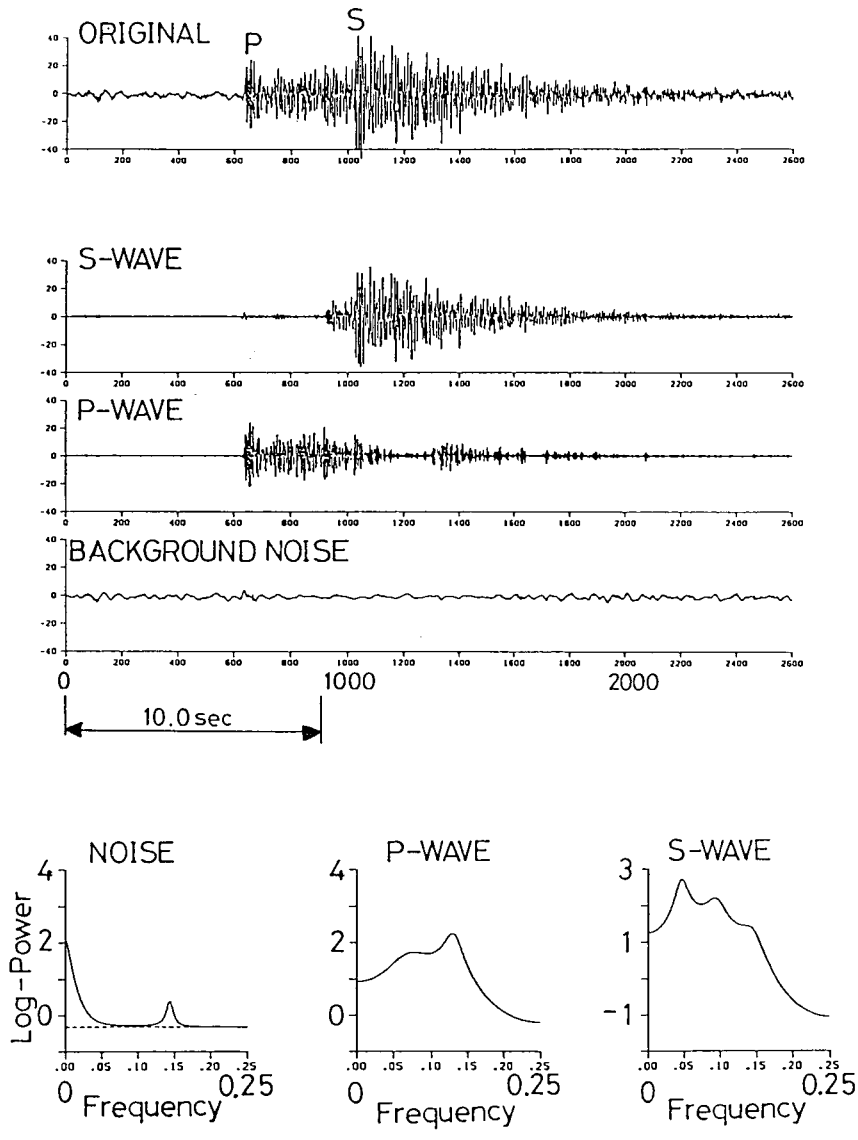


Fig. 10.1. Decomposition by local likelihood: an east-west component seismogram at station Moyori (MYR). Parameters of hypocenter shown as event of No. 1 fore-shock in Table 4.1. From top to bottom: original data, S-wave, P-wave, background noise, and theoretical spectra of component models (background noise plus observational noise (····), P-wave and S-wave signal models. Frequency normalized by sampling rate).

Table 10.1. Minimum AIC model for background noise: Autoregressive coefficients, system noise variance and observational noise variance.

k	a (k)
1	1.29979
2	−0.87154
3	1.24548
4	−0.69901
$\sigma^2 = 0.48329$	
$\tau_1^2 = 0.15774$	

Table 10.2. Minimum AIC model for P-wave: Autoregressive coefficients.

k	a (k)
1	1.29979
2	−0.87154
3	1.24548
4	−0.69901

Table 10.3. Minimum AIC model for S-wave: Autoregressive coefficients.

k	a (k)
1	1.74520
2	−2.31796
3	2.08180
4	−1.68239
5	0.88497
6	−0.41998

are capable of carrying more than three times as much information as the P-waves (Crampin, 1985), analysis of S-waves has the potential to reveal far more about the ray-path and the source than from the analysis of P-waves. Detailed study of three component records of near microearthquakes may support these suggestions. P-wave signals are usually simple, and consequently carry very little information. S-waves, in contrast, are complicated, typically possessing

several different vector polarizations, and are rich in important information about the source and the ray-path.

10.3 Conclusion

A model for the decomposition of time series into four components, P-wave, S-wave, background noise, and observational noise is shown. The estimation of the parameters of the model is crucial problem. The component AR models are fitted by the minimum AIC procedure as shown in chapter 9. The method, which is one of the two methods for the estimation of changing variance proposed in the preceding chapter 9, is used. As an example, P and S-wave signals are screened out from noisy wave data by using the fitted model. And there arises a related incident result, which has been found just before the main S-wave. This decomposition analysis is surely useful in detail study of shear waves, which have an important information about the ray-paths and the sources. Further we have applied the model for decomposition of time series to

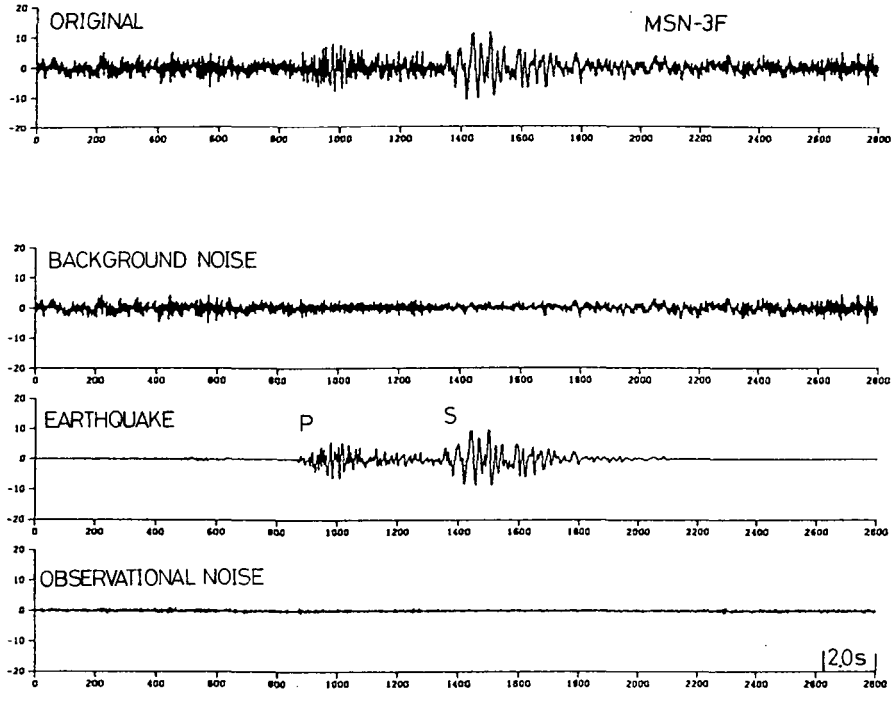


Fig. 10.2. Decomposition of the noisy seismic data shown at the top of Fig. 6.1 (c). From the top: original, background noise, earthquake signal, and observational noise.

the noisy seismogram which is used in chapter 6. In Fig. 10.2, the model for the decomposition of time series into background noise, seismic signal, observational noise is shown. The screening out heavy noisy seismic signal is also shown to exemplify the proposed procedure. A thoroughly screened seismic signal gives us correct information about onset time of P-waves and S-waves.

11. Summary

In this paper, the methodologies for the objective determination of arrival times and for the extraction of necessary information from the observed seismic waves are studied. The efficient and reliable computer programs are developed for seismic analysis of the data obtained by the Hokkaido University Earthquake Recording System.

Before going on to develop the above methodologies, some conventional methods for extracting P and/or S-waves from the microseismic noisy data by using the particle motion analyses (a direct visual method and two statistical methods based on the three principal components analysis), that is, they are the orbit spectrum analysis, the moving window analysis, and the REMODE filter analysis. These preliminary examinations show that the features of seismograms of microearthquakes can be seen to be complicated. Therefore, we have tried to develop a new efficient method by autoregressive modeling to determine P and S-wave arrival times regardless of their ambiguities. Besides, a state space model with autoregressive component models is used for the decomposition of noisy seismograms into seismic signals and unwanted noises. Finally, we can summarize the results of this study as followings.

1. We have first developed an efficient new procedure based on the adoption of Householder transformation for such least squares computations after Kitagawa and Akaike (1978). This algorithm provides a very simple procedure for handling additional observations which is useful for the on-line fitting of locally stationary autoregressive models.

2. Besides, we have developed a three dimensional autoregressive model for determining onset times of S-waves. It is based on the algorithm extensive to three-variate autoregressive modeling.

3. The arrival times determined by the computer system developed in this paper compare favorably with hand picked arrival times. The 2032 P-arrivals picked by both seismologists and the machine provide a reasonable measure of the accuracy of the arrival times with a discrepancy of 0.04 ± 0.011 seconds (95% confidence limit).

4. About 200 S-arrivals picked by both a seismologist and the machine show discrepancies of 0.01 ± 0.067 seconds, 0.06 ± 0.076 seconds and 0.02 ± 0.094 seconds for EW, NS, and UD components, respectively. The picker used in this paper is much more lenient in what it will accept as an arrival and the final tuning of parameters is left up to the location stage of the system. The real power of the system comes from combining the picking and locating processes.

5. Next, we have developed a model for decomposition of time series into several components. In the model each component is expressed by an autoregressive model. The crucial problem of estimating changing variance of the model is solved by the techniques of piece-wise modeling and local modeling. The extractions of microearthquake signal and P and S-wave signals are shown to exemplify the power of the proposed procedures. The stability of the procedure and a possible simplification of the procedure are also considered by using these microearthquake signals.

Acknowledgements

The author expresses gratitude to Prof. Hiroshi Okada for his support, patience and guidance throughout this investigation. The author is also deeply indebted to Associate Prof. Genshiro Kitagawa of the Institute of Statistical Mathematics, Tokyo, and Associate Profs. Tsutomu Sasatani, Ichiro Nakanishi, and Dr. Yoshinobu Motoya who provided many fruitful hours of discussion on his study. Dr. Motoya also helped him to provide the reading data of S-phase onset times for statistical tests and the several useful programs for the real-time examinations. In particular, the author thanks to Associate Profs. Genshiro Kitagawa, Tsutomu Sasatani, and Ichiro Nakanishi for careful reading and useful comments to improve the manuscript.

The author would also like to thank Associate Prof. Yasunori Nishida, President Inatani of the Katsujima Manufactory Co. LTD, Tokyo, and Prof. Takakichi Kaneko of the Senshu University, Bibai for their encouragements and all staffs of the Research Center for Earthquake Prediction for their supports throughout this study.

The computations were made by HITAC M682H at the Hokkaido University Computer Center and by MS190 at the Research Center for Earthquake Prediction. The program package of TIMSAC-78 which was made by Prof. Hirotugu Akaike and his colleagues of the Institute of Statistical Mathematics was very helpful.

This paper is part of the author's doctorial thesis submitted to the Hok-

kaido University in March, 1990.

References

Aggaewal. Y.P., L.H. Sykes, J. Armbruster, and M.L. Sbar, 1973. Premonitory changes in seismic velocity and prediction of earthquakes. Nature, **241**, 101-104.

Akaike, H., 1970. Statistical predictor identification. Ann. Inst. Statist. Math., **22**, 203-217.

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. 2nd International Symposium on Information Theory, ed. B.N. Petrov and F. Csaki (Budapest Akademiai Kiado), 267-281.

Akaike, H., 1979. Likelihood and Bayes procedure. International Meeting on Bayesian Statistics, May 28~June 2, 1979, Velencia, Spain, 1-34.

Akaike, H., G. Kitagawa, E. Arahata, and F. Tada, 1979. TIMSAC78. Computer Science Monographs, A publication of the Institute of Statistical Mathematics, **11**, 1-275.

Aki, K., 1967. Scaling low of seismic spectrum. J. Gephys. Res., **72**, 1217-1231.

Aki, K., 1968. Seismicity and seismological method. Tectonophysics, **6**, 41-58.

Aki, K., 1976. Signal to noise ratio in seismic measurements. in Volcanoes and Tectonosphere, Tokai Univ. Press, Tokyo, 187-192.

Allen, R.V., 1978. Automatic earthquake recognition and timing from single traces. Bull. Seism. Soc. Am., **68**, 1521-1532.

Anderson, K.R., 1978. Automatic processing of local earthquake data. Ph. D. Thesis, Massachusetts Institute of Technology, Cambridge, 173 pp.

Archambeau, C.B. and E.A. Flinn, 1965. Automated analysis of seismic radiations for source characteristics. Proc. I.E.E.E., **53**, 1876-1884.

Archambeau, C.B., E.A. Flinn, and D.G. Lambert, 1969. Fine structure of the upper mantle. J. Geophys. Res., **74**, 5825-5865.

Asada, T., 1982. Earthquake prediction techniques, Their application in Japan. University of Tokyo Press, 317 pp.

Basham, P.W., and R.M. Ellis, 1969. The composition of P codas using magnetic tape seismograms. Bull. Seism. Soc. Am., **59**, 473-486.

Born, M., and E. Wolf, 1965. Principles of optics. 3rd ed., Pergamon Press, New York, 544-555.

Box, G.E.P. and G.M. Jenkins, 1970. Time series analysis, Forecasting and control. Holden-Day, San Francisco.

Brune, J.N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes. J. Geophys. Res., **75**, 4997-5009.

Buland, R., 1976. The mechanics of locating earthquakes. Bull. Seism. Soc. Am., **66**, 173-187.

Crampin, S., 1985. Evaluation of anisotropy by shear-wave splitting. Geophysics, **50**, 142-152.

Dziewonski, A., S. Bloch, and M. Landisman, 1969. A technique for the analysis of transient seismic signals. Bull. Seism. Soc. Am., **59**, 427-444.

Eaton, J.P., W.H.K. Lee, and L.C. Parkiser, 1970. Use of microearthquakes in the study of the mechanics of earthquake generation along the San Andreas fault in Central California. Tectonophysics, **9**, 259-285.

Flinn, E.A., 1965. Signal analysis using rectilinearity and direction of particle motion. Proc. I.E.E.E., **53**, 1874-1876.

Fowler, R.A., B.J. Kotick, and R.D. Elliott, 1967. Polarization analysis of natural and artificially induced geomagnetic micropulsations. J. Geophys. Res., **72**, 2871-2883.

Freeman, H.W., 1966a. The "little variable factor" a statistical decision of the readings of seismograms. Bull, Seism. Soc. Am., **56**, 593-604.

Freeman, H.W., 1966b. A statistical discussion of Pn residuals from explosions. Bull. Seism. Soc. Am., **56**, 677-695.

Freeman, H.W., 1968. Seismological measurements and measurement error. Bull. Seism. Soc. Am., **58**, 1261-1271.

Gersch, W., and G. Kitagawa, 1983. The prediction of time series with trends and seasonalities, Journal Business and Economic Statistics. **1**, 253-264.

Golub, G.H., 1965. Numerical methods for solving linear least squares problems. Numer. Math., **7**, 206-219.

Griffin, J.N., 1966a. Application and development of polarization (REMODE) filters. Seismic Data Laboratory Report **141**, Teledyne Inc., Alexandria, Va. (AD-630-515).

Griffin, J.N., 1966b. REMODE signal/noise tests in polarized noise. Seismic Data Laboratory Report **141**, Teledyne Inc., Alexandria, Va. (AD-800-039).

Gutenberg, B., and C.F. Richter, 1941. Seismicity of the Earth. Geol. Soc. Am., Spec. Pap., **34**, 133 pp.

Hagiwara, T., 1964. Brief description of the project proposed by the earthquake prediction. Research group of Japan, Proc. U.S.-Japan Con. Res. Relat. Earthquake Prediction Probl., 10-12.

Hamaguchi, H. and Z. Suzuki, 1979. Automatic detection of onset time of microearthquake and its confidence. Report of Grant-in-aid for Scientific Research Project on Natural Disaster (No.-54-2), Ministry of Education, Science and Culture, 62-83, (in Japanese).

Hamaguchi, H. and Y. Morita, 1980. Second-order autoregressive representation of short-period seismic waves. Zisin, Ser. 2, **33**, 131-140, (in Japanese with English abstract).

Hasegawa, A., N. Umino, A. Yamamoto, and A. Takagi, 1986. Automatic event detection and location system of microearthquake observation network. Zisin, Ser. 2, **39**, 381-395, (in Japanese with English abstract).

Iizuka, S, 1979. Seismic wave velocity changes as an earthquake precursor and some related phenomena. Ph. D. Thesis, Toukai University, 147 pp.

Iyengar, R.N., and K.T.S.R. Iyengar, 1969. A nonstationary random process model for earthquake accelerograms. Bull. Seism. Soc. Am., **59**, 3, 1163-1188.

Kaneko, T., and N. Watanabe, 1982. Multiple filter orbit analysis of seismograms. Proc, 6th Japan Earthquake Eng. Symp., Tokyo, 409-416.

Kasahara, M., 1976. Seismic and geodetic observations through the digital PCM telemetering system in Hokkaido, Japan. Jour. Geodetic. Soc. Japan, **22**, 2, 292-294.

Kerr, R.A., 1978. Earthquake prediction proving elusive. Science, **200**, 419-421.

Kitagawa, G. and H. Akaike, 1978. Procedure for the modeling of non-stationary time series. Ann. Inst. Statist. Math., 30.

Kitagawa, G., 1981. A nonstationary time series model and its fitting by a recursive technique. Journal of Time Series Analysis, **2**(2), 103-116.

Kitagawa, G., 1983. Changing spectrum estimation, Journal of Sound and Vibration. **89**(4), 433-445.

Kitagawa, G., and W. Gersch, 1984. A smoothness priors-state space modeling of time series with trend and seasonality. Journal of American Statistical Association, **79**(386), 378-389.

Kitagawa, G., and W. Gersch, 1985. A smoothness priors-time varying AR coefficient modeling of nonstationary covariance time series. I.E.E.E., Trans. Automatic Control, **30**(1), 48-56.

Kitagawa, G. and T. Takanami, 1985. Extraction of signal by a time series model and screening out micro earthquakes. Signal Processing, 8, 303-314.

Kubo, T., and J. Penzien, 1975. Characteristics of the three-dimensional ground motions, San Fernando Earthquake. Proceeding of the Review Meeting U.S. Japan Cooperative Research Program in Earthquake Engineering with Emphasis on Safety of School Buildings, Honolulu, Hawaii, 35-51.

Kullback, S. and R.A. Leibler, 1951. On information and sufficiency. Ann. Math. Statist., **22**, 79-86.

Lewis, B.T.R., and R.P. Reyer, 1968. A seismic investigation of the upper mantle to the west of Lake Superior. Bull. Seism. Soc. Am., **58**, 565-596.

Lindh, A.G., G. Fuis, and C. Mantis, 1978. Seismic amplitude measurements suggest foreshocks have different focal mechanisms than aftershocks. Science, **201**, 56-59.

Lindh, A.G., D.A. Lockner, and W.H.K. Lee, 1978. Velocity anomalies: An alternative explanation. Bull. Seism. Soc. Am., **68**, 721-734.

Loh, C.H., 1985. Analysis of the spatial variation of seismic waves and ground movements from SMART-1 array data. Earthquake Eng. Struct. Dyn., **13**, 561-581.

Maeda, I., 1978. Telemetry system for observation of microearthquake and crustal movement. Note of the Research Center for Earthquake Prediction, Faculty of Science, Hokkaido University **3**, 1-7, (in Japanese).

Maeda, I., 1982. Analysis of incident directions of seismic rays by the principal component analysis. J. Fac. Sci. Hokkaido Univ., **VII**, 7, 169-184.

Maeda, N., 1985. A method for reading and checking phase time in auto-processing system of seismic wave data. Zisin, Ser. 2, **38**, 365-379, 1985 (in Japanese with English abstract).

Maeda, N., 1986. On reliability of the results obtained by the method for automatic determination of phase times from seismic wave data. Zisin, Ser. 2, **39**, 555-566, (in Japanese with English abstract).

Matsu'ura, R.S., 1986. Precursory quiescence and recovery of aftershock activities before some large aftershocks. Bull. Earthq. Res. Inst. Univ. Tokyo, **61**, 1-65.

McNally, K.C., and T.V. McEvilly, 1977. Velocity contrast across the San Andreas fault in Central California: Small-scale variations from P wave nodal plane distortion. Bull. Seism. Soc. Am., **67**, 1565-1576.

Mims, C.H., and R.L. Sax, 1965. Rectilinear motion direction (REMODE). Seismic Data Laboratory Report **118**, Teledyne, Inc., Alexandria, Va. (AD-460-631).

Miyamachi, H. and T., Moriya, 1984. Velocity structure beneath the Hidaka mountains in Hokkaido, Japan. J. Phys. Earth, **32**, 13-42.

Mogi. K., 1977. Seismic activity and earthquake prediction. in Proceeing of the Symposium on Earthquake Prediction Research, 203-215, (in Japanese).

Montalbetti, J.F., and E.R. Kanasewich, 1970. Enhancement of teleseismic body phases with a polarization filter. Geophys. J. Roy. Astr. Soc., **21**, 119-129.

Morita, Y and H. Hamaguchi, 1981. Automatic detection of S-onset times using two dimensional autoregressive model fitting. Zisin, Ser. 2, **34**, 223-240, (in Japanese with English abstract).

Morita, Y and H. Hamaguchi, 1984. Automatic detection of onset times of seismic waves and its confidence interval using the autoregressive model fitting. Zisin, Ser. 2, **37**, 281-293, (in Japanese with English abstract).

Moriya, T and M. Ookame, 1972. Structure at the Hidaka and the Kamuikotan Belts in Hokkaido. Programme and Abstracts, The Seismological Society of Japan, 1972 (2), **91**, (in Japanese).

Motoya, Y., 1969. Earthquake observations at the Urakawa Seismological Observatory (I). Geophysical Bulletin of Hokkaido University, **22**, 39-48, (in Japanese with English abstract).

Motoya, Y., 1988. Data processing employing an optical disk system — Analysis of vibration

modes of P wavelets—. Zisin, Ser. 2, **41**, 3, 411-417, (in Japanese with English abstract).

Nakajima, T., 1974. Spatial and sequential distribution of focal mechanisms before and after Tokachi-oki earthquake of May 16, 1968. Geophysical Bulletin of Hokkaido University, **32**, 25-42, (in Japanese with English abstract).

Nakanishi, I., 1989. Determination of depth and fine structure of upper mantle discontinuities. Report of Grant-in-aid for Scientific Research Project (No. 6254085), Ministry of Education, 99 pp.

Ohtake, M., 1973. Change in V_p/V_s ratio with occurrence of some shallow earthquakes in Japan. J. Phys. Earth, **21**, 173-184.

Okada, H., 1984. Re-examination of the detection technique of premonitory change in seismic wave velocity. National committee for Seismology, Science Council of Japan, Seismological Society of Japan, 217-220, (in Japanese).

Ozaki, T and H. Tong, 1975. On the fitting of non-stationary autoregressive models in the time series analysis. Proceeding of the 8th Hawaii International Conference on System Science, Western Periodical Company.

Paulson, K.V., A. Egeland, and F. Eleman, 1965. A statistical method for the quantitative analyses of geomagnetic giant pulsations. J. Atmosphere and Terrestrial Physics, **27**, 943-967.

Pearce, R.G., and B.J. Barley, 1977. The effect of noise on seismograms. Geophys. J. Roy. Astron. Soc., **48**, 543-547.

Rikitake, T., 1976. Earthquake prediction. Elsevier, Amsterdam, 357 pp.

Rikitake, T., 1979. Classification of earthquake precursors. Tectonophysics, **54**, 293-309.

Robinson, R., R.L. Wesson and W.L. Ellsworth, 1974. Variation of P-wave velocity before the Bear Valley California earthquake of 24 February 1972. Science, **184**, 1281-1283.

Sakai, Y., 1883. General Report on the Urakawa-Oki, Japan, Earthquake, of March 21, 1982. Report of Grant-in-aid for Scientific Research Project on Natural Disaster (No. 56020014 and No. 57020018), Ministry of Education, Science and Culture, (in Japanese).

Sakamoto, Y., M. Ishiguro, and G. Kitagawa, 1983. Information Statistics. JYOHOU KAGAKU KOUZA A.5.4, ed. by Kyouritsu Press, Japan, 236 pp, (in Japanese).

Sakata, N., T. Maekawa, and Hm. Okada, 1981. The 3-dimensional velocity structure at Usu-Volcano. Programme and Abstracts, The Seismological Society of Japan (I), **93**, (in Japanese with English abstract).

Semenov, A.N., 1969. Variations in the travel times of transverse and longitudinal waves before violent earthquakes. Izu. Aca. Sci., USSR, Phys. Solid Earth, **3**, 245-248, 1969.

Shirai, K. and I. Tokuhiro, 1979. Detection of seismic onsets. Zisin, Ser. 2, **32**, 141-148, (in Japanese with English abstract).

Stewart, S.W., 1977. Real time detection and location of local seismic events in Central California. Bull. Seism. Soc. Am., **67**, 433-452.

Sugiyama, T., 1989. Upper mantle P-wave structure of the northwestern margin of the Pacific Ocean. The master thesis of science, Hokkaido University, 1-80.

Suzuki, S., T. Takanami, Y. Motoya, M. Kasahara, and I. Nakanishi, 1986. Automatic processing system for microearthquake netwok of Hokkaido University. Program. Abstr., Seismol. Soc. Japan, April 1986, 287 pp, (in Japanese).

Takagi, A., 1982. How small earthquakes occur. Earthquake Prediction Techniques, University of Tokyo Press, 63-88.

Takanami, T., 1982. Three-dimensional seismic structure of the crust and upper mantle beneath the orogenic belts in the southern Hokkaido, Japan. J. Phys. Earth, **30**, 87-104.

Takanami, T., 1984. Applications of AR model to automatic processing for determining onset times of seismic phases. Center News, Oogata Computer Center, Hokkaido

University, **16**(1), 16-30, (in Japanese).

Tjøstheim, D., 1975a. Some automatic models for short-period seismic noise. Bull. Seism. Soc. Am., **65**, 3, 677-691.

Tjøstheim, D., 1975b. Autoregressive representation of seismic P-wave signals with an application to the problem of short-period discriminants. Geophys. J. Roy. Astron. Soc., **43**, 269-291.

Utsu, T., 1973. Temporal variations in travel time residuals of P waves from Nevada sources. J. Phys. Earth, **21**, 475-480.

Vidale, J.E., 1986. Complex polarization analysis of particle motion. Bull. Seism. Soc. Am., **76**, 5, 1393-1405.

Whittaker, E.T., 1923. On a new method of graduation. Proceedings of the Edinburgh Mathematical Society, **41**, 63-75.

Yokota, T., S. Zhou, M. Mizoue, and I. Nakamura, 1981. An Automatic measurement of arrival time of seismic waves and its application to an on-line processing system. Bull. Earthq. Res. Inst., **55**, 449-484, (in Japanese with English abstract).

Yoshii, T., 1982. Variations in seismic velocity, Earthquake prediction technique, Their application in Japan. University of Tokyo Press, 89-102.