Interpretation of Broad-band Seismograms

CHENG TONG

March 1996

A thesis submitted for the degree of Doctor of Philosophy
of the Australian National University

To my husband, Chi Ma

献给我的先生马驰

STATEMENT

This thesis is an account of my research undertaken during the period October 1992 to March 1996, while I was a full-time student in the Research School of Earth Sciences at the Australian National University.

Except as otherwise indicated in the text, the work described is my own.

This thesis has never been submitted to another university or institution. Much of the material contained in this thesis has, however, been published or is submitted for publication. (The correspondence of published papers or manuscripts with chapters is given in chapter 2.)

Cheng Tong

Cheng Tong
Canberra
March 1996

ACKNOWLEDGMENTS

First I would like to express my heartfelt thanks to my supervisor, Professor Brian Kennett, who gave me great help and willing support all the time, offering sound guidance and encouragement, and making critical and detailed review of each chapter.

I also wish to thank Dr Oli Gudmundsson who has been a much valued teacher and adviser during my study. I have enjoyed and benefited a lot from the conversations and discussions with Dr Malcolm Sambridge and Dr Phil Cummins. I am grateful to Mrs Jan Weekes not only for sharing her many years' experience in the reading of seismograms, but also for being an invaluable friend who introduced me to the Australian culture.

During the course of my study at the ANU, there have been many other people who have contributed their expertise and time, and I am greatly indebted to them. Special thanks to the staff and students of the Research School of Earth Sciences for their assistance.

Finally, I express my deepest appreciation to my parents, Xuan-Zhen Tong and Zong-Qiong Zhang, for being a never ending source of support and encouragement.

ABSTRACT

The vector recordings of surface motion achieved by broad-band seismograms contain a wealth of information about the seismic source and structure of the Earth. This information has not yet been fully utilised. This thesis is concerned with the interpretation of three-component broad-band seismograms and describes two distinctive studies of seismogram interpretation on both structural and source interpretation.

The first study presents a structural study exploiting observations of shear-wave splitting on broad-band seismograms. It provides the first clear evidence for seismic anisotropy from refracted S waves travelling in the upper mantle under Northern Territory of Australia.

The major part of this thesis lies in the second project. An automatic seismogram interpretation system is developed for seismic source interpretation. This work represents the first attempt to develop an intelligent tool which can help to extract the useful information provided by broad-band seismograms. The real-time processing capability of this tool has particular significance in the context of monitoring a comprehensive (nuclear) test-ban treaty. At the first stage of this work, a structural pattern recognition technique is applied and a seismogram is analysed like a stream of language. A context-sensitive analysis provides a good estimate of the local dominant frequency for a seismogram. Seismic phase detection is then made adaptive to the local frequency, and five parameters are extracted to characterise the waveform structure of a detected phase. At the second stage of this project, the process in the first stage is extended to a linked analysis of the three components (vertical, east-west, north-south) of a vector seismic record. Techniques are constructed for classification of the physical character of seismic arrivals, so that it is possible to identify the propagation path for the particular form of seismic energy. At the third stage, the simultaneous association of a number of extracted phases is investigated using an expert system technique which exploits codified seismological expertise to test likely assumptions about phase character and hence epicentral distance and depth. Some hypotheses about the nature of the event will be rejected as implausible, and for the remainder an assessment is given of the likelihood of the interpretation based on the fit to the character of all available information. Finally, the whole processing system is tested on a wide range of events recorded on multiple broad-band stations with good and stable results.

CONTENTS

Part one

Introduction

1

General Introduction

1.1 BACKGROUND

The history of seismology is a history of seismogram interpretation. Recordings of ground motion as a function of time, or seismograms, provide the basic data of seismological research. Our knowledge of the velocity structure of the Earth and of the various types of seismic sources is the result of interpreting seismograms. When an earthquake or underground explosion occurs, its waves travel through the body of the Earth and are recorded on seismograms at earthquake observatories. Thus, every wiggle on the seismogram has significance and contains information on the Earth structure through which they have passed, and the information about the source. Seismologists strive to extract all possible information from the seismogram by understanding each wiggle.

There are basically two aspects of work on the interpretation of seismograms.

First of all, seismological observations, seismograms, offer what are undeniably the best opportunities for resolving the detailed interior structure of the Earth. This is because seismic waves have the shortest wavelengthes of any "geophysical wave", and the physics that governs them localises their sensitivity spatially and temporally to the precise path traveled by the energy. This kind of interpretation work which leads to a better understanding of the Earth structure can be called "structural interpretation".

On the other hand, the basic character of seismograms depends strongly on the epicentral distance, which can be used to locate seismic sources. A seismogram at any particular distance will record the corresponding time sequence of arrivals, although source radiation and depth differences may make seismograms at the same distance appear dissimilar. The location of seismic events in space and time is an important part of seismological practise but has particularly significance in the context of monitoring a comprehensive test-ban treaty. Seismology provides a remote-sensing technique for monitoring nuclear testing, because underground explosions produce seismic waves that can be detected at great distances. The interpretation work which leads to the information about seismic source, location and mechanism, can be called "source interpretation".

The history of advances in seismology is one of alternating progress in describing the properties of the source or in improving models of Earth structure. In order to understand the structure of the Earth we require information on the location and characteristics of a source. However, such information is dependent on having a model of Earth structure. The development of seismology has involved a continuous interaction between the classes of "structural " and "source" interpretation.

As improved estimates of seismic travel times developed, earthquake locations were improved and velocity models for the Earth could be produced. By the time of the Jeffreys and Bullen tables (1940) the main features of the seismic velocity distribution were known and a reasonable density model had been developed by Bullen. Improved instrumentation then led to observations of the Earth's free oscillations and refinement of Earth structure including the beginnings of mapping 3-D structure. At the same time free oscillation studies changed the character of work on seismic sources through the introduction of the moment tensor representation.

A major impetus for further change came with the introduction of high-fidelity broad-band recording using feed-back seismometers. Early work on broad-band system by Galitsin and then Kirnos in Russia, where microseismic noise is relatively weak, were limited by signal amplification. The new instrumentation – with the availability of portable systems – has broken down the separation imposed by photographic recording into high-frequency body wave seismology and low-frequency work on surface waves and free oscillations. The force-feedback strategy greatly extends the bandwidth and linearity of a seismometer, and allows retrieval of much more waveform information.

Using the new instrumentation, people can now exploit waveform information in the frequency ranges which were excluded from earlier recordings, e.g., medium frequency S waves. The new observations also bring with them the need for new analysis techniques to cope with the changing appearance of the record as later phases arrive, hence the need for adaptive techniques. Broad-band recording has only been readily available for 12 years and common in the last 5 years – so there has not been very long to change the style of approach. Since many earlier techniques work quite well with the new records.

1.2 THESIS SCOPE AND APPROACHES

The vector recordings of surface motion achieved by broad-band seismograms contain a wealth of information about the seismic source and structure of the earth. This information has not yet been fully utilised. Most current methods for seismogram analysis work with each component of the record separately, and because of the uncertainty in identifying later phases in seismograms only a number of easily identified phases are routinely used. For example, source mechanisms are currently determined by working with the vertical component records of the initial P phases and the SH component of the S arrivals, and methods of event location at teleseismic distances depend on the use of primarily P wave observations from a large number of stations. In this thesis, we try to use the full 3-component records more effectively, and apply techniques to the full vector wavefield.

Nowadays, with fast computers with large memories, the precision as well as speed of seismogram interpretation can be improved by combining new computational technologies. In this thesis, we demonstrate that a number of computer techniques can contribute greatly to seismology research in various ways.

This thesis is concerned with the interpretation of 3-component broad-band seismograms and describes two distinctive studies of seismogram interpretation which belong to the two categories of structural and source interpretation respectively.

Part two, chapter 3, presents a structural study exploiting observations of shear-wave splitting on broad-band seismograms of refracted S wave travelling in the upper mantle under Northern Territory of Australia. The use of broad-band recording enables the use of intermediate period S waves which would not be recorded very well by conventional short- or long- period instruments.

In part three, chapter 4 through to chapter 8, an automatic seismogram interpretation system is described for seismic source interpretation. This represents the first attempt to develop an intelligent tool which can help to extract the wealth of information provided by broad-band seismograms. The real-time processing capability of this tool has particular significance in the context of monitoring a comprehensive (nuclear) test-ban treaty.

In part two, analysis has been applied in order to obtain a clearer understanding of the detailed structure in the Earth, and computational techniques are employed to process the data and improve the quantitative precision of the interpretation. While in part three, Artificial Intelligence techniques (in particular, Pattern Recognition and Expert System) are applied to implement the automatic seismogram interpretation system. This system was initially planed to imitate an eye of a seismologist. To our surprise, the results from the current system show that in many aspects this system performs more efficiently than a human expert.

2

Thesis overview

This chapter provides a brief overview of the contents of the subsequent chapters and the interactions of the various techniques for processing broad-band seismograms.

2.1 AN OBSERVATION OF SHEAR-WAVE SPLITTING ON BROAD-BAND SEISMOGRAMS

In part two (chapter 3), full 3-component broad-band recordings from Warramunga (WRA) in the Northern Territory of Australia have been studied from the vector-field point of view. By processing seismograms to give out clear shear wave polarisation, the first clear evidence for seismic anisotropy from refracted waves travelling in the upper mantle (down to 700 km) has been provided.

Eleven polarisation analyses of waves returned from the transition zone yield an average time shift of 2.3 seconds with the fast direction scattered about the transverse direction. Nine polarisation measurements of waves returned from the top of the lower mantle yield an average time shift of 1.7 seconds, again with the fast direction near the transverse. No appreciable time differences are observed between the radial and transverse polarisations for paths refracted within the lithospheric lid. Because the observations of shear-wave splitting in waves passing through the low-velocity zone, transition zone, and the top of the lower mantle are not coherent in their absolute polarisation, the cause cannot lie in azimuthal anisotropy at shallow depths under the WRA station. The most plausible explanation is transverse anisotropy in shear within the low-velocity zone under the unusually thick mantle "lid" under Australia. A possible contribution may come from transverse anisotropy in $\beta - spinel$ olivine at the top of the upper-mantle transition zone. Evenly distributed transverse anisotropy in the 200 km thick layer below the lithosphere down to the transition zone must be of order 1% with a slow vertical polarisation to explain the data. This is a reasonable level of anisotropy and the polarisation is consistent with lateral flow. The absence of discernible splitting in waves contained within the mantle "lid" implies that the "lid" does not possess a significant component of transverse anisotropy.

This was a cooperative project carried out with O. Gudmundsson and B. L. N. Kennett and my major contribution was the processing and analysis of seismograms. The material in this chapter has been published as "Shear wave splitting in refracted waves returned from the upper mantle transition zone beneath northern Australia" in the *Journal of Geophysical Research* (**99**, 15783-15797).

2.2 AN AUTOMATIC INTERPRETATION SYSTEM FOR BROAD-BAND SEISMOGRAMS

In part three, chapter 4 through to chapter 8, an automatic seismogram interpretation system is described for seismic source interpretation.

2.2.1 Phase detection and phase appearance characterisation

To process broad-band seismogram efficiently, an important step is to automate phase-detection and phase-characterisation. In chapter 4, an automatic analyzer which characterises, as well as detects, seismic phases has been constructed using concepts from work in artificial intelligence. A seismogram is recognised as a hierarchical structure of subunits and the features characterising a phase are extracted via a structural analysis. The automatic analyzer detects the onsets of phases, separates the relevant waveform segments, fits phase structures to standard wavelets, and generates a set of five parameters ("a quintuple") to characterise a detected phase. The whole analysis system requires just a few input parameters and can be accomplished swiftly enough to allow continuous real time operation.

The analysis procedure is adaptive, including updating of the estimate of the local frequency so that it can be applied to broad-band records with a wide range of frequency content. By coupling the analysis procedure with real time filtering, very good results for phase detection and characterisation can be achieved for weak distant events. The characterisation of phase segments also provides useful information for the comparison of seismograms.

The material in this chapter has been published in the *Geophysical Journal International* (**123**, 937-947) in December 1995.

2.2.2 Towards later seismic phase identification

The location of seismic events can be improved if accurate picks can be assigned for later seismic phases, which requires both the detection of an arrival and the recognition of its character. Such phase identifications are particularly valuable if they can be provided in real-time as the seismic disturbance passes across a broad-band seismic recording station.

Chapter 5 shows that a simple but promising scheme for characterising arrivals can be constructed by analysing the energy content of the seismic trace as a function of time. Such an approach can be used to detect arrivals by using a method comparing the short-term average energy to a long-term average, with averaging windows that are adaptive to the local frequency of the seismic disturbance. The phase detector can be tuned to different classes of arrivals by utilising three-component records. By comparing the energy on the vertical component of motion to that in the horizontal plane, it is possible to start to separate P and S arrivals. Phase assignments can be refined by the use of adaptive filtering and by including polarisation information.

With an estimate of the azimuth of propagation it is possible to use approximate projection methods which attempt to compensate for the influence of the free surface, since the surface corrections are not a strong function of slowness for teleseismic arrivals. By this means an instantaneous estimate can be made of the relative contributions of P, SV and SH arrivals which can be very helpful in determining the phase assignment for a particular arrival.

This material has been published as "Towards the identification of later seismic phases" in the *Geophysical Journal International* (**123**, 948-958).

2.2.3 Automatic seismic event recognition and phase identification

The goal of automated analysis of seismic records is to provide a means of recognising the pattern of seismic arrivals associated with the presence of a seismic event in real time accompanied by an identification of the individual phases. For a single station such a real-time analysis can be used to provide a preliminary estimation of the location of the event.

In chapter 6, the seismic phase analyser developed in the previous two chapters has been used as the basis of a seismic interpretation system for three-component broad-band records. The inputs are the values for the phase features produced by the phase analyser, the combinations of these features are investigated using a classification system which exploits expert information to test likely assumptions about phase character and hence epicentral distance and depth. Some hypotheses about the nature of the event will be rejected as implausible and for the remainder an assessment is given of the likelihood of the interpretation based on the fit to the character of all available information.

This event recognition procedure provides an effective and feasible means of characterising information between hundreds of different possible classes of patterns, even when the observation is incomplete. The procedure based on "assumption trees" provides a useful tool for classification problems in which a number of factors have to be identified. The control set of expert knowledge used in testing hypotheses is maintained separately from the computational algorithm used in the assumption search; in consequence the information base can be readily updated.

A paper based on this chapter was submitted to the *Bulletin of the Seismological Society of America* in November 1995.

2.2.4 Automatic seismic event recognition using multiple broad-band stations

The automated analysis of seismograms for a single seismogram is of greatest utility when applied to a network of stations. Also interesting is the robustness of the event recognition procedures using similar data.

These two issues have been addressed in chapter 7 by applying the automated analysis procedure described in the previous chapter to multiple recordings of the same event at different portable broad-band stations in Australia. The stations cover a significant span in epicentral distance. The use of a number of different recordings for the same events provides a test of the consistency achievable with the automated analysis procedure.

2.2.5 A review of the automatic interpretation system

As a review of the automatic interpretation system developed from chapter 4 through to chapter 6, chapter 8 gives a structural description of the whole work from the view point of Artificial Intelligence. It gives a syntactic description of a seismogram, and explains how the AI techniques were applied in the implementation of the interpretation system.

Part two

An observation of shear-wave splitting on broad-band seismograms

3

Seismic anisotropy in the mantle beneath northern Australia

3.1 INTRODUCTION

Most seismic body-wave studies of upper mantle structure have been based on the analysis of P wave arrivals on vertical component instruments or on the use of naturally polarised SH waves on the transverse component to the path from the source. Such natural polarisation is not easy to find and has limited the extent of S body wave studies of the upper mantle. Very little work has been done with SV wave records from the radial component because the waveforms are usually complicated due to coupling between the SV arrivals and P waves in the shallow structure near the recording station.

The Proterozoic rocks of the Tennant Creek inlier in the Northern Territory of Australia lie at an appropriate distance range to investigate upper mantle structure using sources in the major earthquake belt running through Indonesia and New Guinea (Figure 3.1).

The granites of the Warramunga (WRA) group provides excellent recording sites with high-velocity material at or very near the surface. Drilling at the site indicates that the weathered layer is less than 30 m thick or much thinner than one wavelength of the wave we record on the WRA broad-band instrument. The presence of high P

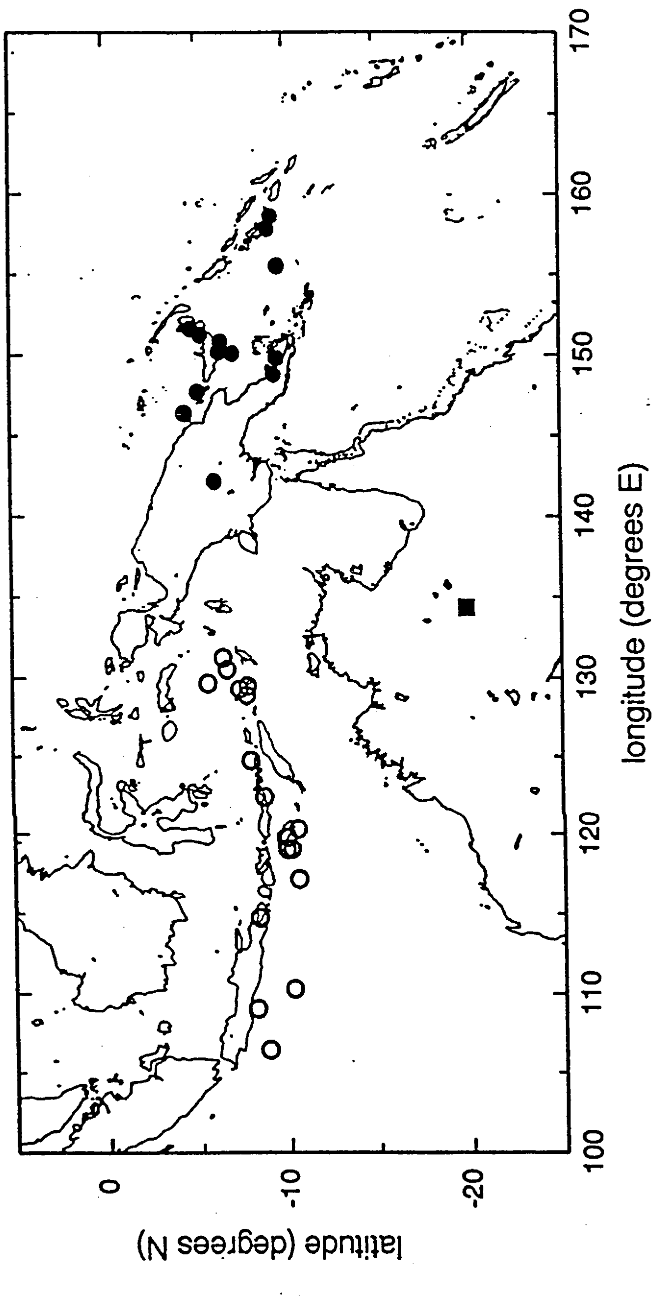


Fig. 3.1. The locations of the events used in the analysis of shear-wave splitting, and the WRA broad-band instrument in the Northern Territory of Australia. A distinction is made between events in the Indonesian region (open circles) and New Guinea and Solomon Islands region (solid circles).

wave velocities at the surface (around 6 km/s) minimises the contamination of the SV wavefield. In consequence, it is possible to exploit SV recordings on both the radial and vertical components of broad-band instruments.

A three-component set of broad-band seismometers (Guralp CMG-3) has been operated since 1988 at the Warramunga array (WRA) in the Northern Territory of Australia, 35 km to the southeast of the small mining town of Tennant Creek. The site is installed on granite and has yielded very good data for both P and S waves propagating through the upper mantle transition zone. Kennett, Gudmundsson and Tong (1994) have presented record sections for P, SV and SH propagation and have illustrated the value of the broad-band recording by deriving both S and P velocity models from the same set of events. The SV and SH wave record sections are very consistent in their general features, but when they are examined in detail, a perceptible time advance may be observed for many of the SH wave arrivals relative to their SV wave counterparts.

The object of this work is to quantify the extent of shear-wave splitting for these S waves refracted back from the upper mantle, and to explain it in terms of the location and orientation of anisotropy. It should be noted that the geometry of the propagation in the present study is very different from measurements of shear-wave splitting for phases such as SKS (Silver and Chan 1991, Vinnik et al. 1992) or ScS (Fukao 1984, Ando 1984). In the upper mantle transition zone the propagation paths are near horizontal so that the S waves are affected by vertical anisotropy rather than horizontal anisotropy. Only in the passage through the uppermost mantle (the mantle "lid") are the propagation paths close to vertical and more comparable to the usual SKS configuration. The geometry of the propagation paths (see Figure 3.2) enables us to put direct constraints on the likely location of anisotropy.

3.2 ANISOTROPY IN THE EARTH'S MANTLE

Evidence for anisotropy of varied geometry in the earth's mantle continues to grow. In general, such anisotropy would need to be described by 21 elastic moduli, but

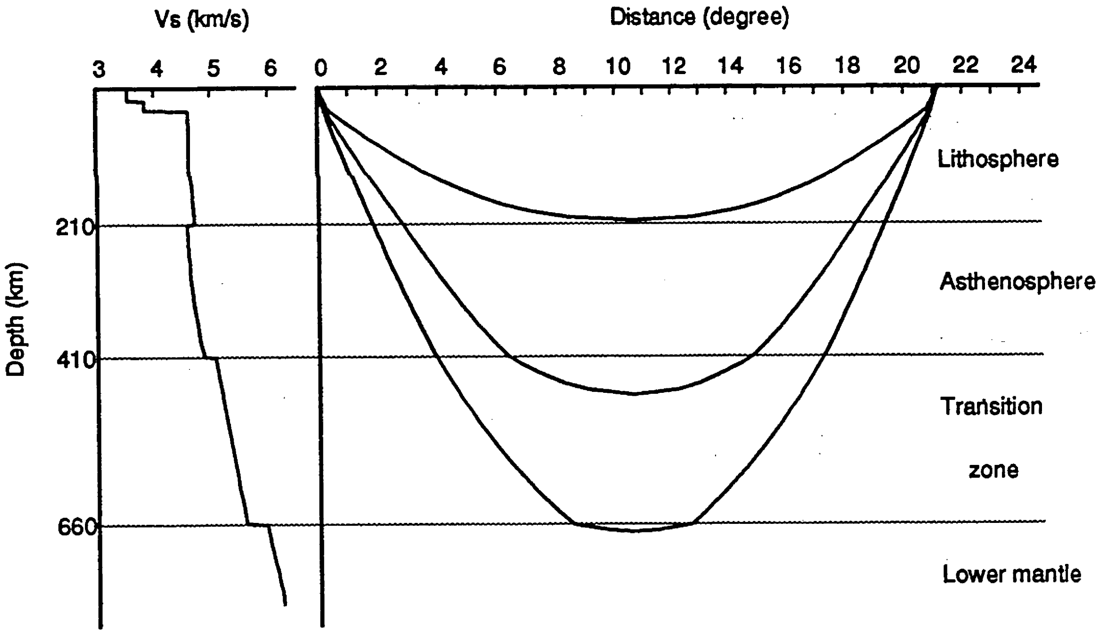


Fig. 3.2. The shear-velocity structure under northern Australia and a schematic drawing of ray paths in this study and their relation of the main structural units of the upper mantle.

the available configurations for seismic observations have tended to limit the class of models which have been considered in the upper mantle. In particular, because S body wave studies have focused on waves traveling close to vertical, attention has been concentrated on the variation of the S wave speeds for different polarisations in the horizontal plane.

It is useful to make a distinction between two basic geometries of anisotropy. Transverse anisotropy, on one hand, is a term which has been applied to the case where a vertical symmetry axis is assumed and a difference found between the elastic properties of vertically polarised waves and horizontally polarised waves. Azimuthal anisotropy, on the other hand, is used to refer to observations of a directional dependence on the rotation angle in a horizontal plane. Of course, anisotropy can manifest itself in a much more complicated fashion than either of those two cases; a mix of the two possibly with a complicated functional dependence on orientation. But, seismic observations are generally limited in their directional resolution. Therefore, interpretation is usually limited to one of those two idealised or simplistic geometries.

Two manifestations of anisotropy in seismic data have long since been recognised. Firstly, Pn velocities from refraction profiles in the oceans revealed in the sixties a directional dependence consistent with 5% azimuthal anisotropy in compression in the subcrustal oceanic lithosphere (Hess 1964, Raitt et al. 1969, Shearer and Orcutt 1986). Similar observations have subsequently been made in continental settings (Fuchs 1983, Hearn 1984). Secondly, the Love and Rayleigh wave discrepancy was recognised in the early sixties (Anderson 1961) and explained in terms of transverse anisotropy in shear of the order of a few percent in a significant volume near the top of the mantle (e.g. Yu and Mitchell 1979, Journet and Jobert 1982). Dziewonski and Anderson (1981) produced a spherically symmetric global model, PREM, on the basis of normal modes, surface-wave dispersion data, and summary travel-time data which is transversely anisotropic by 3% between depths of 25 and 220 km. Regan and Anderson (1984) found regionalised models for age provinces in the oceans with varied degrees of transverse anisotropy. They also found this anisotropy to be confined to the asthenosphere, at least in the oceans.

More recently, measurements of shear-wave birefringence or splitting of near vertically travelling body waves (S, ScS, SKS) have revealed azimuthal anisotropy in the back-arc wedge of subduction zones (Ando et al. 1983, Fukao 1984, Bowman and Ando 1987) and under the continents (Vinnik et al. 1984, Silver and Chan 1991, Vinnik et al. 1992). The time separation of the two split waves is generally of order 1 second which implies anisotropy of the order of 3% if it is evenly distributed throughout a 150 km thick lithosphere. The total set of stations for which SKS-splitting measurements exist now spans most continents fairly densely, sufficiently so to reveal some geometrical patterns and a general correlation with surface tectonics (Vinnik et al. 1992).

The last decade has also seen surface-wave tomographers attempt to map the lateral variation of transverse anisotropy (e.g. Nataf et al. 1986) and azimuthal anisotropy (e.g. Tanimoto and Anderson 1984, Montagner and Tanimoto 1991). Although tentatively interpreted by the authors, those studies were successful to some degree. Nataf et al. (1986) found a correlation of regions with a fast vertical polarisation with the locations of mid-ocean ridges. Assuming the crystal alignment which flow would induce in α-olivine at high temperature this is consistent with vertical flow. Tanimoto and Anderson (1984) find a high degree of correlation between their map of the orientation of azimuthal anisotropy with the lateral flow pattern predicted at the relevant depth by Hager and O'Connell (1979). Again, this is consistent with the alignment of olivine crystals by flow at high temperature.

The interpretation of in situ seismic observations of anisotropy has been greatly aided by work done under the microscope in the labs of petrophysicists. Petrofabric studies of ophiolites (Christensen and Salisbury 1979) and mantle xenoliths (Nicholas et al. 1971) together with studies of fabric and texture induced by forced deformation (Nicolas et al. 1973) have revealed the distribution of the orientations of anisotropic crystals and its relation to strain. The mobility of olivine crystals as a function of temperature is known on the basis of the work of Goetze and Kohlstedt (1973). Nicolas and Christensen (1987) review the state-of-the art in this field. Mainprice and Silver (1993) summarise much of the relevant information from petrophysics to the problem of interpreting seismic anisotropy observations (with particular reference to SKS splitting results). Anderson (1989) gives a comprehensive review of anisotropy in general.

It is now clear that at least parts of the upper mantle are significantly anisotropic and that this anisotropy can have a significant effect on wave propagation in diverse ways. For simple models of anisotropy the effects on travel times (Pn), phase velocities (surface waves) and shear-wave splitting in simple models is well understood. Recent developments in wave theory in anisotropic media implies that anisotropic gradients can cause significant waveform distortions through coupling of the splitting waves and ray bending through anisotropic regions can result in surprising levels of multipathing (Thomson et al. 1992). This implies that elaborate analysis involving forward waveform modelling will be necessary.

More fundamentally, most experiments sensitive to anisotropy have limited directional resolution. SKS-waves are blind to the transverse component of anisotropy. Furthermore, they provide no depth resolution. Love- and Rayleigh waves taken together do potentially provide 3-dimensional directional resolution, but strong tradeoffs exist with heterogeneity (Anderson and Dziewonski 1982).

Measurements of shear-wave splitting in more complicated geometries than the SKS geometry should be explored since they can provide some of the 3-dimensional directional coverage needed to resolve the full anisotropy of the mantle, although such measurements are admittedly difficult. This is what we attempt to do in the present study, motivated by observations of a few seismograms of shear waves refracted through the upper mantle which appear to be naturally rotated to separate two split waves.

It must be acknowledged that the greatest difficulty with measurements of shear-wave splitting is potential contamination by phase conversions giving rise to precursors to the true shear wave. This is much more likely to occur on the radial component given that the earth is to zeroth order a spherically symmetric body. Thus, when shear waves on the radial component appear to be advanced relative to shear waves on the transverse component scattering must be considered as a potential cause. In the present study virtually all the observations are of advanced shear waves on the transverse component or a fast polarisation near the transverse direction.

3.3 SEISMIC DATA SELECTION AND ANALYSIS PROCEDURE

The events selected for shear-wave splitting analysis span the distance range from 12° to 30° from the Indonesian and New Guinea earthquake belt and have been recorded at WRA. We have selected shallow events with body wave magnitudes between 4.9 and 5.5 for which there is good signal strength for S compared to the P wave coda. With this type of event we try to secure as simple a waveform as possible, and also minimise the influence of subduction zone structures near the source.

For epicentral distances out to 2100 km the onset of the SV and SH waveforms show high frequencies associated with arrivals propagating within a thick lid extending to 210 km depth. At larger distances the onset of S, representing waves refracted back from beneath the 410 and 660 km discontinuities, has a much lower frequency content (0.25-0.3 Hz). Such intermediate period arrivals also characterise the S waves returned from the 410 and 660 km discontinuities at shorter distances. The frequency differences have been interpreted by Gudmundsson, Kennett and Goody (1993) in terms of a zone of enhanced attenuation below 210 km and above the 410 km transition.

In all we have looked at about 120 seismograms which satisfy the above selection criteria, recorded at WRA between December 1989 through May 1992. In many cases no shear waves are discernible from the P-wave coda. In other cases the signal-to-noise ratio is insufficiently high to obtain a reliable measurement of shear-wave splitting. In the end we have 29 useful events from which we obtain 34 measurements of splitting. The geographical configuration of the WRA recording station and the earthquakes which provide useful shear-wave splitting measurements is displayed in Figure 3.1, the locations of the events are also presented in Table 3.1. The turning points in the mantle corresponding to these propagation paths lie mostly beneath the northern margin of the Australian continent. Studies of both short period arrivals (Dey et al. 1993) and broad-band records (Kennett et al. 1993) have indicated the presence of regional variations in mantle structure under Northern Australia, and so we have divided the source region into two parts: a) along the Flores Arc, Indonesia, with propagation under northwestern Australia, and b) in New Guinea, with paths to the NNE of the array. The events in each class are identified by open or closed symbols in Figure 3.1.

For each event, a set of time windows is selected to span the expected time for the arrivals for different classes of propagation path. These windows are derived from the interpretation of composite record sections using all available records for both SV and SH arrivals (Kennett, Gudmundsson and Tong 1993). The process is illustrated in Figure 3.3 for an event in the Flores Arc at 17.25° (Figure 3.3a) and an event south of Java at 29.18° (Figure 3.3b).

The two time windows used for the analysis of the seismogram in Figure 3.3a correspond to S-wave propagation in the mantle "lid" extending down to 210 km and to the arrivals from the mantle transition zone below 410 km depth. The seismogram has been bandpass filtered between 2 and 20 seconds periods with a causal, 4-pole, Butterworth filter in order to remove some of the incoherent, high frequency, scattered energy associated with the shallower branch. The time window used for the trace in Figure 3.3b corresponds to waves turned at the top of the lower mantle. The widths of the time windows are chosen so that they span of the order of four periods in the waveform (Note that correlations are computed over half that time length).

For each of the selected time windows a correlation analysis was performed fol-

Table 3.1. *Hypocentres and magnitudes of events for which shear-wave splitting was successfully measured*

Event	Date	Origin Time UT	Latitude deg	Longitude deg	Depth km	Magnitude
1	Jan. 5 1990	10 10:21.80	-8.800	106.442	29.0	5.3
2	Feb. 2 1990	07 58:14.07	-10.233	110.290	45.8	5.8
3	Feb. 9 1990	21 33:40.39	-9.886	119.050	41.7	4.6
4	Feb. 10 1990	13 12:14.20	-5.263	151.271	10.0	5.3
5	Feb. 10 1990	13 58:07.95	-10.506	120.304	33.0	4.6
6	Feb. 26 1990	18 13:59.80	-9.577	149.798	33.0	5.2
7	March 3 1990	17 11:39.72	-5.561	129.633	33.7	5.1
8	April 4 1990	19 51:48.95	-4.738	151.645	31.2	4.9
9	April 13 1990	22 46:55.32	-6.638	130.513	17.7	5.2
10	April 25 1990	15 32:21.70	-7.077	150.112	23.0	5.4
11	May 7 1990	14 52:51.60	-9.648	155.537	25.0	5.1
12	May 15 1990	20 35:34.60	-7.797	129.467	57.0	4.8
13	May 21 1990	13 24:36.60	-8.137	109.043	28.0	5.5
14	June 1 1990	04 45:49.85	-5.119	147.722	33.0	5.4
15	June 6 1990	15 04:50.90	-6.422	131.240	62.0	5.0
16	July 17 1990	22 05:55.02	-7.746	128.893	33.0	4.7
17	Sept. 29 1990	11 16:10.90	-8.701	122.385	33.0	5.0
18	Oct. 18 1990	03 06:52.30	-7.348	129.279	48.0	5.3
19	Dec. 16 1990	20 19:48.10	-6.002	142.167	33.0	5.4
20	May 17 1991	06 37:47.00	-9.941	119.780	21.0	5.2
21	May 7 1991	17 07:04.00	-7.954	124.716	30.0	5.3
22	Sept. 26 1991	09 14:50.40	-9.283	158.620	26.0	5.2
23	Oct. 7 1991	19 56:21.10	-10.553	117.161	50.0	5.2
24	Dec. 27 1991	17 14:30.70	-9.093	157.840	33.0	5.2
25	Dec. 28 1991	03 40:31.40	-6.424	150.840	24.0	5.4
26	Dec. 28 1991	11 05:20.30	-6.325	150.189	33.0	5.4
27	Jan. 24 1992	12 09:24.10	-9.402	148.795	33.0	5.4
28	March 29 1992	15 12:30.60	-10.161	119.063	33.0	5.2
29	March 30 1992	13 42:19.10	-8.307	114.752	33.0	5.2

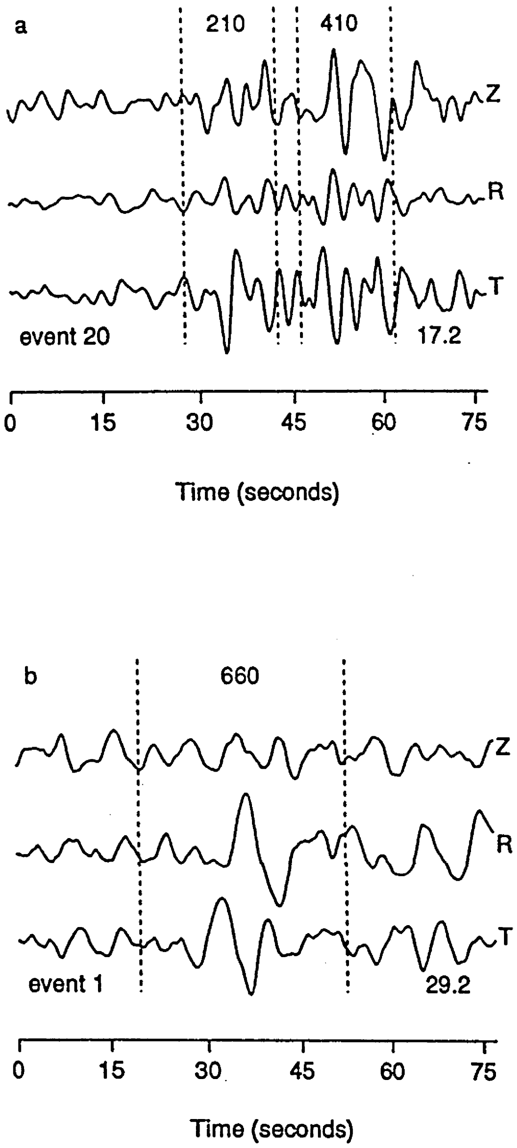


Fig. 3.3. Three-component shear wave records for a pair of events along the Indonesian Arc recorded at WRA. (a) Event 20 in Table 1 at a distance of 17.2°. (b) Event 1 in Table 1 at a distance of 29.2°. The time windows used for polarization analysis of shear-wave splitting are marked for both the "lid" arrivals (210) and the phases associated with the 410 km transition (410) in Fig. 3.3 (a) and for phases turning under the 660 km discontinuity (660) in Fig. 3.3 (b). The traces in Fig. 3.3 (a) have been filtered with a bandpass from 0.05 to 0.5 Hz.

lowing a similar procedure to Bowman and Ando (1987) demonstrated in Figures 3.4, 3.5, and 3.6 for the three time windows shown in Figure 3.3.

We start with the appropriate time window in the R/T (radial/transverse) coordinate system (shown in the upper left-hand corner of Figures 3.4-3.6). Due to significant deviations from the vertical of the propagation direction incident at the recording site, our data are significantly affected by the free surface of the earth. We used techniques developed by Jepsen and Kennett (1991) and Kennett (1992) based on a model of a single plane-wave component for each body-wave phase in order to correct for the effect of the free surface. Thus, we generate the SV and SH traces in the upper middle of Figures 3.4-3.6. When the velocity structure near the surface is well known this correction should be able to remove most of the phase distortion associated with the interaction of SV waves with the free surface. This correction is most significant, and hence uncertain, for waves turning from a shallow depth (the "lid").

We then assume that the corrected SV and SH seismograms can be represented by:

$$\begin{aligned} SV(t) &= A \cdot \cos\phi \cdot f(t) + B \cdot \sin\phi \cdot f(t+\delta t) \\ SH(t) &= A \cdot \sin\phi \cdot f(t) - B \cdot \cos\phi \cdot f(t+\delta t) \end{aligned} \qquad (3.1)$$

which represents a superposition of two orthogonal shear waves rotated by an angle ϕ out of the R/T polarisation reference and separated in time by δt. The amplitudes of the two waves are assumed to be arbitrary, but their waveforms are assumed to be identical. This is admittedly a simplistic model in light of the rich phenomena that can arise due to anisotropy with nontrivial geometry and anisotropic gradients, such as multipathing suffered by waves bending through an anisotropic region and waveform distortions due to coupling of the splitting waves in an anisotropic gradient zone (Thomson et al. 1992). Given the above model we estimate the parameters ϕ and δt using a cross-correlation scheme. Given values for ϕ and δt we decompose the SV and SH traces into two rerotated and time-shifted traces which we cross-

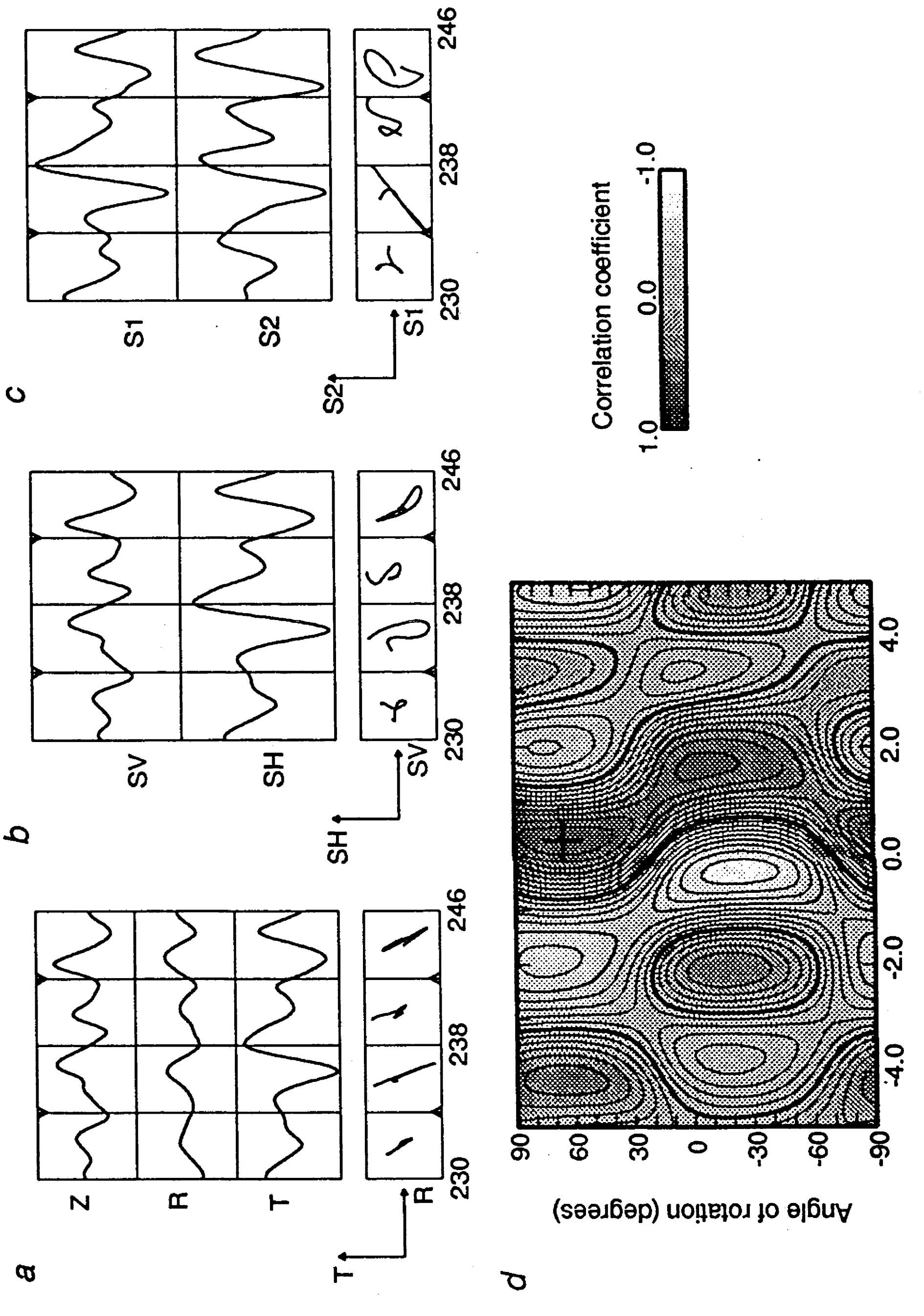


Fig. 3.4. Correlation analysis to determine the time advance and rerotation angle from the radial/transverse geometry for the "lid" arrivals in Fig. 3.3 (a). a) Vertical (Z), radial (R) and tangential (T) traces. b) SV and SH traces incorporating compensation for the influence of the free surface. c) Rotated and time shifted traces to optimize their cross correlation. d) Correlation as a function of time shift and rerotation angle, the cross indicates the maximum correlation found.

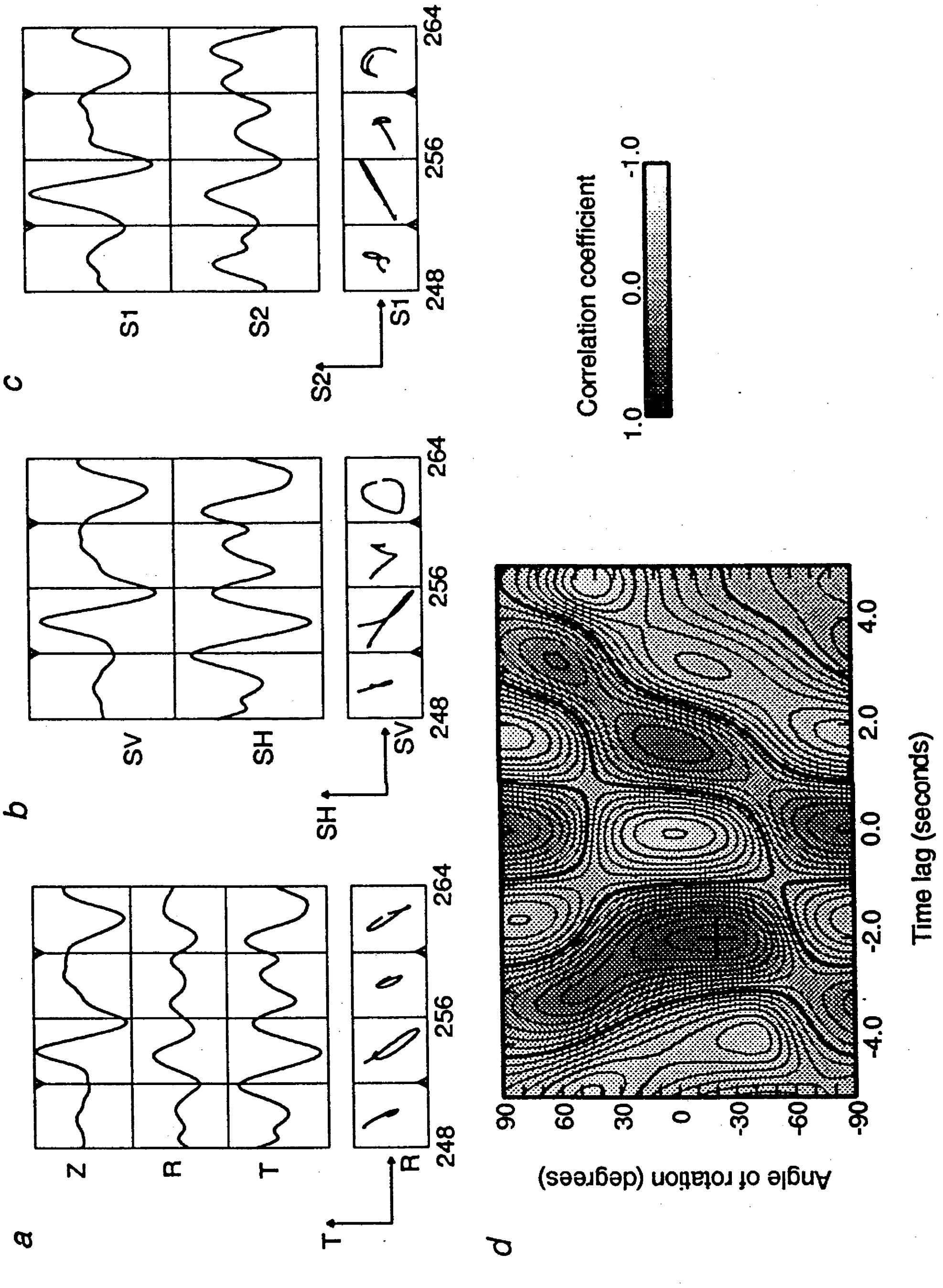


Fig. 3.5. Correlation analysis to determine the time advance and rerotation angle from the radial/transverse geometry for the transition zone arrivals in Fig. 3.3 (a). For explanation see Figure 3.4.

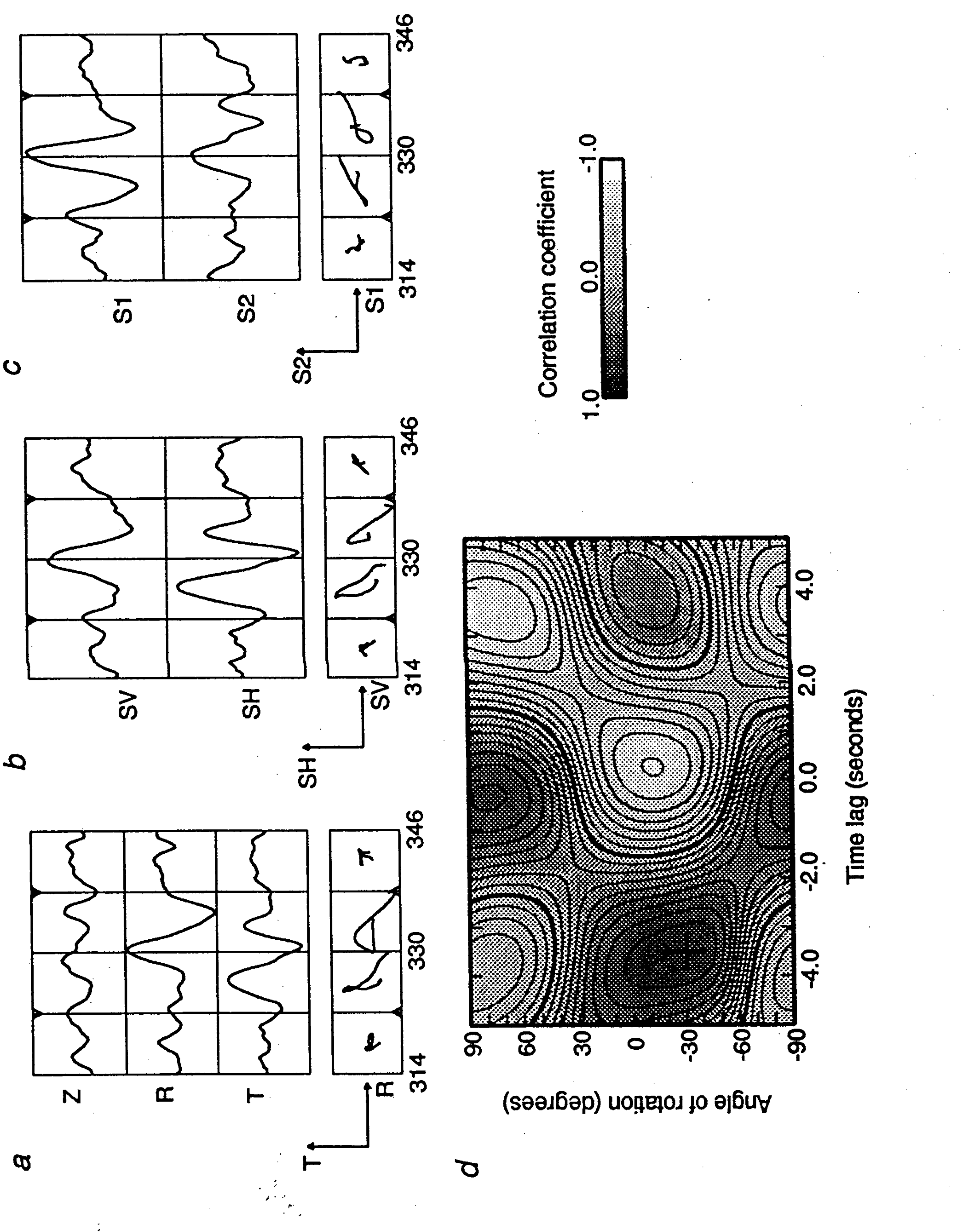


Fig. 3.6. Correlation analysis to determine the time advance and rerotation angle from the radial/transverse geometry for the upper mantle arrivals in Fig. 3.3 (b). For explanation see Figure 3.4.

correlate. Should our choice of ϕ and δt be correct the processed traces would correlate perfectly according to our model. We span a range of values for ϕ and δt and choose the combination which results in the highest positive or negative correlation. In each case the correlation is computed for the centre half of the time windows shown in Figures 3.4-3.6. The correlation is shown as a function of ϕ and δt in the lower half of Figures 3.4-3.6 in a contour diagram. The chosen combination of ϕ and δt is indicated by a solid cross in each figure. The rerotated and time-shifted traces for the optimal parameters are shown in the upper, right-hand corner of each figure.

It is worth noting that in the contoured correlation diagrams of Figures 3.4-3.6 there is some degree of nonuniqueness because the rotation angle, ϕ, spans half a circle. A rerotation by 90° is equivalent to interchanging the traces while changing the polarity of one. Thus, for each positive correlation peak there is a corresponding negative correlation peak at a time shift reversed in sign and a 90° shift in rotation angle. In Figure 3.4 the peak correlation is found at $\phi = 65°$, $\delta t = 0.3s$, but there is an equivalent negative peak at $\phi = -25°$, $\delta t = -0.3s$. It is only because the correlation is computed within a time window of finite length that this equivalence is not perfect in the computed correlation function. Taking this nonuniqueness into account all of the correlation functions presented in Figures 3.4-3.6 have a clear maximum at a correlation approaching 0.9. The widths of the correlation peaks vary considerably. The width along the time axis is in large part controlled by the dominant period of the waveforms. Taking the half width of the chosen correlation peak along the time axis as an uncertainty of the measurement of time shift the value would be 0.5 - 1.0 seconds in Figures 3.4 and 3.5, and 1.5 seconds in Figure 3.6. In general this number is of order 1 second. This may be a somewhat overpessimistic evaluation of uncertainty. The width of the correlation peak along the angle-of-rotation axis is primarily controlled by the rotation kernels, $\cos\phi$ and $\sin\phi$ as is evident from the fact that in general the correlation function spans one circle in the interval from $\phi = -90°$ to $\phi = 90°$. It is not obvious that we can translate the half width of the correlation peak into an absolute measure of uncertainty in this case, but it is clear that the uncertainty in the determination of the optimal rotation angle is considerable, perhaps of the order of ±20°.

3.4 RESULTS

The results of the polarisation analysis are summarised in Table 3.2. 34 measurements are presented out of about 200 attempted. A measurement was deemed useful when a clear arrival was seen in a low-passed seismogram at 2 seconds period, the signal-to-noise ratio exceeded a value of 2.0, and correlation and polarisation diagrams such as presented in Figures 3.4 - 3.6 were unambiguously interpretable (a single correlation peak of height exceeding 0.8 and a linear polarisation achieved throughout a significant portion of the main arrival). In some cases arrivals on the shallow "lid"-branch contain little energy beyond a period of 2 seconds in excess of noise (P coda), but a high level of energy around 1 Hz where scattering effects destroy coherency between the three recorded components. In such cases polarisation analysis was not undertaken. Instead, a travel-time pick was made from both the radial and the transverse components where this could be achieved with accuracy. The time shift was then defined as the timing difference between the SH wave and the SV wave. Entries in Table 3.2 with blanks entered for the rotation angle and the correlation coefficient represent measurements of this type. We argue that the inclusion of these measurements is justified because the true polarisation measurements in most cases result in a fast direction near the transverse direction (irrespective of azimuth).

In all but a few cases the fast polarisation is within 30° from the transverse direction. This, in the context of the likely uncertainty of measurement (±20°), implies that deviations from a purely SH/SV splitting geometry are not significant. A considerable scatter exists in the measurements of time delay, δt. For paths contained within the mantle "lid" (210 branch) we have 14 measurements averaging 0.0 seconds, with a standard deviation of 0.3 seconds and no systematic trend with

Table 3.2. *A summary of the results of polarisation and splitting measurements*

Event	Δ, deg	δt, s	ϕ, deg	R	branch
12	13.03	0.3	...	...	210
16	13.29	0.0	-25.0	0.89	210
18	13.51	-0.3	...	...	210
9	13.81	-0.5	...	...	210
15	13.85	-0.3	...	...	210
7	15.09	0.3	...	...	210
21	15.19	0.0	...	...	210
19	15.87	0.0	-15.0	0.89	210
17	16.13	-0.5	...	...	210
5	16.50	0.2	...	...	210
20	17.25	0.3	65.0	0.88	210
27	17.48	0.0	...	...	210
3	17.86	0.0	...	...	210
23	19.03	0.0	...	...	210
20	17.25	-2.0	-20.0	0.87	410
27	17.48	-1.3	10.0	0.86	410
28	17.69	-0.6	-10.0	0.88	410
23	19.03	-3.0	50.0	0.96	410
10	19.97	-0.9	60.0	0.92	410
26	20.53	-3.0	-15.0	0.97	410
25	20.95	-3.5	0.0	0.92	410
4	22.05	-1.9	-25.0	0.96	410
29	22.24	-1.6	30.0	0.83	410^*
11	22.88	-4.4	-15.0	0.94	410
24	25.20	-3.2	-15.0	0.95	410
6	18.15	-2.6	-35.0	0.94	660
14	19.71	-0.4	-40.0	0.93	660
29	22.24	-1.6	30.0	0.83	660^*
8	22.68	-0.9	5.0	0.94	660
2	25.10	-0.7	-20.0	0.94	660
24	25.20	-2.0	0.0	0.88	660
22	25.74	-0.3	-30.0	0.92	660
13	27.18	-3.3	30.0	0.86	660
1	29.18	-3.5	-30.0	0.89	660

Δ is epicentral distance in degrees, δt is the measured time shift in seconds, ϕ is the measured rotation angle of the slow polarisation clockwise away from the radial direction, and R is the peak correlation for each measurement. Where measurements were made directly from high frequency seismograms without polarisation analysis that is indicated by blank entries for the polarisation direction and correlation. Superscript $*$ indicates where two interfering arrivals cannot be separated while a good polarisation measurement can still be made.

epicentral distance. Clearly, no resolvable anisotropic effect is seen here. This is at odds with tentative statements made by Goody (1991) and Dey et al. (1993) about observations of timing differences between SH and SV waves on this branch based on only a few observations. Eleven measurements based on arrivals returning from the transition zone (410 branch) have an average time shift of 2.3 seconds and a standard deviation of 1.2 seconds. Nine measurements based on arrivals returned from the top of the lower mantle (660 branch) have an average of 1.7 seconds and a standard deviation of 1.2 seconds. We have a resolved anisotropic effect although the scatter among the delay-time measurements appears to be somewhat high. As was discussed in the previous section uncertainties of the order of $\pm 0.5 - 1.0$ second are anticipated, which is less than the observed scatter. Incorrect identification of phase may contribute to this scatter as well as poorly understood effects of noise in the seismograms.

Figure 3.7 summarises the results in Table 3.2 in terms of branch on the travel-time curve and epicentral distance. Again, it is clear from the figure that a significant level of anisotropy is found on the 410 and 660 branches, while no discernible anisotropic effect is found on the shallow 210 branch. The results on the 410 and 660 branches show a hint of a weak increase of δt with epicentral distance. A linear regression fit results in both cases in $d\delta t/d\Delta = 0.1$ second/degree and the standard deviation of the scatter is reduced from 1.2 seconds to 1.1 seconds.

In Figure 3.8 we summarise the results geographically. Only those events contributing a measurement on either the 410 or the 660 branch are shown. They are connected with the WRA recording site by a thin line to show the radial direction for each event. The polarisation results are shown as time bars. Each time bar is of length proportional to δt and oriented along the inferred fast polarisation direction. A scaled bar is shown in the lower left-hand corner of the figure for reference. The results on the 410 branch are plotted in grey at an arbitrary point along each ray at a distance of 30% of the total ray length from WRA. Results on the 660 branch are also plotted at an arbitrary point along each ray, halfway along the ray and drawn in

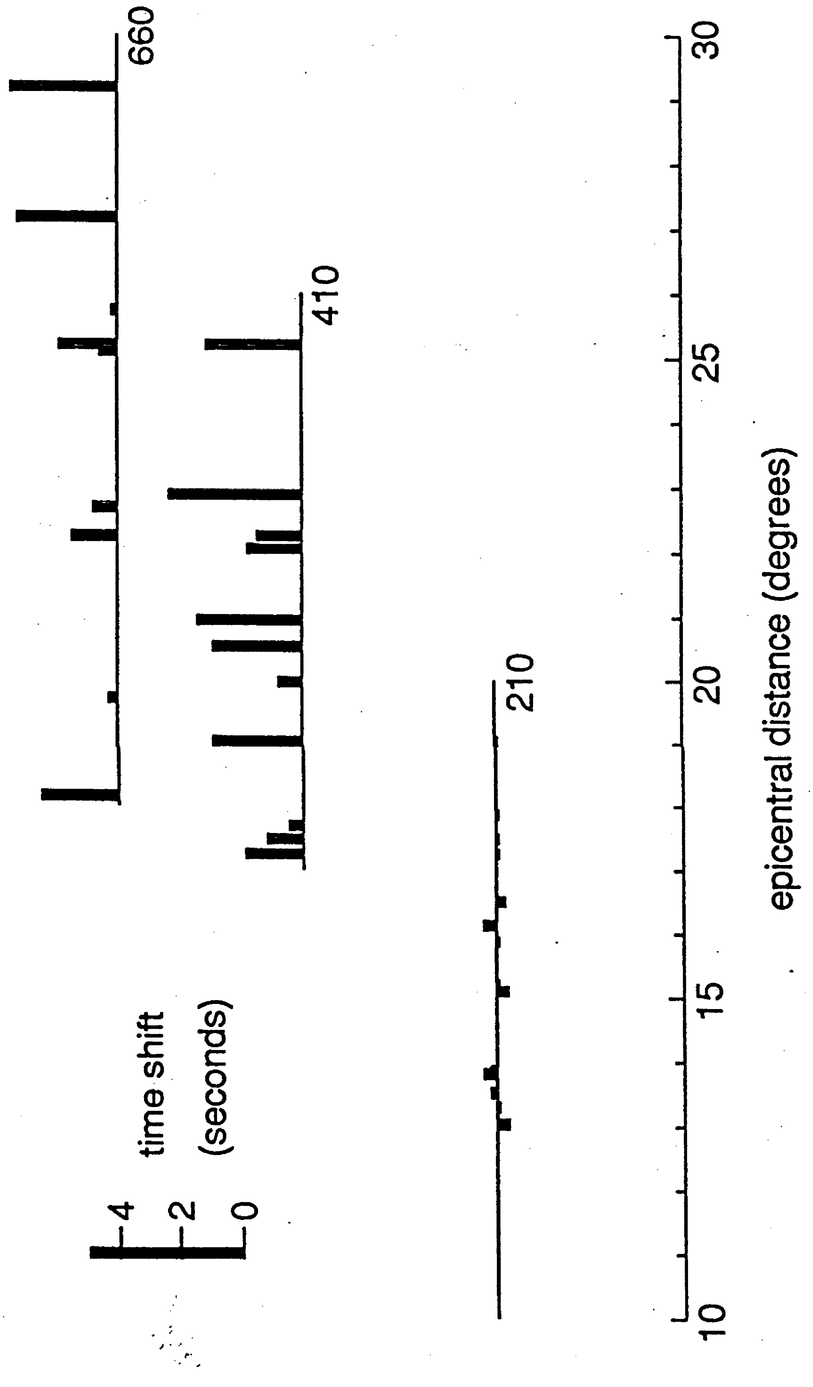


Fig. 3.7. A summary of measurements of time shift as a function of epicentral distance and branch on the travel-time curve.

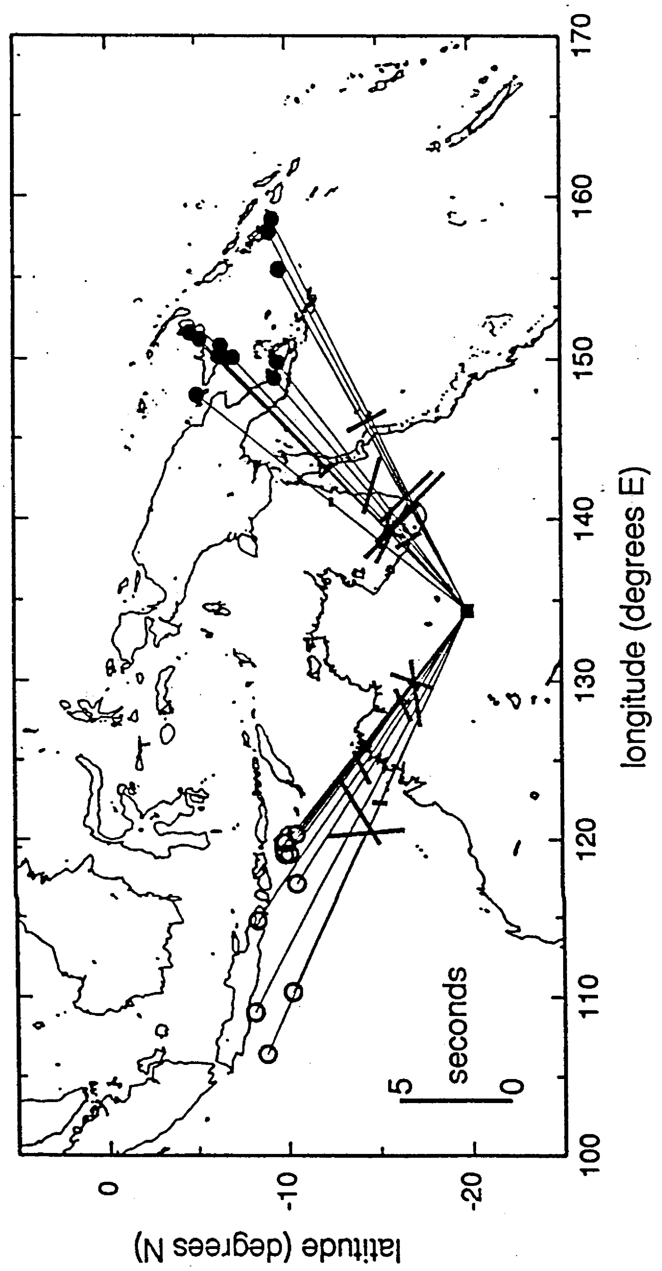


Fig. 3.8. A summary of polarization and time shift measurements. The result of each splitting measurement is represented by a time bar oriented along the fast polarization (as measured at the recording site). The length of the bar is proportional to the measured time shift as indicated by the bar in the lower left hand corner. Only earthquakes contributing a measurement on the 410 and 660 branches are included. Measurements on the 410 branch are plotted in gray at a quarter of the distance of each event from WRA. Measurements on the 660 branch are plotted in black at the midpoint between each event and WRA.

black. This is done for display purposes only, i.e. in order to separate the two classes of measurements in the figure. All the measurements are done at the WRA site and are azimuthal in geometry. The features of the results that become evident in this figure are the distinction between the two classes of events used (Indonesian on one hand and New Guinean on the other). The absolute polarisation is fairly consistent within each cluster, but very different for the two clusters. The fast polarisation is close to the transverse direction to each ray, particularly so for the cluster of events in New Guinea and the Solomon Islands. The relative values for the time delays of the two branches are different in the two regions. The average time delay is about the same on the 660 branch as on the 410 branch for the Indonesian events, but vary by almost a factor of two for the New Guinean events, the 410 branch having larger values. It is not clear if much can be made of this difference because the number of observations on each branch from each cluster is quite low.

In summary, no indications of anisotropic effects on the propagation of shear waves within the mantle "lid" were found. A clear indication was found of shear-wave splitting for propagation paths through the lower parts of the upper mantle, somewhere along the path. Polarisations are generally fast in the transverse direction and slow in the radial direction. The sizes of the measured time delays are of the order of 2 seconds on average with considerable scatter of about 1 second. The time shift may increase with epicentral distance. The time shift is on average somewhat larger on the 410 branch than the 660 branch, particularly for events in the New Guinea region.

3.5 DISCUSSION AND CONCLUSIONS

We hope to have convinced the reader that we have observed quantifiable manifestations of shear-wave splitting in the propagation of shear waves which are refracted through the lower part of the upper mantle. What remains is to isolate the source of this anisotropic effect. That is both a poorly constrained and non-unique exercise. However, we can successfully discuss this problem in terms of idealised geometries of anisotropy and the simple layering of the upper mantle, which is a manifestation of temperature as it varies with depth and perhaps more importantly a manifestation of mineralogy. First, we will consider two idealised anisotropy geometries, azimuthal and transverse. A vertically propagating shear wave is sensitive to azimuthal anisotropy, but blind to transverse anisotropy, while a horizontally travelling shear wave is sensitive to transverse anisotropy and blind to azimuthal anisotropy. Second, we will consider three basic hypotheses; 1) anisotropy near the sources, 2) anisotropy near the recording station, or 3) anisotropy in between the two. Since we can rule out the first two of these we are left with attempting to make a crude and general statement about the depth to the anisotropic region, which can be constrained by what is known about the mineralogy and the anisotropy of minerals.

Since the difference in the take-off angle from the source between waves which travel within the mantle "lid" out to about 15° on one hand and waves which dive deeper into the upper mantle on the other is small, the fact that no splitting of shear waves is observed for arrivals from the "lid" argues against a source effect. There is however limited overlap. Our observations from short distance, $12 - 18°$, come from the Banda sea and the eastern end of the Flores arc while observations of the deeper penetrating branches are from geographically distinct areas (New Guinea, Java). Considering the relative conformity of the polarisation within the two clusters of events in Figure 3.8 and the lack of an apparent relationship with the orientation of the island arcs where the events occurred, it must be considered unlikely that the source region is the cause of the splitting. All the events occurred in the front-arc region of various Island arcs (Solomon Islands, New Britain, Indonesia), but the ray paths bear no uniform relation with the strike of the arcs. We therefore rule out the source region as a hypothesis.

It is much simpler to rule out azimuthal anisotropy in the lithosphere under the WRA site. The simple fact that the absolute polarisation is different for the two clusters of events from Indonesia on one hand and from New Guinea on the other decisively rules out this hypothesis. This calls for more detailed consideration because significant SKS splitting is observed everywhere on the earth's continents (Vinnik et al. 1992). Three stations on the Australian continent (CTAO, NWAO, and TAU) have an average SKS delay time of 1 second and a relatively consistent fast polarisation at N60E (Vinnik et al. 1992). Given the tectonic stability of the Australian continent one might expect WRA to have similar SKS splitting properties. But, as can be seen in Figure 3.2, ray paths for refracted arrivals from the upper mantle deviate significantly from the vertical within the lithosphere (mantle "lid") although they are generally steep at the surface. Therefore, our data set has limited sensitivity to azimuthal anisotropy in the lithosphere.

We are therefore left with those segments of the propagation paths of the split shear waves which are significantly removed from the endpoints and where the path is not too far from horizontal. The splitting is thus due to anisotropy of the transverse type. We have two observations to constrain the depth of the anisotropic region. The absence of splitting in the 210 branch for which propagation is confined to the unusually thick (210 km) mantle "lid" (lithosphere) constrains the region from above. The lack of significantly anisotropic minerals which are stable below a depth of about 480 km (Mainprice and Silver 1993) constrains the region from below. Olivine is the most anisotropic of the most abundant minerals in the upper mantle and apparently the most ductile mineral, i.e. the mineral which most easily aligns with flow (Anderson 1989). Its anisotropy and alignments of its orientation are generally believed to cause the azimuthal anisotropy which has been observed in Pn waves and vertical S, ScS, and SKS waves. It is also believed to be the main cause of the transverse anisotropy which manifests itself in the inconsistency between Love and Rayleigh waves in terms of isotropic models (Anderson 1989). At or near the 410 discontinuity the low pressure form of olivine is transformed into a spinel-like structure, $\beta - spinel$ olivine. It is almost as anisotropic a mineral as the low pressure form of the mineral, $\alpha - spinel$ olivine, but nothing is known about its petrofabrics (Mainprice and Silver 1993). The olivine content of the transition zone may be lower than nearer to the earth's surface (Duffy and Anderson 1989). Nevertheless,

$\beta - spinel$ olivine must be considered a plausible source of seismic anisotropy in the mantle. Below a depth of about 480 km $\beta - spinel$ olivine transforms into $\gamma - spinel$ olivine (e.g. Ringwood, 1991) which is only slightly anisotropic (Anderson 1989). Thus, the lower part of the transition zone is not a candidate region for anisotropy. Mainprice and Silver (1993) estimate the maximum shear-wave anisotropy of the main constituents of the lower mantle, perovskite and *wüstite*, to be around 10%. They point out, however, that little is known about the petrofabrics of these minerals, and that indications are that deformation does not induce petrofabric (Karato and Li 1992). The lower mantle is thus a weak candidate region for anisotropy.

We can determine with some confidence where below a depth of 210 km the significant splitting on the 410 and 660 branches is likely to occur. Since the measured time delays are only weakly a function of epicentral distance, if at all, it is unlikely that the splitting occurs in the region of turning for either branch. The arguments of Mainprice and Silver (1993) concerning possible anisotropy in perovskite at the top of the lower mantle are equally applicable to this case of transverse anisotropy as their case of azimuthal anisotropy. Transverse anisotropy at the top of the lower mantle is unlikely. It is not trivial to rule out anisotropy in the $\beta - spinel$ olivine regime at the top of the transition zone. In order to predict an observation where the time shift does not increase with epicentral distance on the 410 branch, the anisotropy would have to decrease sharply with depth. Allowing for a 70 km thick layer at the top of the transition zone with a linearly decreasing level of anisotropy results in about 10% anisotropy at the top of the layer in order to satisfy the 1.7 second delay on the 660 branch. The maximum level of shear anisotropy in $\beta - spinel$ olivine is 14% (Anderson 1989) which assumes perfect alignment in an aggregate of 100% $\beta - spinel$ and sampling of the maximal polarisation. 10% anisotropy is therefore unreasonable. Transverse anisotropy at the top of the transition zone cannot be the major contributor to our splitting measurements, although some contribution is possible.

We are then left with the depth interval between 210 and 410 km as the source region for the anisotropic effect. Rays corresponding to arrivals on the 410 and 660 branches traverse this region at an angle of about 65° and 50° from the vertical, respectively. Thus, the polarisation of the SV waves is 65° and 50° from the horizontal. Assuming a simple variation of velocity with polarisation angle from the horizontal, θ, of the form

$$v(\theta) = v_o \cdot (1 + \alpha \cos 2\theta + \beta \cos 4\theta)$$

and uniform anisotropy with depth between 210 and 410 km, we can match the average time delay on the 410 and 660 branches with the parameters α and β. They come out to be $\alpha = 0.0036$ and $\beta = 0.0048$. We cannot interpret those numbers with much confidence individually, but the maximum anisotropy of 1.4% between a horizontally polarised SH wave and an SV wave polarised at an angle of 50° from the horizontal is indicative of the level of anisotropy needed in this region in order to explain our data. An anisotropy of the order of 1% is entirely feasible. A higher level of anisotropy in a thinner layer is also possible.

The transverse anisotropy of the upper mantle was probably studied in most detail by Regan and Anderson (1984). They find this form of anisotropy to be confined to the vicinity of the low-velocity zone under the oceans from their predominantly oceanic surface-wave dispersion data. The oceanic lithosphere appears to be only weakly transversely anisotropic. Our finding that no splitting occurs in waves refracted through the lithosphere is consistent with the models of Regan and Anderson (1984) in this sense. However, very different depths are involved. The low-velocity zone under the oceans occupies the depth range from about 50 to 250 km, whereas the lithosphere under northern Australia is 210 km thick. A weak low-velocity zone extends almost to the 410 km depth discontinuity (Kennett et al. 1993) (see Figure 3.2). The fact that we measure splitting of waves returning from depth in the upper mantle indicates that transverse anisotropy persists under Australia to a depth approaching 400 km which may be associated with the low-velocity zone. The PREM reference earth model of Dziewonski and Anderson (1981) includes a layer of transverse anisotropy down to a depth of 220 km and is thus similar to the regionalised oceanic models of Regan and Anderson (1984). Since, the earth's surface is predominantly covered by oceans, any true global model would be biased towards oceanic structure. It is therefore not surprising that the anisotropic structure under a continent such as Australia is much different from a global model.

We conclude that we have made clear measurements of splitting in shear waves returned from the mantle transition zone and below. Considerable scatter persists in the results of our measurements, but in terms of averages they are consistent with transverse anisotropy in the low-velocity zone under northern Australia. This is not a unique explanation from a seismological point of view, but one which avoids conflict with conventional wisdom about mineralogy, anisotropy of minerals and petrofabrics. The resulting model of about 1% anisotropy in a 200 km thick low-velocity zone at depth compares well with results from surface waves concerning transverse anisotropy in the upper mantle. The relationship of the anisotropy with the main features of the velocity-depth profile are the same.

Part three

An automatic interpretation system for broad-band seismograms

4

Characterisation of seismic phases – an automatic analyzer for seismograms

4.1 INTRODUCTION

The frequent demand for comparisons between the appearances of seismic phases motivates us to develop an efficient analyzer to recognise differences in the appearances of phases. To be an automatic system, it should also include an automatic phase picker. Therefore, two different pattern recognition problems are required – phase detection and phase feature extraction. First, we need to separate seismic signals from background noise, i.e., we consider two pattern classes. Then, we try to recognise the differences between phases, which requires working with many pattern classes. The comparison of different phases is expedited if the waveform information for the seismic phases is held in an efficient compressed form of characteristic measurements (called features).

The automatic phase-detection problem has been well studied (see e.g. Earle & Shearer, 1994); but there has not been a corresponding exploration of the automatic extraction of the features characterising a seismic phase.

During the past decades, most of the techniques dealing with pattern recognition problems in seismic studies (see e.g. Chen, 1982) belong to the class of *statistical approach*, in which features are normally extracted as statistical measurements, and the recognition of each pattern class is made by a partition of the feature space. For example, a correlation coefficient is calculated and used as a means of measuring the similarity between two phase-waveforms (see e.g. Tong et al 1994), and the classification into "non-similar" and "similar" is decided by partition of the space of the measurement into $[-1,\ c]$ and $[c,\ 1]$, where c is a predetermined constant.

It would be difficult to apply the statistical approach to the problem of feature extraction for an arbitrary phase. Because the patterns under consideration, the appearances of phases, are quite complex and the number of pattern classes is large, the problem of phase characterisation is essentially an artificial intelligence problem. Therefore, we introduce an alternative approach from the field of pattern recognition which defines features with a clear practical interpretation, and where the pattern to be distinguished is described in terms of simpler subpatterns and each simpler subpattern may again be described by even simpler subpatterns. This is termed a "structural (or syntactic) approach". By using this approach, the phase detection and feature extraction processes can be accomplished efficiently by performing analyses of the structure of a wavetrain.

4.2 SOME PREVIOUS SYNTACTIC APPROACHES FOR A GENERAL WAVEFORM-FEATURE-EXTRACTION

A syntactic approach for general waveform feature extraction has been tried by people working in the field of artificial intelligence. Pavlidis (1971) used piecewise linear sections to simplify waveforms, and later Horowitz (1977) described a peak recognition procedure which uses as its input a polygonal approximation (i.e. the piecewise sections produced by Pavlidis' algorithm). People have also tried to utilise *a priori* knowledge for some special waveforms. In the field of medical research, Stockman et al (1975) have developed a wave parsing system for the carotid pulse wave which is generated from the carotid artery in the neck. They attempted a functional approximation by merging waveform points into either lines or parabolas.

The common procedures in these methods can be summarised as the following:

- First, the waveform is segmented into simple sections (called "pattern primitives") according to some predetermined standards.
- Then, the structure of the input pattern is constructed on top of the pattern primitives.
- Features are extracted by an analysis on the structure of the waveform.

We will adopt a similar approach for analyzing seismograms but will need to adapt our approach to the character of seismograms. We will first describe the way in which we can divide a seismogram into segments so that we can recognise the onset of a seismic phase; and then turn our attention to the specification of the features of the phase.

4.3 SEGMENTATION OF A SEISMOGRAM

In the case of a long section of seismogram, especially for a broad-band seismogram, phases with very different frequencies and different magnitudes may appear in turn and the background situation changes all the time (e.g., the tail of a previous phase will be the local background for another phase). Therefore, it is not appropriate to have a predetermined standard for the procedure of initial segmentation. Consequently, for seismic phase analysis we cannot simply adapt any of the previous algorithms which were designed for general waveform analysis.

In human phase recognition, a phase is recognised when an obvious incoherence (in amplitude and frequency) between a waveform section and its neighbourhood is observed. In other words, the recognition of a phase needs the information of its neighbourhood. This property of a structure is called "context-sensitive", and represents a knotty problem in pattern recognition.

However, in the case of a seismogram, a promising scheme for a simple context analysis can be constructed by using a dynamically updated standard for the initial segmentation.

As we know, frequency is an important character of a seismic waveform. A number of methods have been used to assess the frequency content of a seismogram as a function of time including a sliding window Fourier analysis. Since the rate of zero crossings on a seismogram can approximately represent the local frequency content, this suggests a natural approach to the initial segmentation.

4.3.1 Segmentation by zero crossings

A seismogram can be easily divided into consecutive sections according to its zero crossings. Twice of the time period between zero crossings provides a rough measure of the local period, which is the inverse of local dominant frequency.

Figure 4.1 (a) gives an example of segmentation of a portion of a wavetrain by zero crossings. The dashed lines indicate the segmentation points. We notice that small peaks at the zero crossing can lead to an apparent frequency which is much higher than the true dominant frequency. To avoid this kind of mistake, the procedure of segmentation should have the ability to neglect small peaks. This requires us to pose the two problems: what is the standard for "negligible peaks" and how should they be neglected?

As we know, a small peak (in amplitude as well as duration) among large peaks must be a negligible peak. But if the same peak appears in another local environment of comparable or even smaller peaks, it can never be neglected. For this reason, instead of setting a fixed predetermined standard for the peaks to be neglected, we shall design a system which can dynamically update the standard using current knowledge of the local frequency and amplitude.

The second problem, how should small peaks be neglected, can be solved by the process of peak recognition.

4.3.2 Peak segmentation

The problem of peak recognition in general waveforms was studied by Horowitz (1977). Let a waveform be represented by a discrete set of points (x_1, y_1), (x_2, y_2), ..., (x_i, y_i), ..., (x_n, y_n), with $x_i < x_{i+1}$ for all i=1, ..., n. Since the local extrema $EX(x_{EX},\ y_{EX})$ should be a sequence of maxima and minima which appear alter-

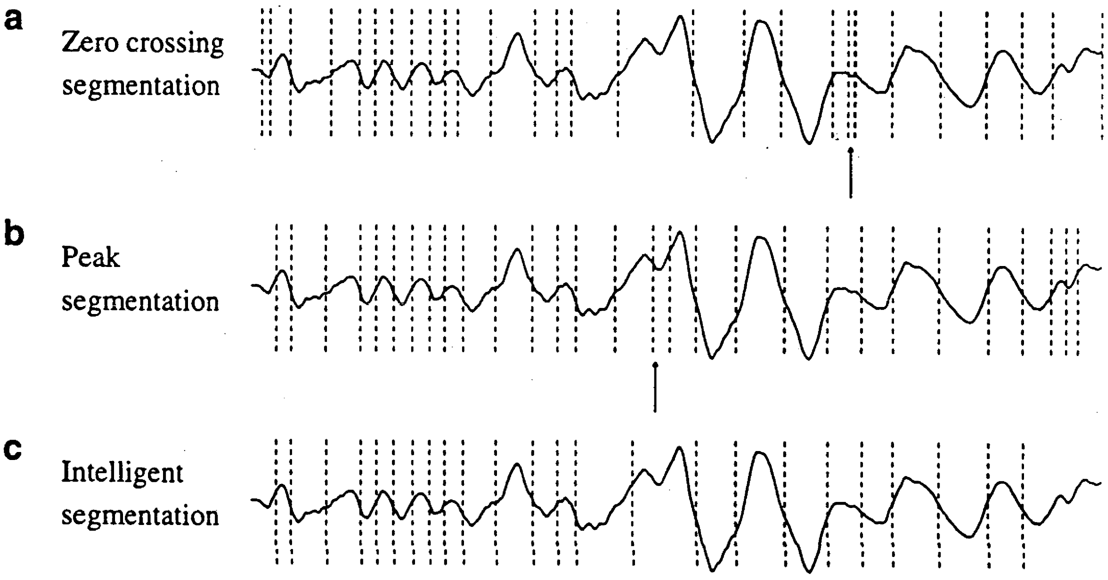


Fig. 4.1. Comparison of the results generated by three segmentation methods. The dashed lines indicate the resulting segmentation points. (*a*) Zero crossing method. At the place indicated by an arrow, frequent zero crossings do not represent the dominant local frequency. (*b*) Peak method. This method uses predetermined standards to neglect small peaks. At the place indicated by an arrow, the branch peaks in a big pulse do not represent the dominant frequency. (*c*) Intelligent segmentation. Segments generated by this segmentation method can better reveal the true dominant frequency.

nately, the local maximum $EX(2j+1)$ and the local minimum $EX(2j)$ can be defined by:

$$\forall i | x_{EX(2j)} \leq x_i \leq x_{EX(2j+2)} \; (y_{EX(2j+1)} \geq y_i);$$

$$\forall i | x_{EX(2j-1)} \leq x_i \leq x_{EX(2j+1)} \; (y_{EX(2j)} \geq y_i).$$

To neglect small spikes caused by high frequency noise, we have to furnish some more restrictive criteria:

$$\begin{aligned} |y_{EX(2j+1)} - y_{EX(2j)}| &\geq c_a, \\ |y_{EX(2j+1)} - y_{EX(2j+2)}| &\geq c_a, \\ x_{EX(2j)} - x_{EX(2j-1)} &\geq c_d, \\ x_{EX(2j+1)} - x_{EX(2j)} &\geq c_d, \end{aligned}$$

where c_a and c_d are predetermined constants. c_a is the lower limit of the acceptable amplitude for peaks, and c_d is the lower limit of duration. The appropriate choices of c_a and c_d depend upon the practical requirements in a particular case.

If a suitable standard for "negligible peaks" is available, using the peak recognition method above, we can find acceptable-size peaks with an alternate positive and negative excursion. The mid-points between consecutive extrema of these dual peaks can therefore provide another kind of segmentation – the peak segmentation.

Figure 4.1 (b) shows an example for peak segmentation using predetermined c_a and c_d. The mistakes in Figure 4.1 (a) are avoided here. However, at another place indicated by the arrow the branch peaks in a large pulse are separated, and so we miss the true dominant frequency. This indicates that it is also necessary to take account of the pattern of zero crossings.

As segmentation points, we prefer the mid-points proposed in this section to the zero-crossings of section 4.3.1. The zero crossing segmentation is very sensitive to the particular baseline chosen. For example, if the chosen value for baseline is lower than the true value, a segment below the baseline will be much smaller than its two neighbours. This kind of uneven segmentation can be avoided by choosing the mid-points between extrema of neighbouring peaks.

4.3.3 An intelligent segmentation

Based on the previous discussion, an intelligent segmentation procedure has been designed to reveal the dominant frequency. The processing procedure can be summarised as following:

(1) The time duration between zero crossings gives a first estimate (denoted as d_0) for the duration of current segment. The local maximum amplitude a_0 is also extracted in the same process.

(2) If the local maximum amplitude a_0 is large compared with the previous value a_p, i.e. $a_0 > 3a_p$, the preceding waveform will have little effect on the local waveform, therefore, set the two parameters for peak recognition (see section 4.3.2) c_d and c_a as: $c_d = d_0/5$, $c_a = a_0/5$; otherwise, the preceding waveform information will contribute to the recognition of the current segment, i.e. $c_d = (d_p + d_0)/10$, and $c_a = (a_p + a_0)/10$, where d_p and a_p are duration and amplitude of the preceding segment respectively.

(3) With the parameters c_d and c_a estimated, a peak recognition procedure as described in section 4.3.2 is applied to the waveform to find the appropriate next local extrema (with an alternate positive and negative excursion).

(4) The mid-point in time is found between the most recent two consecutive extrema and is taken as the current segmentation point. The final decision on the current segment is thus obtained as the waveform section between the most recent two segmentation points.

(5) The parameters d_p and a_p are updated with the duration and amplitude of the current segment respectively. Go to step (1) and search the next segment.

In the procedure above and some other procedures which appear later in this chapter, some of the parameters chosen are not expected to be generally and absolutely correct, but have been found to be sufficiently effective by experiment. For example, in (2) above, we choose a factor of 3 for amplitude maxima and a factor of 5 for peak sizes.

The segments of the seismogram generated by the intelligent segmentation scheme are called "units" in this chapter.

An example of the results generated by the intelligent segmentation is shown in Figure 4.1 (c). Comparing it with the results in Figure 4.1 (a) and (b), we can see that the mistakes in the two upper panels are avoided. The units obtained by the intelligent segmentation scheme are a sequence of pulses, positive (hill-shape) and negative (valley-shape) alternately. Twice of the time duration of each unit can approximately represent the local period which is the inverse of the local dominant frequency. Therefore, this segmentation procedure is used for the initial structural division in our phase characterisation system.

The intelligent segmentation procedure corresponds to the first block in the flow diagram of the whole system (see section 4.4.4). As described above, this procedure itself requires continuous feedback of the most recent knowledge for local frequency and amplitude. Therefore, the feedback cycle in the block diagram will be performed once every unit as the input waveform passes through the whole system.

4.4 PHASE DETECTION AND CHARACTERISATION

Units, the segments generated by the intelligent segmentation, constitute the lowest level substructures for a seismogram. The higher level substructures are seismic phases. The appearance of a seismic phase in a seismogram is normally a section of waveform which has greater amplitude than its background. For one arrival of seismic phase, the waveform section should have a consistent frequency content and a regularly changing amplitude cycle after cycle. The phase may last several cycles for a body wave phase, and even more for a surface wave phase.

4.4.1 Phase detection

People have been using "STA/LTA ratio methods" to detect seismic phases for a long time (see e.g. Earle & Shearer, 1994). These methods use the ratio of a short-term-average (STA) over a long-term-average (LTA) to trigger on the onsets of phases. The optimum lengths for STA and LTA windows depend on the frequency content of the seismogram. For example, long-period waveforms require larger averaging windows than do short short-period waveforms. In a long section of broad-band seismogram, phases with all different frequencies can appear in turn. Using specific window lengths, only those phases whose frequencies fit the pre-selected windows or the phases whose amplitude are very prominent will be picked out.

4.4.1.1 A real time detector

Having developed a real time procedure for intelligent segmentation, we can improve the ratio methods by dynamically updating the lengths of STA and LTA windows using current knowledge of the local dominant frequency.

A technique of "updating STA & LTA" based on the unpublished work of Dr. S. Horiuchi at Tohoku University provides the convenient evaluations for the STA and LTA values:

$$\begin{aligned} STA &= C_S(STA + a^2) \\ LTA &= C_L(LTA + a^2) \end{aligned} \tag{4.1}$$

where, a is the current amplitude of the seismic signal; C_S and C_L are predetermined constants which control the rates of attenuation of the contributions from preceding amplitudes, and thus control the lengths of STA and LTA windows. The simple updating scheme for STA and LTA evaluations avoids the tedious procedures of keeping a memory for recent seismic signal.

In our real time detector, the values of C_S and C_L are changed with the local frequency.

The two expressions in (4.1) have the common form

$$A_i = C(A_{i-1} + B_i). \tag{4.2}$$

In the evaluation of the current A_i, the contribution from the current amplitude is $C \times B_i$; the contribution from the preceding amplitude is $C^2 \times B_{i-1}$; ...; and the contribution from a previous amplitude B_{i-m} $(m \geq 1)$ is $C^{m+1} \times B_{i-m}$.

Suppose a contribution is considered as negligible when its value is smaller than one tenth of the original, i.e.,

$$C^{n+1} \times B_{i-n} \ < \ 0.1 \times B_{i-n}. \tag{4.3}$$

Let k be the smallest number of time steps which satisfies the equation above. The effective contributions are $C \times B_i$, $C^2 \times B_{i-1}$, ..., $C^k \times B_{i-k+1}$. Therefore, the length of the window in consideration is measured as k, which represents the time duration of k data points. The value of k satisfies

$$C^k \times B_{i-k+1} \ = \ 0.1 \times B_{i-k+1}. \tag{4.4}$$

Then, we have

$$C \ = \ 10^{-1/k}, \tag{4.5}$$

which gives the relation between the attenuation parameter C and the length of the window k.

In the new phase detection system, the length of STA window is set to be one dominant period, i.e., $k \ = \ 2m$, where m is the number of data points in the current unit. Therefore, the parameter C_S is updated as

$$C_S \ = \ 10^{-1/(2m)}. \tag{4.6}$$

We choose the length of LTA window as 12 times as long as the STA window, then

$$C_L \ = \ 10^{-1/(24m)}. \tag{4.7}$$

By using dynamically updating values for attenuation parameters C_S and C_L as indicated in equations 4.6 and 4.7, the new detector automatically chooses the lengths of STA and LTA windows, therefore, it has a better adaptability than the previous phase detectors.

Figure 4.2 demonstrates the applications of some typical "STA/LTA ratio detectors" and our new detector on the same record. All these detectors use the technique of "updating STA & LTA" to evaluate the STA and LTA . The typical detectors use fixed values for C_S and C_L, while the new detector chooses the values for C_S and C_L according to equations 4.6 and 4.7.

There are six panels of waveforms on display: the original Z component waveform, the local period (i.e., the inverse of the local dominant frequency) given by the the procedure of intelligent segmentation, then the measurements of three typical phase detectors (each time value in parentheses indicates the length of the window for STA, the corresponding LTA window is 12 times as long), and last, the quantity designed to act as the real-time-detector.

A deep event recorded at a distance of 29° is displayed which has particularly clear later arrivals - note the benefit of broad-band recording since the PP frequency is about 0.25 Hz and would not be recorded effectively on a short period instrument. At this distance, there are two pP arrivals associated with the upper mantle triplication.

The measurement of local period indicates that the prominent P phase has a shorter period than the background noise. As the influence of the P phase reduces with time, the value of local period gradually rises, and reaches its highest near the arrival time of phase PP.

The three typical ratio detectors represent a high frequency phase detector, a medium frequency detector, and a low frequency detector respectively. We can see that the high frequency phase detector is not sensitive to some long period phases (PP in this example), while a low frequency detector is not sensitive to some high frequency arrivals (pP in the figure). In the last measurement of the new detector, phases with different frequencies (PP and pP) are picked out at the same time.

Note 1. The dashed horizontal lines in Figure 4.2 indicated as "trigger1" and "trigger2" are two choices of the threshold for phase detection. $trigger1$ is used to detect the first phase of an event, while $trigger2$ is applied for the detection of later phases. Both of them are predetermined constants, and their values are affected by the ratio of the length of STA window compared with the length of LTA window.

Note 2. In our automatic system, when the measurement of the real time phase

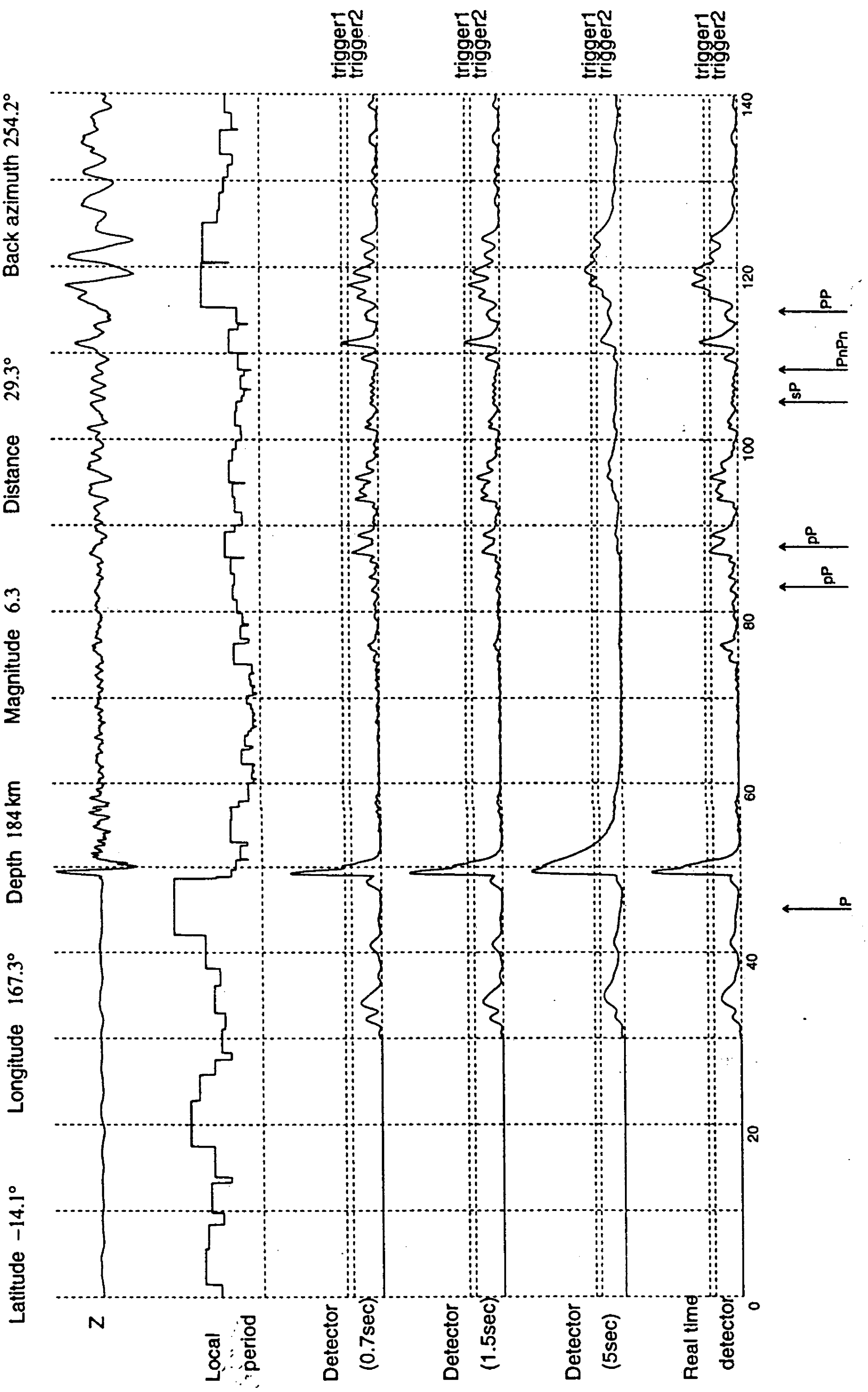


Fig. 4.2. Comparison of the results generated by different ratio detectors.

detector is greater than a predetermined threshold, a phase onset is detected, a "phase-flag" is set "on", and a procedure called the "phase separator" (see section 4.4.3.1) is triggered. The flag will be set "off" when the phase section is declared as completed by the phase separator.

4.4.1.2 An improvement : real time filters

The phase detector described above tries to detect any phase which is dominant in energy among the arrivals happening at the same moment. But a phase which has a different frequency from the local dominant frequency will be neglected.

To improve the performance of the phase detector, the use of real time filters is proposed. Two variable filters are used to filter the seismogram one unit at a time as the waveform stream passes through: the first filter is a low-pass filter which accepts frequencies at or below the dominant frequency, to produce a low frequency component of the waveform; and the other is a high-pass filter which accepts frequencies higher than the dominant to produce a high frequency component. Two detectors, one for each composite seismogram, will be used to indicate low frequency arrivals and high frequency arrivals at the same time.

As implemented in our system, the high frequency component has been filtered with a high pass from $2f$ (where f is the local dominant frequency) using a causal, four-pole, Butterworth filter. In the action of such a high pass filter, there is a convolution procedure, which always produces a time delay. Fortunately, the time delay is tolerable for a high-pass filter. To avoid the considerable time delay produced by convolution in a low pass filter, a subroutine *smooft* (Press et al, 1988) has been used to smooth the waveform and act as the low-pass filter. Before this routine is called, a parameter n is set as $\frac{4}{5}m$, in which m is the number of data points in the current unit. The parameter n controls the level of smoothing over n adjacent points. Because the routine *smooft* is applied on the waveform one unit by one unit, the boundaries between units will not be smooth. Fortunately, this has little effect on the results of later phase detection and phase characterisation.

Some examples which demonstrate the application of this approach will be given in section 4.5.

4.4.2 Models for phase waveforms

To separate and characterise a phase which has been detected, we have to consider the possible appearances of phases. A function $f(t)$

$$f(t) \ = \ e^{-(2\pi f_M/\gamma)^2 t^2} cos\ (2\pi f_M t \ + \ \nu),$$

with proper parameters f_M, γ and ν, can simulate most observed seismic wavelets (see e.g. Červený et al, 1977). Examples of wavelets produced by this function are shown in Figure 4.3.

We have chosen to use the three parameters f_M, γ and ν as features which describe the appearance of a phase. With the addition of two further parameters, the arrival time of the "zero" (the zero time in a model, see Figure 4.3) in the input waveform stream (denoted as t_0), and the maximum amplitude in the phase waveform (denoted as B), we obtained a set of five features (t_0, B, f_M, γ, ν) to represent fully a seismic phase. The five-parameter model for a phase is $f'(t)$

$$f'(t) \ = \ B\ e^{-(2\pi f_M/\gamma)^2 (t-t_0)^2} cos\ (2\pi f_M (t-t_0) \ + \ \nu)$$

which depends nonlinearly on the five unknown parameters.

A typical procedure for fitting data into a model (Press et al, 1988) is to define a merit function which measures the difference between the data and the model with a particular choice of parameters, and determine best-fit parameters by the minimisation of the merit function. With nonlinear dependence, however, the minimisation must proceed iteratively. Therefore, even a good nonlinear fitting algorithm can sometimes cost a lot of calculation.

With the syntactic approach, the fitting problem can be solved directly with the help of practical interpretation of these parameters.

By inspecting Figure 4.3, we can see that the five parameters can be interpreted as:

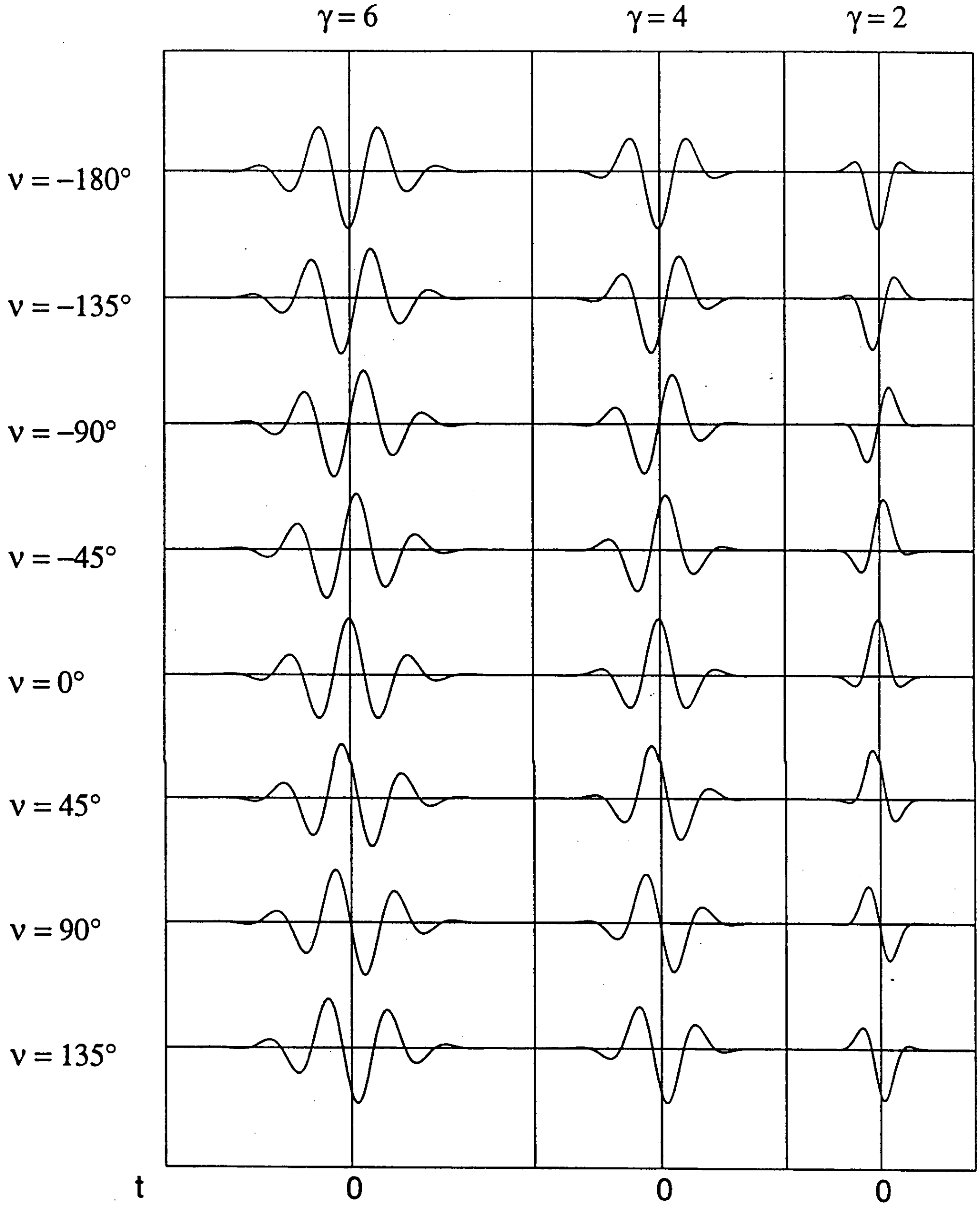


Fig. 4.3. The wavelets produced by the function $f(t) = e^{-(2\pi f_M/\gamma)^2 t^2} cos\ (2\pi f_M t + \nu)$, for γ = 2,4,6 and for ν = −180°, −135°, −90°, −45°, 0°, 45°, 90°, 135°.

(1) γ approximately equals to the number of units in the wavelet;

(2) ν is an angle which controls the relative energy distribution between the units in the wavelet.

(3) f_M gives approximately the dominant frequency of the waveform. A change of the value of f_M will re-scale the waveform on the time axis.

(4) t_0 will shift the waveform on the time axis;

(5) B will re-scale the waveform on the amplitude axis.

The parameters t_0, B and f_M decide the scale and position of the waveform, while γ and ν control the general shape of the wavelet.

Now what we have to do is to build a phase structure, then analyze this structure to obtain estimates of the waveform parameters. The detailed procedures are described in the following section.

4.4.3 The procedures for phase feature extraction

Phase feature extraction is accomplished by a two stage process:

- The construction of a phase structure – a phase separator.
- Structural analysis – a phase feature extractor.

4.4.3.1 The construction of a phase structure – a phase separator

As mentioned in section 4.4.1.1, when a phase is detected, a phase separator will be started. It takes the consecutive units (that is the sections generated by the intelligent segmentation) as its input, constructs a phase structure, which contains the important information: amplitude and duration of every compositional unit, as follows:

- A phase section starts from an onset-unit (U_o) which is the unit where a phase is detected by a detector.
- The present unit is recognised as a unit in the left part of a whole model (denoted as U_l), if: (1) the preceding unit is a U_l or U_o, (2) the present amplitude is greater than the preceding, (3) and the present duration is consistent with the preceding.

- The present unit is recognised as a "right unit" (denoted as U_r), if the the present amplitude is smaller than the preceding and the present duration is consistent with the preceding.
- When a unit is recognised as U_o, U_l or U_r, a corresponding structural node is constructed with the information of its magnitude, polarisation, and relevant times.
- The phase section will end when a present unit can neither be fitted into U_l nor U_r.

Therefore, the structure of a phase section will be constructed as a sequence of nodes in the form of "$U_o(U_l)^*(U_r)^*$", where $(P)^*$ means that P may appear many times (including zero) .

4.4.3.2 Structural analysis – a phase feature extractor

When a phase waveform is recognised as complete, a "phase feature extractor" will be started.

Because of contamination by noise or other phases, the observed appearance of a phase may not include all the units which appear in the model pattern. Fortunately, the character of a phase can be preserved in the the central part of the wavelet (including three central units), which has greater energy and will be less deformed by other effects. Once the phase structure has been constructed, a "centre unit" which has the maximum amplitude of the phase can be easily found between the last U_l and the first U_r. The five feature parameters are then estimated from their practical interpretations:

- The parameter t_0 is roughly set to be the boundary position (time) before the first U_r.
- B is set to be the amplitude of the centre unit.
- f_M is set to be the inverse of the duration of the central two units (the last U_l and the first U_r).
- ν will be set as one value among $\{-180°, -135°, -90°, -45°, 0°, 45°, 90°, 135°\}$. If the centre unit is positive, the value is chosen according to the following criteria:

(i) If the structure of the waveform is roughly symmetric between the left and the right part of the phase (i.e., the two units beside the centre unit have an equal amplitude), the parameter ν is set to be 0°.

(ii) If the structure of the waveform is roughly antisymmetric between the left and the right parts (i.e., the last U_l and the first U_r have an equal amplitude), the parameter ν is set to be 90°.

(iii) When the structure of the waveform is neither symmetric nor antisymmetric, if the unit before the centre unit has smaller amplitude than the unit after the centre one, ν is set to be 45°.

(iv) When the structure of the waveform is neither symmetric nor antisymmetric, if the unit before the centre unit has greater amplitude than the unit after the centre one, ν is set to be −45°.

If the centre unit is negative, a first estimate for ν, denoted as ν_0, is obtained using the same criteria as above, then,

$$\begin{aligned} \nu &= \nu_0 - 180^\circ \quad \text{if } \nu_0 \geq 0^\circ \ , \\ \nu &= \nu_0 + 180^\circ \quad \text{if } \nu_0 < 0^\circ . \end{aligned}$$

- Inspecting the standard wavelets in Figure 4.3, we can get an idea of the possible number of units for the complete wavelet by using a rough estimate of $\gamma = 2\ h_m/dh$, where dh is the amplitude difference between two consecutive units and h_m is the maximum amplitude of the phase.

Experiments show that the pattern model calculated from the five parameters estimated by the procedure above matches very well with the structure of the waveform, but the scale and position may need an adjustment.

By comparing the structures of the original waveform and the model calculated, each of the three parameters t_0, B, f_M can be enlarged or reduced respectively, then the estimates of the five phase feature values are ready for output.

4.4.4 System summary

The whole system of the automatic analyzer is represented in a block diagram in Figure 4.4, and can be viewed as a system for seismic data compression which runs in real-time. The input of the system is a digitised on-line seismic waveform stream, while the output is a phase-feature-stream in which every five feature-values describe a detected seismic phase. The only necessary parameter required when the system starts is an approximate position of the baseline.

4.5 EXAMPLES (EVENT - ZAIRE)

We will illustrate the performance of the phase characterisation system with the application to the same event recorded on two portable broad-band instruments deployed at two different sites in the Northern Territory of Australia. The data were recorded on Reftek 72A-05 disc recorders (16-bit resolution) with Güralp CMG-3ESP seismometers (flat to ground velocity from 0.03 - 30.0 Hz) with Omega timing. The seismometers were deployed in simple pits in consolidated gravels with a concrete base slab, but the quality of recording is high.

Because the two sites are fairly close, only several hundred kilometers apart, and use the same kind of portable broad-band seismic instruments, we expect that for the same event the two records would have similar phases. Using these two examples, we can compare the phase characterisation results for both similar and non-similar phases. We apply the automatic seismogram analyzer on the vertical (Z) components of the two records, which are indicated as "trace 3" and "trace 12" respectively.

The event is at a great distance, approximately 105°, and is in the core shadow for P. The seismograms for this shallow and distant event are rather noisy and therefore provide a good test of the phase characterisation procedure.

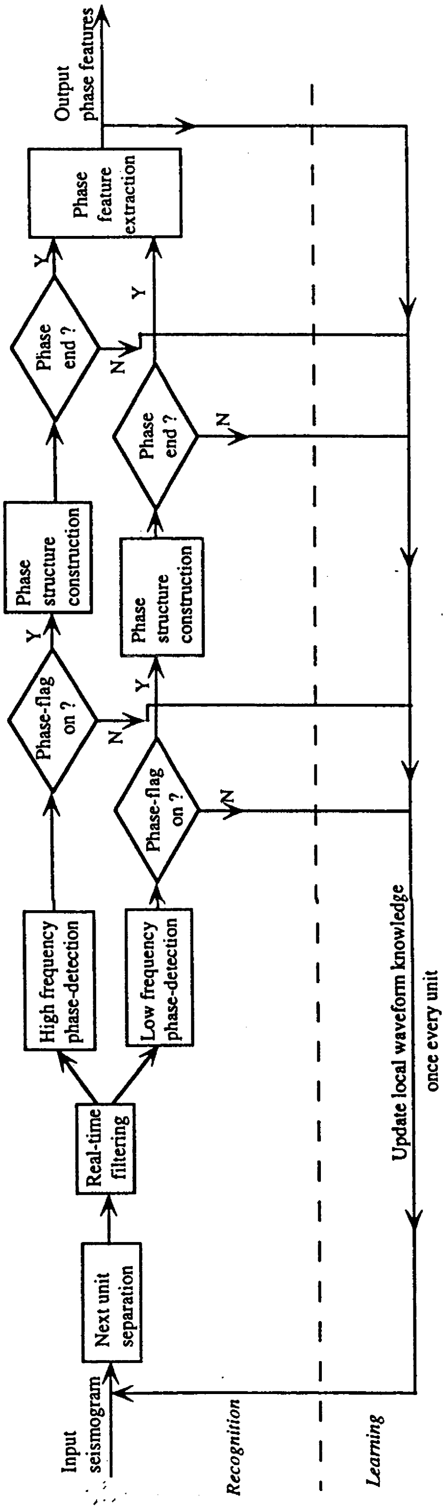


Fig. 4.4. The flow diagram of the automatic analyzer for seismograms.

4.5.1 Phase picker

The phase detection procedures carried out on trace 3 and trace 12 are illustrated in Figures 4.5 and 4.6. There are six panels in each figure. In the upper two panels, the original waveform and the real-time phase detection curve are displayed. The following four panels show the high-pass filtered waveform, the corresponding phase detector, the low-pass filtered waveform and the phase detector for the low frequency part of the waveform. Beneath these panels, the arrows labelled with phase names give us the theoretical phase times, derived from the *iasp*91 table (Kennett & Engdahl 1991), for reference.

For this event, the first arrival represents the weak diffracted P and is closely followed by diffracted sP and diffracted pP. After a while, there comes a complex package of the $PKiKP$ cluster interfering with the converted phase PP.

The use of the waveform decomposition is very successful for both of the two examples. It gives a surprisingly good isolation of the interfering phases. We can see that the detector on the low-pass filtered waveform performs similarly to the detector on the original waveform, while the detector on the high-pass filtered waveform reveals additional phases. It should be noticed that the low-pass filtered waveform is a smoothed waveform on which the phase model-fitting procedure works more efficiently than on the original waveform.

In both of the examples, the phases picked up by the two detectors on the two composite waveforms correspond very well with the theoretical expectation (i.e., the phase times given by *iasp*91 table).

4.5.2 Phase feature extractor

Figures 4.7 and 4.8 display the results of the feature extraction procedures for the traces 3 and 12.

(1) The upper panel on each figure shows the result produced by the phase picker described in section 4.4.1.2. The full trace of the original waveform (Z component) is displayed together with the phase times, indicated as arrows labelled with

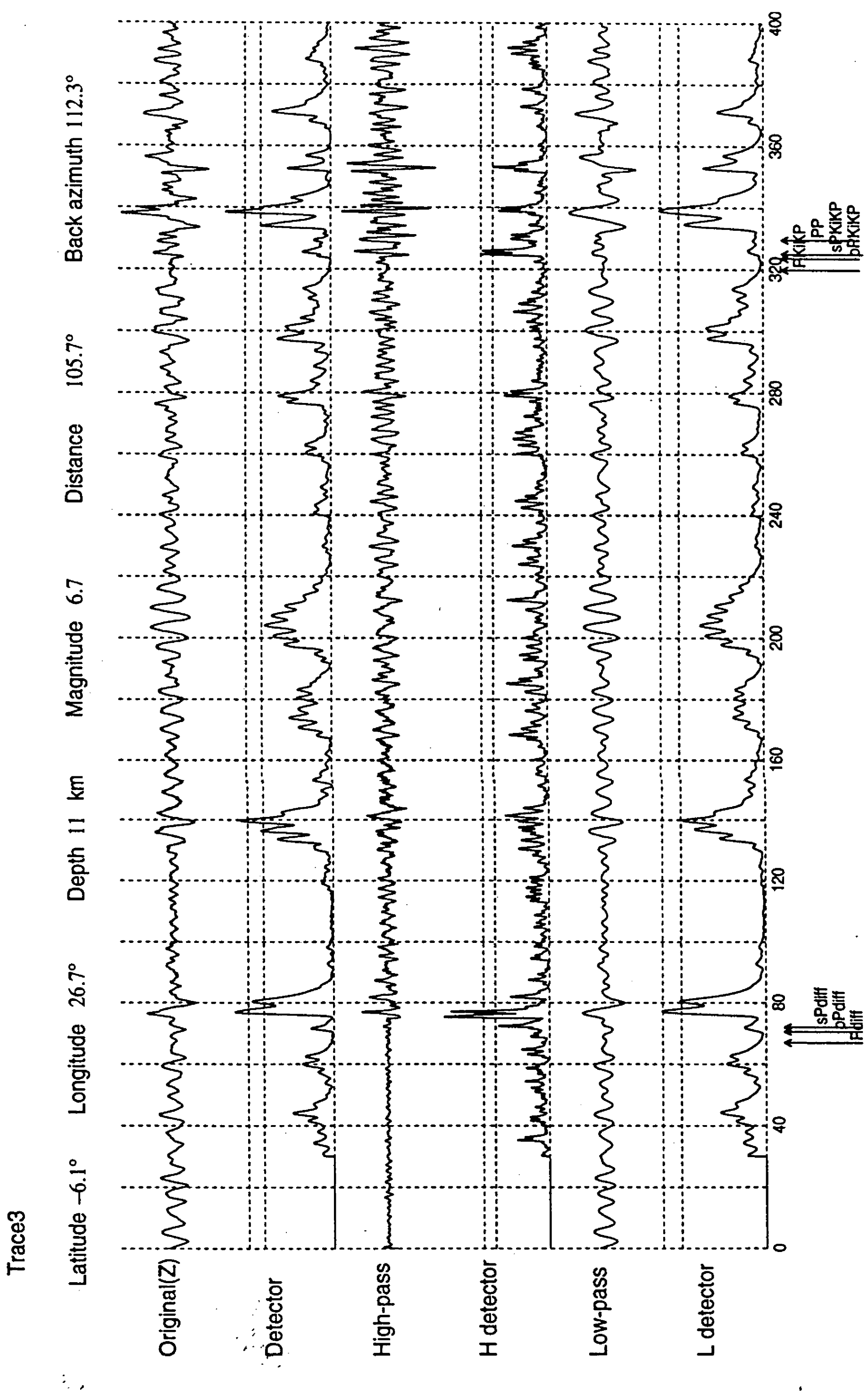


Fig. 4.5. Phase detections for trace 3, Zaire event.

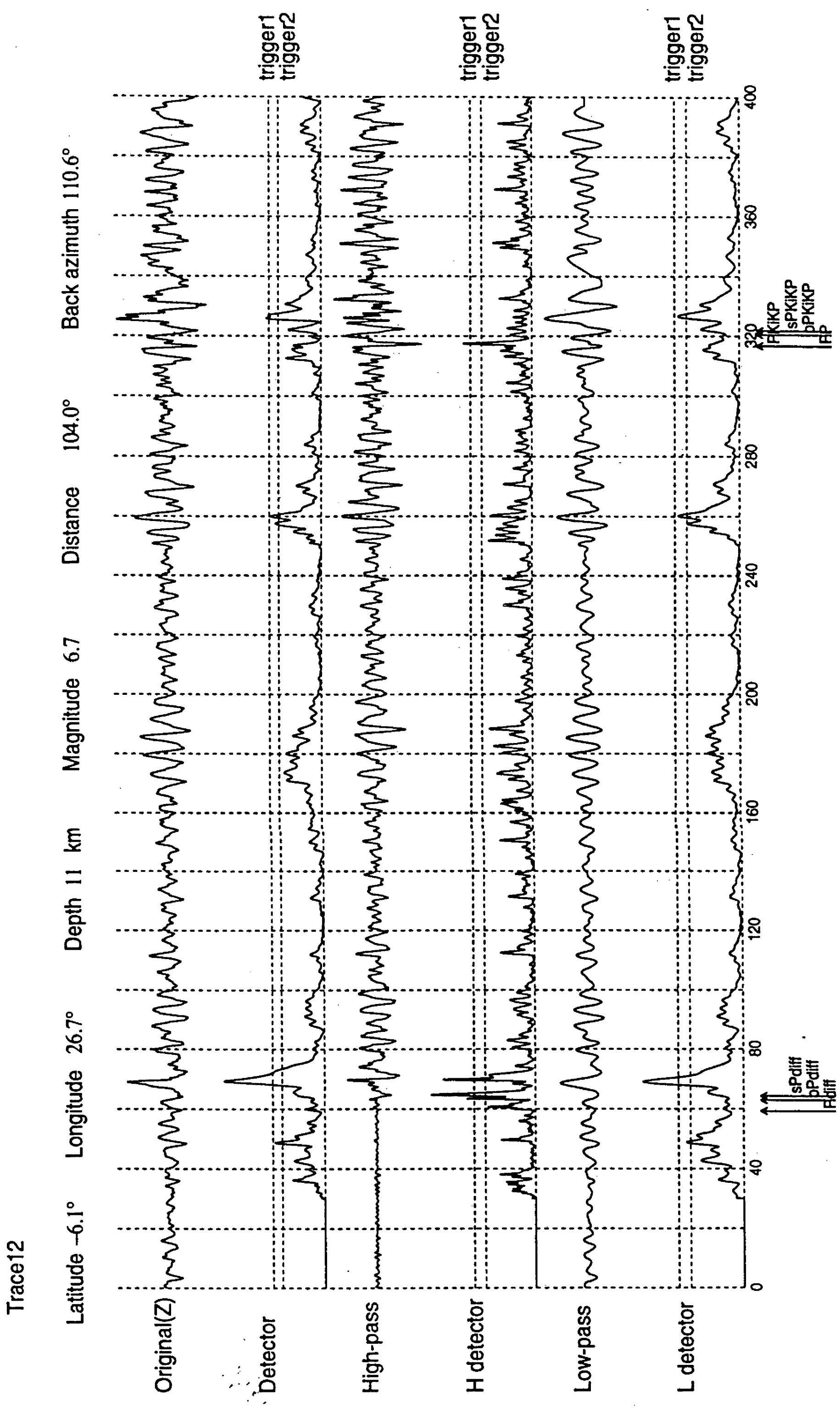


Fig. 4.6. Phase detections for trace 12, Zaire event.

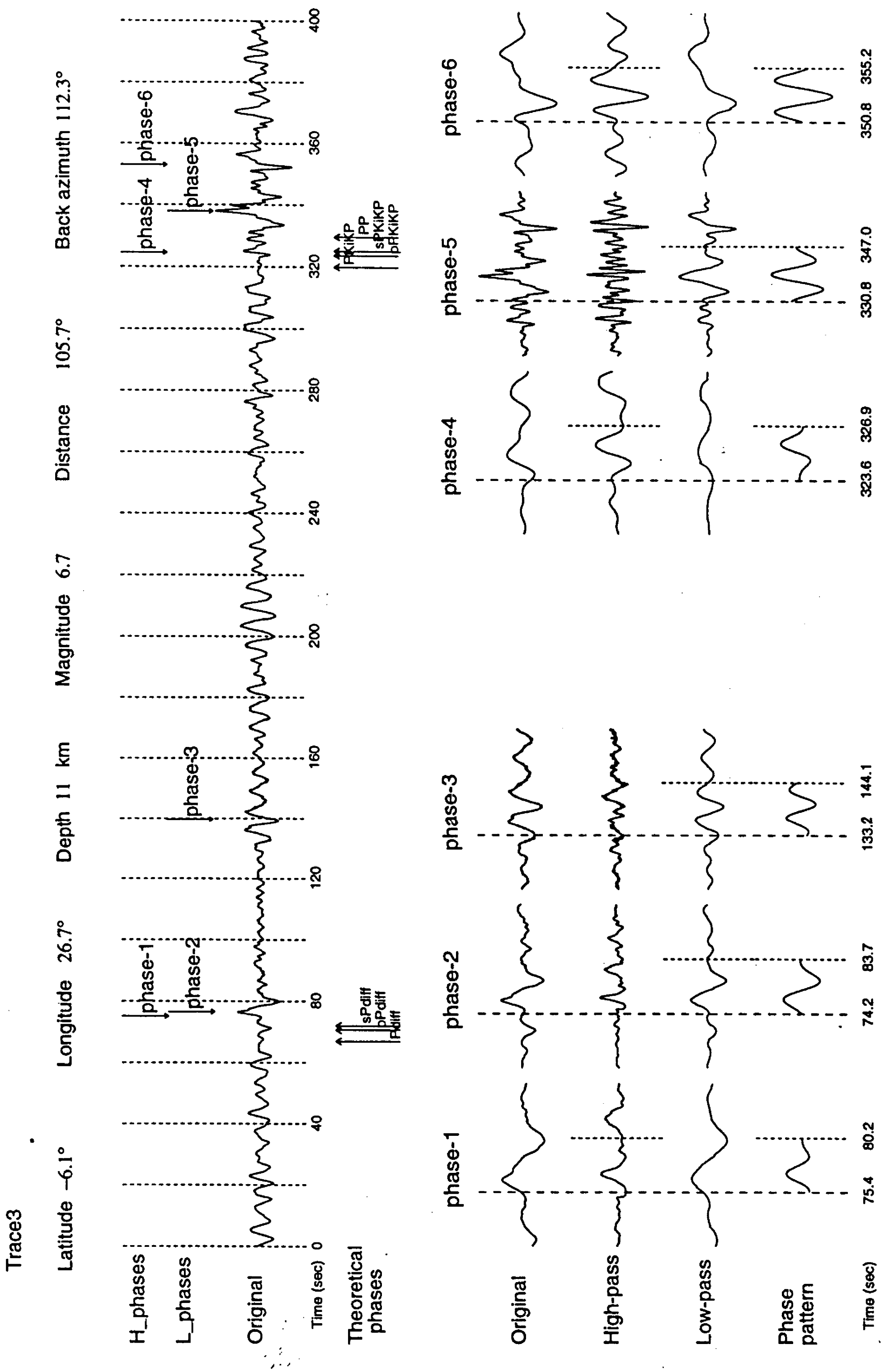


Fig. 4.7. Phase characterisation for trace 3, Zaire event.

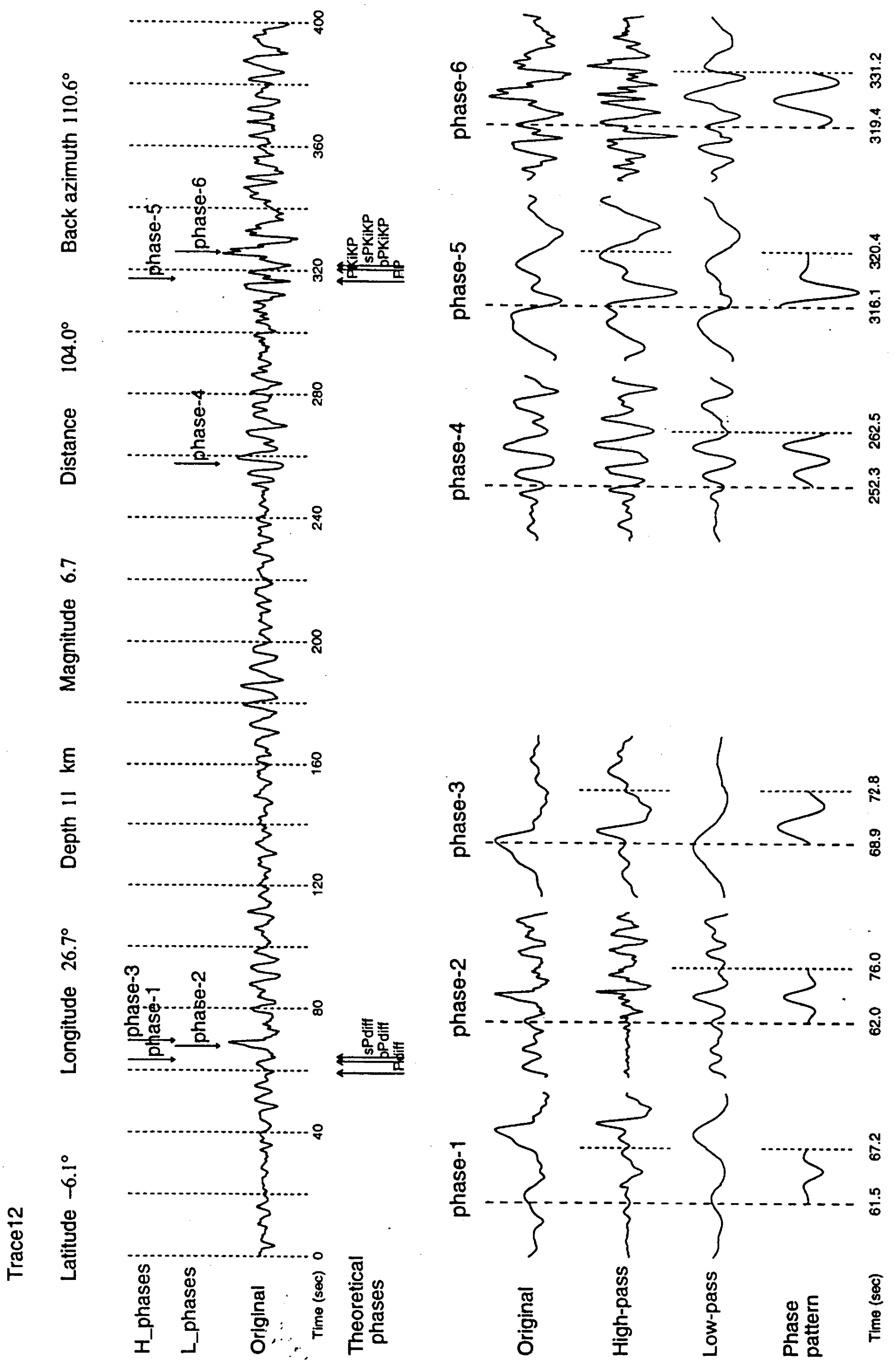


Fig. 4.8. Phase characterisation for trace 12, Zaire event.

phase names beneath the waveform. Above this waveform, there are two sets of arrows labelled with "phase-i", which indicate where phases are detected by the automatic phase picker. The set indicated as "*H_phases*" are the phases detected on the high-pass filtered waveform, and "*L_phases*" are those detected on the low-pass filtered waveform.

(2) The procedure of phase-feature-extraction is displayed in the lower panel of each figure. There are several sections from the left to the right, and the waveform segments in each section correspond to the waveforms of one of the detected phases in the upper panel. In each section we display a set of four panels of waveforms: the original waveform, high-pass filtered waveform, low-pass filtered waveform, and derived phase-pattern model. The four panels are shown at the same time scale. The beginning of a phase is indicated by the long vertical dashed line which passes through the four waveforms. For each of the "*H_phases*", the phase-feature-extraction procedure was carried out on the high-pass filtered waveform, while for the "*L_phases*", it was performed on the low-pass filtered waveform. The relevant filtered waveform is indicated with a second dashed line at the end of the phase waveform.

In these two figures we can see that the pattern model estimated by the phase feature extractor is very similar to the original waveform segment.

From visual inspection, the corresponding phases detected on the two traces are listed in Table 4.1.

4.5.3 Phase feature values

Every phase model displayed in Figures 4.7 and 4.8 is a waveform generated from five phase-feature parameters. The values for the parameters for all the phase models in Figures 4.7 and 4.8 are listed in Table 4.2.

On comparison of the feature values for corresponding detected phases as listed in Table 4.1, we can see that the frequency f_M and the parameter γ are quite consistent for the same phase recorded at the two different sites. The amplitude,

Table 4.1. *Corresponding detected phases between Trace 3 and Trace 12 for Zaire event*

Trace 3	phase-1	phase-2	phase-4	phase-5
Trace 12	phase-3	phase-2	phase-5	phase-6
Relevant	H	L	H	L
T-Phase	pP_{diff}	sP_{diff}	$PKiKP$	$sPKiKP$

In the line "Relevant", "H" indicates the phase is detected on the high-pass filtered waveform while "L" indicates the phase is from low-pass filtered waveform. "T-Phase " represents the possible phase expected according to the *iasp*91 table.

Table 4.2. *Feature values for all the phases detected in trace 3 and trace 12 for Zaire event*

	Detected phase	t_0	B	f_M	γ	ν
Trace 3	phase-1	77.00	92.38	0.274 (Hz)	3.0	0°
	phase-2	78.28	114.47	0.118 (Hz)	2.5	90°
	phase-3	139.10	91.70	0.162 (Hz)	6.0	135°
	phase-4	325.55	117.32	0.378 (Hz)	2.5	-45°
	phase-5	337.75	164.98	0.108 (Hz)	5.0	-45°
	phase-6	352.90	185.71	0.358 (Hz)	6.0	-180°

	Detected phase	t_0	B	f_M	γ	ν
Trace 12	phase-1	64.85	62.69	0.258 (Hz)	2.5	-180 °
	phase-2	68.65	101.27	0.112 (Hz)	2.5	0°
	phase-3	70.47	136.44	0.258 (Hz)	2.5	45°
	phase-4	259.34	102.56	0.186 (Hz)	7.0	-45°
	phase-5	317.06	230.13	0.340 (Hz)	2.5	135°
	phase-6	326.10	147.22	0.116 (Hz)	6.0	45°

the parameter B, is consistent for a low-pass filtered phase (indicated as "L" in Table 4.1), while not consistent for the high-pass filtered phase. This is because the convolution procedure used for high-pass filtering can change the amplitude of the original waveform. The convolution also causes a negligible time delay. The high-pass filter may generate more accurate values for parameter B if a tuning procedure is added.

It is worth noting that the parameter ν is not consistent. One plausible explanation is that this parameter might contain the information related to the recording site. Another possibility is that the parameter ν is more sensitive to the interference of phases and the precise choice of segmentation.

4.6 DISCUSSION

4.6.1 A seismogram – a context-sensitive structure

We have discussed in section 4.3 that the construction of a present substructure (the units) cannot be performed without the information of its neighbours, which represents a simple case of a "context-sensitive" system.

The procedure of phase picking by eye always proceeds in the direction of increasing time. This tells us that subsequent information has less effect on the present unit than the preceding one. Similarly, in the automatic analyzer, only the information of the preceding units is used for the analysis of a current unit. The reference to the predecessors occurs in the initial segmentation as well as in the construction of phase structures. Therefore, in the case of a seismogram, although the structure is context-sensitive, the structural analysis can still be processed in the direction of increasing time, which allows continuous real-time processing.

4.6.2 The automatic analyzer – an interpreter for seismograms

The translator for a programming language (a compiler or an interpreter) is a recognition device for the source language. Analogously, the seismic phase characterisation system described in this chapter can be considered as a translator for the special language: seismograms.

The difference between a compiler and an interpreter is that, using a compiler, the program can only be executed after the the whole program is recognised; while using an interpreter, the program is executed at the same time as the recognition.

Since our seismic-phase characterisation system has been designed as a real-time system, it is more like an interpretor than a compiler.

5

Towards the identification of later seismic phases

5.1 INTRODUCTION

The vector recording of surface motion achieved by a broad-band seismic recording system includes a wealth of information about seismic noise and the character of seismic arrivals. The procedures for phase detection and characterisation described in this chapter are based on the simultaneous use of a number of different simple schemes based on the physical character of the seismic arrivals which can be applied in real-time to a three-component record. The combination of several different pieces of information provides a good basis for the identification of many different classes of phases.

The first need is to devise a detection system which will find the onset of a seismic wavetrain against the background of microseismic and cultural noise. For broad-band records such detections can be achieved by using a modification of the common procedure in which the local behaviour is compared to a longer term trend e.g. by declaring a trigger when the short-term average of the trace energy (STA) reaches a certain fraction of the long-term average of the trace energy (LTA). In order to cope with the variations in frequency content along the record the effective windows in the averaging process are chosen in a adaptive manner depending on the local frequency content.

Once a *P* wave has been detected it is desirable to recognise and identify later phases, even though this needs to be done against the background of the *P* wave coda. With conventional three-component recording, the orthogonal components are directed vertically, north-south and east-west and so polarisation purely along one component is fortuitous. Nevertheless, we know that the vector properties of a *P* wave are relatively simple, and not affected greatly by interaction with the free surface. It is therefore possible to get quite good estimates of apparent azimuth and angle of incidence from an extracted phase. For *S* waves, we expect both vertically and horizontally polarised energy to be produced by the source so that any analysis based on the assumption of a single wavetype at any time is likely to be misleading.

Even when we do not know the direction of arrival of the energy we can make use of intrinsic properties of the wavefield to start to refine the assignment of seismic phases. For distant events, *P* waves appear dominantly on the vertical component and *S* waves in the horizontal plane. We can therefore consider a *P* wave detector based on a measure of the proportion of energy on the vertical component and a *S* wave detector based on the proportion of energy in the horizontal plane. Such detectors can be profitably coupled to the STA/LTA detection procedure to provide a range of detection procedures with focus on a specific class of phases.

A further property of the seismic record which can be used to indicate the onset of a new phase is a change in the frequency content of the seismogram. Such a change may be either the arrival of fresh high frequency energy, as in the onset of the *P* wave train or a later phase such as *SS* or shear coupled *PL* with a lower frequency than the coda of preceding phases. Monitoring the dominant local frequency of the waveform through the use of pattern recognition techniques can help to pick up the lower frequency arrivals. The onset of small high frequency arrivals can often be enhanced by the use of a measure based on the complexity of the seismogram.

5.2 THREE-COMPONENT ANALYSIS

With a seismic array it is possible to estimate the instantaneous azimuth and slowness of an arrival, based on the assumption of a plane wave crossing the array, and so gain some immediate hints as to the likely character of the associated seismic phase, particularly where three-component information is also available. For example, on the regional arrays of the Noress type the detection of *Sn* is carried out on an incoherent horizontal beam and confirmed via a phase velocity analysis (see e.g. Mykkeltveit et al 1990).

A single three-component record contains a good deal of information about the propagation path but information on azimuth and slowness can only be extracted based on some model of the instantaneous character of the wavefield. Different results will be obtained for the azimuth of propagation depending on whether the assumption is made of a *P* wave, an *SV* wave or an *SH* wave, and any estimate of horizontal slowness will be strongly affected by the way in which the azimuth is extracted.

5.2.1 Influence of wavetype

We therefore need to define the expected behaviour for the different wavetypes, recognising that they will have different patterns of interaction with the free surface. We will consider a medium with surface wavespeeds α_0 for P and β_0 for S, and assume that the azimuth of propagation is ϕ and that the angle of incidence is i for P waves and j for S waves. The connections between these angles of incidence and horizontal slowness are then given by Snell's law

$$\begin{aligned} p_\alpha &= \sin i/\alpha_0 \quad \text{for } P \text{ waves,} \\ p_\beta &= \sin j/\beta_0 \quad \text{for } S \text{ waves.} \end{aligned} \tag{5.1}$$

For an incident plane P wave with amplitude P the three components of the surface displacement including free surface effects can be written in the form:

$$Z = -\alpha_0 q_\alpha C_1 P,$$

$$N = \alpha_0 p_\alpha C_2 \cos\phi\, P, \qquad E = \alpha_0 p_\alpha C_2 \sin\phi\, P, \tag{5.2}$$

where $q_\alpha = \cos i/\alpha_0$ is the vertical slowness for the P wave at the surface. C_1, C_2 are slowness-dependent quantities representing the effect of free surface amplification on the P–SV wave system

$$\begin{aligned} C_1 &= 2\beta_0^{-2}(\beta_0^{-2} - 2p^2)/[(\beta_0^{-2} - 2p^2) + 4p^2 q_\alpha q_\beta], \\ C_2 &= 4\beta_0^{-2} q_\alpha q_\beta/[(\beta_0^{-2} - 2p^2) + 4p^2 q_\alpha q_\beta], \end{aligned} \tag{5.3}$$

where q_β is the vertical slowness for SV waves for horizontal slowness p.

For an incident plane SV wave with amplitude S the surface displacements are

$$\begin{aligned} Z &= \beta_0 p_\beta C_2\, S, \\ N &= \beta_0 q_\beta C_1 \cos\phi\, S, \qquad E = \beta_0 q_\beta C_1 \sin\phi\, S, \end{aligned} \tag{5.4}$$

with the vertical S wave slowness $q_\beta = \cos j/\beta_0$.

The presence of the factors C_1, C_2 in the free-surface displacements means that estimates of incidence angle based on the ratio of the vertical and radial components are distorted and tend to approximate the behaviour for lower wavespeeds at the surface than are actually present.

The situation for plane SH waves is simpler because there is complete reflection at the free surface with no conversion, the surface displacements for a wave of amplitude H are

$$\begin{aligned} Z &= 0, \\ N &= -2\sin\phi\, H, \qquad E = 2\cos\phi\, H. \end{aligned} \tag{5.5}$$

For both a P and a SV wave the apparent azimuth ψ estimated by $\tan^{-1}(E/N)$ will lie along the true azimuth ϕ, but for SH waves ψ will be displaced by 90° from the true azimuth.

Although the free-surface amplification effects are rather complex, Kennett (1991) has demonstrated that the inverse relation for correcting back to the amplitudes P, S and H in an infinite medium is comparatively simple. These inversion formulae can be applied to three-component records when the azimuth is known so that an initial rotation can be made in the horizontal plane to radial (R) and transverse (T) components:

$$\begin{aligned} R &= N\cos\phi + E\sin\phi, \\ T &= -N\sin\phi + E\cos\phi, \end{aligned} \tag{5.6}$$

where ϕ is the azimuth from north.

The correction for the effect of the free surface then rests on the assumption that an adequate model of the wavefield at any instant is a plane wave with a single slowness p but a mixture of P, SV and SH character. From the vertical component Z, radial component R and transverse component T the amplitudes P, S, H can then be found as

$$\begin{aligned} P &= -Z(1-2\beta_0^2p^2)/2\alpha_0 q_\alpha + Rp\beta_0^2/\alpha_0, \\ S &= Zp\beta_0 + R(1-2\beta_0^2p^2)/2\beta_0 q_\beta, \\ H &= {}^1\!/_2 T, \end{aligned} \tag{5.7}$$

The transformation coefficients are relatively slowly varying functions of slowness p for the smaller slownesses appropriate to teleseismic arrivals (see Figure 5.1, Kennett 1991). A reasonable estimate of the relative proportions of P, SV and SH waves can be therefore be made for teleseisms without knowledge of a precise slowness value. Such a decomposition into wavetype components can be quite helpful in assigning a phase identifier to a particular arrival.

For a P wave the apparent angle of incidence ι at the surface can be found from the behaviour in the vertical plane from $\tan^{-1}(R/Z)$ to give

$$\begin{aligned} \iota &= \tan^{-1}[\tan i\,\{C_2/C_1\}] \\ &= \tan^{-1}[\tan i\,\{2q_\alpha q_\beta/(\beta_0^{-2} - 2p^2\}] \end{aligned} \tag{5.8}$$

so that the apparent angle ι will be systematically smaller than the true angle i.

The models of the wavefield for the different types of plane wave propagation now enable us to define suitable measures of the wavefield which we can expect to have some degree of specificity for different types of arrival. For distant events for which

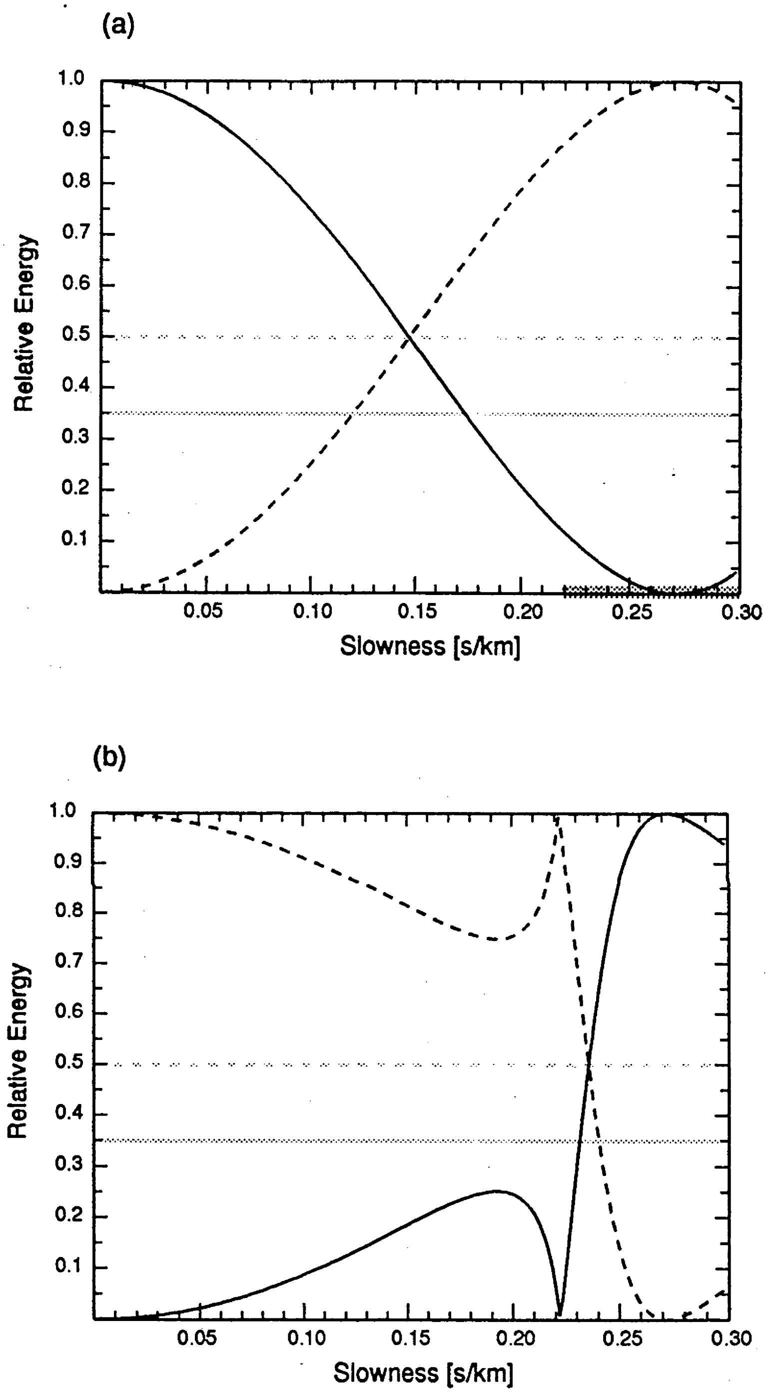


Fig. 5.1. The relative energy distribution between the vertical and horizontal components as a function of slowness p for incident P and SV waves on a free surface with surface velocities $\alpha_0 = 4.8$ km/s and $\beta_0 = 2.6$ km/s. The solid traces indicate V_E/E_3 and the dashed traces H_E/E_3. (a) incident P waves, evanescent P is indicated by shading (b) incident SV waves.

the slowness is small, the P waves will appear dominantly on the vertical component with little expression in the horizontal plane, whilst S waves will be largely confined to the horizontal plane. We therefore construct three measures of the energy on the traces:

the total energy	$E_3 = Z^2 + N^2 + E^2$,
the vertical component energy	$V_E = Z^2$,
the energy in the horizontal plane	$H_E = N^2 + E^2$,

which do not depend on any knowledge of the azimuth of the arrivals.

In Figure 5.1 we display the energy measures V_E, H_E as a proportion of the total energy E_3 for incident P and SV waves as a function of slowness for surface velocities $\alpha_0 = 4.8$ km/s and $\beta_0 = 2.5$ km/s. As expected, we note the dominance of vertical energy for P waves and horizontal energy for SV waves for smaller slownesses, which we can exploit to generate quantitative measures of phase character. For a wide range of surface conditions, including hard rock sites, we can characterise a P wave arrival by:

$$V_E > 0.35E_3, \qquad H_E < 0.5E_3 \tag{5.9}$$

for slownesses less than 0.12 s/km, a range which encompasses both far-regional and teleseismic arrivals. For the same slowness range, we can recognise SV wave arrivals when:

$$H_E > 0.5E_3, \qquad V_E < 0.35E_3 \tag{5.10}$$

and the dominance of horizontal energy will be reinforced by the presence of SH waves. For regional to far-regional S arrivals the energy patterns are complex with commonly significant energy on both the vertical and tangential components.

P wave arrivals at regional distances show significant energy on the horizontal component and can often be recognised by simultaneous vertical and horizontal energy with

$$V_E > 0.35E_3, \qquad H_E > 0.5E_3. \tag{5.11}$$

Kennett (1993) has shown the value of such simple energy measures for three-component data in a study of the distance dependence of the amplitudes of regional phases. Booker & Mitronovas (1964) had previously used a three-component energy measure in a study of amplitude ratios on short-period instruments of the LRSM network.

5.2.2 Adaptive phase detection

In chapter 4, we have described an automatic analysis system for seismic traces based on the use of pattern recognition techniques. The segmentation of the trace in this procedure provides a good estimate of the local dominant frequency. For three-component records this context-sensitive analysis is applied simultaneously to each of the three-components and a robust estimate of the local frequency f_ℓ is provided by the average of the values for the different components. This time-varying dominant frequency f_ℓ is then used to guide an adaptive phase detection scheme.

The detection scheme is based on the use of the STA/LTA ratio technique (see e.g. Earle & Shearer 1994) in which the local behaviour is compared to a longer term trend, by declaring a trigger when the short-term average of the trace energy (STA) reaches a certain fraction of the long-term average of the trace energy (LTA).

5.2.2.1 *Adaptive STA and LTA measures*

Rather than use an explicit time-window for the calculation of the STA and LTA measures, we follow Tong (1995) (see chapter 4) and use instead an approach of continual updating, based on the unpublished work of Dr Horiuchi at Tohoku University. We will illustrate the process for the total energy E_3: at each time step we increment the STA and LTA values as

$$S = C_S(S + E_3), \qquad L = C_L(L + E_3). \tag{5.12}$$

The parameters C_S and C_L are chosen smaller than unity and control the rates of attenuation of the contributions from preceding time steps, and thus the effective length of the STA and LTA windows. We take this window to be the time interval before the influence of a sample drops to one tenth of its original value. For a local frequency f_ℓ we choose the STA parameter C_S so that the effective window matches the dominant period

$$C_S = 10^{-f_\ell \Delta t}, \tag{5.13}$$

where Δt is the sampling interval. The LTA window is chosen to be 12 times as long as the current STA window so that

$$C_L = 10^{-f_\ell \Delta t/12}. \tag{5.14}$$

We declare a detection trigger when the ratio $D = S/L$ of the STA and LTA measures exceeds a pre-specified value.

In the application of this approach to three-component records, the LTA measure is calculated for the total energy E_3 (i.e. the sum of the contributions from the three-components) while STA terms are calculated for the energy on each of the Z, N and E components as well as the horizontal energy H_E, and the total energy E_3. This gives a set of five different different STA/LTA triggers, which can not only be used to detect phases but also give an indication of their character via the combination of triggers which are activated, using the energy relationships for P and S waves we have discussed above.

5.2.2.2 *Adaptive noise filtering*

A useful complement to the phase detection procedure we have just described is to make use of selective frequency filters. We keep track of the local dominant frequency f_ℓ as a function of time, and once a P phase detection is declared we use the dominant frequency f_n in the interval just before the detection to characterise the background noise for the particular set of subsequent arrivals. We then employ real-time filtering to produce two sets of three-component traces:

- a high pass (SP) band with corner frequency $2f_n$ (implemented via recursion equivalent to a fourth order Butterworth filter)

- a low pass (LP) band with corner frequency $0.5f_n$ (implemented via trace smoothing and equivalent to a second order Butterworth filter).

Such a separation into higher and lower frequency components is more flexible than using a fixed SP and LP band response on either side of the microseismic peak. The real-time filter approach can cope with a wide variety of types of noise of different frequency.

The set of five STA/LTA phase detectors are also applied to the high-pass and low-pass filtered data to enhance the possibility of further phase phase detections and also to provide a check on the results from the raw broad-band records.

5.2.3 Phase attributes

Following the detection of an arrival on any of the three-component sets (high-pass, low-pass, broad-band) we endeavour to extract as broad a range of phase attributes as possible, and also utilise the information on phase character contained in the energy triggers.

For every phase detection, the original three component waveforms are analysed. The time of the first onset is picked and an adaptive window around the maximum amplitude on any component close after this onset is used to extract a vector to parametrise the phase - a "phase vector". This phase vector is constructed to represent the average behaviour over a quarter cycle of the dominant period and is then represented in terms of an amplitude component and a unit vector (z, n, e) with $z^2 + n^2 + e^2 = 1$. The apparent azimuth angle ψ can then be estimated by $\tan^{-1}(n/e)$ and an apparent incidence angle in the vertical plane for P waves ι can be found from $\cos^{-1}(z)$. The azimuth estimates are better for P waves (commonly within 10°) than for S waves, because of the complex polarisation pattern for S arising from the interactions of SH and SV waves.

At the beginning of the seismic wave train, the dominant contribution comes from the P wave and then it is not unreasonable to adopt the model of just an incident P wave to try to estimate the azimuth of the arrival. We have found that an effective procedure for a short interval after the first P detection on a broad-band record is to find the azimuthal angle which minimises the ratio of the transverse to the radial component in a weighted average over a time interval of 2-5 s. The weighting scheme is based on the total energy content at each time step. With an azimuth sweep at 10° intervals, it is possible to recover the azimuth to within $10-20°$ when the noise level is reasonable. This provides a useful check on the estimate from the phase vector.

5.2.3.1 Energy, S and P indicators

In addition to the use of energy measures for phase detection triggers it is useful to monitor the vertical energy V_E, horizontal energy H_E and total energy E_3 distributions as a function of time to provide a insight into the physical character of parts of the seismogram which are not sufficiently energetic to provide a phase trigger. We define simple indicators of wavetype character,

$$\mathcal{P} = V_E - 0.35E_3, \qquad \mathcal{S} = H_E - 0.5E_3, \tag{5.15}$$

which will be useful for teleseismic waves and often for far-regional events (particularly for $\mathcal{P}$).

We have noted above that regional P arrivals are associated with significant energy in both the vertical and horizontal planes; such arrivals can be recognised by their simultaneous presence on both $\mathcal{P}$ and $\mathcal{S}$. A similar pattern occurs for other arrivals such as PP out to 30° because, once again, the propagation has been confined to the shallower structure and the incidence angle to the vertical is large.

Such wavetype indicators can be applied to the original records or to the real-time filtered versions, and also directly in conjunction with phase detectors.

5.2.3.2 Removal of free-surface effects

For a set of phase detections for which a stable azimuth can be found we can attempt to remove the influence of the free surface when the surface velocities are known for the particular receiver. This process is built on the simple model of an incident plane wave with a common slowness but a mixture of P, SV and SH character, and can be implemented by using (5.7) with a suitable value of slowness. For distant events the requisite coefficients are slowly varying functions of slowness and so useful results can be obtained by using generic slowness values appropriate to teleseismic P and S waves. The resulting separation into approximate P, SV and SH contributions at each time instant can provide a valuable check on phase attribution. Such a free-surface correction procedure can be applied to the trace as a whole or also to the portions of the waveforms following a detection trigger.

5.2.4 A complexity measure

Small high frequency onsets in the presence of much larger noise are often quite difficult to pick up automatically even with high-pass filtering, whereas the eye uses a change in frequency as an indicator of the presence of a phase. A convenient procedure which can indicate an increase in the complexity of the seismic trace corresponding to the arrival of higher frequencies is to work with a difference trace in time. The length of the trace between successive time samples is monitored as a function of time $\mathcal{L}(t)$ and the envelope of this difference trace gives a good indication of the onset of higher frequency disturbances. The behaviour of the envelope can be tracked automatically by applying an STA/LTA detector approach of the type we have previously described with a trigger when the detector $\mathcal{C}$ exceeds a certain level.

The detection of a change in the complexity of the seismogram can be very effective in picking high frequency arrivals at regional distances which are commonly characterised by a slow decline in frequency content in the coda after a very high frequency onset. The complexity measure also will indicate extraneous sources of noise, due e.g. to vehicles or other cultural activities, which are typically of compact duration and have a different frequency content to natural arrivals.

Table 5.1. *Events displayed*

Event	Latitude	Longitude	Depth	Distance	Azimuth	Magnitude	Location
A	22.56° S	174.97° W	38 km	48.13°	284°	5.6	Kermadec
B	14.05 ° S	167.27° E	184 km	32.57°	267°	6.3	Vanuatu
C	7.08° S	129.47° E	57 km	13.03°	161°	4.8	Banda Sea

5.3 ANALYSIS OF BROAD-BAND RECORDS

We will consider the application of the three-component analysis techniques discussed above to seismograms from the Northern Territory of Australia which is well placed for events at a wide range of distances in azimuths from the seismic belt to the north and east of Australia. We illustrate the techniques for a number of broad-band records from the Warramunga array (WRA) and from portable instruments installed within 300 km of WRA on the Tennant Creek inlier.

This set of broad-band records shows a variety of noise conditions and signal to noise ratios and enables us to examine the value of the processing techniques in a range of environments. The chosen events as listed in Table 5.1 represent propagation paths of rather different character and provide a useful comparison set for phase identification techniques.

We will illustrate aspects of the phase recognition techniques by reference to the different events and will compare the way in which the various techniques contribute to an understanding of the wavefield. The combination of several different measures of wavefield character can then help to provide a guide to the nature of the arrivals along a three-component record.

Even though the nature of the events are known, we apply the range of techniques without specific tuning so that we can judge how they could perform in a real-time application.

5.3.1 Broad-band detection

We illustrate the direct application of detectors based on energy measures to event A in the Kermadec islands (see Table 5.1), a moderate size event (m_b 5.6) at a range of 48° at a time when microseismic noise is relatively high (Figure 5.2).

The upper panel of the display (Figure 5.2a) shows the unfiltered three-component broad-band records, The expected arrival times for the major phases calculated using the *iasp91* tables (Kennett & Engdahl 1991) are indicated by arrows at the bottom of this panel. The second panel (Figure 5.2b) shows the application of a set of five phase detectors based on the adaptive STA/LTA approach applied to energy measures, with windowing parameters derived by averaging the apparent local frequencies of the three-components. These five detectors are based on: the single component energies, (i) V_E, (ii) N^2, (iii) E^2, the energy in the horizontal plane (iv) H_E, and (v) the total energy.

The phase detections based on the adaptive STA/LTA detectors applied to this set of energy measures do an effective job of extracting the main arrivals for this Kermadec event (Figure 5.2b). The *P* onset is very clear, as also is *PP*. As might be expected the steeply incident arrival *PcP* is most evident on the vertical component. The *S* arrival and also the *ScS* phase can be recognised via the increase in energy in the horizontal plane. The long period arrival following *S* is the shear-coupled *PL* phase which is often easier to recognise than *S* itself.

For comparison in the bottom panel (Figure 5.2c) we show the result of applying single-component adaptive STA/LTA detectors to the individual component records. In this case the window parameters are derived from the individual components and so the results for the single component detections $d(z)$, $d(n)$, $d(e)$ are based on different detection parameters than $D(V_E)$, $D(N^2)$, $D(E^2)$. In particular we note the influence of the higher frequencies present on individual components which lead to a plethora of minor phase indications. In consequence when we seek to identify the main phase arrivals we gain by using the three-component measures of Figure 5.2b rather than single-component detections.

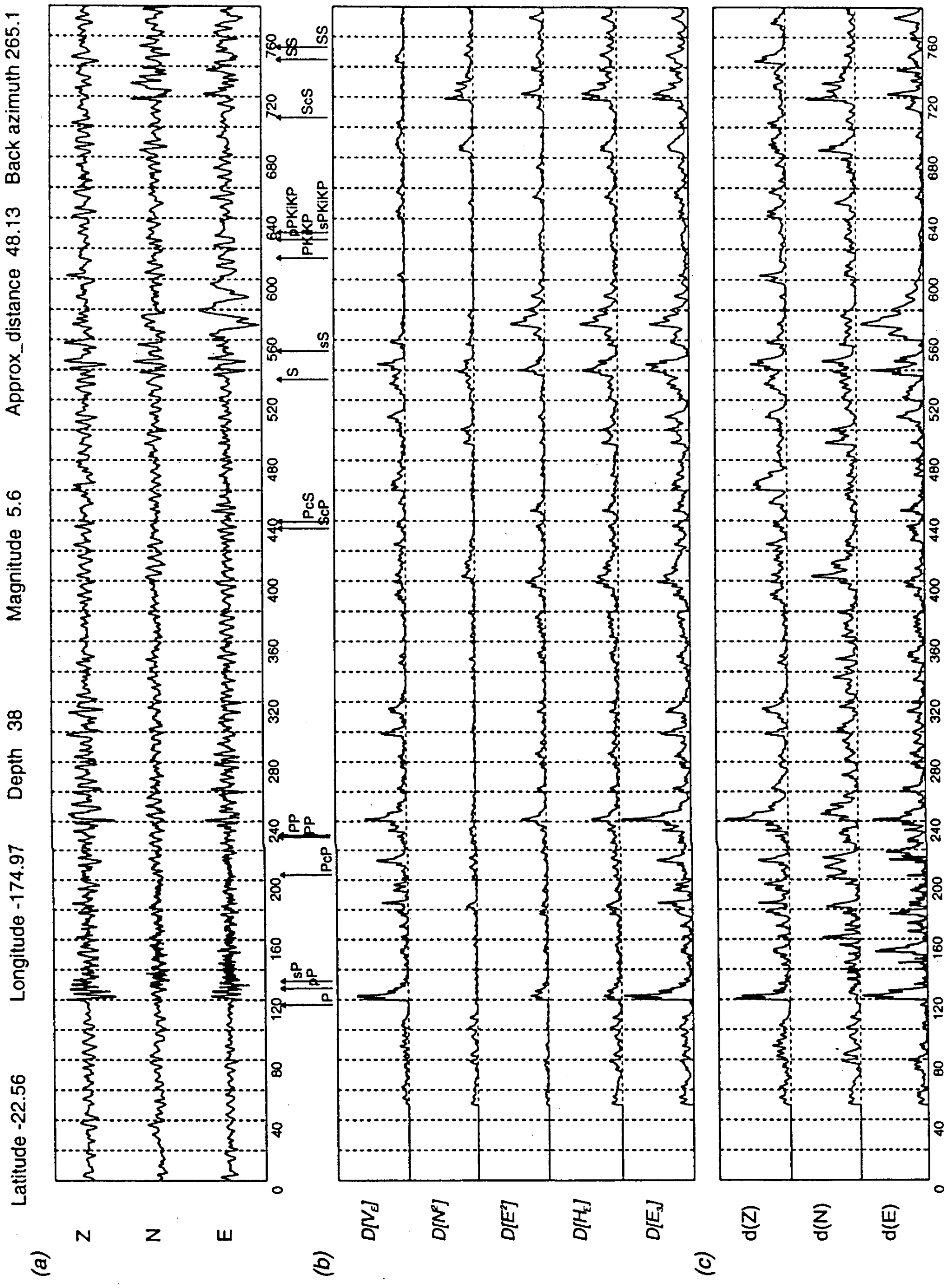


Fig. 5.2. Analysis of broadband records for event A - a shallow Kermadec event recorded in Northern Australia (time scale in seconds): (a) unfiltered broadband records, with indications of the predicted times for phase arrivals using the *iasp91* traveltime tables, (b) phase detectors based on an LTA measure based on total energy and STA measures for various energy contributions from the three-somponent traces, (c) single component STA/LTA detectors applied to each of the three components separately.

5.3.2 Extraction of phase characteristics

In Figure 5.3 we illustrate the extraction of phase characteristics by using noise-adaptive filtering techniques.

The upper panel (Figure 5.3a) once again shows the unfiltered broad-band records together with the arrival times for the major phases predicted from the *iasp91* traveltime tables. The next panel (Figure 5.3b) shows the low-pass filtered traces generated by the noise-adaptive filtering procedure, together with STA/LTA detectors for the three low-pass filtered components (Z,N,E). The arrows indicate where phase detections have been made. For each phase detection the phase vector is extracted as discussed in section 5.2.3; the relevant waveform segments are analysed in Figure 5.3d. In Figure 5.3c the high-pass filtered traces are displayed for each of the three-component traces along with the equivalent STA/LTA detectors. The numbering of the phase detections in Figure 5.3c is linked to Figure 5.3b so that the sequence is in order of arrival time.

The bottom panel (Figure 5.3d) shows the waveform segments corresponding to each detection denoted by L or H depending on the traces from which it is was derived, together with the amplitude and normalised phase vector. The apparent azimuths for each phase detection are displayed with each waveform time interval.

For this Kermadec event the overlap of the S frequencies and the noise makes it difficult to eliminate the noise without major loss of signal in the low-passed records (Figure 5.3b). Nevertheless we can detect the shear-coupled *PL* arrivals [phase 6] and there is a suggestion of converted *P* precursor to *S* with significant energy on the vertical component. The *P* arrival is particularly clear on the high-passed records (Figure 5.3c) and the onset [phase 1] is followed by a detection on the low-passed records [phase 2]. The *PP* phase segment can be extracted from the low-passed record. For shallow events at distances beyond 30° there is usually little high frequency *S* so that the phase extraction from the high-passed records are usually of *P* character. The phase detection [phase 8] on the vertical component does not following weak *S* arrival does not have a clear association with this event.

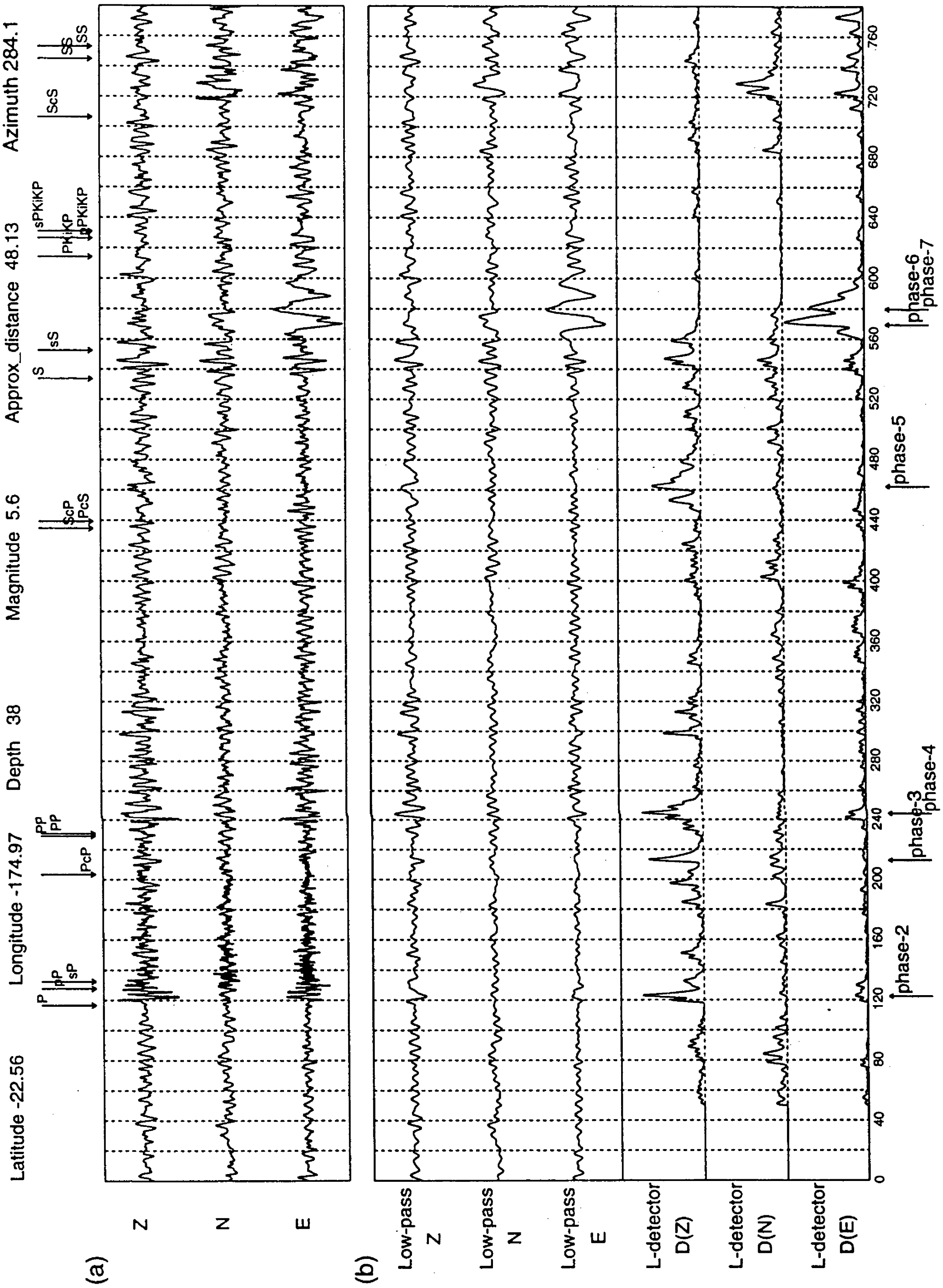


Fig. 5.3. Analysis of broadband records for event A - a shallow Kermadec event recorded in Northern Australia (time scale in seconds). (a) unfiltered broadband records with indicators of the arrivals from the *iasp91* tables, (b) Low-pass filtered records and STA/LTA detections, (c) High-pass filtered records and STA/LTA detections, (d) associated waveform segments and phase vectors.

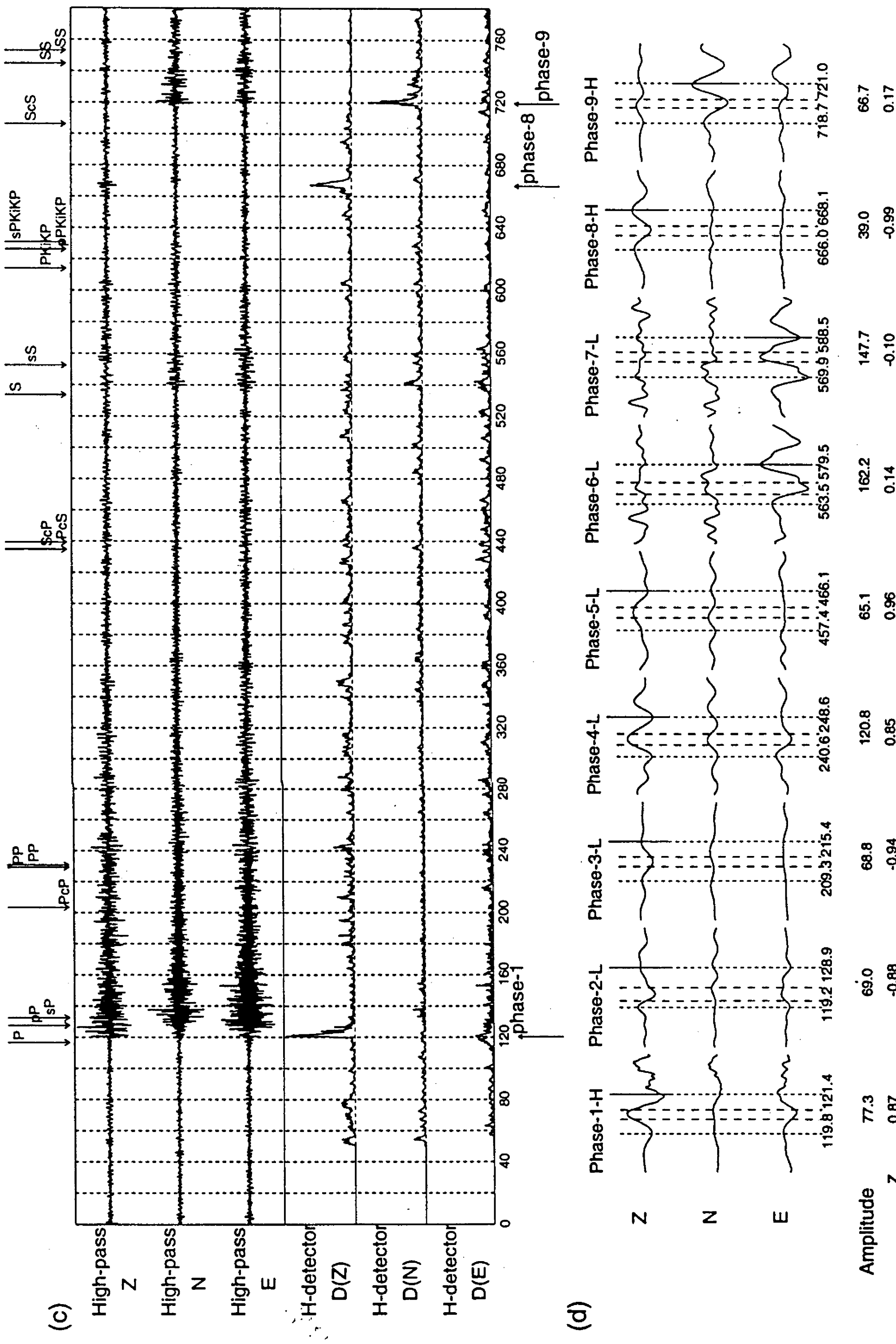


Fig. 5.3. – *continued.*

The waveform segments extracted for analysis following the noise-adaptive filtering are shown in Figure 5.3d, together with the amplitude and normalised phase vector and the estimated azimuth. The longer dashed lines outline the portion of the waveform from which the averaged amplitude and phase-vectors were constructed. The solid line indicates the component on which the detection was made. The extracted phases 1-5 have *P* character with a high proportion of energy on the vertical component, as also does the isolated arrival phase 8. Phases 6,7 and 9 display the energy characteristics expected for *S* waves. The azimuth estimates extracted from the *P* arrival [phase 1] and the shear-coupled *PL* arrival [phase 6] are close to the great-circle azimuth of 284°.

As we have seen from Figure 5.2, we can get effective phase detections directly from unfiltered broad-band records, but if we want to extract further information on the phase attributes, e.g. the azimuth or apparent angle of incidence, we need to remove the main noise component in order to stabilise the shape of the relevant waveform segments.

5.3.3 Energy-related measures for phase character

In Figure 5.4 we show an illustration of the application of energy-related measures of phase character for the three-component wavefield $\mathcal{P}$ and $\mathcal{S}$ for event B (Table 5.1), an intermediate depth event in Vanuatu at an epicentral distance of 32.3°. The upper traces (Figure 5.4a) show the unfiltered broad-band records together with the phase arrival times for the *iasp91* model, and beneath we show traces designed to give a visual indication of the *P* and *S* character on the wavetrain. We can note the benefit of broad-band recording because of the substantial difference in frequency content the high frequency *P* wave and the *S* wave with frequency about 0.25 Hz.

In the middle panel (Figure 5.4b) the P indicator is displayed when the $\mathcal{P}$ measure averaged over a 2 second interval is non-zero, the intensity of tone represents the size of $\mathcal{P}$ and the display is modulated by the square root of the total energy in order to give a measure of the significance of each arrival. The S indicator is displayed

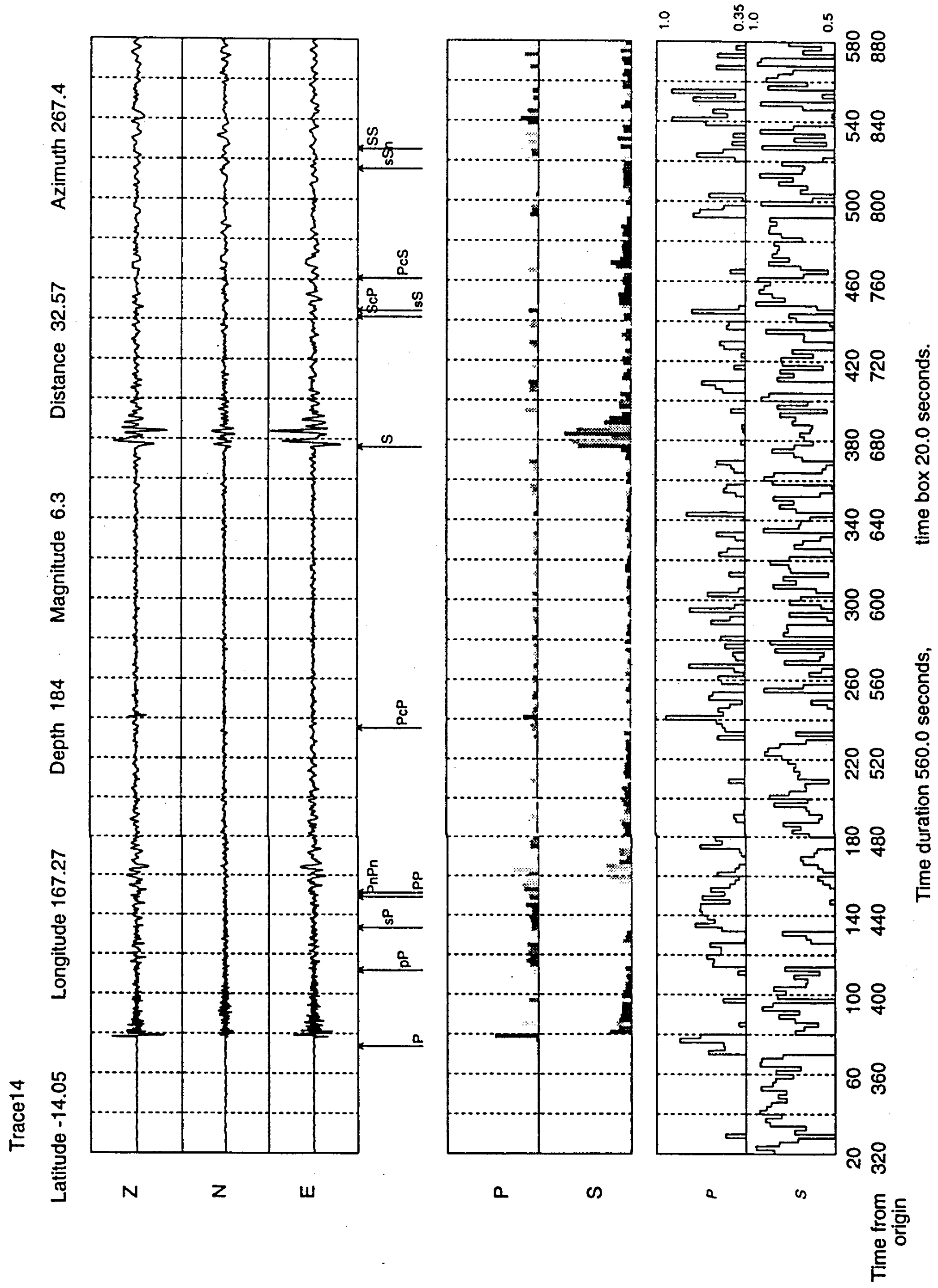


Fig. 5.4. Analysis of broadband records for event B - an intermediate depth Vanuatu event recorded in Northern Australia (time scale in seconds): (a) unfiltered broadband records with indicators of the arrivals from the *iasp91* tables, (b) waveform characterisation using wavetype classifiers P and S with modulation by the amplitude of arrivals, (c) the wavetype indicators $\mathcal{P}$ and $\mathcal{S}$.

when $\mathcal{S}$ averaged over a 2 second interval is non-zero. Once again the intensity of shading indicates the size of $\mathcal{S}$ and the display is modulated by the square root of the total energy to show phase arrivals. The lower panel (Figure 5.4c) displays the direct values of the $\mathcal{P}$ and $\mathcal{S}$ measures as a function of time.

The P and S indicators perform well for this intermediate depth event (Figure 5.4b). The $\mathcal{P}$ indicator is just triggered soon after the onset of S which gives a faint registration on the P trace. It is interesting to note that the high frequency coda immediately following P has a largely S character. The PP arrival is picked up successfully by the P indicator despite a much shallower angle than P - the shallow propagation character can be recognised by the simultaneous presence of $\mathcal{P}$ and $\mathcal{S}$ as in (5.11). The PcP indication is very clear and can also be picked up from the change in frequency content of the vertical component. Even in the S train the P energy indicator draws attention to features which might otherwise be missed such as ScP. The S energy indicator also performs quite well in flagging the major arrivals.

5.3.4 Approximate corrections for free-surface effects

In Figure 5.5 we show a display of estimates of the P, SV and SH wave components in the wavefield for the intermediate depth Vanuatu event used in Figure 5.4, after approximate correction for free-surface effects.

The rotation used is for the great circle azimuth. Two separate sets of estimates for the P and SV wave content using (5.7) are shown

(i) for a typical P wave slowness of 8.5 sec/deg,
(ii) for a typical S wave slowness of 15.0 sec/deg.

These slowness values have been chosen to be suitable for the distance span associated with a range of events from far-regional through to teleseismic. Because the corrections are fairly slowly varying functions of slowness the choices for the generic slownesses are not critical.

The wavefield decomposition (Figure 5.5b) has been reasonably successful for

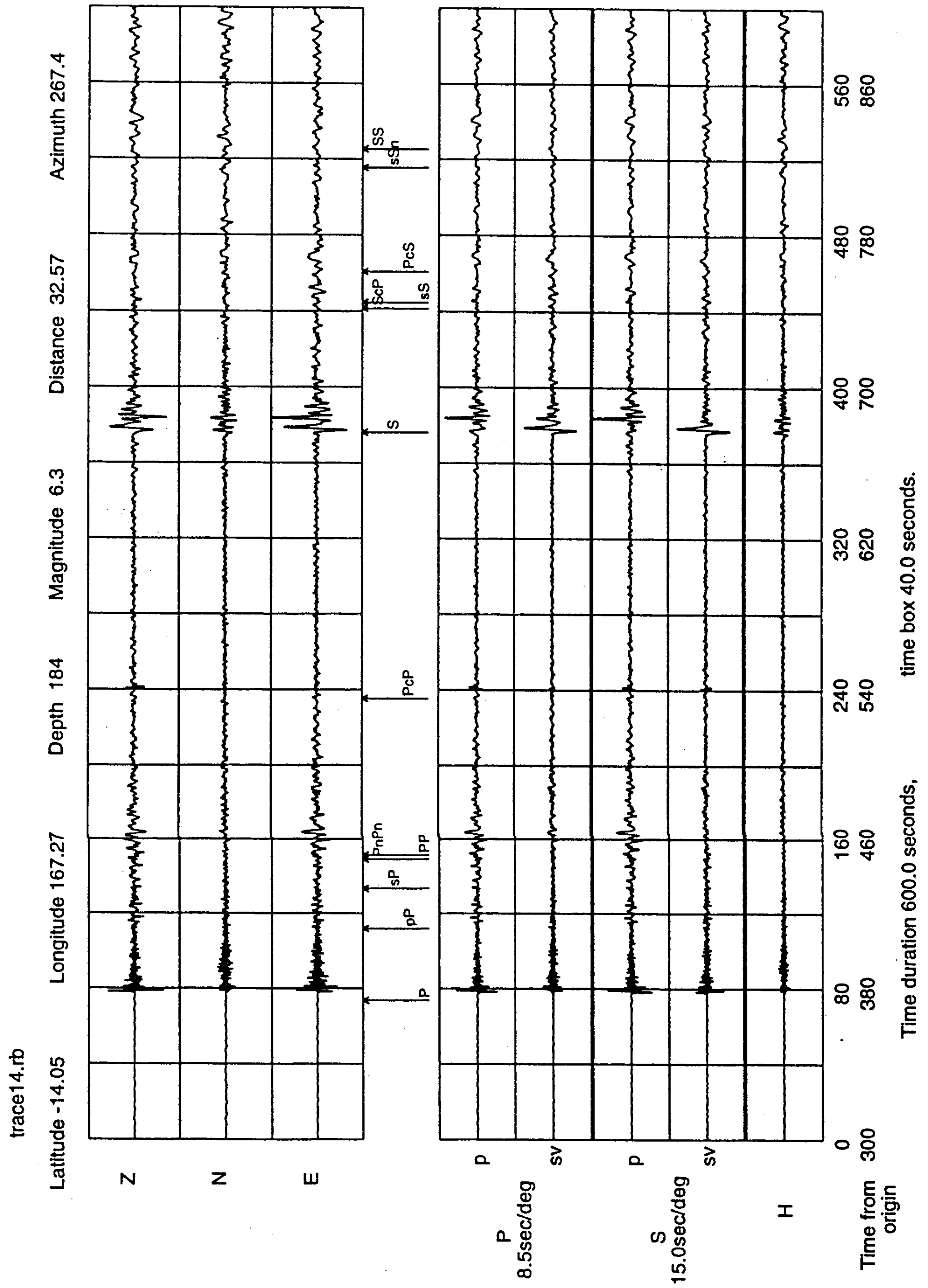


Fig. 5.5. Analysis of broadband records for event B - an intermediate depth Vanuatu event recorded in Northern Australia (time scale in seconds): (a) unfiltered broadband records with indicators of the arrivals from the *iasp91* tables, (b) Approximate wavefield decomposition in P, SV and SH parts, by rotation to the great-circle azimuth and correction for free-surface effects using generic slownesses.

separating the P and SV components especially for such phases as PP, which is prominent whichever of the slowness values is considered. An unexpected result is the separation of the clear S wave arrival into two parts with separate SV and P wave character, which is echoed to some extent in the energy measures; this may represent near-receiver conversion but we note that the SH arrival is also doubled.

5.3.5 Complexity measure

Figure 5.6 illustrates the use of the complexity measure introduced in section 5.2.4 for a regional arrival from the Banda Sea (event C - Table 5.1) in the presence of large amplitude microseismic noise.

The upper traces (Figure 5.6a) display the unfiltered three-component broad-band records in which the P onset can be discerned by the change in frequency content. In the second panel (Figure 5.6b) we display the difference trace in time for each component e.g. $\mathcal{L}[Z]$ representing the length of the trace on each component between successive time samples. The lower panel (Figure 5.6c) shows the application of STA/LTA detectors to the time difference traces to indicate changes in the character of the record, notably by the arrival of high-frequency energy. The phase arrival times from the *iasp91* tables are indicated at the base of the figure. We note that this simple procedure gives a very clear indication of the onset of P without any filtering, and the arrival of the S wave train leads to a second trigger. The considerable residual between the observed arrival and the time estimated from the *iasp91* tables is characteristic of the fast cratonic paths from the Banda Sea region to sites in northern Australia

5.4 DISCUSSION

The phase identification tools that we have discussed in this chapter have been applied to previously recognised events, but give an indication of the power of energy-based measures even when the direction of arrival is unknown. All of the techniques

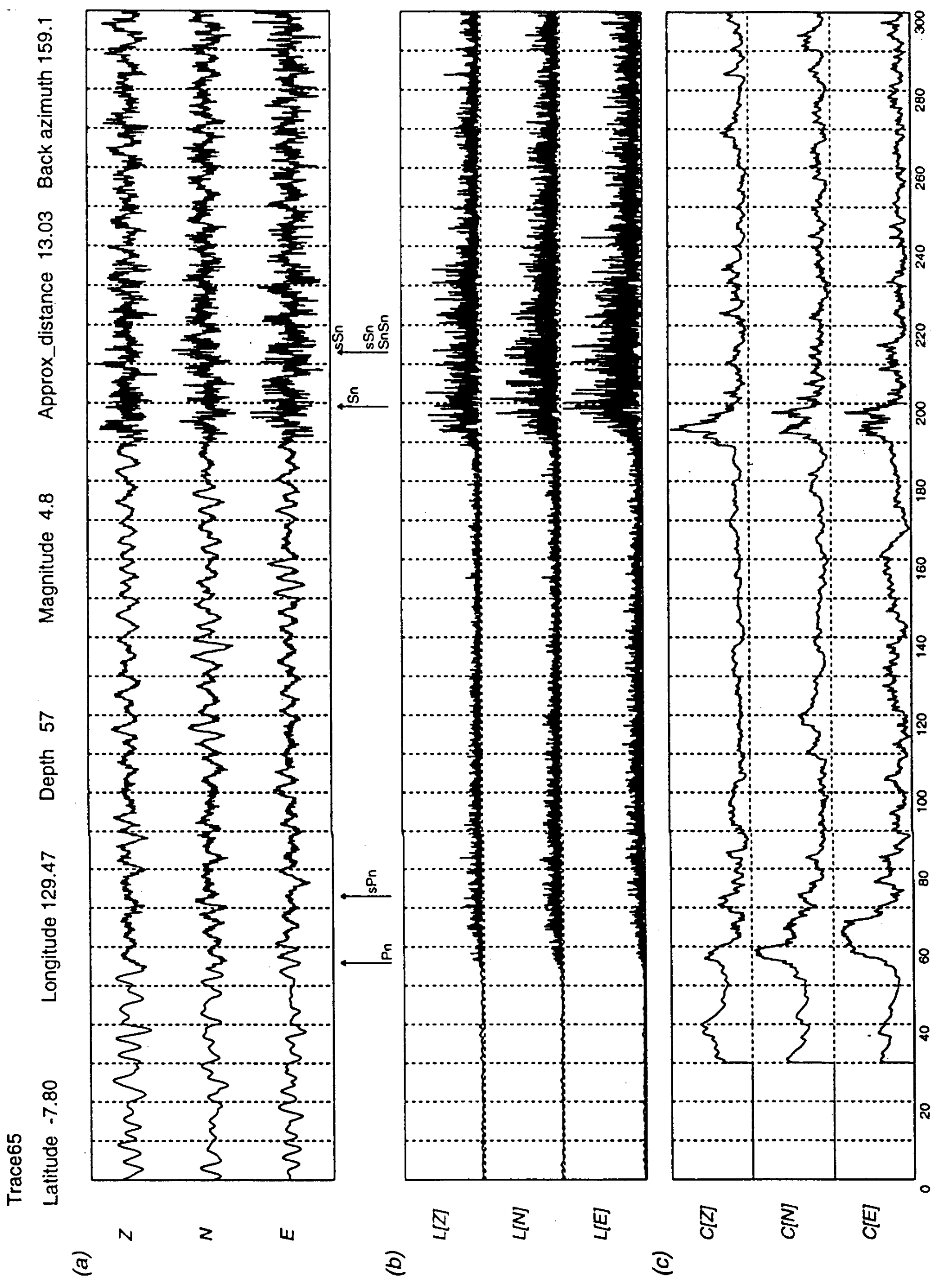


Fig. 5.6. Analysis of broadband records for event C - a Banda Sea event recorded at the Warramunga array in Northern Australia (time scale in seconds): (a) unfiltered broadband records, (b) difference traces based on the length of trace between time samples, (c) Complexity measures derived from the action of STA/LTA detectors applied to the difference traces.

are simple and can be directly applied in real-time as the seismograms are being recorded.

We have illustrated the different techniques with records that span a range of distances, and show the cumulative power of using a number of different criteria, each of which helps to indicate the character of a phase arrival. We have taken advantage of the estimation of the local frequency content derived from the pattern recognition procedure of Tong (1995) which can be extracted as a record unfolds to both give information on the character of arrivals and to guide a variety of adaptive processing steps.

From the application of the various procedures to a wide range of events ranging from regional to far-teleseismic distances we can draw some general observations on the effectiveness of the different methods.

The use of multiple STA/LTA phase detectors based on the energy distribution across three-component records works well except when the main signal is buried in microseismic noise or the coda of a preceding event. In such cases effective phase detection can nearly always be made from the use of the simple complexity measure introduced in section 5.2.4.

The utility of the STA/LTA detections for phase characterisation can be enhanced where the energy detectors are combined with selective filtering to suppress the dominant noise source. The use of adaptive windows for the STA and LTA calculations based on the local dominant frequency gives very good results for broad-band records even when there are large variations in frequency content along the trace. Recognition of the *P* onset and identification of *P* character is normally very good indeed. Later *P* phases can generally be well characterised for events beyond 20°, though good results can be obtained at far-regional distances when the signal-to-noise ratio is above 2:1. Characterisation of *S* arrivals works well for teleseismic events, using the low-pass filtered records. For shorter distances a strong signal is needed before unambiguous identification can be made.

The very simple S and P wave indicators provide a useful basis for phase classification, and do not depend on knowledge of the azimuth. They can be used effectively on the unfiltered broad-band records when the background noise is not too large but can also be usefully applied to either the low-passed or high-passed records. The complexity measures are also independent of the azimuth of the arrival and can be applied directly to individual components or to energy traces. They are very effective indeed at regional and far-regional ranges, but tend to be slightly less effective beyond 40° as the frequency of the P and S arrivals is reduced by the influence of anelasticity on their passage through the Earth. The complexity measure can work well even where there is a very high noise level as we have seen in Figure 5.6, provided that there is some degree of separation between the frequency content of the arrivals of interest and the main energy on the traces.

In many cases a reasonable estimate for the azimuthal direction of the incident P wavefront can be obtained from a P detection and this can then be used to help to explore the character of the wavetrain that follows. We have illustrated the use of generic slownesses to extract information on the relative contributions of P and S waves to the three-component seismograms, as the wavetrain arrives. The use of generic slownesses has proved to be effective for teleseismic events especially when the signal-to-noise ratio exceeds 2:1. When specific information is available about the location of the event it is possible to improve on the technique we have used to generate the wavefield contributions by approximate removal of the effects of the free surface effects. P, SV and SH estimates can be derived using a linear sweep in slowness in time from the value for the P wave arrival through the corresponding slowness for S at the expected arrival time. Such a procedure will certainly misestimate the free surface corrections for steeply travelling waves like ScS but is normally more realistic than the use of generic slownesses as in section 5.3.

In this chapter we have illustrated the use of the various tools as a visual guide to phase identification but their obvious success suggests that it will be worthwhile to begin to develop a set of "rules" for each major class of arrival as a function of epicentral distance. For example, a simultaneous detection trigger of V_E and E_3

is a good indicator of P character for a phase, particularly if the $\mathcal{P}$ indicator at the time of onset is also positive. For S phases a simultaneous detection on H_E and E_3 can be effective indicator linked to the $\mathcal{S}$ measure. At regional distances a detection on the complexity measure can be coupled with phase character indicators applied to the high-pass filtered records to distinguish P and S. Since we do not know in advance what class of arrivals to expect, we need to use a number of different procedures to ensure that we have the best chance of characterising a phase. The process of developing satisfactory phase "rules" will be helped by utilising broadband theoretical seismograms including a full range of arrivals such as in the *complete ray expansion* procedure of Clarke (1993).

6

Automatic seismic event recognition and later phase identification for broad-band seismograms

6.1 INTRODUCTION

The continuous stream of data recorded by a modern high-fidelity seismic station is only useful when the portions of the record associated with different events can be isolated for further analysis. Subsequent interpretation depends on being able to characterise the event generating the pattern of observed arrivals. With a multi-station network an approximate classification of an event in terms of epicentral distance, azimuth and depth can provide a good starting point for refinement of event location (see e.g Kennett 1995). Particularly for a sparse global network, as is planned for the monitoring of a comprehensive test ban treaty, the quality of the initial location estimate is critical to the quality of event characterisation. In such a context it is very important to be able to determine whether the arrivals detected at different stations are likely to have arisen from a single event or from two or more geographically distinct events in a limited time period.

The procedure described here is designed to recognise the pattern of arrivals associated with an event from a single three-component broad-band record, and to provide a preliminary estimate of the epicentral distance, depth and azimuth to that event. The method builds on the work in the previous chapters 4 and 5 where we have introduced an automatic system for seismic phase detection and analysis which produces, in real-time, a stream of phase attributes characterising each detected phase. In this chapter we show how these attributes can be exploited with the aid of a new automatic reasoning method (the "assumption tree" method) to combine the information from many phase detections to form event segments which can then be characterised in terms of the properties of the source.

6.2 RECOGNITION OF A SEISMIC EVENT

If we are to be able to characterise portions of a seismic data stream in terms of events we must be able to:

(1) recognise related seismic phases as components of an event, and
(2) provide possible interpretations of the nature of the phases based on their observed attributes.

As a working definition of a seismic event we have used the expectation that there will be a number of P phases followed by a number of S phases during a certain time interval (say, 20 minutes). The procedures suggested in chapter 5 provide a means of recognising the character of teleseismic phases in terms of P and S arrivals using, for example, the $\mathcal{P}$ and $\mathcal{S}$ detection scheme based on the relative distribution of the energy in the waveform between the vertical and horizontal components. However, such a measure of waveform character has to be supplemented by other attributes to suggest possible identifications for the seismic phase. In chapter 5, we have shown how a waveform segment associated with a phase detection can be specified in terms of a set of parameters based on a model of the seismic wavelet. These basic parameters can then be used to extract five attributes for each detected phase, e.g. for the ith phase we would have:

(1) t_i - the arrival time
(2) a_i - the amplitude
(3) ν_i - the local frequency

(4) ϕ_i - the azimuth in the horizontal plane

(5) ψ_i - the angle of incidence to the vertical

We can represent the stream of detected phases as a collection of specifications of phase attributes,

$$..., F_i(t_i, a_i, \nu_i, \phi_i, \psi_i), F_{i+1}(t_{i+1}, a_{i+1}, \nu_{i+1}, \phi_{i+1}, \psi_{i+1}), ... \quad (6.1)$$

from which an event is to be constituted. Thus if there are N_p phases detected for a presumed event we have a set of $5N_p$ features (5 attributes per phase) as the input to the classification process.

The pattern classes we have to recognise are seismic events, and since different epicentral distance or depth can produce totally different sets of observed phases, we will endeavour to classify the events by range and depth. In order to provide a reasonably comprehensive coverage of possible wavefields we work with 90 possible distances (dividing the range from 0° to 180° into 2° intervals) and a selection of 4 depths (0 km, 100 km, 300 km and 600 km) as in the *iasp91* travel-time tables (Kennett 1991). In this way, with 360 pattern classes, we are able to provide a characterisation into shallow, intermediate and deep events and to provide a reasonable sampling of the dependence on epicentral distance.

The problem of event recognition can then be viewed as a classification problem in a $5N_p$ dimensional vector-space, with 360 possible destination classes. In order to understand the complexity of this problem we have to recognise that the phase summaries extracted from the observed data may be distorted due the presence of noise or incomplete (e.g. the amplitude may lie below the detection threshold). For each of the 360 pattern classes we can generate complete information for each of the phase attributes for a broad range of possible phases using the *iasp91* model. However, the N_p phases will frequently represent only a subset of the possible phases and so we will be faced with a classification based on incomplete data.

One obvious approach to the event classification problem would be to set up 360 model patterns and compare any set of observed phase attributes with all these model patterns to find the best match. The phases associated with a particular source depth and distance will have different degrees of importance; thus, some phases must be observed for this combination to be identified whilst others might or might not be observed. The attributes for each phase must also be regarded as having differences in significance, since e.g. the relative amplitudes of phases will vary depending on the source mechanism. When the comparison between the phase attributes and the values for a model source is made, the weights and required level of fit for different attributes need to be flexible. This means that there will be of the order of $3N_p \times 360$ features to be individually considered, i.e. typically several thousand cases. In addition any modification to the system (e.g. to increase the sampling of the model space) will impose a reconsideration of an already complex algorithm.

Instead, we employ a new strategy based on "assumption trees" which allows the progressive inclusion of seismological expertise to isolate the most probable combination of epicentral distance and source depth for the available phase attributes. This procedure provides a more graceful and efficient procedure than exhaustive matching.

6.3 THE USE OF ASSUMPTION TREES

6.3.1 Inclusion of seismological expertise

In order to reduce the computational effort required to find the appropriate source distance and depth, we try to build in as much seismological expertise as possible.

In particular we need to take account of the way in which the patterns of seismological phases vary with epicentral distance. For example, up to 84° the major observed phases are P and S, but beyond 84° SKS precedes S and appears as the first phase with an "S-wave" character. In other cases significant information is highly distance specific; thus $PKKP$ has a focus near 120°, whilst near 132° SKP and PKS are usually the largest phases in the early part of the seismogram.

In order to identify phases, we need information about the epicentral distance.

Table 6.1. *P-S models and related ranges*

P-S identity	Time interval range (minutes)	Distance range (°)
P - S	2.18 - 10.43	12 - 85
P - SKS	9.37 - 10.65	82 - 99
P_{diff} - *SKS*	9.39 - 10.69	100 - 129
PP - SKS	3.96 - 7.22	82 - 129
PP - PS	9.20 - 10.14	104 - 125
PKP - SKS	5.84 - 7.21	114 - 143
PKP - SKKS	7.97 - 12.66	126 - 180
PKP - SS	20.05 - 27.66	136 - 180
PKIKP - PKS	3.35 - 3.60	126 - 141

On the other hand, to obtain the estimation of epicentral distance from a single record we have to identify at least two phases - we can then use the differential time between the identified phases to determine the distance. In order to make the process tractable we make a set of hypotheses about the identities of two observed phases. Thus we select the first "*P* phase" and the first "*S* phase" in a sequence as the key phases. The phase character assignment will be based on measures such as $\mathcal{P}$ and $\mathcal{S}$ as discussed in chapter 5. We then try to associate the pair of phases with the most commonly observed combinations of *P* and *S* phases for different epicentral distance ranges as listed in Table 6.1. The dependence of the differential times for these *P-S* combinations on epicentral distance and depth are illustrated in Figure 6.1. We can see that most of these pairs of phases have sufficient variation in differential times to provide some constraint on range and depth. However, there will often be ambiguity in the appropriate combination of epicentral distance and depth associated with a particular differential time which can only be resolved by the matching of other phases in the record.

We will employ this set of nine *P-S* pairs as the first stage of hypothesis testing as indicated in Figure 6.2, which represents the classification hierarchy. Since there is some degree of overlap between the operative distance intervals for a number of the *P-S* pairs the tree structure in Figure 6.2 is not a strict classification. The number

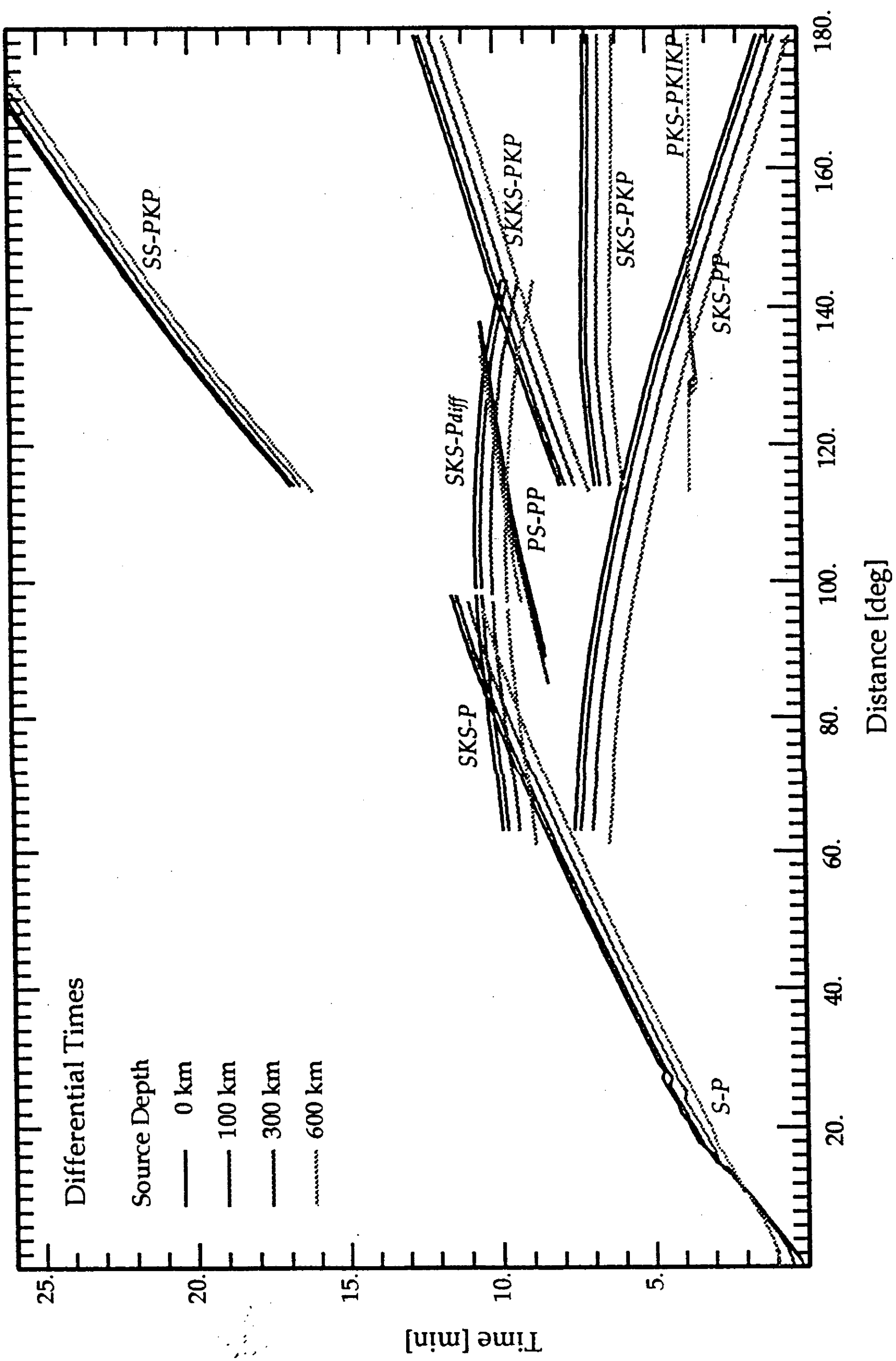


Fig. 6.1. The distance and depth dependence of *P*-*S* differential times.

of leaves on the tree is actually more than 360, whilst the total number of different pattern classes that they represent remains 360. In the following section we develop the new strategy of an assumption tree based on the structure illustrated in Figure 6.2.

6.3.2 An assumption tree

Before we introduce our "assumption tree" we will briefly review the "decision tree" strategy which is frequently used in pattern classification. Decision trees use a sequential decision making strategy to classify a vector of features. At each branching point in a decision tree the branch with which the feature vector will be associated is determined by the evaluation of a test. A commonly used test is the hyperplane test (L. Breiman et al. 1984) in which the feature vector $\mathbf{x}$ is tested with respect to a hyperplane in model space. The branch is then determined by which side of the hyperplane $\mathbf{x}$ lies, leading to a binary tree structure. Building a decision tree requires the construction of appropriate hyperplanes at the branching points.

In our system a multi-branch tree is constructed; this tree grows its branches as a set of assumptions are tested and so we call it an "assumption tree". Such an assumption tree is designed to deal with a classification problem in which there are a number of factors to be identified (see Figure 6.3).

At each level of the assumption tree, a set of assumptions about a single factor are tested. The set of stems springing from the same branching point will cover the full range of possible hypotheses connected with this factor. Every branching point (node) in the tree has an associated package of information. At the root level the information consists of just the observed data. Every new node in the assumption tree will inherit all the information of its parent node, and include extra information. The new information includes the premise which has just been made to produce this node, and the inferences based on the particular assumption drawn from the summary of expert knowledge held in a separate knowledge bank. The information package at a node is examined to look for any contradictory results. If

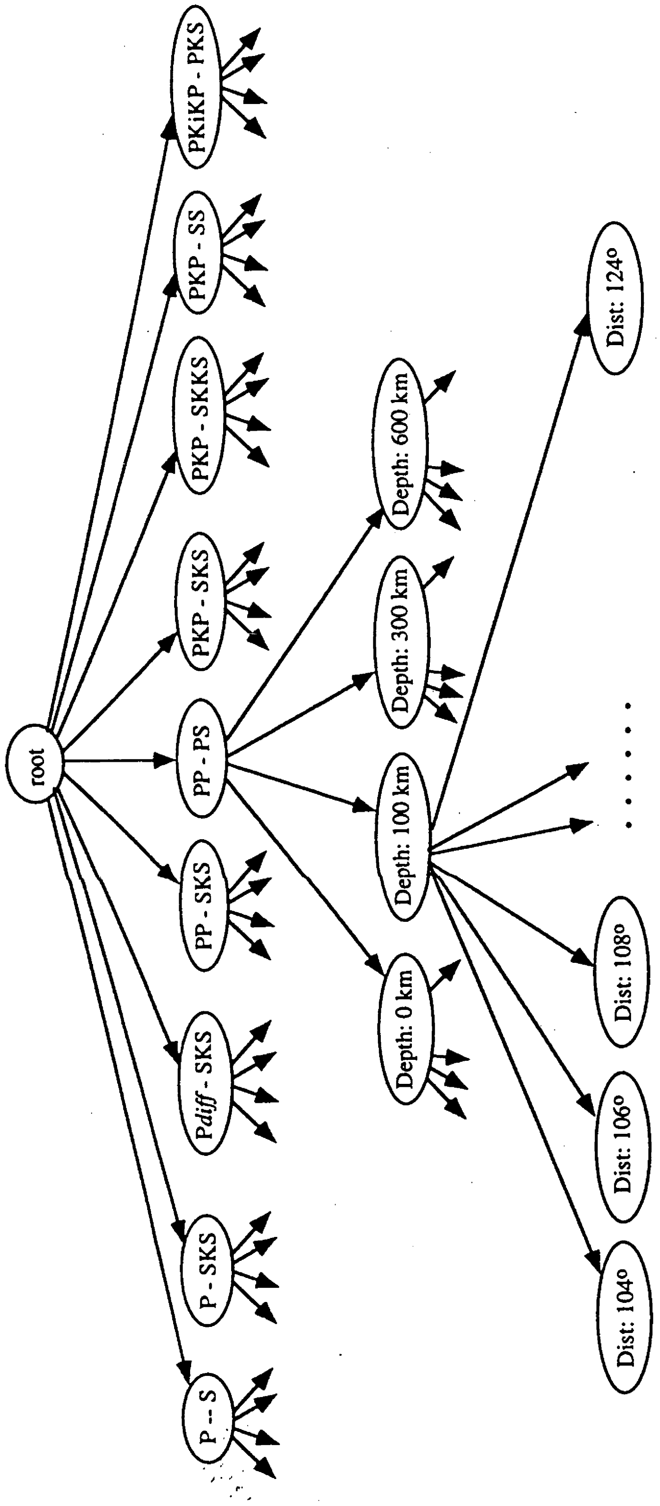


Fig. 6.2. Seismic event classification hierarchy.

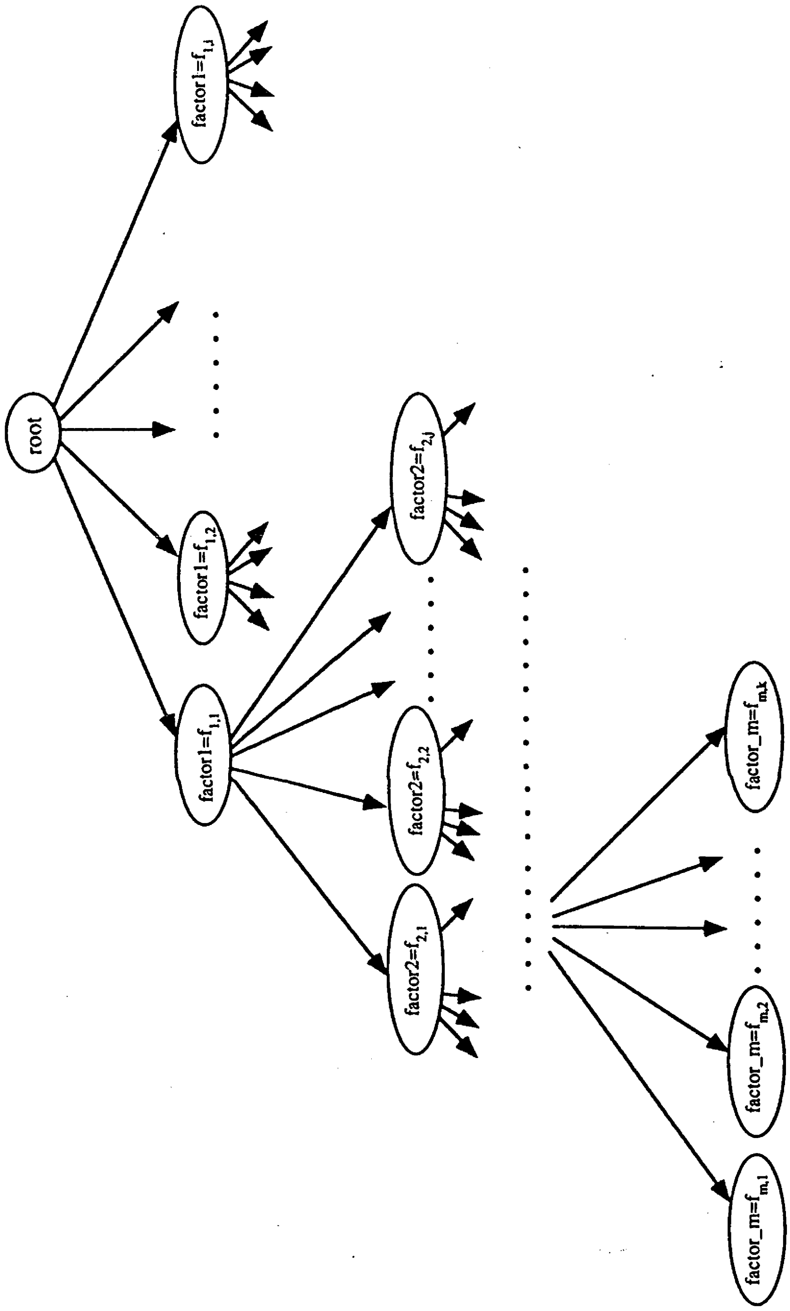


Fig. 6.3. The structure of an assumption tree.

any contradiction is found, the branching associated with the current node will be terminated. Otherwise the branching process will continue from the node and allow the evaluation of further hypotheses. At any node, the new information introduced from the knowledge bank may limit the choices for other currently undetermined factors. As a result the pattern of branching at a node will depend on the current information package. When no further assumptions are to be tested we reach a leaf of the tree representing a solution based on a set of hypotheses about the factors describing the data.

The difference between an assumption tree and a decision tree is that at each step the decision tree is constructed using a specific decision criterion (knowledge), whereas in an assumption tree the construction depends on the testing of hypotheses. In the decision tree strategy expert knowledge is used to determine which branch to choose, while in an assumption tree such expert knowledge is applied to terminate a branch associated with a particular premise. A decision tree is pre-designed, so that adding a new decision criterion (additional knowledge) requires modifying the decision tree and hence modifying the pattern classification program with other potential side-effects. For the assumption tree procedure employed in this chapter, adding an additional item of expertise requires the addition of an item in the separate knowledge bank; this process is much simpler and safer than reconstructing the program.

6.4 IMPLEMENTATION OF THE ASSUMPTION TREE PROCEDURE

6.4.1 Phase detection and feature extraction

In chapter 5, we have described a number of techniques for seismic phase feature extraction particular for three-component records. Also described is a new technique based on a complexity measure which can detect an increase in frequency content for phases whose signal-to-noise ratio is less than one.

In the application of these techniques to the problem of event identification we try to ensure that the maximum opportunity is made to extract the relevant phases. The first stage uses a detection system based on the complexity measure to detect the approximate position for a phase package. Then the original seismogram is divided into a number of segments which contain different phase packages for each of which the local frequency can be estimated as described in chapter 5. Each seismogram segment is then filtered into a high-frequency trace and a low-frequency trace with filter parameters based on the local phase frequency. Finally, a set of detections are made on both the low-pass and high-pass filtered traces using adaptive STA/LTA detectors which combine energy measures on one or more components of ground motion with a local frequency analysis. To simplify the input for event identification, we choose to detect only the first phase where there are complicated phase packages.

In order to be able to gain reliable information about a phase we want not only to detect the arrival but also to characterise the phase in terms of a phase vector comprising a summary of the information in the arrival. By filtering segment by segment we avoid problems associated with a wide variation between the frequencies on the three-components and by using a common filtering for all three-components improve the reliability of the phase vector information.

6.4.2 Preprocessing - separating relevant phases

In chapter 5, we have shown that P or S character can be assigned to a detected phase based on the relative strength of the energy on the vertical component compared with the total energy on all three components. For a teleseismic event, a detected phase which has more than one third of its total energy on the vertical component can be recognised as a "P phase", while otherwise it will be assigned as an "S phase".

Since an event on a seismic record can be recognised as a number of P phases followed by at least one S phase and possibly a number of other P or S phases within a 30 minute window, a seismic event can be formally represented as:

$$\langle\text{an event}\rangle \longrightarrow \langle\text{first P phase}\rangle\langle\text{P phase}\rangle^*\langle\text{first S phase}\rangle(\langle\text{S phase}\rangle|\langle\text{P phase}\rangle)^*,$$

where the notation $(X)^*$ means that X may appear many times (including zero), while $X|Y$ means "either X or Y ".

This event definition is used to assemble a set of detected phases (and their associated parameters) into a list which will be used as the basis for event classification and further phase identification.

6.4.3 Seismic event interpretation

The system of event processing is based on an assumption tree with a similar structure to that illustrated in Figure 6.2. The assumption tree grows at run time as we test a set of hypotheses about the observed seismic phases, branches that lead to contradictory results will be terminated and not grow to leaves representing a viable event classification.

6.4.3.1 Stage 1 - choice of P-S pair

The initial information package comprises the set of information on the detected phases at the root node. Nine stems are then grown to child-nodes representing different assumptions about the identity of the combination of first P and first S phases (as indicated in Figure 6.2). The consequences of these assumptions are then tested against the expert knowledge held in the separate knowledge base.

If the assumption about the P-S phase pair is true, the event must lie within the distance range for this combination (see Table 6.1), and the differential time between the first P and the first S must lie inside the possible P-S time interval. The possible distance range is thereby reduced from 0° – 180° to the distance range appropriate to the particular choice of phases. The observed P to S differential time can then be compared with the expected range for this phase combination. If the observed differential time does not lie inside the expected range, we have a contradiction of the assumption about the phase pair, and the branch of the assumption tree is terminated. The phase pairs in Table 6.1 have some overlap in differential times so

Table 6.2. *P-S pairs and related feature properties for earthquakes*

P-S identity	Expected features
P - *S*	The frequency of *S* should be lower than the frequency of *P*; *P* should not be too steep (as *PKP*); there should not be any low frequency S phase before the *S*.
P - *SKS*	*SKS* should not be a high frequency phase
P_{diff} - *SKS*	*P* has a smaller amplitude compared with *SKS*
PP - *SKS*	*PP*: long period and shallow incidence
PP - *PS*	*PP*: long period and shallow incidence
PKP - *SKS*	*PKP*: short period and steep incidence
PKP - *SKKS*	*PKP*: short period and steep incidence
PKP - *SS*	*PKP*: short period and steep incidence; *SS*: long period
PKIKP - *PKS*	*PKIKP*: short period and steep incidence

The specification "steep" or "shallow" depends on the apparent angle of incidence; the smaller the incidence angle, the steeper. *PKP* and *PKKP* are observed for very distant events for which the incidence angle is very steep, normally less than 16°. Thus, an observed phase should have an incidence angle smaller than 16° to be matched with *PKP* or *PKKP*. The value of 16° has been obtained by trial and error, and is subject to further adjustment.

that at most 5 possible assumptions about the *P-S* phase character can survive this test.

At this stage we can apply additional seismological information to test the different premises about the *P-S* pair. The properties which have employed are tabulated in Table 6.2.

6.4.3.2 Stage 2 - the depth of the event

At this stage, every surviving first generation node grows four new stems to child-nodes (see Figure 6.2), where each child-node is associated with a different assumption about source depth. Based on the assumptions about the identity of the *P* and *S* phases and depth and the observed *P-S* differential time we are able to isolate a single possible epicentral distance. Nine differential-time tables (one for each of the *P-S* pairs) connect the *P* to *S* time interval with epicentral distance. A search in the appropriate table will return the distance which corresponds to the smallest difference between the theoretical and the observed differential time.

Since we find the distance by table search we do not implement the third generation assumptions for distance (indicated in Figure 6.2), and as there are no more assumptions to test, we have reached a leaf of the assumption tree.

With the estimates of the depth and distance, we can compare the patterns of observed phases with those expected for the nine different assumptions on the *P*-*S* pairs. As described above, at the most 5 models for the *P*-*S* pair can survive from the first generation. Once four possible depths are included there would be at most 20 live nodes in the second generation. Since 20 is not a large number, the analysis procedure accepts all the second generation nodes as possible interpretations, and then we try to assess the likelihood of each interpretation rather than look for contradictions.

6.4.3.3 Identification of other frequently observed phases

So far we have used seismological information on the likely character of the two main phases (the first *P* phase and the first *S* phase) which are required to form an event pattern. For a choice of source distance and depth, we can make use of a set of "other frequently observed phases" as summarised in Table 6.3, the arrival times and other properties can then be calculated.

Then by comparison between the sequence of observed phases (expression 6.1) and the phase predictions for the assumed distance and depth we may be able to identify phases, and further the presence of expected phases can provide support for the assumed depth and distance.

The specific choice of *P*-*S* pair and the other expected phases constitute the information set to be compared with the observed set of phases. We therefore need to match the observations against the expectations for the proposed distance and depth. For each expected phase we have an arrival time determined by the phase identity; and each detected phase has a set of observed phase attributes which includes the arrival time. The first *P* and first *S* phase in the observed data have

Table 6.3. *Expected phases for choice of P-S pairs, for a certain distance and depth ranges*

P-S identity	Other expected phases
P - S	Deep event: depth phases *pP, sP*; 25° - 65°: *PcP, ScP, PcS*;12° - 39°: *ScS*; 31° - 79°: *PP*.
P - SKS	For all: *PP*; 95° - 99°: *PKKP*.
P_{diff} - *SKS*	For all: *PP*; 115° - 125°: *PKKP, PS*.
PP - SKS	115° - 125°: *PKKP, PS*; 95° - 99°: *PKKP*; For all: *SS*.
PP - PS	For all: *PKKP, SKS, SKKS*.
PKP - SKS	Deep event: depth phases *pPKP, sPKP*; For all: *PP*; 135° - 143°: *SKKS*; 115° -125°: *PKKP, PS*; 125° - 143°: *PKS*; 114° - 131°: *SS*.
PKP - SKKS	Deep event: depth phases *pPKP, sPKP*; For all: *PP*; 125° - 139°: *PKS*; 137° - 180°: *SS*; 159° - 180°: PKP_{ab}.
PKP - SS	Deep event: depth phases *pPKP, sPKP*; 141° - 180°:*PP*; 159° - 180°: PKP_{ab}.
PKIKP - PKS	Deep event: depth phases *pPKIKP, sPKIKP*; For all: *PP, PKKP, SKKP*; 137° - 143°: *SS*.

already been associated with two phases; we can therefore establish a simple time mapping between the observed data and the expected arrivals. We search among the observed arrivals for candidates for an expected phase whose arrival time lies within a predefined tolerance of the expected time. The observed attributes of the phase are then compared with a summary of phase properties in the knowledge base (see Table 6.4). If there is no match then that hypothesis for the phase identification is rejected, but if more than one phase match is possible, the one with the smallest time difference between observed and expected times is selected.

Table 6.4. *Expected properties of other frequently observed phases*

Phase identity	Expected features
pP, sP	*pP, sP* are similar to *P* both in frequency and incidence.
PcP, ScP	*PcP, ScP* are high frequency P phases which is steeper than *P*.
PP	*PP* is shallower than *P* and lower frequency.
PKKP	*PKKP* is a steep and high frequency P phase.
PKP	*PKP* is a steep P phase.
ScS	*ScS*: high frequency steep S phase, strong on tangential component.
PcS	*PcS*: high frequency steep S phase.
PS	*PS* is shallow and low frequency S phase.
SS	*SS* is a low frequency S phase.
SKS, SKKS	*SKS, SKKS* can not be high frequency.

6.4.3.4 Estimation of likelihood of hypotheses

In order to estimate the likelihood of any particular interpretation, we consider both the observed data and the expectation for a particular combination of *P*-*S* pair, distance and depth. When we compare the two sets of phases we have three different cases: matched phases, unmatched phases in the observed data, and unmatched phases in the expected data.

For every matched phase, a time error e_i is used to represent the mismatch between the corresponding observed and expected phases. We set

$$e_i = \frac{\mid t_o - t_e \mid}{t_s - t_p}, \tag{6.2}$$

where t_o is the phase time in the observed data; t_e is the phase time for the expected phase; t_s is the observed time of the key S phase; and t_p is the observed time of the key P phase. For every unmatched phase (both in the observed data and the expected phase list), we set $e_i = 1$.

In order to provide an overall measure of the quality of the match between the observed phases and those expected for the particular distance and depth, we weigh each observed phase with its amplitude (i.e. we place more weight on large arrivals). For those expected phases which have not been successfully matched with any observed phase, the smallest amplitude of the matched components is applied as the weight. We denote the weight for each phase as w_i, and then define the misfit measure δ as

$$\delta = \frac{\sum_{i=1}^{n} e_i w_i}{\sum_{i=1}^{n} w_i}. \tag{6.3}$$

where the sum is taken over n, which is the number of phases in all the three cases.

6.4.3.5 Details of the implementation

We take a pragmatic approach to the selection of phases for the key *P*-*S* pair since we have used the most commonly observed phases; rather than insist on the "first" arrival of particular type we look for a prominent arrival, especially for *S*. After the detection of a "first" *P* phase, we choose the first *S* phase encountered in the processing scheme as the candidate for the key *S* phase. However, if a second *S* phase occurs within 5 minutes without any intervening phase and is larger, it will replace the first candidate.

In general, as pointed out in chapter 4, a low-passed seismogram gives a better definition of phase attributes than the corresponding high-passed seismogram. Therefore, when we select the two key phases, we choose low-frequency detections.

We use the azimuth estimate for the key *P* phase as the azimuth of the whole event. For very distant events from 110° upwards the first *P* arrival is usually somewhat weak and the resulting azimuth estimate is somewhat doubtful.

6.5 EXAMPLES OF EVENT RECOGNITION

The event-recognition system described in the previous section has been applied to a wide range of observed seismic records with considerable success. We have selected a set of four examples which represent events from different distances and depths.

We use seismic records from portable broad-band instruments deployed in the Northern Territory of Australia. The data were recorded on Reftek 72A-07 disc recorders (24-bit resolution) with Güralp CMG-3ESP seismometers (flat to ground velocity from 0.03 - 30.0 Hz) with Omega timing. The sampling rate was 25 samples

Table 6.5. *Events used for illustration*

Event	Year	Day	Time	Station	Distance[°]	Depth [km]	Azimuth[°]	Mb
A	1994	281	21:45:55.8	SC03	15.4	17	175	6.4
B	1994	220	21:15:36.6	SC01	52.0	122	133	6.0
C	1994	244	15:20:18.8	SC08	113.8	10	260	6.6
D	1994	160	00:46:55.9	SC09	137.1	631	213	7.0

per second. Such data is of good quality but not of observatory standard and so represents a useful test for both event recognition and phase identification.

The four events we use for illustration are listed in Table 6.5. We note that the event locations for these events as given in the table had been determined using information from many stations with a broad azimuthal coverage whereas our estimates are based on just a single three component record. The time listed in the table is the beginning time of the record from which the automatic system is applied.

In every case, the process of phase-detection and feature-extraction is represented in a sequence of display panels:

(a) The original broad-band record is displayed together with the arrival times for the major phases predicted from the *iasp91* travel time tables as a reference for the phase interpretation process.

(b) The STA/LTA detectors generated for the three low-pass-filtered components (Z, N, E). The arrows indicate where phase detections have been made. For each phase detection the phase features are extracted as described in chapter 5. The relevant waveform segments and the extracted phase attributes are displayed in (d).

(c) The STA/LTA detectors on the three high-pass-filtered traces are displayed, along with the arrows which indicate phase detections. The numbering of the phase detections in panel (b) and panel (c) is linked to panel (d), so that the sequence is in order of arrival time.

(d) The last panel shows the waveform segments corresponding to each detection, together with the amplitude and and other attributes of the phase (see chapter 5). In the waveform segment, the time interval between the two labeled times represents the assigned period of the detected phase.

The figures for the phase detection processes are accompanied by a summary table of the stages in the event recognition and phase identification procedure. This table is organised in terms of the four major steps in the analysis

Step 1: identification of the key *P* and *S* phases

Step 2: comparison of the differential time between the key phases with the expected times for *P*-*S* pairs.

Step 3: checking of expected *P*-*S* pair phase attributes

Step 4: matching of observed phases against expected patterns for different models of event character.

The final step is selection of the combination of distance and depth which leads to the least misfit between the observed and expected phases.

6.5.1 Event A – Halmahera (Indonesia)

The event shown in Figure 6.4 is a regional event at the distance of 15°. In panel (a), the original seismic record is dominated by *Pn* and *Sn* packages. Later phases are mixed with surface waves. Since the *P* arrival shows little energy on the *E* component, the seismogram is almost naturally polarised, i.e., the *E* component is the tangential component which reveals *SH* wave arrivals.

In panel (b) and (c), the *Pn* and *Sn* arrivals are detected on both high-frequency and low-frequency sets. Since *ScS* is strong on the tangential component, the high-frequency set *E* component clearly reveals the *ScS* arrival (see [phase-5] in panel c). The detections [phase-1] and [phase-2] occur at almost the same time on the two differently filtered sets, so also do [phase-3] and [phase-4] The automatic system uses the attributes for the phase on the low-frequency set from each pair, we therefore choose [phase-1] as the key *P* phase and [phase-3] as the key *S* phase for the differ-

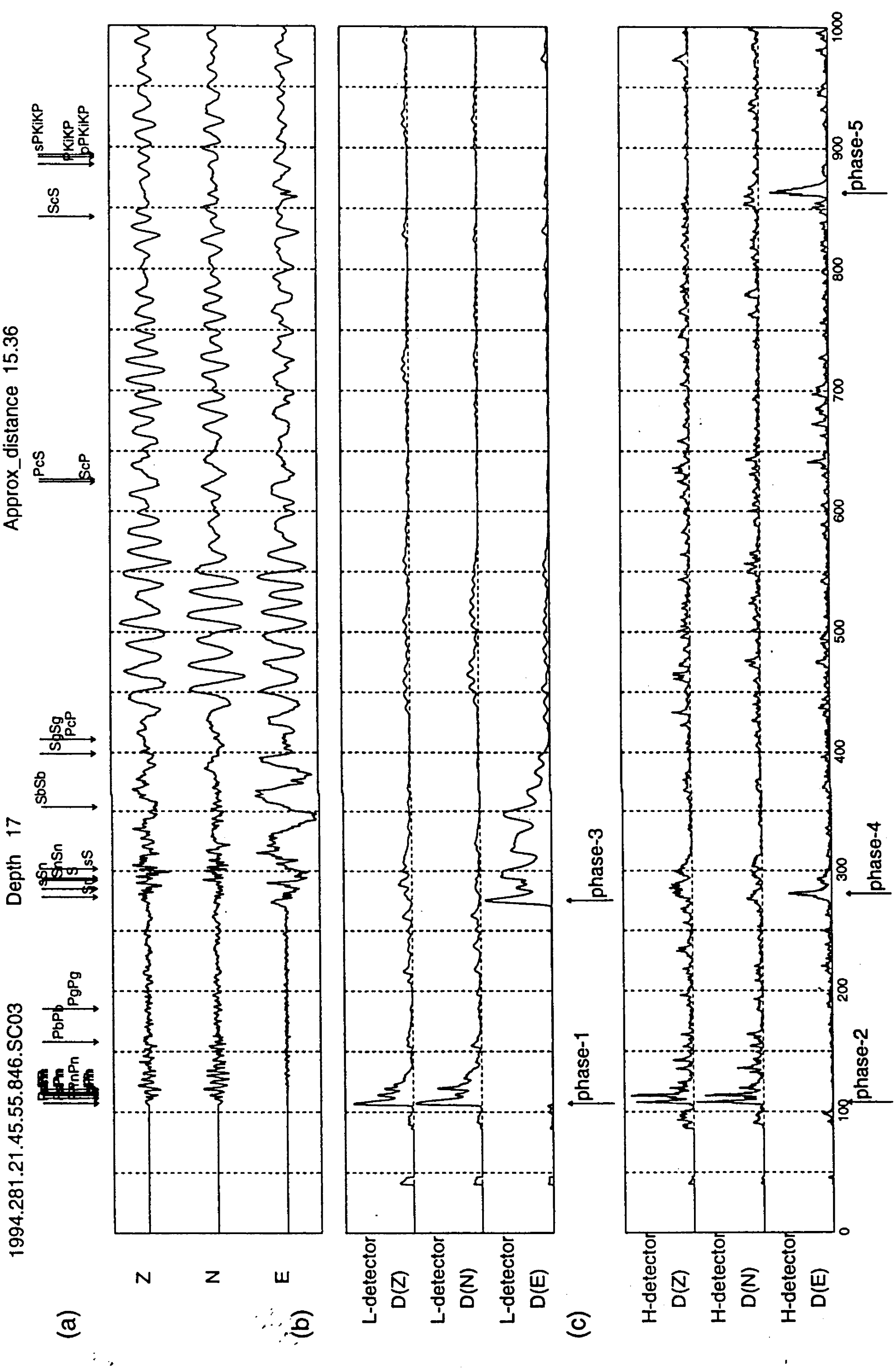
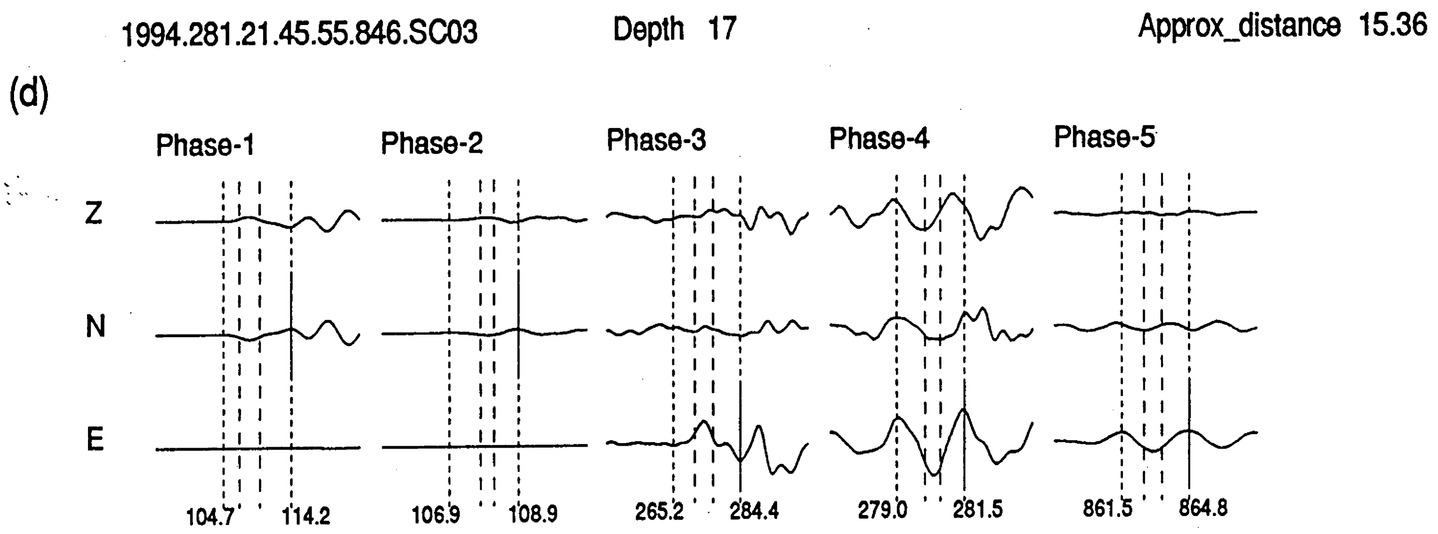


Fig. 6.4. Phase-detection and feature-extraction for event A – Halmahera (Indonesia).



	Phase-1	Phase-2	Phase-3	Phase-4	Phase-5
Ref. time	107.2	108.0	275.0	280.6	863.4
Amplitude	0.0600	0.0299	0.1752	0.3436	0.1018
Frequency	0.1359	0.5682	0.0568	0.4310	0.3205
Azimuth	180.26	180.50	85.74	250.17	263.02
App. incidence	50.16	49.58	9.22	11.89	1.37
P/S character	P	P	S	S	S

Fig. 6.4. – *continued.*

Table 6.6. *Processing sequence for Event A*

Step 1: Choice of Key Phases
Key *P*: phase 1, Key *S*: phase 3

Step 2: Test on Differential Time
167.8 s - surviving models: *P-S*

Step 3: P-S Pair Feature Match
Surviving models: *P-S*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
P S	0	14	180	phase 1: *P*, phase 3: *S*, phase 5: *ScS*	0.0033
	100	16	180	phase 1: *P*, phase 3: *S*	0.3333
	300	16	180	phase 1: *P*, phase 3: *S*	0.3333
	600	16	180	phase 1: *P*, phase 3: *S*	0.3333

ential time analysis. Table 6.6 summarises the progress of the event interpretation scheme. After the test of the differential times, only the *P-S* pair survives. This *P-S* model also survived the test on the expected *P-S* features. Therefore, there are only four nodes in the second generation corresponding to different interpretation of the event depth.

For every choice of seismic source interpretation in step 4 of Table 6.6, the distance range and *P-S* pair lead us to expect the arrival of the phase *ScS*. For the first case with a surface source (depth 0 km), the expected time for *ScS* corresponds to the observed [phase-5] and the expected phase features for *ScS* are matched with the attributes of [phase-5]. Therefore, in addition to the two key phases a further phase is identified. For each of the other three cases, the expected time range for *ScS* does not cover any observed phases, as a result the expected phase *ScS* is not found and the misfit measures are larger than that for a surface source.

When we compare the best interpretation (Distance 14°, Depth 0 km, Azimuth 180°) with the event specification given in Table 6.5, we can see that the estimations of distance, depth and azimuth are very close to the true values. At the same time, the automatic system has correctly identified three phases: [phase-1] as *P*, [phase-3] as *S*, and [phase-5] as *ScS*.

6.5.2 Event B – Myanmar (Burma)

This is a typical teleseismic event which is dominated by clear *P*, *S* and core reflections. In Figure 6.5 panel (b) and (c), we see that *P* and *S* arrivals are detected on both the high-frequency and the low-frequency set. Other detections are [phase-3] for *PcP* and [phase-4] for *PcS*. In Figure 6.5 (b), the detection traces also indicate the presence of *pP*, *PP* and *sS* arrivals. However, these detections are not flagged with an arrow because they lie below a pre-defined threshold. The threshold has been set to a common value for all the illustrations, and appears to be a little too high for this unusually clear event.

The phases identifications above can be made because we have external information which is not available to the automatic event recognition procedure. Now, let's see how the automatic system determines the event location and identifies the phase detections. The automatic system chooses the low frequency detections [phase-2] as the key *P* phase and [phase-6] as the key *S* phase. The subsequent steps in the event interpretation scheme are summarised in Table 6.7. After the test on the differential time between the key *P* and *S* phases, the three pairs *P-S*, *PP-SKS* and *PKP-SKS* have survived. Since the apparent angle of incidence of the key P phase is 27.8°, which is not as steep as required for *PKP* (see the note in Table 6.2), the choice of *PKP-SKS* fails in the test on *P-S* features. The other two *P-S* combinations survive this feature test and so there are eight nodes in the second generation in the assumption tree corresponding to different combinations of distance and depth as well as phase choice.

The smallest measure of misfit occurs for the second choice in step 4 of Table 6.7 which corresponds to an interpretation of the key phases as *P* and *S*. The estimates of the event parameters are epicentral distance 52°, depth 100 km, and azimuth 316°

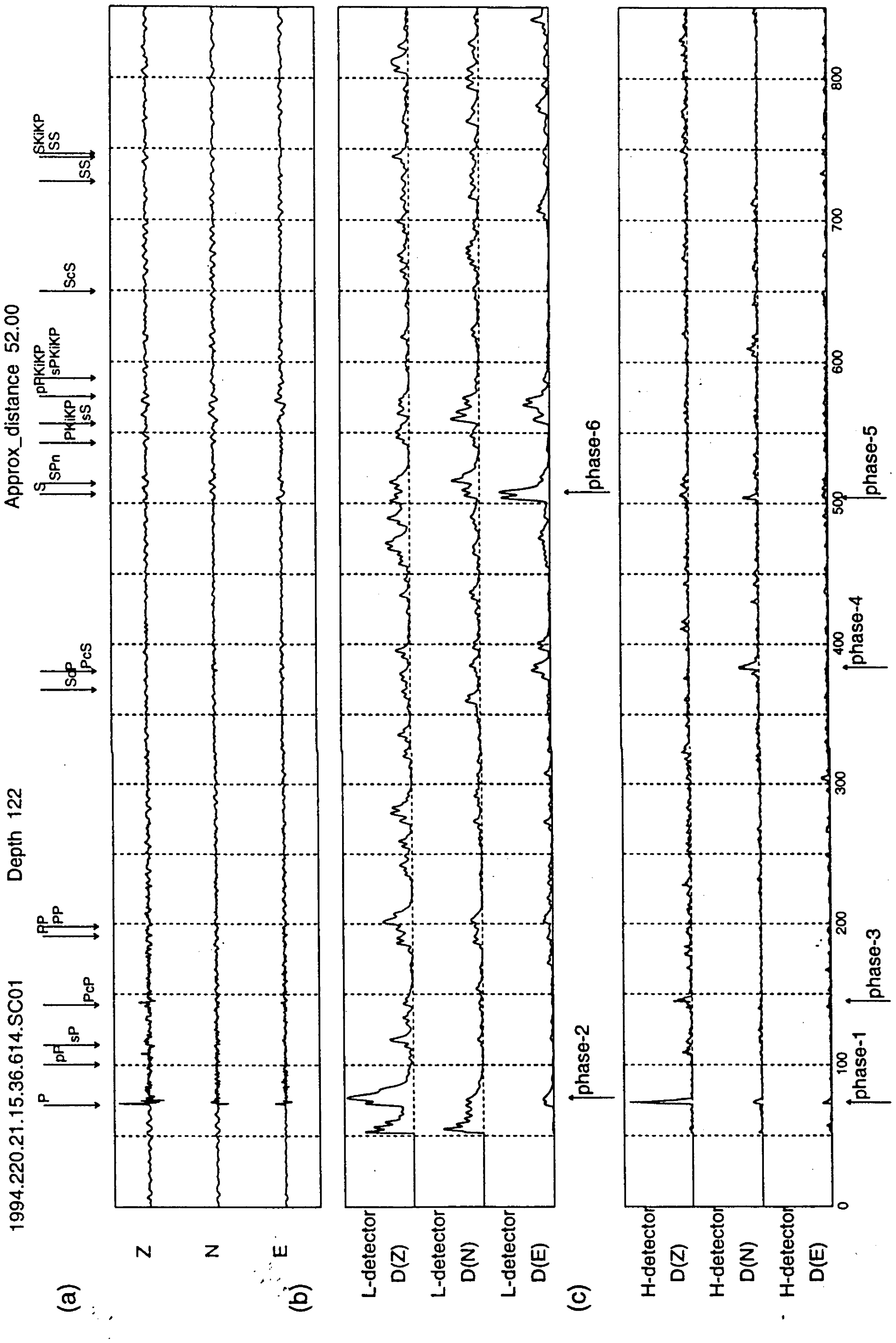
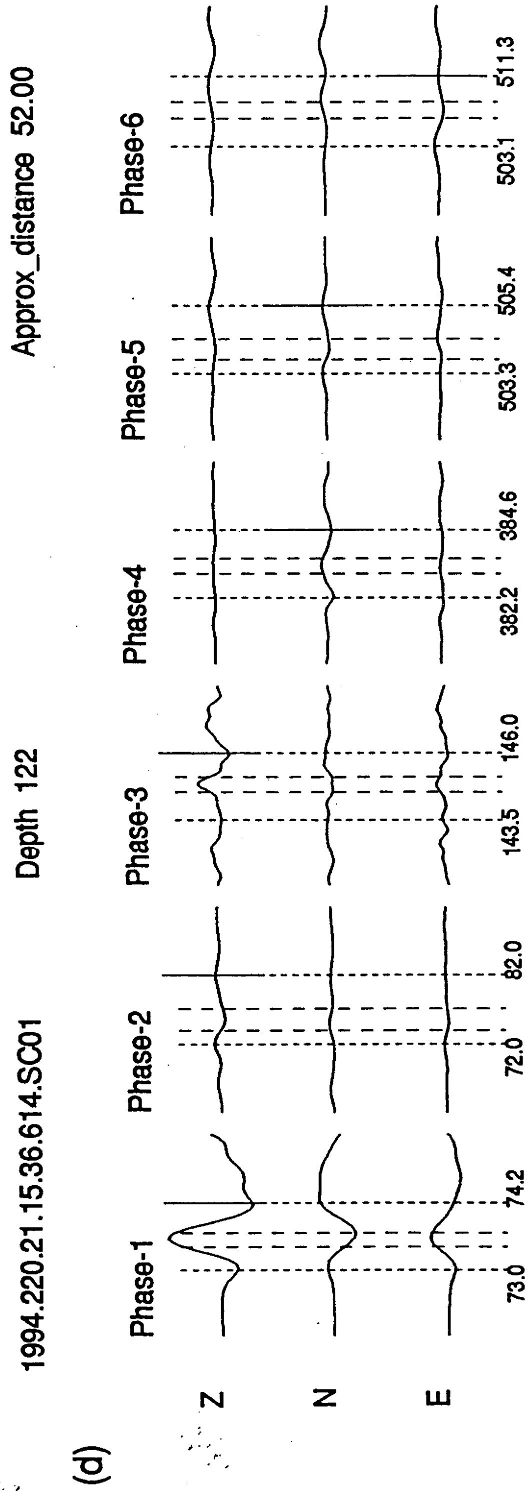


Fig. 6.5. Phase-detection and feature-extraction for event B – Myanmar (Burma).



	Phase-1	Phase-2	Phase-3	Phase-4	Phase-5	Phase-6
Ref. time	73.5	76.8	144.9	383.5	504.2	508.1
Amplitude	0.5901	0.0589	0.1508	0.0571	0.0404	0.0516
Frequency	0.8929	0.1404	0.4464	0.4310	0.6250	0.1276
Azimuth	140.72	316.43	131.46	10.81	170.32	286.85
App. incidence	26.97	27.82	25.37	9.54	33.64	5.06
P/S character	P	P	P	S	S	S

Fig. 6.5. – *continued.*

Table 6.7. *Processing sequence for Event B*

Step 1: Choice of Key Phases

Key *P*: phase 2, Key *S*: phase 6

Step 2: Test on Differential Time

431.3 s - surviving models: *P-S, PP-SKS, PKP-SKS*

Step 3: P-S Pair Feature Match

Surviving models: *P-S, PP-SKS*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
P S	0	50	316	phase 2: *P*, phase 3: *PcP*, phase 6: *S*	0.00196
	100	52	316	phase 2: *P*, phase 3: *PcP*, phase 4: *PcS*, phase 6: *S*	0.00015
	300	54	316	phase 2: *P*, phase 3: *pP*, phase 4: *PcS*, phase 6: *S*	0.00031
	600	58	316	phase 2: *P*, phase 3: *PcP*, phase 4: *PcS*, phase 6: *S*	0.00105
PP SKS	0	84	316	phase 2: *PP*, phase 6: *SKS*	0.33333
	100	82	316	phase 2: *PP*, phase 6: *SKS*	0.33333
	300	82	316	phase 2: *PP*, phase 6: *SKS*	0.33333
	600	82	316	phase 2: *PP*, phase 6: *SKS*	0.33333

(which also implies 136°), which matchs the information provided in Table 6.5 very well.

6.5.3 Event C – off coast of Northern California

Event C is at a much greater distance and so is in the core shadow for *P*. The expected first arrival, diffracted *P*, is so weak that it can not be seen on the seismogram (Figure 6.6a). Core phases (e.g., *PKP* and *SKS*) and surface multiples (e.g., *PP* and *PS*) are the main features on seismic records at this distance range.

The set of detections displayed in panels 6(b) and 6(c) pull out all the major arrivals on the seismogram. We notice that the detection [phase-7] and another high

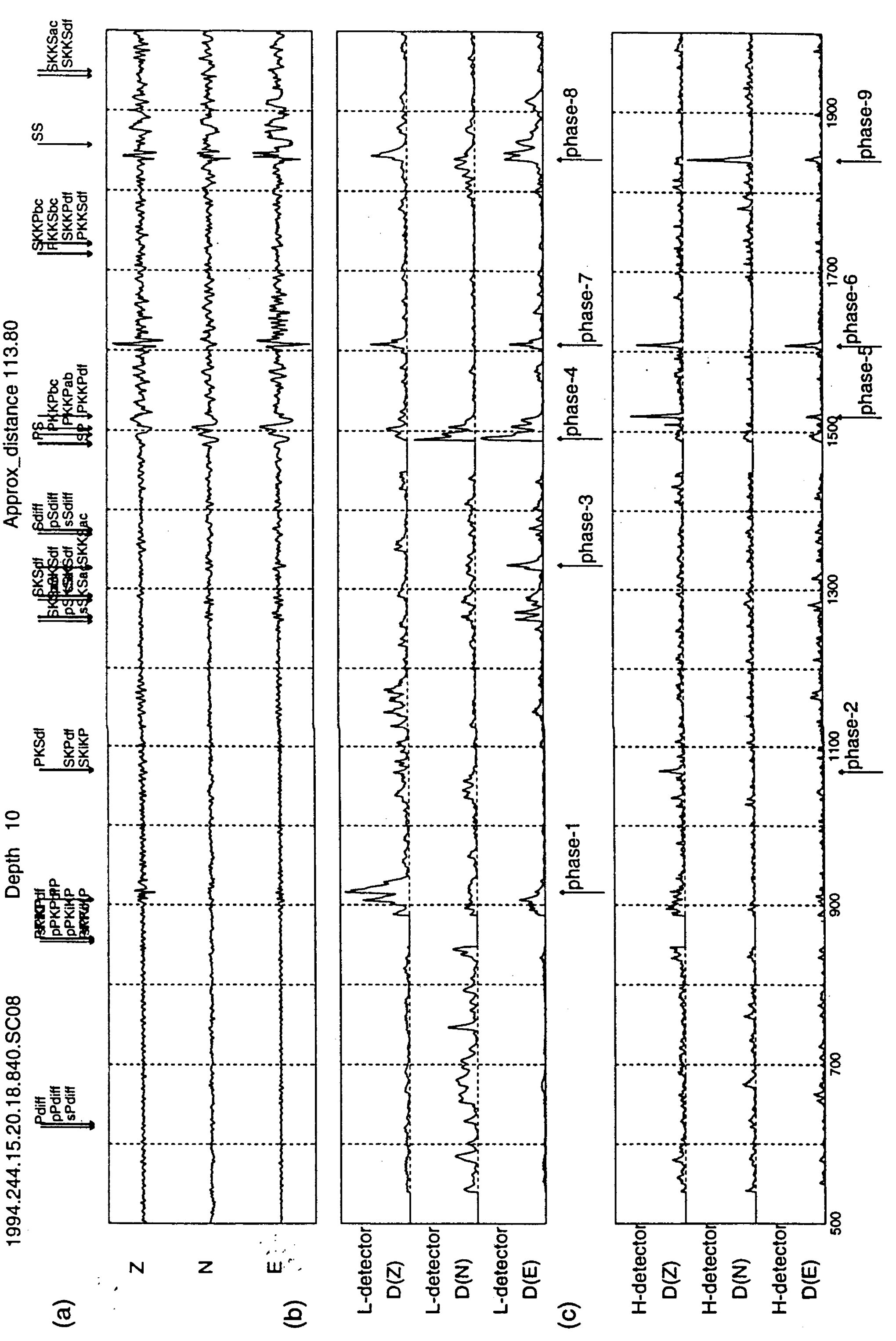
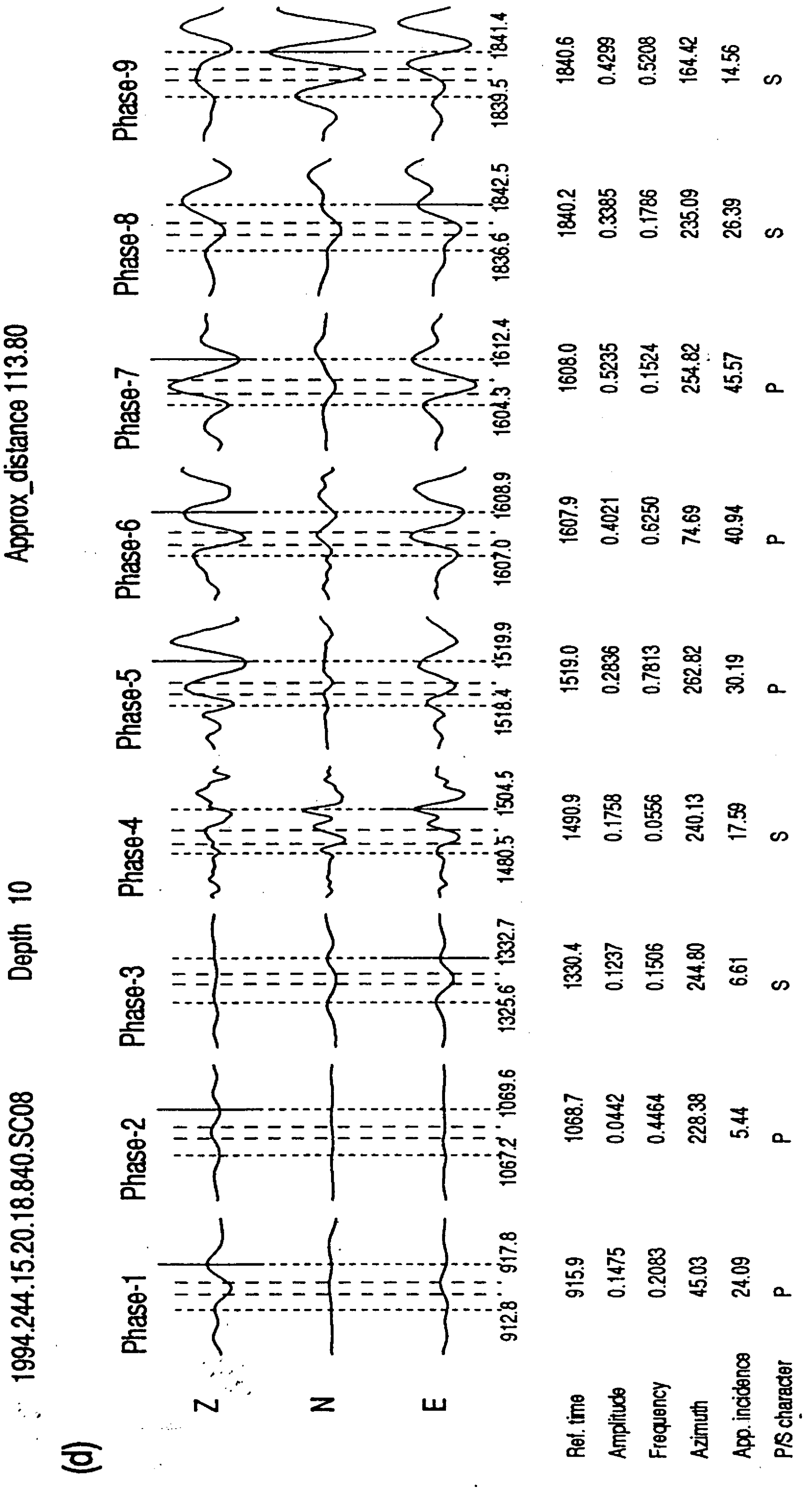


Fig. 6.6. Phase-detection and feature-extraction for event C – off coast of Northern California.



Ref. time	915.9	1068.7	1330.4	1490.9	1519.0	1607.9	1608.0	1840.2	1840.6
Amplitude	0.1475	0.0442	0.1237	0.1758	0.2836	0.4021	0.5235	0.3385	0.4299
Frequency	0.2083	0.4464	0.1506	0.0556	0.7813	0.6250	0.1524	0.1786	0.5208
Azimuth	45.03	228.38	244.80	240.13	262.82	74.69	254.82	235.09	164.42
App. incidence	24.09	5.44	6.61	17.59	30.19	40.94	45.57	26.39	14.56
P/S character	P	P	S	S	P	P	P	S	S

Fig. 6.6. – *continued.*

frequency S arrival [phase-9] have no obvious correlation with the phases expected for this even and may well come from some different source.

The detection on the low frequency set [phase-1] is selected as the key *P* phase. Note that since [phase-4] is bigger than [phase-3] and there is no detection in between, [phase-4] is selected as the key *S* phase in place of [phase-3] The progress of the automated event recognition system is presented in Table 6.8. In the first generation in the assumption tree, five set of *P-S* pairs survived the test on the differential time between the key phases. For the choice *P-S*, we do not expect to see another low frequency *S* phase before the key *S* phase, therefore this choice is unsuitable. In the case of P_{diff}-*SKS*, the amplitude of the key *P* phase is expected to be much smaller than the key *S* phase. However, the observed behaviour is that the key *P* phase is almost comparable with the key *S* phase, therefore this choice is also rejected. For the choice *PKP-SKS*, the angle of incidence of the key *P* phase is not suitable for a *PKP*, as also occurred for event B. Finally, the surviving phase pairs at the first generation nodes are *P-SKS* and *PP-PS*.

In step 4 of Table 6.8 we note that the most likely interpretations are for sources at either the surface or 100 km depth, at a distance of 112° with *PP-PS* identified as the pair of key phases. The level of misfit is very close for the two cases. When we compare the event parameter estimates with the event specification in Table 6.5, we can see that the distance is very well determined, the estimate of depth is reasonable good, while the estimated azimuth is far from the true value because of the distance. However, it is interesting to see that the azimuth 263° for [phase-5] in Figure 6.6d (identified as *PKKP*),is very close to the true azimuth 260°.

The larger misfit for event C compared with the previous cases is caused by the uncorrelated big arrivals (i.e., [phase-7] and [phase-9]), which most likely come from an overlapping event (note that there is no obvious association with the phases predicted from the *iasp91* travel times).

Table 6.8. *Processing sequence for Event C*

Step 1: Choice of Key Phases

Key *P*: phase 1, Key *S*: phase 4

Step 2: Test on Differential Time

575.0 s - surviving models: *P-S, P-SKS,* P_{diff}*-SKS, PP-PS, PKP-SKKS*

Step 3: P-S Pair Feature Match

Surviving models: *P-SKS, PP-PS*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
P SKS	0	82	45	phase 1: *P*, phase 4: *SKS*	0.8399
	100	82	45	phase 1: *P*, phase 4: *SKS*	0.8399
	300	82	45	phase 1: *P*, phase 4: *SKS*	0.8399
	600	92	45	phase 1: *P*, phase 4: *SKS*	0.8399
PP PS	0	112	45	phase 1: *PP*, phase 3: *SKKS*, phase 4: *PS*, phase 5: *PKKP*	0.4662
	100	112	45	phase 1: *PP*, phase 3: *SKKS*, phase 4: *PS*, phase 5: *PKKP*	0.4658
	300	110	45	phase 1: *PP*, phase 4: *PS*	0.7161
	600	108	45	phase 1: *PP*, phase 4: *PS*	0.7161

6.5.4 Event D – Northern Bolivia

The epicentral distance is even larger for this event and the seismogram is now dominated by core phases, such as *PKP*, *PKKP*, *SKS*, and *SKKS* etc. Figure 6.7(a) shows two complicated packages of core arrivals in the early part of the broad-band record.

The phase detection system is designed to pick out the first onset of a phase package. Therefore the detections [phase-2] and [phase-4] are indicated at the beginning of the two packages. From the low-frequency detections, [phase-2] is selected as the key *P* phase and [phase-6] as the key *S* phase. As before we can follow the progress of the automatic system in Table 6.9. Three choices for the pair of key phases survive the test of differential time. Since the incidence angle for [phase-2] is 9.78°, which is

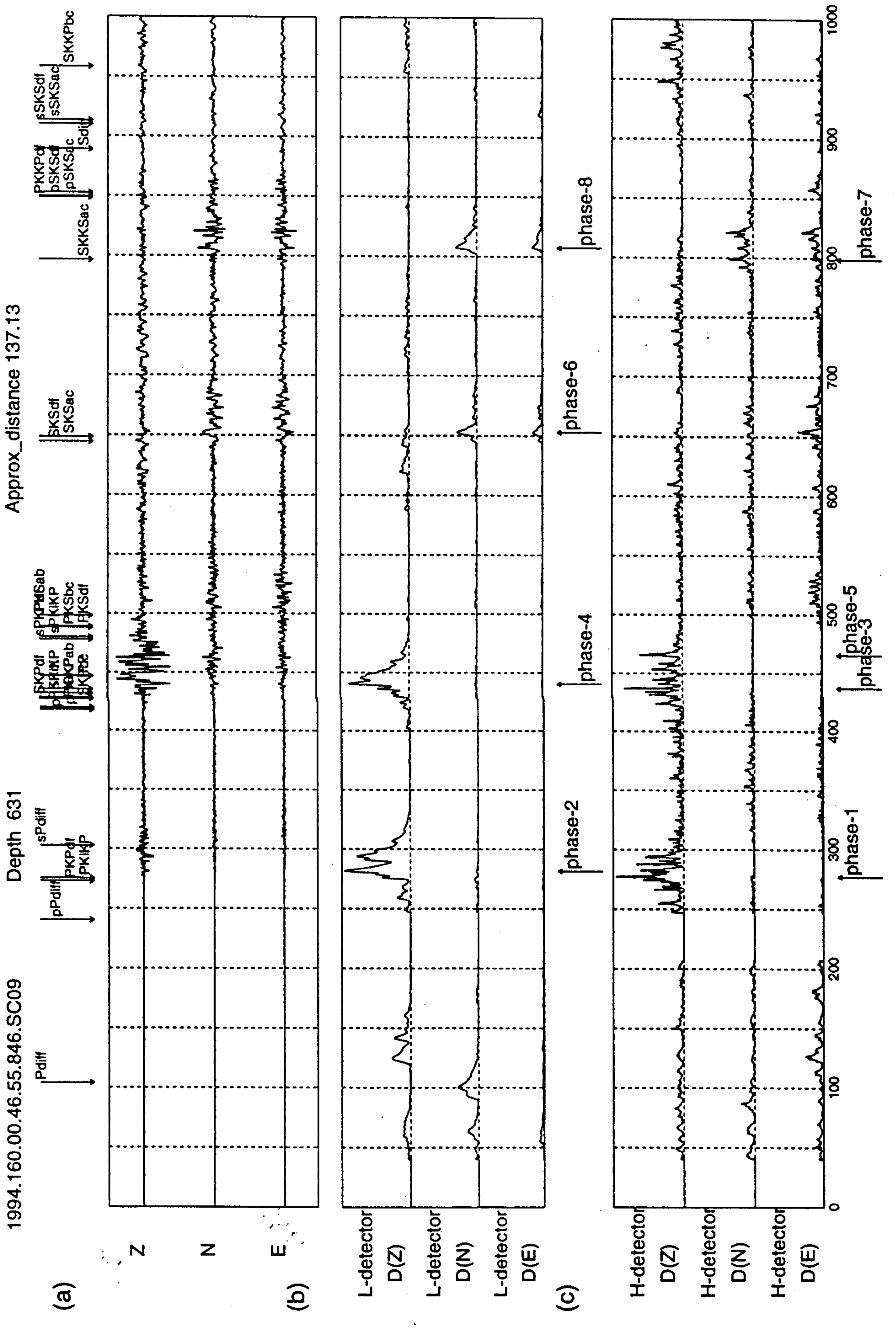
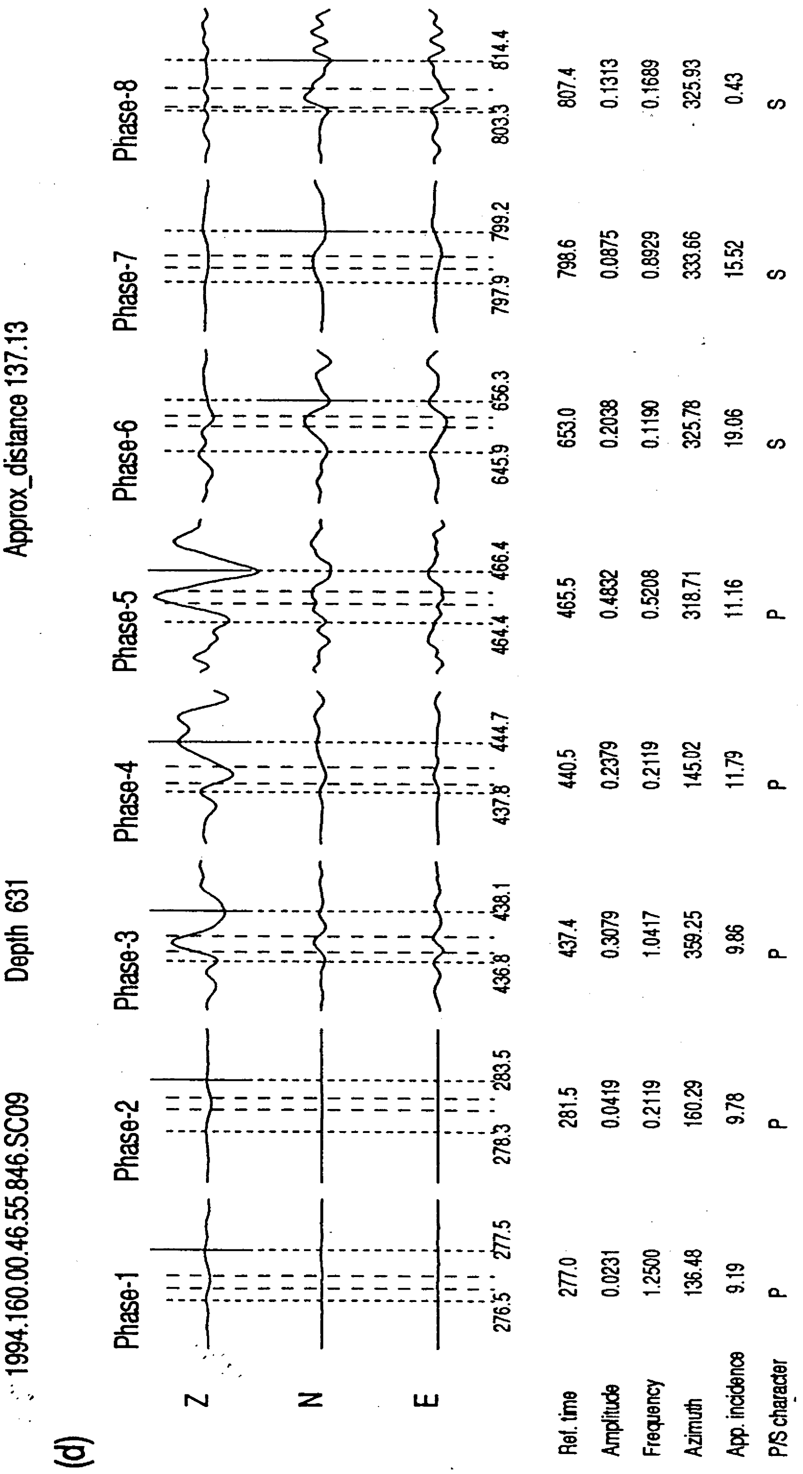


Fig. 6.7. Phase-detection and feature-extraction for event D – Northern Bolivia.



	Phase-1	Phase-2	Phase-3	Phase-4	Phase-5	Phase-6	Phase-7	Phase-8
Ref. time	277.0	281.5	437.4	440.5	465.5	653.0	798.6	807.4
Amplitude	0.0231	0.0419	0.3079	0.2379	0.4832	0.2038	0.0875	0.1313
Frequency	1.2500	0.2119	1.0417	0.2119	0.5208	0.1190	0.8929	0.1689
Azimuth	136.48	160.29	359.25	145.02	318.71	325.78	333.66	325.93
App. incidence	9.19	9.78	9.86	11.79	11.16	19.06	15.52	0.43
P/S character	P	P	P	P	P	S	S	S

Fig. 6.7. – *continued.*

Table 6.9. *Processing sequence for Event D*

Step 1: Choice of Key Phases
Key *P*: phase 2, Key *S*: phase 6

Step 2: Test on Differential Time
371.5 s - surviving models: *P-S, PP-SKS, PKP-SKS*

Step 3: P-S Pair Feature Match
Surviving models: *PKP-SKS*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
PKP SKS	0	114	160	phase 2: *PKP*, phase 6: *SKS*	0.6423
	100	114	160	phase 2: *PKP*, phase 6: *SKS*	0.5763
	300	114	160	phase 2: *PKP*, phase 6: *SKS*	0.5763
	600	128	160	phase 2: *PKP*, phase 3: *pPKIKP*, phase 5: *sPKIKP*, phase 6: *SKS*	0.2058

too steep to be a *P* in the choice of the *P-S* pair, the branch corresponding to *P-S* is terminated. Similarly, the incidence angle and the frequency of [phase-2] is not suitable for a *PP* in the *PP-SKS* pair. Therefore only one choice of *P-S* phase pairs model *PKP SKS* is left after the *P-S* feature test.

The most likely interpretation identified in step 4 of Table 6.9 is the last choice of a very deep event at a distance of 128°. Once again the large epicentral distance means that the azimuth of the first *P* phase is not a reliable estimate. When we compare with the event specification in Table 6.5, the estimated depth is quite accurate. However, the estimated distance is 9° away from the true value, which is not as good as in the previous examples. This difference is actually caused by the procedure described in section 6.4.3.2, because the mapping between differential time and distance is not one-to-one for the case of *SKS-PKP* (see Figure 6.1). When only one choice of distance is made from the mapping (differential time to distance), the true distance can be missed.

6.5.5 Performance of the event recognition system

As can be seen from these examples, the automatic system can successfully restrict attention to a limited number of possible interpretations, and then a quantitative assessment can be made by using the misfit measure between the observed and expected phases. The misfit measure performs very well in selecting the most likely interpretation and results of the phase identification process give estimates for the distance and depth of the event which would be very valuable for preliminary event location.

6.6 DISCUSSION

6.6.1 Extensions of the processing scheme

The present event recognition system has been designed to be simple and robust so that we can concentrate on the processing strategy. We recognise that, at present, we have not included the full range of available seismological information on the expected character of the seismic wavefield. For example, for some source orientations and azimuths it is possible for the depth phases e.g. *pP, sP* and *pS, sS* to be substantially larger than the equivalent direct phases and so a depth phase could be picked as a key phase with consequent distortion of the present phase identification process. Fortunately, an important feature of the assumption tree approach is that it is relatively easy to make revisions and improvements in the future.

Firstly, the seismological expertise for the nature and properties of commonly observed phases (Table 6.3, 6.4) is kept separately from the algorithm. Thus the addition or modification of the summary of this seismological expertise will not affect the algorithm.

More significantly, the information required for constructing the assumption tree is also held separated from the algorithm itself. For example, the growth of the stems to the first generation nodes, representing the set of *P*-*S* pairs, is implemented by reading a separate list of *P*-*S* choices (Table 6.1). The automatic system will construct a first generation node for every item in this *P*-*S* list. Thus the inclusion of further *P-S* pairs can be readily accomplished by adding more items to the *P-S* list. A similar approach can be applied to the second generation, to revise or extend the possible source depths.

The sampling in distance could be increased by a employing a denser travel-time table for each depth. The current number of depth/distance combinations (360) has been chosen to allow rapid testing for a possible real time environment whilst still providing a useful definition of the nature of an event for preliminary location. However, the number of classes could be increased without a very large processing overhead to provide an increased density of states.

6.6.2 Further developments

6.6.2.1 Refinement of source distance and depth

As described in section 6.4.3.2, the source distance is obtained by inference from the *P-S* model and the differential time interval. For a typical combination of *P-S* phase pairs, the differential time has an almost linear dependence on the source distance. This property is exploited in the search to find the closest match to the observed differential time.

There are two factors which need to be taken into consideration, firstly we have to expect some differences between the observed differential time and that for an Earth model such as *iasp91* and so we need to allow for some tolerance when matching times. Secondly, the mapping between differential time and distance is not always one-to-one, as can be seen for the {*PKP, SKS*} and {*PKIKP, PKS*} pairs in Figure 6.1; this occasional problem could be resolved by using a third generation in the assumption tree (Figure 6.2) to determine distance.

The distance and depth estimates can be refined after the main event recognition process by introducing a local assumption tree to search for the best solution in the distance/depth neighbourhood of the postulated event parameters .

6.6.2.2 Continuous operation and overlapping events

As we have described in section 6.4.2, the phases corresponding to an event sequence are recognised and separated from the input stream by comparison with a certain expected pattern for an event. The input is a continuous stream of phase detections so that we need to also be able to run the event-interpretation process continuously. The normal procedure in pattern recognition is use information in time order to separate groups of phases as events, event by event, without going back or reusing the previous data. However, in the case of seismic event recognition, taking into consideration the likelihood of arrivals from different events overlapping in time, we would suggest that the unmatched arrivals in previously observed data be combined with the rest of the input stream and reused for recognising an event sequence, to yield the input to the next event process. In this way, we will be able to separate overlapping events to allow full phase recognition and interpretation.

7

Automatic seismic event recognition using multiple broad-band stations

The automated analysis of seismograms provides useful information from a single seismogram but is of greatest utility when applied to a network of stations. For global monitoring, the network of stations is sparse and the ambient conditions at each site will vary so that tuning of signal detection thresholds will be required so that comparable operation can be achieved at each site.

A related issue is the robustness of the event recognition procedures using similar data. This issue has been addressed by applying the automated analysis procedure described in the previous chapter to multiple recordings of the same event at different portable broad-band stations in Australia. The stations cover a significant span in epicentral distance (although generally the azimuths will be similar).

The use of a number of different recordings for the same events provides a test of the consistency achievable with the automated analysis procedure. The portable sites cover a range of geological conditions and ambient noise environments and so provide a useful test of the reliability of the algorithms. With accurate GPS timing and positioning the extraction of information on range and depth can be very effective. However, a limitation of the use of portable sites is that the accuracy of seismometer alignment is limited and so there can be systematic (site-dependent)

Table 7.1. *Events used for illustration*

Event	Year	Day	Time	Latitude	Longitude	Depth	Mb
A	1994	281	21:44:07.2	-1.26	127.98	17	6.4
B	1994	220	21:08:31.6	24.72	95.20	122	6.0
C	1994	244	15:15:53.0	40.40	-125.68	10	6.6
D	1994	160	00:33:16.2	-13.84	-67.55	631	7.0
E	1994	234	17:26:37.5	-11.51	166.45	142	6.2
F	1994	232	04:38:50.5	44.66	149.18	24	6.2
G	1994	240	18:37:20.6	44.78	150.06	19	6.1
H	1993	133	11:59:49.2	55.18	-160.46	32	6.4

errors in azimuth. By varying the sets of stations it is possible to get a good coverage in epicentral distance.

Figure 7.1 shows the stations which have been used for the events discussed in this section. Most of the events use the stations SC01-SC10 in the Northern Territory of Australia, in conjunction with station YB01 in the southeast. One event uses the SA stations in Queensland.

Figure 7.2 shows the locations of events which have been used. The events have been chosen with a minimum of four records and normally at least a 5 degree span in epicentral distance. The events are displayed in Table 7.1.

The first group (Events A-D) are those from which individual stations have been illustrated in the previous chapter. The second group (Events E-H) covers a range of epicentral distances not represented in the earlier set (30-40°, 60-70°, 80-90°). The same set of trigger levels have been used for all stations.

The results of the automated procedures are displayed in a set of summary tables (Tables 7.2-7.9) in which the estimated ranges, depths and azimuths are compared with the values determined from published hypocentral information derived from global observations. Note that, the azimuth estimate for the key P phase is used as the azimuth of the whole event. Since the estimated azimuth actually implies two possible directions (with a 180° difference in between), both the two possible

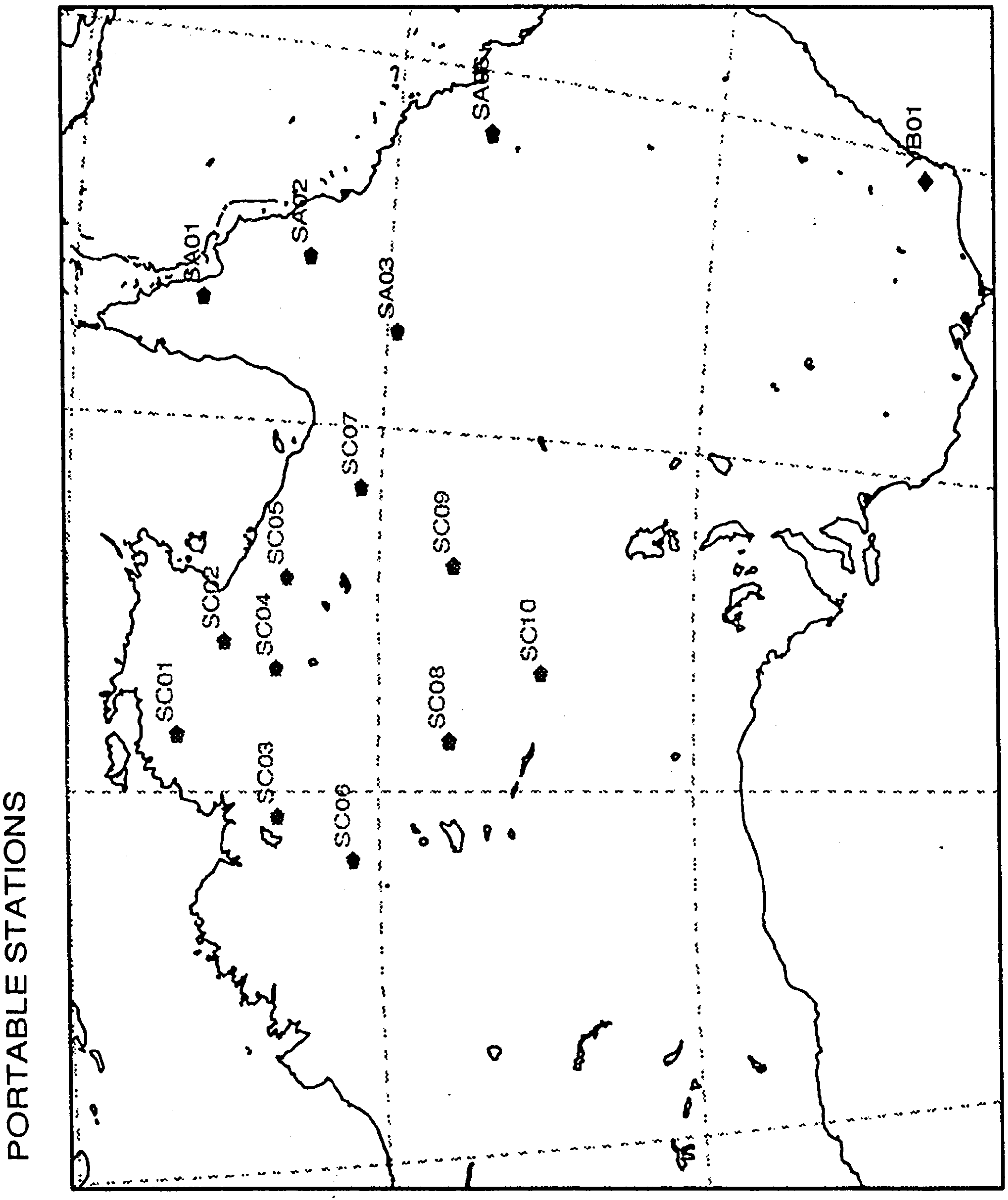


Fig. 7.1. The stations which have been used for the extensive testing.

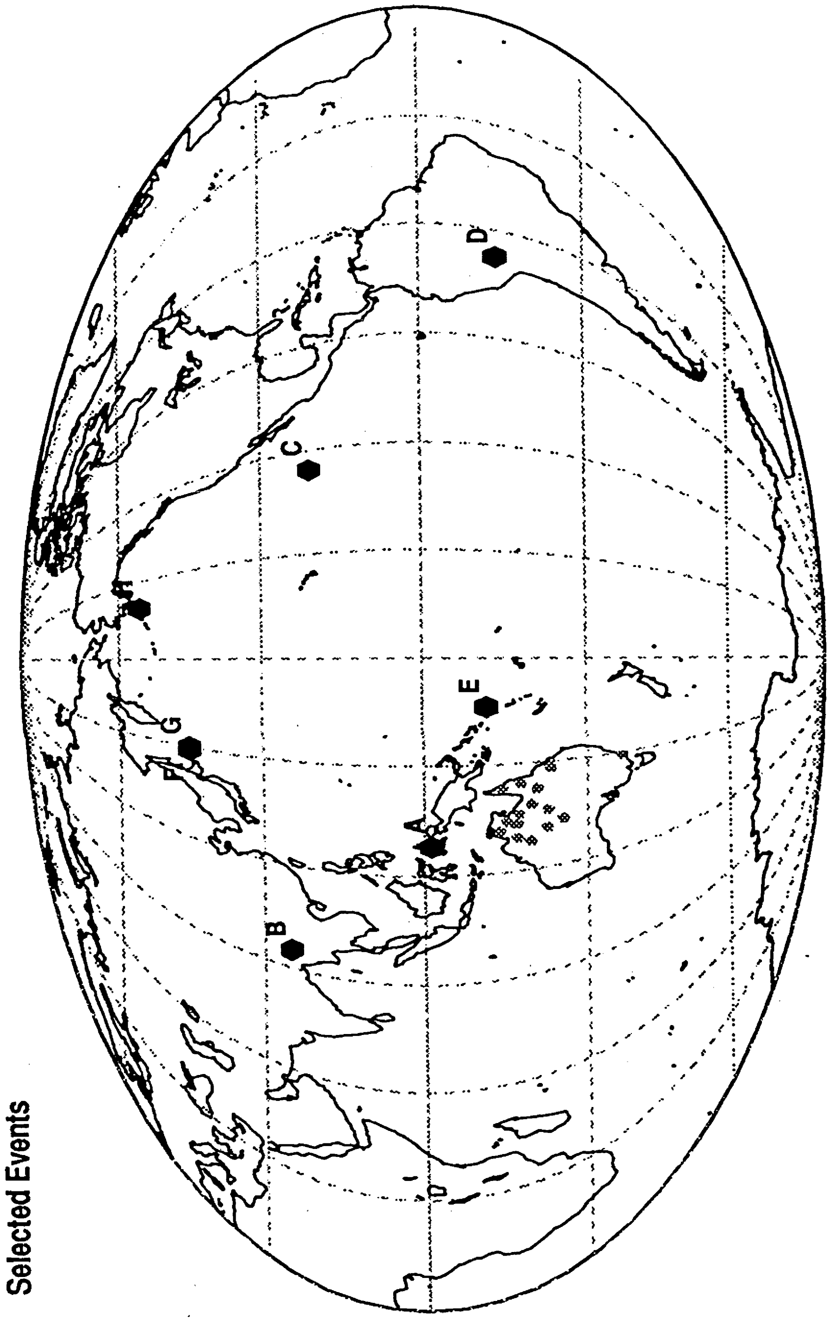


Fig. 7.2. The locations of events which have been used for the extensive testing.

Table 7.2. *Event A - Depth 17 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SC03	P-S	14	0	0, 180	15	175
SC08	P-S	26	600	170, 350	21	171
SC10	P-S	26	0	163, 343	24	168
	PP-SKS	128	0			
YB01	P-S	40	0	158, 338	40	153

Table 7.3. *Event B - Depth 122 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SC01	P-S	52	100	141, 321	52	133
SC04	P-S	54	300	139, 319	55	134
SC02	P-S	54	100	133, 313	55	132
SC03	P-S	54	300	150, 330	53	138

values for azimuth are given in the Tables 7.2-7.9. In the case when the values of the "misfit" measure (see chapter 6) do not show a clear preference between two interpretations, both of them are included in the summary table (see e.g., in Table 7.2, there are two interpretations for station SC10).

The two closest events are Event A (Table 7.2) and Event E (Table 7.6) for which the phase pair used for event characterisation is *P-S*. For event A the depth estimate was good when an additional phase *PcP* (SC03) or *ScS* (YB01) could be associated with the main arrivals. For station SC08 the expected arrival time of *PcP* is very close to that of the *S* waves interacting with the upper mantle triplications and cannot be resolved. For SC10 the discrimination provided by just the $P-S$ differential time was limited and the possibility arose of a much larger epicentral distance which can be discounted by comparison with the other stations. For event E the range estimates are quite good and the event is recognised to lie at depth even though the tendency is for too great a depth to be estimated. The rather large depth at station SC06 arises from an attempt to associate *PcS* very close in time

Table 7.4. *Event C - Depth 10 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SC08	PP-PS	112	100	45, 225	114	260
SC10	PP-PS	114	300	32, 212	114	257
SC01	PP-SKS	108	0	55, 235	108	268
	P-S	42	100			
SC03	PP-SKS	112	0	38, 218	112	266

This event suffers from interference from another event of larger amplitude.

Table 7.5. *Event D - 631 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SC09	PKP-SKS	128	600	160, 340	137	213
SC08	PKP-SKKS	132	600	162, 342	140	208
SC01	PKP-SKKS	148	600	161, 341	147	216
SC06	PKP-SKKS	134	100	142, 322	144	206
SC04	PKP-SKKS	134	100	145, 325	143	215
SC05	PKP-SKKS	134	100	145, 325	142	217

SKS is visible in the records for SC04 and SC05 but is not large enough to meet the preset trigger.

Table 7.6. *Event E - Depth 142 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SC05	P-S	32	300	86, 266	30	256
SC04	P-S	34	300	91, 271	33	257
SC01	P-S	34	100	87, 267	34	263
SC08	P-S	36	300	87, 267	35	248
SC03	P-S	36	100	76, 256	36	258
SC06	P-S	42	300	84, 264	38	254

Table 7.7. *Event F - Depth 24 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SC01	P-S	60	100	16, 196	60	200
SC05	P-S	62	0	6, 186	63	194
SC04	P-S	64	100	0, 180	63	197
SC03	P-S	64	100	18, 198	64	201
SC06	P-S	66	100	9, 189	66	202
SC08	P-S	68	0	7, 187	69	198
	PKP-SKKS	130	100			
SC10	P-S	72	100	173, 353	71	195
YB01	P-S	84	100	163, 343	81	180

Table 7.8. *Event G - Depth 19 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SC04	P-S	64	100	44, 224	63	198
SC01	P-S	60	100	15, 195	63	201
SC05	P-S	64	100	11, 191	63	195
SC03	P-S	64	100	19, 199	64	202
SC06	P-S	66	100	29, 209	67	203
SC08	P-S	70	100	6, 186	69	199
	P-S	78	600			
SC10	P-S	72	100	2, 182	71	196

Table 7.9. *Event H - Depth 32 km*

Station	Phase pair	Est Range	Est Depth	Est Azimuth	Range	Azimuth
SA03	PKP-SKKS	154	100	166, 346	89	232
	P-SKS	98	0			
SA06	PKP-SKKS	154	100	21, 201	89	226
	P-SKS	98	0			
SA02	P-SKS	98	100	27, 207	86	232
SA01	PKP-SKKS	150	100	34, 214	84	234
	P-S	84	100			

to S. It may be desirable to modify the procedure to try to force a minimum time separation between associated phases to minimise such problems.

Event B with ranges close to 50° gives a good match of epicentral distance, the depth estimates are mostly constrained by matching the expected time of PcS (which might in fact be ScP with somewhat different implications for depth).

The two shallow events F, G (Tables 7.7, 7.8), which mostly lie between 60° and 70° from the source, are generally well matched and there is good consistency across the sets of stations. We note that once again there is a tendency for the depth to be overestimated. For shallow sources this can occur when a feature on the record is associated with a near source reflection (e.g. pP, sP) since these cannot be correlated with a surface source. Because of the sparse sampling in depth, the procedure is forced to make a decision between 0 km and 100 km. Complexity in the source pulse (in the case of these two examples, the interference from the closely following PcP) can be mistaken for a depth phase and hence suggest a deeper origin than appropriate. Also, in order to match the observed differential time with the 2 degree sampling of the field, it may be possible to improve the match by adjusting the depth. Further the *iasp91* travel time tables have been used in the analysis for which S is now recognised to be a little slow (Kennett, Engdahl & Buland, 1995) and this can be compensated by increasing the depth of the source.

A detailed inspection of the automated process for event F from four stations (SC01, SC06, SC10, and YB01) is presented in Figures 7.3-7.6 and Tables 7.10-7.13. The procedure of phase-detection and characterisation is represented in the same form of figure with four panels as described in chapter 6 (section 6.5). Each of these figures is accompanied by a summary table of event interpretation in the same form as the tables in chapter 6. The figures and tables are associated with the individual stations (SC01 – Figure 7.3, Table 7.10; SC06 – Figure 7.4, Table 7.11; SC10 – Figure 7.5, Table 7.12; SC01 – Figure 7.6, Table 7.13). The original seismograms in panel (a) in Figures 7.3-7.5 show that the waveforms are very similar at stations SC01, SC06, and SC10. Every of the three stations recorded a clear phase P followed by

Table 7.10. *Processing sequence for Event F at station SC01*

Step 1: Choice of Key Phases
Key *P*: phase 1, Key *S*: phase 6

Step 2: Test on Differential Time
488.0 s - surviving models: *P-S, PKP-SKKS*

Step 3: P-S Pair Feature Match
Surviving models: *P-S, PKP-SKKS*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
P S	0	58	16 or 196	phase 1: *P*, phase 6: *S*	0.5158
	100	60	16 or 196	phase 1: *P*, phase 3: *pP*, phase 6: *S*	0.1875
	300	64	16 or 196	phase 1: *P*, phase 3: *PcP*, phase 6: *S*	0.2148
	600	68	16 or 196	phase 1: *P*, phase 4: *pP*, phase 6: *S*	0.2903
PKP SKKS	0	126	16 or 196	phase 1: *PKP*, phase 6: *SKKS*	0.5616
	100	126	16 or 196	phase 1: *PKP*, phase 3: $pPKP_{\mathrm{df}}$ phase 6: *SKKS*	0.2456
	300	126	16 or 196	phase 1: *PKP*, phase 6: *SKKS*	0.4818
	600	128	16 or 196	phase 1: *PKP*, phase 4: $pPKP_{\mathrm{df}}$ phase 6: *SKKS*	0.2515

a stronger depth arrival in about 15 seconds. The depth arrival comes a little later than expected from the *iasp91* travel time table. At the station YB01, the event distance is somewhat greater than for the other three, and signal to noise ratio is also lower (around 1/1 for the phase *P*) so that the phase *P* can only be picked out by the high-frequency detector. The phase-detection works very well in the four cases, and has actually detected *PKiKP* at SC01 and *SKiKP* at YB01; these core phases have not been identified by the interpretation system, simply because they have not been included in the preliminary pattern (i.e., Table 6.3).

The performance of the automated procedures is very good out to about 80 ° when

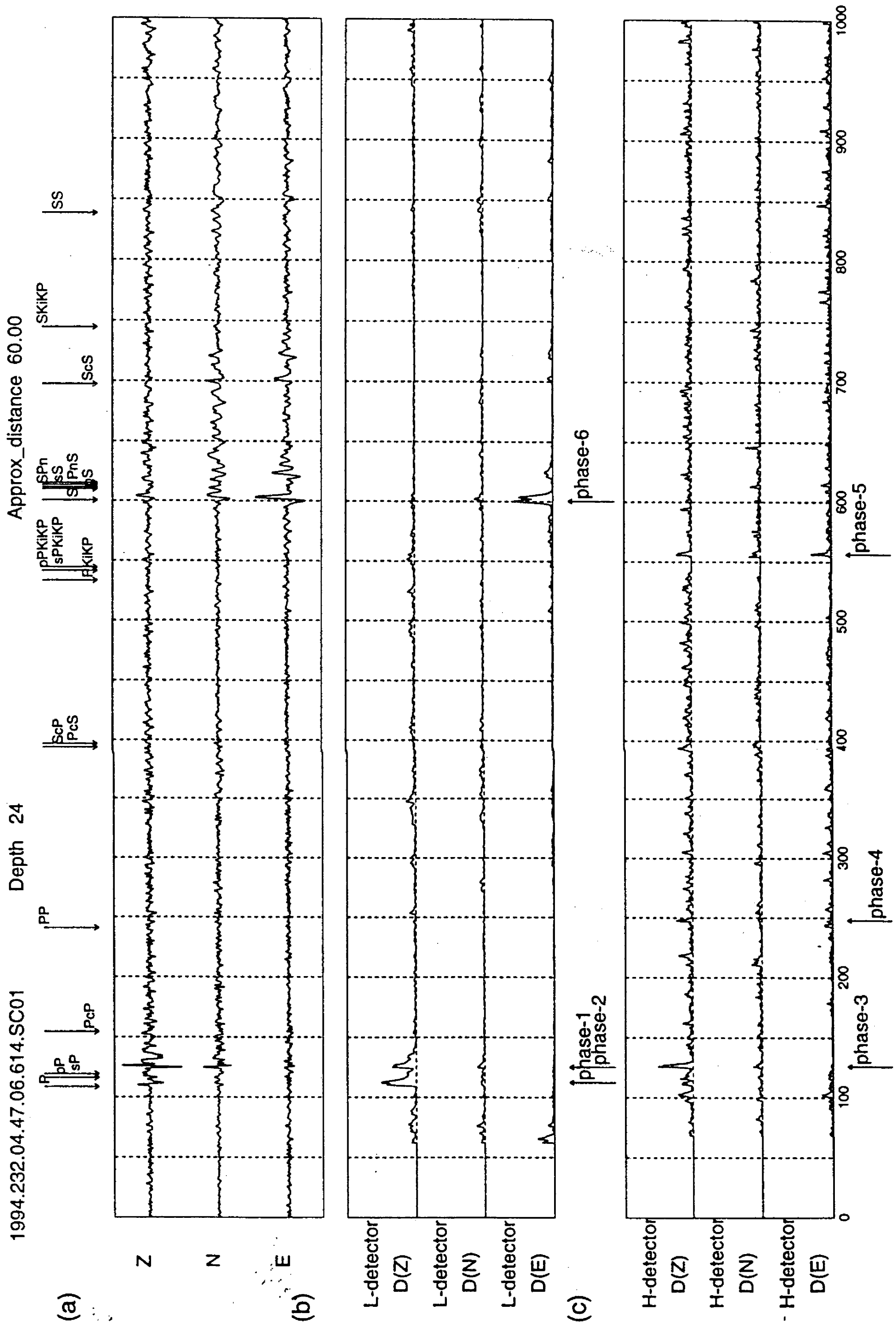
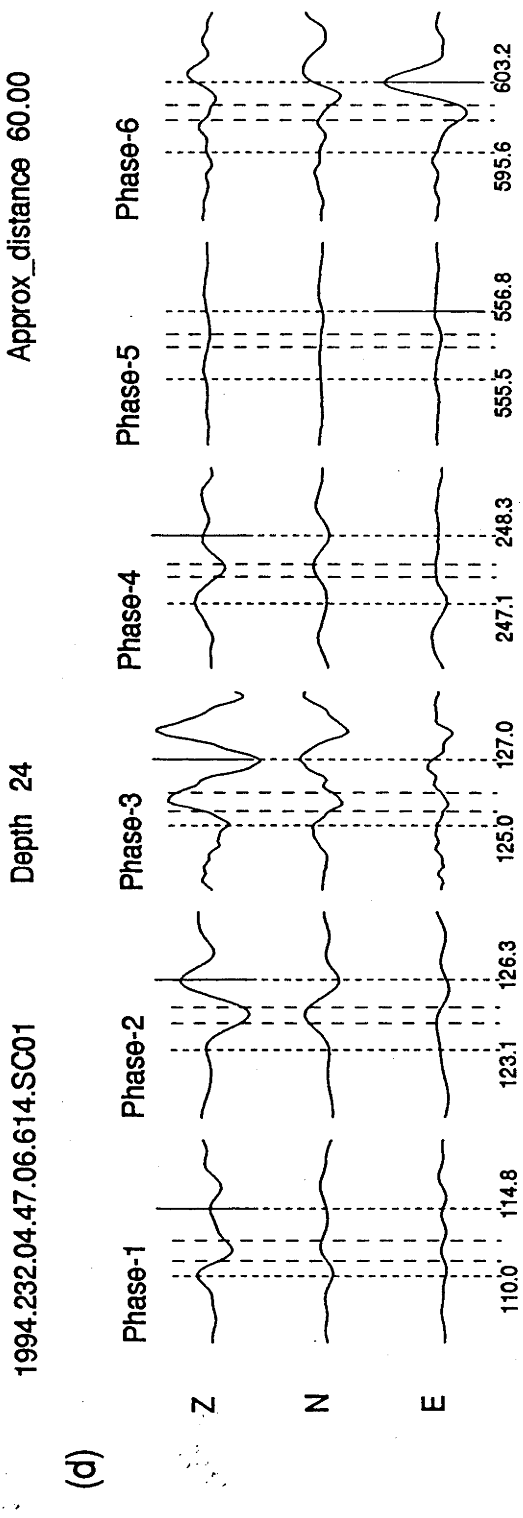


Fig. 7.3. Phase-detection and feature-extraction for event F at the station SC01.



Ref. time	112.3	125.0	126.0	247.9	556.4	600.3
Amplitude	0.1786	0.4350	0.3884	0.1899	0.0477	0.3020
Frequency	0.2717	0.3205	0.6579	0.8333	0.9615	0.1524
Azimuth	16.34	10.52	202.20	19.10	268.23	262.24
App. incidence	17.74	28.57	25.31	26.92	54.36	6.27
P/S character	P	P	P	P	P	S

Fig. 7.3. – *continued.*

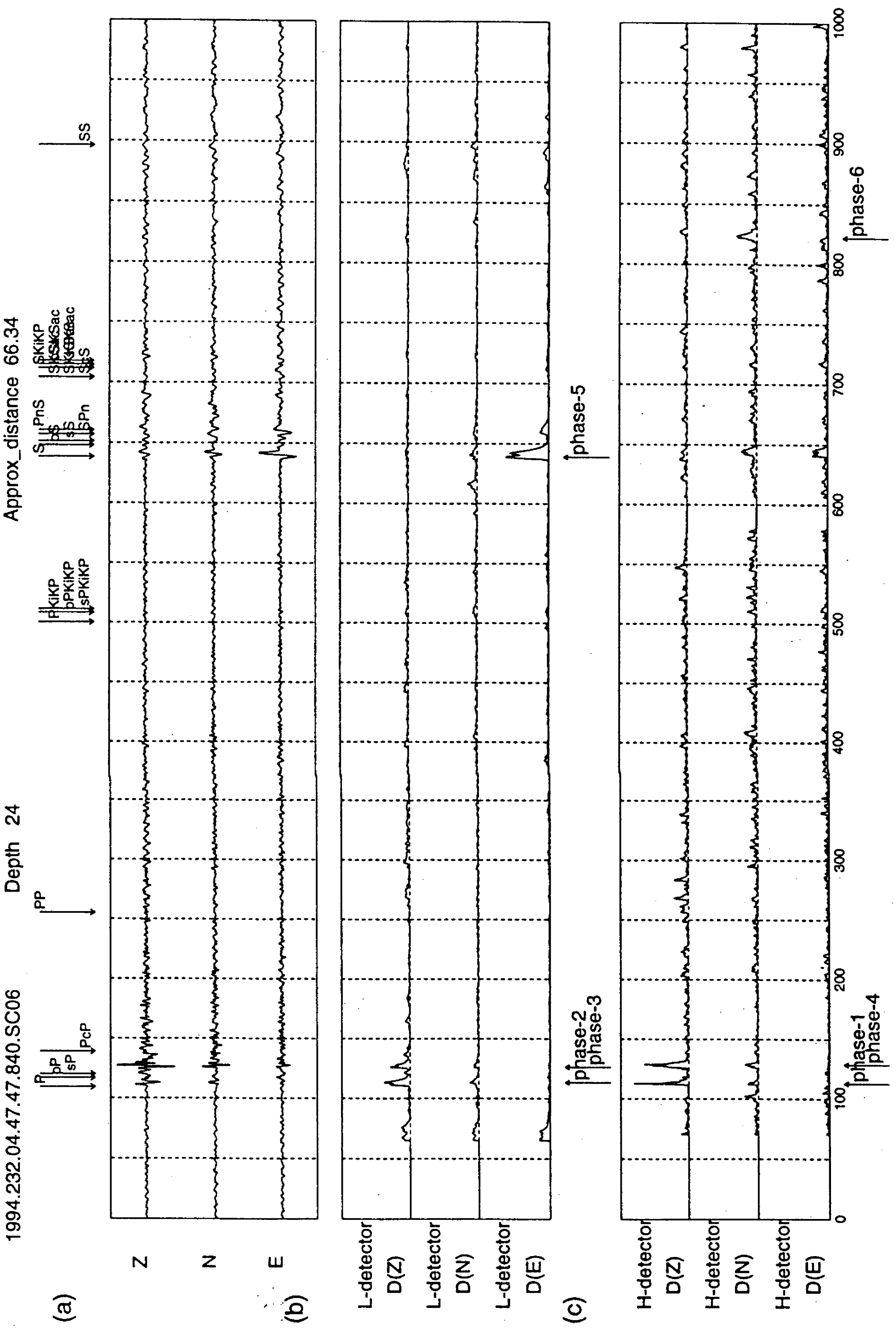
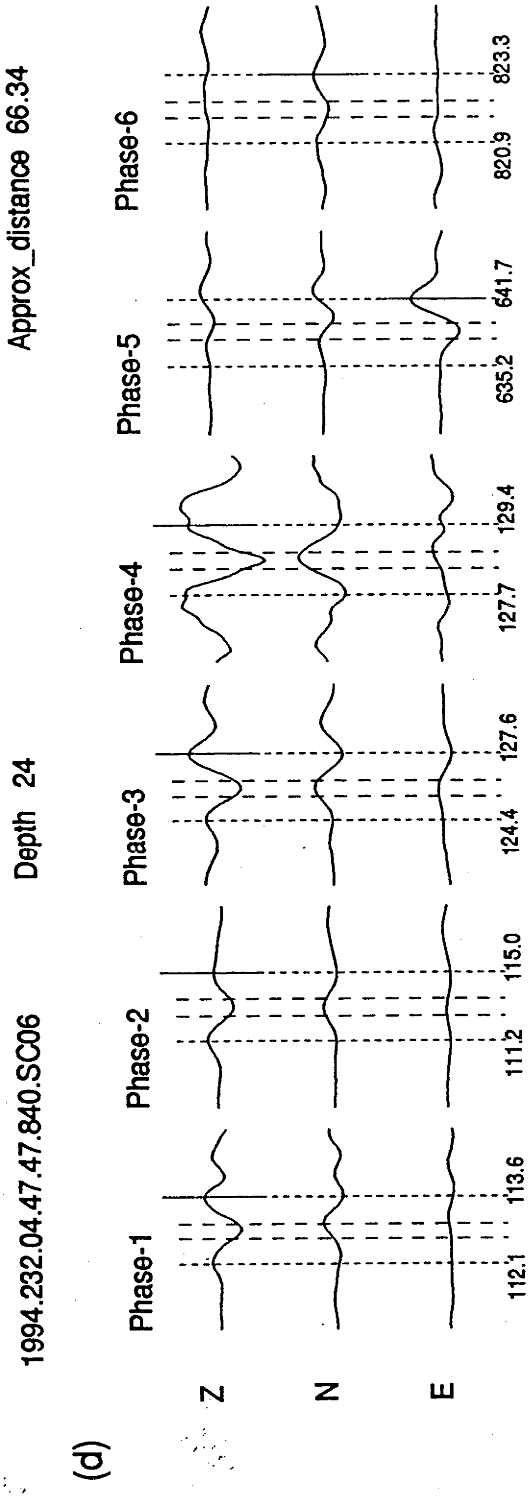


Fig. 7.4. Phase-detection and feature-extraction for event F at the station SC06.



	Phase-1	Phase-2	Phase-3	Phase-4	Phase-5	Phase-6
Ref. time	112.9	113.4	126.3	128.6	638.9	822.4
Amplitude	0.2184	0.1749	0.3002	0.5480	0.2017	0.0729
Frequency	0.6944	0.2604	0.3125	0.5952	0.1603	0.4167
Azimuth	8.26	9.18	24.13	14.01	266.79	185.90
App. incidence	24.73	30.28	28.21	29.88	6.90	5.90
P/S character	P	P	P	P	S	S

Fig. 7.4. – *continued.*

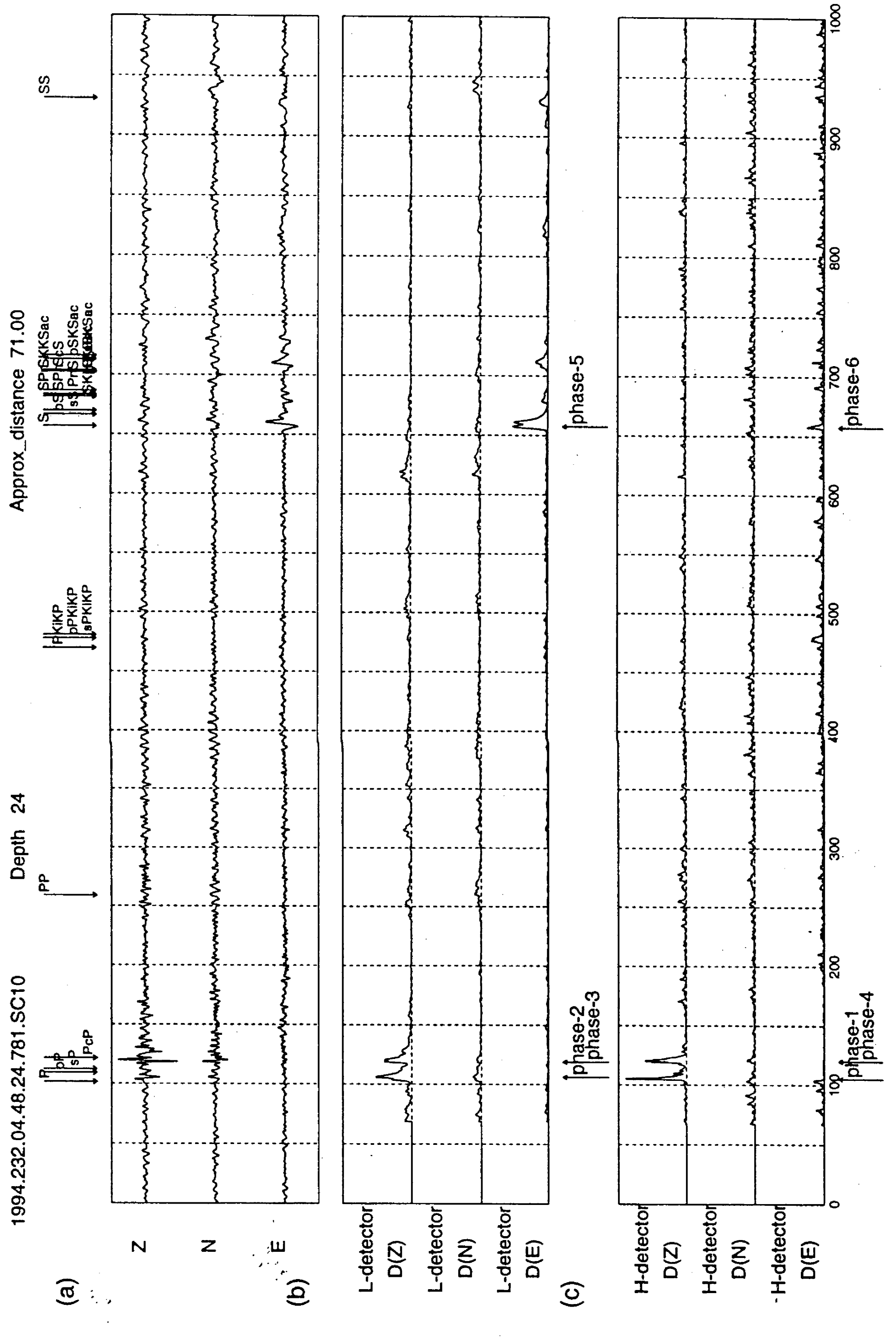
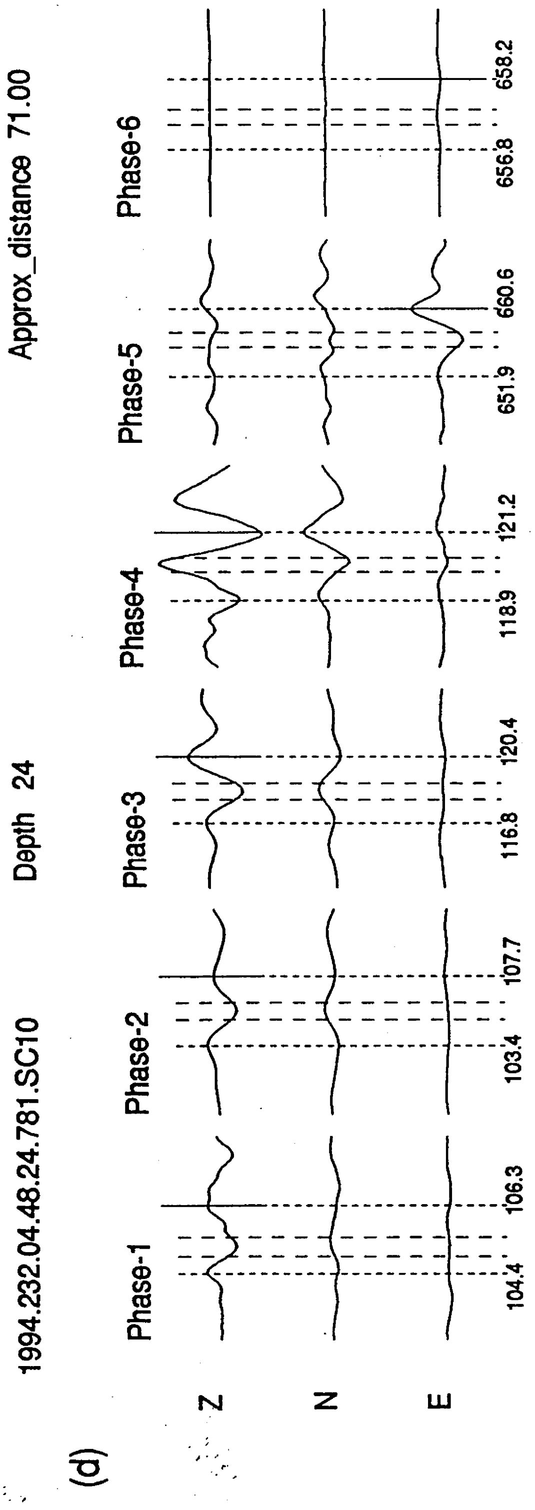


Fig. 7.5. Phase-detection and feature-extraction for event F at the station SC10.



Ref. time	105.3	106.1	119.0	120.2	657.4	657.6
Amplitude	0.1459	0.1962	0.2912	0.5324	0.2273	0.0178
Frequency	0.6250	0.2358	0.2841	0.4630	0.1263	0.6944
Azimuth	357.43	353.12	2.35	194.34	257.13	103.85
App. incidence	15.39	24.92	20.79	22.36	0.32	8.25
P/S character	P	P	P	P	S	S

Fig. 7.5. – *continued.*

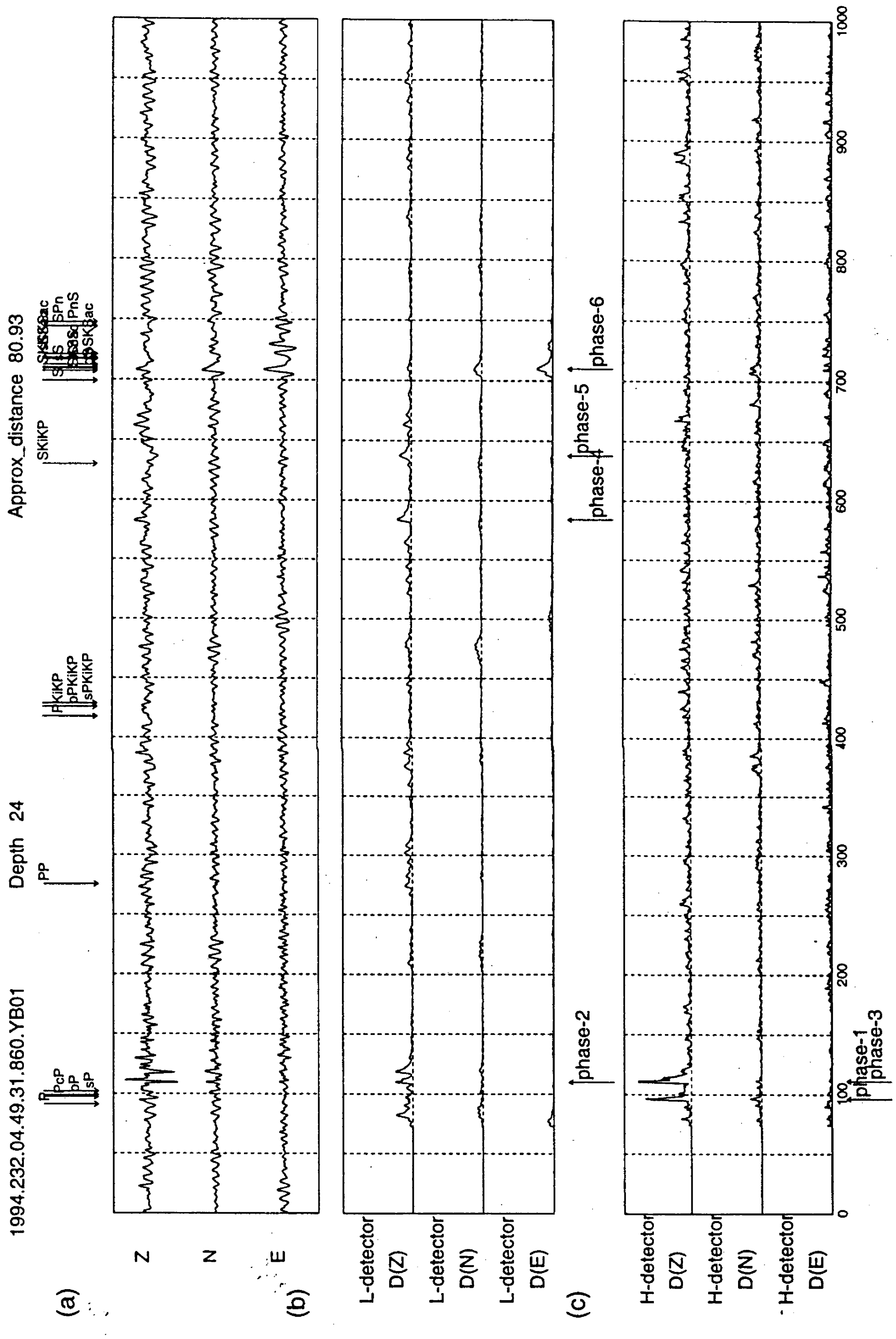
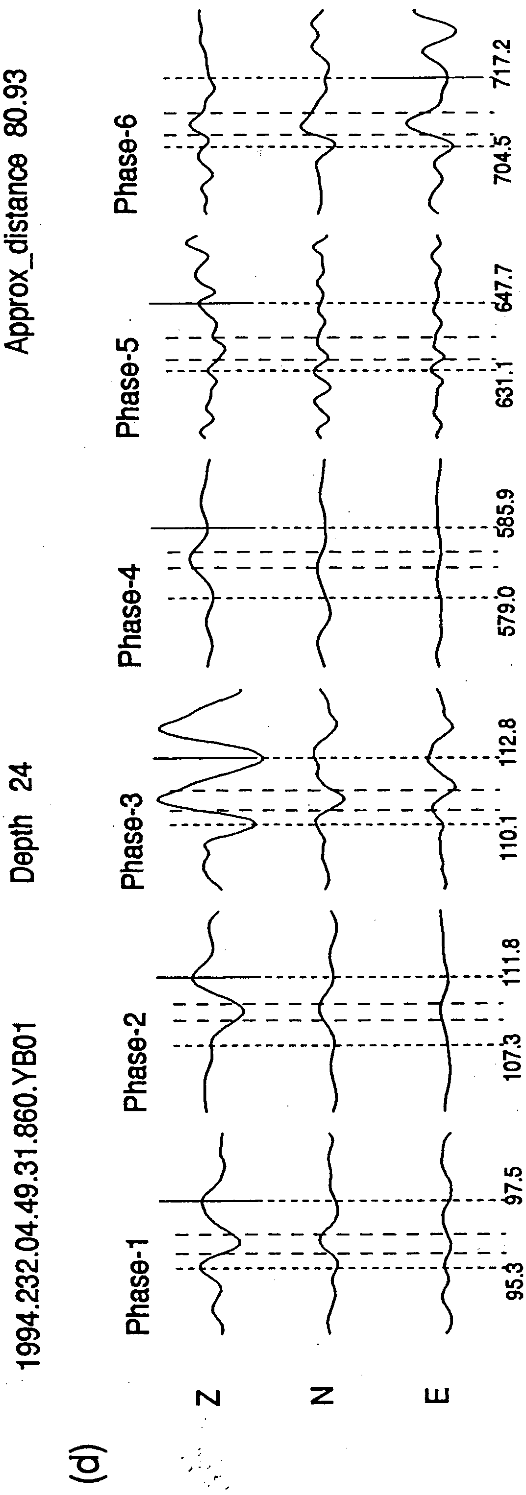


Fig. 7.6. Phase-detection and feature-extraction for event F at the station YB01.



Ref. time	96.3	110.0	111.3	583.9	637.6	710.0
Amplitude	0.2151	0.3005	0.4773	0.2001	0.1448	0.2917
Frequency	0.5952	0.2232	0.4808	0.1603	0.0926	0.1157
Azimuth	342.65	23.08	180.01	346.01	234.61	56.58
App. incidence	26.44	18.87	17.97	12.94	3.89	22.10
P/S character	P	P	P	P	P	S

Fig. 7.6. – *continued.*

Table 7.11. *Processing sequence for Event F at station SC06*

Step 1: Choice of Key Phases

Key *P*: phase 2, Key *S*: phase 5

Step 2: Test on Differential Time

525.5 s - surviving models: *P-S, PKP-SKKS*

Step 3: P-S Pair Feature Match

Surviving models: *P-S*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
P S	0	66	9 or 189	phase 2: *P*, phase 5: *S*	0.5970
	100	66	9 or 189	phase 2: *P*, phase 4: *pP*, phase 5: *S*	0.2543
	300	70	9 or 189	phase 2: *P*, phase 5: *S*	0.4918
	600	74	9 or 189	phase 2: *P*, phase 5: *S*	0.4918

Table 7.12. *Processing sequence for Event F at station SC10*

Step 1: Choice of Key Phases

Key *P*: phase 2, Key *S*: phase 5

Step 2: Test on Differential Time

551.3 s - surviving models: *P-S, PKP-SKKS*

Step 3: P-S Pair Feature Match

Surviving models: *P-S*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
P S	0	70	173 or 353	phase 2: *P*, phase 5: *S*	0.5643
	100	72	173 or 353	phase 2: *P*, phase 4: *pP*, phase 5: *S*	0.2388
	300	74	173 or 353	phase 2: *P*, phase 5: *S*	0.4657
	600	80	173 or 353	phase 2: *P*, phase 5: *S*	0.5058

Table 7.13. *Processing sequence for Event F at station YB01*

Step 1: Choice of Key Phases

Key *P*: phase 1, Key *S*: phase 6

Step 2: Test on Differential Time

613.7 s - surviving models: *P-S, P-SKS,* P_{diff}*-SKS, PKP-SKKS*

Step 3: P-S Pair Feature Match

Surviving models: *P-S, P-SKS*

Step 4: *Event Pattern Match*

P-S pair	Depth	Dist	Azim	Identified Phases	Misfit
P S	0	82	163 or 343	phase 1: *P*, phase 6: *S*	0.6630
	100	84	163 or 343	phase 1: *P*, phase 3: *pP*, phase 6: *S*	0.2975
	300	84	163 or 343	phase 1: *P*, phase 6: *S*	0.5102
	600	84	163 or 343	phase 1: *P*, phase 6: *S*	0.5102
P SKS	0	126	163 or 343	phase 1: *P*, phase 6: *SKS*	0.5693
	100	126	163 or 343	phase 1: *P*, phase 6: *SKS*	0.5693
	300	126	163 or 343	phase 1: *P*, phase 6: *SKS*	0.5102
	600	128	163 or 343	phase 1: *P*, phase 6: *SKS*	0.5102

in general the phase pair employed is *P-S*. Once *SKS* overtakes *S* the combination *P-SKS* has less distance and depth discrimination, as can be seen in the rather poor results for event H (Table 7.9). This can be expected from the rather flat *SKS-P* differential time curve as displayed in Figure 6.1. As discussed in the previous chapter, the resolution could be improved by chosing a number of choices of distance within an error tolerance rather than chosing only one in our present program.

At distances beyond 90° such combinations as *PP-PS* or *PP-SKS* can often be useful, as can be seen from the results for event C (Table 7.4). As indicated in the previous chapter, event C suffers from interference from another event of large amplitude. At the stations SC01 and SC03, the successful identification of the phase *SS* helps to produce quite accurate solutions of depth and distance. At the stations SC08 and SC10, the low frequency phase *SS* is contaminated by a high frequency phase from the other event. The failure to identify *SS* causes a relatively poor depth resolution.

For the largest ranges useful coarse estimates of range can be obtained (event D - Table 7.5) although as noted in the previous chapter the non-monotonic dependence of the differential time on range tends to force too small a range. Another contribution to the poor results for event D comes from the rather complicated phase packages on the seismograms. The situation of too many phases interfering with each other leads to a relatively poor phase detection, which consequently affects the resolution of event recognition. Improvements for phase detection in the presence of interfering phases will be a target for further development.

The various tests indicate that in general the performance of the automated analysis procedure is good. However, the original set of trigger levels prove to be somewhat conservative and some clear arrivals can be missed which, if included, would improve the resolution of the seismic parameters. A refinement of the current procedure to include a preliminary analysis with standard trigger parameters, followed by a second pass with tuned parameters and a finer sampling in epicentral distance and depth may well prove to be advantageous.

The examples also suggest that, a simultaneous application of this analysis procedure to a seismic array with several stations could improve the accuracy of the solution. The correlation of the results obtained from different stations can avoid problems arising at only one or two stations. The full power of the array can be used to provide accurate slowness and azimuth estimates for each of the identified phase. For a distributed network, real-time event recognition can be achieved by collecting preliminary results from each station, and then all the different possible solutions can be fed back to every individual station, and a final round of matching of later phases cross all stations can be used to determine the best solution. In such a way, efficient use can be made of later phases, so that the number of stations needed to locate a seismic event can be greatly reduced.

8

AI techniques applied in the automatic interpretation system

As a review of the automatic interpretation system developed from chapter 4 through to chapter 6, this chapter provides a general description of the whole work from the view point of Artificial Intelligence (AI). It gives a syntactic description of a seismogram, and explains how the AI techniques are applied in the implementation of the seismological interpretation system.

8.1 OVERVIEW OF THE AUTOMATIC INTERPRETATION SYSTEM

The whole system developed from chapter 4 through to chapter 6 is designed to simulate the interpretation by a human seismologist.

When an experienced seismologist reads a seismogram, she would proceed in two stages: phase picking and event interpretation. At the first stage, she picks out phases based on her experience. She could feel the difference in the appearances of phases, and tell which phases are similar. This kind of experience or skill can not be precisely explained. It is similar to the case of speech recognition, which is a task that we all perform extremely well, but none of us have much idea how we do it. Then, after the phases have been picked out, she would associate the different phases to a particular seismic event, and give out interpretations on the location of the event as well as the identities of the phases. In the second stage, a considerable amount of expert knowledge is applied. The expert knowledge includes: frequently observed phases generated by an event in a certain range of distance and depth; different appearances of phases associate with different identities. The seismologist might also need to check seismic travel time tables during reasoning.

In the automatic system, for the first stage, we apply Pattern Recognition techniques. The waveform of a phase is recognised as a hierarchical structure and the features characterising a phase are extracted via a structural analysis. For the second stage of event recognition, we apply Expert System techniques and pay close attention to the implementation details of how expert knowledge is represented, accessed (or organised) and applied.

The interpretation procedures performed by a human seismologist is compared with that used by the designed automatic system in Figure 8.1.

8.2 PATTERN RECOGNITION TECHNIQUES IN THE WORK OF PHASE CHARACTERISATION

In the work presented in chapter 4, a seismogram is analysed as a hierarchical structure. The technique of structural pattern recognition is also known as "syntactic pattern recognition".

8.2.1 The background in syntactic pattern recognition

I shall give a brief description of the background to syntactic pattern recognition based on the textbook by Fu (1974).

In order to represent the hierarchical structural information of a pattern, the syntactic approach draws an analogy between the structure of the patterns and the syntax of languages. Patterns are specified as building up out of subpatterns in various ways of composition just as phrases and sentences are built up by concatenating (or linking) characters. Evidently, for this approach to be advantageous, the

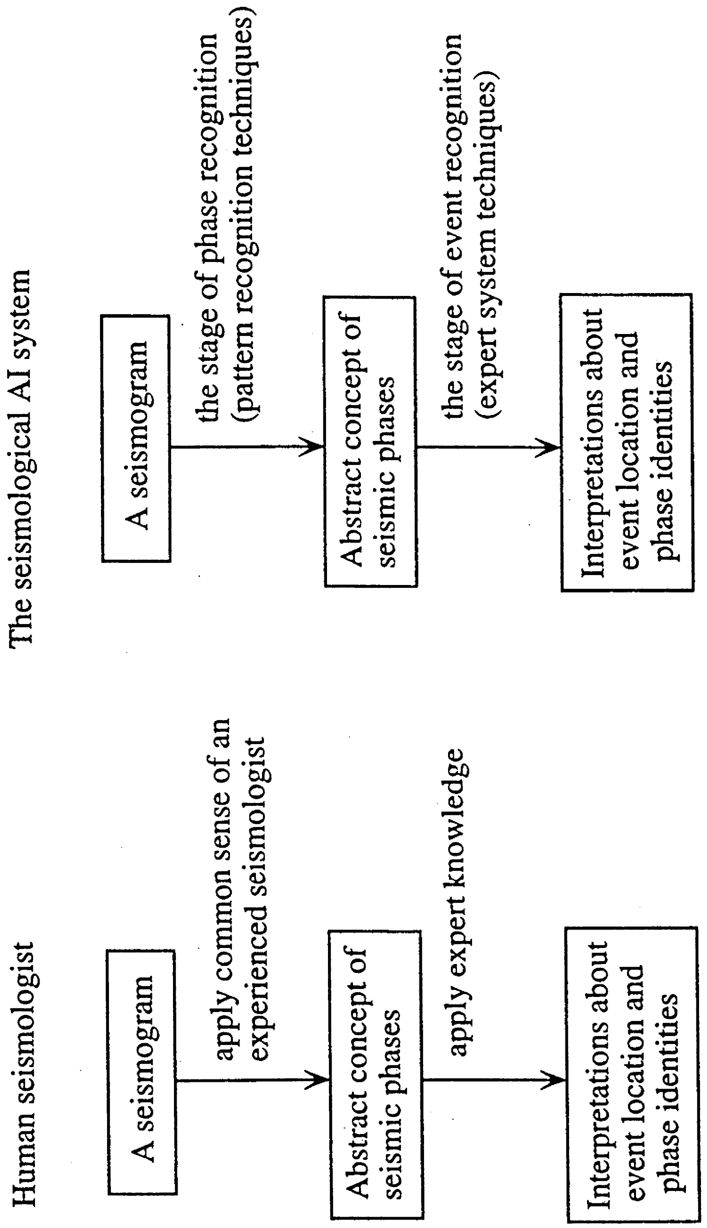


Fig. 8.1. Comparison of seismogram interpretation procedures performed by human seismologist and by the designed automatic system.

simplest subpatterns selected, called "**pattern primitives**", should be much easier to recognise than the patterns themselves.

The "language" which provides the structural description of patterns in terms of a set of pattern primitives and their composition operations, is called "**pattern description language**". The rules governing the composition of primitives into patterns are usually specified by the so-called "**grammar**" of the pattern description language.

After each primitive within the pattern is identified, the recognition process is accomplished by performing a syntax analysis (**parsing**) of the pattern to determine whether or not it is syntactically (or grammatically) correct with respect to the specified grammar. In the mean time, the syntax analysis also produce a structural description of the input pattern.

8.2.1.1 Phrase-structure grammar

The concepts of phrase-structure grammars originated in the parsing of a simple English sentence. Consider the parsing of the sentence, "The roses bloom luxuriantly". Let the symbol "$\longrightarrow$" mean "can be rewritten as", the previous sentence is formed by performing the following transcription (or "rewrite") rules:

$$\begin{aligned}
\langle\text{sentence}\rangle &\longrightarrow \langle\text{noun phrase}\rangle\ \langle\text{verb phrase}\rangle \\
\langle\text{noun phrase}\rangle &\longrightarrow \langle\text{adjective}\rangle\ \langle\text{noun}\rangle \\
\langle\text{verb phrase}\rangle &\longrightarrow \langle\text{verb}\rangle\ \langle\text{adverb}\rangle \\
\langle\text{adjective}\rangle &\longrightarrow \text{the} \\
\langle\text{noun}\rangle &\longrightarrow \text{roses} \\
\langle\text{verb}\rangle &\longrightarrow \text{bloom} \\
\langle\text{adverb}\rangle &\longrightarrow \text{luxuriantly}
\end{aligned}$$

The structure above can be abstracted to define the formalised phrase-structure grammar.

Definition 2.1 A phrase-structure grammar G is a four-tuple $G = (V_N,\ V_T,\ P,\ S)$

in which V_N and V_T are the nonterminal and terminal vocabularies of G, respectively. P is a finite set of transcription rules, called "**productions**", denoted by $\alpha \longrightarrow \beta$, where α and β are strings over V $(V = V_N \cap V_T)$, and with α involving at least one symbol of V_N. $S \in V_N$ is the starting symbol of a sentence.

So, in the example sentence above,

$$\begin{aligned} V_N &= \{\langle\text{sentence}\rangle,\ \langle\text{noun phrase}\rangle,\ \langle\text{verb phrase}\rangle,\ \langle\text{adjective}\rangle,\ \langle\text{noun}\rangle,\ \langle\text{verb}\rangle, \\ &\quad\ \langle\text{adverb}\rangle\} \\ V_T &= \{\text{the, roses, bloom, luxuriantly}\} \\ S &= \langle\text{sentence}\rangle \end{aligned} \tag{8.1}$$

and the production set P is just the set of transcription rules given previously.

Chomsky divided the phrase-structure grammars into four types according to the forms of the productions. In type 0 (unrestricted) grammars, there is no restriction on the productions, which makes it too general to be useful.

In type 1 (context-sensitive) grammars, the productions are restricted to the form

$$\zeta_1 A \zeta_2 \longrightarrow \zeta_1 \beta \zeta_2,$$

where $A \in V_N$; $\zeta_1, \zeta_2, \beta \in V^*$ (V^* is the set of all finite-length strings of symbols in the finite set of symbols V, including λ, the string of length 0); and $\beta \neq \lambda$.

In type 2 (context-free) grammars, the productions are of the form "$A \longrightarrow \beta$", where $A \in V_N$ and $\beta \in V^+ (V^+ = V^* - \{\lambda\})$.

Regular grammar

The productions of type 3 (regular or finite-state) grammars are of the form $A \longrightarrow aB$ or $A \longrightarrow b$, where $A, B \in V_N$ and $a, b \in V_T$. Note that A, B, a, b all are single symbols. This is the grammar we use in this paper to describe the structure of seismograms.

8.2.1.2 Recognition devices

A grammar, as introduced above, can be regarded as a finite specification of a language from the viewpoint of generating sentences. An alternative way of specifying

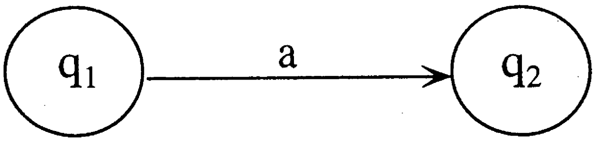


Fig. 8.2. Graphical representation of the transition $\delta(q_1,\ a) \ = \ q_2$.

a language is in terms of the set of strings that are accepted by a certain recognition device (a recognition viewpoint). Here, we introduce a simple recogniser, called a finite-state automaton, which is relevant to our work on phase recognition.

Definition 2.2 A (deterministic) finite-state automaton A is a quintuple

$$A \ = \ (\Sigma,\ Q,\ \delta,\ q_0,\ F),$$

in which

(1) Σ is a finite set of input symbols (alphabet),
(2) Q is a finite set of states,
(3) δ is a mapping of $Q \ \times \ \Sigma$ into Q,
(4) $q_0 \in Q$ is the initial state, and
(5) $F \subseteq Q$ is the set of final states.

A convenient representation for the mapping "$\delta(q_1,\ a) \ = \ q_2$" is the state transition diagram shown in Figure 8.2. The interpretation for this diagram is that, the automaton in state q_1 and scanning the input symbol a, goes to state q_2 and the input device moves one step to the next position (ready to get the next input symbol).

Theorem 2.1 Let G be a regular grammar, then there exists a finite-state automaton A, that the languages generated by G can always be accepted by the automaton A.

8.2.2 Syntactical structure for a seismogram

In chapter 4, a seismogram is recognised as a sequence of "phase" or "non-phase" segments, while a phase is a structure constructed from a number of "units". Note that a "unit" of seismogram is basically a hill shape or valley shape segment which represents half a local cycle. The hierarchical structure has been built from the bottom (pattern primitives) to the top (a complete pattern of seismogram).

8.2.2.1 Pattern primitives in a seismogram

In the construction of the description structure for a seismogram, six pattern primitives have been used. All the units generated by an intelligent segmentation procedure can be classified into the six groups of pattern primitives:

(1) U_b: a unit of non-phase waveform, i.e., background noise;
(2) U_o: the first unit where a phase wavelet begins;
(3) U_l: a unit on the left side of a phase wavelet;
(4) U_r: a unit on the right side of a phase wavelet;
(5) U_a: the last unit of a phase wavelet;
(6) U_n: an uncoherent unit which ends the phase before the pattern of wavelet could complete normally.

As introduced in chapter 4, phase-detections (based on STA/LTA triggers) are continuously generated in real time, which makes available the information of whether the current unit begins a phase wavelet. We use a two-value flag, called a "phase-flag" in the procedure of primitive identification. Its initial value is "off", which indicates no phase is being analysed at the current time. When a phase is indicated by the phase-detector, the phase-flag is set "on", which means a phase is being analysed. It will be turned "off" later when the current phase wavelet is recognised as ended.

These primitives can be identified as following:

(1) When the phase-flag is off, the unit will be recognised as U_o or U_b according to the current phase-detection value. if the phase-detector indicate a detected phase,

the current unit is identified as U_o, and the phase-flag is turned "on"; otherwise the current unit is identified as U_b.

(2) When the phase-flag is on, the current unit will be recognised as:

U_l , if the amplitude of current unit is greater than the amplitude of former unit, the former unit is U_l or U_o, and the duration is consistent with U_o;

U_r , if the current amplitude is smaller compared with the former one but greater than background noise, and the duration is consistent,

U_a , if the current amplitude is within the amplitude range of background noise;

U_n , if the duration of the current unit is not consistent with U_o;

Having the pattern primitives defined as above, we can have a relatively simple pattern grammar to describe the pattern structure.

8.2.2.2 Pattern grammar for a seismogram

A regular grammar G has been constructed to describe any seismogram with any number of phases in it.

$$G = (V_N,\ V_T,\ P,\ \langle \mathit{Seismogram} \rangle),$$

where $\langle \mathit{Seismogram} \rangle$ is the start symbol;

$$V_N = \{\langle \mathit{Seismogram} \rangle,\ \langle \mathit{Phase} \rangle,\ \langle \mathit{Left} \rangle,\ \langle \mathit{Right} \rangle\}$$

is the set of nonterminals;

$$V_T = \{U_b,\ U_o,\ U_l,\ U_r,\ U_a,\ U_n\}$$

is the set of terminal symbols;

and the productions are

$P = \{$

$$\langle \mathit{Seismogram} \rangle \longrightarrow U_b\ \langle \mathit{Seismogram} \rangle,$$

$$\langle \mathit{Seismogram} \rangle \longrightarrow U_o\ \langle \mathit{Phase} \rangle,$$

$$\langle \mathit{Phase} \rangle \longrightarrow U_l\ \langle \mathit{Left} \rangle,$$

$$
\begin{aligned}
\langle Phase\rangle &\longrightarrow U_r\ \langle Right\rangle,\\
\langle Phase\rangle &\longrightarrow U_a\ \langle Seismogram\rangle,\\
\langle Phase\rangle &\longrightarrow U_n\ \langle Seismogram\rangle,\\
\langle Left\rangle &\longrightarrow U_l\ \langle Left\rangle,\\
\langle Left\rangle &\longrightarrow U_r\ \langle Right\rangle,\\
\langle Left\rangle &\longrightarrow U_a\ \langle Seismogram\rangle,\\
\langle Left\rangle &\longrightarrow U_n\ \langle Seismogram\rangle,\\
\langle Right\rangle &\longrightarrow U_r\ \langle Right\rangle,\\
\langle Right\rangle &\longrightarrow U_n\ \langle Seismogram\rangle,\\
\langle Right\rangle &\longrightarrow U_a\ \langle Seismogram\rangle,
\end{aligned}
$$

$\}$.

Note that the previous identification procedures for pattern primitives are actually context-sensitive, while the pattern grammar is context-free. The relatively simple context-free grammar can be easily recognised.

8.2.2.3 Pattern recognition device

The pattern described by the previous regular grammar can be recognised by a deterministic finite-state automaton A.

$$A = (\ \Sigma,\ Q,\ \delta,\ q_0,\ F\ ),$$

where,

$$\Sigma = \{U_b,\ U_o,\ U_l,\ U_r,\ U_a,\ U_n\}$$

is the set of input symbols (alphabet);

$$Q = \{q_0,\ q_1,\ q_2,\ q_3\}$$

is the set of states; the set of transitions δ is listed in Table 8.1; q_0 is the initial state; and the set for final states F is

$$F = \{q_0,\ q_1,\ q_2,\ q_3\}.$$

The state transition diagram of the automaton A is shown in Figure 8.3.

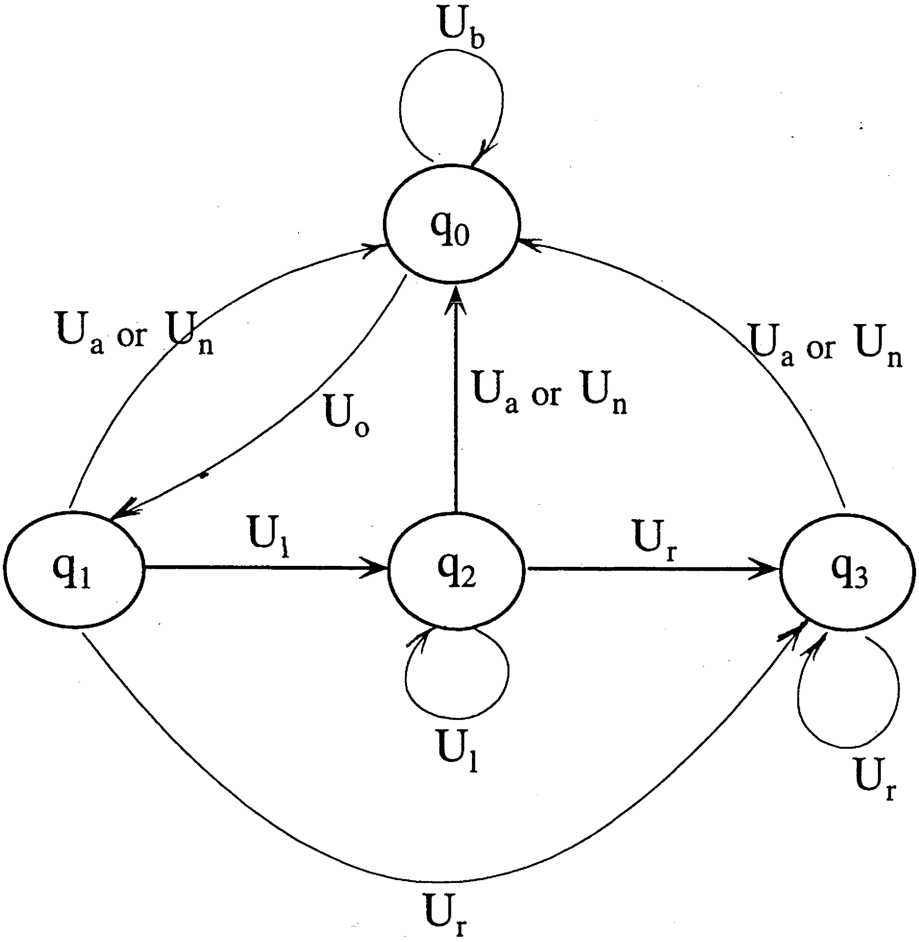


Fig. 8.3. State transition diagram of the phase-recognition automaton *A*.

Table 8.1. *Transitions in the automaton A*

$$
\left\{
\begin{array}{lll}
\delta\,(q_0,\,U_b) = q_0, & \quad & \delta\,(q_0,\,U_o) = q_1, \\
\delta\,(q_1,\,U_l) = q_2, & & \delta\,(q_1,\,U_r) = q_3, \\
\delta\,(q_1,\,U_a) = q_0, & & \delta\,(q_1,\,U_n) = q_0, \\
\delta\,(q_2,\,U_l) = q_2, & & \delta\,(q_2,\,U_r) = q_3, \\
\delta\,(q_2,\,U_a) = q_0, & & \delta\,(q_2,\,U_n) = q_0, \\
\delta\,(q_3,\,U_r) = q_3, & & \delta\,(q_3,\,U_n) = q_0, \\
\delta\,(q_3,\,U_a) = q_0 & &
\end{array}
\right\}
$$

Note 1: a seismogram can be terminated at any state among the four choices in Q.

Note 2: the four states in Q actually correspond to the four nonterminals in the grammar G which is given in the previous section.

Note 3: the two pattern primitives, U_a and U_n have the same effect of ending a phase in this preliminary implementation, therefore they can be combined into one. However, the identification of U_n may help to indicate the onset of an interfering phase in future developments of the system.

The phase recognition system as presented in chapter 4 has been built based on the automaton described above.

8.3 EXPERT SYSTEM TECHNIQUES IN THE WORK OF EVENT RECOGNITION

As described in chapter 6, event recognition is a complicated problem which involves the application of an incomplete domain of expert knowledge. From the AI point of view, the main problems in designing an appropriate expert system are: (1) how to represent the expert knowledge (including experience); (2) how to apply the unordered expert knowledge during the search for a solution. "Knowing what one knows, and knowing when and how to use it", seems to be an important part of expertise; this is usually termed "meta-knowledge", i.e., knowledge about knowledge.

8.3.1 Representation of expert knowledge – parameters with adjustable values

The expert knowledge collected in the automatic system developed in chapter 6 includes travel time tables, differential time tables, and expert experience on the character of seismograms summarised in Table 6.1 – Table 6.4.

Some of the knowledge as obtained from a human expert is not immediately suitable for representation in the computer, such as the term "steep high frequency phase" (which is used to describe the appearance of phase *PcP*). We introduce parameters to represent this kind of knowledge. For example, as indicated in Table 6.2, we use an angle of incidence parameter with a preliminary upper boundary value of 16° to define the term "steep". Similarly, we introduce frequency boundaries for particular phases (e.g., *PcP*, *PP* etc). And an arrival time retrieved from travel time tables is always translated into a time range with a tolerance for small difference. The values for these parameters may need to be adjusted during testing to ensure a reliable performance.

8.3.2 Knowledge application in the procedure of seismic event recognition

The procedure of knowledge application is a procedure of reasoning. It is a kind of engine in an expert system – called an "inference engine" in the field of Artificial Intelligence (AI).

Many expert system shells (software tools which help to construct expert systems) use the architecture of a rule-based system (see e.g. Ford, 1991), while rule-based systems use decision trees to implement their inference engines.

It would be very difficult to apply this kind of architecture to our present application of seismic event interpretation. Because at the present time, we don't have a precise and complete rule set and the limited domain of knowledge is subject to continuous modifications and improvements due to the continuous development of seismological expertise with broad-band data. The identification of seismic phases is by no means a trivial exercise. In fact, many modern-day seismologists have little direct experience in the routine "reading" of seismograms. Even experienced seismologists can misidentify arrivals, given the many possibilities. The ISC (International Seismological Centre) often reidentifies phases picked by station operators who do not have accurate location estimates.

Since the unknown factors (origin time, distance, depth) are interacting with each other through the expert knowledge in a complicated way and the available expert knowledge is limited, using a normal rule-based architecture, the system could be bogged down in endless searching.

To solve the problem of knowledge application in seismic event interpretation, a new technique of "assumption tree" is proposed in chapter 6. It was inspired from a typical human research strategy – making assumptions and then using the inferences arising from the assumptions. It is an attempt to make an efficient application of a limited domain of expert knowledge while there are a number of interacting unknown features to be constrained.

In the specific application of seismic event interpretation, a three level assumption tree has been used to guide the application of expert knowledge.

At the first stage, the tree grows out nine branches which represent nine different assumptions on the feature of "P-S" pair. Then the knowledge base which only depends on the "P-S" choice (represented by Table 6.1 and Table 6.2) can be applied. The inferred result, a constraint on some other features, is used to reject the current choice or to constrain choices for later level unknown features. In other words, the result of the application of knowledge base is used to prevent the tree from growing to its full size.

Similarly, at the second stage, every surviving node from the first level branches into four child nodes, each with a different choice of depth. Then, knowledge which depends on "P-S" choice and depth choice, represented by differential time tables, can be applied to constrain the distance.

When the distance is known, the knowledge which depends on distance, represented by Table 6.1, Table 6.2, and travel time tables, can be applied. And the result of the knowledge-application is used to measure the fitness of the current interpretation.

The separation of domain knowledge from the inference engine makes it easy to update the expert knowledge in the future. The entities in the seismological expert system and their relationships are illustrated in Figure 8.4.

The domain knowledge is partitioned into six groups: Table 6-1 – Table 6-4, travel time tables and differential time tables, which represent distinctive specialists (or independent knowledge sources). The application of these knowledge sources is scheduled by the assumption tree. Solutions are built up on a global data structure, the "information package", which represents a type of "blackboard" (which is a technique in the field of expert system). This particular type of blackboard – the information package – has multiple existences during processing. They are dynamically created and evoluted with the development of the assumption tree. In fact, the assumption tree method combines the blackboard and reasoning scheduler into an integral tree structure.

Thus, by making assumptions and grouping knowledge items according to their premises (or conditions), the assumption tree method has enabled efficient knowledge application in the complicated problem of seismic event interpretation.

8.3.3 A complementary description of the assumption tree method

The assumption tree is a new automatic reasoning method which is presented in this thesis. Although proposed for the design of a seismological expert system, it is a general approach to an efficient automatic reasoning with limited domain expert knowledge.

This section gives a generalised description of the assumption tree method and some important design criteria when applying this method. The material presented may guide to design a new type of expert system shell – "an assumption tree expert system shell" – for a wide range of expert-system applications.

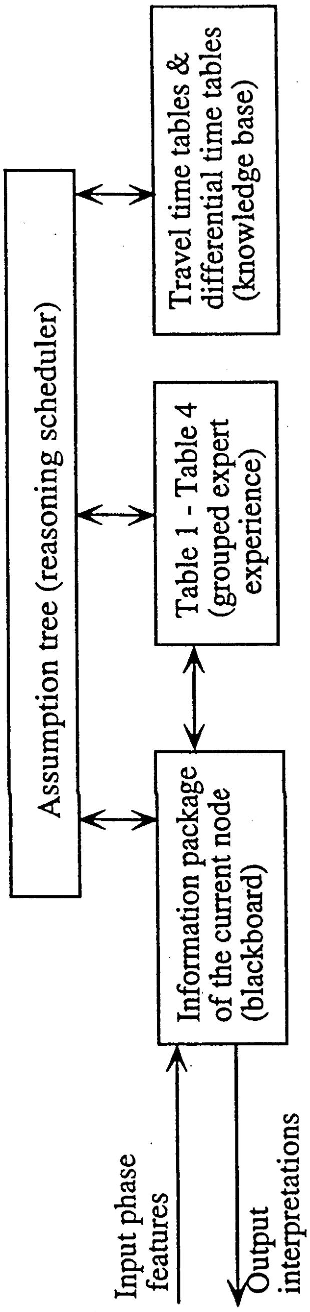


Fig. 8.4. Entities and their relationships in the seismological expert system.

8.3.3.1 Suitable applications of the assumption tree method

The application problems which fall under the umbrella of the assumption tree method have the following common description: given a collection of expert knowledge and a partial description of an unknown object, to constrain the choices for the object. The "partial description" refers to a constraint on part of the features which describe the object. The expert knowledge which is available is a collection of procedures or rules, each of which produces a constraint on a subset of features following from a constraint on another subset of features.

To help later discussions, denote all the features which describe the object as

$$F = \{F_1, F_2, F_3, ..., F_i, ..., F_m\}.$$

The number of choices for F_i is denoted as n_i, and all the possible choices for F_i can be represented as

$$R_i = \{f_{i,1}, f_{i,2}, f_{i,3}, ..., f_{i,n_i}\}.$$

8.3.3.2 The assumption tree: structure and reasoning algorithm

The structure of an assumption tree is given in Figure 6.3. Every feature (or factor) corresponds to a specific level (or generation) in the assumption tree. Branching of the tree at a specific level is made according to the choices of the corresponding feature.

Knowledge application is organised according to the level in the assumption tree. At each level, a different set of expert knowledge is retrieved. As ordered from root to leaves in Figure 6.3 (F_1, F_2, ..., F_m), the domain knowledge items is divided into m groups according to the order they will be applied :

(S_1) = knowledge items which only rely on the choice of F_1,
(S_2) = (knowledge items which rely on F_1 and F_2) - S_1,
(S_3) = (knowledge items which rely on F_1, F_2 and F_3) - S_2 - S_1,
......
(S_m) = (knowledge items which rely on F_1, F_2, ..., and F_m) - S_1 - S_2 - ...... - S_{m-1}.

At the ith level of the assumption tree, knowledge set S_i is applied.

Every node in the assumption tree carries an information package (denoted as I) which records the current constraints on all features:

$$I = \{I_1, I_2, I_3, ..., I_i, ..., I_m\}, \qquad I_i \subseteq R_i.$$

The reasoning algorithm begins from the root node. The information package at the root may contain a constraint on some of the features (which represents individual evidence), and full ranges of choices for the other features. The root node branches into n_1 child nodes according to the number of choices for the first level feature F_1 (which is recorded in the information package), and the information package of the root is copied into every child node. Each child node has a different choice for F_1 (denoted as $f_{1,i}$), therefore, in its information package the constraint on F_1 is updated with the only choice of $f_{1,i}$. Then, every node searches in the knowledge set S_1 for applicable items from the knowledge base. The application of an item of knowledge will produce new constraints on some other features. For a new constraint for the feature F_i, denoted as I_i': when $I_i' \cap I_i$ is an empty set, a contradiction occurs, and the growth of the current node is terminated; otherwise, the old constraint I_i is replaced with $I_i' \cap I_i$. After the procedure of the application of knowledge, a surviving node continues to branch according to the choices for the next level feature, and the procedure will continue recursively until the current node is rejected or a leaf of the assumption tree is reached.

8.3.3.3 How to arrange features in an assumption tree

While applying the assumption tree method to a general application of expert system design, a question arises: how to place the m features on different levels in the assumption tree so that it can work most efficiently ? That is, which permutation to choose from $m!$ possible choices ?

Suppose: (1) all the knowledge items will be used to terminate a current node, (2) the order of F_1, F_2, ..., F_m is chosen as the order from root to leaves in the assumption tree.

For an item of knowledge, the earlier it is applied, the more efficient it is used.

For example, application of an item of knowledge in the first generation will reject $n_2 \times n_3 \times n_4 \times ...... \times n_m$ solutions (note: n_i is the number of choices for the feature F_i), while application of an item in the last generation can only reject one solution. Thus, the number of choices for every feature shall be taken into consideration. The feature with more choices might better be placed at a later generation.

Since all knowledge items compete with each other for earlier application, we shall take the order from the earliest generation to the latest while arranging the features on the assumption tree to achieve the best integral efficiency.

As described previously, the knowledge set which is applied at the ith level is S_i. Let the number of knowledge items in S_i be k_i, and k_i represents the upper limit of the number of knowledge items which could be applied at the ith generation.

As mentioned above, firstly we shall choose a feature (from the m features) for the first generation. The number of solutions which would be eliminated at the first generation can be measured as

$$Q_1 = n_2 \times n_3 \times n_4 \times ...... \times n_m \times k_1.$$

The greater the value for Q_1, the higher the quality of the assumption tree.

For the second generation, the actual number of knowledge items applied at run time would depend on individual applications. This is because that, the first generation applications might have cut off some branches so that the second generation would not grow to full size. However, we can still use the upper limit number k_2 to evaluate the average efficiency. The efficiency of the second generation can still be measured as

$$Q_2 = n_3 \times n_4 \times n_5 \times ...... \times n_m \times k_2.$$

Generally, the quality of the ith generation is given by

$$Q_i = n_{i+1} \times n_{i+2} \times n_{i+3} \times ...... \times n_m \times k_i.$$

We can use this quality measure to choose appropriate features for all the levels of the assumption tree from root to leaves.

8.3.3.4 The choices for the "assumption features"

The features which appear as specific levels in an assumption tree, named "assumption features", are not limited only to the apparent features which describe the object, they can also include some redundant classification features, which could sometime greatly improve the performance of knowledge application. For example, in the previous seismological interpretation system, the classification of "P-S" pair models is a redundant classification which is a combination of a distance constraint and two constraints on phase-identity. When some combinations of constraints on original features frequently appear in the premise part of some knowledge items, they may suggest a good choice of a redundant classification feature. Whether or not to include this kind of choice into the structure of the assumption tree can be decided by using the quality measure which is described above.

8.3.3.5 The new idea in the assumption tree method

Although the assumption tree method has been designed for a particular group of applications, we can compare it with other general reasoning methods in a characteristic aspect.

The major new idea in the assumption tree method is to schedule reasoning according to the expert knowledge which is available.

In the field of expert system, typical examples of the control mechanism for reasoning are "backward chaining" (or "backward reasoning") and "forward chaining" (or "forward reasoning") (see e.g. Jackson, 1990). Backward chaining entails taking a conclusion first, then seeking evidence to support it. Forward chaining works in the reverse direction, the system tries to find a rule whose conditions are now satisfied by the information. In the case of insufficient evidence and insufficient expert knowledge while the choices for the solution are many, using the backward chaining method the reasoning procedure may never terminate; while using the forward chaining the reasoning procedure might not be able to progress to any solution (due to insufficient evidence).

In contrast, in the assumption tree method, the reasoning procedure is chained neither from evidence to solution nor the other way around, it is directed in a predecided order which has taken into consideration the distribution of knowledge with the features (as described in section 8.3.3.3). The new idea is to follow from "what one knows" to decide "when and how to use it". In this way, the assumption tree method could apply knowledge more efficiently in the case of insufficient knowledge and evidence, and the reasoning procedure is guaranteed to terminate in any circumstance.

Bibliography

Anderson, D. L., 1961. Elastic wave propagation in layered anisotropic media, *J. Geophys. Res.*, **66**, 2953-2963.

Anderson, D. L., 1989. Theory of the Earth, Blackwell Scientific Publications, Boston, 366 pp.

Anderson, D. L. & Dziewonski, A. M., 1982. Upper-mantle anisotropy: evidence from free oscillations, *Geophys. J. Royal Astron. Soc.*, **69**, 383-404.

Ando, M., Ishikawa Y. & Yamazaki, F., 1983. Shear-wave polarization-anisotropy in the upper mantle beneath Honshu, Japan, *J. Geophys. Res.*, **88**, 5850-5864.

Booker, A. & Mitronovas, W., 1964. An application of statistical discrimination to classify seismic events, *Bull. seism. Soc. Am.*, **54**, 961-971.

Bowman, J. R., & Ando, M., 1987. Shear-wave splitting in the upper-mantle wedge above the Tonga subduction zone, *Geophys. J. R. Astron. Soc.*, **88**, 25-42.

Breiman, L., Friedman, J., Olshen, R., & Stone, C., 1984. *Classification and regression trees*, Belmont, CA: Wadsworth International group.

Červený, V., Molotkov, I. A. & Pšenčík, I., 1977. Ray method in seismology, Univerzita Karlova, Praha.

Chen, C. H., 1982. Application of pattern recognition to seismic wave interpretation, *Applications of pattern recognition*, 107-119, Boca Raton, Fla : CRC Press.

Christensen, N. I., & Salisbury, M. H., 1979. Seismic anisotropy in the upper mantle: Evidence from the Bay of Islands ophiolite complex, *J. Geophys. Res.*, **84**, 4601-4610.

Clarke, T.J., 1993. The complete ordered ray expansion - I Calculation of synthetic seismograms, *Geophys. J. Int.*, **115**, 421-434.

Dey, S.C., Kennett, B.L.N., Bowman, J.R. & Goody, A., 1993. Variations in the upper mantle velocity structure under northern Australia, *Geophys. J. Int.*, **114**, 304-310.

Duffy, T., & Anderson, D. L., 1989. Seismic velocities in mantle minerals and the mineralogy of the upper mantle, *J. Geophys. Res.*, **94**, 1895-1912.

Dziewonski, A. M., & Anderson, D. L., 1981, Preliminary reference earth model, *Phys. Earth Planet. Inter.*, **25**, 297-356.

Earle, P. S. & Shearer, P. M., 1994. Characterisation of global seismograms using an automatic-picking algorithm, *Bull. Seism. Soc. Am*, **84**, 366-376.

Ford, N., 1991. Expert systems and artificial intelligence, Library Association Publishing, London.

Fu, K. S., 1974. Syntactic methods in pattern recognition, 295 pp., Academic Press, New York and London.

Fuchs, K., 1983. Recently formed elastic anisotropy and petrological models for the continental, subcrustal lithosphere in southern Germany, *Phys. Earth Planet. Inter.*, **31**, 93-118.

Fukao, Y., 1984. ScS evidence for anisotropy in the earth's mantle, *Nature*, **309**, 695-698.

Goetze, C., & Kohlstedt, D. L., 1973. Laboratory study of dislocation climb and diffusion in olivine, *J. Geophys. Res.*, **78**, 5961-5971.

Goody, A., 1991. Broad-band studies of the upper mantle beneath northern Australia, *honours thesis*, Aust. Natl. Univ., Canberra.

Gudmundsson, O., Kennett, B.L.N., & Goody, A., 1994. Broad-band observations of upper mantle seismic phases in northern Australia and the attenuation structure in the upper mantle, *Phys. Earth Planet. Inter.*, **84**.

Hager, B. B., & O'Connell, R., 1979. Kinematic models of large-scale flow in the earth's mantle, *J. Geophys. Res.*, **84**, 1031-1048.

Hearn, T, 1984. Pn travel times in Southern California, *J. Geophys. Res.*, **89**, 1843-1855.

Hendrajaya, A.L., 1981. A study of S-body wave velocity structure in the mantle down to 1100 km, using the Warramunga seismic array, **Ph.D. thesis**, Aust. Natl. Univ., Canberra.

Hess, H., 1964. Seismic anisotropy of the uppermost mantle under oceans, *Nature*, **203**, 629.

Horowitz, S. L., 1977. Peak recognition in waveforms, *Syntactic pattern recognition, Applications*, 31-49, Springer-Verlag Berlin Heidelbeg, New York.

Jackson, P., 1990. Introduction to expert systems, 2nd ed., Addison-Wesley.

Journet, B., & Jobert, N., 1982. Variation with age of anisotropy under oceans, from great-circle surface waves, *Geophys. Res. Lett.*, **9**, 179-181.

Karato, S., & Li, P., 1992. Diffusive creep in perovskite: Implications for the rheology of the lower mantle, *Science*, **255**, 1238-1240.

Kennett, B.L.N., Gudmundsson, O., & Tong, C., 1994, The upper mantle S and P velocity structure beneath Northern Australia from broad-band observations, *Phys. Earth Planet. Int.*, **86**, 85-98.

Kennett, B.L.N., 1991. The removal of free surface interactions from three-component seismograms, *Geophys. J. Int.*, **104**, 153–163.

Kennett, B. L. N. & Engdahl, E. R. 1991. Travel times for global earthquake location and phase association, *Geophys. J. Int*, **105**, 429-465.

Kennett, B. L. N., 1991. *IASPEI 1991 Seismological Tables*, Bibliotech, Canberra.

Kennett, B.L.N., 1993. The distance dependence of regional phase discriminants, *Bull. seism. Soc. Am.*, **83**, 1155–1166.

Kennett, B. L. N., 1995. Event location and source characterisation, in *Monitoring a Comprehensive Test Ban Treaty*, eds E. S. Husebye & A. M. Dainty, Kluwer, Dordrecht.

Mainprice, D. & Silver, P. G., 1993. Interpretation of SKS waves using samples from the subcontinental lithosphere, *Phys. Earth Planet. Int.*, **78**, 257-280.

Montagner, J-P., & Tanimoto, T., 1991. Global upper-mantle tomography of seismic velocities and anisotropy, *J. Geophys. Res.*, **96**, 20337-20351.

Mykkeltveit, S., Ringdal, F., Kvaerna, T., & Alewine, R.W. (1990) Application of regional arrays in seismic verification research, *Bull. Seism. Soc. Am*, **80**, 1777–1800.

Nataf, H-C., Nakanishi, I., & Anderson, D. L., 1986. Measurements of mantle-wave velocities and inversion for lateral heterogeneities and anisotropy, 3, Inversion, *J. Geophys. Res.*, **91**, 7261-7307.

Nicolas, A., Bouchez, J. L., Boudier, F. & Mercier, J. C., 1971. Textures, structures and fabrics due to solid state flow in some European lherzolites, *Tectonophysics*, **12**, 55-86.

Nicolas, A., Boudier, F. & Boullier, A. M., 1973. Mechanisms of flow in naturally and experimentally deformed peridotites, *Am. J. Sci.*, **273**, 853-876.

Nicolas, A., & Poirier, J. P., 1976. Crystalline Plasticity and Solid-State Flow in Metamorphic Rocks, Whiley, London, 437 pp.

Nicolas, A., & Christensen, N. I., 1987. Formation of anisotropy in upper-mantle peridotite - a review, *Rev. Geophys.*, **25**, 111-123.

Pavlidis, T., 1971. Linguistic analysis of waveforms, *Software Engineering Vol. 2*, (J. Tou, ed), Academic Press, New York.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling W. T., 1988. Numerical Recipes, Cambridge University Press.

Raitt, R. W., Shor, G. G., Francis, T. J. G., & Morris, G. B., 1969. Anisotropy of the Pacific upper mantle, *J. Geophys. Res.*, **74**, 3095-3109.

Regan, J, & Anderson, D. L., 1984. Anisotropic models of the upper mantle, *Phys. Earth Planet. Inter.*, **35**, 227-263.

Shearer, P. M., & Orcutt, J., 1986. Compressional and shear-wave anisotropy in the oceanic lithosphere, *Geophys. J. Royal Astron. Soc.*, **87**, 967-1003.

Silver, P. G., & Chan, W. W., 1991. Shear-wave splitting and subcontinental mantle deformation, *J. Geophys. Res.*, **96**, 16429-16454.

Stockman, G., Kanal, L. & Kyle, M. C., 1976. Structural pattern recognition of carotid pulse waves using a general waveform parsing system, *Comm. Assoc. Comp. Mach.*, **19**, 688-695.

Tanimoto, T, & Anderson, D. L., 1984. Mapping convection in the mantle, *Geophys. Res. Lett.*, **11**, 287-290.

Thomson, C. J., Kendall, J-M., & Guest, W. S., 1992. Geometrical theory of shear-wave splitting: corrections top ray theory for interference in isotropic/anisotropic transitions, *Geophys. J. Int.*, **108**, 339-363.

Tong, C., Gudmundsson, O. & Kennett, B. L. N., 1994. Shear wave splitting in refracted waves returned from the upper mantle transition zone beneath northern Australia, *J. Geophys. Res*, **99**, 15,783-15,797.

Tong, C., 1995, Characterization of seismic phases - an automatic analyser for seismograms, *Geophys. J. Int.*, **123**, 937-947.

Tong, C., & Kennett, B. L. N., 1995. Towards the identification of later seismic phases, *Geophys. J. Int*, **123**, 948-958.

Tong, C., & Kennett, B. L. N., 1996. Automatic seismic event recognition and later phase identification for broad-band seismograms, *Bull. Seism. Soc. Am.*, submitted in November 1995.

Vinnik, L. P., Kosarev, G. L., & Makeyeva, L. I., 1984. Anisotropy in the lithosphere from the observations of SKS and SKKS, *Dokl. Acad. Nauk SSSR*, **278**, 1335-1339.

Vinnik, L. P., Makeyeva, L. I., Milev, A., & Usenko, A. Y., 1992. Global patterns of azimuthal anisotropy and deformations in the continental mantle, *Geophys. J. Int.*, **111**, 433-447.

Yu, G. K., & Mitchell, B. J., 1979. Regionalized shear-velocity models of the Pacific upper mantle from observed Love- and Rayleigh-wave dispersion, *Geophys. J. R. Astron. Soc.*, **57**, 311-341.