

AD-A286 889



IMPROVING SEISMIC EVENT CHARACTERISATION

B.L.N. Kennett, O. Gudmundsson & C. Tong

The Australian National University
Research School of Earth Sciences
Canberra ACT 0200
AUSTRALIA

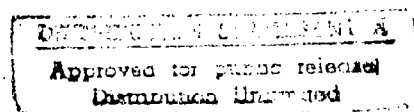
AFC512TR 916

C405



1996 July 22

Final Technical Report
1995 March 1 - 1996 June 30



Grant: F49620-94-1-0110

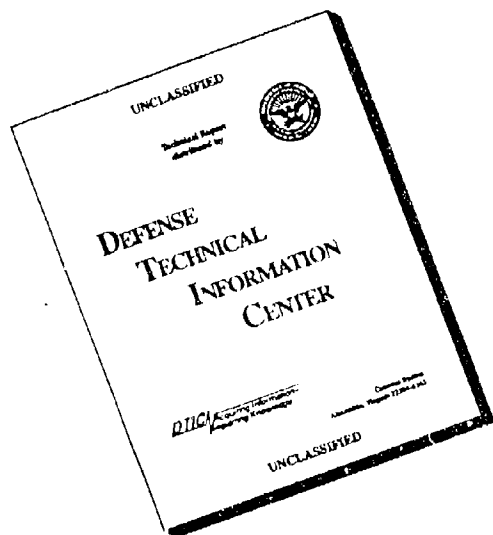
A 1

96-01686



8 5

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE COPY
FURNISHED TO DTIC CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 1996 July 22		3. REPORT TYPE AND DATES COVERED Final Technical 01 Mar 95-30 July 96	
4. TITLE AND SUBTITLE Improving seismic event characterisation				5. FUNDING NUMBERS G:F49620-94-1-0110	
6. AUTHOR(S) B.L.N. Kennett, O. Gudmundsson and C. Tong					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research School of Earth Sciences The Australian National University Canberra ACT 0200 Australia				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NM 110 Duncan Avenue Suite B115 Holling AFB DC 20332-001				10. SPONSORING MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES					
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This work is directed towards improved identification and characterisation of the later phases in the seismogram so that they can be exploited in location and source mechanism studies. Section 1 is directed towards the use of seismic travel times. Corrections are produced for the effect of ellipticity on a wide range of seismic phases. A regionalised tomographic procedure is introduced for determining upper mantle structure with the object of improving systematic corrections for 3D structure in location techniques. The influence of stochastic heterogeneity on intermediate scales on seismic waveforms is investigated by numerical modelling. Section 2 is concerned with rapid automatic classification of seismic arrivals with the object of even recognition and phase identification. Using techniques derived from Artificial Intelligence a classification scheme allows the testing of hypotheses about the location of an event and the character of the phases. The output from a single seismogram is an epicentral distance, approximate depth and azimuth as well as a phase association. This information is useful for preliminary event location and is enhanced with use by a network.					
14. SUBJECT TERMS Seismic models, travel times phase identification				15. NUMBER OF PAGES	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT 500		

Global Tomography with Irregular Parameterisation
(Towards regionalised travel-time corrections
for body waves)

Olafur Gudmundsson
and
Malcolm Sambridge

Introduction

Global models of velocity heterogeneity have emerged in the past decade which show significant deviations from spherical symmetry in Earth's structure. As the quality of data and their coverage increases more detail is included in the models. Global models currently claim to resolve structures as small as 1000 km in lateral extent, but regional models resolve structures to considerably smaller scales, i.e. a few hundred kilometres. The strength of heterogeneity in the models, i.e. the peak-to-peak variation of velocity, increases as more detail is included. The strength of heterogeneity in the P-wave velocity field reaches $\pm 5\%$ at the top of the upper mantle at the scales which are resolved regionally. It is clear that such strong heterogeneity will significantly affect the travel times of the body-wave phases most commonly used in earthquake location. Accounting for this effect can therefore significantly improve the quality of event location which is of obvious importance for nuclear monitoring.

It is clear from models derived from body waves and surface waves that the level of velocity heterogeneity in the Earth varies dramatically with depth and is strongly concentrated in the top 300-400 km (e.g. Gudmundsson et al. [1990]). Most of the effect of structure on travel times is therefore due to structures at relatively shallow depth through which teleseismic body waves have a relatively simple ray geometry (steep). The image of upper mantle heterogeneity that has emerged is one that clearly correlates with the overall tectonic processes on the surface. Young oceanic regions have a low velocity as do regions beneath tectonically active continents, while old oceanic regions and old stable continents are underlain by a mantle of high velocity.

Accounting for the effects of velocity heterogeneity on travel times and thus event location must include the main features of the velocity structure and remain a simple procedure in order to be implementable. We attempt to meet these prerequisites with the following strategy:

We use the high level of correlation of global velocity models with surface tectonics to define the main features of Earth structure. A detailed regionalisation defines irregularly shaped bodies on the Earth's surface which in turn define discrete model parameters. With irregular sampling of the regionalisation we can represent irregular bodies and variable scales. The sampling is dense where needed (continental margins, island arcs and tectonic continent), and sparse where that suffices (intraregional, oceanic).

We use Delaunay tessellation and Voronoi cells around each sample of the regionalization to define discrete regions belonging to the various tectonic types. This allows us to collapse the number of parameters needed to represent the complex regions significantly. Efficient algorithms are available for navigation and book keeping in an irregular grid (Delaunay tetrahedra or Voronoi cells) (Sambridge et al., 1995).

We use a select data set of travel times from well located events and nuclear explosions in order to alleviate the mapping of mislocation into structure. We invert these data for an irregularly parameterised, regionalised upper-mantle model (RUM). Our objective is not to find the best possible model, but to find the simplest model that still contains enough information about structure to be useful for generating corrections to travel times used in event location.

Teleseismic travel times can then be corrected for this model by a relatively simple look-up table based on a tabulation of the tectonic regions in which stations and events fall.

Method

Figure 1 shows our regionalisation. It is primarily based on maps published by Sclater et al. (1980) and Jordan's (1981) GRT1 regionalization, but also incorporates information from e.g. Miyashiro et al. (1982). Eight tectonic regions are defined. The ocean is divided into three age provinces, the same as the GRT1 regionalization of Jordan [1981]. This regionalization has a resolution of two degrees compared to five degrees in GRT1. The continents are divided into 5 regions according to age, four regions of low volcanic and earthquake activity, in addition to a region termed tectonic continent where activity is high. We have termed the regions Young Continent, Intermediate-age Continent, Old Continent and Ancient Continent. The chosen age boundaries are listed in Figure 1 and were selected by comparing the age zonation of the continents with tomographic images of the upper mantle (e.g. Woodhouse and Dziewonski, 1984). We have perhaps unconventionally termed the Iceland-Faeroe-Greenland ridge continental because of its highly anomalous crustal thickness for oceanic crust (25-35 km according to Bott and Gunnarson, 1980).

The regionalization is represented by 16200 parameters. Roughly a quarter of those are selected to parameterise the boundaries between the regions (4100 parameters). The boundaries are all parameterised with a two-degree resolution except the boundaries between oceanic regions. Thus, the density of parameterisation is high where we expect potential sharp contrasts, i.e. continental margins and boundaries between tectonic and atectonic regions, and low where we don't, i.e. within the oceans. Figure 2a shows the Voronoi cells around the selected samples of the regionalization. The figure contains about 4100 Voronoi cells. We reproduce the full complexity of the regionalization in Figure 1 with only about a quarter of the number of parameters in the original (see Figure 2b). Each Voronoi cell is taken to encompass a region of uniform velocity at each depth. The regionalization is extended from the surface to a

depth of 660 km to span the whole of the upper mantle.

Subducting slabs are significant structures in the upper mantle which are clearly oblique and cannot be included in a two-dimensional regionalization as above. Most of the earth's seismicity occurs in or around subducting slabs. Slabs are thus likely to strongly affect travel times in general. We have included oblique slab structures in our parameterization of the upper mantle as the ninth region. First we contoured slab-related seismicity world wide. Examples of slab contours are presented in Figure 3a for the slabs of the northwest Pacific and the Tonga/Kermadec slab. The contours are drawn at fifty km depth intervals and such that they lie near the top of the seismogenic region. We used the relocated NEIC catalog of Engdahl, van der Hilst and Buland to define the seismicity. The catalog is not complete and contains significant gaps which are difficult to interpret. Such gaps are not interpolated if they exceed a few hundred km in size. The slab contours obviously do not define a slab where there is no seismicity. It is difficult to envisage the thermal anomaly of a subducting slab to terminate as abruptly as the seismicity in many places. Thus our model of slabs is likely to be incomplete. However, slabs are defined where we most need them, i.e. where the earthquakes occur for which travel-time corrections are needed.

The slab contours define a surface when interpolated which we take to define the top of the slab (or what used to be the top of the oceanic lithosphere before subduction). We then define a complementary surface assuming the slab is 200 km thick. This is a simplification of slab geometry, particularly at depth. Projecting the original surface 200 km along its local normal results in the folding of the original surface. Such folds were simply edited out interactively with a graphics tool. We sample the contours discretely at a spacing of about 10 km. Our samples are not internal to the slabs as we define them, but on their surface. Thus constant velocity Voronoi cells are not suitable to describe their complex geometry. Instead we place the slabs in a global coordinate system and construct the three-dimensional Delaunay tessellation for the 23,000 slab samples for all 24 individual slab structures. Those Delaunay tetrahedra which join the phases of the same slab structure then define the volume of slabs in the upper mantle. Two examples are shown in Figures 3b and 3c. Those are two of the larger and deeper slabs which display significant complexity in their shape.

In summary the parameterization of the upper mantle consists of two distinct elements: Delaunay tetrahedra to define the slabs and a two-dimensional grid of Voronoi cells to define tectonic regions. The slabs are overlaid on the regionalization. To determine which type of environment a given point (along a ray) is in we first evaluate if it is within a slab. If it is not, its position in the regionalization is determined. These tools allow us to represent complex structures efficiently and economic algorithms are available to locate any arbitrary point in the irregular grid.

We use a select data set of travel times from well located events and nuclear explosions in order to alleviate the mapping of mislocation into structure. The events are the 105 events of the test-data set used by Kennett and Engdahl (1991) supplemented with 92 large earthquakes to get a more even sampling of the globe and the 9 tectonic regions. The criterion for the selection of these events was that they be

recorded by a large number of globally distributed stations. Their locations are taken from the relocated NEIC catalog of Engdahl, van der Hilst and Buland (1996) where they have reassociated ISC data, included pP and pwP phases in the location, used IASP91 (Kennett and Engdahl, 1991) and subjected locations to stringent fitness criteria. We use only teleseismic P and PKP arrivals (first arrivals) and picks reported to the ISC as impulsive. We invert these data for an irregularly parameterised, regionalised Earth model.

Figure 4 shows a map of the 197 events used together with the regionalization. The events are grouped into three categories. First there are 21 nuclear and chemical explosions displayed as white stars in the figure. Then there are 39 earthquakes which are well located with local observations in a local velocity model. They are displayed as white dots. The remaining 137 events are located using teleseismic observations. The locations for the 21 explosions and 39 well located events are taken fixed. For the remaining 137 events we use locations from Engdahl et al. (1996).

We reference travel-time residuals to AK135 (Kennett et al., 1995) and reassociate all impulsive picks reported to the ISC for the above events. We find a surprising number of events with a significant mean residual. This is surprising because AK135 is almost identical in the mantle to IASP91 which Engdahl et al. (1996) used for their locations. The explanation must lie in a difference in data selection. We for example include PKIKP and not pP. We do not associate pwP. The large number of events with a significant apparent origin-time shift led us to allow origin time to be a free parameter in our inversion for regionalized structure.

Results and conclusions

We solve a linearized inverse problem. A higher order approach is not warranted because ray bending critically depends on velocity gradients which are artificially represented by our parameterization. We sample depth ten times at 66 km spacing and seek 9 independent velocity profiles, one for each of the eight tectonic regions as well as slabs. We use a standard damped-least-squares inversion procedure. In regions with limited depth distribution of sources teleseismic rays give limited depth resolution in the upper mantle. We cannot well resolve ten independent parameters except in subducting slabs, where there is a distribution of events in depth. Since the level of heterogeneity in the upper mantle is strongly concentrated near the surface (e.g. Gudmundsson et al. 1990), in the top 300-400 km, we have projected the depth profiles on to a set of 4 Gaussian basis functions which are not allowed to put structure in the lower half of the upper mantle. This is with the exception of slabs which are allowed to have structure to the base of the upper mantle. The result we present are moderately damped. An undamped inversion allowing for 10 independent depth parameters in each region achieves a 19% variance reduction. Limiting the degrees of freedom to four per region via the Gaussian basis reduces the variance reduction to 16%. We have damped the solution to yield a variance reduction of 14%.

The results of inversion are shown in Figure 5. Each profile represents a given tectonic region. The curves have error bars equivalent to one standard deviation assuming that the error variance of the data is about 1 second (Gudmundsson et al. 1990). The results are encouraging. Slabs appear as a continuous fast anomaly throughout the upper mantle except at the top where we have tectonic crust rather than a true slab and where the result is probably affected by the slow mantle wedge. The amplitude of the slab structure is 2-3% velocity perturbation. This is assumed to be uniform over a thickness of 200 km and would thus be consistent with a peak perturbation in the core of the slab of 4-5%. Other features of the results are also encouraging. The oceans are slow with old oceans developing a fast lithosphere. Tectonic continent and young continent are also slow and essentially indistinguishable while intermediate-age continents have developed a fast lithosphere. There is no resolvable structure underneath old continents suggesting that AK135 (and IASP91) are representative of continent in that age range. Ancient continents are fast, but not to a great depth, only about 200 km. It should be noted that we have not corrected for crustal structure so the model should be interpreted to include crustal signature. We refer to this model as the RUM model (Regional Upper Mantle).

Our procedure of data reduction concludes with the calculation of station corrections (where sufficient numbers of picks are reported) computed after correction for the above model. We can then divide the reduction process into three steps: relocation (origin time), regionalized tomography, and subsequent station corrections. Table 1 summarizes the variance and mean reduction achieved by these three steps in the analysis. Relocation accounts for much of the variance (more than 50%). The structural model achieves 14% variance reduction. It is worth noting that we work with individual observations and have not subjected the data to any averaging procedures prior to inversion (such as summary rays). Finally, station corrections accomplish a 16% variance reduction. This indicates that our model leaves significant structural signal in the data. Histograms of residuals before and after analysis are presented in figure 6.

Figure 7 presents the results for the station corrections. Clearly a significant signal is left in the data after inversion for regionalized structure. Note for example that a fast anomaly remains within the Laurentian craton. On the other hand the African cratons are left with slow anomalies. Some tectonic regions are left with systematically fast residuals while others are slow. Clearly all regions of the same type are not the same. Our parameterization in effect imposes complex spatial constraints on this inversion for upper-mantle structure which we can think of as a priori information, i.e. the presumption that the upper mantle can be regionalized as we have done. The success of the inversion justifies those constraints. The remaining station residuals point to the fact that this regionalized description is not complete, however, and that on order of half the structural signal in the data is explained by it.

We have described upper-mantle structure with 23 degrees of freedom and manage to explain about half of the apparent structural variance in travel time data. The data are individual, not summary ray data. Preliminary results from S-wave

residuals give results similar to the ones presented here for P-wave velocity variations.

References cited:

Bott, M.H.P. and Gunnarsson, K., [1980]. Crustal structure of the Iceland-Faeroe Ridge, *J. Geophys.*, **47**, 221-227.

Engdahl, van der Hilst, and Buland, [in preparation?]

Gudmundsson, Davies, and Clayton, [1990]. Stochastic analysis of global travel time data: Mantle heterogeneity and errors in the ISC data, *Geoph. J. Int.*, **102**, pp 25-43.

Jordan, [1981]. Global tectonic regionalisation for seismological data analysis, *Bull. Seism. Soc. Am.*, **71**, pp 1131-1141.

Kennett and Engdahl, [1991]. Travel times for global earthquake location and phase association, *Geophys. J. Int.*, **105**, pp 429-466.

Kennett, B.L.N.K., Engdahl, E.R., and Buland, R., [1995]. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, **122**, pp 108-124.

Miyashiro, A., Aki, K., and Sengor, A.M.C., [1982]. *Orogeny* Wiley, Chichester, England, 242p.

Sambridge, Braun, and McQueen, [1995]. Geophysical parameterisation and interpolation off irregular data using natural neighbours, *Geophys. J. Int.*, **122**, pp 837-857.

Sclater, Jaupart, and Galson, [1981]. The heat flow through oceanic and continental crust and the heat loss of the Earth, *J. Geophys. Res.*, **86**, pp 269-311.

Woodhouse, J.H. and Dziewonski, A.M., [1984]. Mapping the upper mantle: Three-dimensional modeling of Earth's structure by inversion of seismic waveforms, *J. Geophys. Res.*, **89**, pp 5953-5986. Wiley, Chichester, England, 242p.

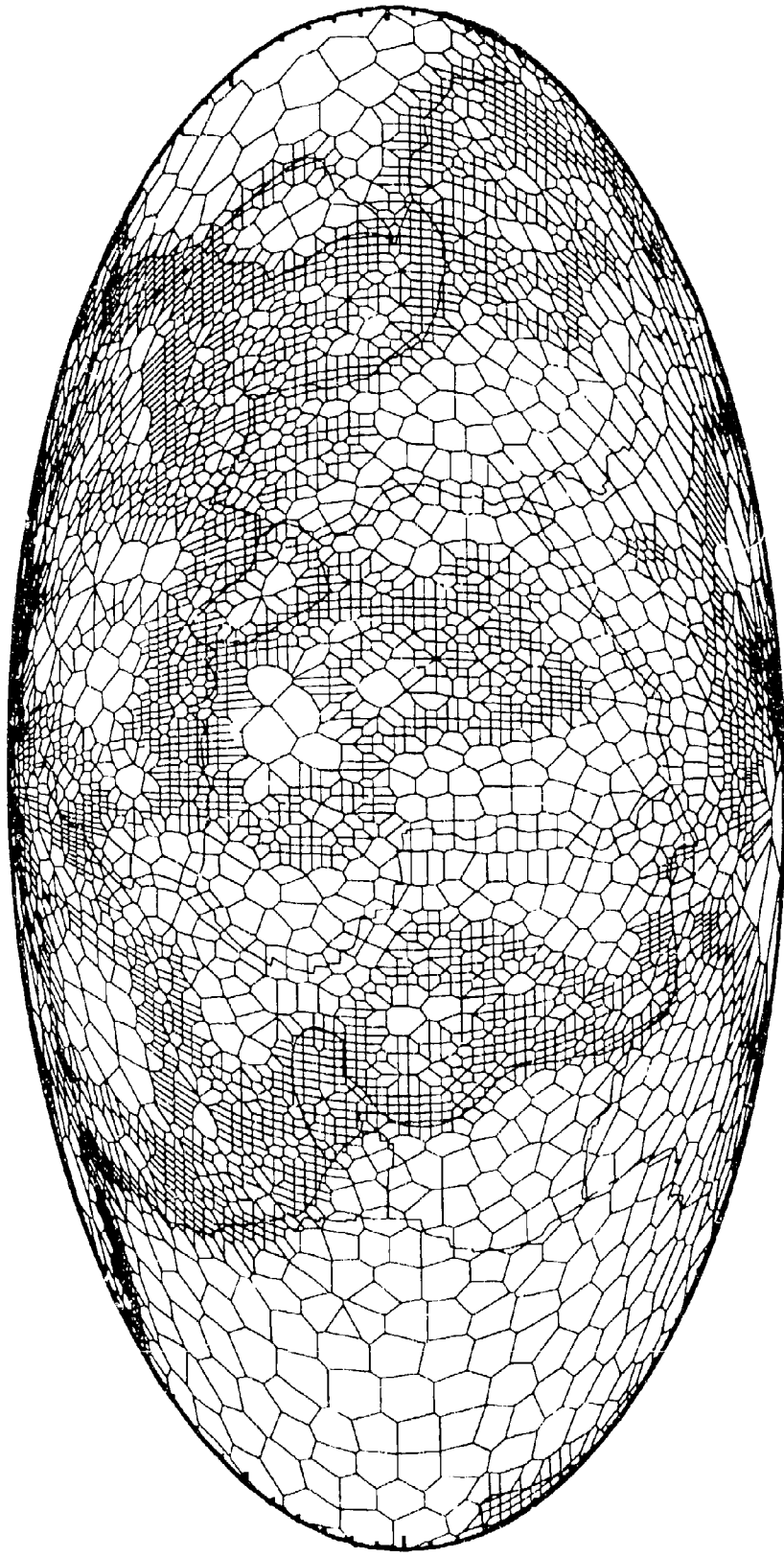
Tectonic Regionalization



- O1: Young ocean (< 25 Myr)
- O2: Intermediate ocean (25-100 Myr)
- O3: Old ocean (> 100 Myr)
- C1: Tectonic continent
- C2: Young continent (< 150 Myr)
- C3: Intermediate continent (150-800 Myr)
- C4: Old continent (800-1700 Myr)
- C5: Ancient continent (> 1700 Myr)

fig 1

Voronoi diagram



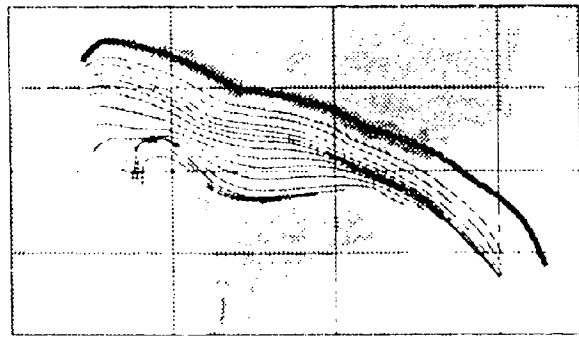
The Voronoi diagram produced by 4274 nodes on the surface of the Earth.
Here the individual cell boundaries can be seen and many cells are joined
together to form the 126 tectonic regions shown above.

fig 2a

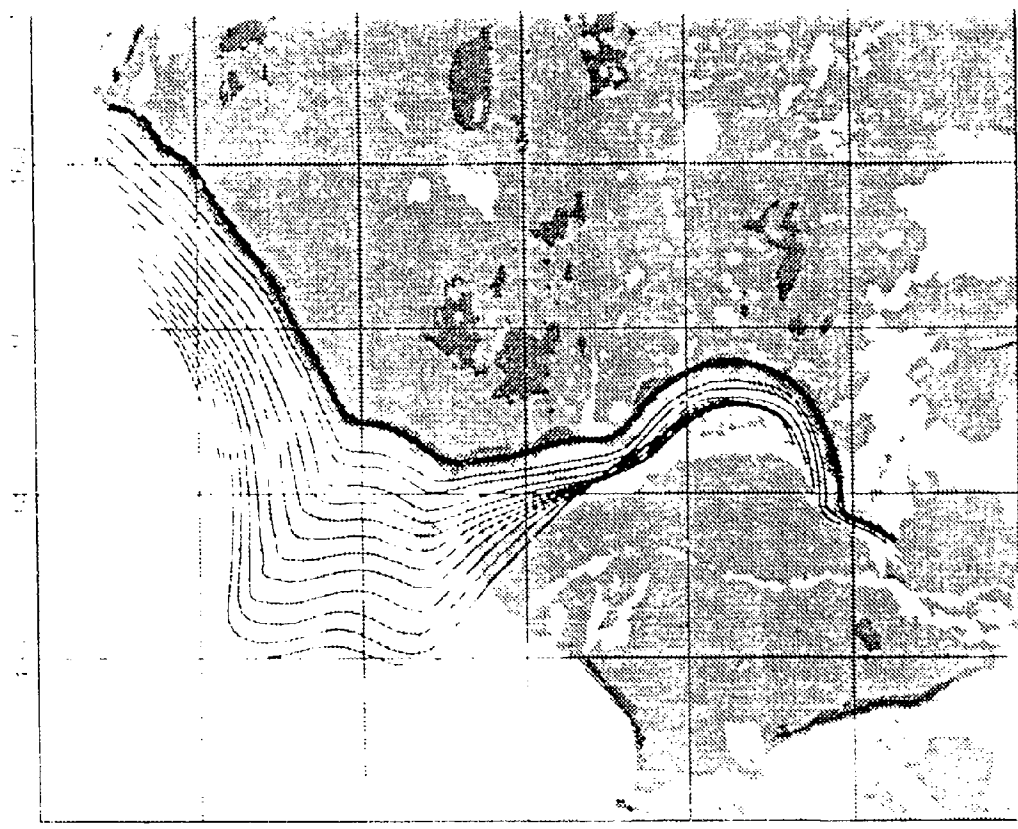
Voronoi representation



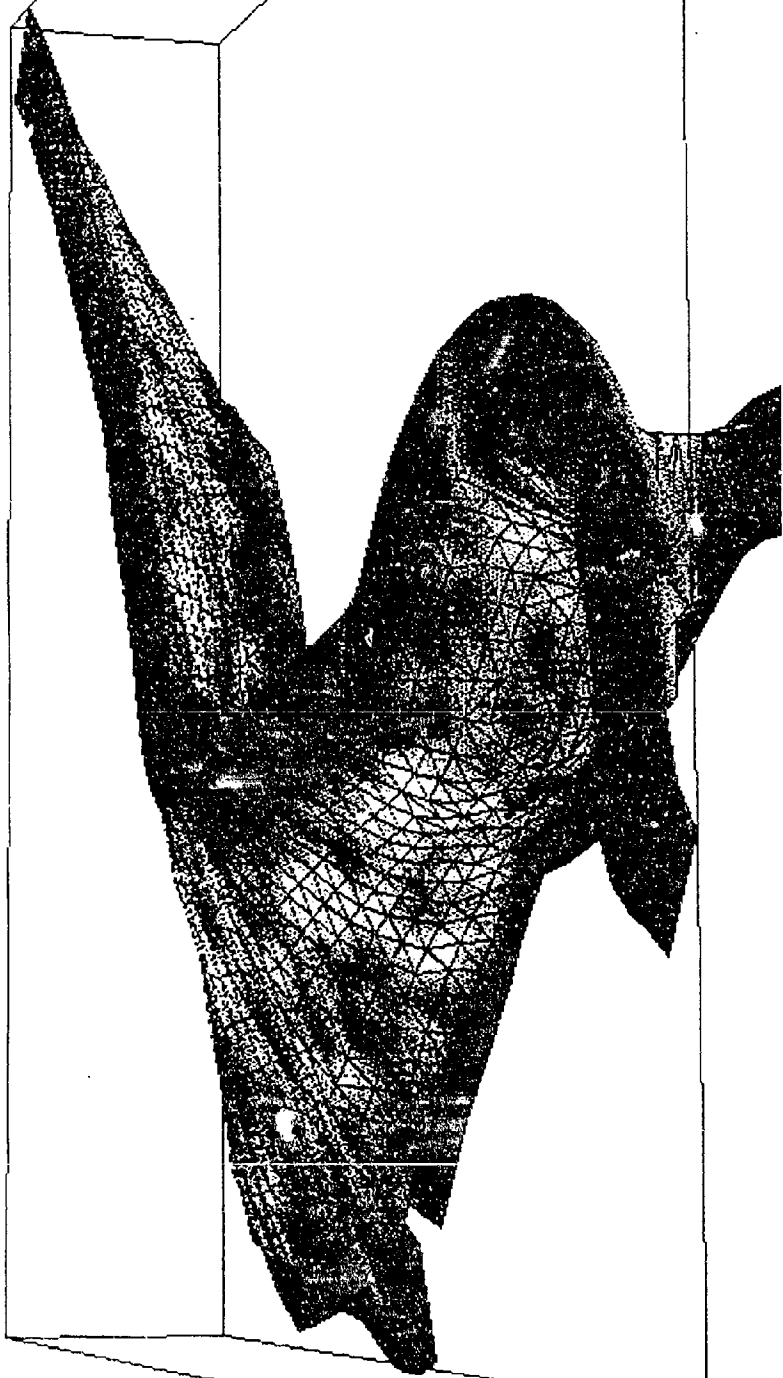
Each tectonic region is formed by grouping together Voronoi cells of the same type. The cells are coloured according to tectonic type. All regions with the same colour contribute to same regional velocity model in the inversion.



Bathymetry / topography of the northwest Pacific and the Tonga-Kermadec region with slab contours every 50 km in depth. The contours represent the top of the sinking lithosphere.

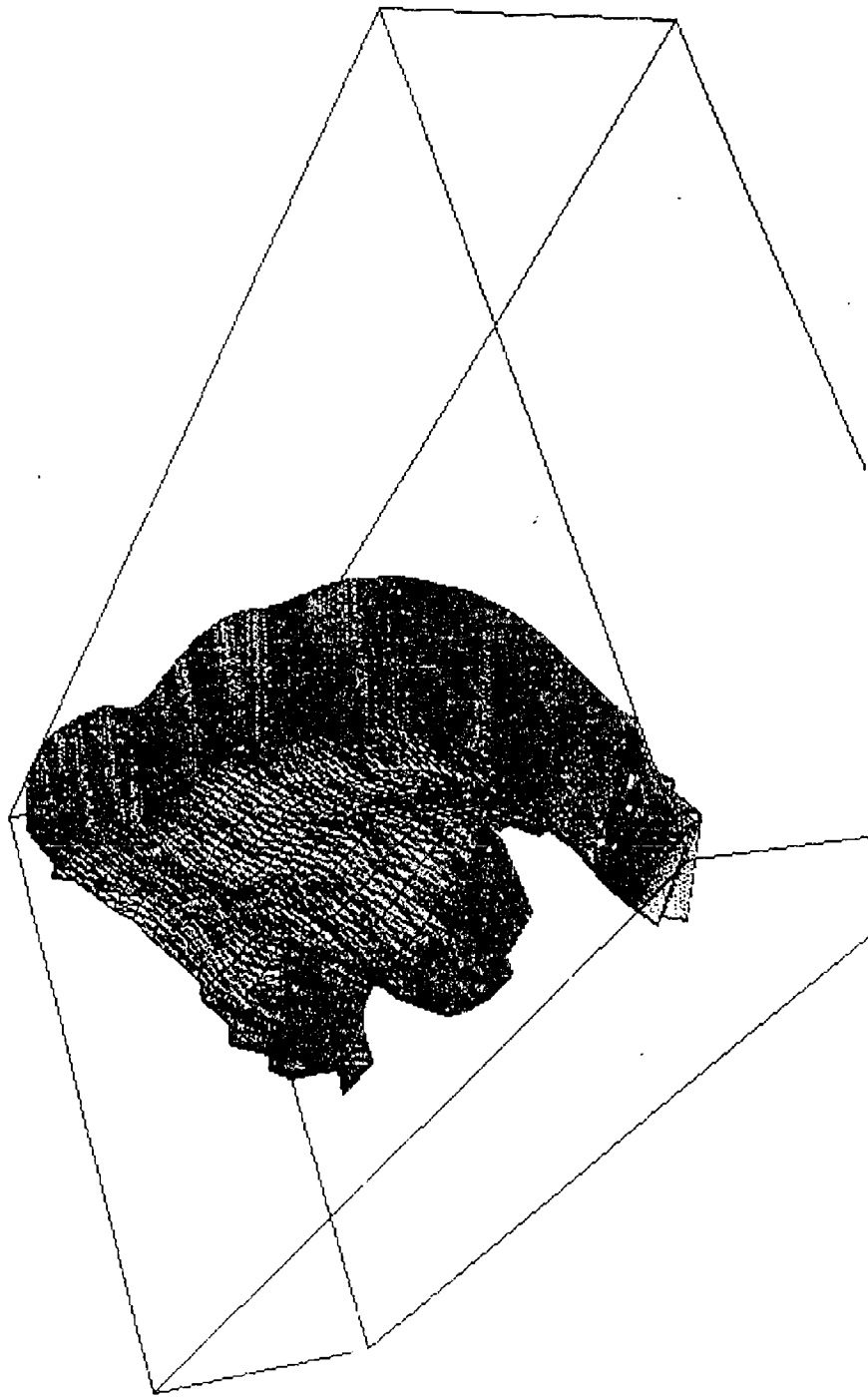


fig



Perspective view of the Marianas/Japan/Kuriles subduction-zone model derived from contours of slab seismicity. The 3-D structure consists of 1420 nodes and 3764 Delaunay tetrahedra. In the inversion all rays intersecting these tetrahedra contribute to the global slab model.

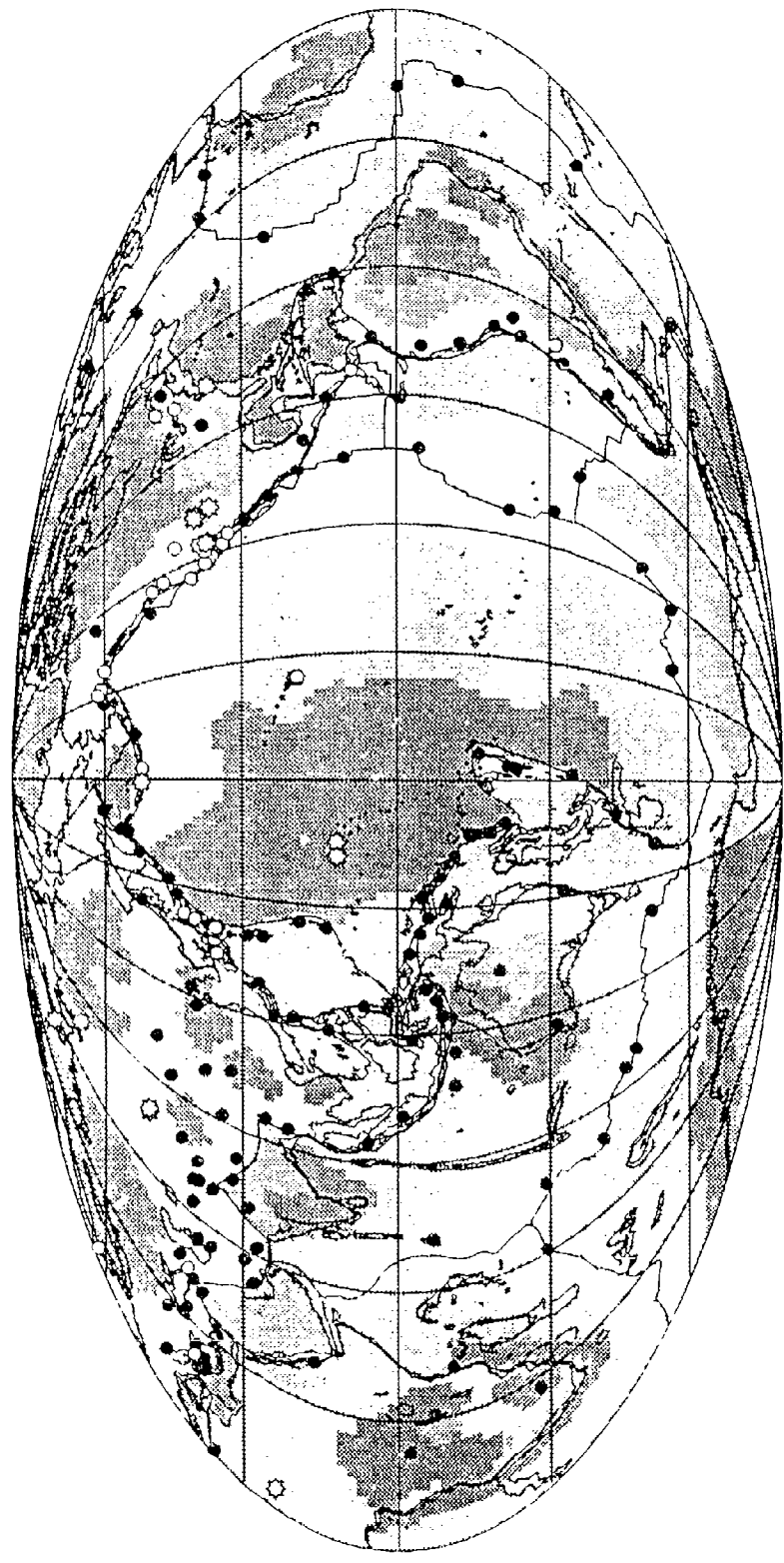
fig



Perspective view of the Tongan subduction-zone model derived from contours of slab seismicity. The 3-D structure consists of 1453 nodes and 3679 Delaunay tetrahedra. In the inversion all rays intersecting these tetrahedra contribute to the global slab model.

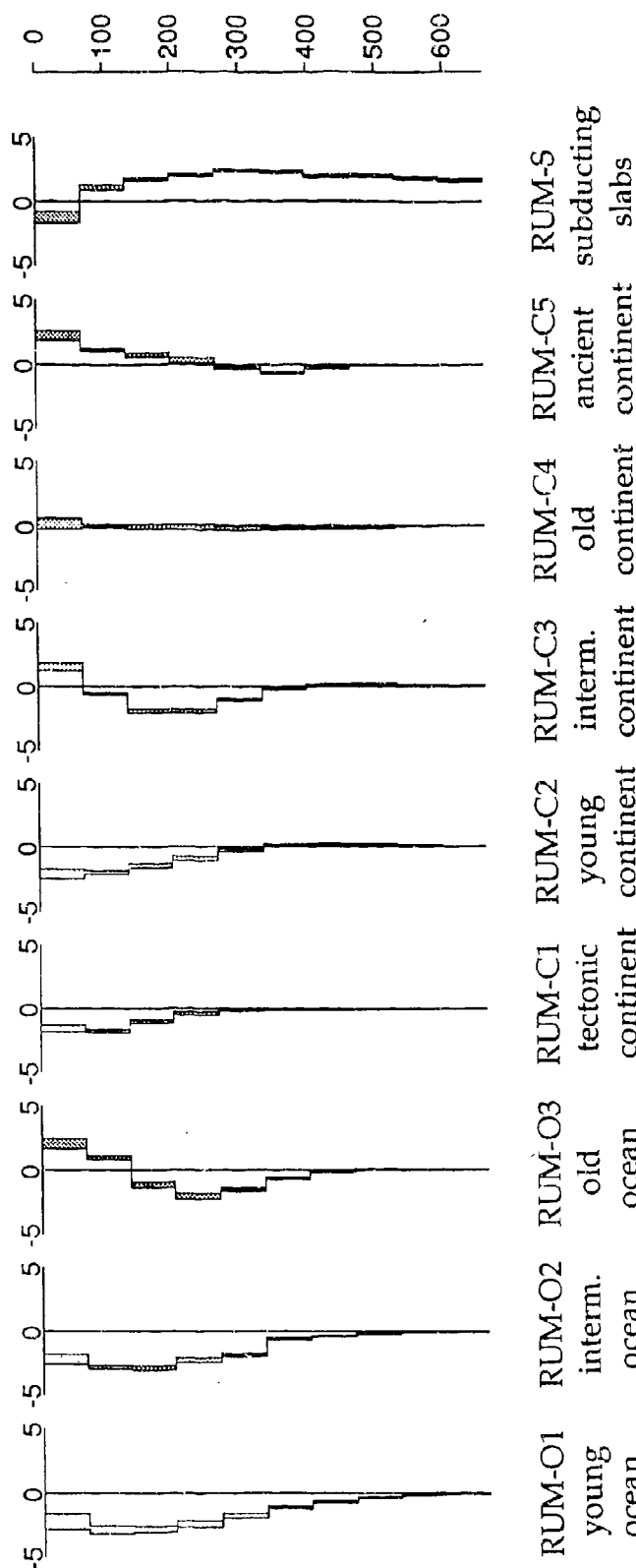
fig 2

Regionalization and events



Epicentres of the events used. Stars are nuclear or chemical bombs. White dots are earthquakes with tight constraints on location from local observations. Black dots are earthquakes with no local constraints, but which have been located with a large number of global observations subject to fitness constraints.

velocity perturbation [percent]

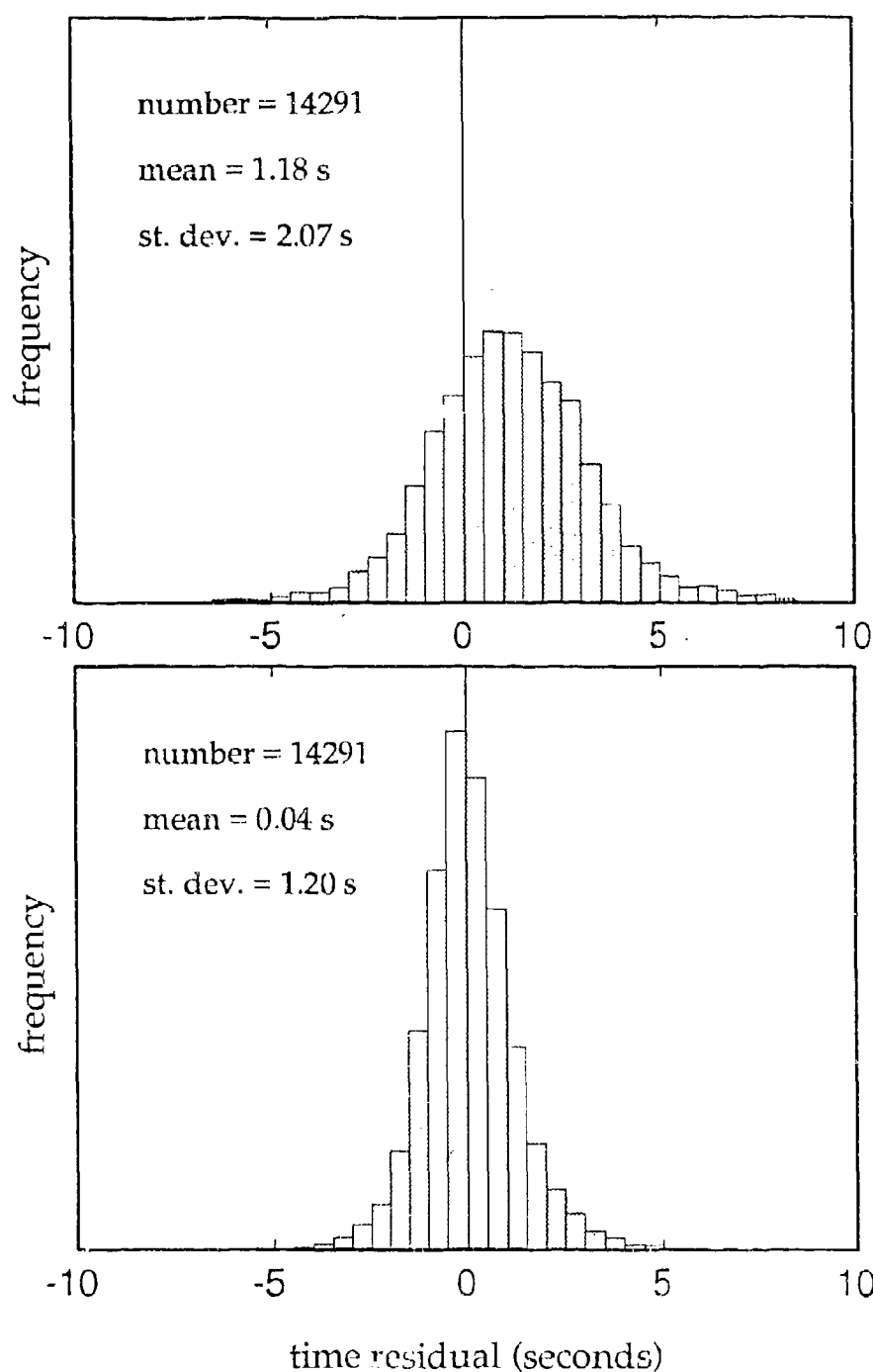


One-dimensional, velocity-perturbation profiles for the nine tectonic regions. The error bars are one standard deviation assuming that the standard deviation of data errors is 1 second. Colours are the same as in maps. This structural model explains about 15 percent of the travel-time data.

Variance and mean reduction

	mean [s]	variance [s ²]	variance reduction
Initial	1.18	4.30	
After relocation	0.69	1.99	54%
After velocity perturbation	0.19	1.71	14%
After station correction	0.04	1.44	16%

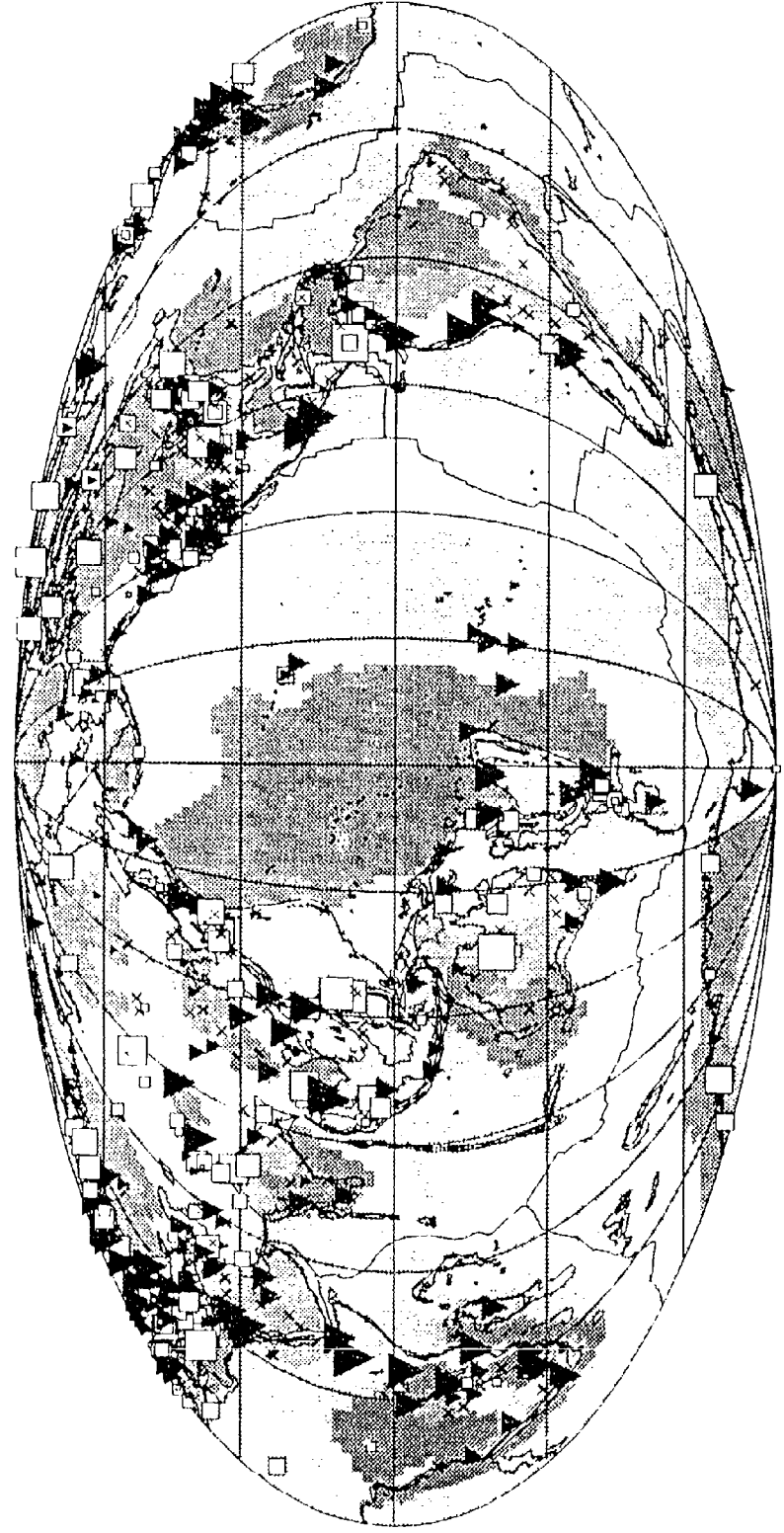
Statistics of travel-time residuals at progressive stages in the analysis. Much of the variance of the data lies in earthquake mislocation. A significant level of variance reduction is achieved by the modeled velocity perturbation although no averaging [e.g. summary rays] has been applied [we use individual residuals]. The variance reduction achieved by station corrections is probably due to structure which is either on scales smaller than allowed for, or which reflects inconsistencies between regions categorized as of the same tectonic type [see map of station corrections].



Distributions of impulsive, first-arrival, time-residuals before (above) and after (below) inversion for origin-time shift, regionalized structure and station corrections.

fig

Regionalization and stations



The stations providing impulsive picks used in study. Symbols show magnitude and sign of station corrections computed after inversion.

- Negative correction (fast) 1 second.
- ▼ Positive correction (slow) 1 second.
- Events with too few picks to define a robust correction.

Ellipticity Corrections for Seismic Phases

B.L.N. Kennett & O. Gudmundsson,

Research School of Earth Sciences The Australian National University Canberra ACT 0200, Australia

SUMMARY

The advent of broad-band seismology has meant that use is being made of a wide range of seismic phases for many of which ellipticity corrections have not been readily available. In particular, when many seismic phases are used in location schemes, it is important that the systematic effects of ellipticity are included for each phase.

An efficient and effective procedure for constructing ellipticity corrections is to make use of the ray-based approach of Dziewonksi & Gilbert (1975), as reformulated by Doornbos (1988), in conjunction with the rapid evaluation of travel times and slowness for a given range using the tau-spline procedure of Buland & Chapman (1983).

Ellipticity coefficients have been tabulated for a wide range of seismic phases and are available in electronic form. The ellipticity correction procedures have been extended to include an allowance for diffraction phenomena e.g. P_{diff} , S_{diff} diffracted along the core-mantle boundary. Corrections for additional phases can be generated by building the ellipticity coefficients from suitable combinations of the coefficients for different phase segments.

Key words: Ellipticity corrections, travel times

Running Title: Ellipticity corrections for seismic phases

1 INTRODUCTION

As methods for the delineation of the three-dimensional structure from seismic arrivals have been developed, a wider range of seismic phases have begun to be used. It is important therefore that the influence of the ellipticity of the figure of the Earth on seismic travel times be taken into account in a comparable way for each seismic phase. The same need arises in the exploitation of later seismic phases in rapid earthquake location using a sparse global array; the systematic effects of ellipticity must be included for all phases used in the location procedure.

The normal model for the calculations of travel times for seismic phases is to use a spherical earth with a radial distribution of velocity; aspherical effects are then included via a perturbation treatment. An important component of such asphericity arises from the ellipticity of the figure of the Earth. The resulting influence on the travel time of a seismic phase will then depend on the location of the source and receiver points, rather than just the epicentral distance Δ between them. For a source at co-latitude ϑ_0 , and a path with azimuth ζ to the receiver (measured clockwise from North), Bullen (1937a) introduced triple entry correction tables $\delta t = \delta t(\vartheta_0, \Delta, \zeta)$ to evaluate the ellipticity correction for P waves; these were later extended to polar latitudes (cf Engdahl & Gunst, 1966).

Dziewonski & Gilbert (1976) have developed an elegant representation of the influence of ellipticity on travel times and have demonstrated the significance of source depth, z_s , in determining the appropriate corrections. For a seismic phase the correction can be built up from three spherical harmonic components of order 2,

$$\delta t = \sum_{m=0}^2 \sigma_m(z_s, \Delta) P_{2,m}(\cos \vartheta_0) \cos m\zeta. \quad (1.1)$$

The coefficients $\sigma_m(z_s, \Delta)$ represent ray-path integrals in terms of the radial velocity distribution $v(r)$ as an integral over epicentral distance. Introducing $\lambda_m(\theta)$ as a scaled associated Legendre function,

$$\lambda_m(\theta) = \frac{2}{3} P_{2,m}(\cos \theta), \quad (1.2)$$

the coefficients $\sigma_m(z_s, \Delta)$ take the form

$$\begin{aligned} \sigma_m(z_s, \Delta) = & \frac{1}{p} \int_0^\Delta d\theta \eta^3(r) \frac{dv}{dr} c(r) \lambda_m(\theta) \\ & - \sum_j \left\{ c \lambda_m(\theta) \left[(\eta^2 - p^2)^{1/2} \right] \right\}_j + \sum_k \left\{ c \lambda_m(\theta) \left[(\eta^2 - p^2)^{1/2} \right] \right\}_k \end{aligned} \quad (1.3)$$

where

$$\eta(r) = v(r)/r, \quad (1.4)$$

p is the ray slowness, $c(r)$ is Earth's ellipticity of figure, and the effect of source depth is implicit in the dependence of r on θ along the ray. The summations represent the influence of discontinuities in velocity (v) in transmission (j) and reflection (k). Doornbos (1988) has shown how the integrals for $\sigma_m(z_s, \Delta)$ can alternatively be formulated in terms of integrals over radius without sacrificing the convenience of the representation

(1.1). Note that the ellipticity coefficients were designated by τ_m by Dziewonski & Gilbert (1975), we have changed to the notation σ_m to avoid conflict with the well established usage $\tau(p)$ for the delay time which is also used in section 2.1.

Dziewonski & Gilbert (1975) presented a tabulation of ellipticity corrections using the $\sigma_m(z_s, \Delta)$ formulation for the phases *P*, *PcP*, *PKP*_(ab,bc,df), *S*, *ScS*, *SKS* using the *pem* model of Dziewonski, Hales & Lapwood (1975). Corrections were given for depths of 0, 300 and 650 km. The perturbations in travel time are fortunately relatively insensitive to the particular velocity model which is employed. Kennett (1991) extended the $\sigma_m(z_s, \Delta)$ tabulation using the *iasp91* model of Kennett & Engdahl (1991) to include a finer tabulation in depth ($z_s = 0, 35, 50, 10, 200, 300, 500$ and 700 km) and additional phases *SKS*_(ac,df), *ScP* and *SKP*.

In this paper we demonstrate the convenience of combining the Buland & Chapman (1983) tau-spline procedure for calculating travel times, as used by Kennett & Engdahl (1991), with the Doornbos (1988) calculation of the σ_m ellipticity coefficients using ray integrals. An extensive range of ellipticity coefficients have been tabulated and are available in electronic form. The correction procedures have been extended to include an allowance for diffraction phenomena, as e.g. at the core-mantle boundary in *P*_{diff}, *S*_{diff}. Further, we demonstrate how to combine the σ_m coefficients for different phases to produce corrections for cases where the coefficients have not been tabulated.

2 ELLIPTICITY CORRECTIONS

In terms of the $\sigma_m(z_s, \Delta)$ coefficients the explicit representation of the time correction to be added to the value calculated for a spherically symmetric Earth model is

$$\begin{aligned} \delta t(z_s, \vartheta_0, \Delta, \zeta) = & \frac{1}{4} (1 + 3 \cos 2\vartheta_0) \sigma_0(z_s, \Delta) \\ & + \frac{3}{2} \sin 2\vartheta_0 \cos \zeta \sigma_1(z_s, \Delta) \\ & + \frac{3}{2} \sin^2 \vartheta_0 \cos 2\zeta \sigma_2(z_s, \Delta), \end{aligned} \quad (2.1)$$

where, as before, ϑ_0 is the epicentral co-latitude, Δ is the epicentral distance, ζ is the azimuth from the epicentre to the receiver, and z_s is source depth.

2.1 Ray based corrections

Doornbos (1988) has presented an effective algorithm for the computation of the σ_m ellipticity coefficients which depends on a specification of the ray parameter for any particular phase. This approach is based on computation of ray integrals for fixed slowness and in consequence needs to be coupled to an auxiliary routine if corrections are to be produced for fixed epicentral distance, Δ .

Buland & Chapman (1983) have introduced the tau-spline method for rapid construction of travel times as explicit functions of range. This method is based on the properties of the delay time τ

$$\tau(p) = T(p) - p\Delta(p). \quad (2.2)$$

The stationary values of

$$\theta(p, x) = \tau(p) + p \cdot x = T(p) + p[x - \Delta(p)] \quad (2.3)$$

correspond to geometrical arrivals for the range, x . With a tau-spline representation of $\tau(p)$ for each travel-time branch the condition $\partial\theta(p, x)/\partial p = 0$ leads to the solution of a quadratic to find the slowness, p_x , and then a rapid evaluation of the travel time for range x from

$$T(x) = \tau(p_x) + p_x x. \quad (2.4)$$

The tau-splines can be established for many phases at a time by using a common slowness discretisation and the superposition of τ segments for different ray segments. The appropriate software has been widely distributed with the release of the *iasp91* model (Kennett & Engdahl, 1991) and can be readily adapted to different radial velocity models.

The output from the Buland & Chapman (1983) algorithm is ideally suited for use with Doornbos' (1988) ellipticity corrections since the slowness is known explicitly for a given range and phase specification. Thus, provided a ray description is available, the ellipticity coefficients, σ_m , can be calculated rapidly and a cross check made of the accuracy of the calculation by comparison of the estimated and desired range. For the model *ak135* of Kennett, Engdahl & Buland (1995), which is specified by linear gradients in radius, the convergence is better than 0.025% for all the major phases.

This combination of ray algorithms has been used to develop an extensive set of ellipticity corrections in the form of σ_m coefficients as a function of depth, z_s , and range, Δ , (see Table 1). These corrections are available in electronic form as detailed in Appendix B.

In the application of such ellipticity corrections care must be taken where interpolation is applied to ensure that the corrections correspond to the same class of propagation path. It is important to separate the upgoing P and S waves from deep events (P_{up} , S_{up} in Table 1) from the main P and S which have a separate ray specification.

As noted by Dziewonski & Gilbert (1976) the ellipticity corrections for P and S phases having comparable paths scale very closely with a constant of proportionality which varies from 1.801 to 1.834, i.e. 1.82 ± 0.015 , despite the considerable variation in the ratios of P to S wave speed with depth. The constant of proportionality is very close to the value 1.8 proposed by Bullen (1937b). The simple rescaling works well for mantle phases, such as PcP and ScS , but cannot be applied where the propagation paths have different character, as e.g. core phases such as PKP and SKS .

2.2 Diffracted phases

The ellipticity coefficients for a diffracted phase can be determined from the values at the grazing point on the appropriate interface. For a given source depth, z_s , and epicentral distance for the grazing ray, Δ_g , we construct the coefficients for the grazing ray: $\sigma_0^g(z_s, \Delta_g)$, $\sigma_1^g(z_s, \Delta_g)$, $\sigma_2^g(z_s, \Delta_g)$.

As detailed in Appendix A, the ellipticity coefficients for diffracted waves are specified

by two integrals along the upward leg of the grazing ray:

$$X = \frac{1}{p} \int_{\Delta_a}^{\Delta_g} d\theta \cos 2\theta \epsilon \eta^3 \frac{d\eta}{dr} - \sum_j \left\{ \epsilon \cos 2\theta \left[(\eta^2 - p^2)^{1/2} \right]_-^+ \right\}_j \pm \sum_k \left\{ \epsilon \cos 2\theta \left[(\eta_{\pm}^2 - p^2)^{1/2} \right]_-^+ \right\}_k, \quad (2.5)$$

$$Y = \frac{1}{p} \int_{\Delta_a}^{\Delta_g} d\theta \sin 2\theta \epsilon \eta^3 \frac{d\eta}{dr} - \sum_j \left\{ \epsilon \sin 2\theta \left[(\eta^2 - p^2)^{1/2} \right]_-^+ \right\}_j \pm \sum_k \left\{ \epsilon \sin 2\theta \left[(\eta_{\pm}^2 - p^2)^{1/2} \right]_-^+ \right\}_k. \quad (2.6)$$

The change in arc length due to ellipticity is second order in ϵ and may therefore be ignored.

The σ_m^d coefficients for the diffracted phase may then be constructed from those at grazing, σ_m^g , by

$$\sigma_0^d(z_s, \Delta_g + \delta) = \sigma_0^g(z_s, \Delta_g) + \sin \delta (X \sin \delta + Y \cos \delta), \quad (2.7)$$

$$\sigma_1^d(z_s, \Delta_g + \delta) = \sigma_1^g(z_s, \Delta_g) + \frac{2}{\sqrt{3}} \sin \delta (Y \sin \delta - X \cos \delta), \quad (2.8)$$

$$\sigma_2^d(z_s, \Delta_g + \delta) = \sigma_2^g(z_s, \Delta_g) - \frac{1}{\sqrt{3}} \sin \delta (X \sin \delta + Y \cos \delta). \quad (2.9)$$

where δ is the arc distance of diffraction. Only two numerical integrals need to be evaluated to generate the new coefficients. The integrals X and Y are weakly dependent on source depth through the variation in epicentral distance of the grazing point.

The set of σ coefficients for P_{diff} and S_{diff} diffracted at the core-mantle boundary are illustrated in figure 1 as functions of epicentral distance, together with those for P and S . As we expect there is a smooth transition from the phase defined by a geometric ray to the diffracted phase. For short diffracted legs a simple extrapolation of the geometric results provides an adequate approximation.

The σ coefficients for P_{diff} and S_{diff} are tabulated in Tables 2,3 in comparable form to that employed in Kennett (1991) for P and S .

3 COMBINATION OF ELLIPTICITY CORRECTIONS FOR COMPOSITE PHASES

The form of the ellipticity correction (1.2) is the same for all phases with the coefficients σ_m being the only phase-dependent quantity. Thus, if we can construct the wave path of a particular phase as a combination of two or more elementary phases (e.g. $PKKP$ as a combination of PKP and PcP), we can not only construct the correction for the composite phase from the corrections for the elementary phases, but also construct the σ_m tables for the composite phase by a linear combination of the σ_m tables for the elementary phases. Consider, for example, a surface reflected phase with two legs, a and b , where the point of reflection lies at Δ_a so that the two propagation legs are $(0, \Delta_a)$ and (Δ_a, Δ) (see figure 2). The corrections could be determined by calculating the coefficients for the Δ_a leg and then introducing an apparent source at the reflection point to build a

further correction for the arc length $\Delta_b = \Delta - \Delta_a$. Alternatively, we can use spherical trigonometry to develop an expression for the composite σ_m coefficients:

$$\begin{aligned}\sigma_0^c(z_s, \Delta) &= \sigma_0^a(z_s, \Delta_a) + \frac{1}{2}(3 \cos^2 \Delta_a - 1)\sigma_0^b(0, \Delta_b) \\ &\quad - \frac{\sqrt{3}}{2} \sin 2\Delta_a \sigma_1^b(0, \Delta_b) \\ &\quad + \frac{\sqrt{3}}{2} \sin^2 \Delta_a \sigma_2^b(0, \Delta_b),\end{aligned}\quad (3.1)$$

$$\begin{aligned}\sigma_1^c(z_s, \Delta) &= \sigma_1^a(z_s, \Delta_a) + \frac{\sqrt{3}}{2} \sin 2\Delta_a \sigma_0^b(0, \Delta_b) \\ &\quad + \cos 2\Delta_a \sigma_1^b(0, \Delta_b) \\ &\quad - \frac{1}{2} \sin 2\Delta_a \sigma_2^b(0, \Delta_b),\end{aligned}\quad (3.2)$$

$$\begin{aligned}\sigma_2^c(z_s, \Delta) &= \sigma_2^a(z_s, \Delta_a) + \frac{\sqrt{3}}{2} \sin^2 \Delta_a \sigma_0^b(0, \Delta_b) \\ &\quad + \frac{1}{2} \sin 2\Delta_a \sigma_1^b(0, \Delta_b) \\ &\quad + \frac{1}{2}(1 + \cos^2 \Delta_a) \sigma_2^b(0, \Delta_b),\end{aligned}\quad (3.3)$$

where

$$\Delta_b = \Delta - \Delta_a. \quad (3.4)$$

The derivation of these expressions is presented in Appendix A.

As a quick check on the composition rules we compared a direct calculation for the coefficients for PP at 60° , which is composed of two P legs of 30° each, with the results from the application of (3.1)-(3.4). The resulting coefficients differ by no more than 0.006 s, which reflects the truncation errors associated with tabulated values.

More complex combinations of phases can also be made by exploiting the linear superposition of the corrections for different propagation paths. Consider for example $PKKP$, which can be constructed from the information from PKP and PcP . For a full propagation path, Δ , we can build the travel time for $PKKP$ by a combination of two PKP legs with the subtraction of a PcP segment:

$$T^{PKKP}(z_s, \Delta) = T^{PKP}(z_s, \Delta_a) + T^{PKP}(0, \Delta_b) - T^{PcP}(0, \Delta_c), \quad (3.5)$$

where

$$\Delta_c = \Delta_a - (\Delta - \Delta_b). \quad (3.6)$$

Δ_a and Δ_b are the arc lengths of the two PKP legs which have to be found by matching the ray parameter (see figure 5).

The σ_m coefficients for $PKKP$ are then to be found from a linear combination of the coefficients for the PKP and PcP components:

$$\begin{aligned}\sigma_0^{PKKP}(z_s, \Delta) &= \sigma_0^{PKP}(z_s, \Delta_a) + \frac{1}{2}(3 \cos^2 \Delta_a - 1)[\sigma_0^{PKP}(0, \Delta_b) - \sigma_0^{PcP}(0, \Delta_c)] \\ &\quad - \frac{\sqrt{3}}{2} \sin 2\Delta_a [\sigma_1^{PKP}(0, \Delta_b) - \sigma_1^{PcP}(0, \Delta_c)] \\ &\quad + \frac{\sqrt{3}}{2} \sin^2 \Delta_a [\sigma_2^{PKP}(0, \Delta_b) - \sigma_2^{PcP}(0, \Delta_c)],\end{aligned}\quad (3.7)$$

$$\begin{aligned}\sigma_1^{PKKP}(z_s, \Delta) &= \sigma_1^{PKP}(z_s, \Delta_a) + \frac{\sqrt{3}}{2} \sin 2\Delta_a \sigma_0^b(0, \Delta_b) \\ &\quad + \cos 2\Delta_a [\sigma_1^{PKP}(0, \Delta_b) - \sigma_1^{PcP}(0, \Delta_c)] \\ &\quad - \frac{1}{2} \sin 2\Delta_a [\sigma_2^{PKP}(0, \Delta_b) - \sigma_2^{PcP}(0, \Delta_c)],\end{aligned}\quad (3.8)$$

$$\begin{aligned}\sigma_2^{PKP}(z_s, \Delta) = & \sigma_2^{PKP}(z_s, \Delta_a) + \frac{\sqrt{3}}{2} \sin^2 \Delta_a \sigma_0^b(0, \Delta_b) \\ & + \frac{1}{2} \sin 2\Delta_a [\sigma_1^{PKP}(0, \Delta_b) - \sigma_1^{PCP}(0, \Delta_c)] \\ & + \frac{1}{2} (1 + \cos^2 \Delta_a) [\sigma_2^{PKP}(0, \Delta_b) - \sigma_2^{PCP}(0, \Delta_c)].\end{aligned}\quad (3.9)$$

4 DISCUSSION

Although the significance of the Earth's ellipticity on travel-times has long been recognised, applications have been restricted to a few major phases. The extension of the corrections for diffracted phases and for combinations of seismic phases means that it is now feasible to generate corrections for most observable body-wave phases.

We hope that the provision of a wide-range of ellipticity coefficients in readily useable electronic form will encourage the systematic inclusion of the effects of ellipticity in both regional and global studies.

APPENDIX A: CORRECTIONS FOR DIFFRACTED AND COMPOSITE PHASES

A1 Diffracted phases

The geometry of a phase diffracted around a spherical boundary within the Earth is shown and defined in figure A1. Energy is radiated downward from the source at S at a particular angle defining a ray parameter, $p = p_g$, such that when the wave refracts toward the boundary it grazes it at point G, an arc length Δ_a away from the source. The ray path then refracts upwards toward point P on the surface a further arc length Δ_b from G. The continuous solid curve in the figure (SGP) defines the grazing ray to the boundary. Some energy may diffract along the boundary from G to H, where it is refracted upwards again to reach the surface at R, an arc length $\Delta = \Delta_a + \Delta_b + \delta$ away from the source. At high frequencies the diffracted energy is confined to the upper edge of the boundary and has a predictable travel time, which is constructed by a path integral of the slowness of the Earth along the path (SGHR) shown in figure A1, albeit not a "ray" path. The geometry of this path and the slowness encountered along it, and consequently the predicted travel time, will be perturbed by the Earth's ellipticity in a manner much like a true "ray" path as formulated by Dziewonski and Gilbert (1976).

We can divide the path SGHR (see figure A1) into three segments, SG, GH, and HR, and thus construct the perturbation due to ellipticity by combining three contributions. The first and third path segments are "ray" paths and may be treated as such using the results of Dziewonski and Gilbert (1976). The middle segment, GH, encounters a slowness equal to the slowness at the top side of the boundary in the spherical reference Earth model used. It therefore suffers no effect other than that due to change of arc length along the boundary. This in turn is second order in ϵ , the Earth's ellipticity of figure, and may be ignored. We are then left with the "ray" path effects along segments SG and HR.

Before commencing the derivation from the formulation of Dziewonski and Gilbert (1976) it is convenient to simplify their result (eq. 1.3) by defining:

$$f(r) = \frac{1}{p_g} \eta^3(r) \frac{dv}{dr} \epsilon(r), \quad (A1)$$

and

$$\Delta F_i = - \left\{ \epsilon \left[\left(\eta^2 - p_g^2 \right)^{1/2} \right]_{-}^{+} \right\}_i \quad \text{or} \quad \Delta F_i = \pm \left\{ \epsilon \left[\left(\eta_{\pm}^2 - p_g^2 \right)^{1/2} \right]_{-}^{+} \right\}_i, \quad (A2)$$

depending upon the nature of the interaction with the particular boundary in the Earth, which here is labelled by i . We then have a more compact expression for the coefficients σ_m :

$$\sigma_m(z_s, \Delta) = \int_0^{\Delta} d\theta f(r) \lambda_m(\theta) + \sum_i \Delta F_i \lambda_m(\theta_i) \quad (A3)$$

We can now follow Dziewonski and Gilbert (1976) and write for the diffracted phase:

$$\begin{aligned} \sigma_m^d(z_s, \Delta_g + \delta) &= \int_0^{\Delta_g} d\theta f(r) \lambda_m(\theta) + \sum_{i, SG} \Delta F_i \lambda_m(\theta_i) \\ &\quad + \int_{\Delta_g + \delta}^{\Delta_g + \delta} d\theta f(r) \lambda_m(\theta) + \sum_{i, HR} \Delta F_i \lambda_m(\theta_i) \end{aligned} \quad (A4)$$

where $\Delta_g = \Delta_s - \Delta_t$ is the epicentral distance of the grazing ray. The sums labelled SG and HR represent sums over all discontinuities in velocity (slowness) encountered by the SG and HR paths respectively. The θ_i under $\sum_{HR}$ refers to the angular distance from S along the HR path segment. Note that the HR path is a replica of the GP path shifted by δ in angular distance from S. A simple change of variable in the last two terms of (A4) to $\theta' = \theta - \delta$ thus transforms HR to GP and we have:

$$\begin{aligned} \sigma_m^d(z_s, \Delta_g + \delta) &= \int_0^{\Delta_g} d\theta f(r) \lambda_m(\theta) + \sum_{i, SG} \Delta F_i \lambda_m(\theta_i) \\ &\quad + \int_{\Delta_g}^{\Delta_g} d\theta' f(r) \lambda_m(\theta' + \delta) + \sum_{i, GP} \Delta F_i \lambda_m(\theta_i' + \delta) \end{aligned} \quad (A5)$$

Now recall that $\lambda_m(\theta) = -\frac{2}{3} P_{2,m}(\cos \theta)$, where $P_{2,m}$ are associated Legendre functions of degree 2. Simple algebra leads to:

$$\begin{aligned} \lambda_0(\theta + \delta) &= \lambda_0(\theta) + \sin \delta (\sin \delta \cos 2\theta + \cos \delta \sin 2\theta) \\ \lambda_1(\theta + \delta) &= \lambda_1(\theta) + \frac{2}{\sqrt{3}} \sin \delta (\sin \delta \sin 2\theta - \cos \delta \cos 2\theta) \\ \lambda_2(\theta + \delta) &= \lambda_2(\theta) - \frac{1}{\sqrt{3}} \sin \delta (\sin \delta \cos 2\theta + \cos \delta \sin 2\theta) \end{aligned} \quad (A6)$$

Thus, if we define:

$$X = \int_{\Delta_g}^{\Delta_g} d\theta f(r) \cos 2\theta + \sum_{i, GP} \Delta F_i \cos 2\theta_i$$

and

$$Y = \int_{\Delta_g}^{\Delta_g} d\theta f(r) \sin 2\theta + \sum_{i, GP} \Delta F_i \sin 2\theta_i \quad (A7)$$

we can write

$$\begin{aligned}
 \sigma_0^d(z_s, \Delta_g + \delta) &= \int_0^{\Delta_g} d\theta f(r) \lambda_0(\theta) + \sum_{i,SG} \Delta F_i \lambda_0(\theta_i) \\
 &\quad + \int_{\Delta_g}^{\Delta_g} d\theta f(r) \lambda_0(\theta) + \sum_{i,GP} \Delta F_i \lambda_0(\theta_i) \\
 &\quad + \sin \delta (X \sin \delta + Y \cos \delta) \\
 &= \int_0^{\Delta_g} d\theta f(r) \lambda_0(\theta) + \sum_{i,SP} \Delta F_i \lambda_0(\theta_i) \\
 &\quad + \sin \delta (X \sin \delta + Y \cos \delta) \\
 &= \sigma_0^g(z_s, \Delta_g) + \sin \delta (X \sin \delta + Y \cos \delta)
 \end{aligned} \tag{A8}$$

and similarly for σ_1^d and σ_2^d :

$$\sigma_1^d(z_s, \Delta_g + \delta) = \sigma_1^g(z_s, \Delta_g) + \frac{2}{\sqrt{3}} \sin \delta (Y \sin \delta - X \cos \delta) \tag{A9}$$

$$\sigma_2^d(z_s, \Delta_g + \delta) = \sigma_2^g(z_s, \Delta_g) - \frac{1}{\sqrt{3}} \sin \delta (X \sin \delta + Y \cos \delta) \tag{A10}$$

Note that relative to the grazing ray, the change in the σ_m tables is expandable in terms of $\sin \delta$, where δ is the arc length of diffraction. Thus, for a short diffraction distance the σ_m tables for the diffracted phase are well represented by a linear extrapolation of the tables for the associated geometrical ray. It is apparent in figure 1 that the ellipticity coefficients are continuous to first order across the transition from an optical path to a diffracted path (dashed vertical lines). It is possible to show analytically that this is the case for surface focus paths (S at the surface) and approximately so in general. The proof of this statement is beyond the scope of this paper. This is potentially useful because if one has the coefficients sufficiently accurately tabulated for the refracted phase one can estimate the integrals X and Y from their derivatives.

It is worth noting that while figure A1 draws the diffraction boundary near the core-mantle boundary and our numerical examples are for diffractions around the core-mantle boundary, the arguments and derivation above apply to diffractions around arbitrary spherical boundaries in the Earth. Our derivation holds true whenever the predicted time of the diffracted phase is constructed by a path integral akin to the path depicted in figure A1. The derivation applies e.g. to diffractions off the caustics of upper mantle triplications.

It is also worth noting that our description of the diffracted path involves a high-frequency assumption. High-frequency diffractions decay rapidly with propagation distance and the travel time of diffracted waves is dispersive. The change of travel time with frequency can be understood in terms of a change in the diffraction geometry. We have not explored the potential deviations from the above derivation due to finite frequency.

A2 Composite phases

If we can construct a wave path of a particular phase as a combination of two or more elementary phases we can evaluate the ellipticity correction for the composite phase by adding the corrections for the elementary phases according to eq. 1.2 using the appropriate epicentral distances, colatitudes and azimuths. Because the form of the ellipticity corrections is the same for all phases we can go further and construct the tables of ellipticity coefficients, σ_m , for the composite phase from the σ_m coefficients of the elementary phases. We start by introducing the shorthand:

$$g_m(\theta, \zeta) = P_{2,m}(\cos \theta) \cos m\zeta. \quad (\text{A11})$$

We can then write the ellipticity correction for the composite phase, δt^c , either in terms of the coefficients for the composite phase, σ_m^c , or in terms of the coefficients for the elementary phases, σ_m^a and σ_m^b (see figure A2 for definitions of geometry):

$$\begin{aligned} \delta t^c &= g_0(\theta_s, \zeta_s) \sigma_0^c(z_s, \Delta) + g_1(\theta_s, \zeta_s) \sigma_1^c(z_s, \Delta) + g_2(\theta_s, \zeta_s) \sigma_2^c(z_s, \Delta) \\ &= g_0(\theta_s, \zeta_s) \sigma_0^a(z_s, \Delta_a) + g_1(\theta_s, \zeta_s) \sigma_1^a(z_s, \Delta_a) + g_2(\theta_s, \zeta_s) \sigma_2^a(z_s, \Delta_a) \\ &\quad + g_0(\theta_p, \zeta_p) \sigma_0^b(0, \Delta_b) + g_1(\theta_p, \zeta_p) \sigma_1^b(0, \Delta_b) + g_2(\theta_p, \zeta_p) \sigma_2^b(0, \Delta_b). \end{aligned} \quad (\text{A12})$$

In order to construct the coefficients σ_m^c we need to transcribe the terms involving θ_p and ζ_p , which arise through the shift of the origin of the path for phase b, in terms of θ_s and ζ_s . We utilize the following relationships from spherical trigonometry (we apply the sine and cosine rules to the left-hand side triangle in figure A2):

$$\sin \zeta_s, \sin \theta_p = \sin(\pi - \zeta_p) / \sin \theta_s = \sin \zeta_p / \sin \theta_s,$$

$$\cos \theta_p = \cos \theta_s \cos \Delta_a + \sin \theta_s \sin \Delta_a \cos \zeta_s$$

$$\cos \theta_s = \cos \theta_p \cos \Delta_a + \sin \theta_p \sin \Delta_a \cos \zeta_p,$$

from which we arrive at

$$\sin \theta_p \sin \zeta_p = \sin \theta_s \sin \zeta_s$$

$$\sin \theta_p \cos \zeta_p = \sin \Delta_a \cos \theta_s - \cos \Delta_a \sin \theta_s \cos \zeta_s$$

by simple algebraic manipulation. We can now rewrite the terms of eq A11 which involve θ_p and ζ_p , the terms $g_m(\theta_p, \zeta_p)$, as:

$$\begin{aligned} g_0(\theta_p, \zeta_p) &= \frac{1}{4}(1 + 3 \cos 2\Delta_a)g_0(\theta_s, \zeta_s) + \frac{\sqrt{3}}{2} \sin 2\Delta_a g_1(\theta_s, \zeta_s) + \frac{\sqrt{3}}{2} \sin^2 \Delta_a g_2(\theta_s, \zeta_s), \\ g_1(\theta_p, \zeta_p) &= -\frac{\sqrt{3}}{2} \sin 2\Delta_a g_0(\theta_s, \zeta_s) + \cos 2\Delta_a g_1(\theta_s, \zeta_s) + \frac{1}{2} \sin 2\Delta_a g_2(\theta_s, \zeta_s), \quad (\text{A13}) \\ g_2(\theta_p, \zeta_p) &= \frac{\sqrt{3}}{2} \sin^2 \Delta_a g_0(\theta_s, \zeta_s) - \frac{1}{2} \sin 2\Delta_a g_1(\theta_s, \zeta_s) + \frac{1}{2} (1 + \cos^2 \Delta_a) g_2(\theta_s, \zeta_s). \end{aligned}$$

We find that they decompose into combinations of terms $g_m(\theta_s, \zeta_s)$ with weights which depend upon the angular distance of the shift of origin of phase b from that of phase c, Δ_a .

It is now straight forward to match terms in eq A11 and write the ellipticity correction coefficients for phase c, σ_m^c , in terms of the coefficients for the elementary phases:

$$\begin{aligned}
 \sigma_0^c(z_s, \Delta) &= \sigma_0^a(z_s, \Delta_a) + \frac{1}{4}(1 + 3 \cos 2\Delta_a)\sigma_0^b(0, \Delta_b) \\
 &\quad - \frac{\sqrt{3}}{2} \sin 2\Delta_a \sigma_1^b(0, \Delta_b) + \frac{\sqrt{3}}{2} \sin^2 \Delta_a \sigma_2^b(0, \Delta_b) \\
 \sigma_1^c(z_s, \Delta) &= \sigma_1^a(z_s, \Delta_a) + \frac{\sqrt{3}}{2} \sin 2\Delta_a \sigma_0^b(0, \Delta_b) \\
 &\quad + \cos 2\Delta_a \sigma_1^b(0, \Delta_b) - \frac{1}{2} \sin 2\Delta_a \sigma_2^b(0, \Delta_b) \\
 \sigma_2^c(z_s, \Delta) &= \sigma_2^a(z_s, \Delta_a) + \frac{\sqrt{3}}{2} \sin^2 \Delta_a \sigma_0^b(0, \Delta_b) \\
 &\quad + \frac{1}{2} \sin 2\Delta_a \sigma_1^b(0, \Delta_b) + \frac{1}{2} (1 + \cos^2 \Delta_a) \sigma_2^b(0, \Delta_b)
 \end{aligned} \tag{A14}$$

In order to implement these expressions one needs to match the ray parameters of phases a, b, and c in order to determine Δ_a and Δ_b from Δ and then the σ_m tables of phases a and b can be combined to calculate the tables for phase c.

APPENDIX B: ELECTRONIC FORMS FOR ELLIPTICITY CORRECTIONS

A Fortran subroutine `ellip` provides access to the ellipticity corrections for a given source location, epicentral distance and azimuth using direct access tables of the σ_m coefficients (as specified in Table 1. This subroutine and the necessary direct access table `ellipcor.tbl` are available using anonymous ftp at the following sites:

- the Research School of Earth Sciences, Australian National University, Canberra, Australia: `rses.anu.edu.au`
- the IRIS Data Management Centre in Seattle, Washington, U.S.A. `iris.washington.edu`
- the National Earthquake Information Centre, U.S. Geological Survey, Golden, Colorado, U.S.A.: `gldfs.cr.usgs.gov/directory/ellip`

A further program `ttimel` provides combines access to rapid calculation of seismic travel times, using the *ak135* model of Kennett, Engdahl & Buland (1995), with ellipticity corrections for the full range of phases in table 1.

ACKNOWLEDGMENTS

This work was supported in part by grant F49620-94-1-0110 from the U.S. Air Force Office of Scientific Research. We are grateful to the various site managers for the arrangements for ftp access.

References

- Buland R. & Chapman C.H., 1983. The computation of seismic travel times, *Bull. seism. Soc. Am.*, **73**, 1271-1302.
- Bullen, K.E., 1937a. The ellipticity correction to travel times of *P* and *S* earthquake waves, *Mon. Not. R. astr. Soc. Geophys. Suppl.*, **4**, 143-157.
- Bullen, K.E., 1937b. A suggested new 'seismological' latitude, *Mon. Not. R. astr. Soc. Geophys. Suppl.*, **4**, 158-164.
- Doornbos, D.J., 1988. Asphericity and Ellipticity Corrections in *Seismological algorithms* (D.J. Doornbos Ed.), 75-85, Academic Press, New York.
- Dziewonski, A.M. & Gilbert, F., 1976. The effect of small, aspherical perturbations on travel times and a re-examination of the corrections for ellipticity, *Geophys. J. R. astr. Soc.*, **44**, 7-17.
- Dziewonski, A.M., Hales, A.L. & Lapwood, E.R., 1975. Parametrically simple Earth models consistent with geophysical data, *Phys. Earth. Planet. Inter.*, **10**, 12-48.
- Engdahl, E.R. & Gunst, R.H., 1966. Use of high speed computer for the preliminary determination of earthquake hypocenters, *Bull. seism. Soc. Am.*, **56**, 325-336.
- Kennett, B.L.N., 1991. *IASPEI 1991 Seismological Tables*, Research School of Earth Sciences, Australian National University.
- Kennett, B.L.N. & Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.*, **105**, 429-465.
- Kennett, B.L.N., Engdahl, E.R. & Buland, R., 1995. Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, **122**, 108-124.

Figure 1. (a) $\sigma_m(\Delta)$ ellipticity coefficients for P and P_{diff} , (b) $\sigma_m(\Delta)$ ellipticity coefficients for S and S_{diff} .

Figure 2. Addition of ray legs to produce a composite phase

Figure 3. Addition of ray legs to construct $PKKP$ from two PKP segments and PcP .

Figure A1. Ellipticity corrections for diffracted phases

Figure A2. Geometric relations on a sphere for phase addition

Table 1. Ellipticity corrections

Phase	$\Delta_{\min}$	$\Delta_{\max}$
P_{up}	0	10
P	5	95
P_{diff}	100	150
PKP_{ab}	145	175
PKP_{bc}	145	155
PKP_{df}	115	180
$PKiKP$	0	155
pP	20	100
$pPKP_{ab}$	145	175
$pPKP_{bc}$	145	155
$pPKP_{df}$	115	180
$pPKiKP$	0	155
sP	5	100
$sPKP_{ab}$	145	175
$sPKP_{bc}$	145	155
$sPKP_{df}$	115	180
$sPKiKP$	0	155
P_cP	0	90
S_cS	0	60
SKP_{ab}	130	140
SKP_{bc}	130	150
SKP_{df}	110	180
$SKiKP$	0	145
$PKKP_{ab}$	235	255
$PKKP_{bc}$	235	285
$PKKP_{df}$	210	360
$SKKP_{ab}$	215	220
$SKKP_{bc}$	215	280
$SKKP_{df}$	205	360
PP'	40	190
$P'P'$	235	360
S_{up}	0	10
S	5	95
S_{diff}	100	150
SKS_{ac}	65	140
SKS_{df}	105	180
pS	60	100
$pSKS_{ac}$	70	140
$pSKS_{df}$	110	180
sS	20	100
$sSKS_{ac}$	65	140
$sSKS_{df}$	110	180
S_cS	0	90
P_cS	0	60
PKS_{ab}	130	140
PKS_{bc}	130	145
PKS_{df}	110	180
$PKKS_{ab}$	215	220
$PKKS_{bc}$	215	280
$PKKS_{df}$	205	360
$SKKS_{ac}$	65	275
$SKKS_{df}$	200	360
SS	40	190
$S'S'$	130	360
SP	55	135
PS	90	135
PnS	65	90

Table 2. Ellipticity coefficients P_{diff}

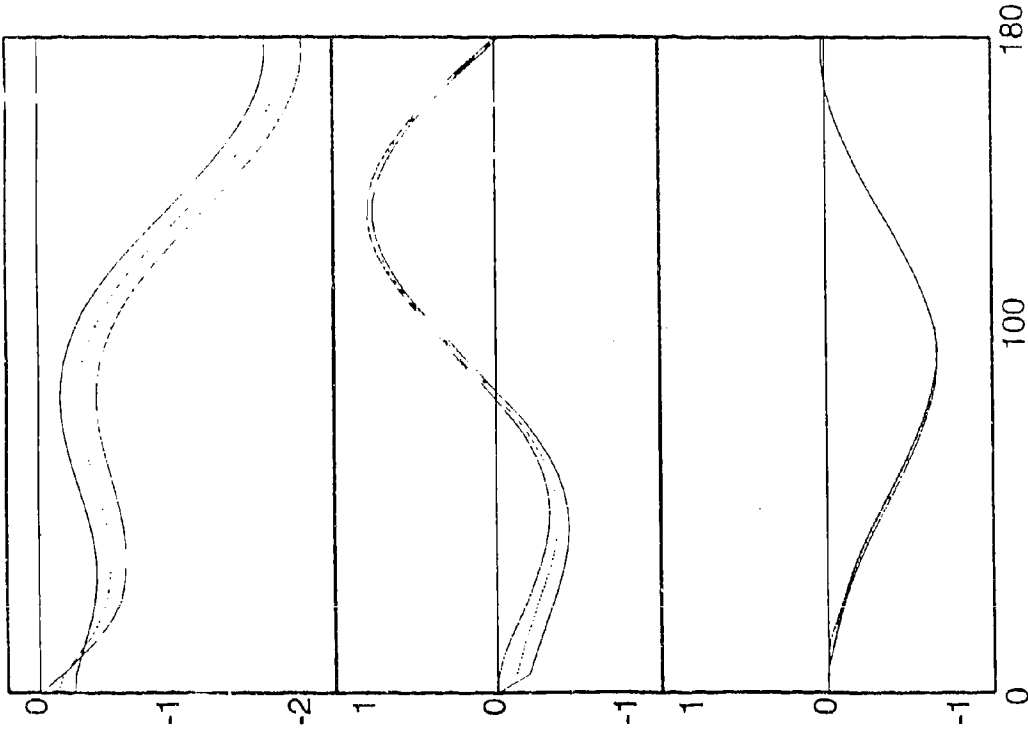
Depth of source [km]												
Δ	0.			35.			50.			100.		
	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2
100.0	-0.568	0.496	-0.808	-0.550	0.493	-0.808	-0.543	0.492	-0.808	-0.518	0.489	-0.808
105.0	-0.637	0.627	-0.768	-0.620	0.625	-0.768	-0.612	0.623	-0.768	-0.587	0.620	-0.768
110.0	-0.726	0.742	-0.716	-0.708	0.740	-0.717	-0.701	0.739	-0.717	-0.676	0.735	-0.717
115.0	-0.830	0.838	0.656	-0.813	0.836	-0.656	-0.805	0.835	-0.656	-0.780	0.831	-0.657
120.0	-0.947	0.911	-0.588	-0.930	0.909	-0.589	-0.922	0.908	-0.589	-0.898	0.904	-0.589
125.0	-1.074	0.960	-0.515	-1.056	0.958	-0.516	-1.049	0.957	-0.516	-1.024	0.953	-0.516
130.0	-1.206	0.983	-0.439	-1.188	0.980	-0.439	-1.181	0.979	-0.439	-1.156	0.976	-0.440
135.0	-1.339	0.979	-0.362	-1.322	0.976	-0.362	-1.314	0.975	-0.363	-1.289	0.972	-0.363
140.0	-1.470	0.948	-0.289	-1.452	0.945	-0.287	-1.445	0.944	-0.287	-1.420	0.941	-0.287
145.0	-1.594	0.891	-0.215	-1.576	0.889	-0.215	-1.569	0.888	-0.216	-1.544	0.884	-0.216
150.0	-1.708	0.811	-0.149	-1.690	0.808	-0.150	-1.682	0.807	-0.150	-1.658	0.804	-0.150
Δ	200.			300.			500.			700.		
	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2
100.0	-0.474	0.482	-0.809	-0.431	0.476	-0.809	-0.354	0.465	-0.810	-0.285	0.455	-0.812
105.0	-0.543	0.614	-0.768	-0.501	0.607	-0.769	-0.424	0.596	-0.770	-0.355	0.587	-0.771
110.0	-0.632	0.729	-0.718	-0.590	0.723	-0.718	-0.512	0.711	-0.719	-0.443	0.702	-0.720
115.0	-0.736	0.825	-0.657	-0.694	0.818	-0.658	-0.617	0.807	-0.659	-0.548	0.798	-0.660
120.0	-0.853	0.898	-0.589	-0.811	0.892	-0.590	-0.734	0.881	-0.591	-0.665	0.871	-0.592
125.0	-0.980	0.947	-0.516	-0.938	0.940	-0.517	-0.861	0.929	-0.518	-0.791	0.920	-0.519
130.0	-1.112	0.969	-0.440	-1.070	0.963	-0.441	-0.992	0.952	-0.442	-0.923	0.942	-0.443
135.0	-1.245	0.965	-0.363	-1.203	0.959	-0.364	-1.126	0.948	-0.365	-1.057	0.938	-0.366
140.0	-1.376	0.934	-0.288	-1.334	0.928	-0.288	-1.256	0.917	-0.289	-1.187	0.907	-0.291
145.0	-1.500	0.878	-0.216	-1.458	0.872	-0.217	-1.381	0.860	-0.218	-1.311	0.851	-0.219
150.0	-1.614	0.797	-0.151	-1.572	0.795	-0.151	-1.494	0.780	-0.152	-1.425	0.770	-0.153

Table 3. Ellipticity coefficients for S_{diff}

Depth of source [km]												
Δ	0.			35.			50.			100.		
	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2
100.0	-1.041	0.902	-1.480	-1.010	0.896	-1.480	-0.997	0.894	-1.481	-0.953	0.886	-1.481
105.0	-1.167	1.144	-1.407	-1.137	1.138	-1.407	-1.124	1.136	-1.407	-1.080	1.128	-1.407
110.0	-1.329	1.357	-1.314	-1.298	1.351	-1.314	-1.285	1.349	-1.314	-1.241	1.341	-1.314
115.0	-1.520	1.534	-1.204	-1.489	1.529	-1.204	-1.476	1.526	-1.204	-1.432	1.518	-1.204
120.0	-1.734	1.671	-1.080	-1.704	1.667	-1.080	-1.691	1.662	-1.080	-1.647	1.655	-1.080
125.0	-1.966	1.762	-0.946	-1.936	1.756	-0.946	-1.923	1.754	-0.946	-1.879	1.746	-0.946
130.0	-2.208	1.805	-0.806	-2.178	1.799	-0.806	-2.165	1.797	-0.806	-2.121	1.789	-0.806
135.0	-2.453	1.799	-0.664	-2.423	1.794	-0.665	-2.410	1.791	-0.665	-2.366	1.783	-0.665
140.0	-2.694	1.744	-0.526	-2.663	1.739	-0.526	-2.650	1.736	-0.526	-2.606	1.728	-0.526
145.0	-2.922	1.642	-0.394	-2.892	1.636	-0.394	-2.879	1.634	-0.394	-2.835	1.626	-0.394
150.0	-3.132	1.496	-0.273	-3.101	1.490	-0.273	-3.088	1.487	-0.273	-3.044	1.479	-0.273
Δ	200.			300.			500.			700.		
	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2	σ_0	σ_1	σ_2
100.0	0.876	0.882	-1.482	-0.796	0.860	-1.482	-0.652	0.824	-1.483	-0.528	0.808	-1.484
105.0	-1.003	1.124	-1.408	-0.923	1.102	-1.409	-0.778	1.065	-1.409	-0.655	1.050	-1.411
110.0	-1.165	1.337	-1.315	-1.085	1.314	-1.315	-0.940	1.278	-1.316	-0.816	1.263	-1.318
115.0	-1.356	1.515	-1.205	-1.275	1.492	-1.205	-1.131	1.456	-1.206	-1.007	1.441	-1.207
120.0	-1.570	1.654	-1.081	-1.490	1.628	-1.081	-1.345	1.592	-1.082	-1.222	1.577	-1.084
125.0	-1.802	1.742	-0.947	-1.722	1.719	-0.947	-1.577	1.683	-0.948	-1.454	1.668	-0.950
130.0	-2.045	1.785	-0.807	-1.964	1.763	-0.807	-1.820	1.727	-0.808	-1.696	1.712	-0.810
135.0	-2.289	1.780	-0.666	-2.209	1.757	-0.666	-2.064	1.721	-0.667	-1.941	1.706	-0.668
140.0	-2.530	1.725	-0.527	-2.450	1.702	-0.527	-2.305	1.666	-0.528	-2.181	1.651	-0.530
145.0	-2.758	1.622	-0.395	-2.678	1.600	-0.395	-2.533	1.564	-0.396	-2.410	1.548	-0.398
150.0	-2.968	1.476	-0.274	-2.888	1.453	-0.274	-2.743	1.417	-0.275	-2.619	1.402	-0.277

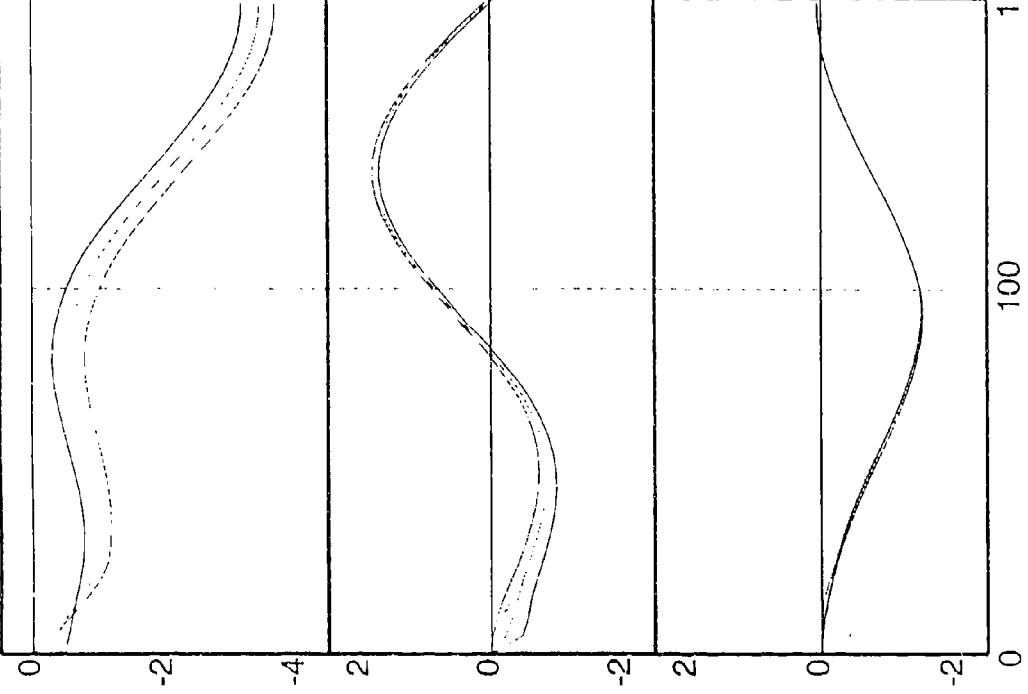
P & Pdiff

time (seconds)



S & Sdiff

time (seconds)



epicentral distance (degrees)

Fig. 1

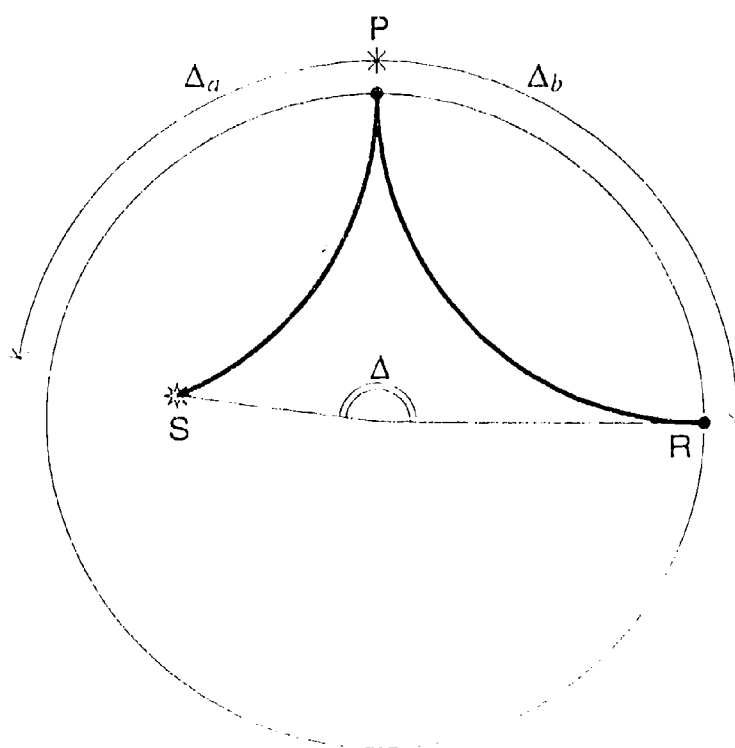


Fig 2

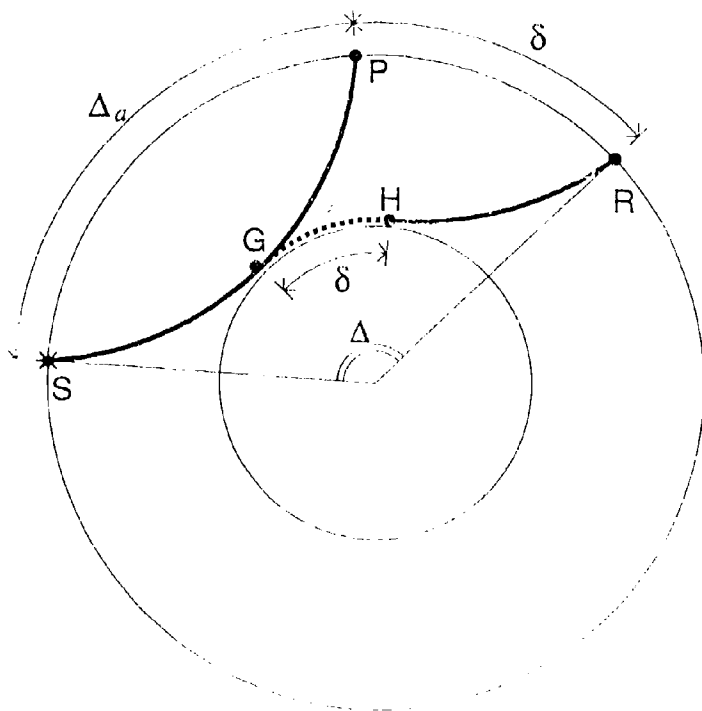


Fig 41

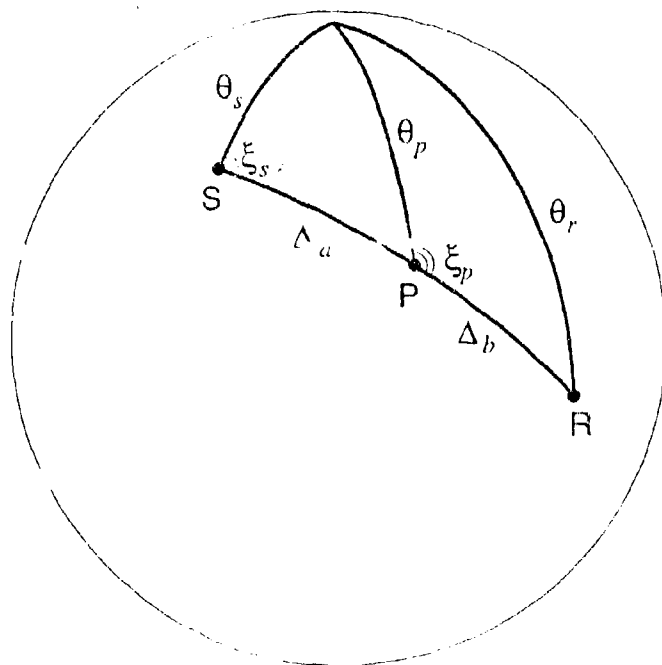


Fig 42

Frequency-dependent effects on travel times and waveforms of long-period *S* waves: implications for the scale of mantle heterogeneity

Heiner Igel, Institute of Theoretical Geophysics, Cambridge, UK
Oli Gudmundsson, Research School of Earth Sciences, Canberra, Australia

Submitted to *FEPL*, July 1996

Abstract

The spectral properties of upper-mantle velocity perturbations are controversial. While global Earth models obtained by tomographic inversion are characterized by the dominance of long-wavelength features, results from regional tomography, and stochastic analysis of travel times suggest the presence of more power at intermediate scales than predicted by global models. We study frequency-dependent effects of long-period *S*-wave propagation through random upper-mantle models with specified spectral properties. Wave propagation is simulated by a finite-difference approximation to the axi-symmetric wave equation in spherical coordinates. For global models with cylindrical symmetry and constant angular increment $\Delta\theta$ the use of spherical coordinates leads to an effective lateral grid spacing (arc length) decreasing with depth. This is contrary to the requirements of global models with low velocities at the top of the mantle, which necessitate a dense grid spacing at small depths and a wider grid spacing at the base of the mantle. We introduce a grid with depth-dependent lateral grid spacing to overcome this inconsistency. Our simulations suggest that (1) the properties of power spectra of travel-time residuals are frequency dependent; (2) power spectra of models obtained from long-period tomography may considerably underestimate the power at intermediate scales; (3) frequency-dependent effects on the waveform are sensitive to the scales and amplitudes of perturbations present in the upper mantle.

Keywords: mantle structure, finite differences, *S*-waves

Introduction

Determining the spectrum of heterogeneities in the Earth's mantle is important to constrain the types of convection present in the mantle and to understand the origin of modeled seismic-velocity anomalies. Spectral properties of mantle structure can be estimated from three-dimensional (3D) models obtained by global tomography or they can be estimated directly from observed travel-time residuals. A detailed comparison of power spectra for different global and regional models can be found in Passier and Snieder (1995). A discussion of global mantle models is given in Ritzwoller and Lavelle (1995).

The spectral properties of mantle-velocity perturbations determined by various methods often differ substantially (Passier and Snieder, 1995). While tomographic images obtained from global-data inversion suggest a sharp onset of spectral decay at harmonic degree $l \approx 8$ (e.g. Su and Dziewonski, 1992; Zhang and Tanimoto, 1991, 1993; Pollitz, 1994) results from regional tomography (e.g. Snieder, 1988; Zielhuis and Nolet, 1994; Spakman et al., 1993) and stochastic analysis of travel times (Gudmundsson et al., 1990; Davies et al., 1992) indicate that there is more power in intermediate-scale heterogeneity than predicted by global inversions. These differences are attributed to (1) varied effects of the inversion schemes (e.g. Laske, Masters and Zürn, 1994; (2) complexity of the regions where regional tomography is carried out (Passier and Snieder, 1995); (3) filtering effects of finite-frequency propagation on travel times (Gudmundsson and Cummins, 1991).

To understand frequency-dependent effects on the determination of spectral properties we perform full wavefield calculations for mantle models with known spectral properties. The numerical method we use is an extension of the finite-difference (FD) approach by Igel and Weber (1995, 1996) to grids with depth-dependent, lateral grid spacing. Grid refinement near the Earth's surface considerably improves the performance of the FD algorithm because (1) low velocities (small wavelengths) at the top of the mantle can be sampled more efficiently, and (2) higher velocities in the lower mantle are sampled less densely, improving stability.

The purpose of this paper is to present an FD algorithm with depth-dependent, lateral grid spacing for global, toroidal wave propagation, to discuss its accuracy, and to simulate frequency-dependent waveform effects for upper-mantle models with specified spectral properties.

Numerical algorithm

The algorithm presented in this paper is an extension of the high-order FD scheme of Igel and Weber (1995) to grids with vertically varying grid spacing. The equations of motion in spherical coordinates r, θ, φ are solved for wavefields and media invariant in φ , thus rotationally symmetric with respect to the axis $\theta = 0$. For spherical or polar grids with equal spacing in the angular domain, the grid is coarse for large radii and dense near the centre. This is contrary to what is required for global, cartesian Earth models. The velocities near the Earth's surface are considerably lower than deeper within the Earth and therefore require dense rather than coarse sampling. In addition, heterogeneities are clearly strongest in the lithosphere and uppermost mantle (possibly with the exception of the lowermost mantle, D") which consequently should be sampled with a dense grid.

Theory

Assuming invariance in φ the equations for toroidal motion u_φ are (e.g. Lapwood and Usami, 1981)

$$\rho \frac{\partial^2 u_\varphi}{\partial t^2} = f_\varphi + \frac{\partial \sigma_{\varphi r}}{\partial r} + \frac{1}{r} \frac{\partial \sigma_{\varphi \theta}}{\partial \theta} + \frac{3}{r} \sigma_{r\varphi} + \frac{2}{r} \sigma_{\theta\varphi} \cot \theta, \quad (1)$$

where ρ is the mass density, f_φ is a volumetric force, σ_{ij} are the components of the stress tensor,

$$\sigma_{r\varphi} = \mu \frac{\partial u_\varphi}{\partial r}, \quad \sigma_{\theta\varphi} = \mu \frac{\partial u_\varphi}{\partial \theta} \quad (2)$$

μ being the shear modulus.

Discretization

All space dependent fields – stress, sources, and displacement – are defined on a spherical, staggered grid. Staggered grids have been used widely in isotropic (e.g. Virieux, 1984; 1986) and anisotropic (Igel et al., 1995) FD wavefield calculations. In staggered grids, the elements of displacement, stress, and strain are not defined at the same locations, allowing the first derivatives of those fields to be centered in between grid points. Due to the antisymmetry

of the difference operator, this improves the accuracy of the finite-difference calculation for a given space increment.

The spherical grid is shown schematically in Figure 1. The grid is defined between the Earth's surface and the core-mantle boundary (CMB). These two boundaries are modeled as free surfaces. To implement this boundary condition, the grid is extended $L/2 + 1$ points above and below the physical boundary, L being the length of the FD differential (and interpolation) operator.

In the toroidal case with constant grid spacing in both the angular and the radial direction this leads to a consistent scheme without requiring additional interpolations. However, to link the two domains with different lateral grid spacing, interpolations are required at or around the depth of the grid change.

The free-surface boundary condition

$$\sigma_{r\varphi} = \mu \frac{\partial u_{\varphi}}{\partial r} = 0 \text{ at } r = 6371\text{km. and } r = 3480\text{km.} \quad (3)$$

is implemented by imposing symmetry of u_{φ} and antisymmetry of the $\sigma_{r\varphi}$ component with respect to the free surface(s).

As shown in Figure 1, the elements defined at $\theta = 0$ are the displacement u_{φ} and the stress element $\sigma_{r\theta}$. At $\theta = 0$ the equation of motion is singular and the displacement field can not be evaluated. However, by symmetry we impose $u_{\varphi} = \sigma_{r\theta} = 0$. Analogous to the situation at the free surface the grid is extended $L/2 + 1$ points across the boundary $\theta = 0$ and the boundary condition is imposed by anti-symmetry of both stress and displacement field.

The source region near $\theta = 0$ is schematically shown in Figure 2. Our goal is to model point-source-like behavior. To achieve this, sources are input at a grid point close to the axis of symmetry. A directional force would thus result in a ring source (pure toroidal motion). The radiation pattern of such a source is unrealistic. Nevertheless, it allows us to study model-dependent wave phenomena where the source radiation pattern is irrelevant.

The motivation for changing the angular grid spacing was given above. Jastram and Tessmer (1994) and Rodrigues (1993) both introduced FD grids with vertically varying grid spacing, condensing the (cartesian) grid near the surface by a factor of 3. The stability of explicit FD algorithms is generally of the form $c_{max} \frac{dt}{dx} \leq const.$, where c_{max} is the maximum velocity, and dt and dx are time increment and grid spacing, respectively. Depending on the actual velocity model, the time step has to be decreased in accordance

with the change in grid spacing to keep the same level of stability. We find a change of lateral grid spacing by a factor of 2 optimal for global models taking account of computation time.

The partial differentials and the interpolations are calculated by high-order operators of length L . The weights of the operators were obtained following the approach of Holberg (1987). The time evolution is carried out by a Taylor expansion. Details of this type of FD algorithm applied to wave propagation on a 3D cartesian grid can be found in Igel et al. (1995). In all simulations carried out in this paper we use 8-point space operators ($L = 8$) for both derivative and interpolation. The time extrapolation is accurate to fourth order.

Comparison with other methods

Before applying our algorithm to arbitrarily heterogeneous models we compare FD synthetic seismograms with those obtained by an *exact* method (Direct Solution Method (DSM), Geller et al., 1994; Cummins et al., 1994) for spherically symmetric media. The FD grid is divided into two domains with depth ranges 0-320 km and 320-2891 km, with grid sizes 2048×64 and 1028×512 , respectively. The angular domain is $0-\pi$ and the time increment is 0.3 seconds. The model is the isotropic part of PREM (Dziewonski and Anderson, 1981).

In Figure 3 seismograms are compared for two epicentral distances at a period of 20 seconds. The source (toroidal ring source) is at 200 km depth. The overall agreement between the entirely different numerical techniques is excellent. Small phase differences are attributed to the discreteness of the FD grid, which can be suppressed by further grid refinements. The relative amplitude of the phases as well as the waveforms are well modeled by the FD approximation.

How does the grid refinement near the surface improve the performance of the FD scheme? To demonstrate the performance of the suggested algorithm we undertake a convergence test. We compare seismograms from three different simulations: I. Reference seismogram obtained with a very dense grid (5120×1024); II: Seismogram for a 1024×256 grid; III: same as II, but with a 2048×32 refined grid at the top of the model.

The algorithm allows us to generate a synthetic Green's function which

we can later convolve with a source wavelet of the desired frequency band. In Figure 4 seismograms are shown for three different dominant periods at a distance of 60° . At the considered frequencies the coarse grid converges only at a period of 60 seconds while the grid refinement enables the calculation of accurate seismograms down to a period of ≈ 25 seconds. The increase of CPU time for the simulation of type III with respect to type II was 30%. However, this increase leads to a substantial improvement in accuracy.

A further demonstration of the accuracy is given by a comparison with ray-theoretical arrival times (Witte et al., 1996) determined for a 2-D model with lateral perturbations added to the isotropic part of PREM. The lateral S -velocity perturbations are shown in Figure 5a. The perturbations are coherent down to a depth of 250km where they are tapered to zero. Travel times for the S (S_{diff}) phase from the FD seismograms were obtained by maximizing cross-correlation between perturbed and unperturbed seismograms for a period of 30 seconds. Comparison of the travel-time perturbations obtained by these different approaches are shown in Figure 5b. The sampling rate of the FD seismograms is 2/3 seconds and the receiver spacing is 1/3 degrees. The maximum difference between the ray-theoretical and FD travel time is ≈ 1 second.

Since the direct S (S_{diff}) phase is arriving at a very steep angle, we expect anti-correlation between velocity and travel-time perturbation, which can be appreciated by comparing Figures 5a and 5b.

The accuracy of our algorithm as far as waveform and travel times are concerned gives us confidence that we can apply our FD method to study finite-frequency amplitude and waveform effects.

Spectra of mantle models

The aim of this study is to investigate the frequency dependence of travel times and waveforms of S and SS waves for upper-mantle models with varied power spectra and the implications for the scale of mantle heterogeneity. A schematic representation of the power spectrum of mantle velocity perturbations is shown in Figure 6. The two types of spectra shown in this Figure (type I and type II) represent extreme cases of Passier and Snieder (1995). They showed that at intermediate scales ($l=30$) the spectra obtained by global and regional inversions differ by as much as a factor of 10-30. One of

their arguments for the presence of significant intermediate-scale heterogeneity is the complexity of long-period surface-wave spectra due to multipathing and the associated interference effects.

We generate models with type I and II statistics shown in Figure 6. Details of the model generation are given in the Appendix. Figure 7 shows 5 random model realizations of each spectral type. Clearly spectral type I contains mainly continent-scale features while type II contains a significant amount of energy at scales down to ≈ 100 km wavelength.

To test the spectral properties of the synthetic models we average the power spectra of several realizations for each type. Numerical tests show that at least 5 power spectra of different realizations have to be summed to achieve an accuracy of 5% for the decay rate. Results are shown in Figure 8. The power spectra display the characteristic corner frequency and the spectral decay which was imposed in the generation of the models.

2-D models of S -velocity perturbations are obtained by superimposing several lateral 1-D functions weighted by depth-dependent orthogonal functions. Details about the 2-D model generation are given in the Appendix. Sections of models thus obtained are shown in Figure 9. The S -velocity perturbations for all the simulations described below are tapered down to zero between 500km and 600km depth. There are no lateral velocity perturbations below 600km depth. It is important to note that the 2-D perturbations vanish near the axis of rotational symmetry $\theta = 0$ to avoid focusing effects. The perturbations are imposed for $\theta > 10^\circ$.

In the following we investigate synthetic seismograms obtained for the model types just described. All models have 2-D S velocity perturbations according to spectral types I or II from Figure 6. The maximum amplitude of the S -velocity perturbations varies between 0% and 12% (root-mean-square variation between 0 and 3.2%).

Synthetic seismograms

We calculate synthetic seismograms for perturbations of types I and II added to the isotropic part of PREM with the following setup. The mantle is sampled in depth by 578 grid points, leading to a depth spacing of 5.0km. The upper part of the mantle (0-320km depth) is defined on a grid of size 2560×64 , while the bottom part (320-2891km depth) is defined on a 1280×512 grid. We focus on S and SS phases, which are frequently used in global tomographic

studies. Seismograms of length 1 hour (12000 time steps with time increment $dt = 0.3s$) are calculated for epicentral distances $50-150^\circ$ with a receiver spacing of $1/3^\circ$. Windows containing S and SS arrivals are extracted using ray-theoretical travel times. A simulation with the parameters described above takes ≈ 15 minutes CPU time on a 128-processor CM5.

Seismograms for S_{diff} at a receiver sampling of 1 degree are shown in Figure 10 for a dominant period of 25 seconds for epicentral distances $100-150^\circ$. In all simulations the wavelet is the first derivative of a Gaussian. The seismograms shown in Figure 10 were obtained for a 2-D realization of spectral types I (top) and II (bottom) for different maximum perturbation amplitudes. In the context of seismic-velocity anomalies and their relation to tectonic features the small and large perturbation amplitudes may represent regions of little (e.g. cratons) and high (e.g. island-arc regions) tectonic activity.

Concerning the S_{diff} -phase we can make the following observations: (1) For model type I the S waveform is essentially undisturbed even for large perturbations; (2) for model type II the waveform begins to be affected by the scattering for perturbation amplitudes greater than 4%.

From the same set of seismograms we extract time windows containing the SS phase. These are shown in Figure 11. For model type I similar observations can be made as for the S phase. The waveform - in the spherically symmetric case the Hilbert transform of the S phase (Choy and Richards, 1975) - is hardly affected by the velocity perturbations even at large perturbation amplitudes (12%). However, the SS -waveforms of model type II are severely distorted for perturbations larger than 4%. It is worth noting that most of the SS coda in Figure 11 (bottom) is precursory.

How does the waveform distortion depend on the dominant period of the wavelet used? For long-period wavelets (75 seconds) the situation is quite different (see Figure 12). Although there still is a considerable difference between the two model types, the waveforms for model type II are much less distorted than at a dominant period of 25 seconds. We try to quantify this frequency dependent SS -waveform distortion in the next section.

Frequency-dependent effects

To describe the frequency-dependent scattering effect we process the SS windows in the following way: (1) Find the time delay τ of the perturbed

phase with respect to a reference (PREM) wavelet. (2) Correct the time delay and calculate a scattering index $s(T)$ (normalized root-mean-square displacement misfit in a time window of length $2T$, T being the dominant period) according to

$$s(T) = \left[\frac{1}{N} \sum_{\text{receivers}} \frac{1}{2T} \sum_{\text{time}} \frac{(u_{\text{pert}}(T, t - \tau) - u_{\text{ref}}(T, t))^2}{u_{\text{ref}}^2(T, t)} \right]^{1/2}. \quad (4)$$

This processing is carried out for the seismograms shown in Figures 11 and 12, for periods from 20-110 seconds and 50 receivers between 50° and 150° epicentral distance. The results are shown in Figure 13. For spectral type I (Figure 13, left) even for large model perturbations (12%) the average waveform distortion does not exceed 50%. In the case of spectral type II (Figure 13, right) there is a rapid increase in waveform distortion at a period of ≈ 30 seconds. The convergence of the perturbed and unperturbed waveforms is more rapid for models of type I.

We now look at the frequency dependence of the spectrum of SS travel-time perturbations measured at the surface. The model considered is of spectral type II (see Figure 6). The maximum perturbation amplitude is 6%, which is a likely underestimate of the strength of heterogeneity in the upper mantle judged by results of surface-wave tomography (e.g. Zielhuis and Nolet, 1991).

The dense and regular receiver sampling (1/3° between 50° and 150°) allows us to achieve high accuracy in the travel-time spectra at intermediate scales ($l \leq 300$). The travel times are determined by maximizing the cross-correlation between perturbed and unperturbed SS waveforms. Travel times are picked from seismograms of three different dominant periods (25, 50, and 100 seconds). Examples are shown in Figure 14. While there is overall a good correlation between the travel-time residuals, they tend to be larger and show more details at shorter periods.

To obtain stable results for the travel-time spectra we sum power spectra from 5 different realizations for each of the three frequency bands considered. The decay rate of the stacked power spectra is determined by fitting a line through the spectrum between harmonic degrees $l = 20$ and $l = 300$. The results shown in Figure 15 demonstrate a clear frequency dependence. At high frequencies (dominant period 25 seconds), the travel time residuals reflect well the properties of the spectrum of the underlying velocity perturbations. However, already at 50 seconds dominant period the slope of the

decay is nearly 2 powers different from the one in the model (see Figure 8). Furthermore, the spectral corner is shifted to lower harmonic degrees, which is even more significant at a dominant period of 100 seconds.

At $l = 100$ ($\approx 400\text{km}$ wavelength at the Earth's surface), the spectral power determined at periods 25 and 50 seconds varies by a factor of ≈ 10 . This clearly indicates that for the model considered in this example, the spectra of SS travel-time residuals reflect the actual spectral properties of the upper mantle only at periods close to or smaller than 25 seconds.

Discussion and Conclusions

The goal of this study was to analyze full-wavefield, synthetic seismograms for random upper-mantle models with specified spectral properties and to investigate frequency-dependent effects on waveforms and travel times of S and SS waves.

The synthetic study we have undertaken is statistical in nature and requires a number of wave simulations to be carried out in order to achieve a stable estimate of spectra. Therefore, an efficient forward technique allowing us to model complete seismograms for arbitrary homogeneous structures was necessary. Accurate 3-D global simulations without assumptions on the amplitude of perturbations are still too expensive computationally. We compromise in using a FD approximation to the wave equation in spherical coordinates where all fields are invariant in φ . This leads to a 2-D problem which can be solved efficiently on present-day parallel computers.

We extended the FD algorithm suggested by Igel and Weber (1995) to grids with vertically-varying, lateral grid spacing. The grid refinement near the Earth's surface is necessary from an algorithmic point of view (the arc length of the angular grid increment increases with distance from the center of the sphere) and from velocity model considerations (low velocities in the crust and lithosphere require smaller grid spacing). The change in grid spacing has been implemented using a high-order approximation. The relatively thin layer ($\leq 200\text{km}$) with refined grid spacing leads to an additional 30% in computation time (as compared to a simulation without refinement). However, to achieve the same accuracy without grid refinement, simulations about four times longer would have been necessary. This is crucial for this study where the investigation of random structures and their spectral prop-

erties requires many simulations to be carried out.

The scale of seismic-velocity perturbations in the Earth's mantle is important in many aspects of geodynamics (e.g. mantle convection, subduction, hot spots, etc). In several studies, conclusions about mantle convection were drawn from power spectra of global tomographic models (e.g. Montagner and Tanimoto, 1991; Montagner 1991) or even directly from the power spectra of SS -traveltime residuals (e.g. Su and Dziewonski, 1991, 1992). In a recent study Passier and Snieder (1995) highlighted the discrepancy between the scales of velocity perturbations obtained from global and regional tomography.

Our synthetic modeling suggests that - as far as S , SS and $SS-S$ studies are concerned - the discrepancies reported by Passier and Snieder (1995) are at least in part due to finite-frequency effects of wave propagation. If the travel-time perturbations observed at the Earth's surface do not reflect the scale of the underlying mantle, it is impossible for delay-time tomography to model the power spectrum of mantle heterogeneity correctly. Similarly, if phase-velocities measured along the Earth's surface do not reflect the details of the structure of the mantle beneath, that detail cannot be recovered by inversion. Waveforms and amplitudes are necessary to model Earth structure in detail.

The results shown in Figure 13 suggest that the frequency dependent effects on the waveform are sensitive to the scales and amplitudes of perturbations present in the upper mantle. This points to possible processing techniques which should be used in combination with travel time processing when estimating the power spectrum of mantle velocity perturbations. Gudmundsson (1996) studied the effects of frequency on waveform distortion and travel times. He concluded that - depending on the bandwidth - waveform distortion could be minor while diffraction effects on timing are significant. His results and our synthetic modeling suggest that even though waveform distortion may be significant, it may not be clear from a single seismogram (e.g. Figure 11).

The power spectra in Figure 15 demonstrate that the characteristics of the power spectra of SS travel-time residuals are strongly frequency dependent. The slope of decay extracted from the power spectrum varies with the frequency band at which the power spectrum was determined. For example, at around $l = 100$ (≈ 400 km wavelength), the power spectrum determined at 25 seconds dominant period contains around 10 times more energy than the

one determined at a dominant period of 50 seconds. This implies that conclusions on the scale of mantle heterogeneities from long-period seismograms may be erroneous if these finite-frequency effects are not taken into account.

Our results suggest that the travel-time residual spectra determined by Su and Dziewonski (1992) do not reflect the scale of mantle heterogeneities, in part at least due to filtering effects of finite-frequency propagation, and that the spectra they estimated are compatible with mantle models with considerably more energy at intermediate scales.

It is important to note that - since our models are invariant in φ and thus two dimensional - the effects we described in this paper are conservatively described. For 3-D models with the same spectral properties we expect the filtering effects to be even more pronounced.

Acknowledgements

We are grateful to the Centre National de Calcul Parallel en Science de la Terre (CNCPST) at the Institut de Physique du Globe de Paris for giving us access to the Connection Machine CM 5. One of the authors (HI) is funded by the NERC Grant GR3/10086. We are grateful to Phil Cummins for calculating the DSM seismograms and to Michael Roth for the Huygens arrival times. We thank G. Müller, F. Pollitz and Adam Schultz for helpful comments.

Appendix: Generation of velocity perturbations

We generate 1-D models with power spectral properties shown in Figure 6. Using the relation between wavenumber k and harmonic order l $kR = l + 1/2$, R being the radius, we describe the amplitude spectrum as a function of harmonic order. Perturbations are then calculated according to the defined spectrum as follows

- (1) generate a random (white) phase spectrum $[-\pi, \pi]$
- (2) modulate complex phase spectrum with predefined amplitude spectrum to define the Fourier spectrum
- (3) inverse FFT to obtain space domain representation
- (4) scale to desired maximum perturbation.
- (5) modulate perturbation with a prescribed function of depth.

Choices of amplitude (power) spectra and the depth modulation are described in the main body of the paper. Examples of perturbations calculated in this way are shown in Figure 7. (The analogy in 3D spherical coordinates would be to predefine the spherical-harmonic amplitude spectrum and distribute the power within each degree randomly among all its orders).

To make the models more realistic, incoherence in depth is introduced. Lateral perturbations are weighted with the set of depth dependent functions

$$\begin{aligned} f_n(z) &= 1, \quad n = 0, \\ f_n(z) &= \sin\left(\frac{n\pi z}{h}\right), \quad n = 1, 3, \dots, 2m + 1, \\ f_n(z) &= \cos\left(\frac{n\pi z}{h}\right), \quad n = 2, 4, \dots, 2m, \end{aligned} \quad (5)$$

and 2-D velocity perturbations $\Delta v_s(z, \theta)$ are obtained as

$$\Delta v_s(z, \theta) = \sum_{n=0}^m w_n g_n(\theta) f_n(z), \quad (6)$$

where $g_n(\theta)$ are the lateral perturbations, and w_n are weights. Examples of 2-D perturbations obtained for the spectra in Figure 6 are shown in Figure 9.

We use $m = 5$ for all 2-D models. For spectral type I the weights in equation (5) are $w_1 = w_2 = 1.0$, $w_3 = w_4 = 0.8$, $w_4 = w_5 = 0.4$, and for spectral type II all weights are unity.

References

- Choy, G.L., and P.G. Richards, Pulse distortion and Hilbert transformation in multiply reflected and refracted body waves, *Bull. Seis. Soc. Am.*, 65, 55-70.
- Cummins, P., R.J. Geller, T. Hatori, N. Takeuchi, DSM complete synthetic seismograms: SH, spherically symmetric case, *Geophys. Res. Let.*, 21, 1663-1666, 1994a.
- Cummins, P., R.J. Geller, and N. Takeuchi, DSM complete synthetic seismograms: P-SV, spherically symmetric case, *Geophys. Res. Let.*, 21, 1663-1666, 1994b.
- Davies, J.H., O. Gudmundsson, R.W. Clayton, Spectra of mantle shear velocity structure. *Geophys. J. Int.*, 108, 865-882, 1992.
- Dziewonski, A., and D.L. Anderson, Preliminary reference Earth model, *Phys. Earth Planet. Int.*, 25, 297-356, 1981.
- Geller, R.J., and T. Ohminato, Computation of synthetic seismograms and their partial derivatives for heterogeneous media with arbitrary natural boundary conditions using the Direct Solution Method. *Geophys. J. Int.*, 116, 421-446, 1994.
- Gudmundsson, O., On the effect of diffraction on traveltimes measurements, *Geophys. J. Int.*, 124, 304-314, 1996.
- Gudmundsson, O., J.H. Davies, R.W. Clayton, Stochastic analysis of global traveltimes data: mantle heterogeneity and random errors in the ISC data. *Geophys. J. Int.*, 102, 25-43, 1990.
- Gudmundsson, O., P. Cummins, The importance of small-scale structure in the earth's upper mantle. AGU Fall Meeting 1991. Technical Programme, 1994.
- Holberg, O., Computational aspects of the choice of operator and sampling interval for numerical differentiation in large scale simulation of wave phenomena, *Geoph. Prosp.*, 35, 629-655, 1987.
- Igel, H., and M. Weber, SH-wave propagation in the whole mantle using high-order finite differences, *Geophys. Res. Let.*, 22, 731-734, 1995.
- Igel, H., M. Weber, P-SV wave propagation in the Earth's mantle using finite differences: Application to heterogeneous lowermost mantle structure, *Geophys. Res. Let.*, 23, 415-418, 1996.
- Igel, H., P. Mora, and B. Rioulet, Anisotropic wave propagation through FD grids, *Geophysics*, 60, 1203-1216, 1995.

- Jastram, C., E. Tessmer, Elastic modeling on a grid with vertically varying spacing, *Geoph. Prosp.*, 42, 357-370, 1994.
- Laske, G., G. Masters, W. Zürn, Frequency-dependent polarization anomalies of long-period surface waves and their implications for global phase-velocity maps, *Phys. Earth Planct. Inter.*, 84, 111-137, 1994.
- Lapwood, E.R., T. Usami, Free oscillations of the Earth, Cambridge University Press, 1981.
- Montagner, J.-P., Can seismology tell us anything about convection in the mantle?, *Rev. Geophys.*, 32, 115-138, 1994.
- Montagner, J.-P., T. Tanimoto, Global upper mantle tomography of seismic velocities and anisotropies, *J. Geophys. Res.*, 96, 20337-20351, 1991.
- Passier M.L., R. Snieder, On the presence of intermediate-scale heterogeneity in the upper mantle, *Geophys. J. Int.*, 123, 817-837, 1995.
- Pollitz, F.F., Global tomography from Rayleigh and Love wave dispersion: effect of ray-path bending, *Geophys. J. Int.*, 118, 730-758, 1994.
- Ritzwoller, M.H., E.M. Lavelly, Three-dimensional seismic models of the Earth's mantle, *Rev. Geophys.*, 33, 1-66, 1995.
- Rodrigues, D., Simulation de la propagation d'ondes sismiques sur machine massivement parallèle, Thèse de l'Ecole Central de Paris, 1993.
- Snieder, R., Large scale waveform inversions of surface waves for lateral heterogeneity: 2. Application to surface waves in Europe and the Mediterranean, *J. Geophys. Res.*, 93, 12067-12080, 1988.
- Van der Lee, R., Van der Hilst, Traveltime tomography of the European-Mediterranean mantle down to 1400km *Phys. Earth Planct. Inter.*, 79, 3-74, 1993.
- Su, W.-J., A. Dziewonski, Predominance of long-wavelength heterogeneity in the mantle, *Nature*, 352, 121-126, 1991.
- Su, W.-J., A. Dziewonski, On the scale of mantle heterogeneity, *Phys. Earth Planct. Inter.*, 74, 29-54, 1992.
- Virieux, J., SH-wave propagation in heterogeneous media: velocity-stress finite-difference method, *Geophysics*, 49, 1933-1942, 1984.
- Virieux, J., P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method, *Geophysics*, 51, 889-901, 1986.
- Witte, O., M. Roth, G. Müller, Ray tracing in random media, *Geophys. J. Int.*, 124, 159-169, 1996.
- Zielhuis, A., G. Nolet, Shear-velocity variations in the upper mantle beneath central Europe, 117, 695-715, 1991.

Zhang, Y.-S., T. Tanimoto, Global love wave phase velocity variation and its significance to plate tectonics, *Phys. Earth Planet. Inter.*, 66, 160-202 1991.

Zhang, Y.-S., T. Tanimoto, High resolution global upper mantle structure and plate tectonics, *Geophys. J. Int.*, 98, 9793-9823.

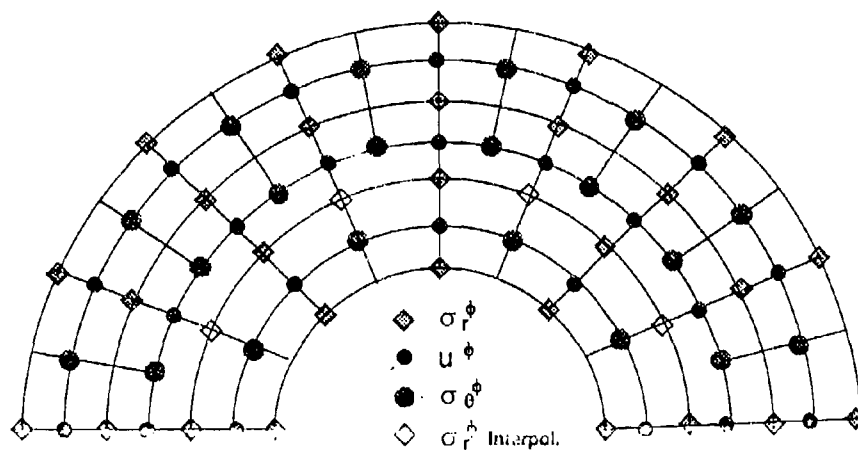


Figure 1: Staggered grid used in the FD scheme. At the top of the model the angular grid spacing $\Delta\theta$ is increased by a factor of 2. The radial grid spacing Δr is constant throughout the grid. The Earth's surface and the CMB are modeled with a free-surface boundary condition. At the level of changing angular grid spacing, $L/2$ levels of interpolated $\sigma_{r\theta}$ components have to be added to connect the two domains (L being the length of the space operator).

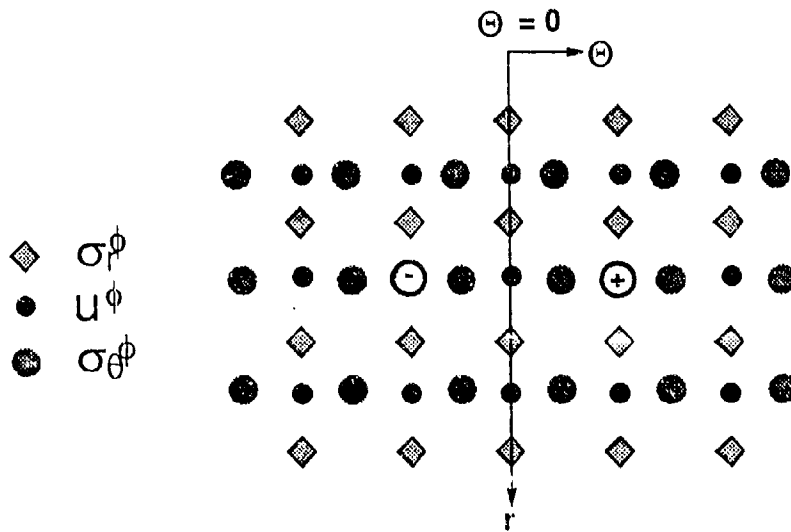


Figure 2: The staggered FD grid close to the axis $\theta = 0$. The equations of motion are singular at $\theta = 0$. At the axis of symmetry $u_z = 0$ and $\sigma_{rz} = 0$ are imposed by symmetry. Note that a directional force input at a grid point (+) next to the axis of symmetry leads to a toroidal ring source. The grid is extended $L/2 + 1$ levels beyond $\theta = 0$, where L is the length of the space operator, to enforce the boundary condition.

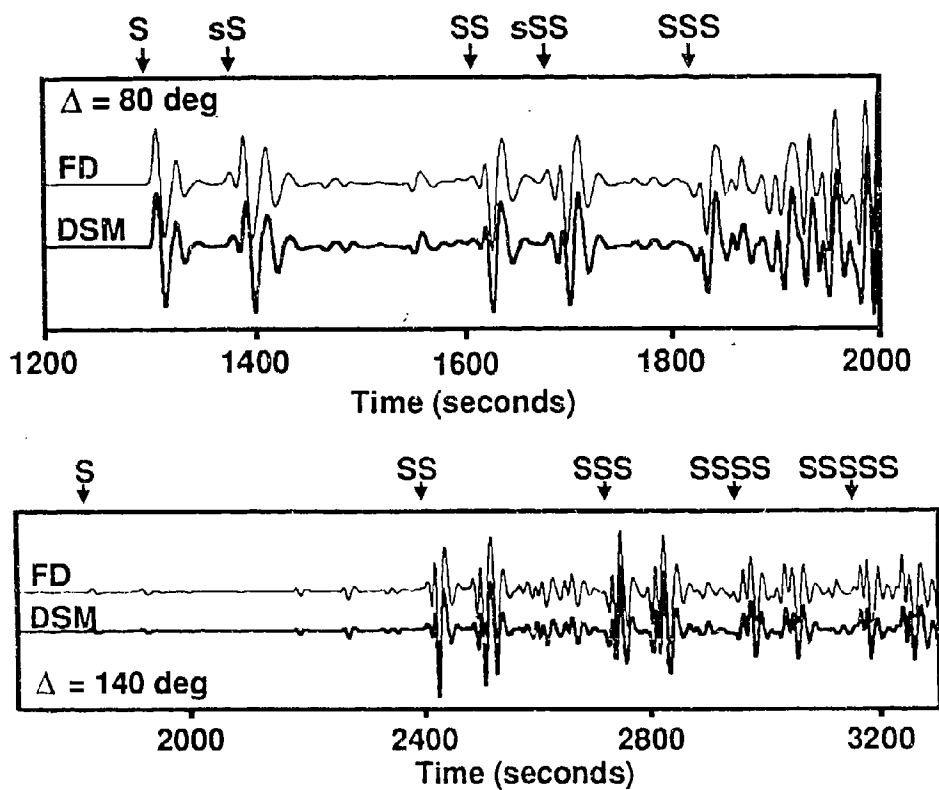


Figure 3: Comparison of DSM (bold) and FD seismograms for the isotropic part of PREM. The dominant period is 20 seconds. Some arrival times of S phases are marked by arrows. Top: Epicentral distance 80 degrees. Bottom: 140 degrees.

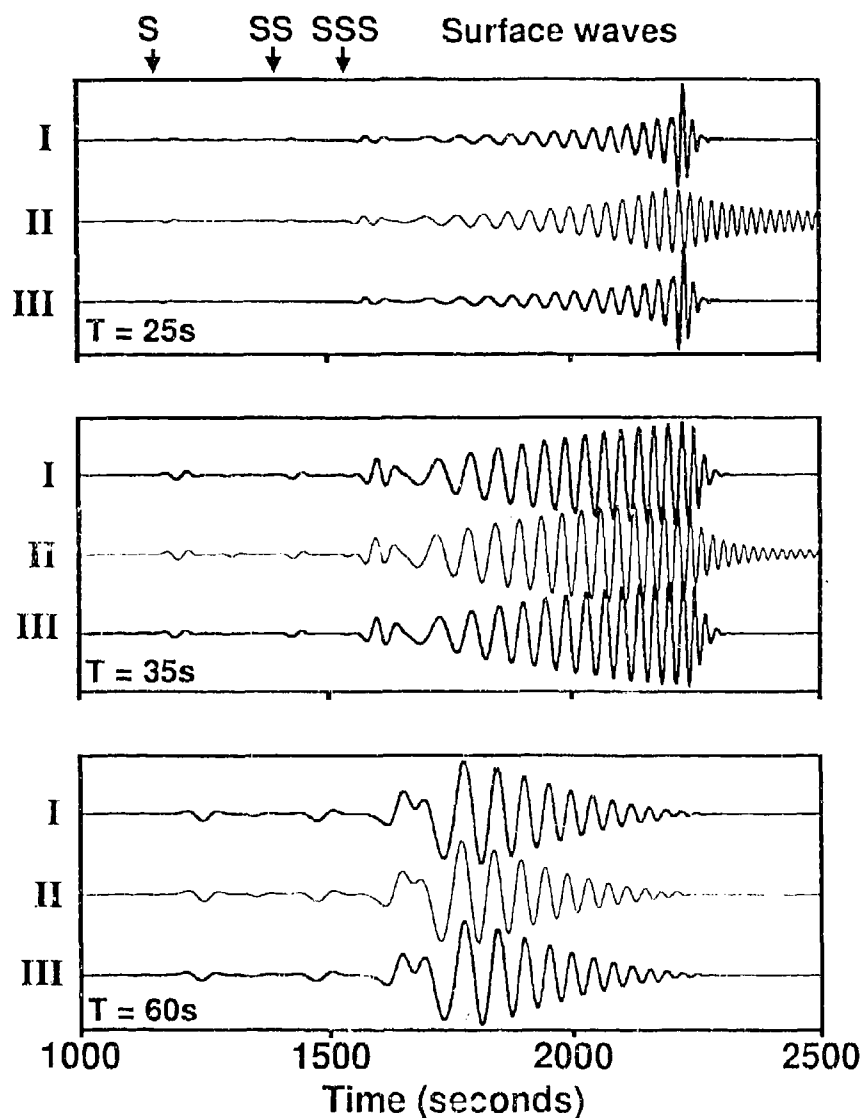


Figure 4: The effect of the change in grid spacing at the top of the model for different frequencies (25, 35, and 60 seconds dominant period). The source is at the surface. The model is the isotropic part of PREM. Arrival times of some phases are marked by arrows. For each frequency band three seismograms at an epicentral distance of 80° are shown: I. Reference seismogram obtained with a very dense grid (5120×1024); II: Seismogram for a 1024×256 grid; III: same as II, but with a 2048×32 refined grid at the top of the model. The accuracy of the surface waves is considerably improved by the grid refinement.

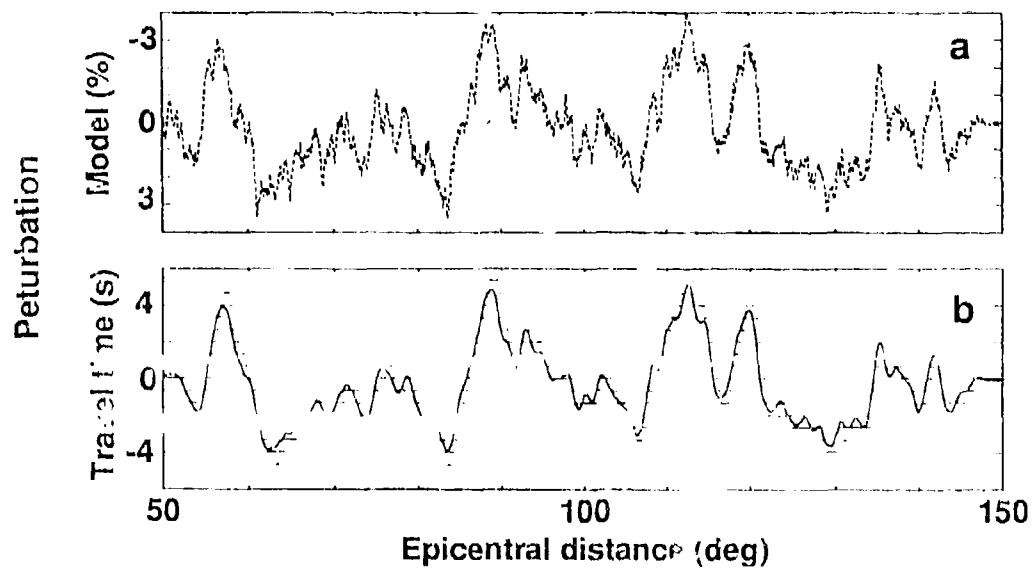


Figure 5: Top: Lateral velocity perturbations which were added to PREM at the top of the mantle. The perturbations are coherent down to a depth of 250 km. Bottom: Travel-time perturbations for S (S_{diff}) obtained by the Huygens method (solid line, Witte et al., 1995), and those obtained with the FD method (dots). The sampling of the FD seismograms was 2/3 seconds.

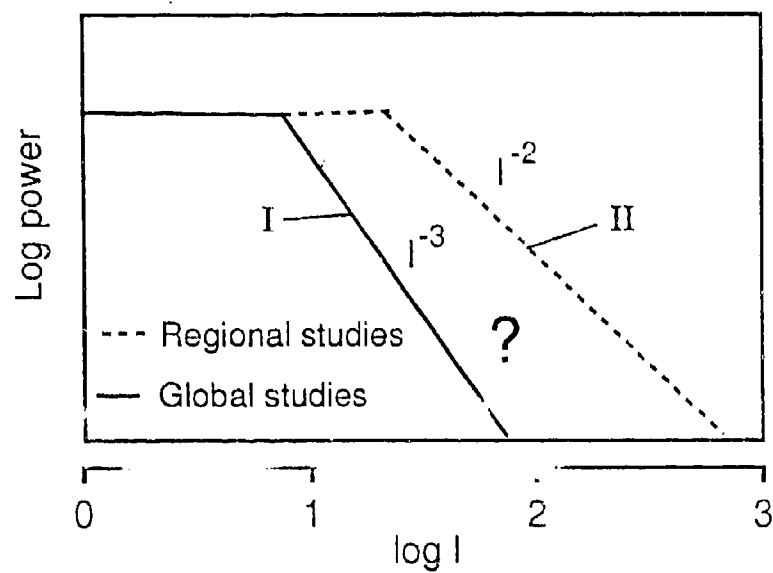


Figure 6: Power spectra of lateral velocity perturbations suggested by Passier and Snieder (1995) as representative of global tomographic models (type I, solid) and regional tomographic (and other) models (type II, dashed).

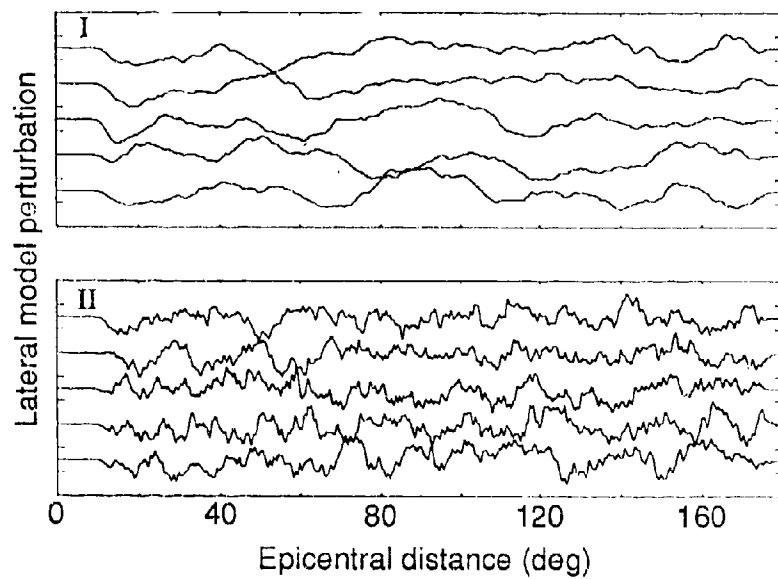


Figure 7: Realizations of models of lateral shear-velocity perturbations for the two types of spectra shown in the previous Figure.

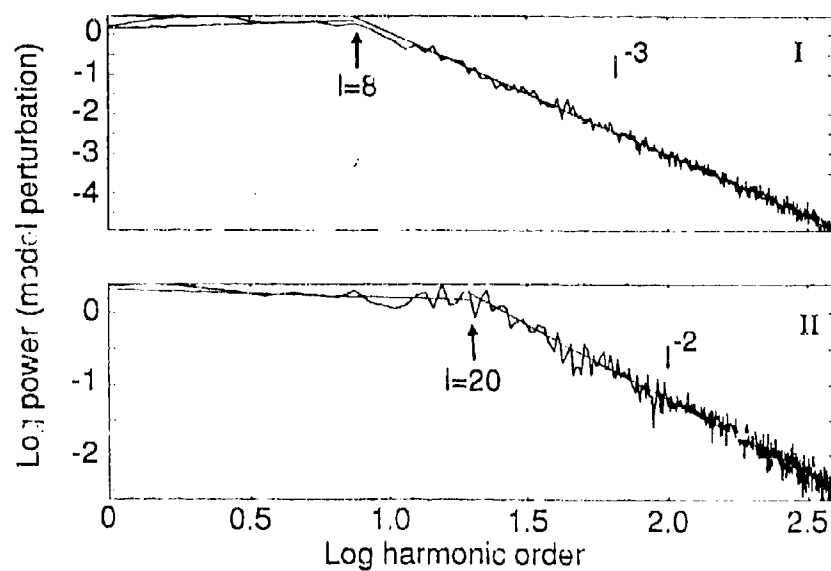


Figure 8: Power spectra obtained by stacking five spectra of the realizations shown in the previous Figure. The corner of spectral decay and the decay rate are well matched (the error of the slope is $\leq 5\%$).

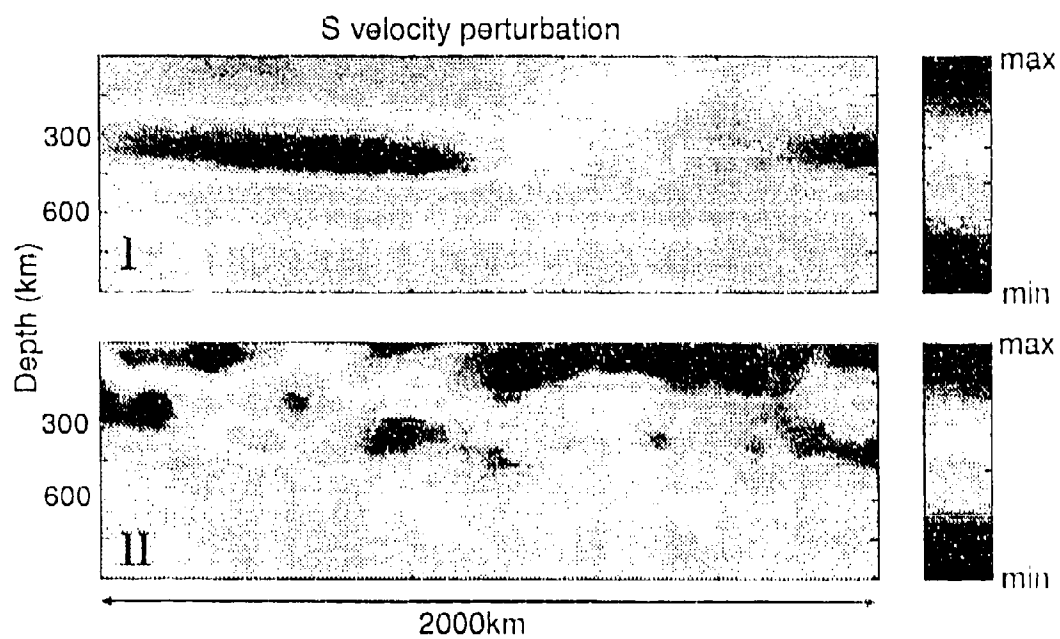


Figure 9. Two-dimensional models for the two types of power spectra. The perturbations are limited to the top 500 km. Top: type I; Bottom: type II. See text for details about the depth dependence.

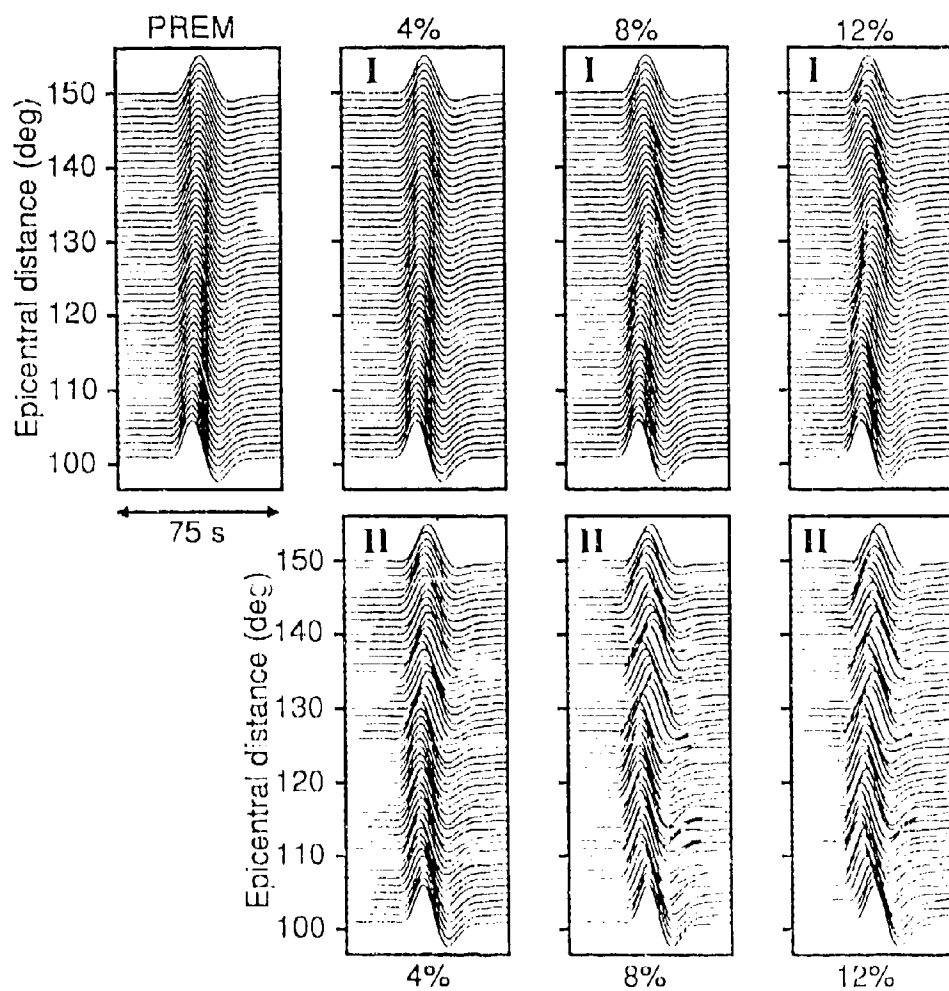


Figure 10: FD seismograms for the S_{440} phase (dominant period 25 seconds) and the PREM model (top left), with heterogeneity according to spectral types I and II superimposed (right). The maximum perturbation is 0% (PREM), 4%, 8%, and 12%. Part of the 2-D models are shown in the previous Figure.

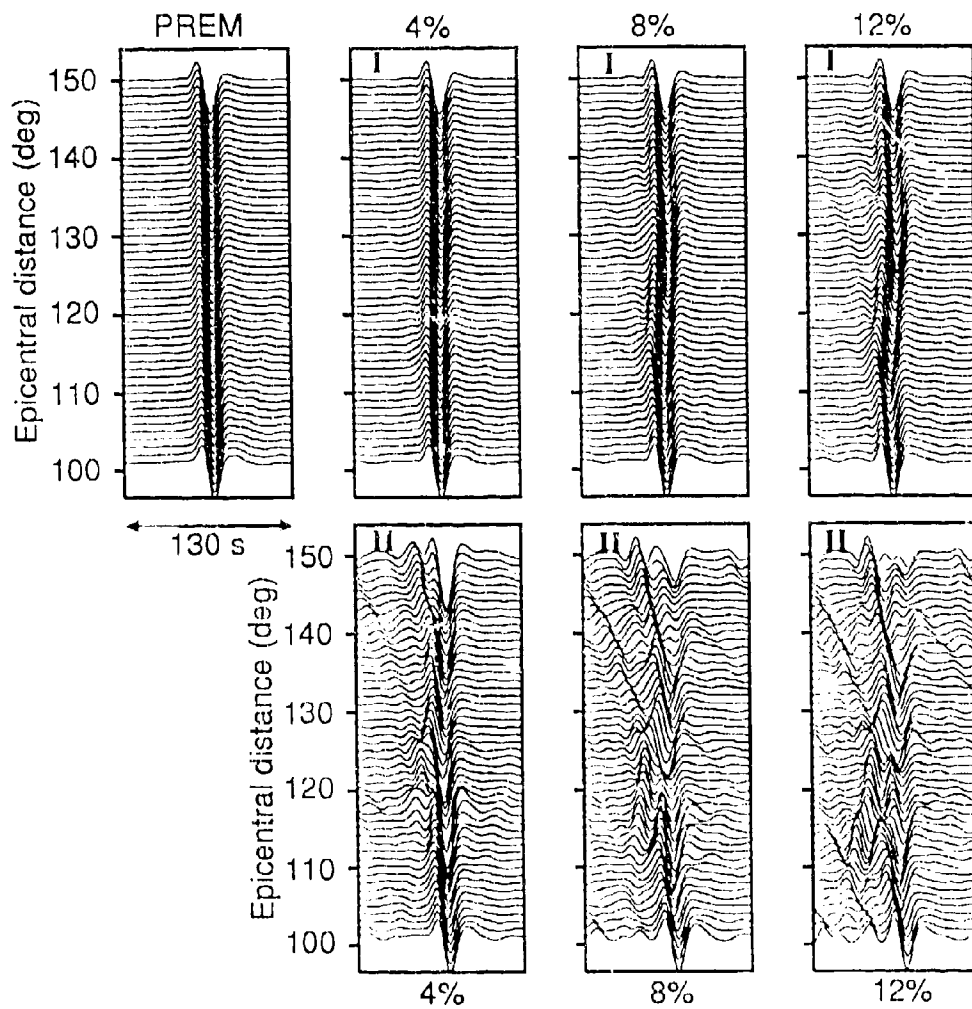


Figure 11: FD seismograms for the SS phase (dominant period 25 seconds) and the PREM model (top left), with heterogeneity according to spectral types I and II, superimposed (right). Same models as in previous Figure.

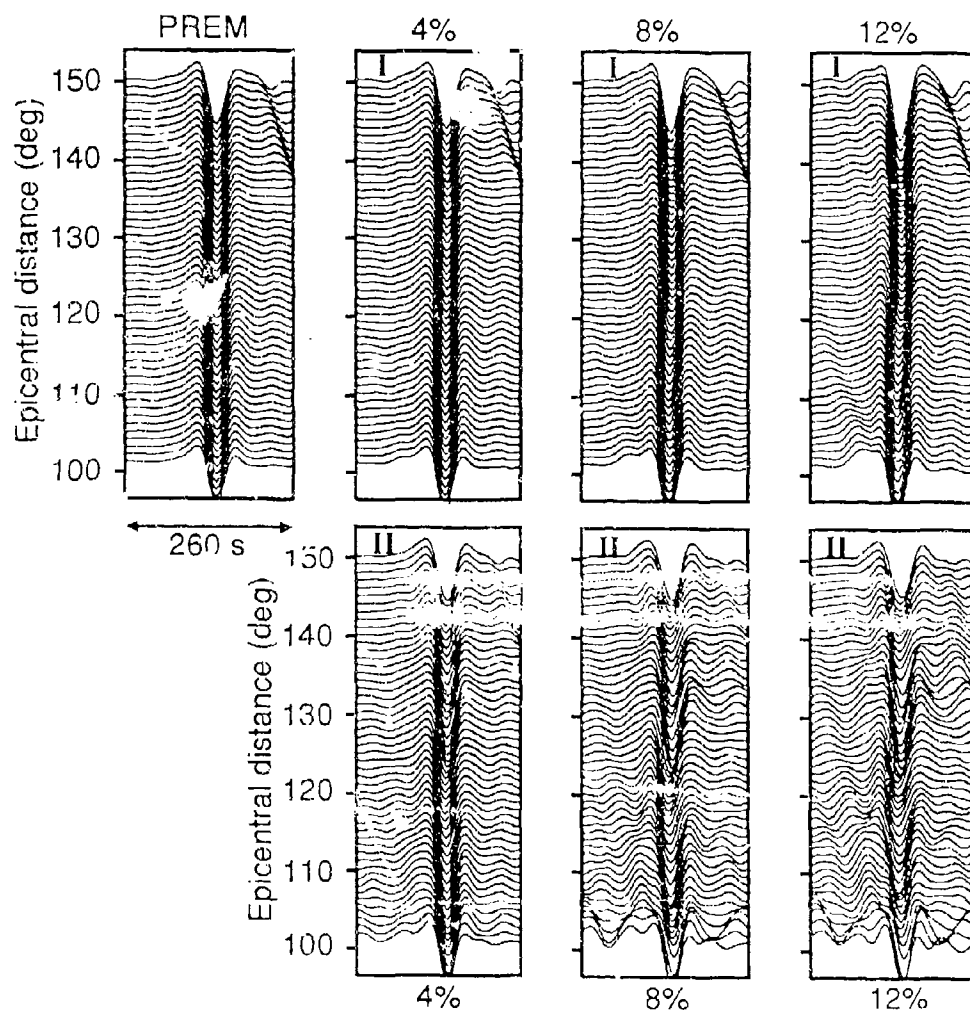


Figure 12. FD seismograms for the SS phase (dominant period 75 seconds) and the PREM model (top left), with heterogeneity according to spectral types I and II, superimposed (right). Same models as in previous figure.

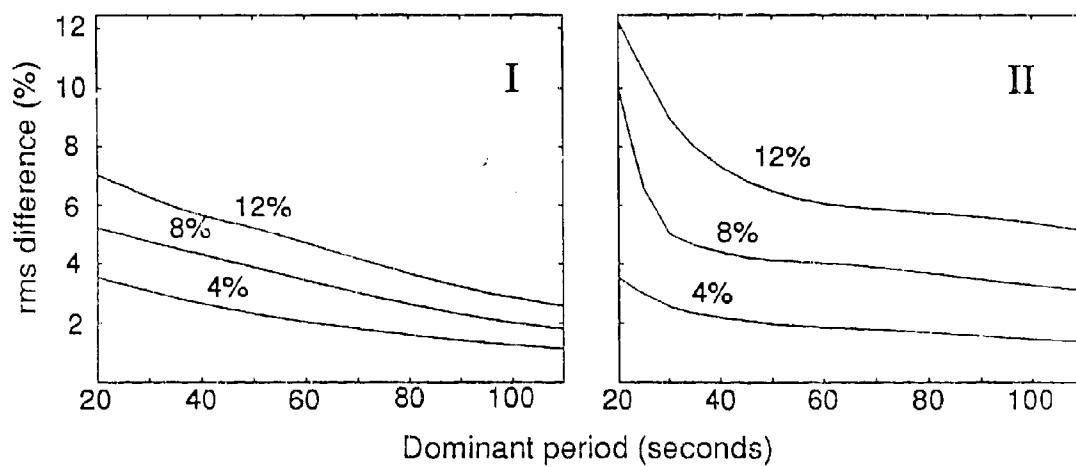


Figure 13: The scattering index $s(T)$ (defined in the text) as a function of dominant period. Results are shown for the two types (I and II) of spectra and different maximum perturbation amplitude (4, 8, and 12%). Spectral type II is characterized by a rapid increase of waveform distortion at a dominant period of ≈ 30 seconds.

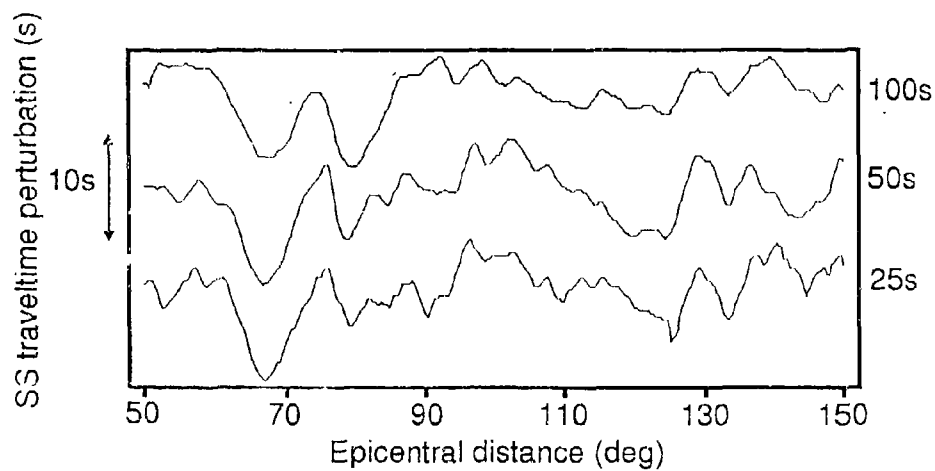


Figure 14: SS travel-time perturbations determined from seismograms of varied dominant period (25, 50, and 100 seconds) between 50 and 150 degrees epicentral distance. The receiver spacing is $1/3$ degrees.

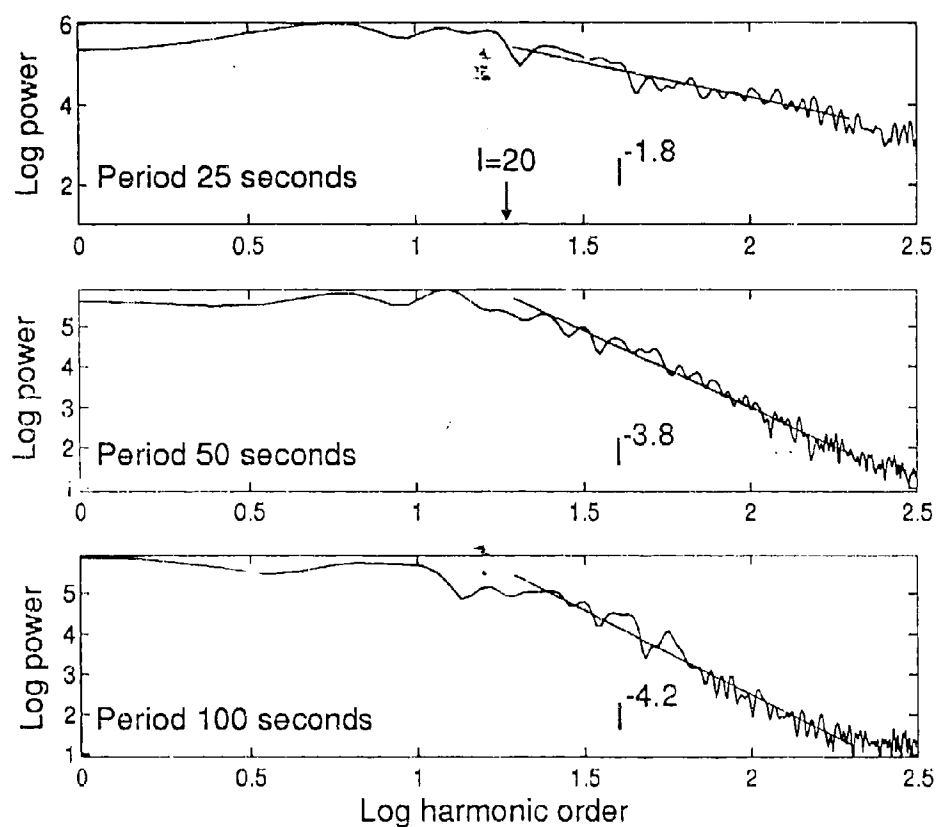


Figure 15: Power spectra of travel-time residuals determined in different frequency bands. In each case spectra for 5 random realizations using Earth models with identical characteristics are stacked. The model spectra are as in Figure 8 (lower). The maximum model perturbation is 6%. The spectral decay determined from *SS* residuals is clearly frequency dependent.

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

by C. Tong

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, c_i for the i th 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,

$$\dots, F_i(t_i, \omega_i, \nu_i, \phi_i, \psi_i), F_{i+1}(t_{i+1}, \omega_{i+1}, \nu_{i+1}, \phi_{i+1}, \psi_{i+1}), \dots \quad (6.1)$$

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 81° the major observed phases are P and S , but beyond 81° 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*

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

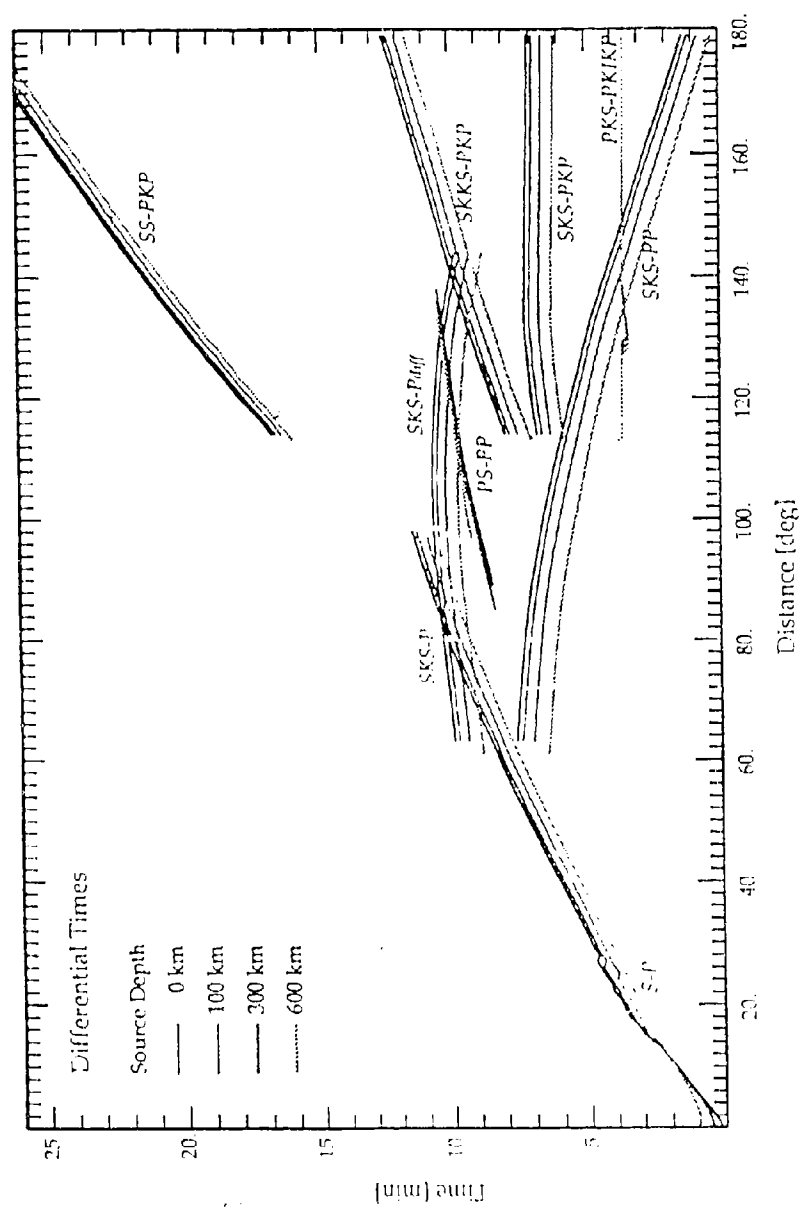


Fig. 6.1. The distance and depth dependence of P - S differential times.

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 x is tested with respect to a hyperplane in model space. The branch is then determined by which side of the hyperplane 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

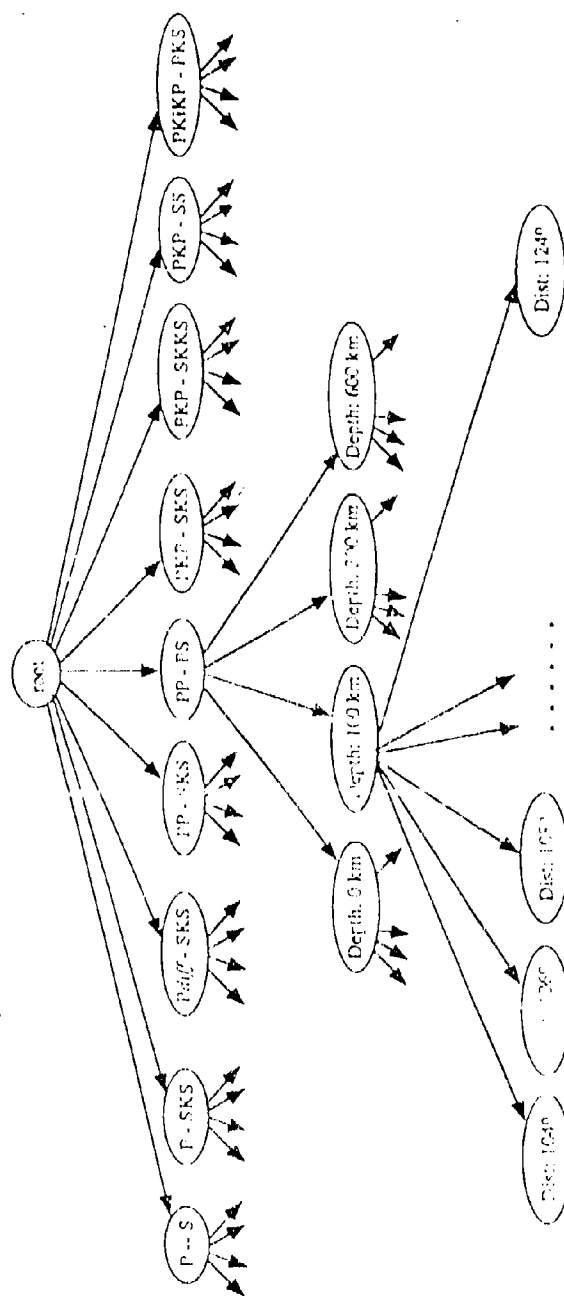


Fig. 6.2. Seismic event classification hierarchy.

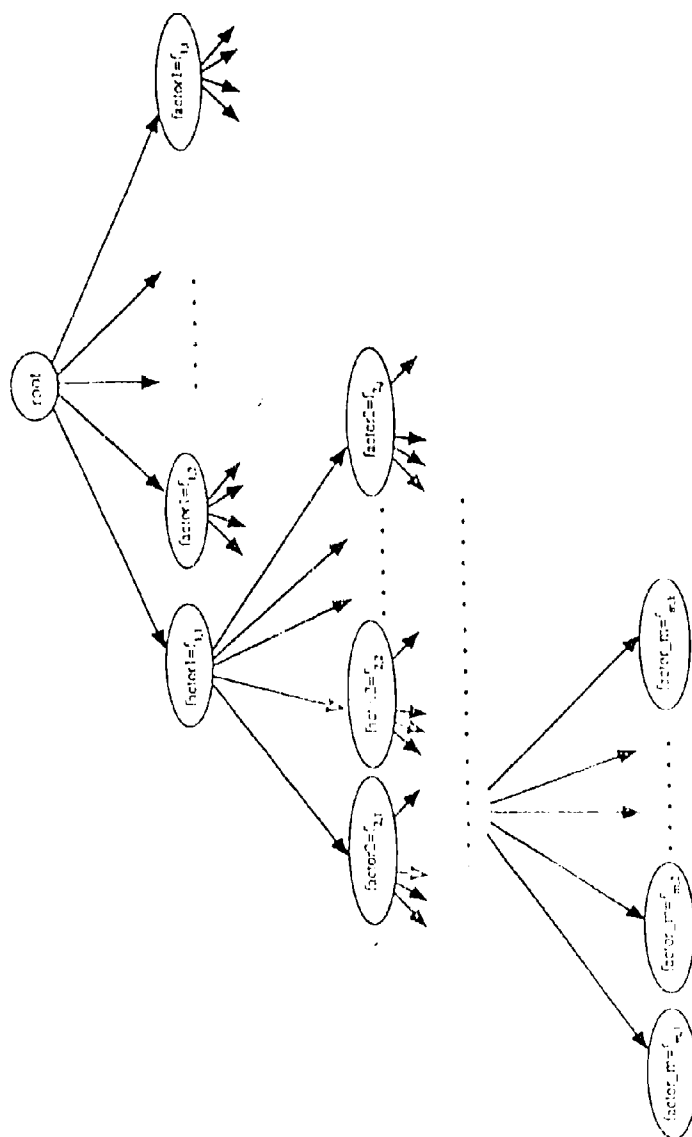


Fig. 6.3. The structure of an assumption tree.

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:

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

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 $0^\circ - 180^\circ$ 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*

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

$$e_i = \frac{|t_o - t_e|}{t_s - t_p}, \quad (6.2)$$

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

$$\delta = \frac{\sum_{i=1}^n e_i w_i}{\sum_{i=1}^n w_i}, \quad (6.3)$$

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 Guralp 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 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 *L* 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

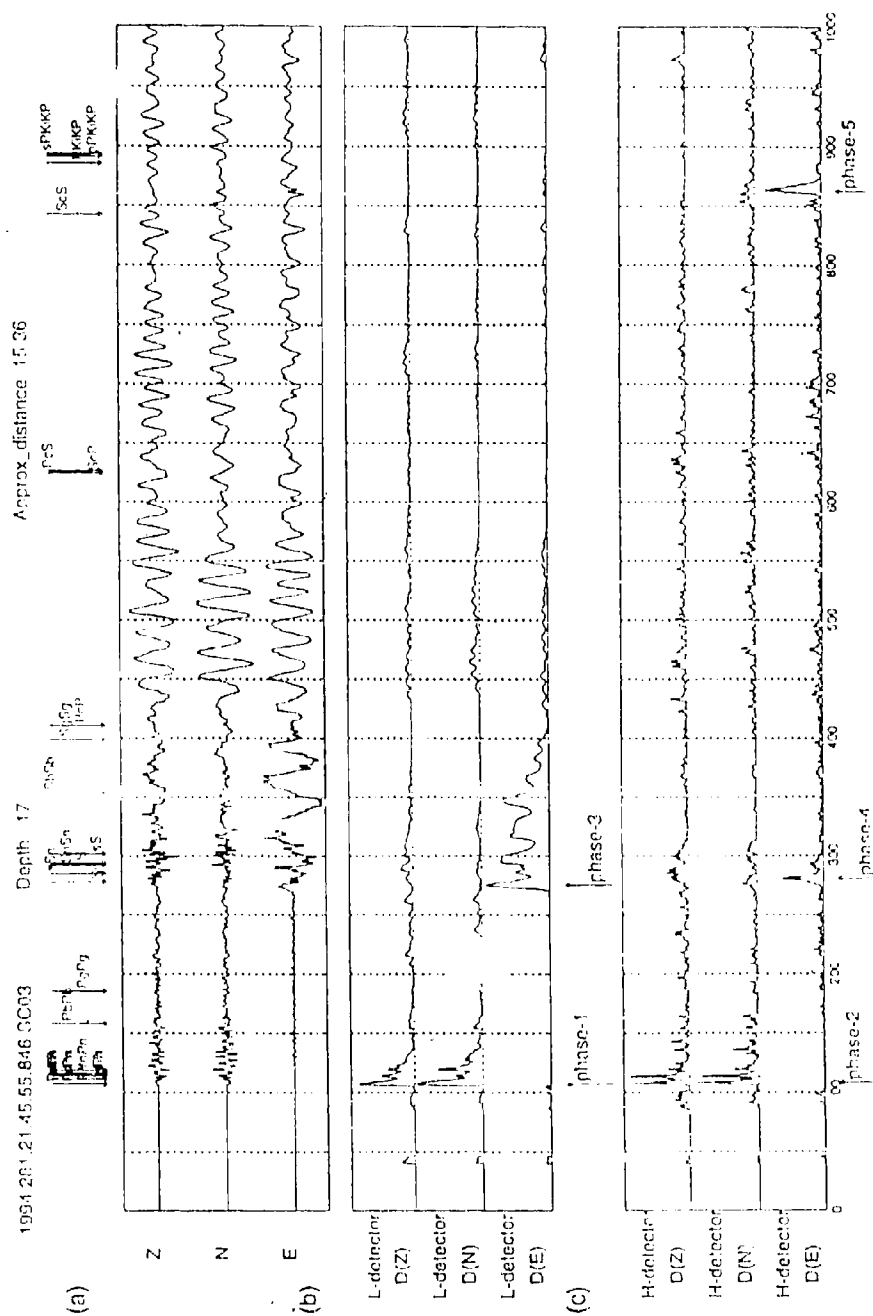


Fig. 6.4. Phase-detection and feature-extraction for event A - Halmahera (Indonesia).

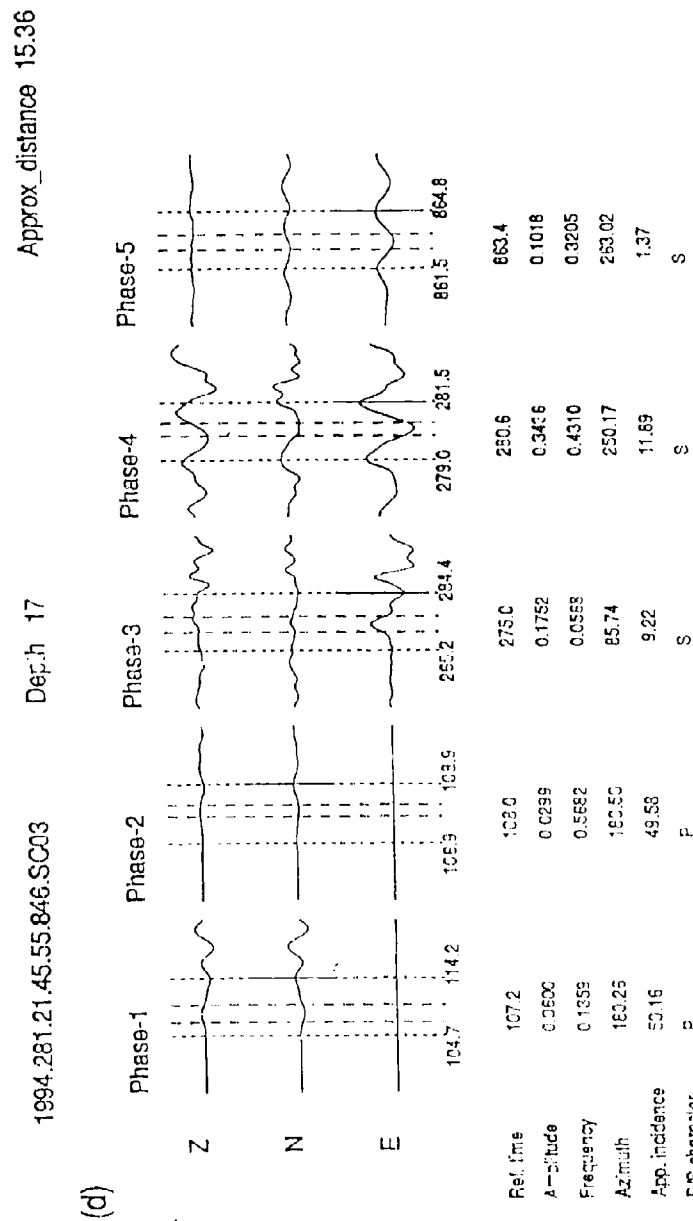


Fig. 6.4. - continued.

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 MatchSurviving models: *P-S*

Step 4:

Event Pattern Match

<i>P-S</i> pair	Depth	Dist	Azim	Identified Phases	Misfit
<i>P-S</i>	0	14	180	phase 1: <i>P</i> , phase 3: <i>S</i> , phase 5: <i>ScS</i>	0.0033
	100	15	180	phase 1: <i>P</i> , phase 3: <i>S</i>	0.3333
	300	16	180	phase 1: <i>P</i> , phase 3: <i>S</i>	0.3333
	600	16	180	phase 1: <i>P</i> , phase 3: <i>S</i>	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 four nodes in the second generation corresponding to different interpretations 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

6.3.2 Event 13 Myanmar Burma 1

^a $\chi^2 = 0.67$, d.f. = 1, $p = .41$.

• • •

• • •

• • • • •

•

100

Each of the four distance and depth combinations give rise to a 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

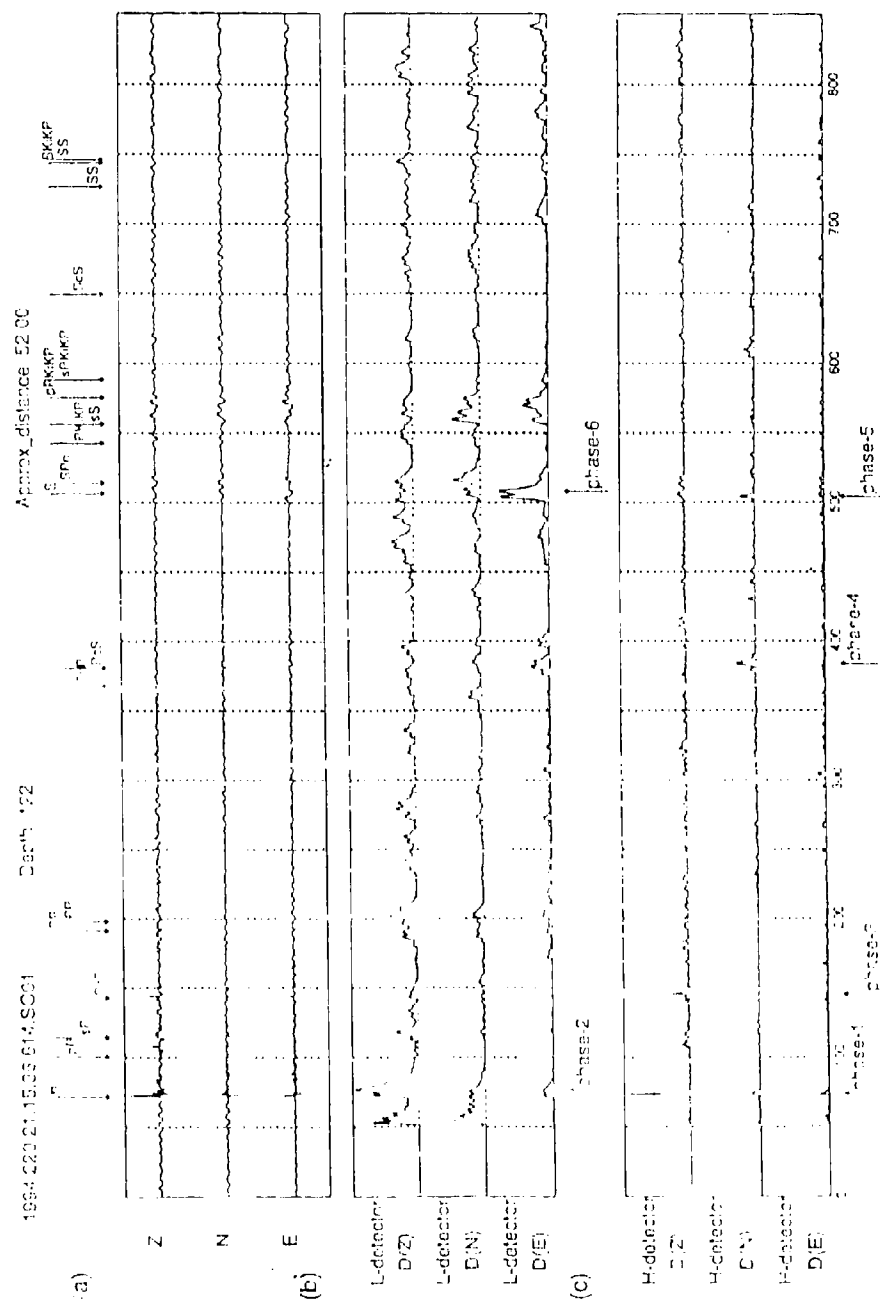


Fig. 6.5. Phase-detection and feature-extraction for event B - Myanmar (Burma).

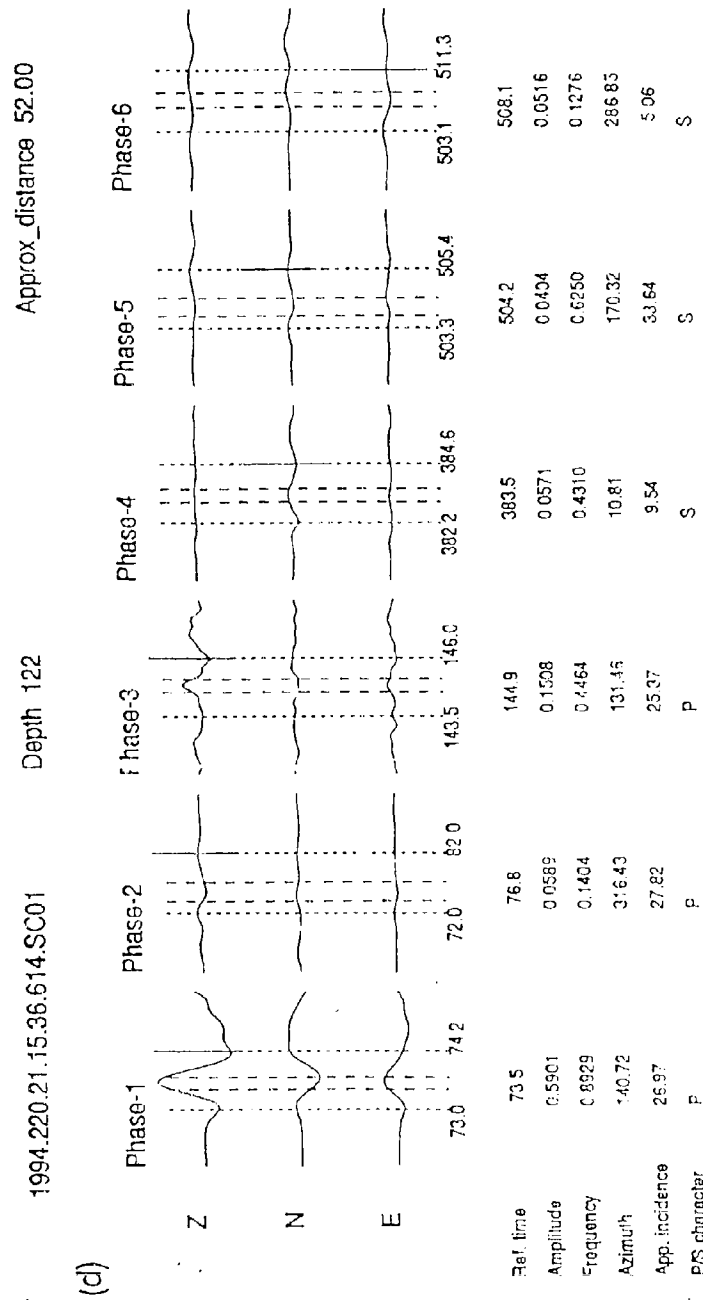


Fig. 6.5. - continued.

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

Step 4:

Event Pattern Match

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

(which also implies 136°), which matches 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

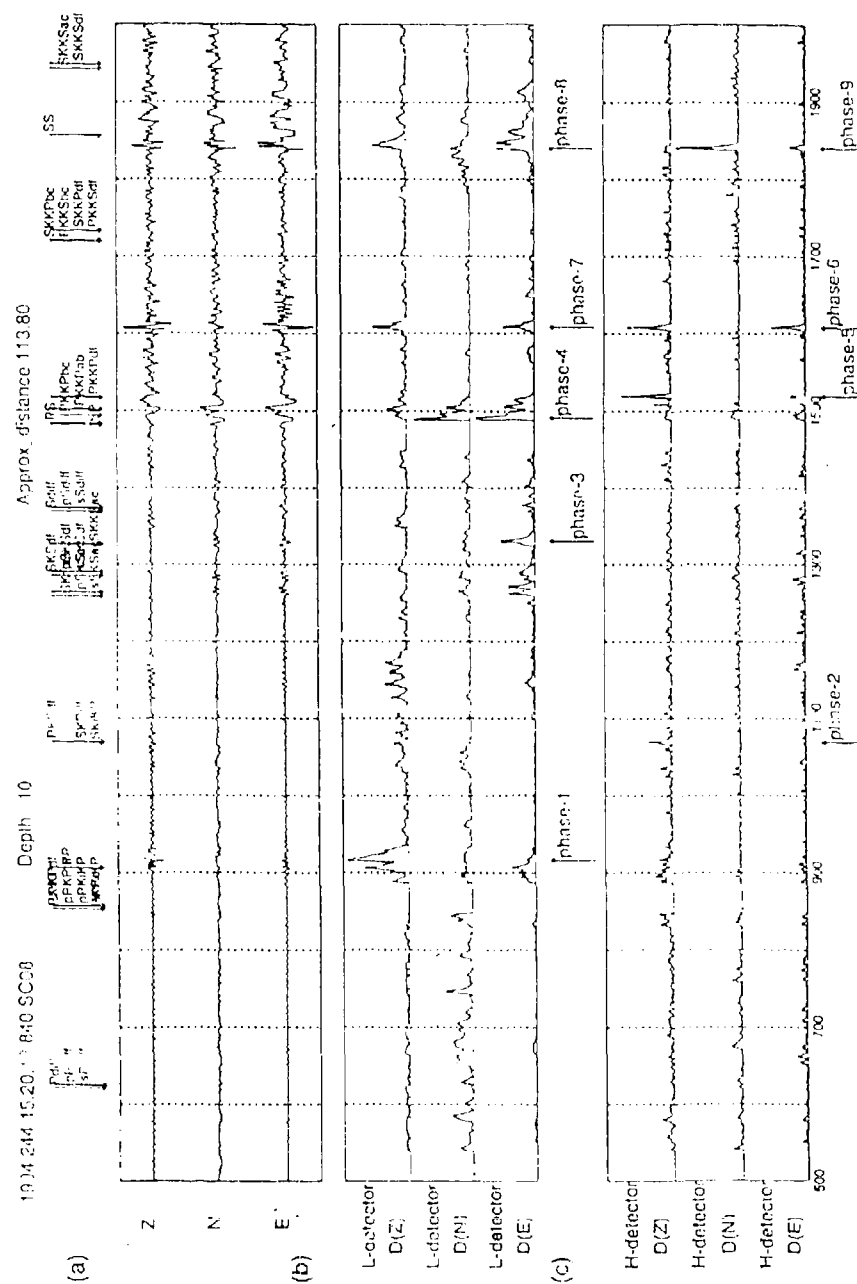


Fig. 6.6. Phase-detection and feature-extraction for event C – off coast of Northern California.

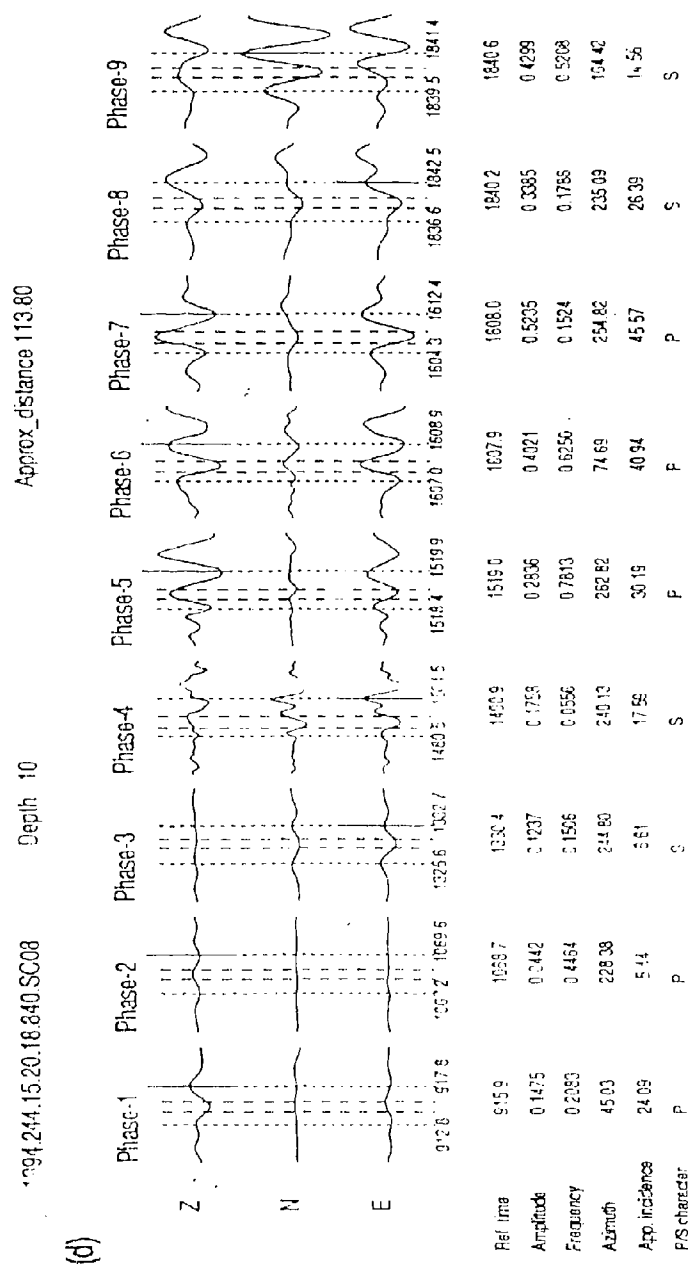


Fig. 6.6. - continued.

frequency *S* arrival [phase-9] have no obvious correlation with the phases expected for this event 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 source 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 reasonably 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 6.6.1 (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*

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

6.5.4 Event D - Northern Bolivia

The epicentral distance is even larger for this event and the seismogram is not 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

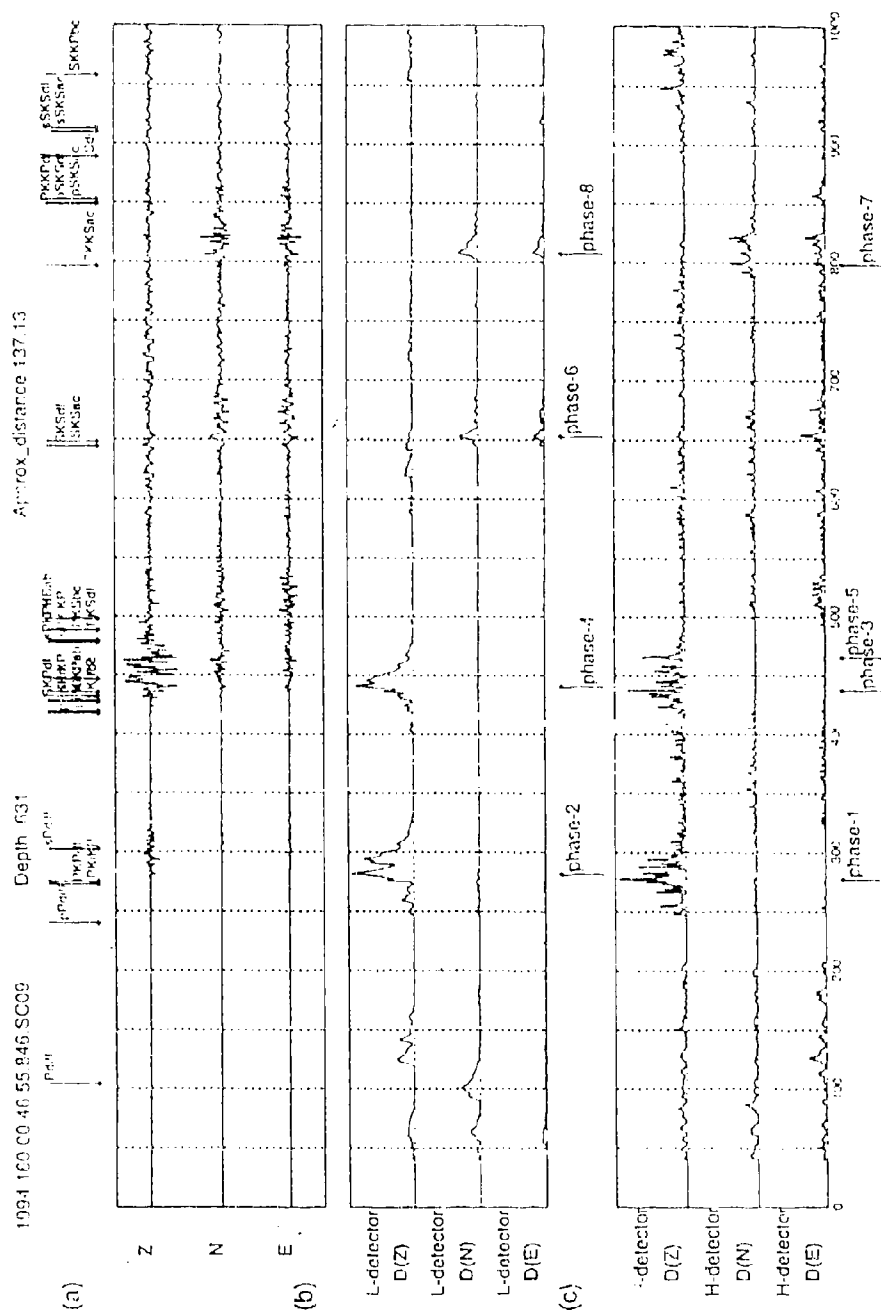


Fig. 6.7. Phase detection and feature-extraction for event D - Northern Bolivia.

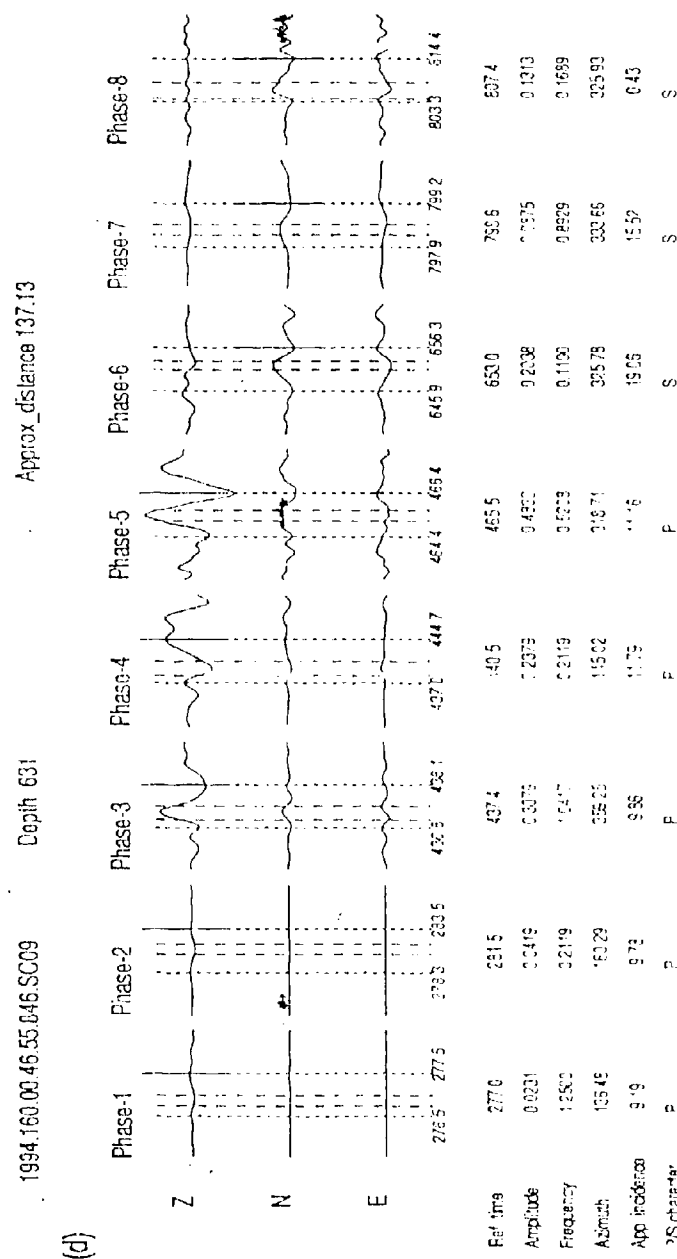


Fig. 6.7. - continued.

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*

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

For step to be a *P* in the choice of the *P-S* pair, the branch corresponding to *P* 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 employing a denser travel-time table for each depth. The current number of depth/distance combinations (36) 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 detection, 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.

Automatic seismic event recognition using multiple broad-band stations

by C. Tong

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 procedure 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	127.93	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	150	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

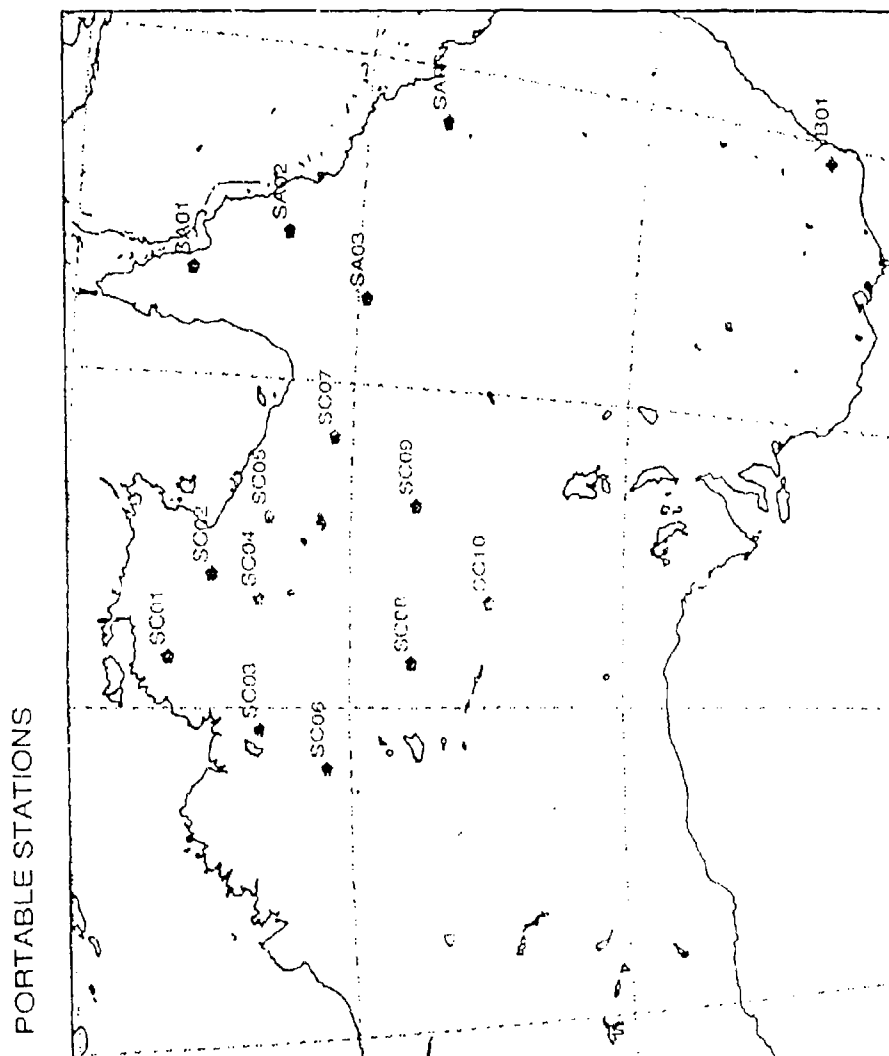


Fig. 7.1. The stations which have been used for the extensive testing.

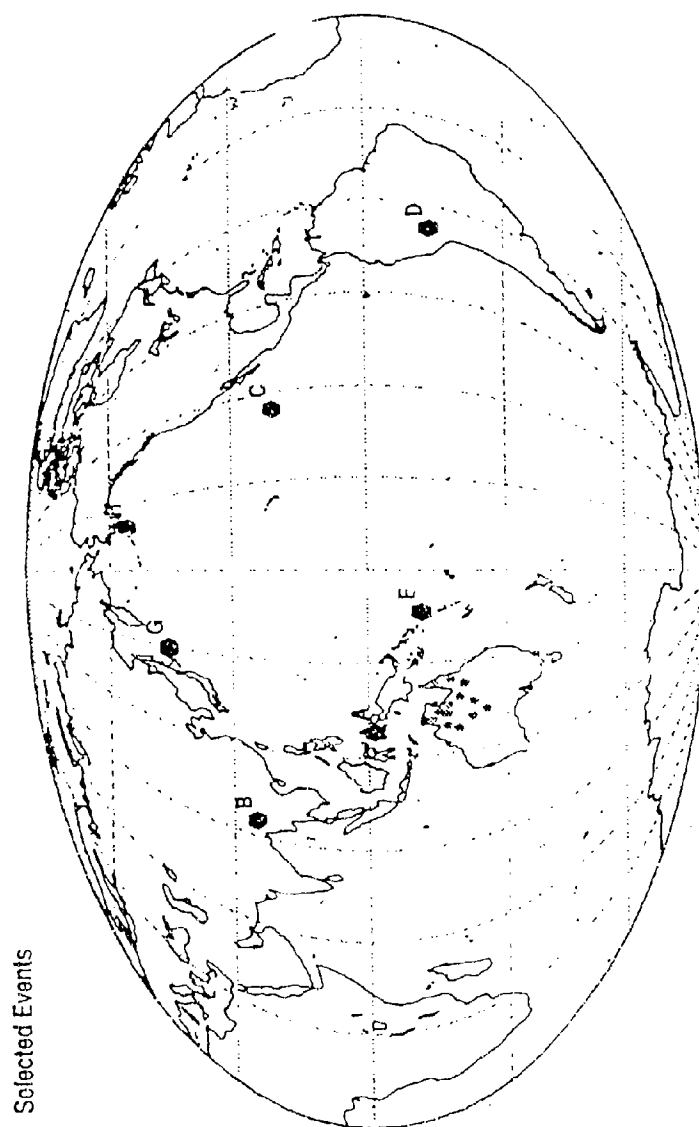


Fig. 7.2. The locations of events which have been used for the extensive testing.

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	000	170, 350	21	171
SC10	P-S	26	0	163, 343	21	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 - 691 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 SC1 and SC5, but is not large enough to meet the present 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	195
SC01	P-S	60	100	15, 195	63	201
SC05	P-S	62	100	11, 191	63	197
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	195

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 MatchSurviving models: *P-S*, *PKP-SKKS*

Step 4:

Event Pattern Match

<i>P-S</i> pair	Depth	Dist	Azim	Identified Phases	Misfit
<i>P S</i>	0	58	16 or 196	phase 1: <i>P</i> , phase 6: <i>S</i>	0.5158
	100	60	16 or 196	phase 1: <i>P</i> , phase 3: <i>pP</i> , phase 6: <i>S</i>	0.1875
	300	64	16 or 196	phase 1: <i>P</i> , phase 3: <i>PcP</i> , phase 6: <i>S</i>	0.2148
	600	68	16 or 196	phase 1: <i>P</i> , phase 4: <i>pP</i> , phase 6: <i>S</i>	0.2903
<i>PKP SKKS</i>	0	126	16 or 196	phase 1: <i>PKP</i> , phase 6: <i>SKKS</i>	0.5616
	100	126	16 or 196	phase 1: <i>PKP</i> , phase 3: <i>pPKP</i> , phase 6: <i>SKKS</i>	0.2456
	300	126	16 or 196	phase 1: <i>PKP</i> , phase 6: <i>SKKS</i>	0.4818
	600	128	16 or 196	phase 1: <i>PKP</i> , phase 4: <i>pPKP</i> , phase 6: <i>SKKS</i>	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

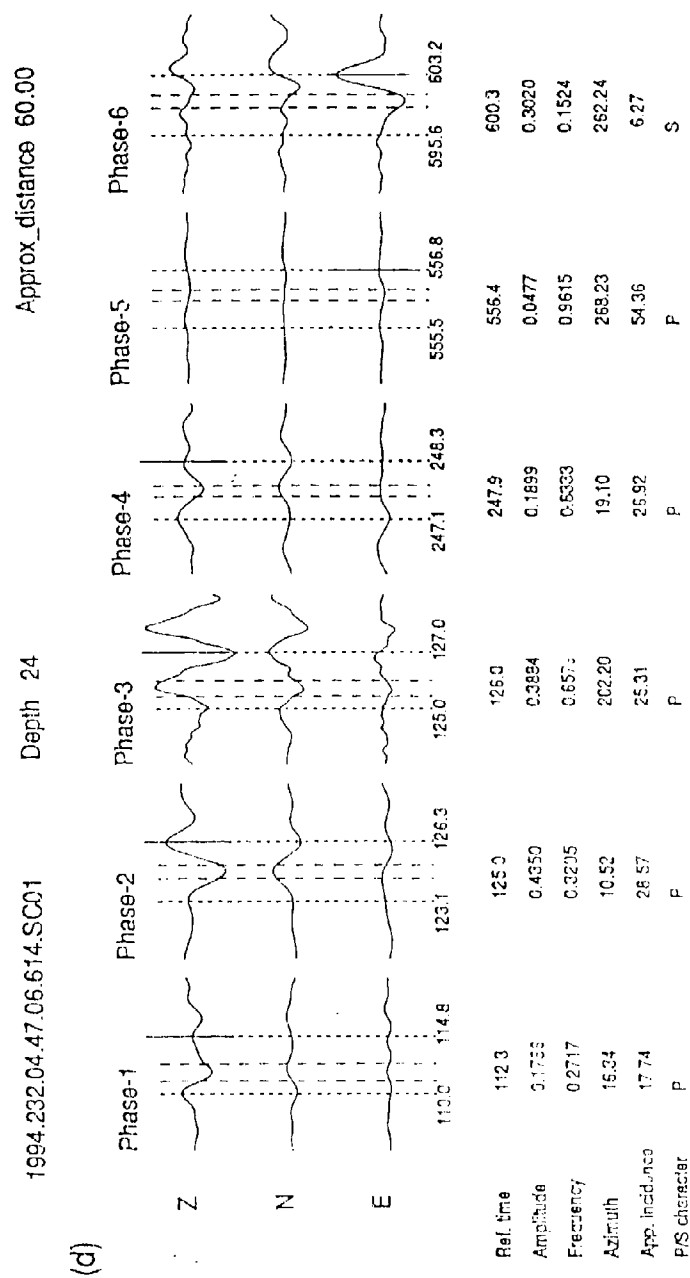


Fig. 7.3. - continued.

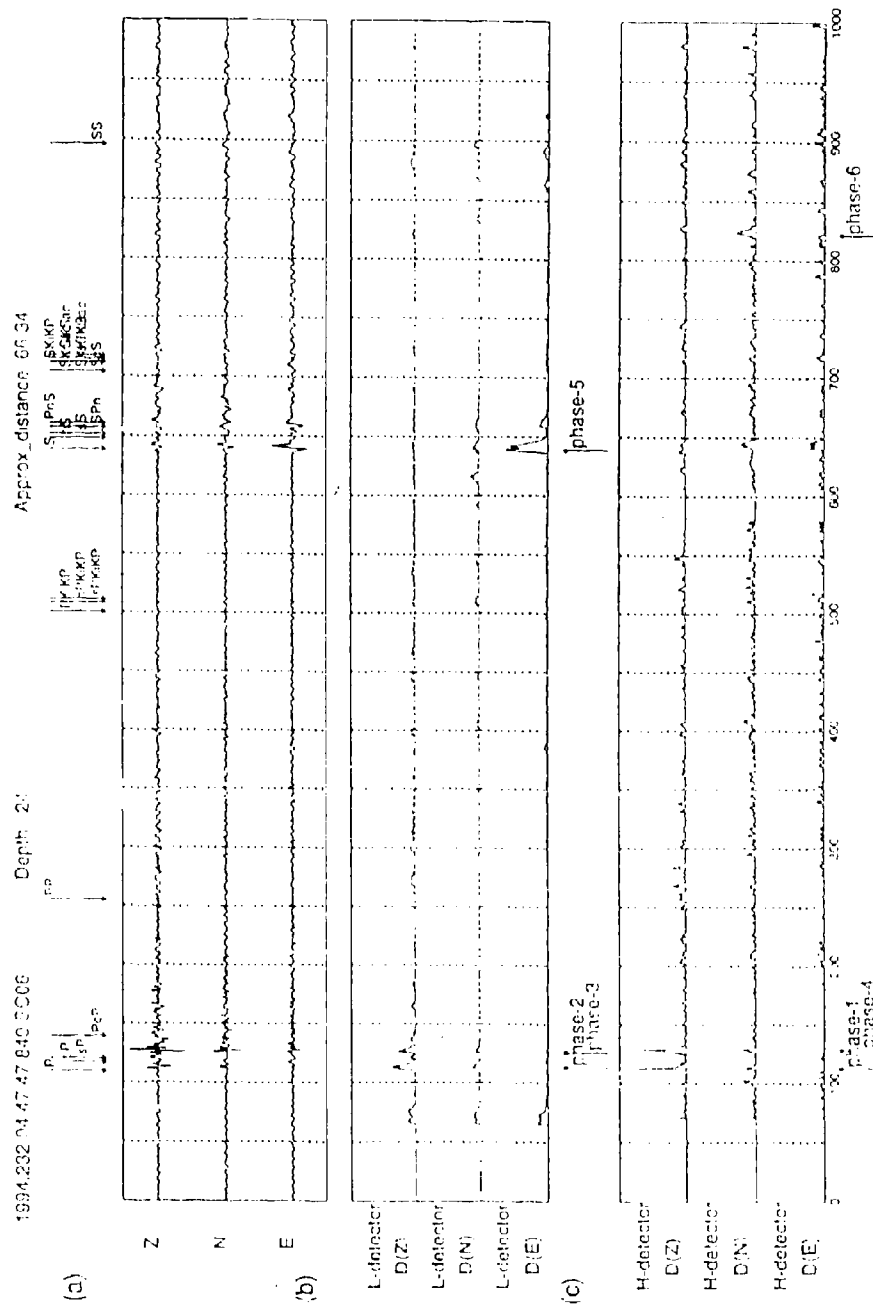


Fig. 7.4. Phase-detection and feature-extraction for event F at the station SC06.

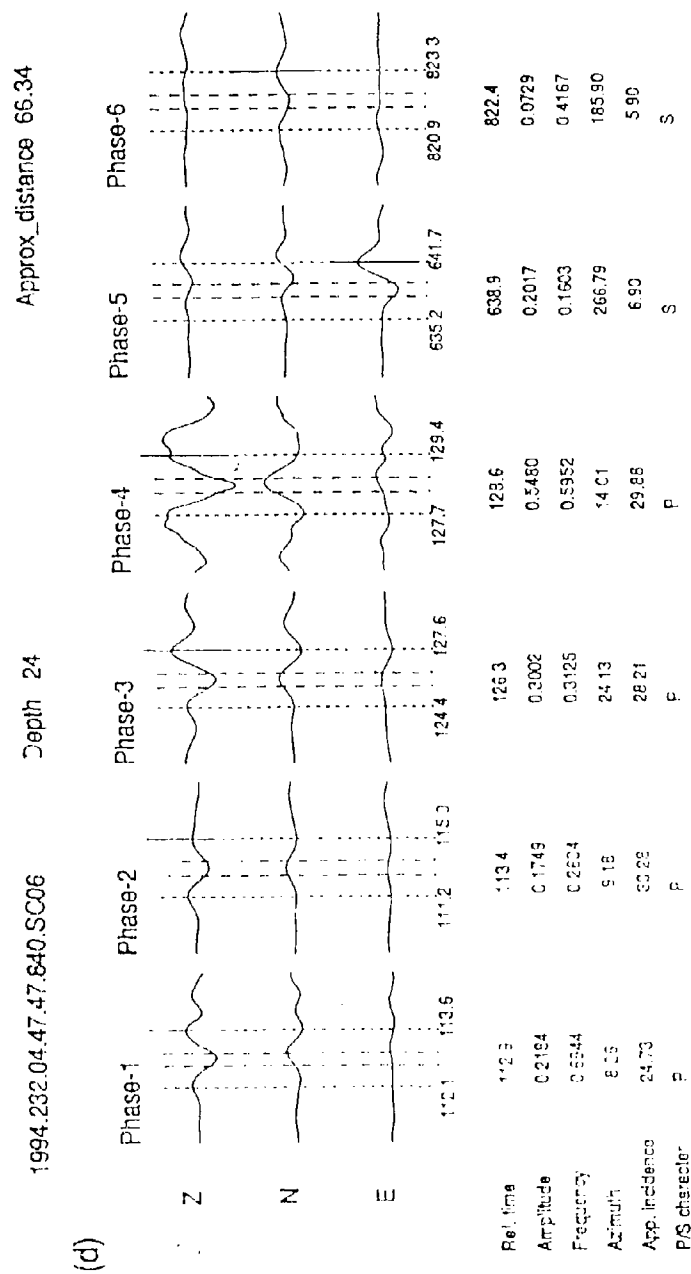


Fig. 7.4. - continued.

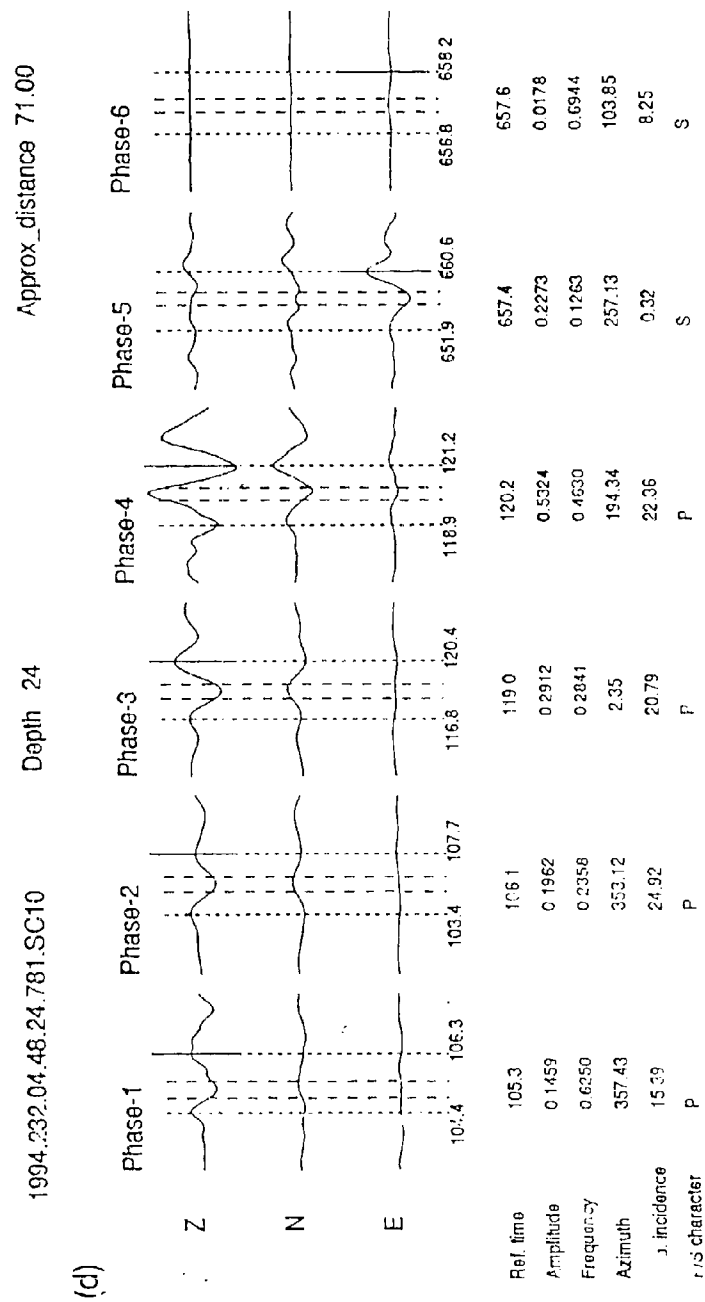


Fig. 7.5. - continued.

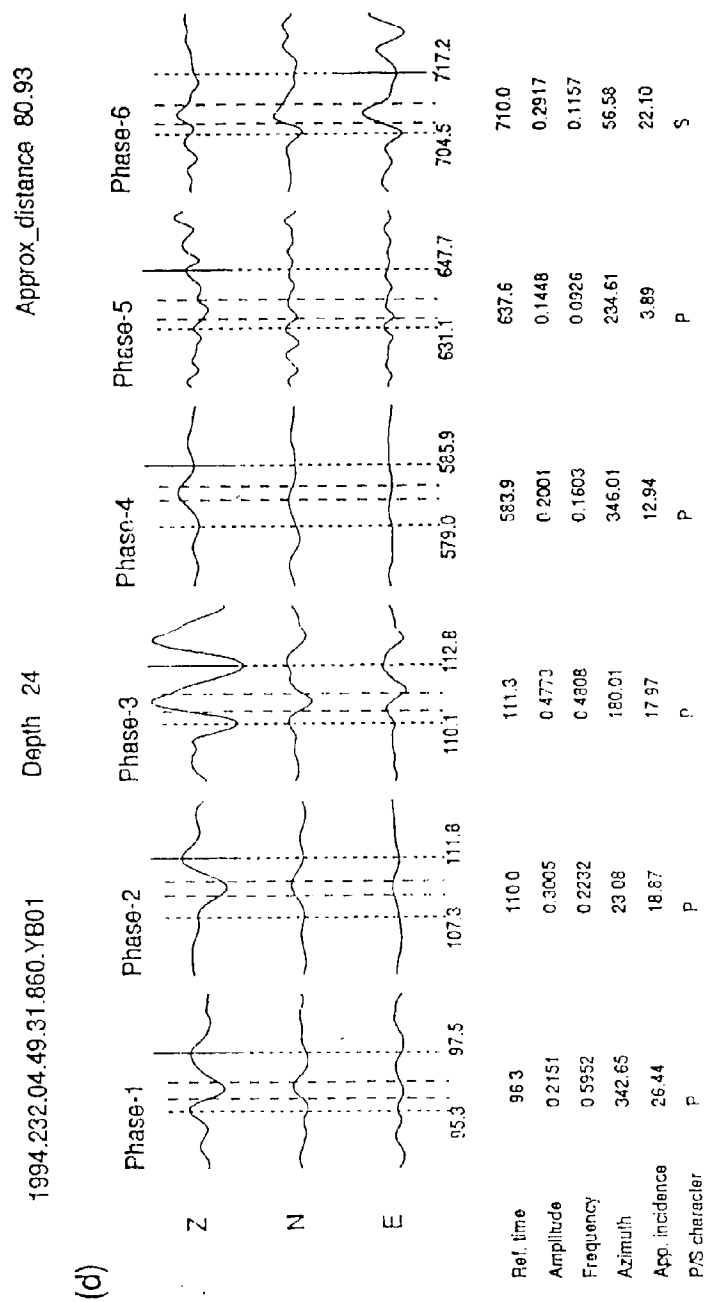


Fig. 7.6. - continued.

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-SKK'S**Step 3: P-S Pair Feature Match*Surviving models: *P-S**Step 4:**Event Pattern Match*

<i>P-S</i> pair	Depth	Dist	Azim	Identified Phases	Misfit
<i>P-S</i>	0	66	9 or 189	phase 2: <i>P</i> , phase 5: <i>S</i>	0.5970
	100	66	9 or 189	phase 2: <i>P</i> , phase 4: <i>pP</i> , phase 5: <i>S</i>	0.2543
	300	70	9 or 189	phase 2: <i>P</i> , phase 5: <i>S</i>	0.4915
	600	74	9 or 189	phase 2: <i>P</i> , phase 5: <i>S</i>	0.4915

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-SKK'S**Step 3: P-S Pair Feature Match*Surviving models: *P-S**Step 4:**Event Pattern Match*

<i>P-S</i> pair	Depth	Dist	Azim	Identified Phases	Misfit
<i>P-S</i>	0	70	173 or 353	phase 2: <i>P</i> , phase 5: <i>S</i>	0.5514
	100	72	173 or 353	phase 2: <i>P</i> , phase 4: <i>pP</i> , phase 5: <i>S</i>	0.2338
	300	74	173 or 353	phase 2: <i>P</i> , phase 5: <i>S</i>	0.4957
	600	86	173 or 353	phase 2: <i>P</i> , phase 5: <i>S</i>	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_{avg}-SKS*, *PKP-SKS*Step 3: *P-S* Pair Feature MatchSurviving models: *P-S*, *P-SKS*

Step 4:

Event Pattern Match

<i>P-S</i> pair	Depth	Dist	Azim	Identified Phases	Misfit
<i>P-S</i>	0	82	163 or 343	phase 1: <i>P</i> , phase 6: <i>S</i>	0.6630
	100	84	163 or 343	phase 1: <i>P</i> , phase 3: <i>pP</i> , phase 6: <i>S</i>	0.2975
	300	84	163 or 343	phase 1: <i>P</i> , phase 6: <i>S</i>	0.5102
	600	84	163 or 343	phase 1: <i>P</i> , phase 6: <i>S</i>	0.5102
<i>P-SKS</i>	0	126	163 or 343	phase 1: <i>P</i> , phase 6: <i>SKS</i>	0.5693
	100	126	163 or 343	phase 1: <i>P</i> , phase 6: <i>SKS</i>	0.5693
	300	126	163 or 343	phase 1: <i>P</i> , phase 6: <i>SKS</i>	0.5102
	600	128	163 or 343	phase 1: <i>P</i> , phase 6: <i>SKS</i>	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 choosing a number of choices of distance within an error tolerance rather than choosing 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 phases. 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.

AI techniques applied in the automatic interpretation system

by C. Tong

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

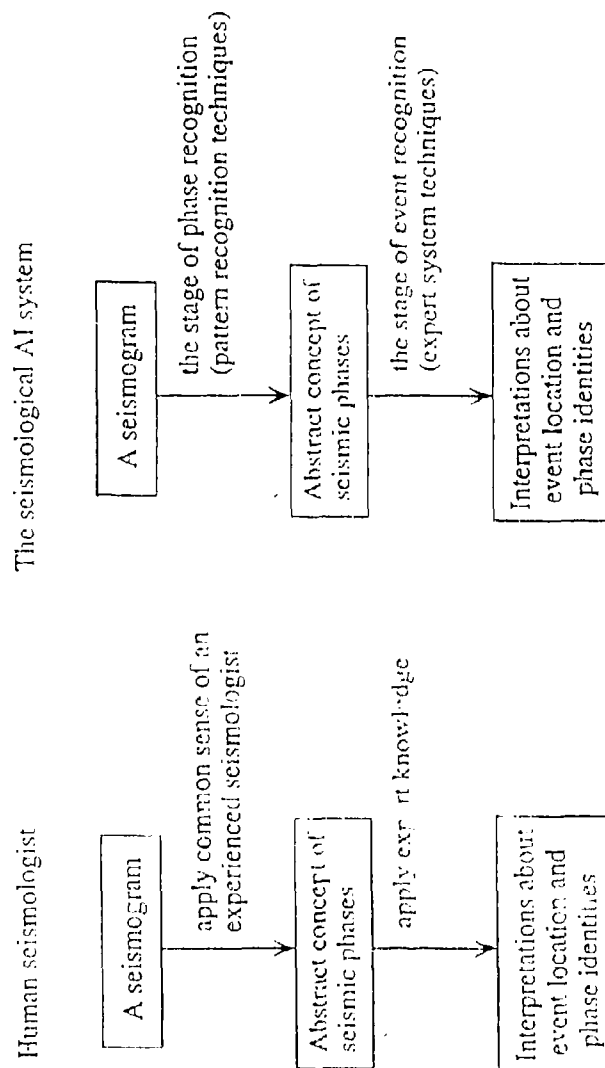


Fig. 8.1. Comparison of seismogram interpretation procedures performed by human seismologist and by the designed automatic system.

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 " $\rightarrow$ " mean "can be derived from", " $\Rightarrow$ " mean "is derived from" formed by performing the following transcription (or "rewrite") rules:

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

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 $\alpha \rightarrow \beta$, where α and β are strings over V ($V = V_N \cup V_T$), 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,

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

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

$$\zeta_1 A \zeta_2 \rightarrow \zeta_1 \beta \zeta_2,$$

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 $|\beta| \geq |\lambda|$.

In type 2 (context-free) grammars, the productions are of the form " $A \rightarrow \beta$ ", 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 $A \rightarrow aB$ or $A \rightarrow b$, 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

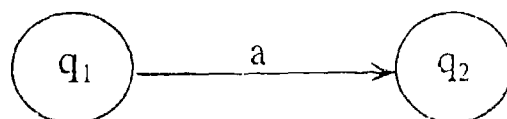


Fig. 8.2. Graphical representation of the transition $\delta(q_1, a) = q_2$.

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

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

in which

- (1) N is a finite set of input symbols (alphabet)
- (2) Q is a finite set of states,
- (3) δ is a mapping of $Q \times N$ 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 the diagram is: the automaton in state q_1 and scanning the input symbol a moves to state q_2 . (The input device moves one step to the next position (treated as 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_a : 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_e : the last unit of a phase wavelet;
- (6) U_c : an incoherent unit which ends the phase but does not complete a wavelet completely.

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_b or U_c 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_n , 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.

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

where $\langle \text{Seismogram} \rangle$ is the start symbol,

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

is the set of nonterminals;

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

is the set of terminal symbols;

and the productions are

$$P = \{$$

$$\langle \text{Seismogram} \rangle \rightarrow U_b \langle \text{Seismogram} \rangle,$$

$$\langle \text{Seismogram} \rangle \rightarrow U_o \langle \text{Phase} \rangle,$$

$$\langle \text{Phase} \rangle \rightarrow U_l \langle \text{Left} \rangle,$$

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

$\}$.

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.

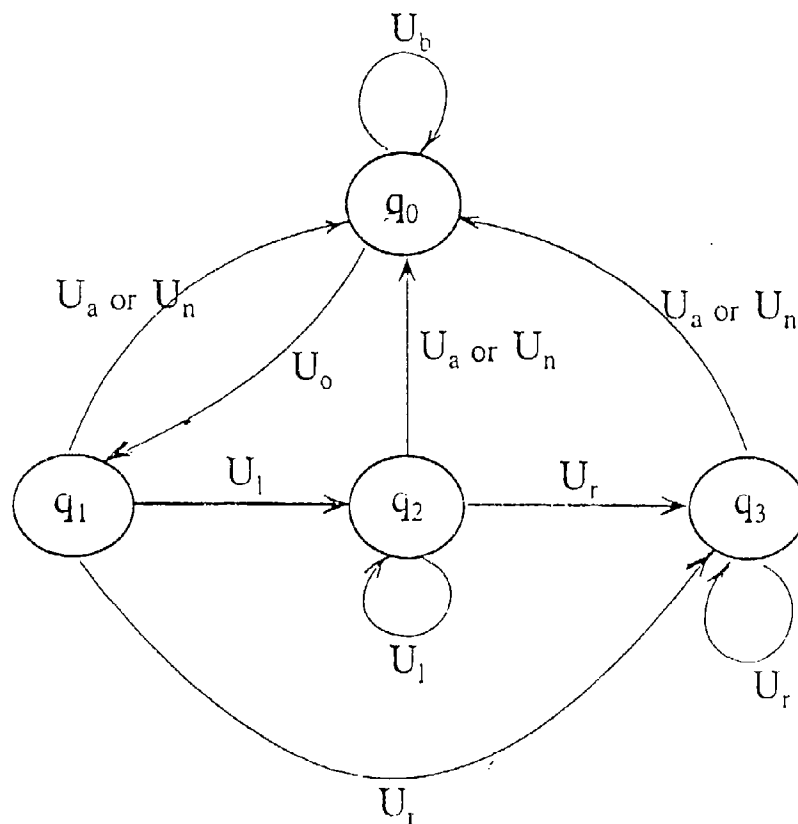


Fig. 8.3. State transition diagram of the phase-recognition automaton A.

Table 8.1. Transitions in the automaton A

$\delta(q_0, U_b) = q_0.$	$\delta(q_0, U_o) = q_1.$
$\delta(q_1, U_l) = q_2.$	$\delta(q_1, U_r) = q_3.$
$\delta(q_1, U_a) = q_0.$	$\delta(q_1, U_n) = q_0.$
$\delta(q_2, U_l) = q_2.$	$\delta(q_2, U_r) = q_3.$
$\delta(q_2, U_a) = q_0.$	$\delta(q_2, U_n) = q_0.$
$\delta(q_3, U_r) = q_3.$	$\delta(q_3, U_n) = q_0.$
$\delta(q_3, U_a) = q_0$	}

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, repre-

sented 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 evolved 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.

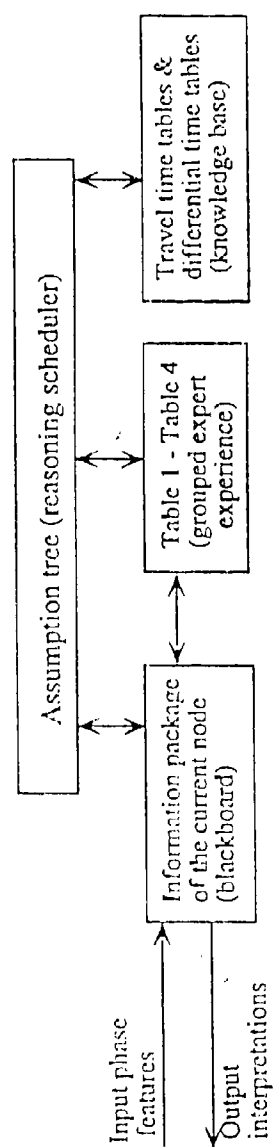


Fig. 8.4. Entities and their relationships in the seismological expert system.

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

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

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

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

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 ($F_1, F_2, \dots, F_m$), 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 $F_1, F_2, \dots$, and F_m) - S_1 - S_2 - - S_{m-1} .

At the i th 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:

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

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 $F_1, F_2, \dots, F_m$ 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 $n_2 \times n_3 \times n_4 \times \dots \times n_m$ solutions (note: n_1 is the number of choices for the feature F_1), 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 i th 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 i th 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

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

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

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

Generally, the quality of the i th generation is given by

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

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 pre-decided 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.
- Cervený, 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**, 1595-1912.
- 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. SeS 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., & Coody, 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 Varananga seismic array. Ph.D. thesis, Aust. Natl. Univ., Canberra.
- Hess, H., 1961. 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 Heidelberg, New York.
- Jackson, P., 1990. Introduction to expert systems, 2nd ed., Addison-Wesley.
- Journet, B., & Jobert, N., 1982. Variation with age of anisotropy under ocean, 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. Inter.*, **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 herzolites, *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**, 853-870.
- Nicolas, A., & Poirier, J. P., 1976. Crystalline Plasticity and Solid-State Flow in Metamorphic Rocks. Wiley, London, 437 pp.
- Nicolas, A., & Christensen, N. I., 1937. 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.
- Rütt, 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 carotic 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 to 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.