

U. S. DEPARTMENT OF THE INTERIOR

U. S. GEOLOGICAL SURVEY

A COURSE ON: PC-BASED SEISMIC NETWORKS

Edited by

W. H. K. Lee

MS 977, 345 Middlefield Road

Menlo Park, CA 94025

and

D. A. Dodge

Department of Geophysics

Stanford University

Stanford, CA 94305

Open-File Report 92-441

August, 1992

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

CONTENTS

PREFACE	iii
1. REGIONAL SEISMIC NETWORKS IN CALIFORNIA <i>by J. P. Eaton</i>	1
2. PC-BASED SEISMIC SYSTEMS <i>by W. H. K. Lee</i>	53
3. SEISMOMETERS THEORY AND PRACTICE <i>by J. P. Eaton</i>	79
4. BASIC TECHNIQUES FOR TELEMETRY <i>by J. R. VanSchaack</i>	100
5. REALTIME SEISMIC DATA ACQUISITION <i>by W. H. K. Lee</i>	115
6. THE XDETECT PROGRAM <i>by W. H. K. Lee</i>	138
7. THE TDETECT PROGRAM <i>by J. R. Evans</i>	152
8. ROUTINE SEISMIC NETWORK DATA PROCESSING <i>by W. H. K. Lee</i>	165
9. PLOTTING SEISMOGRAMS AND MAPS <i>by R. Banfill</i>	190
10. MATHEMATICS FOR EARTHQUAKE LOCATION <i>by W. H. K. Lee</i>	207
11. LOCAL EARTHQUAKE LOCATION PROGRAMS <i>by J. C. Lahr</i>	226
12. COMPUTING TRAVEL TIME AND DERIVATIVES <i>by W. H. K. Lee</i>	251

13. DEVELOPMENT OF EARTHQUAKE MAGNITUDE SCALES	
<i>by J. P. Eaton</i>	281
14. CODA WAVES FOR MAGNITUDE AND Q <i>by W. H. K. Lee</i>	316
15. PCEQ and QCODA <i>by C. M. Valdes</i>	336
16. MATHEMATICS FOR WAVEFORM ANALYSIS <i>by A. Lomax</i>	359
17. THE SEISGRAM PROGRAM <i>by A. Lomax</i>	383
18. FOCAL MECHANISM: THEORY AND HISTORY <i>by P. A. Reasenberg</i>	405
19. THE FPFIT PROGRAM <i>by D. H. Oppenheimer</i>	424
20. MATHEMATICS FOR SEISMIC TOMOGRAPHY <i>by W. H. K. Lee</i>	461
21. APPLICATIONS OF SEISMIC TOMOGRAPHY <i>by H. M. Iyer</i>	479
APPENDIX: Implementing a PC-Based Seismic System: Questions & Answers	
<i>by W. H. K. Lee</i>	512
NAME INDEX <i>by V. Tsai</i>	527
SUBJECT INDEX <i>by V. Tsai</i>	530

PREFACE

A course on PC-based seismic networks was given to about 20 students in Menlo Park, CA from September 9 through 20, 1991. The purpose of this course was to show students how to operate a PC-based seismic network, carry out routine data processing, and perform basic data analysis. A total of 21 lectures were given during the mornings, and laboratory sessions were conducted during the afternoons.

This Open-file Report consists of the 21 lectures given in the above course. These lecture notes were first transcribed by Doug Dodge from video tapes recorded in the class and were then edited by the authors. Minor editing has also been performed by the editors for consistency in the presentation.

We are grateful to John Filson who provided the financial support for conducting this course, and to Carol Lawson for providing the computer training facility for the laboratory sessions. We thank (1) the authors for delivering the lectures and for conducting the laboratory sessions, (2) Randy White and Mary Allen for reviewing the manuscripts, and (3) Virginia Tsai for preparing the Name Index and the Subject Index.

W. H. K. Lee and D. A. Dodge, Editors

1. REGIONAL SEISMIC NETWORKS IN CALIFORNIA

by

J. P. Eaton

U. S. Geological Survey, Menlo Park, CA 94025

ABSTRACT

Short period seismic networks in California have a long history. They have been developed by different institutions with different objectives. Equipment for recording and analyzing earthquakes has undergone several revolutions. What can be done easily and routinely today could hardly have been imagined by the planners of the first extended networks in the decades following the great 1906 San Francisco earthquake. Moreover, the conceptual framework of plate tectonics and the needs of the earthquake hazard reduction program lead to far more detailed and sophisticated questions for the modern network to answer than those addressed by the early networks.

The plan of this paper is to trace the history of the development of seismic networks in California, with emphasis on size, density, instrumentation, and analysis procedures as well as on the purposes that the networks served. The paper is offered to help resolve the impasse that has frozen the networks, prematurely, in their 1982 configuration for nearly 10 years and to encourage a renewed effort to bring the networks to a state of completion that will permit them to fulfill their essential role in earthquake research and hazard reduction.

I. HISTORY OF NETWORK DEVELOPMENT

A. Original local earthquake networks at UC Berkeley and Caltech

The frequent occurrence of earthquakes in California and the need for coordinated networks of seismographs to study them have been recognized since the time of Holden at the dawn of instrumental seismology in the U. S. 100 years ago (Louderback, 1942). The seismic networks that have evolved in northern and southern California over the last century have pressed the limits of available technology; but for many decades the lack of adequate instruments for detecting, recording, and timing earthquake waves and for collecting and analyzing their records placed crippling restraints on the size and effectiveness of seismic networks. From 1887 to the late 1920's, the UC Berkeley stations at Mt Hamilton and Berkeley were the only stations with accurate timing in the state. They operated mechanical seismographs with magnifications of about 100. Even after the development of the Wood-Anderson and Benioff seismographs in the late 1920's and early 1930's, the California networks remained primarily reconnaissance in nature. In 1952 the California networks consisted of only 10 northern (UC Berkeley) and 15 southern (Caltech) widely scattered self-contained seismograph stations, with relatively poor time control, that wrote "paper" records of moderate dynamic range. Collection and hand processing of the records was labor intensive and slow; and the resulting earthquake solutions were generally poorly constrained, especially as regards focal depth.

Significant upgrading of the UC Berkeley northern California network was carried out by Don Tocher in 1959-1961 (Bolt, 1989) with the installation of 8 telemetered short-period stations that were recorded together on a 16 mm film recorder (Develocorder). That

equipment had been developed to serve the U S nuclear test detection program. Seismic network telemetry was introduced to southern California in 1966-1972, when most of the Caltech stations were equipped for telemetering to Pasadena for recording. Both networks remained very sparse and provided essentially reconnaissance coverage of earthquakes of magnitude 3 and larger. In 1968 the northern California network contained about 15 stations and the southern California network contained about 20 stations.

B. Early microearthquake network experiments in California

When the USGS began to develop a program of earthquake research in California in 1966 in response to the challenge posed by the Press Panel report on earthquake prediction (Press, et al., 1965), it brought different experiences with seismic instrumentation and with level-of-detail in local earthquake studies than those underlying the existing California networks. Its study of microearthquakes at Kilauea volcano in Hawaii in the late 1950's and early 1960's, by means of a dense high-gain short-period seismic net that included a small telemetered subnet at its center, had shown the importance of matching seismometer response to the recording environment and the character of the earthquakes studied (Klein and Koyanagi, 1980; Eaton, 1986a, 1986b). Its study of earthquakes produced by injection of waste-water into basement rocks beneath the Rocky Mountain Arsenal in Colorado in the early 1960's, by means of improvised seismic arrays employing truck-mounted, low-frequency seismic systems designed for long-range refraction profiling, had demonstrated the precision of hypocenter determinations that could be obtained with a suitable network (Healy, et al., 1968).

The primary instruments for earthquake studies brought to California by the USGS in 1966 were the 20 portable 3-component seismic systems that recorded on low-power, "10-day", FM tape recorders (Criley and Eaton, 1978). With internal chronometers and WWVB radio time signals recorded on tape along with high- and low-gain tracks for seismic data, these systems provided reliable timing and moderate dynamic range (60+ dB). The electronic response was flat from DC to about 17 Hz; and with the 1-Hz moving coil seismometers employed (EV-17's), the overall system response was flat, for constant peak ground velocity, from 1 Hz to 17 Hz. The shape of the response curve, coupled with the very high electronic amplification available, made these instruments very well suited for recording microearthquakes in the California environment. Earthquake signals that exceeded natural background noise levels in the frequency range 1 Hz to 20 Hz could be detected at virtually any site in the region.

During the next two years these systems were used with great success in exploratory microearthquake studies along the San Andreas fault. The 10-day portable stations were laid out in a dense cluster (5 to 10 km spacing) over the region studied, and refraction profiles were run through the cluster with truck-mounted refraction systems to determine the local crustal structure for interpreting records of earthquakes recorded by the cluster. This work was in response to the Press Panel recommendation for the development of network clusters along major faults for earthquake prediction.

The first experiment was carried out on aftershocks of the 1966 Parkfield-Cholame earthquake (Figure 1). An 8-station, 20-km diameter network of 10-day recorders was deployed around the southern end of the 1966 rupture zone and operated for about 10 weeks.

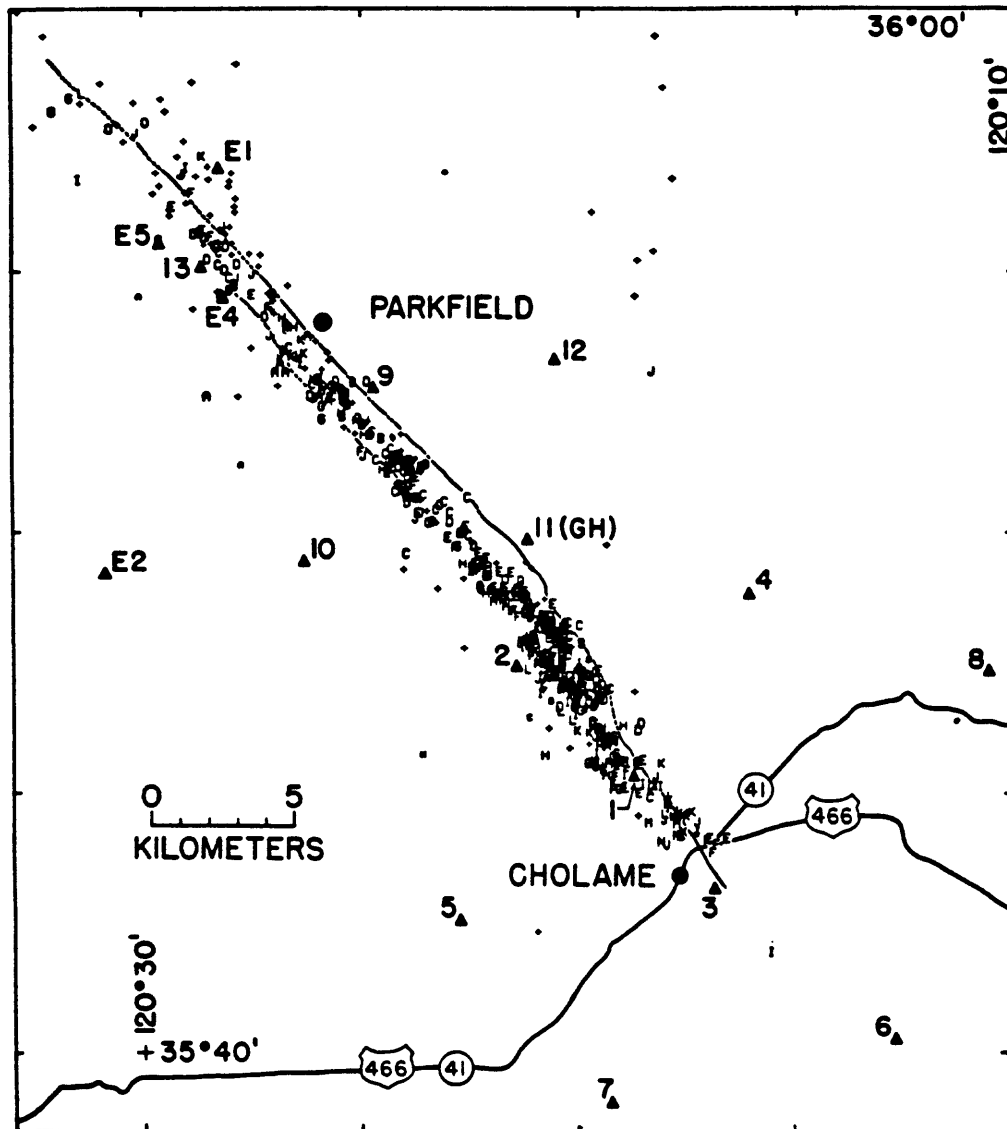


Figure 1. Aftershocks of the 1968 Parkfield-Cholame, California, earthquake. Stations of the portable network are indicated by triangles. Stations E1 through E5 were operated by the Earthquake Mechanism Laboratory of E.S.S.A.; the others by N.C.E.R. Zones of surface fracturing that accompanied the main shock and the aftershock sequence are shown as heavy solid and broken lines extending from the upper left to station 3. The letter symbol that shows the epicenter of an aftershock also indicates its focal depth: 0-1 km = A, 1-2 km = B, and so on. Aftershocks for which focal depths could not be determined are plotted as crosses.

The hypocenters of the hundreds of aftershocks recorded by the net were sufficiently precise (estimated errors less than 1 km) that they mapped out the slip surface of the main shock in great detail (Eaton et al., 1970a). In the second experiment, in 1967, an 18-station portable network about 50 km in diameter was laid out around Bear Valley, south of Hollister, to study microearthquakes on that creeping section of the San Andreas fault (Figure 2). That network, which was operated for about 6 weeks, unexpectedly recorded a shallow M4 earthquake along with hundreds of aftershocks near the center of the network. In addition, it recorded an ongoing background of small earthquakes on the San Andreas fault where it crossed the network. This study demonstrated the detail that such a network can achieve in resolving complex distributions of earthquakes in close proximity to one another (Eaton et al., 1970b).

Concurrent with the portable network experiments, the parameters for a telemetered network were being explored. Because such nets are limited by availability and cost of telemetry, careful thought was given to the selection of a data multiplexing system. A constant bandwidth, IRIG standard, 8-channel audio frequency FM system that operates over a 300 Hz to 3000 Hz voice-grade phone line was selected (Wayne Jackson, written communication; Eaton, 1976). It provides the same frequency response in each channel, DC to about 30 Hz, and can yield 40+ dB dynamic range on all channels if carefully implemented. Data recording was initially on film strip recorders (Develocorders) that permitted about 0.05 sec timing resolution and recorded 16 stations with a dynamic range of 30 to 40 dB. The overall system response was about the same as the 10-day recorder system: flat, to constant peak ground velocity, from about 2 Hz to about 15 Hz.

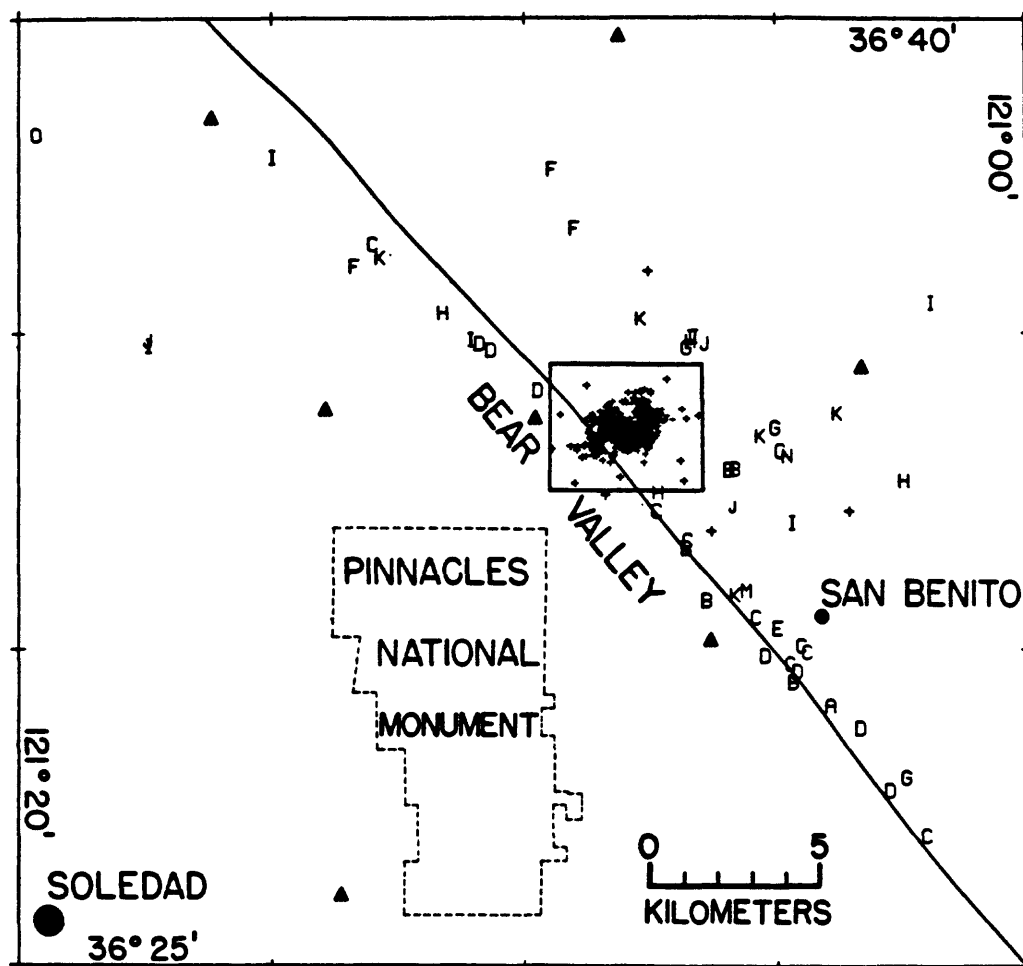


Figure 2. Aftershocks of the July 23, 1967, Bear Valley earthquake. The actively creeping trace of the San Andreas fault is shown by the solid line; but the rift zone is several km wide at Bear Valley and extends from about 1 km southwest of the active trace to about 3 km northeast of it. Portable seismograph stations are shown as solid triangles. Outside of the central rectangle, the letter symbol showing the epicenter of an earthquake also indicates its focal depth: 0-1 km = A, 1-2 km = B, and so on; a large cross indicates a shallow event for which a reliable depth could not be calculated. Inside the central rectangle hypocenters were very closely spaced (more than 300 of them), and they are plotted as small crosses.

Small experimental telemetered clusters were set up on the San Andreas fault near Palo Alto (9 stations) in 1966 and near San Juan Bautista (8 stations) in 1967. In 1967 and 1968 an additional 11 stations were set up between the Palo Alto and San Juan Bautista clusters and a small 4 station cluster was set up at Parkfield. All stations were recorded on Develocorders in Menlo Park. Analysis of 14 months' data (March 1968-May 1969) from the 30+ station telemetered network between Hollister and Palo Alto produced exciting results (Figure 3) (Eaton et al., 1970b). Some sections of the major faults (probably creeping at depth) were marked by dense, narrow zones of microearthquakes between the surface and 10 to 12 km depth, while other sections (probably locked at depth) had virtually no microearthquakes along them. The three-dimensional mapping of microearthquakes made possible by the telemetered network provided new details on the subsurface relationships between faults that were mapped in close proximity at the surface.

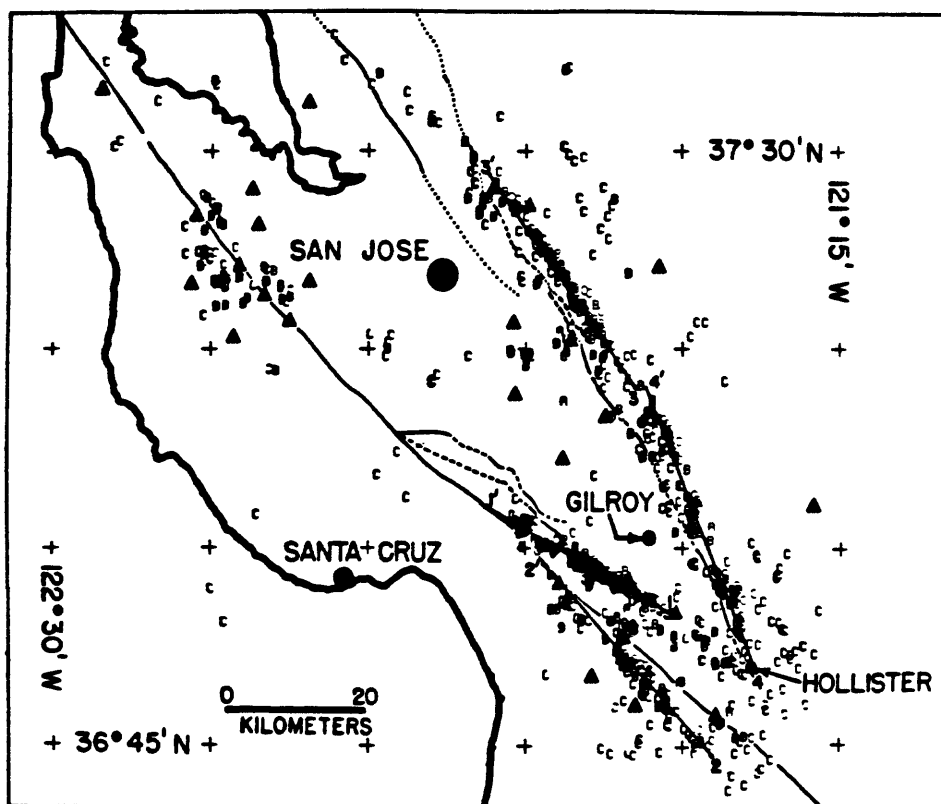


Figure 3. Epicenters of well recorded events within the telemetered network from March 1968 through April 1969. Plotted symbols indicate the reliability of hypocenter determinations: A, well determined epicenter (± 1 km) and focal depth (± 2 km); B, fairly well determined epicenter (± 2.5 km) and focal depth (± 5 km); and C, moderately well determined epicenter (± 5 km) but undetermined focal depth. Zones of hypocenter concentrations marked off by the numbered lines are as follows: 1-1', Sargent fault; 2-2', San Andreas fault west of Hollister; 3-3', Calaveras fault, northern section; 4-4', Calaveras fault, southern section.

C. Growth of a full scale microearthquake network in central California

Lessons drawn from the three experiments described above were: 1) dense microearthquake networks can map faults in three dimensions on the basis of aftershocks of large quakes or ongoing microearthquake activity associated with creeping sections of the faults; 2) the portable nets attain good resolution and are very flexible, but they require considerable effort and time to record, collate, and analyze the data; 3) the telemetered network, with somewhat sparser station spacing, attained results comparable to those of the portable nets, was far simpler to operate and analyze, and could be operated continuously rather than sporadically; 4) a telemetered strip network along the major faults would permit mapping of locked and creeping sections as well as provide a long term record of variations of activity along the faults.

These lessons provided impetus for considerable expansion of the use of telemetered networks over the next decade. The expansion took two forms: gradual expansion of the central California network to cover the Coast Ranges from Cholame to Clear Lake, and deployment of a large number of detached, special-purpose environmental networks that were analyzed separately from the central California network and from each other. Some of the detached networks eventually became important extensions of the central or southern California networks. When the first broad plans for a California prediction network were developed in 1971, the overall network was conceived as a group of strip networks along the major faults with large blank areas between them (Eaton, 1971).

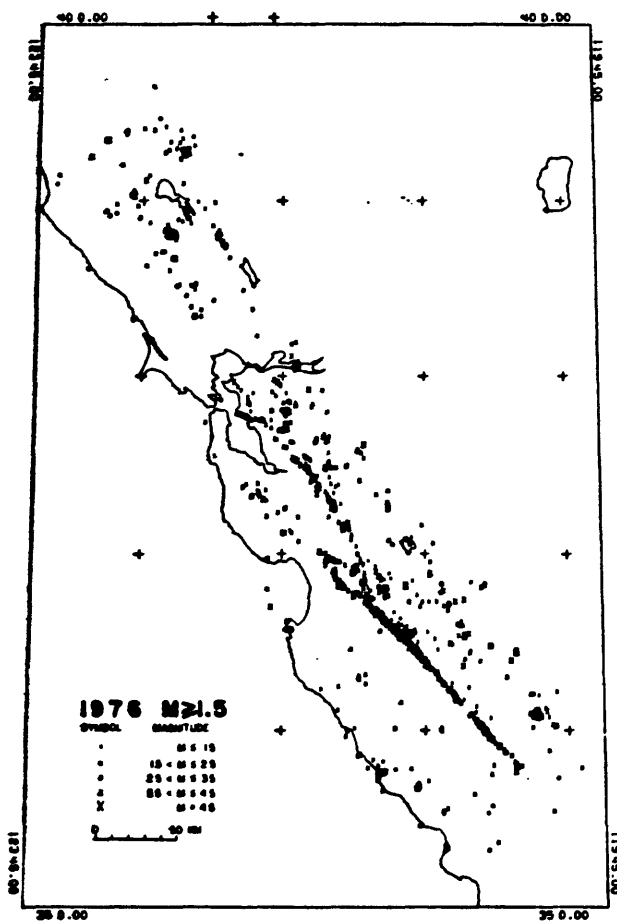
From 1966 through 1979 the central California network was viewed as an experiment to develop a dense network covering the most active part of the San Andreas fault system in

central California and to evaluate what role such a net should play in an earthquake research/hazard reduction program. All stations were recorded on Develocorders (and magnetic tape after the mid-1970's). Events that were detected by scanning the Develocorder films were timed by hand on the viewer screen or on a tabletop digitizer onto which an image of the film was projected. Events that originated significantly outside the network were not processed. Summary results from the network for the years 1970 through 1977 (Eaton, 1985) are as follows (Figure 4):

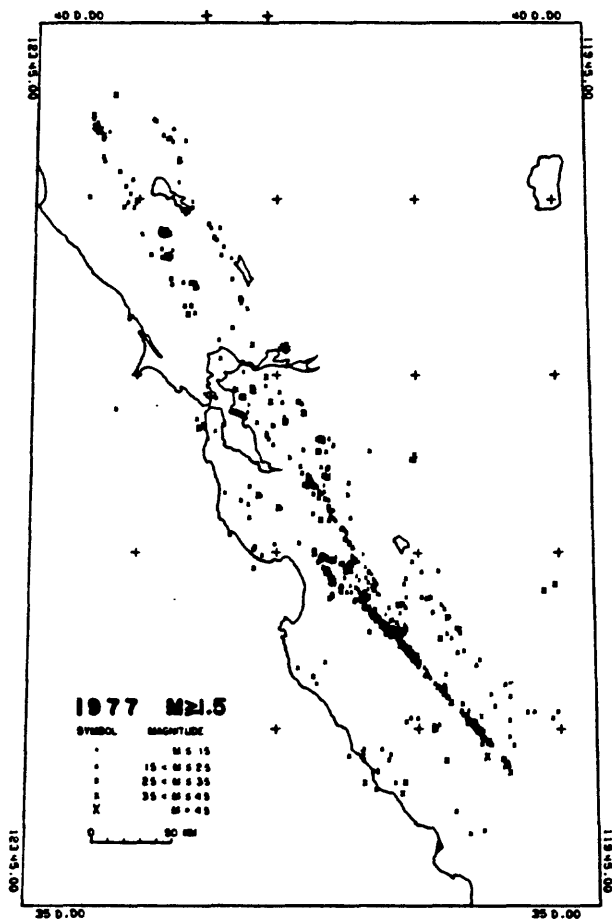
- 1) yearly plots of M1.5 and larger shocks show dense continuous lines of epicenters along creeping sections of the major faults;
- 2) locked sections of major faults, including the sections of the San Andreas fault that broke in 1906 and 1857, are virtually aseismic;
- 3) earthquakes scattered across the Coast Ranges are somewhat concentrated in bands along both flanks of the Coast Ranges;
- 4) focal depths were generally well determined along the major faults near the center of the network but were poorly determined along the flanks of the Coast Ranges where the network was sparse.

D. Emergence of the northern and southern California regional networks

In early 1980 the procedures for analyzing stations telemetered to Menlo Park were revised (Eaton, et al., 1981). All of the northern California environmental networks were added to the central Coast Range network to form a combined northern California network. All stations were recorded on Develocorders, which were scanned to identify events for further processing. The scan lists were supplemented by events from an improved computer-based, real-time processor (RTP) (Allen, 1978, 1982), which detected and located many events in dense parts of the network that fell below the threshold for hand processing. Events continued to be timed by hand from film projected onto a tabletop digitizer. Earthquake phase lists were supplemented selectively by RTP data. With these changes, the



A. Epicenters of earthquakes in central California with $M \geq 1.5$ during 1976.



B. Epicenters of earthquakes in central California with $M \geq 1.5$ during 1977.

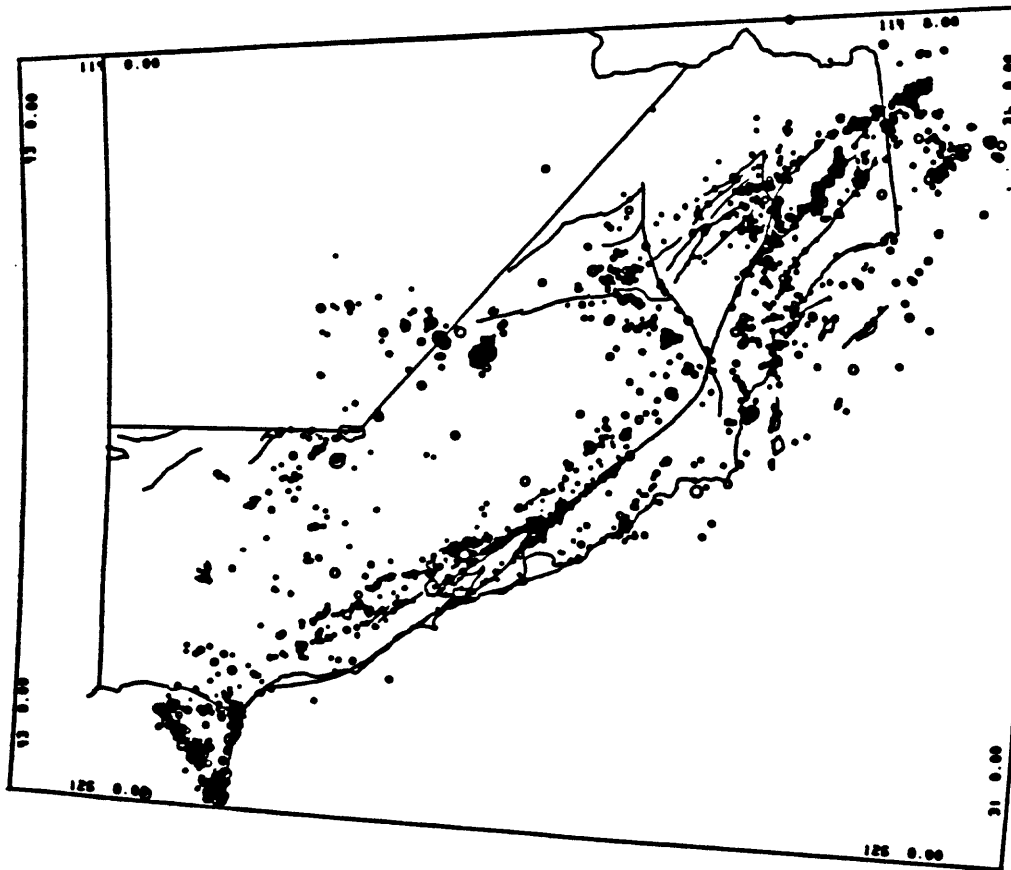
Figure 4. Epicenters of earthquakes in central California.

northern California network took on the character of a true regional network; and by 1982 the number of stations telemetered to Menlo Park exceeded 300.

In southern California, early special-purpose telemetered environmental networks were installed as follows: 1969 - Santa Barbara Channel; 1971 - Los Angeles Basin; 1973 - Oxnard/Ventura Basin and Imperial Valley; 1974 - eastern Mojave Desert. An agreement between the USGS and Caltech for cooperation in the operation and analysis of the southern California nets led to integration and further expansion of the network from 1975 onward. A computer-based system for recording and analyzing the network data was developed at Caltech by Carl Johnson during the late 1970's (Johnson, 1979). By 1982 the number of southern California stations recorded and analyzed at Pasadena exceeded 200.

In an attempt to present a broader picture of California/Nevada seismicity than was possible from the isolated regional networks, summary seismic results for the years 1978-1981 were combined from the four contiguous networks in northern California (USGS, Menlo), southern California (Caltech/USGS, Pasadena), central Nevada (UNR, Reno), and southern Nevada (USGS, Denver). The catalogs were combined to provide best coverage, without overlap and duplication of events, of the four subnet regions; and yearly seismicity maps for the California/Nevada region were prepared. The maps for 1980 and 1981 (Figure 5), when the networks were most extensive, were most interesting. These maps showed the seismicity associated with the entire San Andreas fault system in some detail - from Mexico to Cape Mendocino and from the Pacific Ocean to western Nevada, and they helped to put seismicity of individual parts of the region in better perspective with that of the region as a whole (Eaton, 1982). They also showed that the network was too sparse

CAL-NEV QUAKES 1980
M>1.5



CAL-NEV QUAKES 1981
M>1.5

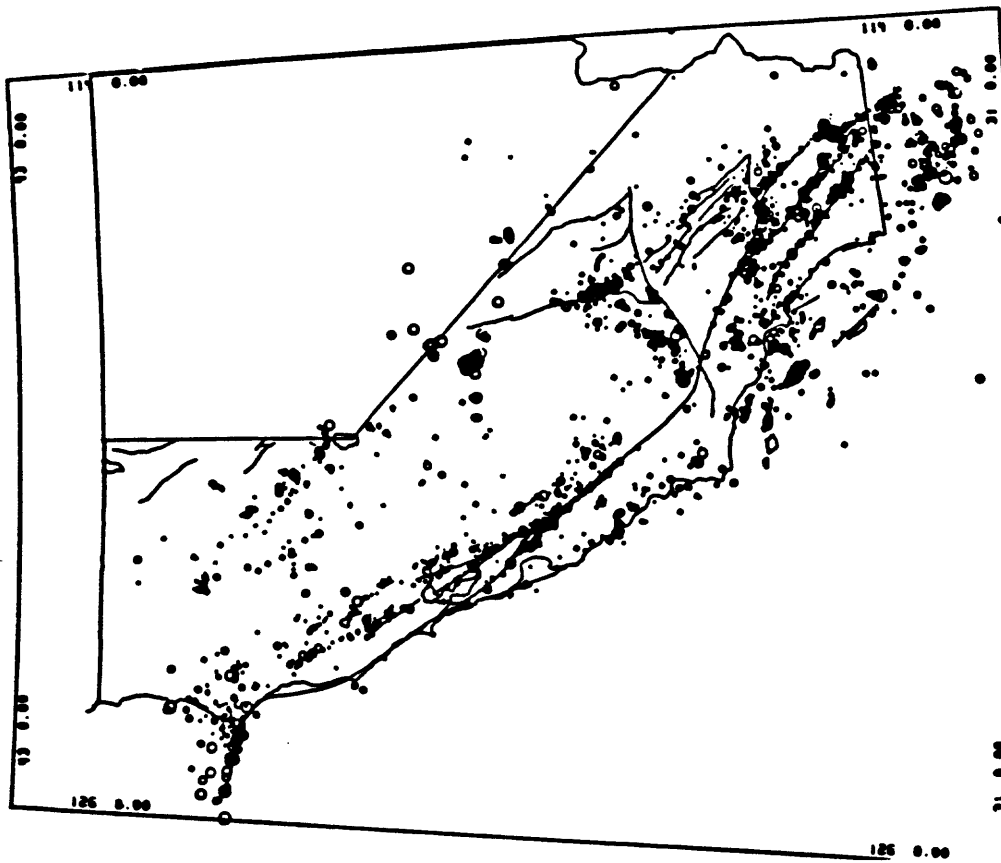


Figure 5. Epicenters of earthquakes in California and Nevada, 1980 and 1981.

in the Great Valley and southern Sierra Nevada to delineate the seismicity in those regions.

The most significant change in the networks after 1982 was the application of the CUSP computer-based recording and analysis system to the northern California network in 1984. That system, which is an outgrowth of the earlier system (CEDAR) developed by Carl Johnson at Caltech for the southern California network, greatly simplifies the collection and analysis of network data. The entire network is digitized and screened by computer for the occurrence of earthquakes in real time. Only the portions of the record corresponding to detected earthquakes are preserved; so the CUSP system requires better network configuration and performance to avoid loss of earthquakes than did the older procedure based on hand analysis. Although the analog FM signals of the entire network are still recorded on magnetic tape so that missed events can be recovered, the tape recorders and associated playback equipment are obsolete and expensive to maintain and use; so that backup facility must be updated, or it will be lost eventually.

Since 1982, network expansion has been limited mostly to small environmental networks that reduce the size of holes in the net or extend it a little farther into seismically active regions around its margins. The largest addition was the network in the Long Valley region to monitor seismicity in Long Valley caldera and the surrounding region.

E. Impact of network results versus network coverage

The examples of results from the network at successive stages in its development summarized above show that the scope of problems addressed by the network expanded rapidly as the network grew and its analysis became more comprehensive. The limited

portable network studies at Parkfield (1966) (Figure 1) and Bear Valley (1967) (Figure 2) demonstrated the resolution of a dense network and showed details of earthquake processes on small sections of individual faults. The prototype telemetered network between Palo Alto and Hollister (1968-1969) (Figure 3) resolved activity on individual faults at the junction of the San Andreas, Sargent, and Calaveras faults. The telemetered strip network between Clear Lake and Parkfield (1976-1977) (Figure 4) documented the very different seismic behavior of locked and creeping sections of the San Andreas Fault and placed them in the context of seismicity in adjacent parts of the Coast Ranges. Even though the network was 500 km long and 100 km wide at that time, it covered only a fraction of the greater San Andreas fault system; and it offered limited insight into the broader relationships among the tectonic elements composing that system.

A much more comprehensive picture of seismicity and the associated crustal deformation emerged when the results of the contiguous California and Nevada networks were combined and plotted together for 1980 and 1981 (Figure 5). Contrasting tectonic styles across the region were matched by contrasting patterns of seismicity. The slipping sections of the major faults were outlined clearly on the annual seismicity maps, but patterns of seismicity in less active regions were not, however.

By the end of 1986, the northern and southern California networks had operated with few changes in station configuration for seven years. Combined maps of earthquakes from the California and Nevada networks for 1980-1986 resolved patterns of seismicity that were not clear on the annual plots. The 1980-1986 seismicity maps and supporting catalog were analyzed and compared with the principal tectonic features of northern California by Eaton

(1989) (Figure 6) and of all of California by Hill, Eaton, and Jones (1990) (Figure 7).

These two papers deal primarily with aspects of the catalog that document the seismicity (and, by inference, the deformation) of the entire San Andreas fault system and its major tectonic subdivisions. Analyses at such a scale are required to place sections of the faults that generate M7+ earthquakes in context with the complex system of which they are parts.

The networks serve interests with a broad range of spatial and temporal scales. The comprehensive regional coverage coupled with the timely, systematic analysis of their data place the microearthquake networks first among our tools for detecting and interpreting significant events and trends within the fault system as well as for preserving a detailed historical record of them. The seismic and strain networks fulfill a statewide observatory function by capturing and preserving the earthquake and strain histories associated with the ongoing movement between the Pacific and North American plates and the inexorable preparation for future major earthquakes. The single thing that we can do today that our successors will not be able to do better is to record and preserve those histories. The cost of failing to do so could be years, perhaps decades, of unnecessary delay in developing a sufficient understanding of the San Andreas fault system to permit prediction of major events within it.

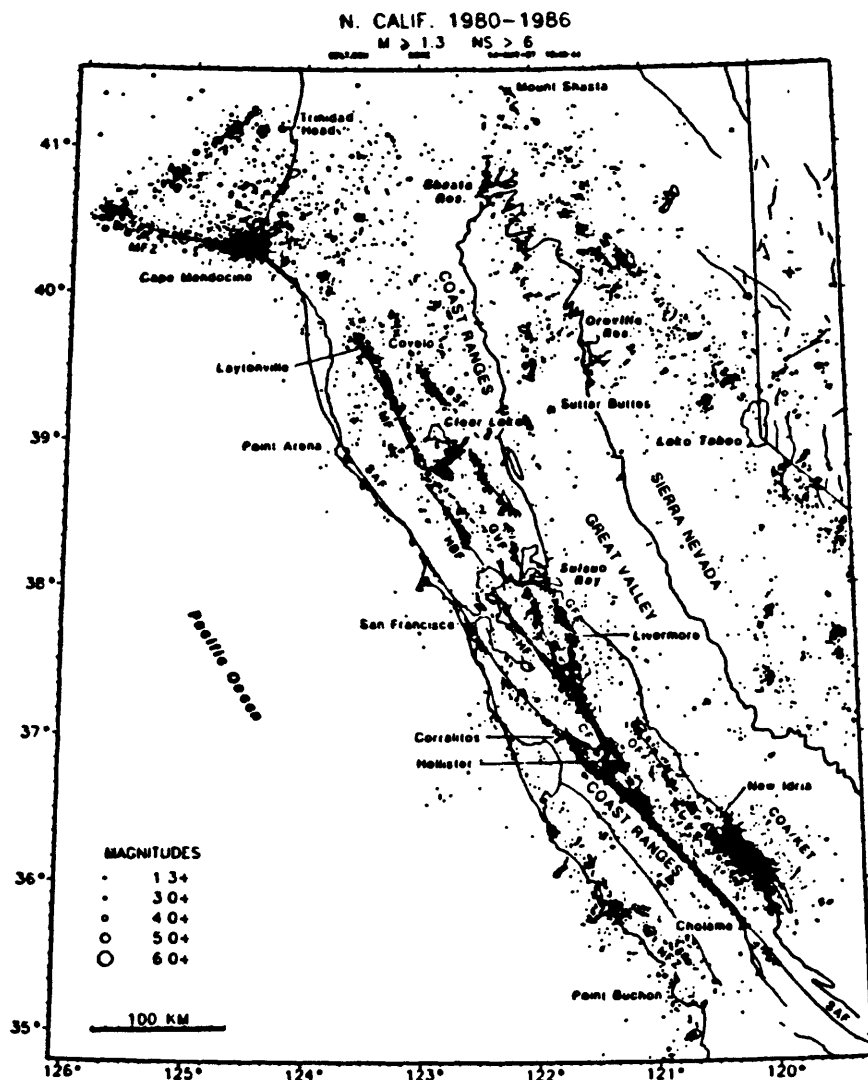


Figure 6. Northern California seismicity: 1980-1986. Symbol sizes are scaled according to magnitudes. Only events with magnitudes greater than or equal to 1.3 and with seven or more stations in the hypocentral solution were included in the plot. Abbreviations: SAF = San Andreas fault, NFZ = Nacimiento fault zone, OF = Ortigalita fault, CF = Calaveras fault, HF = Hayward fault, GF = Greenville fault, GVF = Green Valley fault, BSF = Bartlett Springs fault, HBF = Healdsburg fault, MF = Maacama fault, MFZ = Mendocino fracture zone, COA/KET = Coalinga/Kettleman aftershocks region.

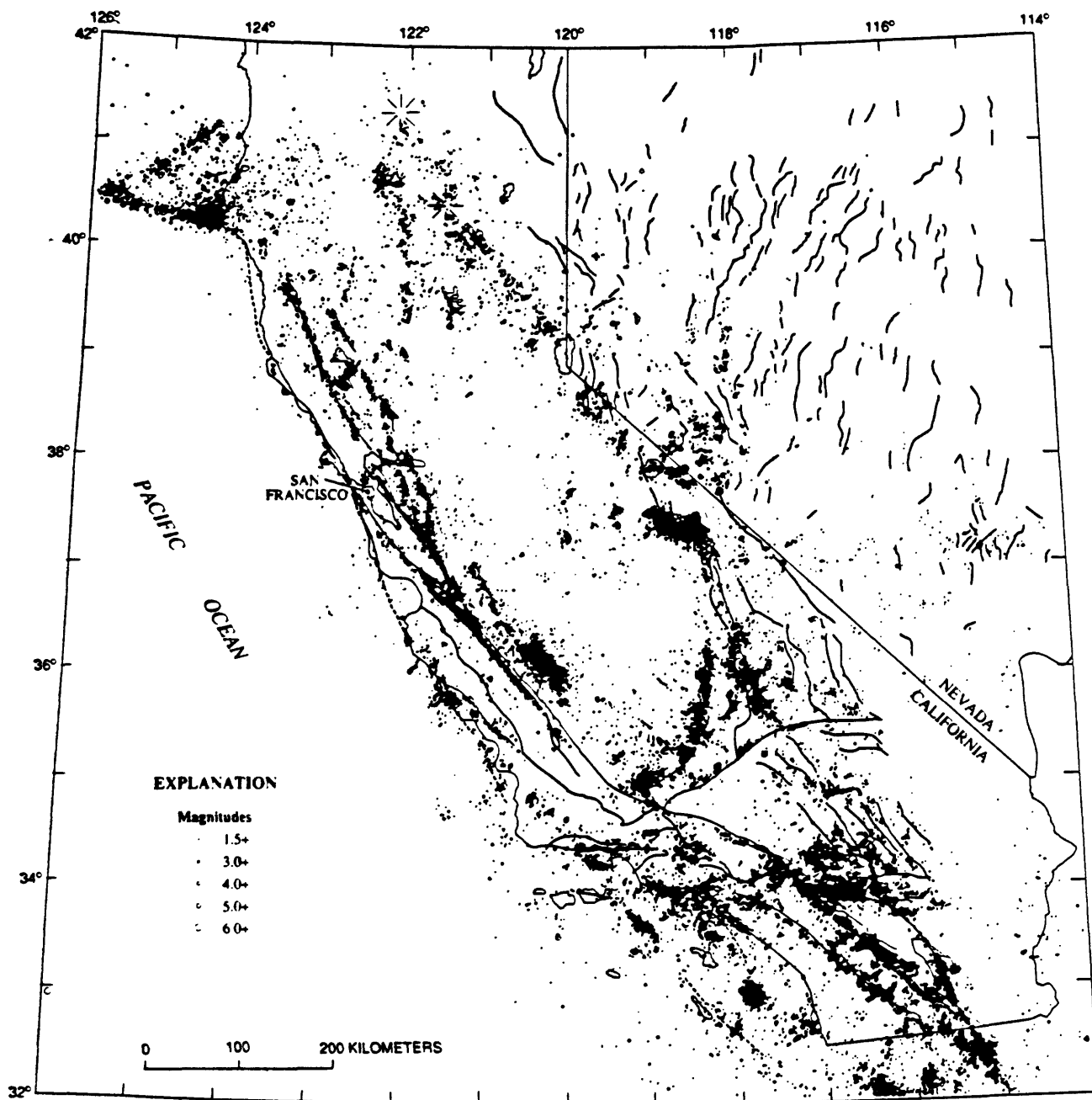


Fig. 7. Locations of 64,000 $M \geq 1.5$ earthquakes in California and western Nevada during 1980-86 and mapped Holocene faults (dotted where concealed; major branches of the San Andreas fault system marked in red).

II. FACTORS UNDERLYING THE DESIGN AND IMPLEMENTATION

OF THE NORTHERN AND SOUTHERN

CALIFORNIA SHORT-PERIOD SEISMIC NETWORKS.

For more than 20 years the U.S. Geological Survey has been a leader in the development and application of modern microearthquake networks for detailed studies of geologic processes in the earth's crust. Although this work had important beginnings at HVO in Hawaii and in the Crustal Studies Branch in Denver, it has been pursued most vigorously under the earthquake prediction research program within the Office of Earthquake Studies in Menlo Park. Selection of the seismic systems and network configuration employed has been driven by a combination of factors, including:

- 1) The USGS mission to monitor and elucidate active geologic processes in the crust, such as volcanic activity and active faulting, at a scale commensurate with that of regional geologic framework mapping and analysis,
- 2) the amplitude and spectral characteristics of seismic signals from small earthquakes ($1 < M < 3$) in relation to background microseisms and cultural noise in the regions studied,
- 3) the number, quality, and distribution of observations required to obtain the needed precision in epicenter location and focal depth of shallow earthquakes ($0 < h < 15$) in the heterogeneous earth's crust,
- 4) the intrinsic limitations of the instrumental components and communications systems available for use in the system (cost and complexity have been important considerations in determining what was "available"),
- 5) the experience and skills of the staff available to install and maintain the network,
- 6) the level of funding available to install the network and to support its ongoing operations.

Regional networks like those in California and Hawaii could not have been developed

without the advances in electronics and telemetry that have occurred over the last 25 years.

The early telemetered networks, such as LASA, that were employed in nuclear test detection and the sophisticated multichannel seismic systems developed for petroleum exploration were particularly stimulating and helpful. The defining characteristics of the regional networks, however, (seismic response and number and spacing of the stations) have evolved in response to the tasks to which the developing regional networks were applied.

The development of the network and refinement of its characteristics went hand in hand with the development of the seismological research based on its records. Attributes of the regional networks that have been found to be vital for detailed seismicity studies include:

- 1) the system frequency response and gain permit the recording of background earth noise (and everything larger) in the frequency range of about 1 Hz to 20 Hz where small earthquakes ($M < 3\pm$) have the best signal to earth noise ratio. The shape of the response curve approximates the inverse of the quiet site earth noise amplitude spectrum at frequencies above about 0.2 Hz, so the limited dynamic range of the system is utilized effectively.
- 2) the spacing of stations in the network is dense enough so that earthquakes above the network threshold (about M 1.5) are recorded at 6 or more stations to insure enough redundancy to avoid gross location errors. The small station separation is also extremely important for determining reliable focal depths for shallow earthquakes.
- 3) earthquake detection and location thresholds are low enough that the relatively frequent small events in the network can be used to delineate seismogenic structures in a reasonably short time.
- 4) the networks cover large regions with relatively uniform density, so major seismogenic structures such as the San Andreas fault system from Mexico to Cape Mendocino can be studied in their entirety.

Earthquake focal depth plays a special role in the design of regional networks. Focal depth is the most difficult hypocentral parameter to determine reliably; and it depends most critically on network geometry (particularly the distance to the nearest station) and crustal

model. Experience has shown that at least one station at an epicentral distance of one focal depth or less is required for a reliable depth determination. Because California earthquakes rarely exceed 15 km in depth and most are less than 10 km deep, station separations of 10 km or so are needed. It appears that a regional network adequate to monitor the San Andreas fault system should cover virtually all of California. If such a network had a station spacing of only 10 km, more than 4000 stations would be required. Because so many stations appears to be an impractical goal, we must seek a distribution of stations that provides adequate coverage in critical regions, and relaxed coverage elsewhere, with a smaller number of stations. Such a modified network derived by selective augmentation of the present northern and southern California networks would have about 800 stations. If uniformly distributed, an 800 station network covering all of California would have an average station separation of about 23 km.

Another critical issue is the choice of seismic system for the network. That choice must depend on the primary uses the data will serve, on the spectral characteristics of the earthquakes studied and of the background noise, and on the limits on wave propagation imposed by the earth's crust. The frequency response and sensitivity of the standard system employed in the USGS networks have been shown to be well suited to recording M1 to M5 earthquakes in California (Eaton, 1977, 1989). The limited dynamic range of the telemetry system (40 to 46 dB) is a problem that has been offset, in part, by operating a sparse subset of dual-gain stations in the network.

Another issue is the complement of instruments in the stations. Ideally, we would like to record all three components of ground motion at each station, but the number of

components in the network would be unmanageably large if we were to do so. The reasons for recording the horizontal components are 1) to improve the resolution of S waves, 2) to obtain horizontal component amplitudes for computing local magnitudes, and 3) to obtain all three components of ground motion to support further analysis of the recorded waves. These purposes do not require the density of stations that is needed to determine reliable focal depths, however.

Clear S wave arrivals at one or more relatively near-in stations are extremely helpful in determining origin time; and for events outside the network, S wave arrival times are essential for determining accurate epicenters as well. Because S waves stand out most clearly on the seismograms in the distance range of direct arrivals (epicenter to 50 km or so), it is desirable to have one or more stations with horizontal components within that range. Detecting S waves on the records also depends on having sufficient dynamic range so that the record is not "clipped", which makes secondary phases virtually impossible to pick.

The subset of NCSN stations with horizontal component systems operating at 42 dB attenuation has proved to be very effective in providing readable S wave arrivals for M2- to M3+ earthquakes. These systems also provide on-scale amplitude measurements for M2- to M5+ events (the larger ones are on-scale only at larger recording distances). Still lower gain (or higher dynamic range) systems are needed to obtain S wave arrivals at short distances for earthquakes larger than M3.5 or so.

Yet another important issue is the telemetry system employed by the network. Digital telemetry would provide much better dynamic range (96 dB or more) than the FM analog system currently used (40 to 46 dB). The lower cost and greater flexibility of the FM

system made it ideal for the early network that was recorded on Develocorders (<40 dB dynamic range) or analog magnetic tape (about 50 dB dynamic range). When computer based recording and analysis was introduced, however, the FM telemetry system was found to limit the overall dynamic range of the system unnecessarily.

Digital telemetry has several practical drawbacks compared with FM telemetry of the analog signals, however. Combining digital signals from several sources in the field is complicated and expensive, and each digital channel requires greater bandwidth in the communications system than does each FM channel. The advantage of FM telemetry is greatest with single component stations: signals from 8 stations can be combined in the field for transmission via one microwave or telephone channel to the central recording facility by means of simple summing amplifiers. For the multi-component stations used in NCSN that generate four analog signals the advantage of FM over digital telemetry is much reduced. One microwave channel can carry the signals from one 3-component digital station (16 bits at 100 sps per channel) or from two 3-component analog stations (8 channels at 40 to 46 dB dynamic range).

The foregoing analysis suggests the use of a hybrid network that employs analog FM telemetry for the many simple vertical component stations required to insure reliable focal depths and digital telemetry for a subset of 3-component stations, operating at slightly lower sensitivities, that will insure recording of readable S waves and on-scale maximum amplitudes for quakes in the M2+ to M5+ range.

The general structure of our telemetry communications system will readily support such a hybrid network. USGS and cooperating agency microwave systems form the

backbone of the system, and VHF (and UHF) radios bring signals from field sites to the microwave towers. The microwave system carries a sufficient number of channels that a modest number of channels (40 +/-) in both northern and southern California could be devoted to digital stations whose data would be telemetered continuously to the recording site for time stamping and recording.

III. CURRENT STATUS OF THE NORTHERN AND SOUTHERN CALIFORNIA REGIONAL NETWORKS

Both NCSN and SCSN have remained incomplete since their development was arrested in 1982. At that time several factors combined to stop network development: 1) the cost of maintenance, telemetry, and analysis reached the limit that could be sustained by available funding; 2) the analysis systems were saturated by records from stations already operating; 3) the impact of network results had not been felt fully because papers describing those results were slow to appear; 4) there was general concern over signal quality, dynamic range, bandwidth, etc., as well as the lack of reliable magnitudes computed from network records. Unfortunately, both networks had been deployed somewhat opportunistically as region-specific or topic-specific funds were available; and the final states in which both networks were frozen in 1982 were somewhat illogical and unbalanced with regard to coverage, density, and distribution of components.

Many improvements in network equipment and analysis have been made over the last 10 years. These include;

- 1) increased use of microwave telemetry and vhf/uhf radio links has greatly expanded network telemetry range and capacity while reducing its cost,
- 2) improved field units with solar power supplies have improved dynamic range and reduced maintenance visits to field sites,
- 3) pre-recording digitization of network seismic events has largely eliminated the delay, work, and expense of dubbing events from 5 analog tape recorders onto a single library tape for eventual digitization and analysis,
- 4) analysis of digitized events in CUSP is much faster, more accurate, and more comprehensive than the hand reading and analysis previously carried out.
- 5) methods for computing amplitude and duration magnitudes, MX and MF, have been developed and evaluated (Eaton, 1992); and they have been implemented in HYPOINVERSE (Klein, written communication) for routine use,
- 6) the effectiveness of the RTP for providing near-real-time monitoring of events in an aftershock sequence has been proven resoundingly. The ability of the network, through RTP analysis, to provide such monitoring is of vital importance for crisis management after a major earthquake,
- 7) many papers documenting network results have now been published; and those papers have established NCSN and SCSN as the primary sources of information on the seismicity and current tectonics of California (Oppenheimer, et al., 1992).

The problems that halted network deployment in 1982 have been mostly overcome.

Moreover, the earthquake catalog and research papers based on network results, as well as the development of the equipment and analytical procedures required to record and interpret the network data, rank among the very best accomplishments of the earthquake program. It is, therefore, appropriate to identify deficiencies of the present networks and to discuss how those deficiencies might be remedied.

Status of NCSN

For a variety of reasons the distribution of stations in NCSN is very uneven. The original "prediction" network built up between 1969 and 1974 consisted of 30-km-wide strips of stations along the San Andreas, Calaveras, and Hayward faults between Clear Lake and Cholame. This network was designed to "map" earthquakes that occurred on or very close to these faults, and average separation of stations was only about 10 to 15 km.

Further development of NCSN was far less orderly than that of the core network described above. It proceeded along two rather different lines that reflected sources of funding. First, funding from non-prediction sources became available to install and operate small special purpose monitoring networks, some of which were near enough to the core network to be treated as part of NCSN. Such networks included NTS (discontinued), Santa Barbara Channel (transferred to SCSN), Coso (transferred to SCSN), Geysers, Warm Springs Dam, Melones Dam, Auburn Dam, Berryessa Reservoir, Lassen Volcano, Shasta Reservoir, Shasta Volcano, and Long Valley Caldera. Second, as the catalog of earthquakes recorded by the core network and special networks took shape, it became clear that important seismicity extended well beyond the limits of the core network; so prediction funds were used to extend the core network laterally to cover the width of the Coast Ranges, southward to include the 1857 break, and northward to include the Cape Mendocino region (the latter using COE microwave telemetry). A cluster network was installed around Oroville Reservoir following the 1975 Oroville earthquake, the Coso network was extended westward across the southern Sierra Nevada (Walker Pass net, transferred to SCSN), and a sparse Central Valley/Sierra Foothills net (discontinued because of high telemetry costs) was set up between Modesto and Merced. Station separation in the fill-in networks funded from both sources was commonly more than double that in the core network. When the network deployment moratorium took effect in 1982, there remained several large holes in NCSN station coverage as well as the need to increase station density in parts of the network where computed focal depths were unreliable.

Signals from 27 stations operated by other institutions (LLL, DWR, UCB, and UNR)

are also telemetered to Menlo Park and processed with the USGS stations. The number of stations in the combined NCSN now recorded in Menlo Park is about 370. In addition, 33 stations from the north edge of SCSN are recorded and processed with NCSN, bringing the total number of stations recorded in Menlo Park up to about 400 (Figure 8, upper half).

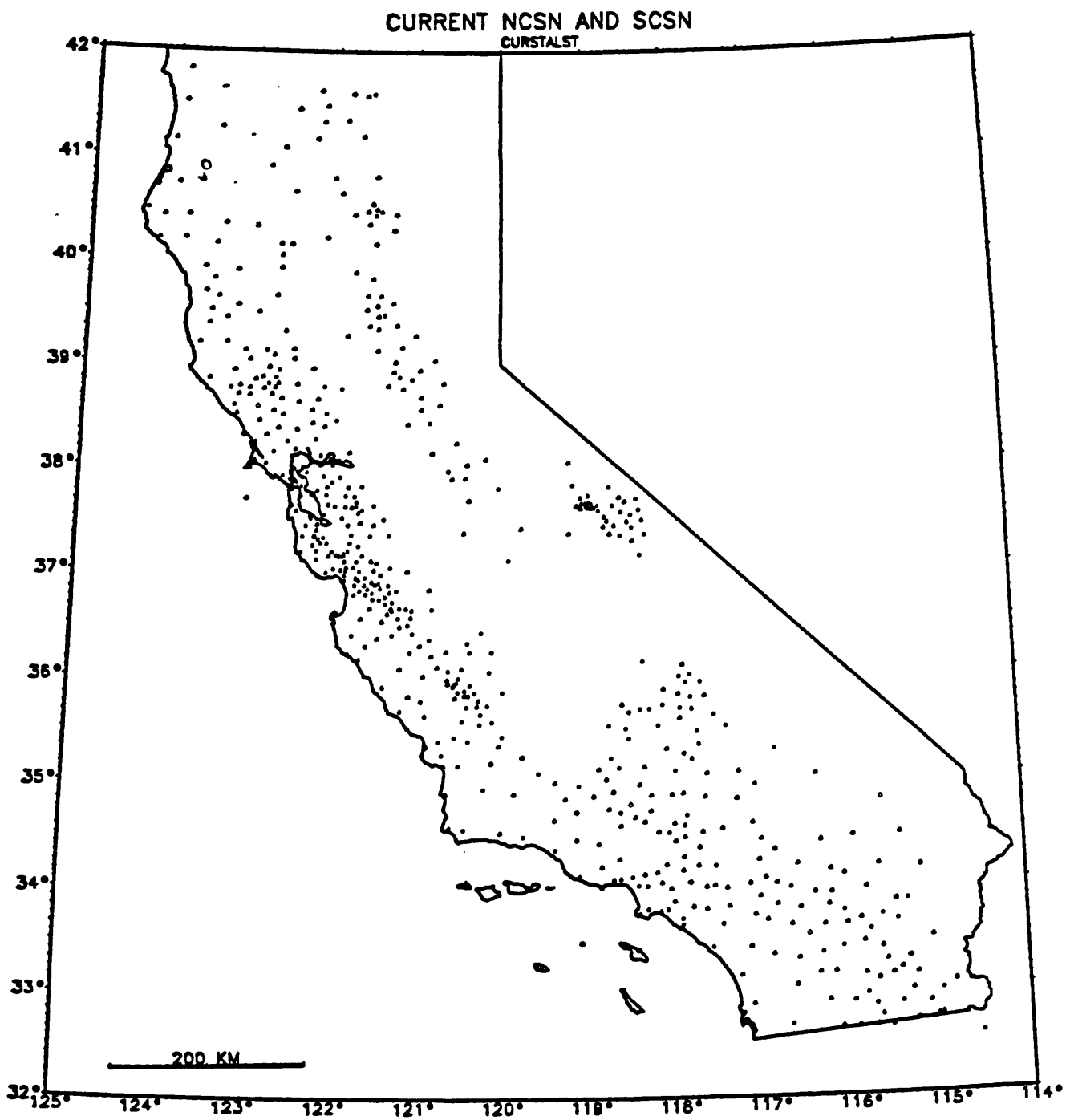


Figure 8. Current NCSN and SCSN stations.

Status of SCSN

The development of SCSN began in 1969 as a piecemeal augmentation of the broad 20-station telemetered Caltech network that had grown over the previous 40 years or so. From the first, however, SCSN took on a character rather different from NCSN. Well defined, narrow linear zones of seismicity were not nearly as apparent in southern California as in northern California; so stations were spread more uniformly over broader areas than in the core of NCSN. Specialized networks were installed approximately as follows:

- 1969 6 stations around the Santa Barbara Channel
- 1971 7 stations around the Los Angeles Basin (Caltech)
- 1973 15 stations in Imperial Valley
 8 stations in the Ventura/Oxnard region
- 1974 17 stations in the eastern Mojave Desert

Beginning in 1975, the USGS/CIT joint effort to complete the network systematically was undertaken.

- 1975 17 stations San Bernardino Mountains
 9 stations Coso Range
- 1976 4 stations Elsinore fault region
 8 stations Carrizo Plains
 13 stations San Bernardino Mountains
- 1979 12 stations Southern Sierra Nevada (Walker Pass)
 5 stations Mojave Desert
- 1981 6 stations Elsinore fault region
 10 stations Mojave Desert
 10 stations Imperial Valley
 13 stations San Bernardino Mountains
 7 stations Transverse Ranges
 5 stations Walker/Coso nets (China Lake)
- 1982-1987 12 stations

Twenty four stations of the Caltech network as well as 11 stations of the USC Los Angeles Basin network (primarily downhole) are also telemetered to Pasadena and analyzed with the USGS stations. Over the years about 30 southern California stations have been discontinued because of the high costs of telemetry and maintenance. The number of stations in the combined SCSN now recorded at Pasadena is about 200. Moreover, 14 stations along the south edge of NCSN are recorded and processed in Pasadena (Figure 8, lower half).

Although station coverage appears to be more uniform in SCSN than in NCSN, it is also much sparser, on average. The most glaring deficiency of coverage in SCSN is the absence of telemetered stations in Owens Valley. Other regions with seriously inadequate coverage are the Elsinore fault to Pacific shore belt and the eastern Mojave/Basin-and-Range boundary region. Moreover, station density over large areas is too low to support reliable focal depth determinations or focal mechanism determinations.

IV. PRINCIPAL FUNCTIONS OF THE REGIONAL NETWORKS, AND DEPENDENCE OF THEIR PERFORMANCE ON NETWORK CONFIGURATION

A. Network purposes

Although the short-period seismic networks in California support a wide range of monitoring and research objectives, their primary purposes are:

- 1) long-term monitoring of local earthquakes throughout the broad zone of seismicity associated with the San Andreas and related fault systems:
 - a) to construct a uniform, long-term earthquake catalog (with supporting phase data and seismograms) to document seismicity of the region,
 - b) to map seismogenic zones and to identify the geologic structures and styles of deformation

with which these zones are associated,
c) to provide a basis for monitoring spatial and temporal variations in seismicity that might presage major earthquakes in the region,

- 2) detailed monitoring and determination of precise hypocentral, magnitude, and focal mechanism parameters of earthquakes along sections of major faults that are expected to produce damaging earthquakes within a decade or so,
- 3) real-time monitoring and analysis of earthquakes to provide timely, reliable information on their locations and magnitudes for crisis management after large earthquakes and to fill the need for general public information on "felt" earthquakes at any time.

Important additional research based on regional network records include:

- 1) determination of improved velocity structures of the lower crust and upper mantle to refine the analysis of local earthquakes,
- 2) tomographic studies of the crust and mantle beneath the network to clarify the relationship of current and past plate tectonic regimes to major structures and seismic zones of the region,
- 3) array analysis of teleseismic body waves to refine our understanding of the velocity structure of the deep interior of the earth.

B. Dependence of network performance on configuration

Network design requirements for fulfilling its primary purposes differ principally in the allowable distance between contiguous stations. This parameter plays a critical role in the calculation of focal depths and in establishing magnitude thresholds for event detection and focal mechanism determinations.

Focal depths

The need for accurate focal depths of events less than 10 km deep sets the most stringent requirement on station spacing. To map out locked patches on a fault surface like the one filled in by the Loma Prieta quake or the one expected to be filled in by the next Parkfield quake, station separation along the fault should be 10 km or less. For station

spacing of 20 km, which insures that no event will be farther than about 10 km from the nearest station, we should be able to determine whether earthquakes are in the lower crust (> 10 km), middle crust (5 km to 10 km), or upper crust (< 5 km); but likely errors in depth for events shallower than 10 km will be quite large. For station spacing of 40 km we should be able to distinguish between quakes in the lower crust or upper mantle and those at mid- or upper-crustal depths. The greater the spacing of stations, however, the stronger will be the dependence of calculated focal depth on the crustal model.

Event detection

Network requirements to insure detection of small events depend on the manner in which the events are detected. An analyst scanning appropriate seismograms can identify an earthquake (or blast) if it is recorded by a single station. Computer detection of events from the network requires that some simple algorithm (e.g. variation in the short-term/long-term ratio of average trace amplitude) be able to detect an "event" more or less simultaneously at a minimum number of stations in the same region. Commonly, that number is set at about 6 to suppress false triggers due to local noise at individual stations.

The number of stations triggered by a small event depends on event magnitude, station spacing, and background noise at the individual stations. As a practical approach, examination of a suite of earthquakes analyzed on CUSP shows that an earthquake of magnitude M1.5 can be read out to different distances in different regions: about 40 km in the central Coast Ranges, about 30 km in the Geysers region, about 50 km in the Cape Mendocino region, and about 60 km in the Lassen/Sierra region. For a square grid of node

spacing L , a circle of radius $1.5xL$ encloses between 4 and 9 nodes; and the probability that it will enclose between 6 and 8 nodes is very high (the area of a circle of radius $1.5xL$ is $7.07xL^2$). Thus, to assure a high probability of recording an $M1.5$ event at 6 or more stations of a network laid out as a square grid, the station spacing for the regions enumerated above should be 27 km in the central Coast Ranges, 20 km in the Geysers region, 33 km in the Cape Mendocino region, and 40 km in the Lassen/Sierra region. The foregoing logic applies to the detection and capture of an event by both the CUSP and RTP systems, but it does not promise that all captured events can be assigned reliable focal depths. For a region of high cultural noise such as the S.F. Bay area, the L. A. Basin, and the Great Valley, station spacing should be decreased to about 20 km to insure detection of $M1.5$ events.

Focal mechanisms

Determination of focal mechanisms sets somewhat different network requirements. For earthquakes of magnitude $M3.5$ and larger, arrivals in the P_n range (beyond 100 km to 120 km in the Coast Ranges) can be used; so rather distant parts of the network come into play. For smaller events, only arrivals within 100 km (perhaps 50 km for $M2$ events) are sharp enough to provide useful first motion data. To insure that observations adequately cover the focal sphere, a moderate number of stations (15 to 20) that are well distributed in azimuth and distance are required. For a square grid network with 25 km station spacing, a 75-km-radius circle centered on a station includes 29 stations within it; and a 50-km-radius circle on the same grid includes 13 stations. Thus, it appears that a homogeneous network with 25 km station spacing would support routine focal mechanism determinations of $M2$ to

M2.5 and larger earthquakes. The quality of focal mechanism solutions depends on focal depth, velocity model, and other factors in addition to the number of observations, however.

C. Comparison of regions of dense network coverage with regions expected to produce damaging earthquakes

The regions in the networks that have a station spacing of the order of 10 to 15 km required for the detailed mapping of the distribution of earthquakes at 5 km depth or less on seismogenic structures in the crust are: 1) a narrow 60-km-long strip along the San Andreas fault centered at Parkfield, 2) a 150-km-long strip along the San Andreas fault from San Benito to Los Gatos, 3) a 20 km by 50 km band of stations from the Geysers to Warm Springs Dam, 4) an 80-km-long cluster of stations from Mammoth Lakes to the north end of Owens Valley, 5) a small cluster of stations at the Coso Range, 6) a small cluster of stations on the San Andreas fault near Palmdale, and 7) a small cluster of stations in the Brawley seismic zone at the southeast end of the Salton Sea. In some of these cases, the network density falls off so rapidly away from the dense zones that the networks do not provide adequate coverage for focal mechanism determinations of M2 to M2.5 earthquakes.

Next, consider the regions that have been identified as having high probabilities of producing M6.5 and larger earthquakes in the next 30 years or so: S. F. Peninsula section of the San Andreas fault, both the southern and northern halves of the Hayward fault, Healdsburg fault, southern section of the San Andreas fault, San Jacinto fault, and the Los Angeles Basin (Figure 9). For the detailed monitoring that these regions require, the network should be augmented so that earthquakes can be mapped on the fault surfaces that

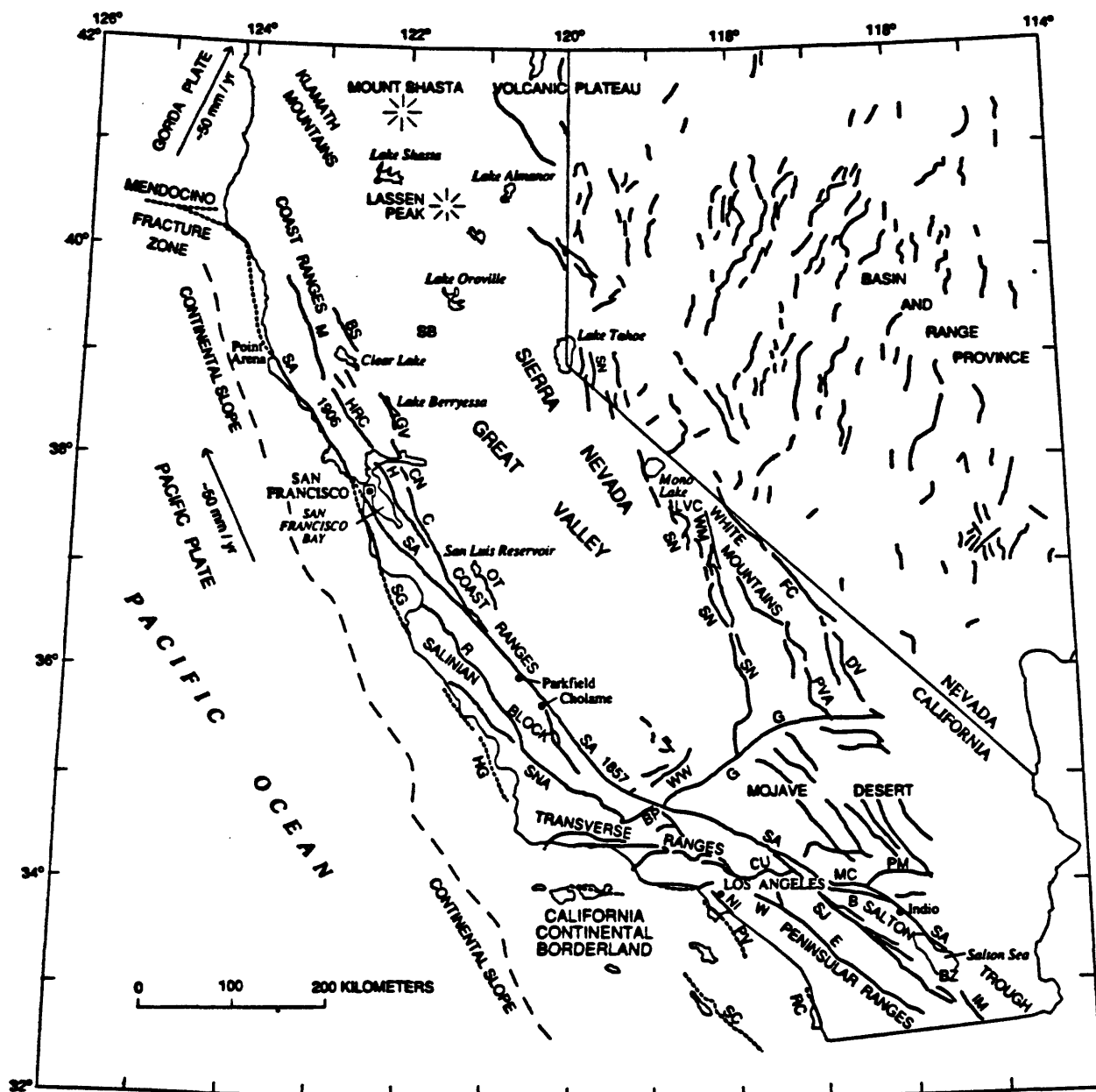


Figure 9. Map showing active faults in California.

are the presumed sources of the impending large quakes. The discussion of network capabilities versus station spacing developed above suggests the need for strip networks with station spacing of about 10 km along the faults flanked by broad areas in which station spacing is not greater than 25 km.

Outside of these immediate high-risk areas the network should be upgraded for more adequate long-term monitoring of earthquakes throughout the San Andreas and related fault systems. Specific targets should include sections of major faults that will produce future large quakes: San Andreas fault north of San Francisco and in the region of the 1857 Fort Tejon break, Sierra Frontal fault in Owens Valley, White Wolf fault, etc. The targets should also include regions of potential large earthquakes where the causative faults are not so obvious: west flank of the Coast Ranges southeast of San Francisco, Great Valley/Coast Ranges boundary at least from Winters to Lost Hills, zone of crustal convergence in the Santa Maria/Santa Barbara/Ventura/San Fernando region, Mendocino Fracture Zone and adjacent subduction zone north of Cape Mendocino, etc.

An overall objective of the broad regional network should be to refine and complete the picture of San Andreas seismicity presented in USGS PP 1515 (Figure 7). An accurate analysis of seismicity, tectonics, and crustal structure on that scale is needed for correlation with the rapidly accumulating information from VLBI and other space-based geodetic techniques on the nature and distribution of deformation in the Pacific Plate/North American Plate boundary zone. Joint analysis of long-term seismicity and deformation of the plate boundary zone is needed to document the accumulation of elastic strain in the source regions of future large earthquakes.

D. Network augmentation to improve coverage of the San Andreas Fault system

On the basis of the map of existing stations (Figure 8), the 1980-1986 seismicity map (Figure 7), the historic record of large earthquakes, and the considerations discussed above, proposed new stations were "added" to the short period seismic networks in California so that they might better meet the needs of the Earthquake Hazards Reduction Program. The needs of the northern and southern networks will be listed separately.

NCSN

Network subregions, number of proposed new stations, and approximate maximum station separations within these subregions are as follows:

Network Subregion	Number of new Stations	Maximum stn Separation
Central Coast Ranges	25	20-25 km
S. F. Bay Area: South	24	10-15 km
S. F. Bay Area: North	24	15 km
Northern Coast Ranges	15	20-30 km
Mendocino Region	14	30-40 km
Shasta/Lassen Region	11	20-30 km
Northern Great Valley	18	25-35 km
Southern Great Valley	21	35-40 km
Northern Sierra	10	30-40 km
Central Sierra	17	30-40 km
TOTAL	179	

In addition to the proposed new sites, all of which should have high-gain vertical seismometers, low-gain horizontal and vertical instruments should be scattered throughout the network to obtain better data for S arrivals and magnitudes. About 40 new low-gain (or high dynamic range) 3-component installations, some replacing single-component low-gain vertical

or horizontal components will be needed.

SCSN

Network subregion, number of proposed new stations, and maximum station separation within each subregion are as follows:

Network Subregion	Number of new Stations	Maximum stn Separation
Santa Barbara/Santa Maria	18	20-25 km
White Wolf	13	20-30 km
So. Sierra/Owens Valley	25	20-40 km
Garlock	9	15-30 km
Basin and Range Borderland	15	40-60 km
Eastern Mojave	17	20-40 km
So. San Andreas/San Jacinto	35	15 km
Ventura	10	15 km
Los Angeles Basin	14	15-20 km
Elsinor/San Diego	13	20-30 km
Offshore	5	20-60 km
TOTAL	174	

In addition to the proposed new sites with high-gain verticals, 40 low-gain (or high dynamic range) 3-component installations should be scattered throughout the network.

The proposed additional stations in both NCSN and SCSN are shown in Figure 10; and a map of the resulting combined network (existing stations plus proposed stations) is shown in Figure 11).

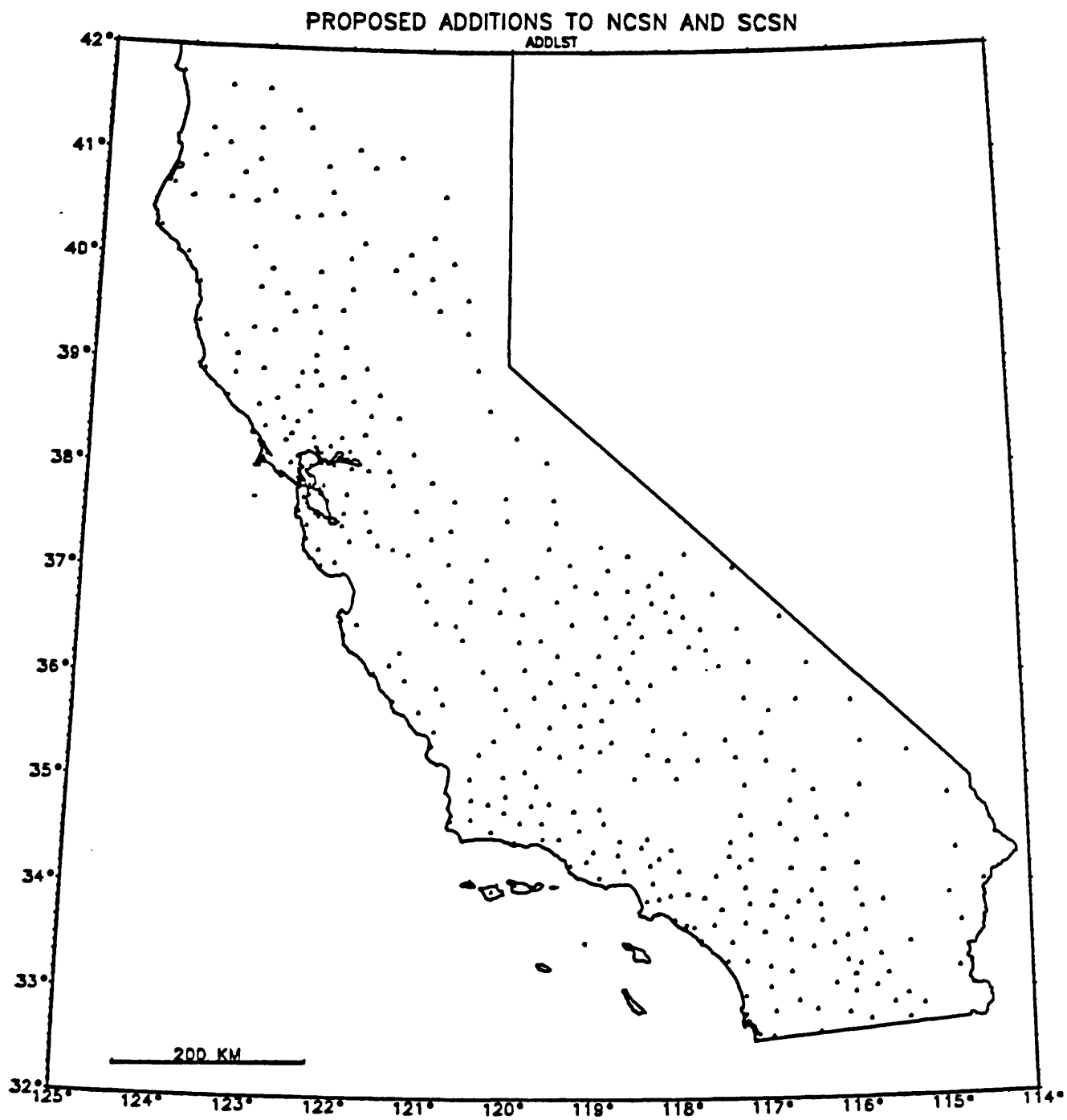


Figure 10. Proposed station additions to NCSN and SCSN.

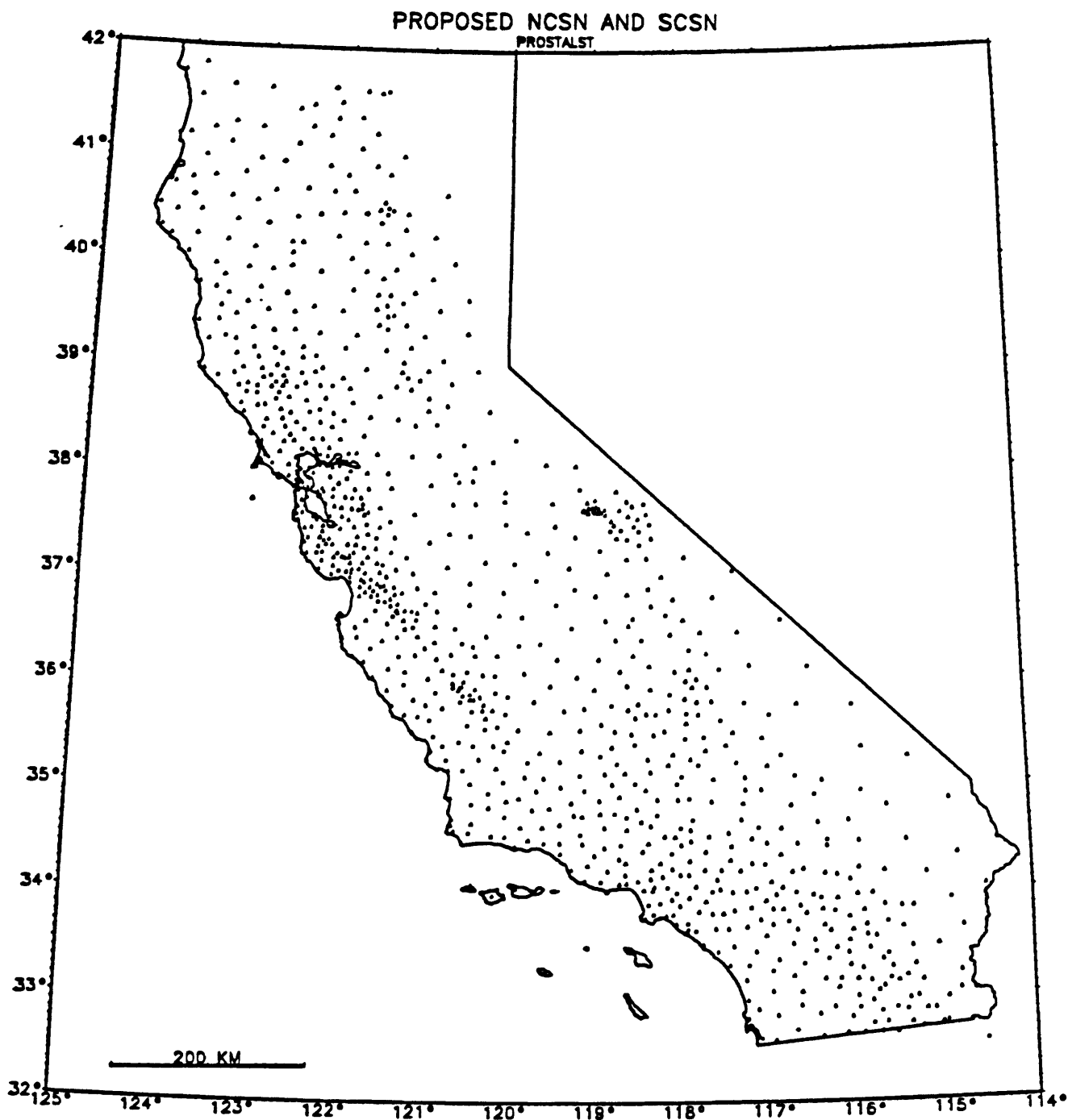


Figure 11. Existing and proposed stations for NCSN and SCSN.

V. FURTHER DEVELOPMENT OF NCSN AND SCSN

The major regional networks have attained a "footprint" that nearly covers the entire zone of seismicity associated with the San Andreas and related fault systems that mark the tectonically active boundary between the Pacific and North American plates in California. The quality of network coverage within that broad region varies considerably, and in some places it is clearly inadequate to fulfill the principal objectives of the network. The statewide map of seismicity in Figure 7 can even be said to be misleading. It suggests a degree of completeness that simply cannot be attained with the present networks. Because of the central role that the California regional microearthquake networks play in developing an understanding of California earthquakes, it is clearly necessary to address the inadequacies of the present networks and to make every reasonable effort to correct them. In decades to come our seismology program will be judged more critically on the quality and completeness of the record of California earthquakes that we pass on to our successors than on any other issue.

The strengths and weaknesses of the network have been described above on a region by region basis; and a general plan to add stations to attain the level of coverage appropriate for each region has been outlined. The overall network augmentation needed is quite large, about 350 additional high-gain short-period vertical-component analog stations plus about 80 three-component short-period digital stations, split about equally between NCSN and SCSN.

Experience over the last 20 years has shown that the task of upgrading the network is closely linked to the ongoing work of maintaining and operating the existing network. The knowledge, skills, and facilities required for both are the same; and changes to improve the

network must be integrated into the operation and analysis of the network as they are made.

To assess the impact of network expansion on the overall network enterprise, it is helpful to identify the primary activities that sustain the network and its operation.

- 1) telemetry - operation and maintenance of the microwave trunks and VHF/UHF radio feeder links,
- 2) seismic systems - operation and maintenance of the seismometers and preamp/VCO's in the field and the discriminators and signal distribution system in the recording center,
- 3) recording and analysis
 - a) backup recording of incoming network signals,
 - b) real-time detection and preliminary location of earthquakes to permit timely response during earthquake emergencies,
 - c) online computer detection of earthquakes and spooling of digitized seismograms,
 - d) offline interactive analysis of earthquakes,
- 4) archiving of seismograms and products of analysis to preserve these materials and to make them available to the seismology community for further exploitation and analysis.

Next, we shall examine how the proposed network augmentation depends upon and impacts these activities.

Recording and analysis

When the network was young, we were far more successful installing stations and gathering data than analyzing the data.

This problem grew more acute as the network approached its present size in the early 1980's. Heavy commitment to the development of improved digital data acquisition and analysis systems during the last 10 years has now tipped the balance in favor of analysis. The CUSP systems now operating in Menlo Park and Pasadena both have the potential capacity (depending on the A/D convertors) to record substantially more stations than they now are. Moreover, these systems are based on modern microcomputer "workstation"

equipment that is much less expensive and more reliable than the equipment used to record and analyze the early networks. In the near future even the backup network recording will be carried out digitally on inexpensive equipment, retiring the bank of half-a-dozen cumbersome, costly, high maintenance analog recorders that have performed that function for the last 20 years. Most impressive, however, is the relative efficiency of data processing in CUSP compared to that of earlier methods: the improvement approaches a full order of magnitude. Thus, the several hundred additional analog stations needed to fill out NCSN and SCSN could be recorded and analyzed on existing equipment with a minimum of additional effort and expense.

Seismic systems

The analog seismic systems employed in the network have been refined over the years to meet the most critical network requirements: simplicity, low cost, low maintenance, reliability, and good data quality (within the bandwidth and dynamic range permitted by analog FM telemetry). Augmentation of the network with this equipment would have a minimum impact on the cost of maintaining the network. One field maintenance technician can take care of about 100 stations. A fifty percent increase in the number of stations would require no increase in the manpower required to operate and maintain the discriminators and signal distribution systems in the recording centers.

The limited dynamic range of the analog FM telemetry system has been offset by the operation of a subnet of low-gain stations, many with three-component seismic systems, with the same frequency response as the high-gain systems. Development of a simple three-

component, 100 sps, 16-bit digital system to replace the low-gain analog systems is nearly complete. That system utilizes a standard 4800 baud communications channel that can be provided by our current microwave and VHF/UHF telemetry system. Time stamping and recording is carried out in a PC-based system, developed by the USGS, that should accommodate up to 48 independent 3-component stations. The data collected by this system will be combined with the CUSP digital network data so that all stations (digitized high-gain analog stations plus low-gain 3-component digital stations) can be analyzed in the CUSP system.

Telemetry

The networks were set up originally to operate over commercial telephone circuits. We were forced to change to a microwave and VHF/UHF radio based system because of excessive cost, inadequate areal coverage, and inadequate data quality of the commercial systems. The remaining long-distance phone circuits that we use will be replaced as soon as microwave facilities can be developed.

Fourteen microwave sites in the Coast Ranges between Eureka and San Luis Obispo constitute the communications backbone of NCSN, and 4 microwave sites in the L. A. Basin and Mojave Desert provide the core of the SCSN communications system. The northern Coast Range sites belong to COE, and the USGS maintains them on a reimbursable basis. The microwave system currently operated by the USGS spans about 1000 km and includes 18 sites. Our access to this system was developed by negotiation with COE, purchase and installation of key USGS links, and considerable self-education in the areas of microwave

electronics and transmission paths over the last decade. A large fraction of the network is now served by this system, but other parts of the network have been beyond its reach.

We have recently gained access to additional microwave facilities, by agreement with COE and FAA, that will provide improved, inexpensive telemetry for much of the rest of the network. The new system covers the Great Valley/Sierra foothills region and the Pasadena to Imperial Valley to southeastern Mojave Desert region. It will also provide a limited number of circuits between Menlo Park and Pasadena and between Menlo Park and Reno, which will replace some of our most expensive phone lines as well as facilitate better exchange of data among these recording and analysis centers. Addition of these new facilities virtually doubles the length of microwave trunk line and number of microwave sites in the overall system that serves the networks.

Although the microwave trunks do not reach the very ends of the networks, they have been "extended" effectively by means of broad-band VHF radio links that can carry four voice-grade channels. Such a system is now bringing stations in northeastern California into the Coast Range microwave system. Similar equipment could extend the southern California microwave system into Owens Valley and into the San Diego region.

In addition to microwave trunks, the network communications system employs several hundred 100-mw VHF and UHF transmitters and corresponding receivers. The low power of the transmitters and the relatively long transmission paths employed in the network, combined with the need for uninterrupted signal transmission, require great skill in the use of these radios.

The impact of our network telemetry system on network coverage, data quality, and

efficiency of data analysis cannot be overemphasized. In an important sense the telemetry system is the network, supplemented by seismic systems in the field and recording and analysis systems at the recording centers. Degraded telemetry leads not only to a serious loss of data but also to a huge increase in the time and effort required to process the noisy events that can be recovered. Assuring adequate maintenance for the telemetry system should have very high priority.

Archiving of seismograms and results of analysis

In the late 1960's when the USGS commenced network seismology in California, methods of preserving seismic data were those that had been used for 100 years: original paper or film seismograms were saved, lists of hypocenters and magnitudes were published in network bulletins, and records of phase arrival times, etc., were filed away for possible future use.

When the regional networks expanded from 15 or 20 stations to several hundred stations and paper or film seismograms were replaced by magnetic tape records, the old methods of preserving the data were completely inadequate. By the mid-1970's the results of analysis, both summary lists of hypocenters and the phase picks on which they were based, were preserved as ascii computer files on digital magnetic tape. The seismograms were preserved both on 16 mm film (Develocorders) and on analog magnetic tape. Recovery of seismograms from the analog tape can be carried out by equipment, now largely obsolete, that is available only in Menlo Park; and it is very time consuming. Moreover, there is considerable apprehension over the stability of the tape records.

From the mid 1980's, for NCSN (and the late 1970's, for SCSN), the primary records of both the results of analysis and the seismograms themselves have been saved on 9-track digital magnetic tape written by the CUSP system. Because the CUSP format is both unique and intractable, recovery of CUSP data has been carried out in a functioning CUSP environment.

Flexibility in analysis of network phase data has been achieved by constructing event phase files, in HYPOINVERSE or HYPO71 format, from CUSP "MEM" files. Summary files of hypocenters as well as the phase files are then preserved in monthly "directories" that are written to 9-track magnetic tape.

Recovery of the seismograms, however, still requires use of the CUSP system, which requires matching "GRM" and "MEM" files for each event recovered. The procedure is cumbersome and slow and has been used only on a limited basis. Alan Walter is currently working on a program to read the CUSP "MEM" and "GRM" files directly on the SUN computer. This program will facilitate access to network data for SUN and other non-CUSP users.

The lack of a uniform, "complete" catalog and supporting phase data has impeded setting up a routine procedure for filling data requests; so such requests have been filled on an ad hoc basis. This situation will improve markedly in the near future when Dave Oppenheimer and Fred Klein complete the massive reprocessing of the NCSN data set that has been underway for several years.

Long-term solutions to the data distribution problem currently are being pursued through cooperation with other institutions: Caltech, UC Berkeley, and IRIS (Seattle).

NCSN and SCSN data in the form of hypocenter summary lists, phase lists, and seismograms will be loaded onto mass-storage devices (eg. optical juke-boxes) and accessed via computer network or magnetic tape.

It took more than a decade to build the network to its present state. It took another decade to develop recording, analysis, and archiving systems that can cope with the data from the existing network; and those systems could handle a 50% increase in the network without significant problems. If we begin an orderly upgrading of the network at this time, the work could be completed before the end of the next decade. If we fail to complete the network, we shall pass an incomplete historical record of earthquakes to our successors and impair their ability to identify and quantify seismic hazards in a California that is even more populous and developed than now.

REFERENCES

- Allen, R.V. (1978). Automatic earthquake recognition and timing from single traces, Bull. Seism. Soc. Am., v. 68, pp 1521- 1532.
- Allen, R.V. (1982). Automatic phase pickers: their present use and future prospects, Bull. Seism. Soc. Am., v. 72, pp S225- S242.
- Bolt, Bruce A. (1989). One hundred years of contributions of the University of California seismographic stations, in Litchner, J.J., ed., Observational Seismology, Univ. of Calif. Press.
- Criley, Ed, and Eaton, J.P. (1978). Five-day recorder seismic system (with a section on the Time Code Generator, by Jim Ellis), U.S. Geological Survey Open-File Report 78-266, 86p.
- Eaton, J. P. (1971). Letter to Jan Foley, Assistant to Senator Alan Cranston, outlining plans for an earthquake prediction research network in California, Adm. Rpt.
- Eaton, J. P. (1976). Tests of the standard (30 Hz) NCER FM multiplex telemetry system, augmented by two timing channels and a compensation reference signal, used to record multiplexed seismic network data on magnetic tape, USGS Open-File Report 76-374.
- Eaton, J. P. (1977). Frequency response of the USGS short-period telemetered seismic system and its suitability for network studies of local earthquakes, USGS Open-File Report 77-884.
- Eaton, J. P. (1982). Reference material on results from regional networks in California and western Nevada, EHRP Asilomar conference on regional networks. Adm. Rpt.
- Eaton, J. P. (1985). Temporal variation in the pattern of seismicity in central California, in "Earthquake Prediction: Proceedings of the International Symposium on Earthquake prediction", held at UNESCO Hq., Paris, 1979.

- Eaton, J. P. (1986a). History of the development of the HVO seismograph and deformation networks. Adm. Rpt.
- Eaton, J. P. (1986b). Narrative account of instrument developments at HVO from 1953 thru 1961, and a summary of further developments from 1961 thru 1978. Adm. Rpt.
- Eaton, J. P. (1989). Dense microearthquake network study of northern California earthquakes, in J. Lithner, ed., "Observatory Seismology: a centennial symposium for the Berkeley seismographic stations", Univ of Calif Press.
- Eaton, J. P. (1992). Determination of amplitude and duration magnitudes and site residuals from short period seismographs in northern California, Bull. Seism. Soc. Am., v. 82.
- Eaton, J. P., M. E. O'Neill, and J. N. Murdock (1970a). Aftershocks of the 1966 Parkfield-Cholame earthquake: a detailed study, Bull. Seism. Soc. Am., v. 60, pp1151-1197.
- Eaton, J. P., W. H. K. Lee, and L. C. Pakiser (1970b). Use of microearthquakes in the study of the mechanics of earthquake generation along the San Andreas fault in central California, Tectonophysics, v. 9, pp 259-282.
- Eaton, J. P., R. Lester, and R. Cockerham (1981). Calnet procedures for processing phase data and determining hypocenters of local earthquakes with UNIX, USGS Open-File Report 81-110.
- Healy, J. H., Rubey, W. W., Griggs, D. T., and Raleigh, C. B. (1968). The Denver Colorado Earthquakes, Science, v. 161, pp 1301-1310.
- Hill, David P., Jerry P. Eaton, and Lucille M. Jones (1990). Seismicity, 1980-1986, Ch. 5, in Wallace, R. W., ed., The San Andreas Fault System, USGS Professional Paper 1515.
- Johnson, Carl Edward (1979). I. CEDAR - An approach to the computer automation of short-period local seismic networks, II. - Seismotectonics of the Imperial Valley of southern California: PhD Thesis, Calif. Inst. of Technology.

Klein, F. W., and Koyanagi, R. Y. (1980). Hawaiian Volcano Observatory network history: 1950-1979, U.S. Geological Survey Open-File Report 80-302, 84 p.

Louderback, George D. (1942). History of the University of California Seismographic Stations and Related Activities, Bull. Seism. Soc. Am., v. 32, pp 205-229.

Oppenheimer, David H., Klein, Fred W., and Eaton, Jerry P. (1992). A review of the CALNET, U.S. Geological Survey Open-File Report 92-xxx.

Press, Frank, chairman (1965). Earthquake Prediction: a proposal for a ten year program of research, Office of Science and Technology, Washington, D. C.

2. PC-BASED SEISMIC SYSTEMS

by

W. H. K. Lee
U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

In this chapter, I want to present an overview of PC-based seismic systems (Figure 1). This is intended to be an introduction for those of you who may not have any experience with these systems. First, I am going to discuss the basic components that make up a seismic network, and how those components have evolved into the present PC-based system. Next, I will explain why we have focused our development around the IBM PC and not some other

hardware configuration. Then I will go into the seismic data acquisition problem. Some of the relevant issues are: sensor type, telemetry method, analog to digital (A/D) conversion, cost, and software features. Next, I will discuss the systems that have been developed at the USGS. There are four of these systems now. The 12-bit, 16 channel system is the first one

<u>PC-based Seismic Networks</u>	
W. H. K. Lee	
1.	<u>Introduction to Seismic Networks</u>
(1)	Components
(2)	History
(3)	Development of PC-based systems
2.	<u>Using IBM-Compatible PCs</u>
(1)	Why?
(2)	Weakness
(3)	Early history
(4)	Current PCs
3.	<u>Seismic Data Acquisition</u>
(1)	Seismic sensors
(2)	Telemetry
(3)	Analog-to-digital
(4)	At what costs?
(4)	Desirable features
4.	<u>PC-Quake Systems Developed at the USGS</u>
(1)	12-bit, 16-channels
(2)	12-bit, 128-channels
(3)	16-bit, 64-channels
(4)	16-bit, 48-channels
5.	<u>Future Directions</u>

Figure 1. Chapter outline.

we developed. The 12-bit, 128-channel system has been running for about two years, and the 16-bit, 64-channel system just got on line this year. The 16-bit, 48-channel version is just being completed. Finally, I want to discuss where we go from here. What are our future directions?

Seismic Networks: Components

A seismic network consists of the components listed in Figure 2. One of the primary components is the sensor. At the low end of the "cost spectrum" is the L-22 seismometer, which is a 3-component velocity sensor. It is about the cheapest seismometer you can buy, at about \$1500 each. This sensor is very rugged and can be tossed into the back of a car and taken anywhere. It is reliable and very easy to work with.

Seismic Networks: Components

- Seismic Sensors: velocity or acceleration
- Telemetry: analog or digital
- Data Acquisition System: mini or micro
- Data Analysis System: mini, PC, workstation

Figure 2. Components of seismic networks.

After the sensor comes some sort of signal conditioning unit and signal conditioning is followed by telemetry. In the lab or lecture room, the telemetry is simple, just wires connecting the amplifier to the PC. Normally, however, you would use telephone or radio for the telemetry link. An important issue related to telemetry is whether to use analog or digital telemetry. We will be discussing that later on.

The data acquisition system scans the continuous stream of data coming from the

sensors and looks for seismic events. When the system identifies a seismic event, it saves the data to disk. In the past we used mini-computers to perform the data acquisition step. Now we are increasingly using PCs for that task.

The final link in the seismic data system is the analysis system. Mini-computers, PCs, and work stations may be used for analysis of seismic data.

Seismic Networks: History

Jerry Eaton has talked about the history of seismic networks so I won't dwell on the issue. However, a little discussion of the history of seismic networks is directly relevant to the development of PC-based systems.

Initially, seismic data were all recorded on drums, using either photographic paper, ordinary paper

with ink and pen, or thermo-sensitive paper and a hot stylus. These drum-based systems have been around since about 1900 and are still in use to some extent today. The records from these systems are very hard to work with because each record has its own timing system, and each earthquake may be recorded in several different records. As a result, timing corrections and data processing become very tedious.

The next step in the evolution of seismic data processing was the creation of the 20

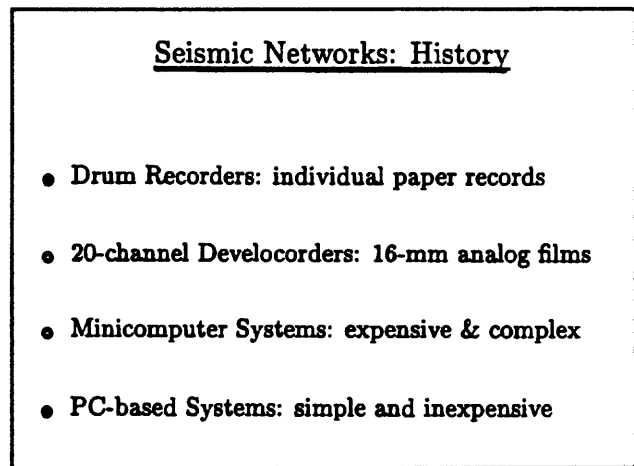


Figure 3. History of seismic networks.

channel developocorder system. These systems recorded seismic data on 16 mm-film, and were very popular from about the 1960's to about 1970.

In the 1970's and 1980's mini-computer-based systems were developed in many different universities and institutions. They typically cost up to millions of dollars, require tens of man years to develop the software, and are very complex. A chief virtue of these systems is that many of them do work quite well and are in routine use today. These are still the backbone of the processing system for most large seismic networks.

Development of PC-based Seismic Networks

In the late 1980's, several groups began to develop PC-based seismic systems. These systems have the advantage of being inexpensive and relatively simple to implement. Many groups outside the United States have also developed PC-based seismic systems. For example, groups in China, Japan, South Africa, and Spain have constructed such systems of varying degrees of complexity. The

French system is actually pretty advanced. For instance, the data telemetry for the system is

Development of PC-based Seismic Networks

- Others: China, Japan, S. Africa, Spain, etc.
- US: UC Berkeley, Weston Obs., USGS, etc.
- 1987: 16-channel system for explosion experiment
- 1988: Systems used for small seismic networks
- 1989: Hardware info and software published
- 1990: System extended to 128 channels
- 1991: System extended to 16-bit; More software
- 1992: Revised editions for IASPEI Vol. 1 and 2

Figure 4. Development of PC-based seismic networks.

by satellite. In the US there are several groups putting together PC-based seismic systems. Principally, these are the Berkeley group, the Weston Observatory group, and the USGS.

The history of the USGS effort began in 1987 with a 16-channel system developed for a totally different purpose than earthquake monitoring. It was used to monitor an explosion experiment in a quarry. At this quarry, explosions were set off regularly as part of the excavation process. However, in the same area, small earthquakes also occurred frequently. We wanted to be able to tell the two types of events apart and I was asked to find a way to do so. As part of this effort we deployed a dense seismic network around the quarry. Although the network was composed mostly of standard equipment used for temporary monitoring purposes, we also wanted to be able to do local real-time monitoring during the blasts. For this purpose, we developed the prototype PC-based monitoring system. The system proved to be quite effective.

In 1988, several groups applied similar systems for small seismic networks. Further development took place and in 1989 we decided to publish information about the system. The idea was to provide executable and source code and system information so that others could copy the system for their own uses. IASPEI Software Library Volume 1 was published in 1989 by the International Association of Seismology and Physics of the Earth's Interior (IASPEI) in collaboration with the Seismological Society of America. So far, about 500 copies have been distributed.

In 1990 the PC-based system was extended to 128 channels by means of an external multiplexer and in 1991 the system has been expanded to provide 16 bit A/D. Also more software has been developed. We realize that the key to a successful system is powerful

software. As Jerry Eaton mentioned, we can collect far more data than we can use. Unless we have powerful software, there is no way we can keep up with the analysis. In 1992, we plan to revise the programs of IASPEI Software Library Volume 1 and Volume 2, which are essentially the network processing software. Volume 3, which came out in early 1991, is on waveform analysis software. I will comment on this later.

USING IBM-COMPATIBLE PCs

Why IBM-Compatible PCs?

Why do we use the IBM PC in preference to a MAC or something else? A major reason is because of the open architecture of the PC. What that means is that anyone can make a PC, not just IBM. The resulting competition has resulted in low hardware prices, lots of peripherals, and worldwide availability. Another good thing about the IBM PC is that the INTEL microprocessor, which is

the heart of the machine has been designed so that all of its future versions are backward compatible. This means that computer code which was written for an early generation machine will execute on a future generation machine. Work stations like the SUN are not

Why IBM-Compatible PCs?

- Open Architecture
 - Cost effectiveness
 - World-wide availability
 - Plenty of peripherals
- Software Binary-Compatible
 - Inexpensive commercial software
 - Compatible with future hardware

Figure 5. Advantages of IBM-compatible PCs.

always backward compatible between generations, so every time they put out a new system, you have to start compiling the code all over again. This can be a real mess. I think that backward compatibility is very important, so that once you have a piece of code that works, you don't have to worry about it again.

You can rest assured that if you choose to develop for the PC you will have compatibility with future hardware. For these reasons, commercial software developers have written many programs for the PC, and the competition in the software market has kept the price of PC software low.

To give you an example, our original PC-based seismic system was developed on an AT (80286) machine. But to make it work on a 386 machine you just plug the A/D card in, install the software, and everything works! You don't have to worry about anything. When the 486 came out things were no different; the system still works just fine on the 486 machine. We know that the 586 is just around the corner and we expect that things will be no different on that machine either.

Weakness in IBM-Compatible PCs

Like anything else, PC strengths are accompanied by weaknesses. Right now, there are about 1000 companies around the world making IBM compatible PCs. There is really a mixture of equipment! In terms of architecture, not all of the "compatibles" are identical. Even though everyone claims to be 100% compatible, not all of them really are. Some machines we have tried will not work with our software. However, we have tested a fair number of PCs, and if you stick with those we have tested, you should have no problems

implementing the system. If you choose a machine we haven't tested, you will just have to try it out yourself.

The other problem is that you have too many choices! Just deciding what machine to use can sometimes be difficult. Since different manufacturers make different pieces, some pieces may not be completely compatible so you may have difficulty putting the system together. At the software end of

things, there also may be too many choices! For example, there are over 100 different word processors. Which one should you use? Another big drawback for a PC-based system is related to the operating system, DOS. As it is now implemented, DOS is a very simple operating system. Although that is in some respects a strength of DOS, there are some important processes which are either not supported or are poorly supported by DOS. Finally, most of the commercial software manufacturers will not distribute their source code. If you don't like the way their program works, you just have to live with it or else replace it with a different one.

Weakness in IBM-Compatible PCs

- **Open Architecture**
 - IBM-Compatibles are not identical
 - Too many choices
 - Hard to fit all pieces together
- **Software Binary-Compatible**
 - Too many choices
 - Limitations of DOS
 - Source code not usually available

Figure 6. Weaknesses of IBM-compatible PCs.

Brief History of IBM PCs

The first PC was introduced ten years ago in 1981. It used the INTEL 8088 chip running at about 5 MHz and came with 128 kilobytes (KB) of RAM. At that time, the designers at IBM considered 128 KB to be a lot of memory. That turned out to be wrong. They thought that the potential RAM configuration of 640 KB was more than anybody would ever use. That turned out to be wrong also. When it first came out, the PC was equipped with a

160 KB floppy disk drive for program storage. This was soon replaced by a 360 KB floppy drive.

In 1983, the IBM PC-XT was introduced. Like the original PC, this machine came with the INTEL 8088 chip running at 4.77 MHz. Standard RAM had been increased to 512 KB, expandable to 640 Kb. The XT also was equipped with a 10 MB hard disk. IBM thought that this improvement would solve the storage problem, but that turned out to be wrong also. In 1984, IBM introduced the AT machine based on the 80286 chip by INTEL. The processor ran at 6 MHz in this machine. As you can see, this did not really improve the clock speed of the machine much. However, the 80286 is a more powerful chip than the

Early History of IBM-Compatible PCs

- **1981: First PC**
 - Intel 8088, 4.77 MHz
 - 128 KB RAM, expandable to 640 KB
 - 160 KB floppy disk
- **1983: PC/XT**
 - Intel 8088, 4.77 MHz
 - 512 KB RAM, expandable to 640 KB
 - 10 MB hard disk
- **1984: PC/AT**
 - Intel 80286, 6 MHz
 - 512 KB RAM, expandable to 16 MB
 - 30 MB hard disk

Figure 7. Early history of IBM-compatible PCs.

8088, so the through-put is significantly improved. In the AT machine, RAM was expandable to 16 MB and the standard hard disk had a capacity of 30 MB. IBM thought that 16 MB was so much memory that there would never be a problem, but that turned out to be wrong also.

Current IBM-Compatible PCs

The 286 chip has been modified by different manufacturers so that it can be operated at higher speeds. Now chips are available with clock speeds of 8, 10, 12, 16, and 20 MHz. This diversity in chip speeds is not due to INTEL. It is due to IBM insisting that they have a second source. When they first negotiated with INTEL to provide the chip, IBM did not want to be stuck with one supplier, so they forced INTEL to license the chip to other manufacturers. That way, IBM could also buy it from someone else. Very soon after, Harris and AMD produced high speed chips. Now it is pretty standard that a 286 PC will come equipped with 1 MB of RAM, and have a 40 to 100 MB hard disk.

IN 1986, INTEL introduced the 386SX AND 386DX processors. These processors

Current IBM-Compatible PCs

- 286 PC
 - Intel 80286: 8, 10, 12, 16, 20 MHz
 - 1 MB RAM, expandable to 16 MB
 - 40 - 100 MB hard disk
- 386sx & 386 PC
 - Intel 80386sx or 80386: 16, 20, 25, 33 MHz
 - 2 MB RAM, expandable to 32 MB
 - 60 - 300 MB hard disk
- 486sx & 486 PC
 - Intel 80486sx or 80486: 25, 33, 50 MHz
 - 4 MB RAM, expandable to 100 MB
 - 100 - 1200 MB hard disk

Figure 8. Current IBM-compatible PCs.

can have speeds of up to 33 MHz. A standard configuration for a 386 machine is to have between 2 MB to 32 MB of RAM and have a 100 MB or larger hard disk. Unfortunately, INTEL has not licensed the design of the 80386 chip so they have been the sole source for this chip. In fact, there is a big legal battle going on between INTEL and AMD over whether anyone besides INTEL can produce a 386 chip. Right now, AMD is making a 386 chip which runs at 40 MHz and is completely compatible with the INTEL chip. You can see that with a little competition, you can always do a little bit better.

In 1988, INTEL introduced the 80486 chip. These are running at 25, 33, and recently, at 50 MHz. Now the feeling is that you need a system with at least 4 MB of RAM, expandable to at least 100 MB. Also, we are talking about hard disks with a capacity of a gigabyte or more. All of this in a span of about 10 years. The processors have improved by a factor of 50 to 100 times over the original 8088 processor. This rapid improvement in technology is due, at least in part, to the open architecture of the PC. Because of the open architecture, there has been a lot of competition. The competition has not only helped to drive down prices, it has resulted in rapid improvements in the available technology.

SEISMIC DATA ACQUISITION: GENERAL ISSUES

Seismic Sensors

One of the first questions to be asked is "What sensors should one use?" There are several types of instruments. Some people like long- period sensors and some people like broadband sensors. Short-period instruments cover a band roughly from one to ten Hz. They are really designed for recording local earthquakes. The long-period instruments respond to

frequencies from about 0.01 second to about 1.0 second. They are good for recording teleseismic events.

Broadband seismometers span the response spectrum covered by these two instruments. They are relatively expensive and are currently used mostly for research.

To give you some idea of the relative costs; a short-period, 3-component seismometer will cost you something like \$1,500. Long-period sensors typically cost something like \$5,000 to \$10,000 each and broadband seismometers will cost from \$10,000 to \$30,000 each. If you are interested primarily in teleseismic events, you don't need that many stations so you could invest more money in your sensors. On the other hand, as Jerry

Eaton pointed out, if you want to pinpoint a lot of small, local earthquakes, then you really need to use inexpensive sensors. Otherwise, you can not afford to deploy hundreds of them.

Seismic Data Acquisition: General Issues

- Seismic Sensors: What is appropriate?
- Telemetry: What is practical?
- Analog-to-digital: What is needed?
- Data Acquisition System: At what costs?

Figure 9. General issues of seismic data acquisition.

Seismic Sensors: What is appropriate?

- Short-period vs long-period vs broadband
 - Short-period (1-10 Hz) for local earthquakes
 - Long-period (0.01-1 Hz) for teleseisms
 - Broadband (0.01-10 Hz) for research
- Velocity vs acceleration for local quakes
 - Short-period velocity sensors for M0 - M6
 - Acceleration sensors for M5 or greater
- Vertical-component vs 3-components
 - Vertical-component for P-waves
 - 3-components for both P- and S-waves

Figure 10. What is appropriate for seismic sensors?

There is also a question about recording velocity versus acceleration. Should you record velocity, or should you record acceleration? The general rule of thumb is that if you want to record in the range from about magnitude 0 to magnitude 6 you would use the velocity sensor. You start using the acceleration sensor for magnitude 5 and larger.

Acceleration sensors are more expensive, typically costing about \$3,000 each. Also, they are not sensitive enough to record the smaller events. In practice, we have a few acceleration sensors in case there is a very large earthquake locally. That way we still have some on-scale records.

There is also a question of whether to use vertical-component sensors or use 3-component sensors. Obviously you would like to use all 3-component seismometers. But if you use all 3-component sensors, your network will cost about three times as much as the equivalent network built using vertical-component sensors. Vertical-component seismometers are really good only for the P-wave. It is not very easy to recognize the S-wave on a vertical waveform trace. The 3-component seismometers respond to both the P- and S-waves. Basically, the choice of sensor depends on what you want to record and how much money you are willing to spend.

Telemetry: What is Practical

The next issue is telemetry. In Chapter 4, John VanSchaack will talk more about that. The problem boils down to what is practical. The main problem is to choose between analog and digital telemetry. The analog telemetry is cheap but will be hard pressed to provide a dynamic range greater than about 48 db or so. To achieve economy, it is necessary to

multiplex analog data before transmission. Right now, with digital telemetry, we typically get a dynamic range of 96 dB with 16-bit digitization. Clearly you have far better dynamic range with a digital system. However, the digital system is several times more expensive because we cannot multiplex as many channels as in an analog system.

Telemetry: What is practical?

- Analog vs digital

- Dynamic range (dB): 48 (analog) vs 96 (digital)
- Cost: digital is several times more expensive

- Cable vs telephone vs radio vs satellite

- Cable: up to about 1 km
- Telephone: convenient but expensive
- Radio: line-of-sight to 100 km, but needs frequency allocation
- Satellite: nice but very expensive

Figure 11. What is practical for telemetry?

There is also a question about how to do the telemetry. There are about four choices. The first is to do the telemetry by cable. That is the cheapest, simplest approach for small distances. Once you buy the cable, you have it literally for forever. There are really no operating costs in the sense that nobody is going to charge you for the use of the cable. However, with this method, any distance beyond about a kilometer will be very difficult. Jerry Eaton has done about 10 kilometers in Hawaii, so he can tell you how hard it is. I give up at about one kilometer in distance.

The next option is to use telephone lines. They are very convenient, and you let the telephone company take care of the maintenance. However, this method is limited to where there are phone lines. Also at the end of each month, you get a big phone bill! Finally, during an earthquake, the phone lines may become inoperative, so if you are dependent on phone lines for all your telemetry, your network goes down.

Another telemetry option is to use radios. But this requires line of sight between stations, and has a maximum distance of about 100 km or so. Beyond that distance, you need repeaters and that brings its own set of problems. Another problem with radios is that there are so many of them in use. You will need to get a frequency allocation for your system. Even then, you may find that others are using the same frequencies and causing interference. This may be especially true in many developing countries, where it is very hard to find a frequency band that no one else is using.

The last option is to use satellite telemetry. That is really nice, but is also extremely expensive. For example, a satellite-based station may cost up to \$1,000,000. As you can see, this is a very expensive proposition.

Analog to Digital Conversion

Every sensor that you might use only produces a voltage change in response to seismic ground motion. It does not produce "numbers" that a computer can use. So you have to do something to turn the analog signal into a stream of digital data that the computer can use. When people talk about digital stations, they are still using an analog front-end. The question is "where is the digitization being done?" Are the signals digitized at the sensor site or at a central recording location? If you need high dynamic range, then you definitely must digitize at the site. That way, you don't get the loss of dynamic range because of the analog transmission. However, in situ A/D is quite costly to install and maintain. If you have 100 stations then you need to maintain 100 A/D converters in the field. That is a lot of work. Another problem with in situ A/D is that it is very difficult to achieve a common time base.

Every station is running on a separate clock, so you need some kind of scheme to synchronize clocks. That can also be very difficult. In short, there is a trade off between fidelity of the digitized signal and cost. A centralized system will be easier and cheaper to implement and maintain but will not produce as high a quality of data as an in-situ method can provide. You have to consider what your network will be

used for. If you really need 16-bit A/D then you have no choice but to use in situ digitization. Otherwise, save yourself time and money and go with analog telemetry.

A related topic concerns how many bits should be used to encode your data. A 12-bit system will give you a dynamic range of about $\pm 2,048$. The current analog network using either phone lines or radio transmission with 8 signals multiplexed onto one carrier has an intrinsic dynamic range of about 40 to 50 dB. So using anything more than 12-bits is really a waste of money. A 16-bit system has a dynamic range of $\pm 32,000$. If you can afford it, this much dynamic range is nice. Even if the system gets some DC drift, you still don't have to worry too much about clipping seismic signals. With a 12-bit system, on the other hand, DC drift can cause serious clipping problems.

The research people really love 24-bit systems. These systems have dynamic ranges

Analog-to-digital: What is needed?

- Centralized vs in situ
 - Dynamic Range: in situ A/D is much better
 - Cost: in situ A/D is more expensive
 - Common Time Base: difficult for in situ A/D
- 12-bit vs 16-bit vs 24-bit
 - 12-bit: ± 2048 range, or 72 dB
 - 16-bit: $\pm 32,768$ range, or 96 dB
 - 24-bit: $\pm 8,388,608$ range, or 144 dB
 - Cost: 12-bit \ll 16-bit \ll 24-bit

Figure 12. What is needed for A/D Conversion?

of $\pm 8,388,608$. With such a system you can record earthquakes in the range magnitude 1 to magnitude 8 on scale. Of course 24-bit systems cost lots of money. Typically the cost of a 24-bit system is an order of magnitude greater than the cost of a 16-bit system.

Data Acquisition System: At What Costs?

What sort of computer should be used for data acquisition? The big debate here is between mini-computers versus PCs. Many people started out with mini-computers and are reluctant to switch. They are used to their present system and all their software is written and debugged, so why should they change? However, sooner or later, the costs will catch up with them. The mini-computers are more expensive to operate and maintain. Also, most mini-computer companies are going bankrupt! Pretty soon, you will not be able to buy replacement parts for mini-computers.

Other issues are related to software. Many people like to have very complex software which can deal with any imaginable situation. But to do that there must be a large investment in money and man hours to create that software. I think it is much better to have simple software. Simple software is easier to develop and maintain. Right now, the cost of

Data Acquisition System: At what costs?

- Mini-computer vs PC
 - PC is much easier to use & maintain than Mini
 - Cost: Mini-computer is more expensive
- Simple vs complex software
 - Simple software is preferred because of high cost of software development and maintenance
 - "Open" software is preferred over "proprietary"

Figure 13. Cost considerations for seismic data acquisition system.

developing software is skyrocketing.

Another software issue is the question of "open" versus "proprietary" software. If your code is open to others then they may build upon it. They don't have to rewrite the code from scratch. This can only happen if you have a common hardware platform and a desire to work together. If people don't want to work together, then even open software won't help.

For these reasons we have designed our system to be based on a PC. When I set out to solve this problem several years ago, that was my goal. I wanted to reduce the time spent on analysis by about a factor of three. That is not easy to achieve. It turned out to be the hardest part of the struggle. I wanted to reduce the system cost by a factor of π^2 , or roughly about ten, and it was easy to do.

I realized that it was necessary to get people to want to help me, so I would have to publish software as it was developed, so others would have access to the code. That is what I did in 1989. Incidentally, when I first proposed this idea to the directors of the Seismological Society of America, one of the directors said that I would be lucky to sell 50 copies in his lifetime. I think I have proved him wrong since about 400 copies of volume 1 and a total of about 1,000 copies of all three volumes have been sold in the last 3 years.

Desirable Features of the Data Acquisition System

We would like to handle one to hundreds of channels. We also would like to be able to sample at rates of from about 10 to about 1,000 samples/second/channel. We want to be able to display the data in real time, so we can see what we are doing. We want as wide a dynamic range as we can afford, and we want to be able to save the data either continuously

or by event. We also want to be able to do a lot of the standard processing automatically. Finally, we want software which will aid in further data analysis. These are the basic parameters around which my system has been built.

Input to the data acquisition system are the channels of analog seismic data and one channel carrying

the IRIG-E time code. The system uses two 386 PCs, one for on-line acquisition and one for off-line analysis. Each PC can handle 128 input signals. With a 25 MHz 386 PC (a middle of the road machine these days) you can digitize 128 channels at a maximum rate of about 300 samples per second per channel. A 486 machine will certainly provide even better performance. The total cost of everything you need for the data acquisition is about \$10,000 to \$15,000. Right now we are using a program called XDETECT for monitoring local earthquakes, and a program called TDETECT for monitoring teleseisms. The XPLAY program allows you to quickly play back events. The FIXTIME program corrects the time, and also determines the actual A/D sampling rate. The program uses the IRIG-E codes, which have been recorded by the system to do the time corrections. That is something that I learned from Jerry Eaton. People think that they can design their hardware to be so smart that they don't need to record the time. However, almost all of the automated timing systems

Data Acquisition System: desirable features

- Handle 1 to 128 channels of input signals
- Digitize 10 to 1000 samples/sec/channel
- Display digitized data in real time
- 72 - 96 dB dynamic range
- Save digitized data continuously or by events
- Automatically pick P-arrivals & locate events
- Off-line analysis support software

Figure 14. Desirable features of a data acquisition system.

have problems of one type or another. Unless you record the time side by side with the seismic signals, you are not really sure what is happening.

Another program is used to demultiplex the data files. The data is recorded in a channel-by-channel fashion so that between two adjacent time samples for one channel are samples from all the other channels. However, when you analyze events you need all the data for each channel to be in a continuous stream. Therefore, you need to demultiplex the data. After demultiplexing the data, you can run it through an automatic picker program which we will discuss later. This may be followed by a hypocenter location program and a plotting program. There are also several waveform analysis programs which can be used on your data.

We feel that this inexpensive system will eventually play a big role in seismic data acquisition and subsequent research on the events collected by the system. As of the fall of 1991, there are about 50 of these systems running around the world and the number is increasing almost every week.

Data Flow

Up to 127 analog seismic signals plus one IRIG-E time code are saved in a waveform file. The waveform file is time corrected and demultiplexed. This is followed by automatic arrival time picking and hypocenter location. The analyst can then review the data and make any necessary changes. When everything is satisfactory, the data is archived on an optical disk.

The automatic picker does a reasonably good job most of the time. Even though a

human analyst could probably do a little better, if you have lots of stations with good arrivals there will be very little difference between the hypocenter locations. The way we use the system is to let it do its work at night. Then in the morning we can look at the paper records it has produced. These show the actual picks and the expected picks according to the computed hypocenter location. Since these are on paper, the analyst can look through the records very quickly, and only follow up on the ones which look like there may be a problem. This method greatly speeds up the routine data processing.

PC-QUAKE SYSTEMS DEVELOPED AT THE USGS

Four PC-Quake systems have been developed at the USGS. In this section, I will describe the basic features of these 4 systems.

PC-Quake System I: 12-bit, 16 Channels

This system uses a standard 8 MHz 286 PC equipped with an off-the-shelf A/D board and software capable of digitizing up to 16 channels with 12-bit resolution at up to 500 samples/second/channel. The system displays digitized data in real time and can save data continuously or by event. The system automatically picks P-arrivals & locates events and off-line software does filtering, FFT, and Coda Q analysis. Principal advantages of this system are that it is inexpensive and available world-wide, (The cost is only about \$5,000.) and all of the required software was published by IASPEI/SSA in 1989. The principal disadvantages are that the system is limited to 16 channels and has only 12-bit resolution. This resolution is further limited by analog telemetry.

PC-Quake System II: 12-bit, 128 Channels

The PC-Quake System II expands the capabilities of System I through the use of a more powerful computer and the addition of an external multiplexer. The system uses standard 25 MHz 386 PC with an off-the-shelf A/D board and software. It can digitize up to 128 channels with 12-bit resolution at up to 500 samples/second/channel when all 128 channels are being used. The sampling rate increases if fewer channels are digitized. Like System I, the system displays digitized data in real time and can save data continuously or by event. The system automatically picks P-arrivals & locates events and uses off-line software to do filtering, FFT, and Coda Q analysis. The principal advantages of the system are that it is relatively inexpensive and available world-wide (Cost is about \$10,000.), and the required software and specifications will soon be published by IASPEI/SSA. Disadvantages of the system are that it is limited to 12-bit resolution and must use an external multiplexer developed at the USGS. The resolution is further limited by analog telemetry.

PC-Quake System III: 16-bit, 64 Channels

The PC-Quake System III has been designed as a resolution system, capable of rapid deployment to the field. It uses a 16 MHz 386sx portable PC with the standard off-the-shelf A/D board and software. The system digitizes up to 64 channels with 16-bit resolution at up to 200 samples/second/channel. As with the previous systems, the System III displays digitized data in real time, saves data continuously or by event and automatically picks P-arrivals & locates events. Off-line software does filtering, FFT, and Coda Q analysis. The principal advantages of the system are its 16-bit resolution, portability and its relatively

low price (\$10,00). Its software and specifications are soon to be published by IASPEI/SSA.

Disadvantages of the system are that it requires a Quesig multiplexer and direct cabling.

Also, it requires AC power (200 watts).

PC-Quake System IV: 16-bit, 48 Channels

The most sophisticated of the PC-Quake systems is the System IV. This unit is based on a 16 MHz 386sx portable PC and is designed so that the A/D conversions are done in the field and the data is transmitted by 16-bit, digital telemetry to the processing site. The system can digitize at up to 100 samples/second/channel when all 48 channels are being used. As with the other PC-Quake systems, the System IV displays digitized data in real time and can save data continuously or by event. The system automatically picks P-arrivals & locates events and has off-line software which does filtering, FFT, and Coda Q analysis. Despite its capabilities, the System IV is relatively inexpensive at about \$10,000. Its low price, portability, and high resolution are its principal advantages. It could be significantly less expensive were it not for the costs of doing field A/D and digital telemetry. Also, the Cutler digital grabber used by the system imposes additional cost and complexity. Another drawback is that the maximum digitization rate is limited by digital telemetry capabilities.

Commercial schemes for digital telemetry involve packaging the data and transmitting. If there is a transmission problem, the data is retransmitted. This requires a relatively complicated field unit. At the central receiving point, things are also more complicated. This is because the order in which data packets are received is not necessarily the order in which they occurred. So the real time detection software has the additional

complication of having to unscramble the data packets.

However, in our scheme, if there is a transmission problem, the system simply does not replace the previous data. Because ground motion is never constant, we know we have a problem if two or more data values are the same. With this design, there is no need for calibrated clocks at the field sites, since the time base is still maintained at the central recording site. Our system is about an order of magnitude simpler than the commercial systems. I think most commercial designs are intended to be so difficult to implement that an individual could not possibly duplicate them. By contrast, our system is intended to be easy to implement.

Right now the 16-bit A/D chips are running about \$100 each. It takes 3 of them per three-component station in our design. In the future, I expect the chip prices to fall, and I hope we will be better able to utilize the chip capabilities so that fewer chips will be required per station. Outside of the digitization hardware, the system is essentially identical to other systems we have produced. The sensors are the same, and the radios are the same as in other of our systems. At the central site, the PC is the same, but in place of the A/D board, there is a digital grabber. Because of the similarities, we hope to evolve a hybrid system that will be able to work with both analog and digital signals.

FUTURE OF PC-BASED SEISMIC SYSTEMS

I expect the hardware cost to continue to go down. This is because of the fierce competition in the electronic and computer industries. The hardware engineers seem to be doing a better job than the software engineers. I expect software costs to continue to increase

as software continues to increase in complexity. I believe that the PC-Quake system will be widely used. I expect that by the end of 1992 there will be about 100 systems in operation. These systems are being used for a lot of different purposes. Some people are using them for monitoring earthquakes. However, the greatest success so far has been in monitoring volcanoes. In the Philippine case, a small network was set up before the eruption using simple gear. The American-Filipino team was able to predict the eruption of Mount Pinatubo very accurately. This allowed all the people to be safely evacuated before the eruption. Potentially, thousands of lives were saved, and the US government was saved billions of dollars in airplanes. The US Bureau of Mines is using the system for monitoring rockbursts. Caltran wants to use the system for monitoring highways and bridges. However, the biggest expansion in system installations will come from building monitoring. Civil engineers are very interested in using the system to monitor the response of large buildings to earthquakes. In the past, there would be at most, 3 accelerographs in a building. This did not provide enough detail about the motion of the building to really understand how it was responding to the earthquake. Now, with a 128 channel system, you can instrument the building well enough to really see what is going on. I imagine that in the next couple of years, the biggest user of the PC-based system will be civil engineering projects, e.g., there are plans in Taiwan to install about 100 systems in the next couple of years.

SUGGESTED READING

- Lee, W. H. K., and S. W. Stewart, (1981). "Principles and Applications of Microearthquake Networks." Academic Press, New York.
- Lee, W. H. K. (Ed.). "Toolbox for Seismic Data Acquisition, Processing and Analysis", IASPEI Software Library Vol. 1, Seismological Society of America, El Cerrito, CA.

3. SEISMOMETERS THEORY AND PRACTICE

by

J. P. Eaton

U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

Commonly we are so tied up in seismology with what we are going to do with the records we have that we don't think about how we are getting those records. More and more of the people who are writing papers in seismology are interested in analysis of waveforms and the physics of the earth, and they get farther and farther away from seismometers. In doing that they don't realize what shaky ground they are on, literally a bed of sand, because seismometers, by and large, don't do what the manufacturers say they do. So it is terribly important that you realize what a seismometer is and what it can do and can't do.

The task that we put seismometers to is breathtaking. One group of seismologists expects them to record down to background earth motion at frequencies as high as 100 Hz while another group expects them to record signals with periods as high as 100 seconds down to the prevailing noise level. This problem is complicated by the earth itself in that things moving on the surface of the earth, weather, pressure patterns, waves, generate a terrible racket in the frequency band of about one to ten seconds. Early on, when the interest was primarily in recording teleseisms, seismologists were well aware of microseisms generated by storms with periods in the range four to ten seconds. As we got more interested in local earthquakes and focused on higher frequencies, the noise generated by wave action with a period of about one second became important, at least along the California coast.

This barrier between the high frequency world and the low frequency world essentially has divided seismology into two camps. More recently there is an effort with the very sophisticated long period instruments, which are essentially force-balance devices driving electronics that make the world look level, as it were, and record with very high dynamic range, to bridge this barrier. So far, this class of instruments has done marvels for long period seismology, but their benefit for short period seismology is questionable.

DEVELOPMENT OF SEISMOMETERS

I will start out by showing you what seismometers used to look like. These are pictures from the classic textbook on seismology by Boris Galitsin, who was in my opinion, one of the most outstanding seismologists that ever lived. He worked in Russia primarily between about 1900 and 1915 when his work was published. Figure 1 is a picture of a German built vertical component mechanical seismograph. The magnification of this instrument was about 100 and its period was several seconds. It was one of the first instruments designed specifically to record P phases of distant earthquakes.

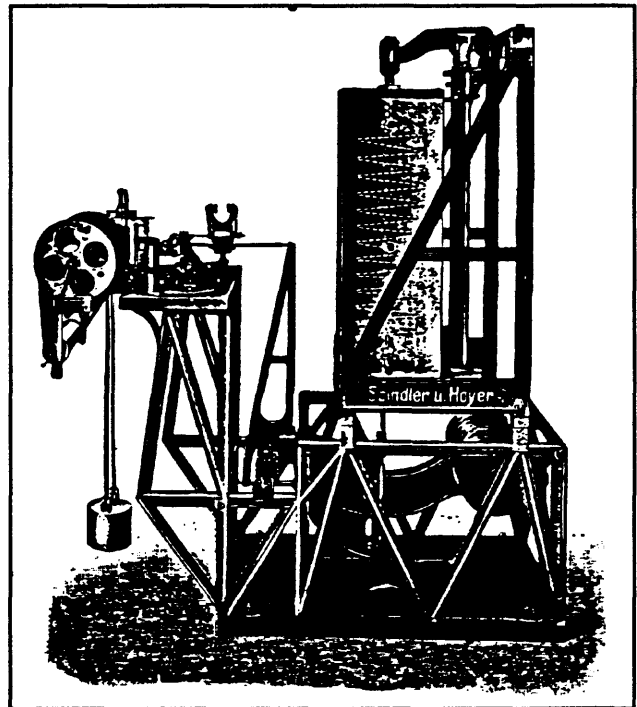


Figure 1. German built vertical component seismometer.

That class of instruments, which depended entirely on mechanical amplification of the ground motion, was superseded by galvanometric recording which was introduced by Galitsin. Figure 2 shows, on the left, a horizontal component seismograph and, on the right, the galvanometer. By using a fairly heavy mass in the seismometer and a light mass in the galvanometer, one could get additional

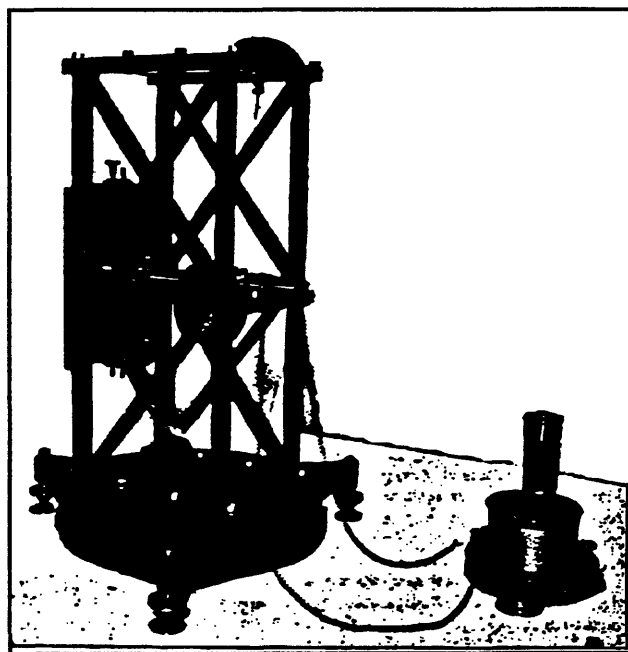


Figure 2. Horizontal component seismometer.

amplification. However, there were problems lurking right under the threshold because in this system the galvanometer reacts on the seismograph and produces a more complicated relationship between the ground motion and the output of the seismometer. This instrument leapfrogged from magnifications of the order of 100 to magnifications of the order of 1000. It came in both horizontal and vertical configurations.

Figure 3 shows a vertical component, galvanometer-assisted seismometer. You can see the mass and the transducer magnet way out on the left. Also, a damping magnet was provided so that you could control the damping. In order to get sensitivity at the long periods for which these were used, the spring and support system were specially configured so that when the boom went up and down a lot, the spring changed length only slightly. This instrument and a modification built by Witip essentially dominated long period seismology

from about 1910 when it first came into use, at least until about 1950. It still is a very respectable instrument that produces excellent records. One of the prices you pay with galvanometric recording is that you record on photographic paper in the dark and you have all of the fuss that goes with that.

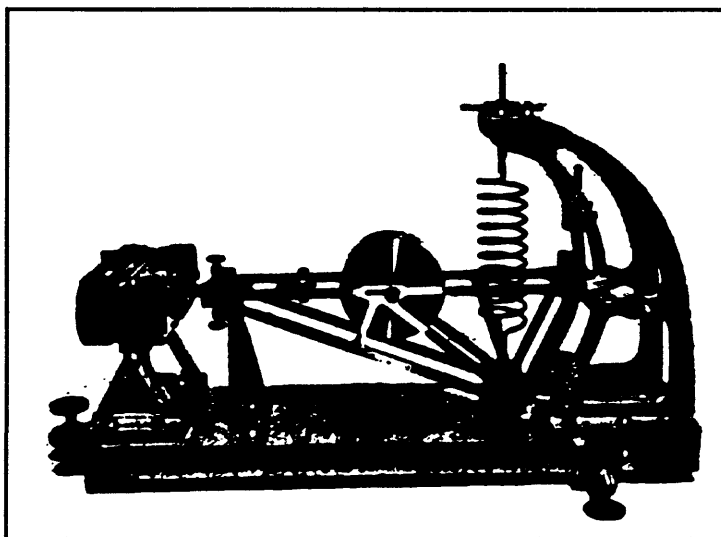


Figure 3. Vertical component, galvanometer-assisted seismometer.

The division of seismology into long period and short period camps is

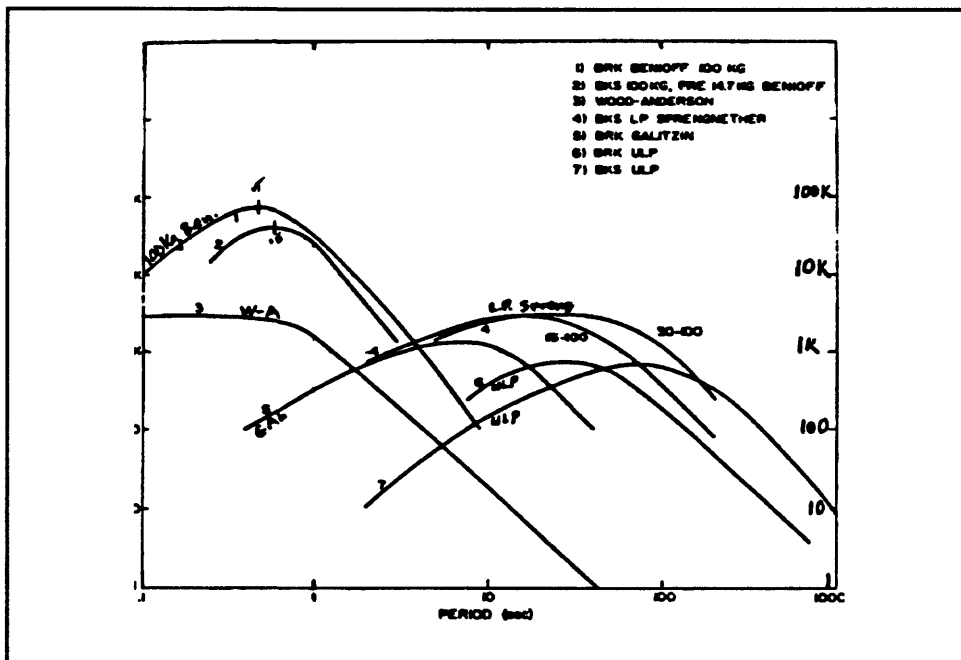


Figure 4. Response curves for several seismometers.

reflected in the next figure. Figure 4 is taken from a summary of earthquakes in northern California put out by Berkeley a few years back. The figure shows response curves for a few instruments. The vertical axes are magnification and the horizontal axis is period of the wave motion. The curve with the high peak in the upper left corner of the chart is a 100 kilogram Benioff. It has magnification of near 100,000 and involves galvanometric recording but the seismometer transducer instead of being a moving coil uses a variable reluctance type transducer which attains magnifications an order of magnitude greater than you can get with a moving coil. So for short period seismology, this instrument invented by Benioff in 1930 dominated until it was superseded by electronic seismographs. On this diagram you can also see the Wood-Anderson curve which has the lowest magnification of the three short period seismographs. The Wood-Anderson also records photographically but is very much smaller than the seismometers shown in the previous figures. For long period work, a variety of special purpose instruments have been built with peak magnifications of the order of 3000 or so, but with the peak magnification displaced out to a period of ten to 100 seconds. They are fairly low gain because the seismic waves in that band are quite large and the noise in some of that band can also be quite large. Their gains drop off as you come back into the vicinity of about one second.

In the higher frequencies, the earth is quite a bit quieter so we run to higher magnifications. The response curves for electronically amplified seismometers are different from that of the Benioff in that the electronic instruments have a much wider range over which their magnification is high. That is very important in getting a high fidelity recording of the earthquake.

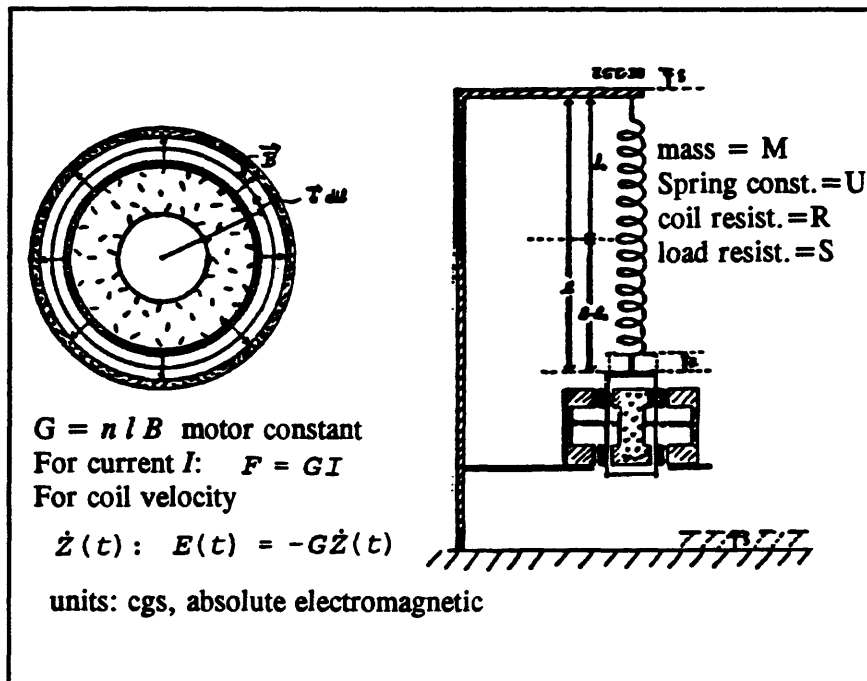


Figure 5. Schematic of moving-coil seismometer.

Now I would like to describe an instrument equivalent to the Mark Products L-4 seismometer. It has a free period of one second and it doesn't have any mechanical output. The signal comes out a set of wires and you have to figure what to do with it from there. A major operational difference between this seismometer and the mechanical ones I have been talking about has to do with testing and calibration. With the other seismometers, one can imagine performing experiments where you put weights on the mass and lift it and play games like that. The L-4 is sealed up so that any test you wish to perform has to be applied through the seismometer leads and interpreted through the signals that come back. That changes the calibration process considerably. Figure 5 shows a schematic view of a moving coil seismometer. The heart of a moving coil seismometer is the coil unit, which is the transducer. As shown in Figure 5, the seismometer mass is suspended by springs and

constrained to move vertically. The mass is suspended inside a transducer which produces a strong magnetic flux through the coils wrapped on the moving part of the instrument. When there is relative motion between the seismometer mass and the frame, one gets an output voltage. The principle is similar to that used in a speaker for an audio system. For a one Hertz instrument, if you have a spring with a free length of L_0 , in order to attain a period of about one second, when you attach the mass the spring must stretch about 25 cm. That poses quite a packaging problem if the seismometer is to be made compact. So in modern seismometers like the L-4 they don't use a simple coil spring but instead use a plate type spring with the property of having different spring constants for different positions of the mass. That poses a problem in terms of the response of the seismometer if the mass sags because of age or, in a horizontal unit, if the seismometer is not level.

In Figure 5, L_0 is the unloaded length of the spring and L is the loaded length. When the seismometer is disturbed, the mass is displaced by an amount Z . That is the coordinate of motion relative to the frame. The whole frame is displaced by an amount ζ , the coordinate of ground motion. The function of the transducer is to convert relative motion between the coil and the frame into an output voltage. The induced EMF produced by the motion is proportional to a motor constant times the relative velocity of the mass with respect to the frame. When a current is fed through the coil, it reacts with the magnetic field to produce a force which moves the mass with respect to the frame. The same motor constant applies in both cases.

EQUATIONS OF MOTION

The equations of motion for the seismometer are developed by calculating the balance of forces on the system described in the last figure. The equations involve gravity, the force due to the spring, damping, and the driving force that comes from the earth. To derive the equations of motion, we first need to calculate a balance of forces for the seismometer system. The force on the mass due to gravity is given by $-Mg$. The restoring force developed by the spring is given by $[(L-L_0)-Z]U$ where Z is the disturbance to the system and U is the spring constant. This force turns out to be equal to $Mg - ZU$.

There are two forms of damping in the system. The coils normally are wrapped on a thin aluminum coil form which acts like a shorted turn. So when the mass moves, even if the external circuit is disconnected, there is fairly large damping, about 0.3 in the L-4. This damping term is given by

$$\text{open circuit damping} = -\Delta \dot{Z} \quad (1)$$

The second part of the damping is due to the current flowing through the seismometer coils and is given by

$$\text{electromagnetic damping} = -GI = -\frac{G^2}{R+S} \dot{Z} \quad (2)$$

So, the total damping is given by

$$\text{total damping} = -\left(\Delta + \frac{G^2}{R+S}\right) \dot{Z} \quad (3)$$

The inertial force is given by

$$inertia = -M\ddot{Z} \quad (4)$$

Finally, there is a virtual force developed by the movement of the earth relative to the χ frame, given by

$$virtual\ force = -M\ddot{\chi} \quad (5)$$

Summing the forces, we get the equation of motion

$$-M\ddot{Z} - \left(\Delta + \frac{G^2}{R+S} \right) \dot{Z} + Mg - ZU - Mg - M\ddot{\chi} = 0 \quad (6)$$

Simplifying, we get

$$\ddot{Z} + \frac{1}{M} \left(\Delta + \frac{G^2}{R+S} \right) \dot{Z} + \frac{U}{M} Z - \ddot{\chi} = 0 \quad (7)$$

It is traditional to further simplify this equation by rewriting the coefficients in terms that relate more naturally to the physical properties of the response. Then we define:

$$\Omega_0 = \sqrt{\frac{U}{M}} = \frac{2\pi}{T_0} \quad (8)$$

$$2\beta\Omega_0 = \frac{1}{M} \left(\Delta + \frac{G^2}{R+S} \right) \quad (9)$$

where β is damping in terms of fraction of critical and Ω_0 is circular natural frequency. Since the damping comes from two terms, β can be separated into two parts; β_0 which is the open circuit damping, and β_1 which results from the current through the coil. Using these coefficients, the equation becomes

$$\ddot{Z} + 2\beta\Omega_0\dot{Z} + \Omega_0^2 Z = \ddot{\chi} \quad (10)$$

Solution of the differential equation of motion can be done a number of ways. Here, I use a Laplace transform method which conveniently deals with the initial conditions of the problem. After transformation, the equation becomes

$$\Phi(\tau) = \frac{\mathcal{L}[\dot{\chi}] + (\tau + 2\beta\Omega_0)Z(0) + \dot{Z}(0)}{[\tau^2 + 2\beta\Omega_0\tau + \Omega_0^2]} \quad (11)$$

The denominator of the RHS can be factored and the resulting equation is

$$\Phi(\tau) = \frac{\mathcal{L}[\dot{\chi}]}{(\tau - \tau_1)(\tau - \tau_2)} + \frac{(\tau + 2\beta\Omega_0)Z(0)}{(\tau - \tau_1)(\tau - \tau_2)} + \frac{\dot{Z}(0)}{(\tau - \tau_1)(\tau - \tau_2)} \quad (12)$$

Here, τ_1 and τ_2 are given by

$$\begin{aligned} \tau_1 &= \Omega_0(-\beta + \sqrt{\beta^2 - 1}) \\ \tau_2 &= \Omega_0(-\beta - \sqrt{\beta^2 - 1}) \end{aligned} \quad (13)$$

One thing we want to know is the response of the system to a long continuing harmonic motion. Such motion can be characterized by

$$\chi(t) = \begin{cases} A \sin(\omega t) & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (14)$$

For this kind of input, equation 13 becomes

$$\mathcal{L}[\dot{\chi}(t)] = \frac{\tau^2 A \omega}{\tau^2 + \omega^2} + \tau \chi(0) + \dot{\chi}(0) \quad (15)$$

There are standard, fairly straightforward inversion techniques for this equation. When we carry those out, we end up with an equation for the motion of the mass relative to the frame of the form

$$Z(t) = \frac{A\omega^2}{\sqrt{(\Omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2\Omega_0^2}} \sin\left[\omega t - \tan^{-1}\left(\frac{2\beta\omega\Omega_0}{\Omega_0^2 - \omega^2}\right)\right] \quad (16)$$

This is the motion of the mass relative to the frame. But we don't measure that. What we measure is an output voltage. This has the form

$$V(t) = \frac{S}{R+S} G \dot{Z}(t) \quad (17)$$

Combining equations 17 and 16 we get the equation for the output voltage

$$V(t) = \frac{S}{R+S} G \frac{A\omega^3}{\sqrt{(\Omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2\Omega_0^2}} \cos\left[\omega t - \tan^{-1}\left(\frac{2\beta\omega\Omega_0}{\Omega_0^2 - \omega^2}\right)\right] \quad (18)$$

SEISMOMETER CALIBRATION

Figure 6 is a schematic diagram of the seismometer and its external circuit. The seismometer is characterized by a motor constant G and an internal resistance R . When there is relative motion between the mass and the frame, an output voltage is developed across the seismometer leads that is given by:

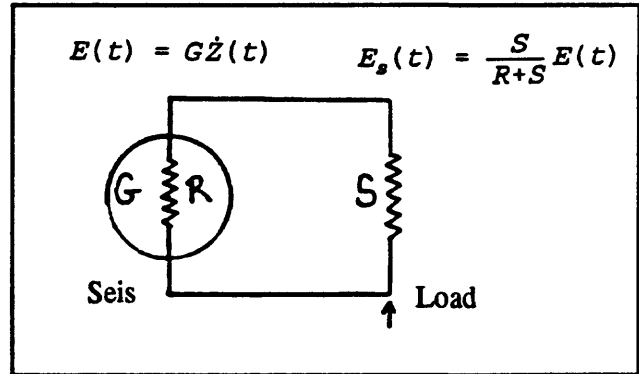


Figure 6. Schematic of seismometer and external circuit.

$$E(t) = G\dot{z}(t) \quad (19)$$

Normally the seismometer is driving some other unit like the input of an amplifier which has an internal resistance S . So the voltage which actually appears the load is given by $S/(R+S)$. This is the relation between the output voltage and velocity.

We need to carry out a number of tests on the seismometer to measure what its constants are. Normally the manufacturer specifies the values of the motor constant, the mass of the seismometer, the free period of the seismometer, the open circuit damping, and the damping corresponding to some particular load value. This invites you to believe that this is all true and that if you set things up according to the specifications you will get the correct results. But seismometers are variable from one to the next. In the L-4 series, getting the free period right is very critical. You specify a range like 5% that you will tolerate. That makes an appreciable difference in some parts of the response curve. Also, the stability of the motor constant depends on the stability of the magnet charge which, with rough handling, can change. So, there are a variety of factors that can cause the seismometer to change with time. You need to be able to protect yourself by measuring those constants so you know what they are.

One way of doing that on a closed seismometer like the L4 is called the release test. In some manner you deflect the seismometer from its equilibrium position for a period of time. Then you release the mass and let it return to equilibrium. In a mechanical instrument, one can envision taking the case off, putting a little weight on the seismometer mass, and then snatching the weight off. That is a perfectly good test, but you can't carry it out on a seismometer with a closed case. Another test that you can use, is to apply a voltage to the

coil which offsets the mass, and calculate what the offset is. Then suddenly cut off the voltage and let the seismometer return to zero while recording its output voltage. That turns out to be a very useful test for this type of instrument. You can use it not only to determine the constants of the instruments, but also, in a cruder form, to tell whether an instrument is centered, or, with a horizontal instrument, to actually carry out the centering of the mass.

The equations governing the release test are derived from the equations of motion shown before. If you set the motion of the earth equal to zero, the initial velocity of the seismometer relative to the frame equal to zero, the initial displacement not equal to zero, and the damping less than critical, then the equation describing the position of the seismometer mass as a function of time is

$$Z(t) = Z(0) \frac{e^{-\beta \Omega_0 t}}{\sqrt{1-\beta^2}} \sin\left(\sqrt{1-\beta^2} \Omega_0 t + \tan^{-1} \frac{\sqrt{1-\beta^2}}{\beta}\right) \quad (20)$$

Since we measure the output voltage rather than the actual motion we need an equation relating output voltage to the motion of the mass. This turns out to be

$$E_s(t) = \frac{S}{R+S} G \frac{Z(0) \Omega_0 e^{-\beta \Omega_0 t}}{\sqrt{1-\beta^2}} \sin\left(\sqrt{1-\beta^2} \Omega_0 t + \tan^{-1} \frac{\sqrt{1-\beta^2}}{\beta}\right) \quad (21)$$

It is necessary to determine the relationship between the offsetting voltage and the value of the first maximum. This is given in equation 22.

$$E_{s_{\max}} = \frac{S}{R+S} G \frac{Z_0 \Omega_0 e^{\frac{-\beta}{\sqrt{1-\beta^2}} \tan^{-1} \frac{\sqrt{1-\beta^2}}{\beta}}}{\sqrt{1-\beta^2}} \quad (22)$$

As we actually carry out this test, we send a current I through the seismometer and it produces an offset equal to the motor constant times I divided by the spring constant U so

$$Z(0) = \frac{G^2 I}{M \Omega_0^2} \quad (23)$$

When we substitute this value for $Z(0)$ into equations 20 and 21 we get the final "working" versions of the equations for output voltage and first maximum for the current release test.

$$E_s(t) = \frac{S}{R+S} \frac{G^2 I}{M \Omega_0} \frac{\Omega_0 e^{-\beta \Omega_0 t}}{\sqrt{1-\beta^2}} \sin(\sqrt{1-\beta^2} \Omega_0 t) \quad (24)$$

$$E_{s_{max}} = \frac{S}{R+S} \frac{G^2 I}{M \Omega_0} \Omega_0 e^{\frac{-\beta}{\sqrt{1-\beta^2}} \tan^{-1} \frac{\sqrt{1-\beta^2}}{\beta}} \quad (25)$$

These equations are essentially what are required to interpret the release test and from it to derive the motor constant, the free period, and the damping. The experimental setup for doing this can be relatively simple. Figure 7 shows in schematic form, the seismometer and the external circuit with a load resistor across it which represents the input to the amplifier.

Here, I have shown that one picks off

the voltage across the external circuit with an oscilloscope. It is much more convenient to run the output voltage directly into a PC so you have a permanent record to study. Also, there is

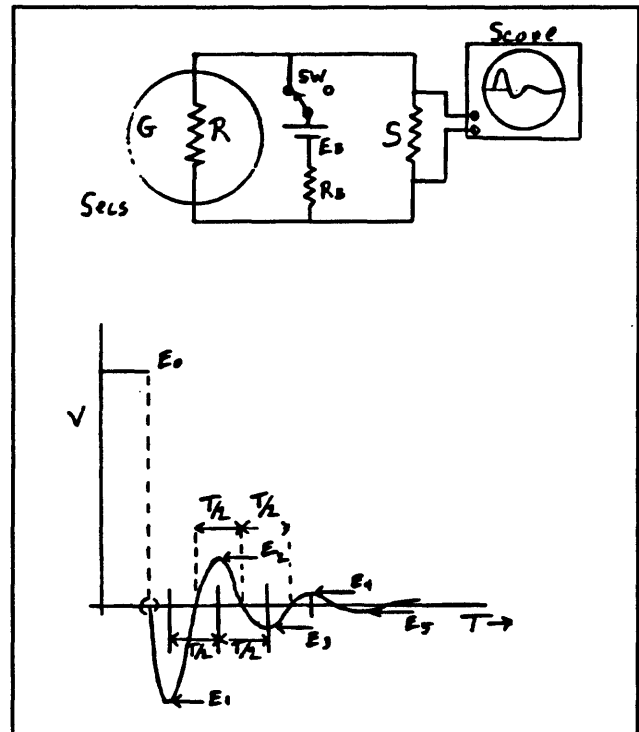


Figure 7. Set-up for Release Test

a battery with a resistance which can be set to a specific level and a switch. When the switch is turned on a current flows which offsets the seismometer mass. Also, current flows through the shunt resistor and the voltage drop across the shunt resistor is equal to the applied voltage on the seismometer. The diagram of the output of the release test is shown in the lower part of Figure 7. Up until the time that you release the mass the voltage across the shunt is E_0 . After that point, the measured voltage is generated by the seismometer as the mass returns to the equilibrium position. You can see the first maximum and then the voltage continues to oscillate if the damping is small enough. This response is characterized by the various peaks and their amplitudes and the time between the peaks. Those are things we have to measure in order to get on with the analysis.

If you differentiate the equation of motion, it takes this form.

$$\frac{dE_s(t)}{dt} = \frac{S}{R+S} \frac{G^2 I}{M} \frac{e^{-\beta \Omega_0 t}}{\sqrt{1-\beta^2}} \sin\left(\sqrt{1-\beta^2} \Omega_0 t - \tan^{-1} \frac{\sqrt{1-\beta^2}}{\beta}\right) \quad (26)$$

From equation 26, you get that the successive zeros and extrema of $E_s(t)$ occur according to

$$\begin{aligned} \sqrt{1-\beta^2} \Omega_0 T &= n\pi & n=1, 2, 3, \dots & : \text{zeroes} \\ \sqrt{1-\beta^2} \Omega_0 t - \tan^{-1} \frac{\sqrt{1-\beta^2}}{\beta} &= m\pi & m=1, 2, 3, \dots & : \text{extrema} \end{aligned} \quad (27)$$

Both successive zeros and successive extrema follow one another at intervals of

$$T = \frac{T_0}{2\sqrt{1-\beta^2}} \quad (28)$$

Where T_0 is the free period of the seismometer and T is the damped free period (what you measure). With the regard to the damping, the ratio of one extremum divided by the

following extremum is given by

$$\frac{E_{s_i}}{E_{s_{i+1}}} = e^{\frac{\pi\beta}{\sqrt{1-\beta^2}}} \quad (29)$$

We define λ as the natural logarithm of that ratio and call it the logarithmic decrement. From that we can get the total damping governing the rate at which the sine wave dies out. This is given by equation 30 as:

$$\beta = \frac{1}{\sqrt{1 + \left(\frac{\pi}{\lambda}\right)^2}} \quad (30)$$

Here β contains both the open circuit damping and the electrical damping. From the definition of β , β_0 , and β_1

$$\beta - \beta_0 = \beta_1 = \frac{1}{2M\Omega_0} \frac{G^2}{R+S} \quad (31)$$

I find it convenient to define a parameter Γ which I call the damping factor as follows:

$$\Gamma = \frac{G^2}{2M\Omega_0} = (\beta - \beta_0) (R+S) \quad (32)$$

From that, one can calculate the motor constant as

$$G = \frac{1}{\sqrt{2M\Omega_0\Gamma}} \quad (cgs \quad emu) \quad (33)$$

Here, Γ has been determined experimentally, Ω_0 has been determined by measurements of successive zeros, and M is the mass of the seismometer. For that, you must either take the seismometer apart and weigh it, or else insist that the manufacturer honestly tell you the exact weight of the mass in each seismometer. Also, we want to operate the seismometer at

some specified level of damping. Commonly, the level chosen is about 0.7 to 0.8 of critical. From the defining equation for Γ we can solve for the value of the external resistance (D) required for a specific level of damping. For a damping value of 0.8 we have:

$$D = \Gamma \left(\frac{1}{0.8 - \beta_0} \right) - R \quad (34)$$

In some experiments, one records the open circuit damping directly by deflecting the seismometer and cutting off the deflection voltage while no shunt resistor is connected across the seismometer. That is not always an easy thing to do correctly. Alternatively, you can carry out two damping tests with different values of shunt resistance S and S' . This leads to the following equations:

$$\begin{aligned} (\beta - \beta_0) (R + S) &= \Gamma \\ (\beta' - \beta_0) (R + S') &= \Gamma \end{aligned} \quad (35)$$

Γ is a physical constant of the system so the only unknown is β_0 which can be solved for as:

$$\beta_0 = \frac{\beta' (R + S') - \beta (R + S)}{S' - S} \quad (36)$$

Table 1 shows the test results from a particular L-4 seismometer. As you can see, 4 different values of shunt resistor were used ranging in value from open circuit to 13280 ohms. The values of Γ obtained this way deviated from each other by about a percent or less. The experiment was carried out rather crudely using a storage oscilloscope. It is difficult to be accurate with this setup, especially for smaller values of resistance since there are only one or two recognizable peaks and the second is rather small. Using the equations relating Γ to the motor constant G we get

Table 1. Damping Test Results on L4-C Seismometer

R = 5580 ohms

M = 980.0 grams

T₀ = 1.039 seconds

S (ohms)	β	$\beta - \beta_0$	R+S (ohms)	Γ	dev (%)
∞	0.271	0.000	∞	----	----
20155.	0.507	0.236	25695.	6064.	+0.6
14999.	0.561	0.290	20579.	5968.	-1.0
13280.	0.592	0.321	18860.	6054.	+0.4
average = 6029.					

$$G = 2.672 \times 10^8 \text{ emu / cm / sec}$$

$$G = 2.672 \text{ volts / cm / sec}$$

There are other ways of going through these equations. You can, indeed, solve them on the basis of a single open circuit damping test if you record the results with something like a PC where you have a very accurate record of the decaying wave train. The difference between the two approaches is that you do have to use the value of the first maximum and the ratio of that to the offsetting voltage to get the motor constant. Since you have essentially carried out an open circuit test, the damping ratio gives you the open circuit damping directly and from that and the measured interval between zeroes you can calculate the free period.

We started out doing it this way because early on I was concerned that cutting off the current in the coil which produces a transient as the inductance tries to keep the current flowing would produce a spike which would confuse the interpretation of the actual height of the first peak. In these damping tests, it doesn't matter how you set the seismometer in motion as long as that activity occurs before the first peaks that you start to use in the

analysis.

Mark Products is now marketing a little card that plugs into a PC and into which you can plug a seismometer, set up the parameters correctly, and punch a button and it gives you back these figures. It solves the problem of calibration. I plan to look into it further and convince myself that it is sufficiently precise. It would be a tremendous boon to be able to set a seismometer on a bench and calibrate in a few seconds, rather than go through these fairly elaborate and difficult experiments.

One of the problems with a seismometer that you can't get inside of, particularly one with a free period as long as one second is that its calibration can change significantly over time. On

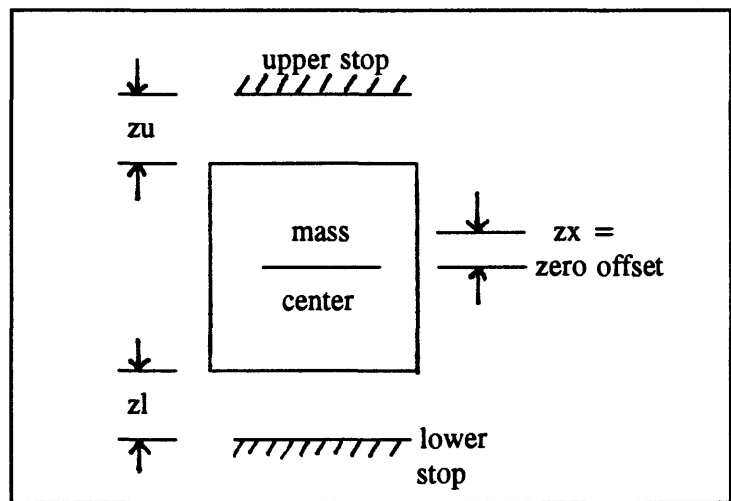


Figure 8. Release-from-stop test.

some seismometers and with some springs, it appears that problems occur, perhaps local work hardening of some parts of the spring. So even though the mass may be perfectly centered when the seismometer is new, over a period of some months or years, the mass gradually settles. Our experience is that after a period of five years in service, a large fraction of the seismometers have periods that are clearly wrong. You need some way of detecting that, or detecting whether the masses are off center.

With the horizontal component seismometers, the problem is even worse because one

tenth of a degree off level is enough to drive the seismometer mass against the stop. There is no indicator on the outside of the case where the mass actually rests. In the three component units a bubble on the case is intended to allow leveling of the internal components, but a little bit of rough handling changes the system axis so the system goes out of adjustment. The release test provides an easy way of getting at that. If the seismometer is symmetrical and if you put a big enough current to drive the seismometer against its stop, first one way and then the opposite way, the amplitudes should be the same. From this, we can derive an equation relating the upper voltage, lower voltage and the zero offset. This relation is:

$$\frac{EU - EL}{EU + EL} = \frac{2 ZX}{EU + EL} \quad (37)$$

One final thing that turns out to be important is that when you calculate the motor constant and other properties of the seismometer, you find that there is a variation in that motor constant. If you want to have standard seismometers, each one performing like every other one, you want to have some way of connecting the seismometer to the input of the amplifier such that all of the seismometers look alike for a given earth motion. You need to be able to adjust the motor constant to some standard value and you want to be able to set the damping to a prescribed value. The diagram and associated equations that are in the handout run through the procedure for doing that.

SUGGESTED READING

Eaton, J. P. (1991). Calibration of a linear spring-supported, vertical-component, moving-coil seismometer by means of damping tests and/or a current release test, U.S. Geol. Surv. Open-file Report, 29 pp.

4. BASIC TECHNIQUES FOR TELEMETRY

by

J. R. VanSchaack
U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

The Branch of Seismology has been operating seismic networks since 1966 in California. We started with about half a dozen telemetered stations made by Teledyne Geotech. However, the quality was not what we wanted, so we discussed specifications with a company in Mountain View. They made a Develco 6202 preamplifier/voltage controlled oscillator (VCO), which was a very good unit for our purposes. However, prices escalated rapidly, and in three years unit prices doubled. At that time, we had installed about 100 units, but our budget would not allow us to buy any more. For the future, we decided that it would be better to not be subject to what the market place had to offer, so we decided to make something of our own. We went through several versions of VCOs numbered J201, J301, J401, J501, and J512, which is where we are now. All of these instruments are basically the same. The seismic signal is amplified and modulates a voltage controlled oscillator with frequency modulation in the audio-frequency range.

What took us 15 years to learn about networks, I am going to try to explain to you in about 30 minutes, so I will miss a lot of the good stuff. I want to concentrate on the general types of things you have to be concerned about with a seismic network whether it is very large or very small. Also I will touch on some of the specifications that turned out to be very

critical, that you have to be very careful of.

All of our sites are telemetered to Menlo Park with the exception of one that is right outside the building, and at each site we use one of two general types of instrumentation. One runs off internal batteries and into telephone lines, and the other requires an external power source and is generally connected to a radio transmitter. The attached station list is intended to give you an idea of what we do here in Menlo Park. It isn't that we do anything that is particularly complex. We just do a lot of it.

NETWORK RELATED TO SENSING AND TELEMETRY

The type of instrumentation that we use is determined by the types of earthquakes we are interested in recording and the areal range of coverage we need. Another factor to consider is the duration of monitoring. If you just are going to be monitoring for a few weeks, you will set up the network quite differently than if you are going to be monitoring for 20 or 30 years. The location of the network has to do with the types of telemetry that you will use to get the data back to a central point. If you have to go to an exceptionally remote location, you may need to take the data processing center to the region you wish to monitor. A PC system is, of course, exceptionally well suited for this sort of application. However, telemetry can be done for long distances. We currently telemeter data from the Oregon border to Menlo Park. For quite a few years we had data coming from Oregon, Washington, Yellowstone Park, and Colorado to Menlo Park. At that time, the telephone lines were very cheap and there wasn't any problem with doing it. These days, because of increased costs, we have had to cut way back on the area that we cover and contract some of

the other networks.

A very important consideration is how you are going to handle network timing. If you want to learn anything about the earthquakes you record, you must have accurate timing. One hundredth of a second accuracy is the standard for our networks. That is adequate for earthquakes with dominant frequencies of 5 to 15 Hz or less.

There are several ways to get timing. We have satellite timing, rubidium frequency standard, and WWVB. Our base timing now is a clock synchronized by satellite. WWVB is commonly used for remote sites. It is very good for a 24-hour period, but in any one second interval, you can have a ten to twenty millisecond error because of the propagation time from Colorado to your station and because of phase changes in the signal. You need to pay attention to that when you are looking at data. You can convince yourself that this really is a problem if you use a storage oscilloscope to look at the timing pulses from WWVB. After you have collected enough samples, there will be a smear about ten to 20 milliseconds in width. So, WWVB is good for a long term average, but you never can tell within about 20 milliseconds where you are if you use just one tick.

INSTRUMENTAL CONSIDERATIONS AT MONITORING SITES

The sensors that we use in the California network are exclusively the 1-Hz Mark Products L-4. This is primarily because of the economics involved in deploying this many sensors. You could also use 2-Hz seismometers for the same price (about \$750). If you are installing a very small, close in network, you could probably use the very small L-22 seismometer, which costs about \$375. Most other sensors cost more than \$1200 per

component.

Another important consideration at the field site is signal conditioning, amplification, and filtering of the seismometer output to prepare the signal for telemetry. The means you choose to employ for this step can have a significant effect on the power requirements at the remote site. If you have a high power system you also need to concern yourself with getting power to the site, installing solar panels, or putting in lots of batteries. This becomes an important factor in the economics of each site.

The telemetry method that we employ exclusively in the northern California network is frequency modulation and multiplexing. However, we are in the process of putting in some digital telemetered stations now. As you know, the analog method has a dynamic range restriction, but you can get many more signals on a carrier. With analog telemetry, on a single voice grade, unconditioned phone line, you can get eight channels easily and a ninth for just a little more trouble. With conditioned lines you can get a somewhat better signal back, but it will cost you about three times as much and the improvement is less than 6 dB in dynamic range.

We don't have much experience with digital telemetry yet. We have been talking with manufacturers about equipment and currently are able to get 9600 baud continuous on a voice grade line. There is one local network in southern California that is employing digital telemetry and transmitting three component, 16 bit, 200 samples per second data. The 15 stations are collected into a microwave link for transmission to the data processing center. This is an expensive and complex system.

The telemetry paths that are available are phone lines, terrestrial radio, satellite radio,

and combined paths. We started off almost exclusively with telephone lines because they were very cheap, about \$20 per month per station. That cost now is about \$150 per month per station. The terrestrial (FM) radios we use operate in either the 160MHz, 200MHz, or 400MHz bands and are low cost, low power systems. We have about 240 to 250 radio pairs in northern California and we have about 100 radios in southern California. The Yellowstone network has about 15 stations using VHF telemetry. Commonly, we use a combined path in which the first link is by radio, the second link is by phone line, and finally the signal is placed on our microwave system.

We installed a microwave system which connects this area to the Parkfield area strictly because of economics. We could save enough money on the telemetry cost to buy more than a link per year. In about three or four years we had the whole network paid for. It takes one technician about 2/3 time to keep the system up. The microwave network has been running about five years now, and has a very good uptime. It is a redundant system with two transmitters and two receivers running all the time, so if one fails the other still carries the load. We have found that the system has a higher percentage of ontime then a telephone line does running through the commercial companies. Also, it turned out that our microwave system worked through the Loma Prieta earthquake, but the phone system did not. The microwave system costs about \$35,000 per link (about 40 miles). It works very well for the San Andreas fault network because it is a strip network and we run the microwave right down the middle of the strip. For areas beyond the range of the microwave system we get the data either by phone lines or by radio link. If you are operating a long network that is configured to monitor something like a fault zone, then microwave communication turns out

to be relatively cheap. Our most expensive telephone lines on the network now are from San Jose to Menlo Park. There are several exchanges on the way and the price is about \$750 to bring eight stations 30 miles. We are in the process of converting most of those to radio stations, both because of the cost and because we think we can keep the system operable during an earthquake by making these changes.

PRACTICAL NETWORK INSTRUMENT SPECIFICATIONS

Table 1 shows my view of what constitutes a practical set of instrument specifications for a seismic network. Experience has taught us that some of the companies manufacturing telemetry equipment do not make equipment which meets these specifications. Choice of an appropriate sensor is at least partly dependent on what it is that you are trying to observe. If you are trying to record strong ground motion you will want a very different sensor than if you are trying to monitor microearthquake activity. Although our network is primarily devoted to monitoring microearthquake activity, we are installing some stations which have a wider dynamic range so that we will be able to record some of the larger events on scale.

Our experience indicates that if the motor constant of your sensor is in the range 1 volt per cm/sec at 0.8 critical damping, then you will have adequate seismic signal to override electronic noise no matter what the gain of the system. The Mark Products L-22 works well in most places where there is a fair amount (10 - 15 μ v) of background noise. However, there seems to be a problem in a well experiment being conducted in Parkfield, in that the Brownian noise of the L-22 seismometer is greater than the earth noise. The L-22 can produce 0.5 volts per cm/sec into a 100 K-ohm load. Practically, an L-22 can achieve

Table 1 Practical Network Instrument Specifications

Sensor

- resonant frequency 1hz \pm 5 %
- motor constant 1Volt/cm/sec at 0.8 critical damping
- dual coil for AC canceling
- waterproof

Signal Conditioners

- amplifier gain, variable 50 - 90 dB
- input, differential 10 - 100K impedance
- filtering 12 dB bandpass, 0.1 Hz - 25 Hz
- system noise < 1 μ V PP, referred to input

Voltage Controlled Oscillator

- sine wave stability 20 - 50 ppm / °C
- sine wave distortion < 1 %
- sensitivity 5V PP --> \pm 125 Hz deviation
- linearity better than 1 %

Power Requirements, Field System \approx 200 mW**Discriminator Specifications**

- signal input level -35 to -15 dB
- discriminator Stability 20 ppm / °C
- bandpass filters 18 to 20 dB attenuation at adjacent band edge
- output level \pm 5V / \pm 125 Hz deviation
- filtering 5 pole lowpass, 20 - 25 Hz
- power requirement 150 - 250 mW

Radio Specifications

- modulation type FM of phase modulation
- frequency bands 160-170 MHz, 216-220 MHz, 406-420 MHz, etc.
- frequency stability .0005 %
- trans., rec. audio dist. < 2 % each
- total distortion < 3 %
- transmitter output power 100-200 mW output
- audio output level -20 - 0 dbm adjustable
- power req. transmitter < 600 mW for 100 mW output
- power req. receiver 150 mW

Overall dynamic range is approximately 50 dB when distortion of carriers is 2 % or less.

about a fourth of the signal possible with the L-4 seismometer for the same movement.

Another problem associated with using a 2-Hz geophone is that because of the low frequency

cutoff, earthquakes above about magnitude 3 will start to all look alike since larger earthquakes have more low-frequency content.

If you are using a coil-wound sensor it is best to use a version with dual windings, since these can be connected to a differential input amplifier which will help provide cancellation of 60-Hz line noise. It is important to buy a sensor that is waterproof, especially if you are leaving it out for several years. We bury Mark Products L-4 seismometers in the earth about a foot deep. We do put the sensor inside a piece of sewer pipe before we bury it. However, that is not for the purpose of waterproofing it. Rather, it is for electrical insulation from the ground. It is preferable to have only one ground point for one of these systems. If a sensor is out twenty or thirty feet, and is in contact with the earth, it makes a very good CW receiver and, at least in this area, you can pick up the continent to continent teletype signals very well.

An amplifier should have gain of the order of 50 to 90 dB with a differential input. Not everyone thinks differential input is critical. However, I am convinced that the differential input configuration gets rid of lots of common-mode noise, particularly in a populated area. In such places there are lots of sources of AC power that cause problems.

The filtering that is used in the J series preamp/vco's is bandpass from 0.1 Hz to 20 Hz with 12 dB rolloff at each end. We selected 20-Hz low-pass to get less noise and to get away from some problems that can occur when you modulate the carrier at 100% at frequencies above 20 Hz on these constant bandwidth systems. The system noise should be less than 1 microvolt peak to peak referred to the input. You can buy a 19-cent integrated circuit that will get you near that range. However, in our units we use an OP-77 amplifier by

Precision Microlithics. This unit has about half a microvolt noise in the range in which we are interested. It provides a simple front end for the system, and if a lightning strike occurs, you just unplug the unit and plug a new one in. They are a couple dollars each. Ease of maintenance is very important. Right now we have two field people working full time maintaining 360 sites. We have one person who works on special assignments to places like Mammoth Lakes or Yellowstone Park. Because of this, it is important that things go together easily.

The voltage controlled oscillator must be very stable in order not to produce drift or false signals. The constant bandwidth multiplexing system allows you to put eight signals on one line with carrier center frequencies spaced at 340 Hz. The first carrier is 680 Hz, the next is 1020 Hz, and so on. The companies manufacturing these multiplexing systems have standardized on this constant bandwidth approach with these frequencies. One of the most critical specifications is the sine wave distortion. We recommend maximum distortion of less than 1% in order to get good dynamic range. If distortion exceeds this in a multiplexed system you start getting beat frequencies. Because each one of these are separated by 340 Hz, you get difference frequencies that can match one of the channel center frequencies. For instance, 1700 Hz minus 1020 Hz gives the channel one frequency of 680 Hz. This difference frequency can beat against the 680 Hz carrier and produce a noise signal. You get a little amplitude modulation in the carrier, which in the discriminators shows up as periodic noise usually. You can easily see this noise on a developocorder record. It is a lot harder to get rid of than it is to identify.

Our VCOs generate sine waves by means of a D/A conversion scheme. An oscillator

operating at 16 times the center frequency of the VCO output produces a square wave which is fed through the D/A convertor to produce the sine wave. The output is very clean and the system is inexpensive to build.

The industry standard for VCO sensitivity is 5V PP input producing ± 125 Hz deviation in the carrier. We have actually cut back on this to about ± 110 Hz to give us a little bit more space between channels to get a better signal to noise ratio.

Linearity is also important. When the carrier swings between its extreme frequencies, you want to have it done in proportion to the voltage levels driving the VCO. We recommend linearity of 1% or better.

Although some of our VCOs have very low power requirements, most draw about 200mW of power. The majority of the units are installed in a field package that contains a radio transmitter which has fairly high power requirements. Therefore most of the sites have a solar panel and batteries which provide enough excess power to run the VCOs which draw 200 mW of power.

Our systems typically contain a 200 mW VCO and a radio transmitter with a power output of about 100mW. We operate both off of a solar panel with 9 Watt peak output and a storage battery. That works very well in California, Oregon, and up into Montana even during the darkest days in midwinter. That solar panel plus the battery costs about \$150.

The discriminator is the demodulator we use at the central recording site. Its output, ideally, reproduces the original seismic signals, which can be digitized or fed to a drum recorder, etc. The discriminator should operate with an input signal level over a range of about -5 dB to -35 dB. The discriminator stability should be about the same as the VCO

stability, even if you are in an air conditioned building. Each discriminator contains a bandpass filter whose function is to only accept signals with the right center frequency for that discriminator. The filter characteristics are important to the overall performance of the discriminator. We recommend that the filters provide 18 to 20 dB attenuation at the edge of adjacent bands. That way even if a channel is modulated at 100%, there will still be an allowable dynamic range of greater than 40dB without cross talk. The output level of the discriminator is not particularly critical, especially if the output is being fed directly into an A/D board. Most companies manufacture discriminators with an adjustable output level. Typically the output signal is fed through a four or five pole low-pass filter with a cut off at about 20 to 25 Hz. On our system we sample at 100 samples per second and use a five-pole filter with a 20-Hz low-pass characteristic and have experienced no problems with aliasing. Typically, discriminators will have a power requirement of about 150 to 250 mW.

Radios are a critical part of a telemetry system. Commercial companies don't generally use the radio bands the Federal government uses. We use the bands from 160 to 170 MHz, 216 to 220 MHz, and 406 to 420 MHz. These bands all work about the same. New radios use a different crystal filter setup from what used to be common, and a consequence is that the selectivity of current radios is not as good as what used to be available. However, they are cheaper. So, in a congested area you may have to put an RF bandpass filter in your system at a cost of about \$150 in order to reduce interference from adjacent channels.

You can purchase radios with frequency modulation or phase modulation. Phase modulation was used by Motorola in their old Handi-talkies years ago. As a matter of fact,

many of the radios in our system use surplus boards from Handi-talkies. They generally meet the specifications we needed. One problem is that they have high audio distortion. We have been trying for several years to replace them, but a radio pair costs about \$1000 these days so we haven't gotten very far in replacing them. Frequency stability is dictated by the FCC for a fixed site. The required frequency stability of .0005% can be achieved easily at the \$1000 a pair level.

Transmitters and receivers should have less than 2% audio distortion each. When distortion reaches a level of about 3% in either unit, you start limiting your dynamic range. When it reaches a level of even 5%, you have serious problems. Distortion is cumulative over the various telemetry devices, so to keep the total distortion low enough each piece must have very low distortion. If you want to have a clean system you must keep the total distortion below 3%. This distortion level is perhaps the single most important specification in the entire telemetry system. It is worth a lot of money and effort. Otherwise you find yourself chasing what you think is telephone line noise, radio noise, or some other kind of noise that isn't really there. It also severely limits your dynamic range. At a total distortion around 1%, the dynamic range is around 60dB. When you get up to 4 or 5 percent distortion, the dynamic range has decreased to around 30dB.

Nearly all our transmitters have a power output of around 100 to 200 mW. For this power output, the input power requirement is around 600 mW. Some companies would like to sell you transmitters for a local network with power ten to 100 times that. However, we have had good success with our low-power systems. Part of the reason may be that we use high gain antennas (9dB Yagi) on both the transmitters and the receivers. If you buy a \$150

Table 2 Computed Receive Signal Strength Level

Antenna gain, receive (dbm)
Antenna gain, transmit (dbm)
coaxial lead loss, receive (dbm)
Coaxial lead loss, transmit (dbm)

Transmitted Power (dbm)

1 mW = 0 dbm
10 mW = 10 dbm
100 mW = 20 dbm
1 watt = 30 dbm
10 watt = 40 dbm

Receive Signal Strength Level (dbm) = transmit power (dbm)
+ antenna gain (Tx)
+ antenna gain (Rx)
- coax loss (Tx)
- coax loss (Rx)
- free space loss

where free space loss (dbm) = $36.6 + 20 \log \text{freq. (MHz)} + 20 \log \text{distance (mi)}$.

Yagi antenna, the effect is pretty much the same as multiplying transmitter power by a factor of ten. Another advantage of using a highly directional antenna is that you may be able to reuse the same frequency off the side lobes of the antenna, since the signal is so strongly attenuated there. We have a network that we installed over in Jordan a few years ago in which a 100 mW transmitter is successfully coupled to a receiver 115 miles away! The network is installed on a flat plateau so there probably isn't even true line of sight between the transmitter and the receiver. We would have recommended an intermediate station, but there was no practical location for it.

The number of repeaters in a telemetry path relates to the earlier discussion about distortion. Every time you repeat a signal you add the distortion of the receiver and the transmitter to the cumulative distortion. This means that with even moderate distortion levels

for individual pieces of equipment, a repeated signal can easily have cumulative distortion that is too high to be usable. In Hawaii the volcano observatory has up to four repeats on some of their telemetry links, which requires using very good equipment. We are limited to one repeat here because of the old Motorola equipment we are using. Occasionally we go up to two repeats, but we know that we are generating noise when we do that.

All of the frequency bands we use require line-of-sight radio paths. There is, however, some bending of the radio wave path to follow the curvature of the earth. It is sometimes possible to receive signals here from south of Parkfield, a total distance of about 280 miles. This also happened in Nevada, where we could transmit 100 mW signals and receive them 250 miles away (but not all the time). Signals of 100 milliwatts will always travel 100 kilometers with 9 dB gain antennae at each end. Receive signal strength of -95 dB is adequate in almost all environments (Table 2).

We have developed our own VCOs and discriminators. Our VCOs operate on voltages of from about 11 to 15 volts DC. They have a built-in calibrator circuit that offsets the seismometer mass once every 24 hours. We manufacture these for our own use, but other government agencies can also get these units from us. If you have one of these built, the cost will be about \$350. We plan to publish all the specifications of the current VCOs in a U.S.G.S. Open File Report soon. There is a completely different VCO-amplifier system in use in Alaska. My understanding is that it is designed to operate well in very cold weather over long periods of time. It is more complex and more expensive.

The discriminator that we manufacture is identical in dimensions to the Develco Model 6203 discriminator. Our present discriminator (model J120) is quite a bit better than

earlier versions we were using. It has bandpass filters which attenuate adjacent channels by 18 to 20 dB at the edge of their allowed band. This is in comparison to attenuations of 12 to 14 dB offered by some of the commercial companies. We are also about to release an open file report describing the discriminator. The cost will be about \$85 to \$90 each in lots of 100. We built about 1000 of the J120 discriminators three years ago to replace the old discriminators in all our operations, and we have had excellent performance from these units.

5. REALTIME SEISMIC DATA ACQUISITION

by

W. H. K. Lee

U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

The first thing I want to cover in this chapter are some general issues related to realtime seismic data acquisition. Since everyone's application is unique, you need to understand these issues to be able to configure your system. The general issues I intend to discuss are: computer hardware, operating systems, software languages, data formats, and cost (See Figure 1).

The question of computer hardware really boils down to "What is available?" This includes the question of affordability and availability. The choice of operating system is largely a consideration of what is practical and fits your needs. You can make all kinds of theoretical arguments, but the bottom line is ease of use and suitability. People can argue endlessly about relative merits of different

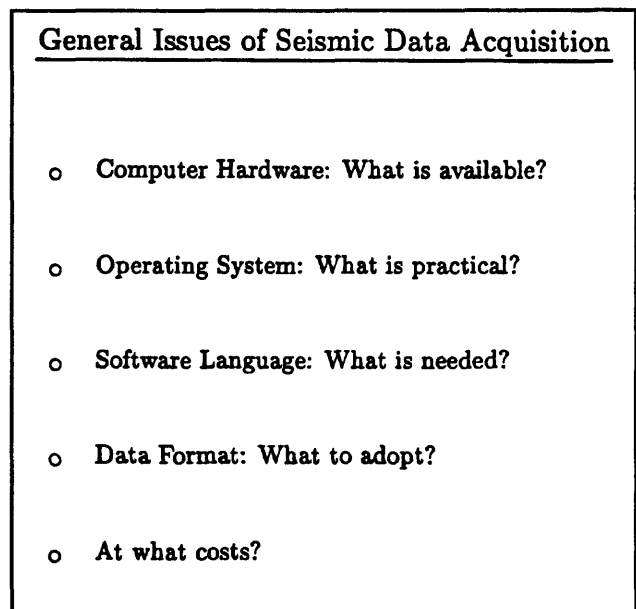


Figure 1. General issues.

software languages, but I think that the important question is "What language will get the job done most efficiently?"

Data formats are also subject to controversy. However, I think that if someone comes up with a format that gets the job done, and if it is adopted by enough users, then it will become the standard which most users will follow. The last general issue I will discuss is the cost of seismic data acquisition. This is something that many people would like to ignore. The cost is not simply the up front purchase price for equipment. Rather, it is the total cost of constructing, maintaining, and operating the seismic data system over its lifetime.

Unfortunately, it often happens that some country will spend a lot of money to install a complicated seismic network. After a few weeks of operation, the network goes down permanently, because the local people just are not able to maintain it. So I keep emphasizing that even though it may be satisfying to have cutting-edge technology, you are going to pay a price for it. Also, most of those latest developed technologies only work in the laboratory environment. It usually takes a few years before they become workable in the field. My advice is to always stay at least one to three years behind whatever the cutting edge of technology is! Let other people become the guinea pigs and work out the solutions for you.

COMPUTER HARDWARE

What computer hardware is available to you depends on the money you have available and also where you live (See Figure 2). In certain countries, for instance, you just can't get SUN workstations and there is no one around who can maintain them. So even if you can afford a SUN, it wouldn't make any sense to use one in such a place.

A question one might ask is "Why not use DEC mini-computers?" These have been available through the 1970's and 1980's. They have been in use for years, so lots of people are familiar with them, and there is a fair amount of software for them. However, there are at least two major objections. First, they have a proprietary architecture, and the second is that they are very expensive. Unless you have a very large budget, they are just too expensive. At several

universities a few years ago, they dropped all their DEC's when Digital Equipment raised the cost of their maintenance contract to the point where they simply couldn't afford to even run the machines anymore. This also illustrates the point that the cost of your system is not just the initial capital cost. It includes operating and maintenance costs as well.

A case can also be made for using SUN workstations. These machines are also very powerful and, in fact, are used by many seismologists. However, these machines also have a proprietary architecture, and are difficult to interface with other hardware. We have a dozen or so SUN workstations in the USGS Seismology Branch at Menlo Park, and we have a full-time person just to keep them going. You have to ask yourself, "Can I afford to have one

Computer Hardware: What is available?

- **Why not DEC Mini-Computers?**
 - PRO: We have been using them for years
 - CON: Proprietary and expensive
- **Why not SUN Workstations?**
 - PRO: Powerful and used by most seismologists
 - CON: Proprietary and difficult to interface
- **Why not Apple Macs?**
 - PRO: Easy to use and cute
 - CON: Proprietary and difficult to interface
- **Why IBM-Compatible PCs?**
 - PRO: Non-proprietary and inexpensive
 - CON: Neither cute nor classy

Figure 2. Computer hardware.

person working full time just to keep the computers going?"

A lot of people these days are using Apple Macintosh computers. They are very easy to use. The problem again is that they have a proprietary architecture and are difficult to interface. You could not even open the case on the original MAC without voiding the warranty. The later MACs are designed more along the lines of the PC. However, you can only buy them from Apple, so you are at their mercy.

As you probably know by now, I am strongly in favor of using IBM-compatible PCs as the computer hardware for seismic data acquisition. The PC does not have a proprietary architecture so anyone can make a PC. This has resulted in them being quite inexpensive for the computing power delivered. Having said that, I still must emphasize that you must make a decision about what hardware to use based on your needs and on your budget. For some applications, the PC just may not be adequate.

OPERATING SYSTEM

The operating system you choose must be practical. Lots of people are using UNIX these days. It has a lot of good features. It supports both multi-tasking and multi-users. However, UNIX is a very complicated operating system with at least 27 different versions, not all of which are compatible with each other. Besides, the UNIX manuals fill up at least three feet of bookshelf space! Do you want to mess with that if you don't have to?

The OS/2 operating system for the PC is a fairly sophisticated multi-tasking operating system which could be used for a data acquisition system. However, OS/2 is still pretty new and because of this, it is not very stable. Also, it has not achieved very widespread

acceptance so the amount of supporting software which runs under OS/2 is rather limited.

We finally settled on using DOS. DOS was originally named QDOS, which stood for "Quick and Dirty Operating System." When IBM decided to market the system they changed the name to DOS which they said stood for "Disk Operating System." The original version of DOS was written by one person. That is

probably one of its greatest strengths, since it is really simple even with all the revisions it has undergone. DOS has been around for a long enough time now that it has become a stable operating system, and it has by far the largest number of users of any other operating system in existence. DOS does not allow multitasking directly. Fortunately, the hardware is cheap enough that for many purposes, you can multitask just by using more than one computer! Another problem with DOS is related to the way it addresses memory. The largest program that can be run under DOS is less than 640 Kbytes in size. It doesn't matter how much physical memory your machine has. That is all DOS is going to use. Also, memory is addressed in 64 kilobyte long segments, so there is a lot of programming overhead involved in specifying a complete memory address. This addressing scheme was forced on DOS by

Operating System: What is practical?

- **Why not UNIX?**
 - PRO: Multi-users and multi-tasks
 - CON: Too complex and too many flavors
- **Why not OS/2?**
 - PRO: Multi-tasks
 - CON: Too new and unstable
- **Why use DOS?**
 - PRO: Simple and stable
 - CON: Can not multi-task
 - CON: 64 KB segments & 640 KB space

Figure 3. Operating system.

the design of the original INTEL 8088 chip, and in order to maintain backward compatibility DOS still uses this method of addressing.

SOFTWARE LANGUAGE

People have been debating the relative merits of software languages for years (See Figure 4). My philosophy is to use whatever language that gets the job done. In practice, multiple languages were used to create the software for the PC-Quake system. The routines that directly control the A/D board were written in assembler. The reasons are (1) faster execution speed can only be achieved with assembly language, and (2) assembler

provides a natural format for manipulating registers and handling interrupts, both of which are integral to controlling the A/D board. In the data acquisition program, the main routines were written in C. Graphic display routines were also written in assembler. Again, this is because of the speed possible with assembler, and the flexibility with which assembler can be used to move information into the display memory. The trigger algorithm was written in C, and the P-picking algorithm and earthquake location program were written in FORTRAN.

Software Language: What is needed?

- o **Use multiple languages out of necessity**
 - o A/D: Assembler and C
 - o Graphic Display: Assembler
 - o Trigger Algorithm: C
 - o P-Picking: Fortran
 - o Event Location: Fortran
- o **Use existing programs to save cost**
 - o A/D: From Data Translation
 - o Graphic Display: From Symmetric Research
 - o Event Location: Modified from HYP071

Figure 4. Software language.

Much of the code used in this system was originally written years ago to run on other computers. Since, historically, FORTRAN has been the programming language of choice for scientists, most scientific code is in FORTRAN. This applies not only to the P-picking algorithm, but also to the event location program.

Whenever possible one should use existing programs to save time and money. For instance, the assembler routines which control the A/D board were provided by the A/D board manufacturer. If you had to write this software from scratch, you would really have a lot of headaches. Not only would you have to deal with all the complexity of assembly language, but you would also have to understand how the A/D board works. Similarly, the graphic display software came from Symmetric Research. Symmetric Research has donated their graphics software to the International Association of Seismology and Physics of the Earth's Interior, so we don't have to pay royalties to them. The event location program is more or less a straight port of the original HYPO71 program (Lee and Lahr, 1975; Lee and Valdes, 1989).

DATA FORMATS

The question of what is the "best" data format is very much an open one (See Figure 5). Right now there are literally hundreds of data formats. Almost every seismologist has created some kind of format for recording earthquake data. My position is that we should not create any new formats until we retire some of the old ones.

One fairly popular format is known as the CUSP format. It has been used for years, there is lots of data recorded in the CUSP format, and many seismologists are familiar with

the format. However, CUSP has one major weakness in my view. The data and indexes are stored in separate files. This means that any time you want to review or reprocess an event you need to access both files and you must be certain that you have matched the correct index file with the data file you intend to analyze. This often can create big problems.

Another candidate for a standard format is the Lamont AH format. This format is widely used on SUN workstations so in that sense, it is somewhat of a standard already. However, in this format, data are stored as floating-point numbers. This is inefficient in terms of storage space requirements.

For our system we have decided to use the SUDS data format developed by Peter Ward (Ward, 1989). This format has been carefully thought out by Peter. I maintain that the most successful designs are often the product of just one or two people. Design by committee just doesn't seem to produce useful products. In the SUDS format, all the information relevant to an event is contained in one file. The file is infinitely expandable, so if you want to add new data to an event, you simply add to its file. You don't have to create a new file.

Data Format: What to Adopt?

- o **Why not CUSP data format?**
 - o PRO: We have been using it for years
 - o CON: Data and indexes are in separate files
- o **Why not Lamont AH data format?**
 - o PRO: Widely used on the SUN workstations
 - o CON: Data stored as floating-point numbers
- o **Why adopt SUDS?**
 - o Carefully thought out by Peter Ward
 - o All event data in one file
 - o Can be expanded indefinitely
 - o Data stored efficiently as binary integers

Figure 5. Data formats.

Also, data are stored very efficiently as binary integers.

Because the SUDS format was written by one person, there is someone to go to when problems come up. In contrast, committees usually get dissolved in a year or two. For instance, when Peter originally designed the format, it was for the SUN-3 computer. When the SUN-4 computer was introduced, a completely different processor was used. Everything changed, so one cannot use SUDS data created on a SUN-3. But I didn't have to do anything about it. Peter worked on the problem every night and on the weekends until the problem was solved. Who knows what would have happened, if the data format had been developed by a committee.

Right now none of the seismic data formats really has more than a small group of adherents. However, I think that the SUDS format may well become the standard in the future. Partly this is just because Peter is so dedicated to maintaining and improving the format. Also, SUDS is more than just a recipe for laying out the data. It involves a philosophy of how data should be handled in general. Since SUDS is the data format for the PC-Quake system, and since there are already about 50 such systems distributed around the world collecting data in SUDS format, there will soon be a lot of data in this format. Also, as the editor of the IASPEI Software Library volumes, I have had the opportunity to insist that all software submitted for publication must support the SUDS format.

We are also working to develop a means of converting all the common data formats into the SUDS format and vice-versa. By increasing the ease with which users can use SUDS data or convert their data into SUDS format, it should become increasingly easy for more and more people to adopt this format. The SUDS that we are using right now is called SUDS

1 (Banfill, 1992). Right now, Peter is working on SUDS 2. He has given it to people to try out in order to get comments or suggestions. By the end of 1992, the format will be frozen and will be called SUDS 3. What Peter and I have agreed on is that SUDS 3 will be a superset of SUDS 1. So, our data acquisition program will still keep on using SUDS 1. SUDS 3 will be totally independent of the computer it is used on. There is some computing overhead involved in making the data files hardware independent. For this reason we have decided not to change our data acquisition program to fully support the SUDS 3 format. We simply don't want to slow the program down. However, in the current system, after the data has been collected, there is a demultiplexing step. In future versions of our system we will give users the option of changing to the SUDS 3 format at that point.

SEISMIC SOFTWARE

One thing not everyone realizes is that software costs far exceed hardware costs (See Figure 6). For example, at least one commercial company sells software roughly equivalent to the IASPEI data acquisition, processing, and analysis volume for \$139,000. By contrast, the first edition of the IASPEI Volume 1 (Lee, 1989) cost only \$50 to SSA

Seismic Software: At what costs?

- Software costs far exceed hardware costs
- Complex software is hard to maintain
- Software is an art depending on individual innovations
- If software is simple to use, then it is hard to write
- Writing good software is harder than writing good papers

Figure 6. Seismic software.

members! Complex software is also hard to maintain. It is nice to have a program that can do everything. But you pay a price for this. Software is an art, and is really dependent on individual innovations. Good software, good format, or even good hardware design is usually traced back to one person. I have never seen a piece of software written by a committee that really worked well. However, one or two people cannot write all the software seismologists need, so when a good piece of software is developed for seismic data analysis, we need to share it. If software is simple to use, then it will be very hard to write. Writing good software is harder than writing good papers. I will "prove" that by contradiction. I can name many good papers. However, I challenge you to name a really good piece of seismological software.

DESIRABLE FEATURES IN SEISMIC DATA ACQUISITION

What I want to do here is describe what I feel are useful features in a seismic data acquisition system (See Figure 7). Not coincidentally, these are more or less the features of our PC-Quake systems. First, the system should handle between one to 128 channels of input data. There are very few networks that require more than 128 channels. The system should be capable of digitizing signals at a rate of at least ten samples per second per channel and occasionally may need to be capable of sampling at rates of over 1,000 samples per second per channel. Currently, a U. S. Bureau of Mines system used to monitor rockbursts is operating at nearly 3,000 samples per second per channel.

The system should also display the digitized data in real time. I think that is very important. The system should have a reasonable dynamic range (12 to 16 bit). The capability

should exist to record data either continuously or by event. It is desirable to have the capability of picking P-wave arrivals and then computing a hypocenter. Of course, there must be off-line analysis support software. We would like the hardware to be inexpensive and easy to maintain.

The PC-quake system is very easy to assemble and get running. I know of one person who assembled

one in 34 minutes, and I routinely assemble a system in less than an hour even while explaining things to other people. It is really not that hard! Even if you are completely new to PCs, it shouldn't take you more than a day or two to get a system up and running.

A final consideration is the software quality. To be really useful the software should be robust and easy to use.

Desirable Features in Seismic Data Acquisition

- Handle 1 to 128 channels of input signals
- Digitize 10 to 1000 samples/sec/channel
- Display digitized data in real time
- 72 - 96 dB dynamic range (12 to 16 bits)
- Save digitized data continuously or by events
- Automatically pick P-arrivals & locate events
- Off-line analysis support software
- Hardware: inexpensive and easy to assemble
- Software: robust and easy to use

Figure 7. Desirable features.

ANALOG TO DIGITAL CONVERSION

In our system, we just use a commercially available A/D board (See Figure 8). The board, manufactured by Data Translation, is available in both 12 bit and 16 bit versions. The 12-bit board (DT2824) comes from the factory set up to do 16 channels. If that fits your requirements, you just plug it in to a PC and go. If you need to monitor more than 16

channels, you will have to connect a multiplexer to the board. In either case, you just plug the A/D board into an empty 16-bit slot in your PC, make the connections to your signal lines, load the software, and you are about ready to begin monitoring! However, I do have to warn you that there are about 40 or 50 jumpers on the A/D board, which are used to configure the board. These are set at the factory and our software uses the board in the default configuration. However, occasionally, a jumper is set differently from the published default setting. So, it is a good idea to check all the jumper settings either using the manual that comes with the board (Data Translation, 1989), or by comparison with a board that you know has been configured properly.

Data Translation provides a set of callable routines that take care of the low level control of the board. Our software uses those routines in all its interactions with the board, since it would have been very difficult to write the control routines from scratch. Also, Data Translation provides a means of using extended memory for data buffers, and we take advantage of their routines in the following way. If the sample rate is high enough relative to

Analog-to-Digital Conversion

- Use Data Translation DT2821 series A/D boards
- Use Data Translation ATLAB software
- Can be either 12-bit or 16-bit resolution
- Use PC's extended memory for data buffers
- Use two DMA channels on 8 MHz AT bus

Advantages

- Inexpensive: \$1,500 for 12-bit; \$3,000 for 16-bit
- Hardware and software commercially available
- Most 286, 386, and 486 PCs with 8 MHz AT bus

Disadvantages

- Must be Data Translation DT 2821 series boards
- Requires a MUX to expand to more channels

Figure 8. A/D conversion.

the processing speed of your computer, then at times the processor may be unable to handle the data as fast as the A/D board sends them. This is called a "buffer overrun", and it will cause the program to stop execution. However, if you have the space in extended memory, the program will just take the data from the A/D board and put them in extended memory in a circular buffer queue. The program then processes buffers of data from the queue as it is able to. As long as the computer catches up before all the available buffers are filled up, no data is lost and everything continues just as though there never were a problem. Of course, if your processor is just too slow relative to the A/D sampling rate, then eventually all the buffers will be filled up and the program will stop.

Another performance feature of the Data Translation board is its use of dual DMA channels. The PC contains a separate chip called a DMA controller chip, which can move blocks of data from one device to another without using much of the CPU processing capability. The Data Translation software allows the continuous movement of data by the DMA controller on alternating DMA channels.

The Data Translation board is relatively inexpensive. The 12-bit version costs about \$1,500 and has a maximum data rate of 50,000 samples per second. The 16-bit version costs about \$3,000. Data Translation also makes a board with a maximum sample rate of 250,000 samples per second. Also, the board is available worldwide and will work with most 286, 386, and 486 IBM-compatible PCs. We did find a few PCs that would not work with the board, and in every case the reason was that the machine did not use the DMA controller chip made by INTEL. We have a list of all the computers that we know work (Lee, 1991).

One disadvantage of our system is that the A/D board must be a Data Translation

DT2821 series board. The problem with having a single source for such a critical part is that you are at the mercy of the manufacturer. However, last week I saw a brochure for another company which is making A/D boards that are compatible with the Data Translation boards. I haven't tried their board, but I think it is a good sign that other people are making Data Translation compatible boards. Recently, Liu (private communication, 1991) has ported the XDETECT program (version 2.02) to work with A/D boards by Microstar and IOtech.

Another disadvantage with our system is that a multiplexer (MUX) is required if you want to do more than 16 channels. Actually, that is not a big disadvantage. It is far cheaper to use a MUX than it is to provide the capability of directly digitizing all those channels.

REAL TIME DISPLAY OF DATA

Our system is unique in that you can see the data that is being collected (See Figure 9). You can display only 32 channels at a time, but you can scroll through the channels in blocks of 16. This makes it quite easy to look at any channel you want. This practice is in accord with the "what you see is what you get" school of thought, that has become so prevalent in commercial software these days. All the important numbers are displayed. Hard disk free space, number of events captured, how long the system has been up, etc. are displayed at the top of the screen.

The program will work with Hercules, EGA and VGA graphics cards. Although we have not supported any higher resolution graphics displays, we are doing something about it now. Symmetric research has added support for 1024 by 768 pixels in their graphics library. We have not yet modified the code to take advantage of this in our realtime software, but we

are beginning to use this resolution in some of our off-line analysis programs.

One slight awkwardness with our data display is that the seismic signals do not scroll across the screen smoothly. Instead, the data is displayed in frames and the screen changes a frame at a time. This means a delay of a few seconds from when the data reaches the PC to when you see it.

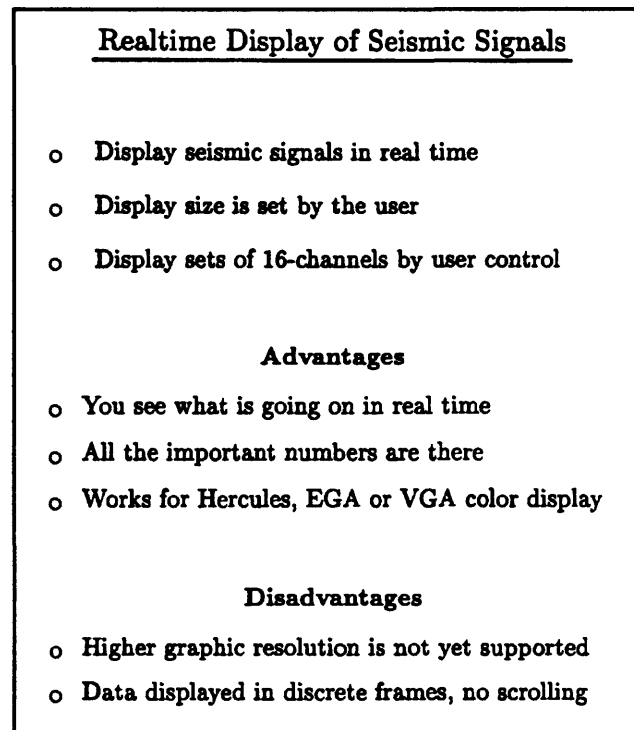


Figure 9. Realtime display.

TRIGGER ALGORITHM

There are two types of algorithm being used (See Figure 10). One is for local events and one is for teleseismic events. To cut down on false triggers we require triggers in multiple channels. We have used two types of trigger algorithm for local events. The first one was a modification by John Rogers of a trigger algorithm that has been in the literature for the last one or two decades (Tottingham et al., 1989). That turned out to be too slow, so we implemented an algorithm that was originally developed in 1969 by Eaton, Lee, and Stewart (See Lee and Stewart, 1981). The teleseismic trigger we use was written by John Evans (Evans and Allen, 1983). Our local trigger algorithm is simple, robust, and efficient.

Recently, we have implemented subnet triggering. Currently, the program only supports one trigger algorithm at a time. This means that there are two versions of the program; XDETECT for local events and TDETECT for teleseismic events.

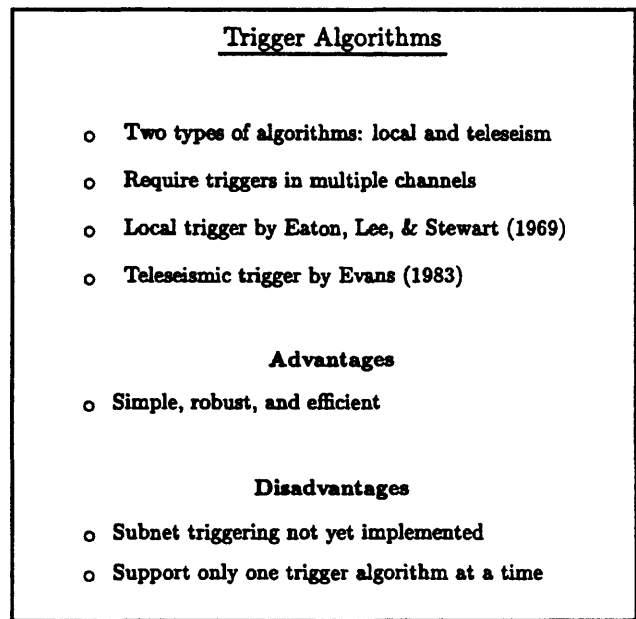


Figure 10. Trigger algorithms.

LOCAL TRIGGER ALGORITHM

The first step in processing is to take the absolute value of the first difference of the incoming time series (See Figure 11). This filters out long-term variation in the signal. The algorithm then calculates recursive short- and long-term averages in a moving window over the data. The algorithm calculates the averages recursively to save on storage requirements and on arithmetic operations. The algorithm continuously compares the ratio of the first difference to the long-term average. Whenever the signal changes abruptly, this ratio will be very high, and if it exceeds some preset level, a trigger flag is set. To confirm the trigger, the ratio of the short-term average to the long-term average is also computed. If it exceeds a different preset level, then the trigger is confirmed.

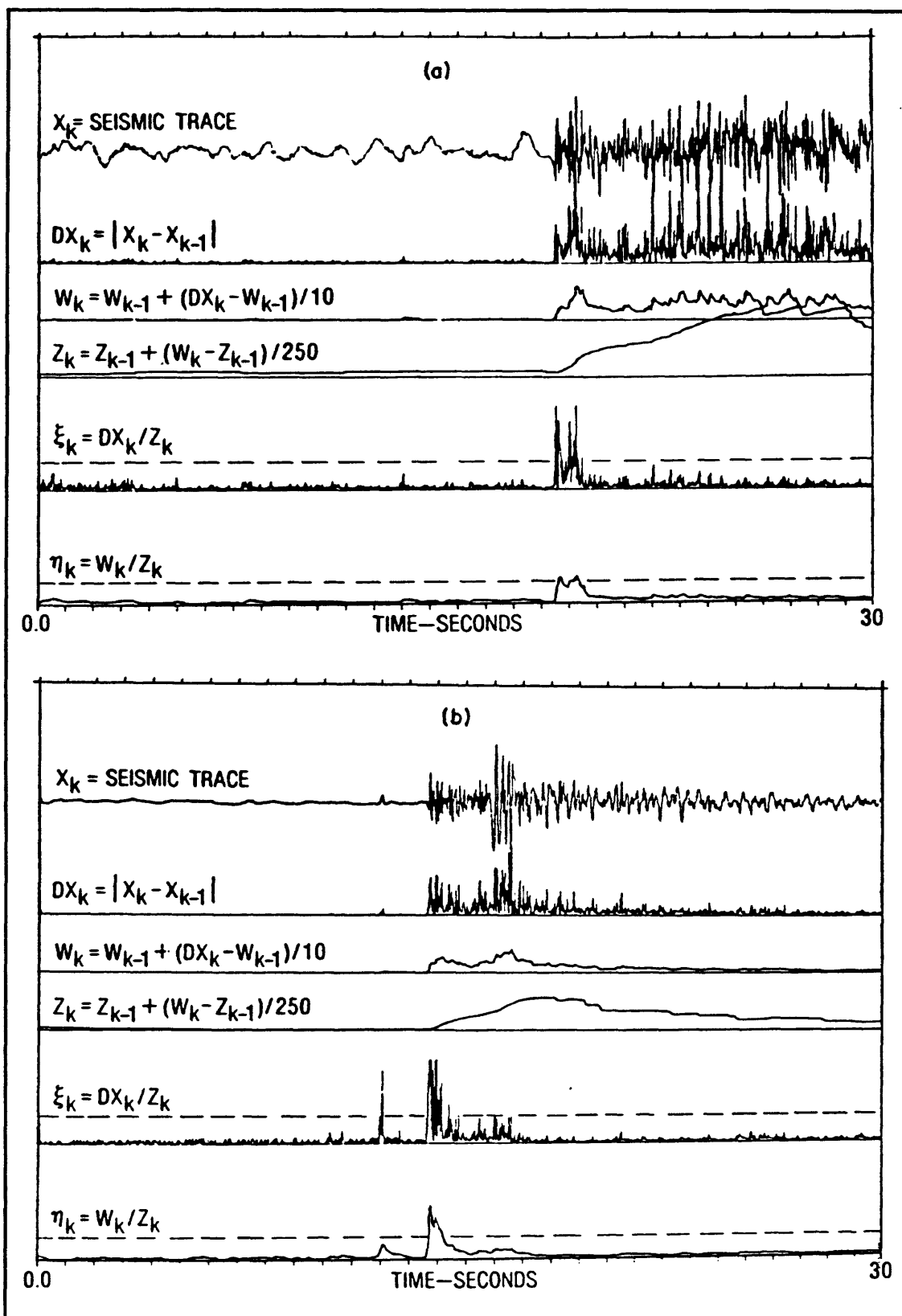


Figure 11. Operation of trigger algorithm.

P-PICKING AND EVENT LOCATION

Although people have been working for decades on the problem of automated P-picking and hypocenter location, there are still no fool-proof algorithms. In our work we have used three different P-picking algorithms (See Figure 12). The original algorithm by Jerry, Sam, and myself was really very primitive. It used the ratio of the instantaneous signal value to the short-term average as the decision criterion and used the onset of the trigger flag as the P-pick. Often that did not work too well. In 1978 Rex Allen published a P-picking algorithm (Allen, 1978), which has been widely used around the world and we have used that algorithm also. More recently, we have been experimenting with a moving window algorithm. This algorithm is much faster than Rex's algorithm and has some other advantages as well.

We use two different hypocenter location programs. One is a slightly modified version of HYPO71, and the other is a simple half-space algorithm written by John Lahr. It is very quick, so it is well suited for realtime hypocenter location.

In the course of developing our system we have come to realize that since hardware is

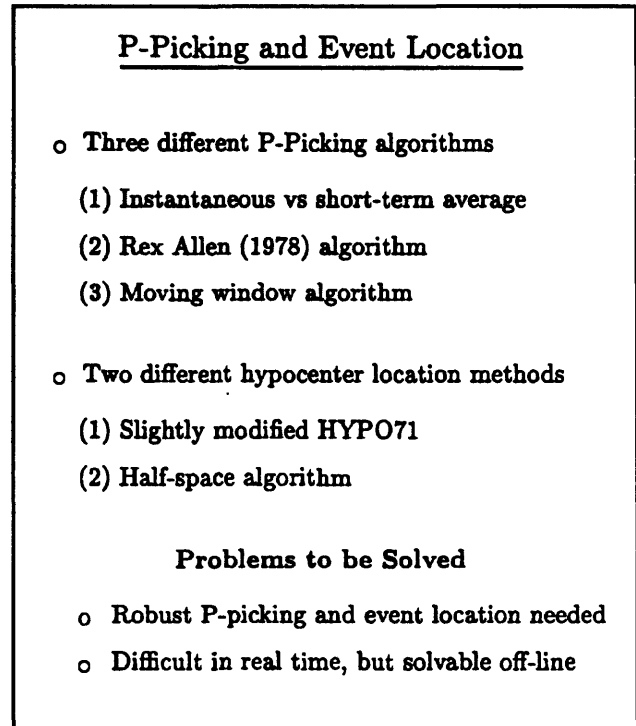


Figure 12. P-picking.

cheap, it is better to divide up the seismic network tasks among several computers, rather than to devise clever schemes which will allow one computer to do everything. Our current approach is to use one machine strictly for data acquisition, and another for all the time consuming data analysis when an earthquake occurred. This works well since earthquakes usually make up only a small percent of the total data stream. This means that the off-line machine has much more time to solve a particular data analysis problem than the on-line machine could afford to devote. Even in the event of an earthquake swarm, as long as the on-line machine is writing to a big hard disk, no data will be lost and the off-line machine will eventually catch up.

BASIC PC-QUAKE SYSTEM

We use two fairly similar PCs in our system, one on-line and one off-line (See Figure 13). Not only does this allow the processing advantages already discussed, but in the event of failure of the on-line machine, the other machine can easily take its place so that there is a minimal loss of data. The current configuration for the on-line machine is a 386 PC with 4 MBytes of RAM and a 300 Mbyte hard

<u>Basic PC-Quake Seismic System</u>	
o	Use 2 PCs: on-line, and off-line
o	On-line PC: 386 PC with 4 MB RAM, 300 MB hard disk
o	Off-line PC: 486 PC with 16 MB RAM, 300 MB hard disk, DSP board, laser printer, WORM drive
o	Software: DOS, IASPEI, KEDIT, PCTOOLS, XTREE
o	Total Cost: less than \$ 20,000

Figure 13. Basic PC-Quake system.

disk. With this size hard disk and if you are recording on 128 channels, you will be able to record 100 to 200 earthquakes before the disk is full. This PC is equipped with an A/D board and multiplexer to allow 128 channels of input data. We typically digitize at rates from 100 to 300 samples per second per channel. The off-line machine is connected to the data acquisition machine through a local area network. For this machine we have been using a 486 PC with 16 Mbytes of RAM and a 300 Mbyte hard disk. This PC is also equipped with a DSP board, a laser printer, and with a WORM drive for data archiving. The reason we use 16 Mbytes of RAM has to do with the maximum expected event duration. The longest event we recorded has a duration of about 300 seconds recorded on 128 channels. This created a file of about 7 megabytes in size. A large RAM disk helps speed up processing of large files like this, if the file is first copied to the RAM disk before processing begins and then accessed from there. This is particularly true if your hard disk is not very fast. The DSP board we use has an AT&T chip which is capable of 25 MFLOPS (million floating-point operations per second). For comparison, a 386 machine with a coprocessor can do about 0.5 MFLOPS, a 486 machine can do about 1 to 1.5 MFLOPS, and the best workstations currently available can do about 8 to 10 MFLOPS. So if you can write code which takes advantage of the DSP board, it is possible to get truly impressive performance out of a PC. For instance, the Symmetric Research DSP board can do a 1024 point FFT in about 15 milliseconds.

We use the following software in our system:

DOS	
IASPEI SOFTWARE	
KEDIT	(text editor)
PCTOOLS	(file & disk management)
XTREE	(file & disk management)

We have also learned the hard way that it is important to have backup power, so our machines are connected to an uninterruptible power supply. During the Loma Prieta earthquake, the system was just plugged into a normal wall outlet, which lost power shortly after the earthquake occurred. Fortunately, John Van Schaack quickly plugged the PC system into an emergency power supply. This allowed the system to record most of the aftershock sequence. The total cost of all this equipment and software (including 2 PC's, A/D boards, etc.) is less than \$20,000.

REFERENCES

- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. Bull. Seismol. Soc. Am., 68, 1521-1532.
- Banfill, R. (1992). SUDS: Seismic Unified Data System, version 1.31, Small System Support, Big Water, Utah.
- Data Translation, Inc. (1989). "User Manual for DT2821 Series High Speed Single-Board Analog and Digital I/O Systems", Data Translation, Inc. Marlboro, MA.
- Evans, J. and S. S. Allen (1983). A teleseismic-specific detection algorithm for single short-period traces. Bull. Seism. Soc. Am., 73, 1173-1186.
- Lee, W. H. K. (Editor) (1989). "Toolbox for seismic data acquisition, processing, and analysis", IASPEI Software Library, Volume 1, Seismological Society of America, El Cerrito.

- Lee, W. H. K., and Lahr, J. C. (1975). HYPO71 (revised): A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes. U.S. Geol. Surv. Open-File-Rep. 75-311, 116 pp.
- Lee, W. H. K. and S. W. Stewart (1981). "Principles and Applications of Microearthquake Networks", Academic Press, New York.
- Lee, W. H. K., and Valdes, C. M. (1989). User manual for HYPO71PC. IASPEI Software Library, Volume 1, Seismological Society of America, El Cerrito.
- Tottingham, D. M., Lee, W. H. K., and Rogers, J. A. (1989) User manual for MDETECT. IASPEI Software Library, Volume 1, Seismological Society of America, El Cerrito.
- Ward, P. L., (1989). SUDS:Seismic unified data system. U.S. Geological Survey Open-File-Report 89-188, 123 pp.

6. THE XDETECT PROGRAM

by

W. H. K. Lee

U. S. Geological Survey, Menlo Park, CA 94025

OLD SOFTWARE FOR REALTIME DATA ACQUISITION

Many people are still using a program which we wrote called MDETECT (Tottingham et al., 1989) (See Figure 1). MDETECT was our first attempt to write a realtime data acquisition program. It proved to be a popular program for small networks. However, we no longer support this program because the code is very difficult to maintain. We wrote XDETECT (Tottingham and Lee, 1989) because we needed a faster data acquisition program than MDETECT.

XDETECT was published in the first IASPEI Software Library Volume. However, it is now superseded by XDETECTV and XDETECTH (See

Old Software for Realtime Data Acquisition

- o Why not use MDETECT?
 - o PRO: Very popular for small seismic network
 - o CON: No longer supported
- o Why not use XDETECT?
 - o PRO: Published
 - o CON: Replaced by XDETECTV or XDETECTH

Figure 1. Old software.

New Software for Realtime Data Acquisition

- o Why use XDETECTV or XDETECTH?
 - o PRO: Currently supported version with SUDS
 - o CON: Poor online P-picking and location
- o Why use TDETECT?
 - o Excellent for teleseismic events
 - o Requires a high-speed PC
- o Why use STPARSE?
 - o Required for preparing STATION.000 file
 - o Input data must be accurate

Figure 2. New software.

Figure 2). These programs removed the requirement that a Hercules graphics card be used. If you want to use a VGA monitor, use XDETECTV. If you still want to use the Hercules card, use XDETECTH. Both of these programs produce their output files in the SUDS data format. Both also can pick P-wave arrival times and do a half-space hypocenter location. However, the quality of these processes still needs to be improved.

There is also a program called TDETECT (See Chapter 7). TDETECT is designed to record teleseisms. However, TDETECT takes about three times as long to process data as XDETECT, so you will need a relatively fast machine to use it. A 25-MHz 386 machine running XDETECT on 128 channels only uses about one third of its computing power, so it would run the TDETECT program successfully.

Both XDETECTV and XDETECTH need an input file which describes the station data. The input file is first created using an ASCII text editor and is then converted into a binary file called "station.@@@", using the program STPARSE. Unfortunately, STPARSE does not do much error checking, so you must be very accurate when you prepare the ASCII version of the input file. Another reason why it is important to prepare the station data file accurately is that any inaccurate information in the station data will be propagated through the entire analysis (since the SUDS data format includes all the station data with the earthquake data). This is one of the ways in which SUDS differs from old earthquake data formats. Instead of having separate earthquake and index files, everything is self-contained in a SUDS file. Overall, this is a strength since there is no possibility of mixing up index files or anything like that. However, any mistakes in the input file will be repeated in all the event files produced under that version of the input file. So be careful!

SETTING UP THE CONFIG.SYS

On most PCs with hard disks there is a file named "CONFIG.SYS." This file is read during initialization of the PC system, and information in it is used to customize the PC system for a particular set of applications. The current version of XDETECT requires that the "CONFIG.SYS" file contains the following (or similar) lines:

```
BREAK ON
SHELL C:\COMMAND.COM /E:1000 /P
DEVICE = C:\XDETECT2\ATLDRV.SYS /E 2048
FILES = 50
FCBS = 16,8
BUFFERS = 30
LASTDRIVE = N
```

The first line will cause DOS to check for the control-break key, so you have a better chance to stop a program which is running. There is always "control-alt-delete" if nothing else works! The second line specifies the command processor to use, the size of the environment each program will have, and specifies that the command processor stays loaded.

The third line is very critical. It tells the system to install the driver required by the Data Translation A/D board, and sets the amount of extended memory to use for the data buffers. The software absolutely will not work if there is an error here. There are different drivers for the 16-bit and 12-bit A/D boards, so make sure you use the correct one. Of course, the amount of extended memory you specify with the /E parameter must exist on your machine, and it must not be used by any TSR programs. Also, you must leave some extended memory free (at least 256 Kbytes) for the XDETECT program to use for pre-event buffers.

The FILES = 50 line sets the maximum number of file handles that DOS allows to

be open at once. During the processing XDETECT may need to have several files open at the same time and this line makes sure that no system errors are generated when XDETECT opens files. You will need the line FCBS = 16,8 if you network the data analysis computer to the data acquisition computer. BUFFERS = 30 sets the number of disk buffers to 30. This is primarily intended to speed up disk accesses.

LASTDRIVE = N should be used if you network the data acquisition computer. This will give you plenty of drive names available on your physical machine and the network "master" drive will, by default, become drive M. That should simplify access to the network drive, since you know it will always be drive M.

If you have problems with your system that you think may be caused by DOS, and if you are currently using MS-DOS, you may be able to correct the problems by switching to IBM-DOS. There are minor differences between the two operating systems, and since we have developed the XDETECT program entirely on IBM-DOS, it is less likely to cause problems than MS-DOS. So far all the incompatibilities we are aware of have been traced to the software provided by Data Translation. If you choose to use IBM DOS version 3.3, you can be sure that you will have no compatibility problems. However, that operating system only supports hard disks up 32 Megabytes. So, if you have a large hard disk you will either have to divide it into several logical drives, or else move up to DOS 4.01 or higher.

SETTING UP THE AUTOEXEC.BAT

Another file used to customize the PC during initialization is called

"AUTOEXEC.BAT." I will show a sample "AUTOEXEC.BAT" used on one of our NEC portables.

```
ECHO OFF
PATH C:\;C:\SYSCOM; C:\DOS;C:\ANALYSIS;
PROMPT $P$G
C:\DOS\SHARE
C:\DOS\TURBO +
C:\SYSCOM\NUMOFF
C:\SYSCOM\MOUSE
CD \XDETECT2
```

The first line just instructs the machine not to display the remaining command lines in the "AUTOEXEC.BAT" file. The next line sets the path. Every time you issue a command, DOS will look first in the current directory for the program name that matches the command. If it is not found there, the DOS will search every directory in the path until it either (1) finds the file and executes it, or (2) fails to find the file and displays an error message. By putting all your executable files in just a few directories and including those directories in the path, you can run any program from any directory.

The PROMPT \$P\$G line causes DOS to change its prompt so that it shows the disk and directory that are currently logged. This is much more informative than the default prompt.

The SHARE command is related to network use and causes the system to support file sharing. The next command relates only to the NEC machine. The system clock can be run at either low speed or high speed on the NEC. Unfortunately, the computer boots up in the

slow-speed mode. The TURBO + command switches the machine into the high-speed mode. The next command reflects my preference to have the numeric key pad function as an arrow key pad. It seems that most machines boot up with the NUM LOCK key on. This little program just turns the NUM LOCK key off. The program is not a part of DOS. It was provided by John Lahr's son, Nils.

If you are going to be running any of the analysis programs you will probably be using a mouse. In that case, you will need to load the mouse driver. That can be done at any time, but in this example the command is in the AUTOEXEC.BAT file. The last line just switches the current directory to XDETECT2. In general, your machine may be used for many different purposes so you would have several "AUTOEXEC.BAT" files and several "CONFIG.SYS" files. You can write a little batch file that will copy a particular "AUTOEXEC.BAT" and "CONFIG.SYS" to the root directory and reboot the system. This makes the process of re-configuring your PC system relatively painless.

PREPARING THE XDETECT.INP FILE

There are two files that XDETECT must have to run. They are the "STATION.@@@" file, and the "STATION.INP" file. Here I will show how to prepare an "XDETECT.INP" file.

Figure 3 shows the first section of the "XDETECT.INP" file. The control parameter keywords are to the left and are terminated with an "=". The setting you choose is to the right of the "=".

When an event is detected, you can have the machine beep or not beep. The first line

controls that behavior. Some people like to have the machine beep, so they can run over to the machine and look at the earthquake. But if you are in a crowded room, this can be very annoying, so we allow you to control this feature. Autotrigger can be switched on or off.

```
EventAlertBell= OFF;
Autotrigger= ON;
Autolocation= ON;
Autoreboot= Off, Time= 00:00:00;

ChannelGain= 1;
ClockSource= INTERNAL;
TriggerSource= INTERNAL;

ChannelBlocksize= 256;
DigitizationRate= 100.00;

PreEventTime= 1.0; #was 15
MinEventTime= 5.0; #was 60
MaxEventTime= 10.0; #was 180
```

Figure 3. Control parameters.

Normally you will leave this feature

switched on. You will normally leave the AutoLocation feature switched on, unless it is not important to you to get real-time hypocenter locations. We normally leave the Autoreboot feature off. If you switch this feature on, then when the PC clock shows the time set after the TIME= parameter, the computer will reboot itself. This can be used for "poor man's multitasking." In Alaska, we are running this system in a very remote village. Once a day the computer reboots, backs up the data files to a cartridge, and then starts XDETECT again. Then the village postman drops by, replaces the cartridge with a new one, and sends the old cartridge to John Lahr.

Channel gain can be set to 1, 2, 4, or 8. The reason that you might want to set the channel gain is to increase the seismic signal levels to match the A/D board. For instance, if the maximum output voltage of the discriminator is ± 2 volts, you should set the gain to be 4. Since the A/D board is set to digitize ± 10 volts at full scale, setting the gain to 4 would drive the board at 80% of its maximum digitization range. This will increase resolution by a factor of 4, without running the risk of over-driving the A/D board. If you are not using a

multiplexer, you can also adjust the gain on individual channels (as we will see soon).

ClockSource and TriggerSource have to be set to INTERNAL at this time, so you must leave them that way.

"ChannelBlocksize" and "DigitizationRate" are related parameters. ChannelBlocksize is the size of the data buffers which XDETECT will handle. If you increase the digitization rate you should increase the ChannelBlocksize, so the machine has a better chance of keeping up with the higher data digitization rate. ChannelBlocksize must be a power of two. The maximum allowable value for ChannelBlocksize is $32,768/nchan$, where nchan is a power of two, equal to or greater than the number of channels in the input station definitions.

Maximum allowable digitization rate is dependent on a number of factors including CPU speed, memory speed, bus speed, and hard disk speed. Right now the bottleneck is the hard disk. With the original AT running at 6 MHz. The best that machine can do is about 100 samples per second (Sps) on 32 channels. However, a 12-MHz AT can run 128 channels at a rate of 100 Sps, because in addition to the faster clock, the hard disk is faster. With a 25-MHz 386 PC we are able to keep up digitizing 128 channels at about 300 Sps. Faster than this, the hard disk will fall behind while an event is being captured.

The PreEventTime is the number of seconds of data to save prior to a trigger. In the example above, it is set to one second; but normally it should be 15 seconds or longer. The MinEventTime is the minimum length of time that data will be written to the disk, regardless of how long the event actually is. The MaxEventTime is the maximum length of time the system will keep recording an event. Normally the program will recognize the end of an event and stop recording. In the case that it doesn't, after MaxEventTime has elapsed, the

system will stop recording. This provides protection against filling the entire disk, because of some anomalous trigger condition.

Figure 4 shows the section of the input file that sets parameters to control how the trigger algorithm works. The first parameter is called `TriggerTimeLimit`. Essentially, this should be set to the time that it takes for a seismic wave to propagate across your network. Any triggers, which do not result in a confirmed event within

`TriggerTimeLimit`, are discarded by the XDETECT program. With a very dense network like the one we have at the USGS, we don't have to wait for a seismic wave to propagate across the entire network. In 6 seconds there might be 50 triggers, so we run numbers like 6 for this parameter. `CriticalNu` is the number of channels required for declaring a valid event. By setting this value relatively high, you can substantially reduce the number of false triggers. In this example file, the parameter is set to 1 since the program is being used for a demonstration session in the lab, and only one seismometer is connected. Normally, this number would be set from 5 to 10, depending on the size of your seismic network.

`STAverageWindow` (STA) and `LTAverageWindow` (LTA) are the number of points in the STA and LTA moving average windows used by the XDETECT trigger algorithm. Unless you have a good reason to change them, just use the default values shown here. `CriticalAlpha` is the threshold value for the ratio of the instantaneous value to the short-term

```
TriggerTimeLimit= 6;
CriticalNu= 1;

STAverageWindow= 16; # default
LTAverageWindow= 256; # default

CriticalAlpha= 7;
TriggerConfirmationCount= 35;

CriticalBeta= 4;
EventContinuationCount= 30;

CriticalGamma= 2;
CriticalMu= 2;
```

Figure 4. Trigger parameters.

average value (STA). If that ratio is greater than CriticalAlpha, then a trigger is declared. The smaller this number is, the more sensitive the trigger algorithm will be. After a successful trigger has occurred, CriticalBeta is compared to the ratio STA/LTA. If the ratio does not exceed CriticalBeta for at least TriggerconfirmationCount samples, then the trigger is discarded. This weeds out spikes and other types of transient signals.

EventContinuationCount, CriticalGamma, and CriticalMu are parameters used to determine when to stop recording an event. If at least CriticalMu channels have gamma values that exceed CriticalGamma for EventContinuationCount samples, then recording continues. Except for the value of CriticalNu, all the values given here are good choices. If you decide to modify them, you should change one parameter at a time. Otherwise, it may prove very difficult to see the effect of changing an individual parameter.

```
CriticalPhi= 4;
HalfspaceVelocity= 6.0;

AuthorityCode= 101;
NetworkName= USGS;
NetworkNodeId="MP, NEC";
PathName= "C:\xdetect2";

# Station Definitions
Ch= 0, StName= IRIG, Component= V, Trigger= OFF, Display=ON;
Ch= 1, StName= L22V, Component= V, Trigger= ON, Display=ON;
Ch= 2, StName= L22N, Component= N, Trigger= ON, Display=ON;
Ch= 3, StName= L22E, Component= E, Trigger= ON, Display=ON;
```

Figure 5. Network information & station definitions.

In Figure 5, CriticalPhi is the minimum number of stations that must have triggered in order to do a source location. This number must be at least 4. HalfspaceVelocity is an assumed P-wave velocity in kilometers per second. The number shown here is a good one to

use, unless you have a better estimate for the region you are working in.

The next three lines of Figure 5 are used to provide information which will be inserted in the SUDS file. The primary justification for this is that since the SUDS file could end up being examined by many different people not necessarily connected with your network, a standard way of identifying the network would help them keep their information straight. If you use an authority code from Peter Ward, then this number alone can be used to identify the data. The Network name also identifies the network; and in the case of an organization with more than one network, tells which of that organization's network collected a particular event. If more than one PC is collecting data from a given network, the NetworkNodeId will identify the particular PC used to collect the data. PathName simply tells XDETECT where to write the data files.

The station definitions provide the link between the station and the physical channel on which the station data is digitized and recorded. It is critical to get the channel numbers correct, or all subsequent analysis will be flawed. The StName parameter is the 4-character station identifier that will be displayed on the screen along with each trace. Component is the seismometer orientation. You can also choose whether to include it in trigger determinations, and whether to display a channel through the Trigger and Display parameter settings, respectively. In this example, only four stations are shown but in general, there will be as many station definitions as there are actual stations in your network (up to the number that your hardware can support).

PREPARING THE "STATION.@@@" FILE

Station.@@@ is the name of the binary file which contains the station information.

The XDETECT program looks for a file by this name when it starts, so there is no other choice on the name of this file. A two-step process is involved in preparing this file. First, you prepare an ASCII file with the station information, and then you use the program STPARSE to convert the ASCII file into binary format. Figure 6 shows part of a sample ASCII station file.

```
#0 IRIG 37.0700 121.8367 0
Network= USGS, StName= IRIG, Component= V, InstrumentType=1,
Azimuth= 0, Incidence= 0, StLatitude= 37.0700, Stlongitude= -121.8367,
Elevation= 0, StDelay= 0, Enclosure= ' ', Annotation= 0,
RecorderType= ' ', RockClass= ' ', RockType= 0, SiteCondition= ' ',
SensorType= d, SensorGain= 1, Polarity= n;

#1 L22V 36.1758N 121.6427W 1219
Network= USGS, StName= L22V, Component= V, InstrumentType=1,
Azimuth= 0, Incidence= 0, StLatitude= 36.1758, Stlongitude= -121.6427,
Elevation= 1219, StDelay= 0, Enclosure= ' ', Annotation= 0,
RecorderType= ' ', RockClass= ' ', RockType= 0, SiteCondition= ' ',
SensorType= d, SensorGain= 1, Polarity= n;
```

Figure 6. Sample ASCII station file.

When you modify either the ASCII station file or the "XDETECT.INP" file, you must make sure that the information, which is common to both files, matches exactly. Otherwise, the program may behave unpredictably when it encounters the inconsistency. Of course the cautions related to the "XDETECT.INP" file also apply here. If any of the information is wrong here, all analysis and all files based on events recorded while this file was in effect will be erroneous.

EXAMINING THE DAILY LOG FILE

Examining the Daily Log File

*** 04:46:24.200 ***

EVENT DETECTED Sat May 04 04:46:30.763 GMT

CH AVG

Short Long Time

3	E	137	30	Sat	May	04	04:46:30.763	GMT
21		244	30	Sat	May	04	04:46:30.174	GMT
26		778	161	Sat	May	04	04:46:29.615	GMT
27		159	36	Sat	May	04	04:46:28.946	GMT
29		328	170	Sat	May	04	04:46:29.475	GMT
33		679	295	Sat	May	04	04:46:28.856	GMT
37		561	60	Sat	May	04	04:46:29.924	GMT
42		143	30	Sat	May	04	04:46:30.693	GMT
46		166	30	Sat	May	04	04:46:29.655	GMT
49		405	125	Sat	May	04	04:46:28.786	GMT
50		151	30	Sat	May	04	04:46:30.863	GMT

*** 04:49:25.399 ***

DATE	ORIGIN	LAT	LONG	DEPTH	MAG	NO	RMS
May 04 1991	04:46:25.750	37.147 N	121.768 W	16.17	2.5	11	0.1

*** 04:49:25.679 ***

Event saved in: C:\tottingh\xdetect\v201\91050400.WVM

G - 10

Figure 7. Example of daily log file.

XDETECT generates two types of output files. During each session of XDETECT, a log file is generated (See Figure 7). Also, there will be an event file made for each

earthquake detected by the system. Whenever an event is detected, XDETECT makes a new entry in the current log file. This will consist of the date and time of the event and for each channel that triggered, the short- and long-term averages for that channel, and the time of trigger for that channel. Then, if a source location is calculated, information from that is also recorded. Finally, the fully qualified file specification of the event file is given. Figure 7 is an example of a daily log file.

REFERENCES

- Tottingham, D. M., and Lee, W. H. K. (1989) User manual for XDETECT. IASPEI Software Library, 1, 89-118.
- Tottingham, D. M., Lee, W. H. K., and Rogers, J. A. (1989) User manual for MDETECT. IASPEI Software Library, 1, 49-88.

7. THE TDETECT PROGRAM

by

John R. Evans
U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

TDETECT is XDETECT with a different trigger module in place of the local-earthquake trigger. TDETECT is optimized for teleseisms, hence the name. These events are very small and low frequency, so they tend to be emergent. Most local earthquake nets get only the very large teleseisms, typically one or two a week on the CUSP system, for example. Teleseisms are useful for global and local structural studies of many kinds, for checking station polarities, and for finding station statics for local earthquakes. Those which are useful in this way are much more common than would be indicated by the number of detections usually made of them. Also, volcanoes often produce low-frequency events probably associated with deep magma movement. These events are very similar to teleseisms and very likely will trigger TDETECT, although we do not yet have an example of that. Indeed, we don't know very much about those deep volcanic events because they are rather rare and the appropriate type of trigger has not been used. An analyst looking at volcanic records typically is not very alert to such candidate events because they are so rare compared to the thousands of local events occurring.

TDETECT can be set to be fairly sensitive to local earthquakes too, so it can be used in some cases for all of the network needs, both local and teleseismic. Eventually, there will

be a capability to run both XDETECT and TDETECT triggers in parallel so there won't be any need to compromise.

The strategy of the TDETECT algorithm is fairly simple. Teleseisms are characterized by low frequencies, typically one Hertz on a short-period network. Figure 1 is a voicegram (amplitude versus frequency and time) for a local or near-regional earthquake. You can see this is a rather broadband signal reaching about 10 Hertz. More distant regional events lose the high frequencies because of attenuation associated with the longer travel path. By about 3 degrees distance (300 km) from the network, in an attenuating environment like the western U.S., all of the

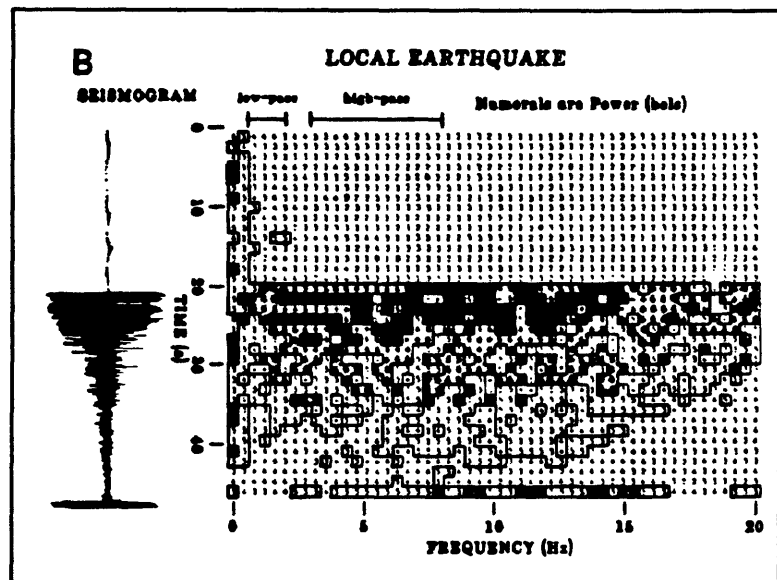


Figure 1. Voicegram of local event.

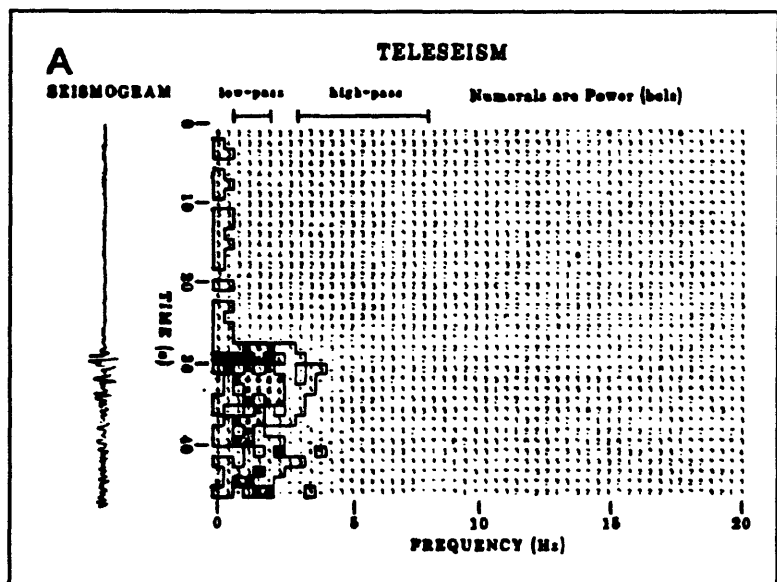


Figure 2. Voicegram for teleseism.

high frequencies are pretty well gone and you have something that looks fairly monochromatic near the corner frequency of the short period seismometer, as in Figure 2.

Even for a relatively high-frequency teleseism, the peak frequency only reaches two or three Hertz in the western U.S. There are a few cases, like New Zealand recording events from Indonesia, where one can get high frequencies 40 degrees (4000 km) away simply because it is a very efficient path, but generally there will be a complete cut off of the high frequencies a few hundred km out.

Figure 3 is a voicegram of a vehicle driving up to the station and stopping. It has the nice characteristic that most of the energy is high frequency.

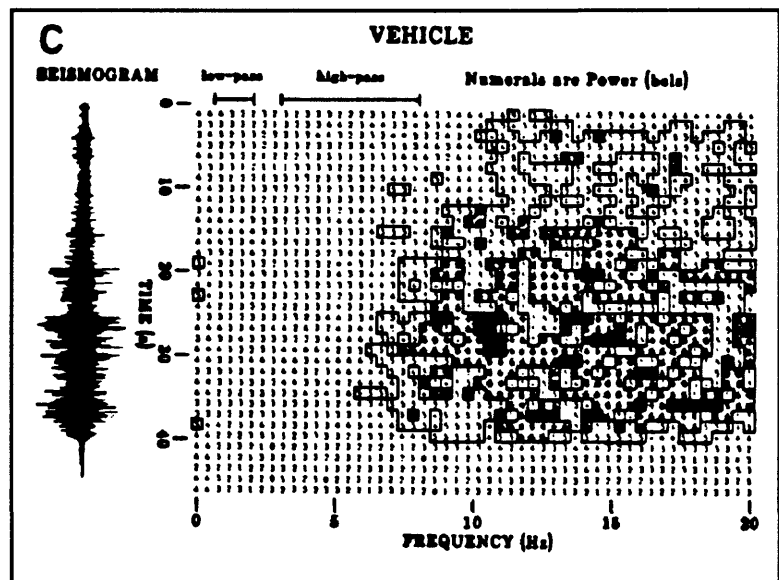


Figure 3. Voicegram for passing vehicle.

LOGIC OF TDETECT PROGRAM

Figure 4 is a flowchart detailing the logic of the TDETECT program. The basic strategy is to run two parallel triggers in two frequency bands, one down around the frequency associated with teleseisms (one Hertz) and one at a higher frequency. A teleseism

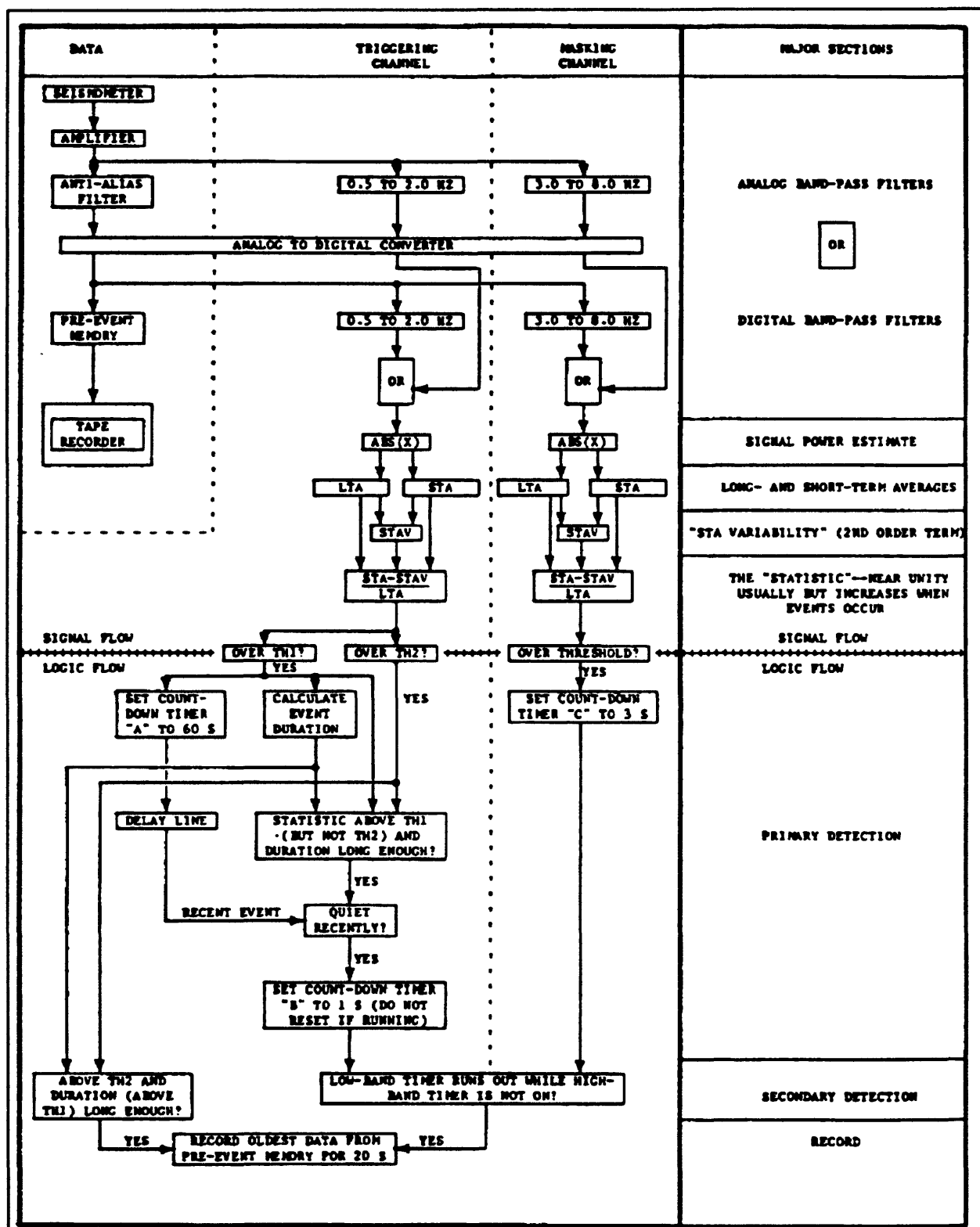


Figure 4. Flowchart of TDETECT operation.

is taken to be an event that has low frequency energy but no high frequency energy. So we want to see a trigger in the low frequency (triggering) band which is not associated with a simultaneous trigger in the high frequency (masking) band.

One divides the incoming signal into a triggering band and a masking band. In each band, typical statistics, like STA and LTA and a higher order term called the "short term average variability" (STAV) are computed and the ratio $(STA-STAV)/LTA$ is computed. In each band the computed ratio is compared to a threshold to trigger within that band. A high frequency "masking" trigger simply turns a timer on for the next few seconds. The idea is that the low frequency trigger may occur somewhat later than the high frequency trigger.

The low frequency triggering threshold is set very low. Typical values may be as low as 2 dB in a network. However, if one has only a single station, the threshold would be higher than this (typically 6 dB). The algorithm pauses for about one second after a low-frequency trigger to allow any masking event that is lagging to come in and mask. This pause also gives time for the machine to evaluate the duration of the low-frequency signal. To throw out certain kinds of spikes, there is a minimum duration in the trigger. This is a very simple requirement that the low-frequency ratio has to remain above its threshold for some minimum amount of time before you accept it as a trigger. When that pause (the "post event quiet period") runs out, typically a second after the initial trigger, the logic simply checks to see if there was any activity in the masking band, and if there was activity it does not record that event. This expedient masks out most kinds of vehicles that happen to come near the sensor. In fact vehicles are rejected with very high reliability by this trigger

algorithm. It also rejects local earthquakes, which are broad band and have simultaneous activity in both bands. Discarding most local earthquakes is necessary because to get teleseisms at adequate sensitivity one must set the threshold level very low. You would be overwhelmed by local earthquakes if the area were at all active.

However, three criteria suggest that we need one other thing going on in the triggering channel--a second, higher, threshold that will accept an event no matter what is going on in the masking channel. First, you don't want to be embarrassed by not recording a magnitude 6 earthquake just because you happen to be running TDETECT. Second, if a teleseism is large enough that clipping occurs, spurious high frequencies will be generated and the event will mask itself. Third, a very large teleseismic signal may simply have enough leakage past the filters that it sets off the masking band. For these reasons, one needs to implement a secondary threshold test. (It is this secondary threshold that allows one to throttle the rate of local-earthquake capture semi-independently of teleseismic triggering.)

Figure 5 shows what is going on in the masking channel and what is going on in the triggering channel. This is an example of a slightly better than average teleseism. Comparison of the low-

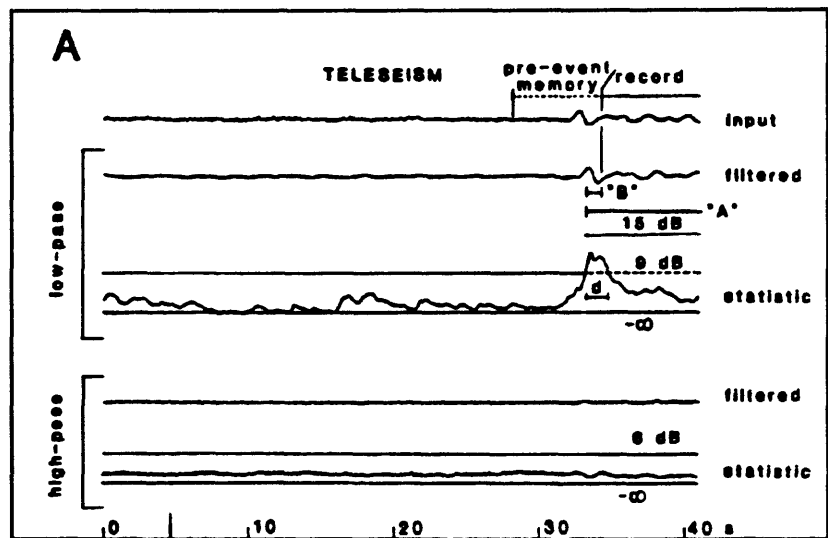


Figure 5. Detection algorithm operating on teleseism.

pass (triggering) filtered trace and the high-pass (masking) filtered trace shows that there is some leakage into the masking channel simply because the filters are imperfect. The third trace down is the (STA-STAV)/LTA ratio for the triggering band. You can see it has the characteristic exponential shapes of the averaging filters that we use. The main threshold level is indicated on the trace also (9 dB in this single-station example). It takes some time for the ratio to climb to the threshold level. The time during which the ratio is above the (9dB) threshold is called the duration. If that number is greater than 0.75 seconds then the duration criterion is satisfied, and the threshold criterion is satisfied, and at the end of one second the algorithm checks the equivalent high frequency statistic. As you can see, the masking statistic isn't doing much, so the algorithm decides that the event is a valid teleseism and saves it.

There is also a cautionary window ("A") that starts up, which for the most part you should ignore because it doesn't work too well. The idea was that sometimes microseisms, oceanic wave noise, large reservoir wave noise, or large waterworks noise, can produce very monochromatic low-frequency signals, and these signals, because of multipathing, can have a repeating, pulse-like characteristic which looks very much like a series of teleseisms. To compensate, one can turn on this cautionary window, which says that if there has been activity in the triggering band recently then don't take anything but big events (in which the ratio exceeds the upper threshold).

Figure 6 shows the signal from a sonic boom. It has characteristics which are very similar to those of a local earthquake, as far as the trigger algorithm is concerned. It has lots of low frequency energy, hence one gets a response of good duration in the triggering

channel. However, the check of the masking band shows that there is high energy there too, so the event is ignored.

If, however, the ratio in the triggering band had touched the upper threshold, even for one sample, then it would accept the event no

matter what was going on in the masking channel, contingent only on a duration criterion in the triggering band. So large local earthquakes, clipping teleseisms, and so on, can be collected by this "big event" type of trigger. You can throttle the number of local earthquakes that are recorded somewhat independently of teleseisms. Since most teleseisms are triggered on the lower threshold and all local earthquakes are triggered on the upper threshold, you can increase the number of local earthquakes recorded by lowering the upper threshold until you have the desired sensitivity. You shouldn't put the upper threshold more than about 6dB above the lower threshold, because you will start missing medium size teleseisms that clip or leak into the masking band. But you can make it less than 6dB above to gather more local earthquakes.

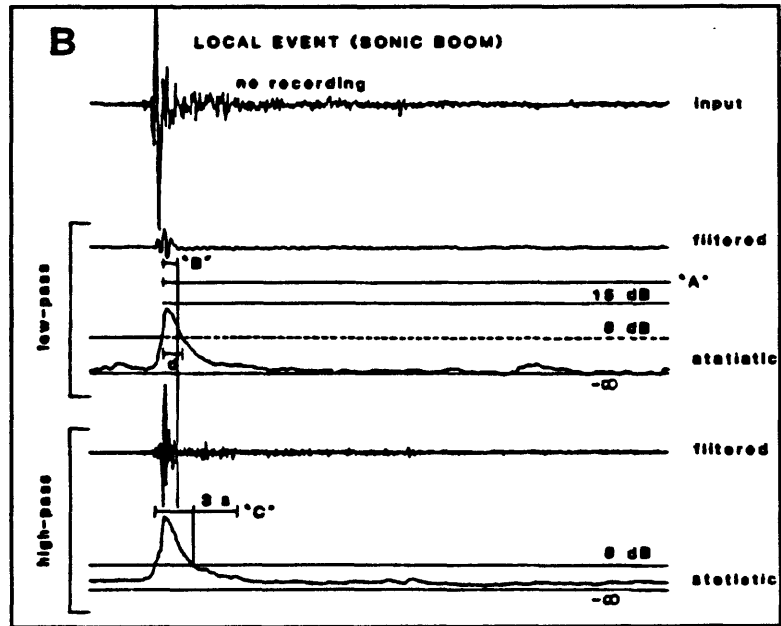


Figure 6. Sonic boom and time series.

ADJUSTING TDETECT FOR BEST PERFORMANCE

Table I is an annotated copy of an input file for TDETECT, with coefficients that have been found to be most effective in capturing teleseisms. The only real difference you will see between the operation of TDETECT and of XDETECT is that the .INP file has a few different settings. In this table, the lines indicated by arrows are the ones that you will most likely have to adjust as you tune your network trigger. Most of the rest of the parameters have been pretty well optimized in lab tests and subsequent field testing so you shouldn't have to change them. Note that the trigger is exquisitely sensitive to the main trigger threshold, Tth1, and that Tth1 admits of fractional values (0.2 dB can be a world of difference). There are something like 18 parameters for this trigger and they are not entirely independent or linear, so it is mostly experience that lets you set them correctly.

The algorithm only works at 100 samples per second at the moment. That is only due to not having extra banks of filter coefficients and a little bit of programming. But 100 sps is perfectly adequate for most situations.

Typically any single channel will take about three seconds after the first arrival to trigger (because of the emergence of the event, the delays for the post-event quiet period, and the time it takes for the statistic to rise above the threshold). So set the PreEventTime parameter to 3 seconds plus the propagation time across your array plus a margin for error. The trigger becomes sensitive to events about three degrees away, given typical values of Q, and you can assume an apparent velocity of about 8 km/s for distant regionals. So if you divide the array aperture by 8 you will get an appropriate transit time. If you are only interested in true teleseisms, from about 20 degrees or further out, the appropriate apparent

Table I TDETECT input file with suggested values of coefficients.

```

#                               TDETECT Parameters for ".INF" File
#
# Operation settings.

EventAlertBell= ON;
Autotrigger= ON;
Autoreboot= ON, Time= 00:00:00;

# ATLAS settings ##### (unit) Comment and suggestions.
ChannelGain= 8;           # 8 is maximum. Set high--teleseisms are very weak.

ChannelBlocksize= 128;    # (samples) 256 OK.
DigitizationRate= 100.00; # (samples per second) Fixed--no option.

...

# Event settings ##### (unit) Description and suggestions.
-> PreEventTime= 15;       # (s) Set > 3 s + (array aperture) / (8 or 10 km/s).
-> MinEventTime= 30;       # (s) Record length of "normal" event.
-> MaxEventTime= 600;      # (s) Record length of "big" event.

-> TriggerTimeLimit= 10;   # (s) Set to (array aperture) / (8 or 10 km/s).
-> CriticalMu= 10;         # (n) number of triggers within TriggerTimeLimit.
-> CriticalNu= 10;         # ... " ... Set equal to CriticalMu.

# Teleseismic trigger settings.
# Timers
Aset= 0;                  # (s) Microseism caution window (a period after any
#                          trigger-band activity during which only "big"
#                          events recorded). Not very effective.
Bset= 1;                  # (s) Pre-event quiet period (in masking band).
-> Cset= 10;               # (s) Pre-event + post-event quiet period.
Eset= 3;                  # (s) Delay onset of microseism caution window.
Sset= 200;                # (s) Start-up trigger delay to let averages settle.

# Masking (high-pass) band. Statistic is (MSTA-MSTAV)/(MLTA-Mminen).
Cmlta= 10;                # (s) LTA 1/e time
Cmsta= 1;                 # (s) STA 1/e time
Cmstavg= 100;             # (s) |STA-LTA|, "STAV", 1/e time
Mminen= 1;                # (counts) "Minimum energy" in masking band.
-> Mth= 5;                 # (dB) MASKING BAND THRESHOLD. Generally >= Tth1

# Triggering (low-pass) band. Statistic is (TSTA-TSTAV)/(TLTA-Tminen).
Ctltav= 10;               # (s) LTA 1/e time.
Ctsta= 1;                 # (s) STA 1/e time.
Ctstavg= 100;             # (s) |STA-LTA|, "STAV", 1/e time.
Tminen= 1;                # (counts) "Minimum energy" in triggering band.

-> Tth1= 2;                # (dB) MAIN TRIGGER THRESHOLD.
Mndur1= 0.75;             # (s) How long the statistic exceeds Tth1.

-> Tth2= 9;                # (dB) BIG-EVENT THRESHOLD. Generally <= Tth1 + 6.
Mndur2= 1.40;            # (s) How long the statistic exceeds Tth1, NOT Tth2.

DCcoef= 10;              # (s) 1/s time of low-cut filter in triggering band.

# Station definitions.
-> (At least 16 triggering stations recommended; maximum is about 48 due to
computational limits.)
-----
("->" means "likely to be changed while tuning PC".)

```

velocity is more like 10 km/s.

The notion of the "normal" event (Tth1) and the "big" event (Tth2) can be used to determine how long the record will be. There is no specific event cutoff criterion in this algorithm. Instead, a "normal" event will always record for the MinEventTime and a "big" event will always record for the MaxEventTime. The TriggerTimeLimit should also be set large enough for the aperture of whichever part of the array you are triggering on.

TDETECT is a very numerically intensive trigger algorithm because of the filters and because you are running two of them for every station. So typically you will run a subset of about 16 of your stations with this algorithm.

The next parameter is called CriticalMu. This and Tth1 will be the numbers you change most often when you are tuning the algorithm. Numbers like 10 (out of 16) seem to work for Calnet. That is partly because we set the main threshold (Tth1) very low. Bset is the number of seconds after an initial trigger in the triggering band during which you must see silence in the masking band (post-event quiet period). Cset equals Bset plus the number of seconds before the trigger that you want the masking band to be silent (pre-event quiet period). For a single station, Cset would typically be set about 3 s. In the Calnet, we found by trial and error that 10 s worked better. Because the algorithm computes very long averages, you want the averages to settle without triggering for awhile after starting TDETECT. The Sset parameter tells the algorithm how many seconds long the settling period should be and it should be about twice the longest averaging period.

The Cmlta, Cmsta, and Cmstav parameters are the $1/\epsilon$ times for the STA, LTA, and STAV windows in the masking band. Ctlta, Ctsta, and Ctstav are equivalents for the

triggering band. There is a minimum energy requirement, in digitizer counts, in both bands. The signals must be above the Mminen or Tminen level before the event is looked at in the masking or triggering band, respectively. Mth is the threshold level in the masking band and Tth1 is the main threshold in the triggering band. Mndur1 is the "normal" event minimum duration. In the triggering band there are also the parameters Tth2, which is the secondary threshold level, and Mndur2, which is the time that the statistic must exceed the main threshold level (Tth1) for a "big" event trigger. Dcoef is a parameter which adjusts the high-pass filter used to remove any DC offset from the signals prior to processing.

Following the parameter list is a list of stations with the information as to which are triggering and which are not, exactly as in XDETECT. For a small network of 16 stations having 10 - 16 stations triggering typically will work. We know that a 25 MHz 386 will handle 48 stations triggering at once and that it will not handle 64. (We now know it also fails eventually for 56.) That will cover just about any net you could consider running, except for us since we have 500 stations running.

We have modified the system so that the log file will come out on an RS-232 stream which can be used to trigger some other system. Our immediate application is to trigger the CUSP system more sensitively for teleseisms than it currently does with its own algorithm. We use a PC that monitors a subnet of CUSP and sends information to CUSP whenever it sees interesting events. CUSP can then evaluate this information and decide whether to keep an event and how much of it to keep. This arrangement is applicable to any network that is running any kind of system. That log file will have a "b" for big events and a "n" for normal events so one can evaluate the event's size and make a determination of recording

duration on that basis. Network glitches and microseisms, for a coastal subarray, are significant problems with Calnet. They generate 3/4 of all triggers. The remainder are valid events, reliably reaching a threshold near $m_b = 5.0$ for teleseismic P. Hence, some manual editing is required, though automatic tests should be relatively easy to develop and should not delete almost any real teleseisms. A 6-month statistical base has been gathered and will be used to design such tests.

SUGGESTED READING

- Allen, R.V. (1978) Automatic earthquake recognition and timing from single traces. Bull. Seism. Soc. Am. **68**, p 1521-1532.
- Allen, S.S., and R.C. Brodine (1980) Development of a teleseism detection algorithm and implementation on a microprocessor based system. Final Report, U.S. Geological Survey Contract no. 14-08-0001-18369
- Evans, J.R. (1981) A teleseism-specific digital detection algorithm for isolated, triggered seismographs. EOS, Trans. Am. Geophys. Union **62**, p 968.
- Evans, J.R., and S.S. Allen (1983) A teleseism-specific detection algorithm for single short-period traces. Bull. Seism. Soc. Am. **73**, p 1173-1186.

8. ROUTINE SEISMIC NETWORK DATA PROCESSING

by

W. H. K. Lee

U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

As pointed out by Jerry Eaton, it is much easier to collect data than it is to process the data. So, one of the basic problems faced by any seismic network is developing an integrated set of procedures to streamline the data processing after data has been collected. In this chapter, I will outline the types of data you can collect and store and the necessary procedures to process that data. Then I will briefly discuss the various IASPEI programs that have been developed to facilitate network data processing and analysis. Many of these programs will be discussed much more thoroughly by other authors later on in this volume.

NETWORK DATA

Seismic network data (See Figure 1) can be grouped into the categories of:

- Instrument data
- Raw Waveform Data
- Phase data
- Earthquake event list
- Reports

The first item is where most people get into trouble: the instrument data. Basically, it is easy to put some instruments out in the field, but most people are not very careful to write down where their sensors are and all of their characteristics. Often times people keep changing seismometers, VCOs, and things like that. After a few months the information gets

all confused. Therefore, it is very important to get all this information recorded before you collect the first byte of data, or else you will regret it from then on.

The volume of the raw waveform data collected is mind boggling. If you digitize a station trace at 100 samples per second, two bytes per sample, and 3×10^7 seconds in a year, you end up with something like

6×10^9 bytes per year. That is 6 gigabytes just for the raw waveform data of one station trace. A big book is about a megabyte, so you are talking about 6,000 big books per year. At a megabyte per floppy disk, you are talking about 6,000 floppy disks for just one station trace. As a result, very few people really collect raw waveform data continuously.

Another problem is that even if you collect all the raw waveform data, it is not easy to move it around. Even writing from a hard disk to another hard disk, the data rate is something like 50 kilobytes per second. In short, it is very hard to rapidly move more than a few tens of megabytes between any two devices. So most seismologists just try to collect those few wiggles that represent earthquakes. That can be very dangerous because you can throw out a lot of data. One classic example is the hole in the ozone layer. That hole was not discovered by NASA with all the satellites and modern instruments. It was discovered by a

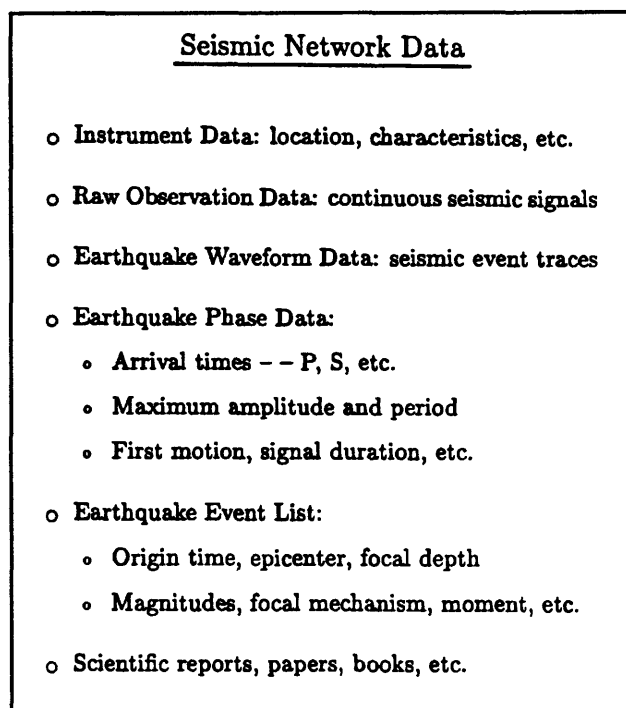


Figure 1. Network data.

few British scientists working in Antarctica using very primitive equipment. The reason NASA didn't discover the hole was that their automatic data collection algorithm was designed to reject "bad data." It turned out that some of the "bad data" was indicating the presence of the hole. So I caution you that even though we only try to save earthquake data, the practice can be very dangerous.

A good thing about the earthquake waveform data is that usually it is no more than a percent or two of the continuous stream of raw data. In other words, the data which must be saved is reduced to something like 6×10^7 bytes per year per station trace. That is something like 60 megabytes, which is not too bad. But if you have 100 station traces, you are back to where you were in the first place. So you can see that in the network we have running here in Menlo Park, we have quite a problem. Since we have more than 500 channels of seismic data coming in, handling all this data is a big problem.

The next step in compressing these waveform data is to just save the phase data. The phase data consist of things like P and S arrival times, maximum amplitude, period, etc. What that means is that each earthquake requires of the order of ten to 20 numbers per station trace to be saved. That uses about 40 bytes, a lot less than a typical seismogram of an event for a station trace which requires 10,000 bytes or more. Consequently, this is what most people have been doing. They just save the phase data and throw the rest away. From the phase data you get the earthquake event list containing origin time, epicenter, focal depth, magnitude, and focal mechanism, etc. If you want to compute seismic moment, then you have to go back to the waveform data again. The final data compression comes when we publish our papers and reports.

SEISMIC NETWORK DATA FLOW

Here I will discuss the flow of data through a PC-based seismic system using the IASPEI software (See Figure 2). The programs which do the processing and analysis are, for the most part, command line oriented and interactive. This allows processing to be done in batch mode while still retaining the ability to intervene in the process.

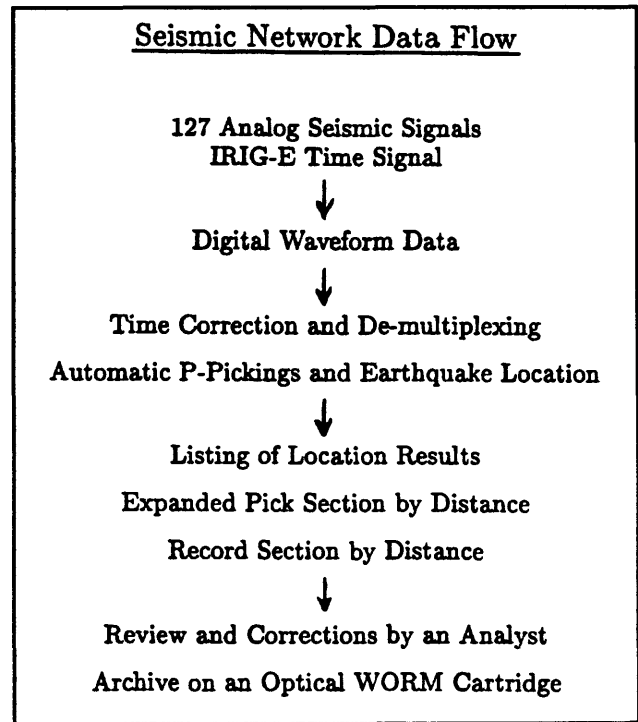


Figure 2. Network data flow.

First we collect 127 channels of seismic data with the IRIG-E time code recorded on one channel. From this, a digital waveform file is created for each seismic event which has been triggered. This file is time corrected and demultiplexed. Then we do the automatic P-picking and earthquake location, and get back an expanded pick section by distance and a record section by distance. Then an analyst has to review this information for errors and make the necessary corrections. Finally the relevant data are archived on a WORM cartridge.

Probably the single most time consuming step in this procedure is manual review of the automatic picks. For this reason, we have put a considerable effort into making the picker program very accurate and also easy to interact with. Figure 3 shows an example of the automatic picks performed by the system.

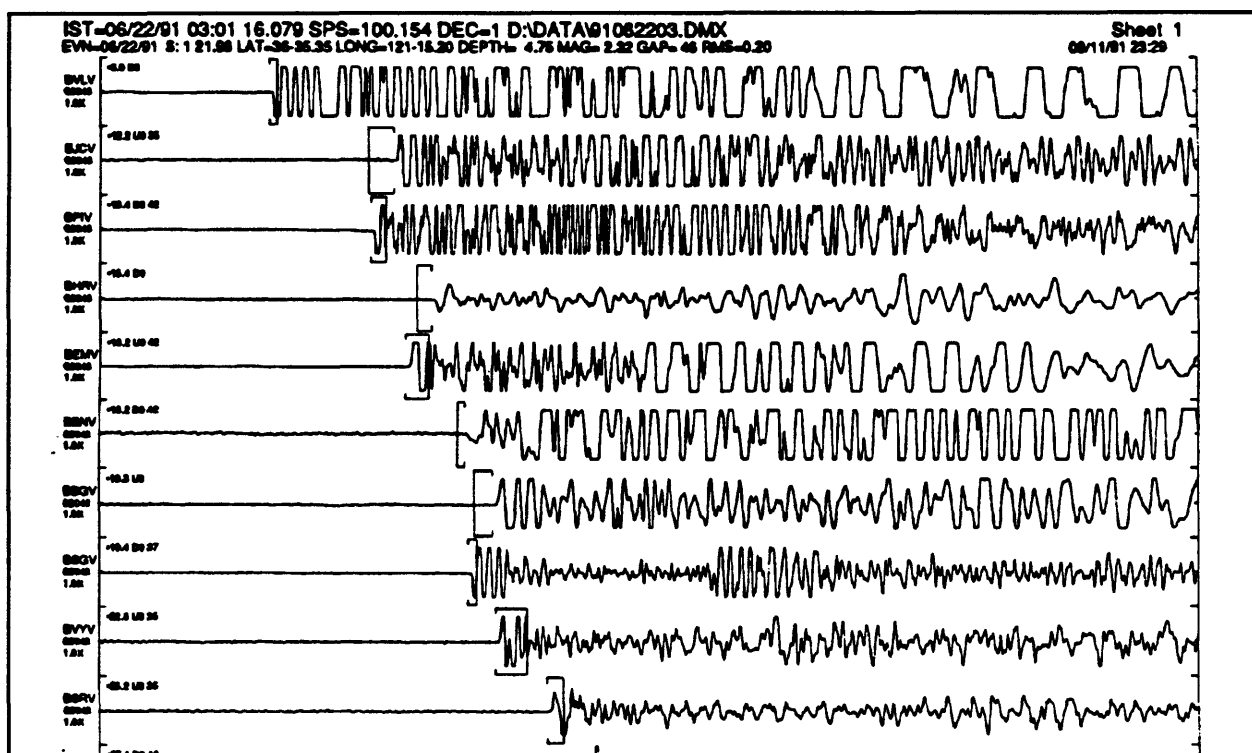


Figure 3. Example of P-wave picks made by SUDSPICK.

As you can see, on the fifth trace the picker appears to have picked the time slightly early.

The solid bar represents the computed arrival time given by the computed hypocenter, and the open bar is the time picked by the automatic picker. None of the stations shown on this section have had any station corrections applied to them. Therefore, if station corrections were applied, the difference between the computed time and the picked time would be less.

Overall, the picks are really quite good, so the analyst would not have to spend much time correcting this set of data. By doing this automatic picking, we hope to reduce the analyst's time by a factor of ten or so. Because in the past he had to look at every trace. Now the analyst needs to look at only one out of ten or so.

IASPEI SOFTWARE

Increases in efficiency for analyzing seismic data are being brought about by a combination of improvements in hardware and in software. In order to aid in the development of efficient and useful software, I feel that the software must be shared. For this reason, the software developed for the PC seismic system has all been published.

<u>IASPEI Software Library List</u>	
o	Vol. 1: Toolbox for Seismic Data Acquisition, Processing, and Analysis (<i>May, 1989</i>)
o	Vol. 2: Toolbox for Plotting and Displaying Seismic and Other Data (<i>September, 1990</i>)
o	Vol. 3: Digital Seismogram Analysis and Waveform Inversion (<i>February, 1991</i>)
o	Vol. 4: Bibliographic References and BSSA Database (<i>December, 1991</i>)
o	Vol. 5: Programmable Interactive Toolbox for Seismic Analysis (<i>April, 1992</i>)
o	Vol. 6: Standard Processing for Earthquake Strong-Motion Data (<i>December, 1992</i>)

Figure 4. IASPEI software library.

Right now, there are three IASPEI volumes that have been published (See Figure 4). The first volume is for seismic data acquisition, processing and analysis (Lee, 1989). Volume 2 is a toolbox for plotting and displaying data (Lee, 1990). Volume 3 is about digital seismogram analysis and waveform inversion (Lee, 1991).

Volume 4, which has not been published, deals with the literature and the BSSA database. The Bulletin of the Seismological Society of America (BSSA) is in its 81st year. The first 80 years of data has now been entered into a database so that you can easily see what papers have been published in BSSA. One big problem these days is that there are just too many books and articles and we don't find time to look through all of them. As a result, we often reinvent the wheel. What we are trying to do in Volume 4 is make it easy for you

to find out what has been done, so that you won't ignore other people's work. For example, when I was a student 30 years or so ago, they were still publishing the collected works of Euler. In fact, they were only about half way through his stuff. Quite often, when someone invents a mathematical technique today, it turns out that Euler, or Gauss, or one of their colleagues, thought about it a long time ago or worse, had already published the idea. Volume 5, which is a programmable interactive toolbox for seismic analysis, is scheduled for publication next year. Volume 6, which will be called "Standard Processing for Earthquake Strong-Motion data", is scheduled for publication at the end of next year.

<u>IASPEI Software Library Volume 1</u>
Purpose: Simple, inexpensive, do-it-yourself, seismic data acquisition, processing, & analysis.
Design Goals:
<ul style="list-style-type: none"> o Reduce analyst's time by a factor of 3. o Reduce system costs by a factor of 10. o Reduce data retrieval time by a factor of 30.
System Characteristics:
<ul style="list-style-type: none"> o Digitize 16 signals at 100+ samples/sec/channel. o Digitized data are displayed in real time. o Automatic P-pick and event location. o Digitize up to 128 channels with a multiplexer.
Programs: Mdetect, Xdetect, Xplay, Hypo71pc, Pceq, and Qcoda.
New Programs: Tdetect, FixTime, Demux, Extract, Export, St2Asc, and SudsPick.

Figure 5. Volume One.

AVAILABLE SOFTWARE

Volume 1 contains the programs MDETECT, XDETECT, XPLAY, HYPO71PC, PCEQ and QCODA (See Figure 5). We are in the process of putting out a new edition of volume 1 with updated versions of these programs. Volume 2 is mostly concerned with plotting and displaying data (See Figure 6). This turns out to be easier said than done. What

we are trying to do is provide an integrated data plotting environment. We also have programs which provide 3-D graphic display of seismic data. We have put a considerable effort into making the programs easy to use. As you know, if a program is easy to use, then it is hard to write.

The programs in these volumes support many earthquake data formats, although we are trying to standardize on SUDS (Banfill, 1992; Ward, 1989).

The programs also support several types of graphics hardware. As you know, one of the biggest problems on the PC is that there are so many different kinds of hardware for the PC that it is difficult to have software to support all of them. The IASPEI programs usually support only the commonly used dot-matrix and/or laser printers.

As an example of the program capability, here is a plot produced by

IASPEI Software Library Volume 2

Purpose: Plotting and displaying seismic data.

Design Goals:

- Integrated data plotting environment.
- Powerful 3D graphic display of seismic data.
- To aid automatic network data processing.

Program Characteristics:

- Easy to use.
- Support many earthquake data formats.
- Support several PC graphics hardware.
- Support dot-matrix and laser printers.

Programs: PixPlot, PenPlot, GoPlot, DoPlot, AcroSpin, Spin*, and RecSec.

New Programs: SudsPlot and QMap.

Figure 6. Volume Two.

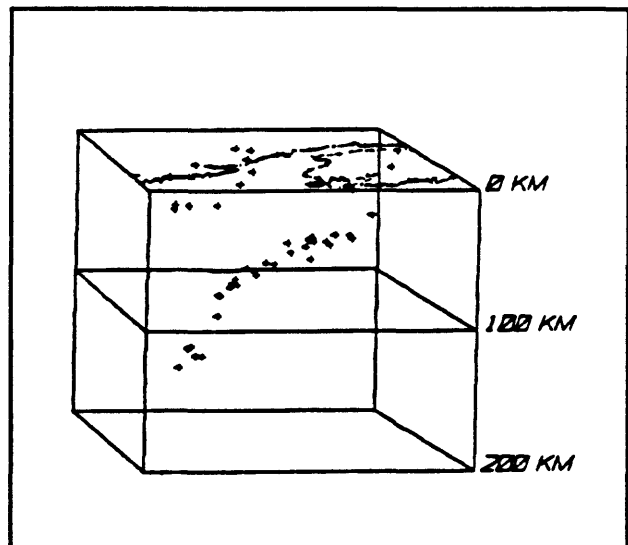


Figure 7. Sample ACROSPIN output.

the program ACROSPIN from Volume 2 (See Figure 7). The program plots earthquake hypocenters on a projection of a 3-D solid, and allows you to rotate the viewing position in 3 dimensions.

Volume 3 provides an interactive data analysis environment for teleseismic events (See Figure 8). There are two major programs in this volume. They are SEISGRAM and SYN4. The SEISGRAM program is described in Chapter 17. SYN4 is used for inversion of teleseismic body waves. Essentially you have recorded some earthquake waveform data, and you are trying to figure out what are the most likely source parameters for that data. You try to match the theoretical waveform generated by a double-couple source model with the observed data. We won't have time to

IASPEI Software Library Volume 3

Purpose: Interactive digital seismogram analysis and inversion of teleseismic body waves.

Design Goals:

- Interactive data analysis environment.
- Powerful graphic display of seismic data.
- To aid seismic network data analysis.

Program Characteristics:

- Easy to use.
- Support many earthquake data formats.
- Support several PC graphics hardware.
- Support dot-matrix and laser printers.

Programs: SeisGram and SYN4.

Figure 8. Volume Three.

talk about body wave inversion so you will just have to read it yourself. Also, Volume 3, which is under revision, will have a program added to do seismic moment tensor inversion.

FUTURE IASPEI VOLUMES

Right now we have three volumes in preparation and three volumes in planning. Planned volumes are: volume 7 will contain seismic ray tracing and synthetic seismograms.

Volume 8 will contain ISOP programs. As you probably know, there is a program called International Seismological Observational Period (ISOP). The idea is that for a period of two years, everyone around the world will

observe earthquakes systematically. This is

very much like the International

Geophysical Year in 1957 that launched

the space satellite program. To facilitate

systematic observations, the necessary

programs will be combined in this one

volume which hopefully will be widely

distributed. We are also thinking of

producing a volume dedicated to

earthquake location and focal mechanism determination, although this could be split into two

volumes, because of the amount of information that needs to be presented. What I am trying

to do is mobilize the international community of seismologists to put a serious effort into

writing programs that are mathematically sound, easy to use, and with a reasonable user

interface. We want the user interface of these programs to be reasonably similar to each

other, so that their use becomes somewhat intuitive.

Software for Processing SUDS Waveform Files

- **Why use FIXTIME?**
 - Required for making time correction
- **Why use XPLAY?**
 - Provides a quick look at the waveform data
- **Why use DEMUX?**
 - Data must be de-multiplexed for processing
- **Why use EXTRACT?**
 - Let you extract selected seismic traces
- **Why use ST2ASC and STDESCR?**
 - Let you examine a SUDS file easily

Figure 9. SUDS file software.

SOFTWARE FOR PROCESSING WAVEFORM FILES

The first of these programs (See Figure 9) is FIXTIME written by Dean Tottingham.

This program is required to make the time corrections in the SUDS files (See Figure 10).

The problem is that the clock in the PC is pretty poor. It drifts about one or more seconds per day. That is not acceptable in seismology. To have accurate timing, we also digitize the IRIG-E time code. The reason for using IRIG-E over any other time codes is that IRIG-E has ten pulses per second, so we can quickly decode what the time is. FIXTIME decodes the digitized IRIG-E codes, and uses them to correct the clock times and the digitization rates recorded in the SUDS file. There is no manual intervention required.

Usage of FIXTIME is:

fixtime IRIG-channel-number filename

IRIG channel number is the channel number on which you record the IRIG-E time code, and filename is the name of the SUDS file to be processed.

The next program which may be used on SUDS files is XPLAY (See Figure 11). This program was also written by Dean Tottingham. It allows you to take a quick look at the data, and make sure that it is not grossly wrong. Usage is

xplayv [filespec]

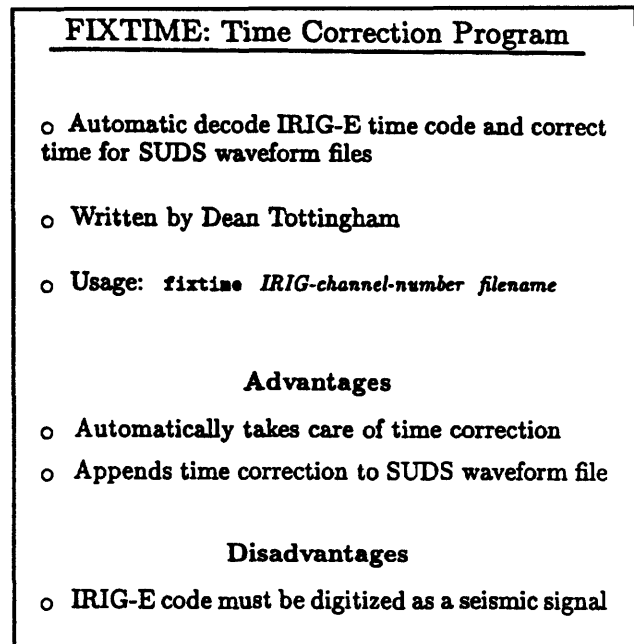


Figure 10. FIXTIME program.

(If you are using a Hercules graphics card then you must use an alternative program called xplayh). The original program published in IASPEI Volume one does not support the SUDS format, but the later versions do. XPLAY has on line help, which may be accessed by pressing the F1 key. The help consists of a summary of the possible valid keystrokes and the effects of those keystrokes. XPLAYV supports EGA and VGA graphics adapters, and also supports IBM-compatible dot

matrix printers for hard copy output. It only displays 16 traces at a time, but you can page down through the traces if you have recorded more than 16 channels. Currently, XPLAYV does not support high resolution graphics modes, but a new version, which is currently being tested, supports resolutions up to 1024 X 1068 pixels.

The DEMUX program, written by Dean Tottingham, is used to demultiplex the SUDS waveform files (See Figure 12). As I explained earlier, the data acquisition program digitizes the data on a channel by channel basis. That is, the first sample is from channel zero, the second from channel one, the third from channel two, and so on. Only after all the channels have been sampled is the next sample taken from channel zero. However, when you

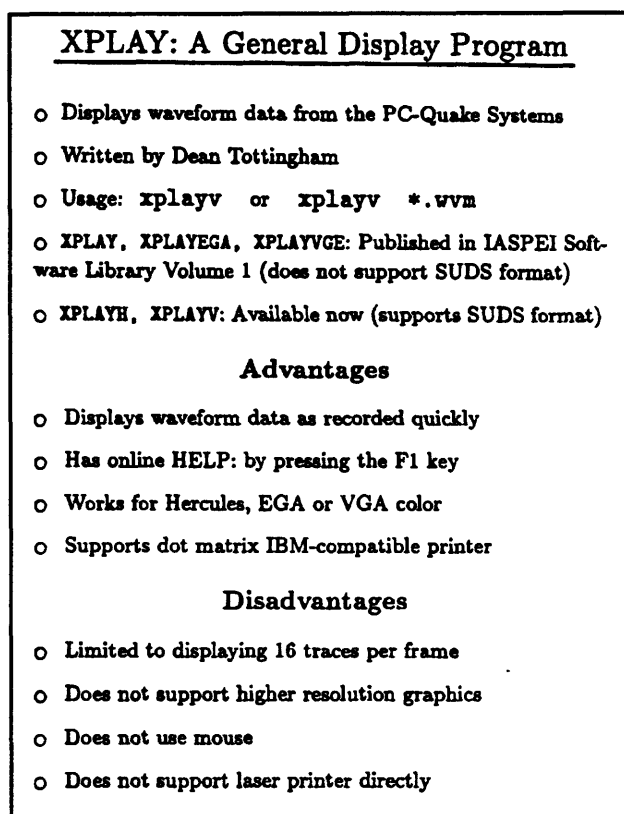


Figure 11. XPLAY program.

process the data, you would prefer to have all the data grouped by channel. The DEMUX program performs this transformation to the SUDS files.

Usage of DEMUX is:

demux input-filename output-filename

DEMUX does not destroy the input file so if you do not want to keep this, essentially redundant, information on disk you will have to erase the input file.

By convention we have been naming the waveform files with the extension .WVM, where WV stands for waveform, and the M stands for

Menlo Park. The demultiplexed file normally has an extension of .DMX. Although this program does not force these conventions on you, some of the analysis programs, which use demultiplexed files as input, have a default extension of .DMX. So if you are using a different extension, then when using those programs, you will have to type the complete filename (including the file extension), rather than just the filename. DEMUX is rather disk intensive, so if you are using a PC which does not have a very fast hard disk, then you may want to create a RAM disk to do your demultiplexing on.

The EXTRACT program by Dean Tottingham, is used to produce a new waveform

DEMUX: A De-multiplexing Program

- Automatically de-multiplexes a SUDS waveform file
- Written by Dean Tottingham
- Usage: **demux input-filename output-filename**

Advantages

- Automatically takes care of de-multiplexing
- Creates a demuxed file needed by many programs

Disadvantages

- Slow: Requires a RAM disk for speed

Figure 12. DEMUX program.

file which contains a subset of all the recorded channels (See Figure 14). For instance, if some channels were not working correctly when an earthquake was recorded, you can remove them from subsequent processing with the EXTRACT program. EXTRACT extracts traces according to a list specified in a control file.

Usage is:

extract [control filename] [input filename] [output filename]

The input file must be a SUDS file, and the output file written by EXTRACT will also be a SUDS file. The control file is an ASCII file, which lists the stations to extract in sequential order. Unfortunately, EXTRACT has limited error handling facilities, so if the control file has an error, unpredictable results may occur.

Two programs which are similar in purpose are ST2ASC AND STDESCR. These programs are written by Peter Ward and are intended to let you easily look at a SUDS file. Since a SUDS file is binary, you can't just type it, or use a word processor, if you want to examine it. The ST2ASC program simply creates an ASCII version of the SUDS file, so that you can easily examine it. Usage is:

st2asc [-v] input filename [> output filename]

The optional -v switch produces a verbose or complete listing. Otherwise, the output is a

EXTRACT: A Seismic Trace Extract Program

- Automatically extract seismic traces from a SUDS demuxed file according to a list specified in the Control file
- Written by Dean Tottingham
- Usage:
`extract control-filename input-filename output-filename`

Advantages

- Automatically extract selected seismic traces

Disadvantages

- Control file must be prepared accurately
- Input data file must be a SUDS demuxed file

Figure 13. EXTRACT program.

summary of the event. Output of ST2ASC is sent to the standard output device, which is usually the screen monitor. If you want to produce a file or else send the output to a printer just use the DOS redirection symbol (>) followed by the device name or the output filename.

STDESCR produces an ASCII file containing all the important structures of the SUDS file. It provides a summary look at the event. Usage of STDESCR is:

stdescr [-o output-filename] input-filename

If the optional output filename switch is not invoked, then output goes to the screen monitor.

Both ST2ASC and STDESCR are disk intensive, so expect some delay even if your machine has a fast hard disk. Also, neither program has extensive error handling built in. If you make a mistake in typing, the program is not going to complain. It will just do something unexpected, or else terminate with little or no explanation.

SOFTWARE FOR EARTHQUAKE LOCATION

Before discussing the software specifics (See Figure 15), let's first talk about what data are needed. The first data are instrument coordinates. Next you need a velocity model,

ST2ASC & STDESCR: Utilities for SUDS files

- ST2ASC: Converts a SUDS data file to an ASCII file
- STDESCR: Gives a summary of the SUDS structures contained in a specified SUDS data file
- Written by Peter Ward
- Usage 1: **st2asc [-v] input-filename > output-filename**
- Usage 2: **stdescr [-o output-filename] input-filename**

Advantages

- Let you examine a SUDS file in ASCII format
- Use STDESCR for a quick look
- Use ST2ASC only if you need an ASCII data file

Disadvantages

- Intensive data I/O and conversion in ST2ASC
- No error handling facility to help you
- Requires a RAM disk for speed in ST2ASC

Figure 14. ST2ASC & STDESCR.

i.e., you must have some idea what the crustal structure is like under the seismic network. You need some control parameters to constrain the program. Finally, you need the phase data from the seismograms. Most of this information is more or less the same for all earthquakes processed by your network, but every earthquake has a unique set of phase data. What is needed is an efficient way to prepare the phase data given by the earthquake waveform data.

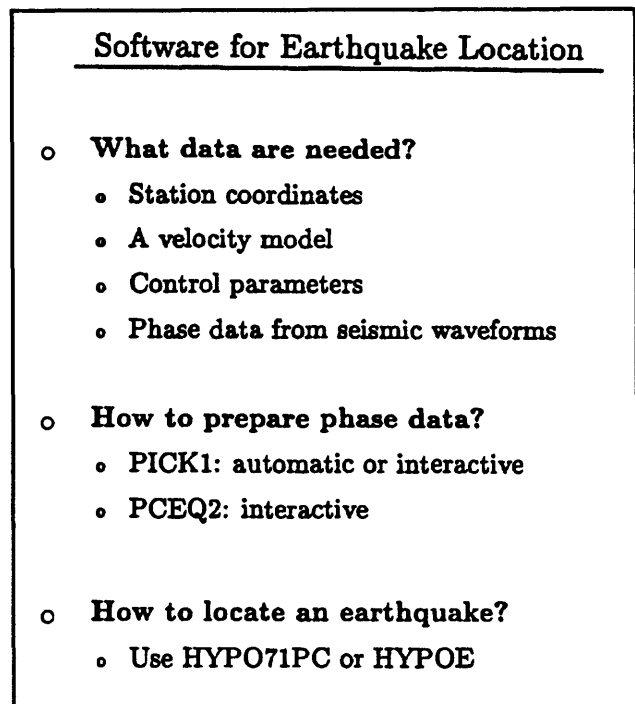


Figure 15. Location software.

One way to do this is with the program PICK1 written by Willie Lee and Robert Banfill (See Figure 16). This program will pick phase arrivals automatically or interactively. It provides all the standard information used for subsequent processing with the exception of amplitude measurements. It automatically determines first P-arrival times, polarities, and coda durations. PICK1 is actually just a temporary name for this program since it is still under development. When the program is completed, it will be called SUDSPICK. Usage of PICK1 is:

pick1 filename [/b]

The filename is the name of the demultiplexed waveform file. and the optional /b switch is for batch processing. The algorithm used in PICK1 is more stable and faster than the one by

Allen (1978). Also, PICK1

automatically determines a weight or quality factor for its picks. Therefore, the picks with the highest confidence are given the most weight in the solution process. As part of the continuing work on PICK1, we are trying other algorithms. The intent is to create a picker which can produce good picks at least 90% of the time.

Another approach is to use the program PCEQ2 written by Carlos Valdes (See Figure 17). PCEQ2 is a program designed to perform waveform analysis on a single seismic trace. It can be used to pick arrivals, determine coda duration, perform filtering, and produce amplitude spectra. The primary disadvantage of PCEQ2 for this purpose is that you have to do all the work. It does not determine any parameters

PICK1: An Offline P-picking Program

- Automatic P-picking using a moving window algorithm in either batch or interactive mode
- Written by Robert Banfill and Willie Lee
- Usage: `pick1 filename [/b]`

Advantages

- Supports either batch or interactive mode
- Creates a phase file for earthquake location
- More stable & faster than Allen (1978) algorithm

Disadvantages

- Still under development and not yet completed

Figure 16. PICK1 program.

PCEQ2: A Waveform Analysis Program

- Waveform analysis on a single seismic trace: windowing, picking arrivals and coda duration, filtering, and FFT
- Written by Carlos Valdes
- Usage: `pceq2 /s:4` for VGA
- PCEQ: Published in IASPEI Software Library Volume 1 (does not support SUDS format)
- PCEQ2: Available now (supports SUDS format)

Advantages

- Efficient analysis on a single seismic trace
- Produces a phase file for earthquake location
- Works for Hercules, EGA or VGA color
- Supports dot matrix IBM-compatible printer

Disadvantages

- Limited to one seismic trace at a time
- Does not support higher resolution graphics
- Requires a mouse
- Does not support laser printer directly

Figure 17. PCEQ2 program.

automatically. Everything is done using a mouse. Program usage is:

pceq2 [/e:4]

The /e:4 switch is used if your system uses a VGA adapter. In that case, PCEQ2 needs the parameter in order to correctly map the display onto the screen. The original program (PCEQ) did not support the SUDS format, but PCEQ2 does. PCEQ2 can be used to print a hard copy of a single trace, if you have an IBM compatible printer connected to your computer. At this time, PCEQ2 only supports a resolution of 640 X 350 pixels even on a VGA monitor. Also, it will not work at all without a mouse. These problems are being worked on now.

The programs available for hypocenter location are HYPO71PC and HYPO2. In Chapters 10 and 11, John Lahr and I will talk more about these two programs, but here I want to outline the HYPO71PC program (See Figure 18). HYPO71PC locates

earthquakes using a layered earth model. There is also a version of HYPO71PC known as HYPO2. HYPO71PC does not support blank data fields, and HYPO2 is the fix for that. HYPO71 has been in use for about 25 years, so it is a very stable and well tested program. However, it is limited to a multiple layer earth model, so for some applications it may be inadequate. Also its punch card heritage is still quite evident, and is the cause of many

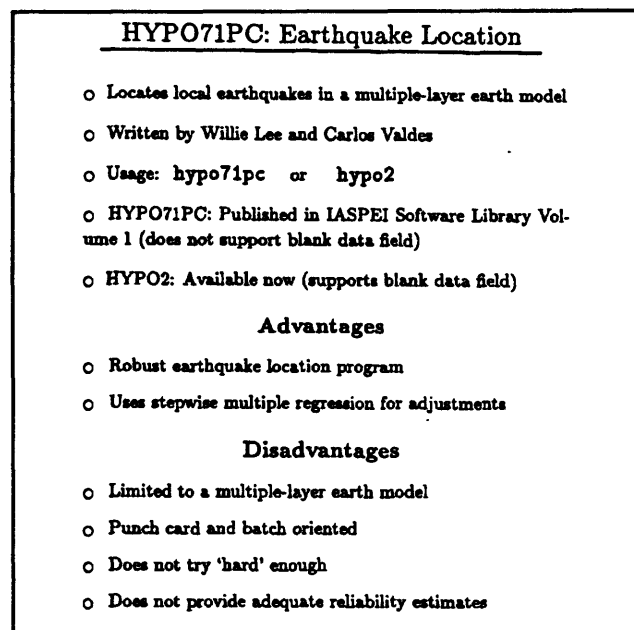


Figure 18. HYPO71PC program.

complaints by users. Many users also complain that HYPO71 does not "try" hard enough to get an optimum solution. This may be true if all the default control parameters are used. However, since its behavior is controlled by parameters in the input file, adjustments can be made to certain parameters which will cause the program to "try" much harder. It also does not provide adequate reliability estimates on its solutions.

The SUDSPLOT program by Robert Banfill was designed to allow you to look at and plot seismic record sections (See Figure 19). He will discuss it in detail in the next chapter, but I will present an outline of it here. SUDSPLOT supports both screen display and an HP laser jet printer as a hard copy device. It is command line oriented. All the control information

for the program is specified on the command line. If you forget how to specify the control

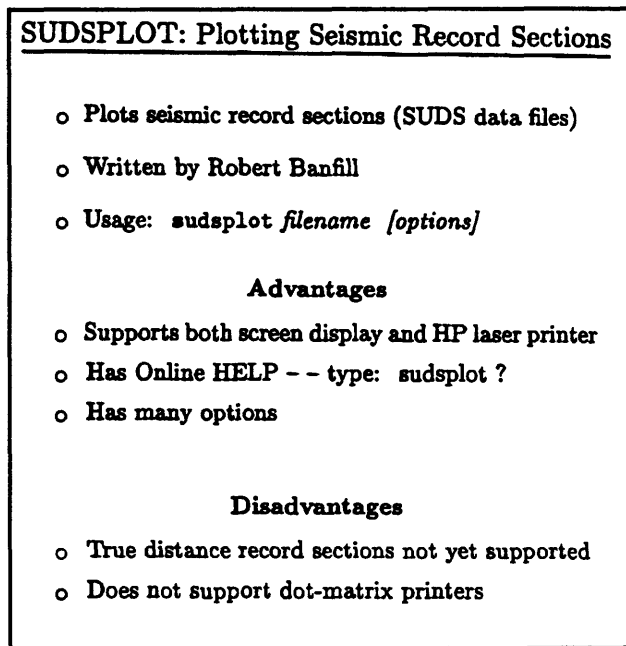


Figure 19. SUDSPLOT program.

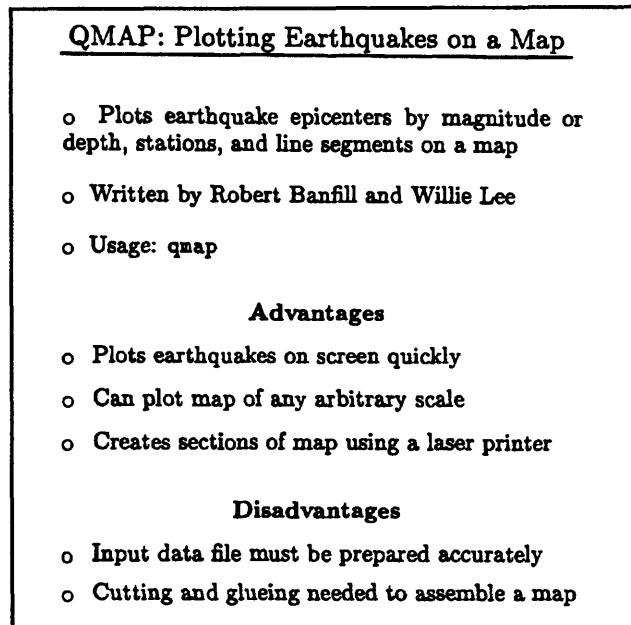


Figure 20. QMAP program.

information just type:

SUDSPLOT

and it will describe how it should be used. At this point the program does not support true distance record sections, but work to support that option is in progress. The program also does not support dot matrix printers, although given the graphics speed of dot matrix printers that is not much of a loss.

The QMAP program by Robert Banfill and Willie Lee is used to plot earthquakes on a map (See Figure 20). It plots earthquakes epicenters by magnitude or by depth, and also plots stations and an arbitrary set of line segments defining geographic boundaries. Usage is:

qmap

After the program starts, it prompts you for some information and then proceeds to produce the requested map. QMAP plots maps on the screen very quickly, and can produce a map of any scale. It does this by printing sections of the map on 8.5 by 11 sheets. The bigger the map, the more sheets are used. You have to manually paste the sheets together to get the complete

map. Although it is inconvenient to have to tape the sheets of paper together, it actually turns out to be much quicker than using a large pen plotter. At this time, there is very little error

ACROSPIN: A 3-D Viewing Program

- o Allows you to view data in 3-dimensions
- o Written by David Parker
- o Published in IASPEI Software Library Volume 2

Advantages

- o Let you zoom and rotate data in 3-dimension
- o Supports Hercules, EGA or VGA color
- o Extremely fast, even on an 8088 PC
- o Supports IBM/Epson dot-matrix printer

Disadvantages

- o Input data must be prepared accurately
- o SPIN* programs for data preparation are clumsy

Figure 21. ACROSPIN program.

handling ability in QMAP. Therefore, if you don't specify the data correctly, you will get unpredictable results.

The ACROSPIN program written by David Parker allows you to view data in 3-dimensions. It lets you "zoom" in on your data or rotate it in 3 dimensions (See Figure 21). The program supports Hercules, EGA, and VGA graphics adapters. It is extremely fast, even on an 8088 PC. In fact, on the fast 386 and 486 machines you may have to slow down the PC in order to keep up with the program. Screen plots can be sent to an IBM/Epson dot matrix printer. As with many of the programs in this collection, there is not very extensive error handling built in. Therefore, the input data must be prepared very carefully. There are some programs to do the data preparation but they are very clumsy.

SOFTWARE FOR WAVEFORM ANALYSIS

We have three programs which fall under this general heading (See Figure 22). These are QCODA2, SEISGRAM, and PITSA. Both SEISGRAM and PITSA are intended to do 3-component analysis. PITSA also is a general signal processing program.

QCODA2 by Carlos Valdes and David Novelo-Casanova, is designed to analyze coda waves to determine estimates of the attenuation factor Q (See Figure 23). It shares much of the interface code of PCEQ2 and its usage is similar:

<u>Software for Waveform Analysis</u>	
o	General processing: PCEQ2, SEISGRAM, PITS
o	3-component analysis: SEISGRAM, PITSA
o	Signal processing: PITSA
o	Coda Q analysis: QCODA2

Figure 22. Waveform analysis.

qcoda2 [/e:4]

Because of the similar code, QCODA2 also does not support graphics at a resolution higher than EGA, it works with only one trace at a time, and it requires a mouse to operate. The original QCODA does not support SUDS data, but QCODA2 does support SUDS data. QCODA2 computes five different estimates of Qc. The philosophy is that: if these estimates agree substantially, then they are probably good estimates.

The SEISGRAM program will be discussed in detail in Chapter 17 by its author Anthony Lomax (See Figure 24). However, I will briefly mention some of its features here. It allows you to perform general waveform analysis procedures on seismic traces. It supports several graphics resolutions including EGA and VGA. Output can be sent to IBM/Epson compatible dot matrix printers or to HP laser printers. The original version of SEISGRAM does not directly read SUDS files. They must be converted into the SEISGRAM format. However, Anthony has created a new version of the program called SGRAM2 that does support SUDS files. SEISGRAM is also limited to 15 traces at a time. This is mostly a function of screen size. Also, because of the DOS limitation on program size, SEISGRAM can not handle very long seismic traces. We plan to eventually develop a version which can

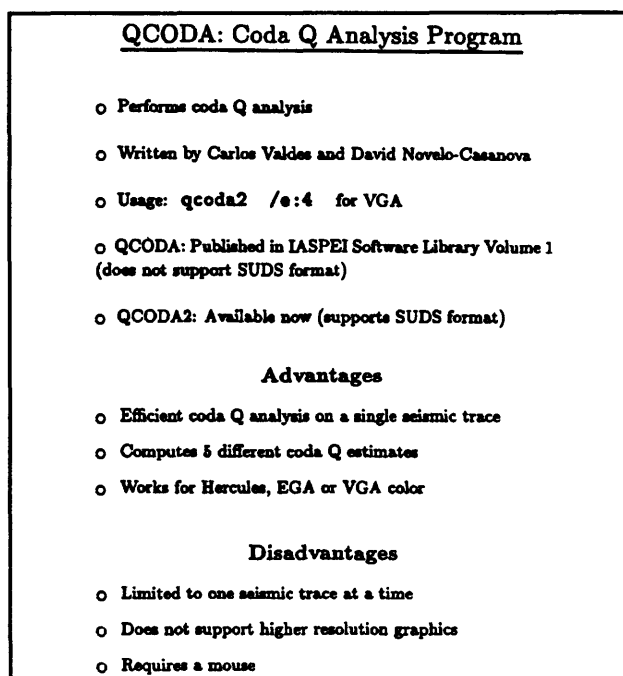


Figure 23. QCODA program.

work with arbitrarily long traces, although this version would be restricted to running on 386 and higher PCs.

PITSA (Programmable Interactive Toolbox for Seismic Analysis) has been under development for the last four or five years by Frank Scherbaum in Germany (See Figure 25). It is a very ambitious program. Essentially, it is a "smart" program

which remembers the way you like to analyze seismic data. Once you lay out a procedure, the program remembers it and from then on you don't have to guide it through the analysis. We plan to publish it in Volume 5 of the IASPEI Software Library in 1992. The program supports the same range of output devices as the other programs I have mentioned and also comes with a digital signal processing tutorial. Also, the published version will support both the PC and the SUN work station. In fact, starting with Volume 5 of the IASPEI Software Series, we plan to make all of our software work on both the PC platform and the SUN platform. It turns out that a lot of people, especially in the United States, are using SUN work stations, so to attract more users to our software library, we have to start supporting the SUN. SUN plans to port their operating system to the PC so that eventually you will be

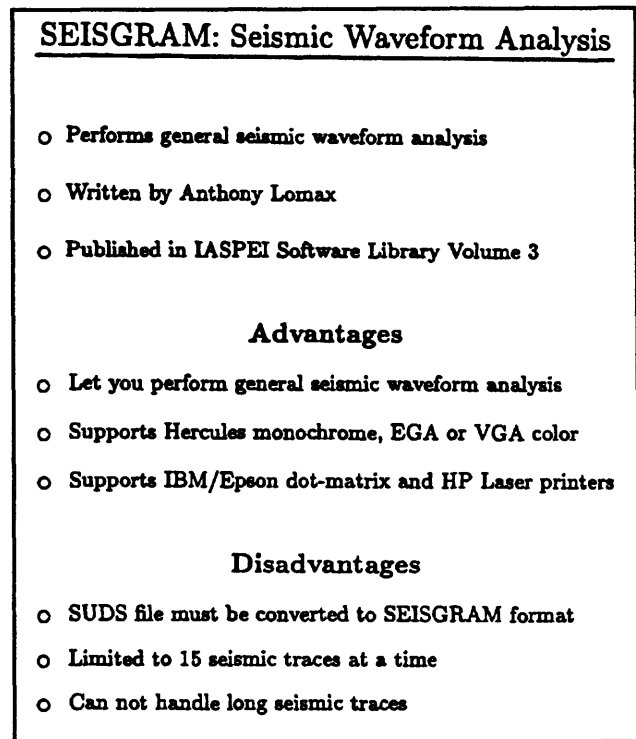


Figure 24. SEISGRAM program.

able to choose either the SUN environment or the DOS environment on your PC. PITSA is a complicated program and my estimate is that it takes about a month to become completely familiar with all its intricacies.

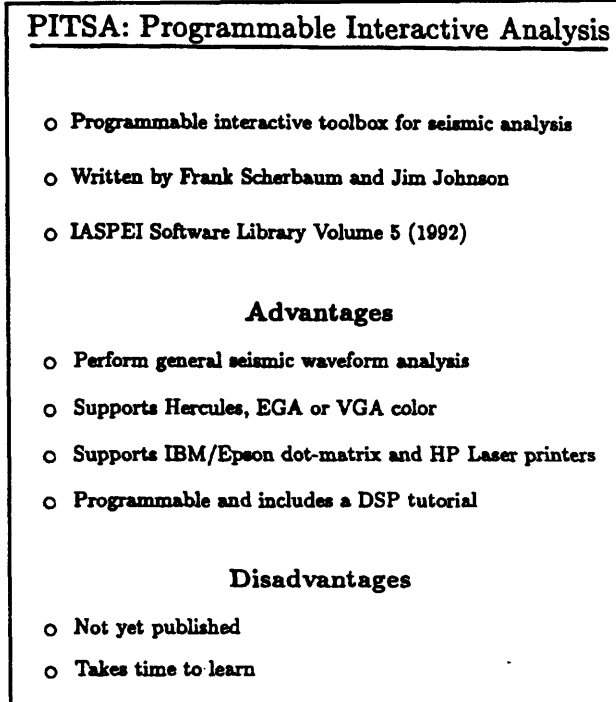


Figure 25. PITSA program.

REFERENCES

- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. Bull. Seismol. Soc. Am., 68, 1521-1532.
- Banfill, R. (1992). SUDS: Seismic Unified Data System, version 1.31, Small Systems Support, Big Water Utah.
- Lee, W. H. K. (Editor) (1989). "Toolbox for Seismic Data Acquisition, Processing, and Analysis", IASPEI Software Library, Volume 1, Seismological Society of America, El Cerrito.
- Lee, W. H. K. (Editor) (1990). "Toolbox for Plotting and Displaying Seismic and Other Data", IASPEI Software Library, Volume 2, Seismological Society of America, El Cerrito.
- Lee, W. H. K. (Editor) (1991). "Digital Seismogram Analysis and Waveform Inversion", IASPEI Software Library, Volume 3, Seismological Society of America, El Cerrito.
- Ward, P. L. (1989). SUDS: Seismic Unified Data System, U. S. Geol. Surv. Open-file Report 89-188.

9. PLOTTING SEISMOGRAMS AND MAPS

by

Robert Banfill
Small Systems Support
P. O. Box 410205, Big Water, UT 84741
(801) 675-5827

INTRODUCTION

I want to start out by outlining a typical processing cycle for a PC system using our software. Figure 1 shows the cycle in schematic form. First XDETECT captures an event and creates a waveform file. The times are adjusted using the time code stored in the file by FIXTIME and demultiplexed using DEMUX. Then SUDSPICK is used to pick the P phases and HYPO71PC is used to locate the event and make a magnitude estimate. Next SUDSPLOT is used to produce various forms of hardcopy records of the event. At that point you would archive your data and perform various housekeeping tasks.

Then you would loop back to the beginning and start over again.

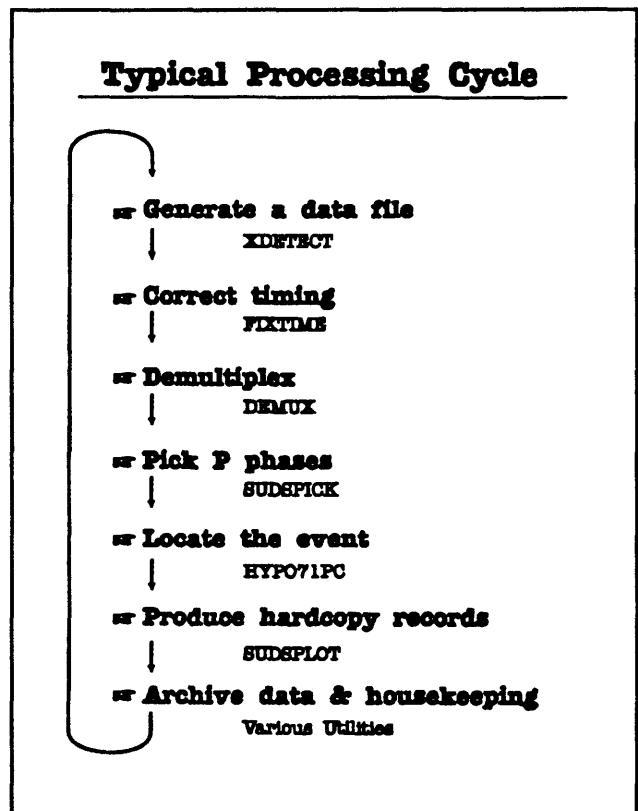


Figure 1. Typical processing cycle for PC-quake system.

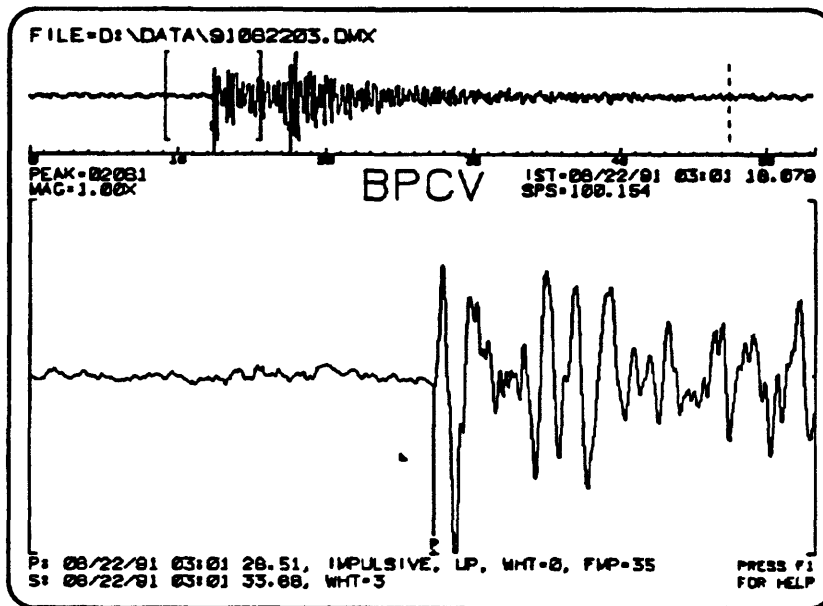
We have developed a program called SUDSMAN which automates the processing procedure from the point where the waveform file has been created to the creation of hardcopy records. To implement this system, you must have XDETECT running on one computer and SUDSMAN running on another computer with the two machines connected by a network. SUDSMAN looks across the network, checking for the existence of a waveform data file. When it finds one, it brings it over and oversees all subsequent processing through the use of standard DOS batch programming techniques. The good thing about this approach is that your data acquisition system never needs to come off line. Data acquisition goes on 24 hours per day and the processing goes on concurrently. Today I want to talk about SUDSPICK and SUDSPLOT. I also will spend some time talking about QMAP.

SUDSPICK

SUDSPICK is a program designed to do interactive or automatic phase picking. When you run it in interactive mode the display looks something like figure 2 on the next page. At the bottom of the figure is the help screen you get if you invoke SUDSPICK without any arguments or with a question mark argument. You can see that we support various super VGA modes up to 1024 X 768 pixels and 256 colors for VGA cards using the Tseng Lab chipsets. We are trying to develop a standardized graphics library which is very general and which supports both 300 dpi laserjet printers, PCL and PostScript, and super VGA graphics adapters. We hope to complete this library and build all our software around in order to simplify the development process.

Figure 2 shows a typical view of the screen when SUDSPICK is in interactive mode.

SUDSPICK



SUDSPICK Version 1.20 - W. Lee & R. Banfill

Usage: SUDSPICK [switches] inputfile
switches:

/B - Batch mode. (OFF)
/Mmode - Display mode:
1 - 2 color 720x348 Hercules Graphics Card
2 - 16 color 640x350 EGA
(3) - 16 color 640x480 VGA
4 - 16 color 800x600 SVGA
5 - 16 color 1024x768 SVGA < Tseng only
7 - 256 color 640x350 SVGA < Tseng only
8 - 256 color 640x480 SVGA < Tseng only
9 - 256 color 800x600 SVGA < Tseng only
10 - 256 color 1024x768 SVGA < Tseng only

/Cchip - VGA Chipset:
(1) - Tseng Labs ET3000
2 - Tseng Labs ET4000

/Ffont - font filename. (C:\SUDS\SIMPLEX.VEC)
inputfile - SUDS format waveform file. (SUDSPICK.DMX)

() indicates the default value.

Arguments may be placed in any order, and are not case sensitive.

Figure 2. Example screens from SUDSPICK program.

This is a VGA plot so it is not very high resolution. If you have a super VGA board that uses the Tseng Lab chip set, you can get very high resolution displays with this program.

At the top of the screen is a window showing the entire waveform and below, in the large window, is an expanded section of the waveform. The location of the expanded trace is shown in the upper window by the square brackets. The lower window is scaled so that there is exactly one pixel per sample. You see every sample in this window. The vertical line with a "P" at the bottom is the P-pick. The pick is also shown on the upper plot although, without the aid of color, it doesn't show up very clearly. This pick was done automatically. Even when you work interactively, the program will still try to pick the phase arrival and show it to you unless you explicitly disable this with the /N (Nopick) command line argument. If you like it, you are done, if not you can change it. The program also allows you to pick S-phases.

Data about the picks is displayed below the waveform in the lower window. You can see that the P phase is marked as impulsive, first motion up, weight is zero, and FWP (coda duration) is 35 seconds. The dashed line in the upper window marks coda duration. When the program picks the P-phase, if the weight is better than one, it estimates the coda duration. If you do an S-pick (which you will have to do interactively) the S pick data is displayed below the P pick data. At the top of the lower window is shown the station name extracted from your input file and in the upper left hand corner is the peak amplitude in digital counts and the magnification factor. In the upper right hand corner are displayed the initial sample time and the digitization or sampling rate. Finally, in the upper left hand corner of the display screen is the data file specification.

To run SUDSPICK, from the DOS command line type:

sudspick [switches] input file

Most of the software that I have written is very "friendly" when it comes to the command line. It doesn't matter whether the switches come before or after the filename or what order they are in and nothing is case-sensitive. When the program is run in batch mode (/b switch) it doesn't require any interaction from the user. However, it does display status information every time it makes a pick such as station name, pick time, etc. When it is in batch mode, the program does as much error trapping as I could squeeze into it. It returns exit codes so that you can use the "IFERRORLEVEL" construct to test for errors. If it encountered any kind of error, it exits with an exit code of one or greater.

The mode switch (/m) followed by a number specifies the video mode to use. The default mode is indicated by parentheses around the number. The /Cchip switch selects which chip set is used. The font file name switch selects the name of your font file. You can see that the default file is a fully qualified file specification. The program determines this name by finding what directory it was loaded from and then prefixing that path to the name of the necessary font file. So as long as there is a path to the directory in which you keep SUDSPICK and as long as the necessary support files are kept in that directory, you can run SUDSPICK from anywhere with no problems.

The default input file name is SUDSPICK.DMX. .DMX is a convention Willie and I established which indicates that the file is a demultiplexed SUDS file. As long as your files have the .DMX extension, you do not have to include the extension as part of the file name on the SUDSPICK command line.

SUDSPLOT

That brings us to SUDSPLOT, a general purpose waveform plotting program. The figure on the next page is out of the SUDSPLOT documentation and is a typical example of SUDSPLOT output. This is a very small local earthquake. There are 128 channels plotted, 64 to a page in this example. By each trace are listed the station name and the peak sample value. At the top of the page are the initial sample time, the sample rate, decimation factor and data file name. Also, in the upper right corner are the time and date that the plot was produced and the sheet number. You can see that this was a small event since only a few of the channels are responding to it.

IST=06/22/91 03:01 16.079 SPS=100.154 DEC=4 D:\DATA\91062203.DMX

Sheet 1
09/11/91 23:00

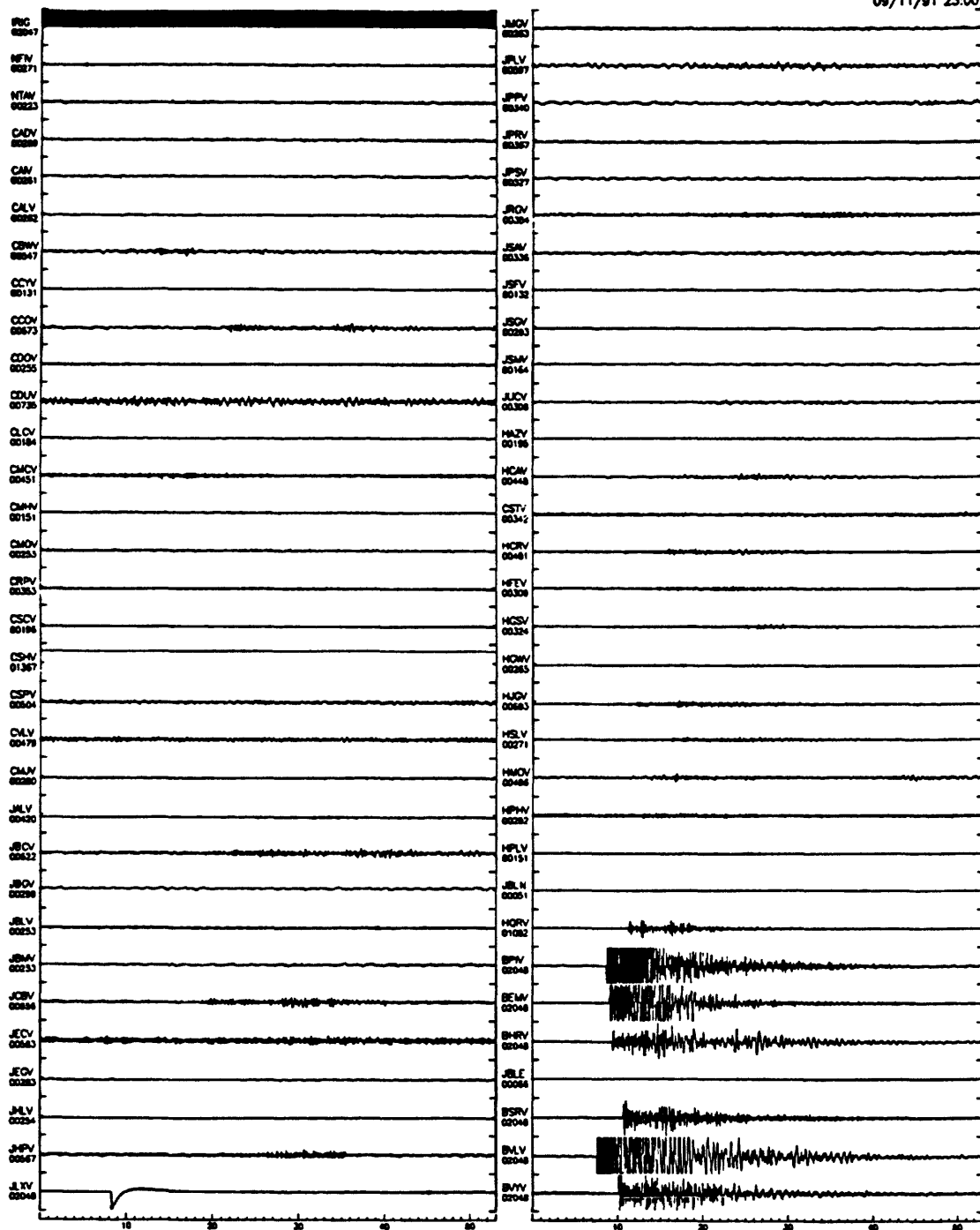
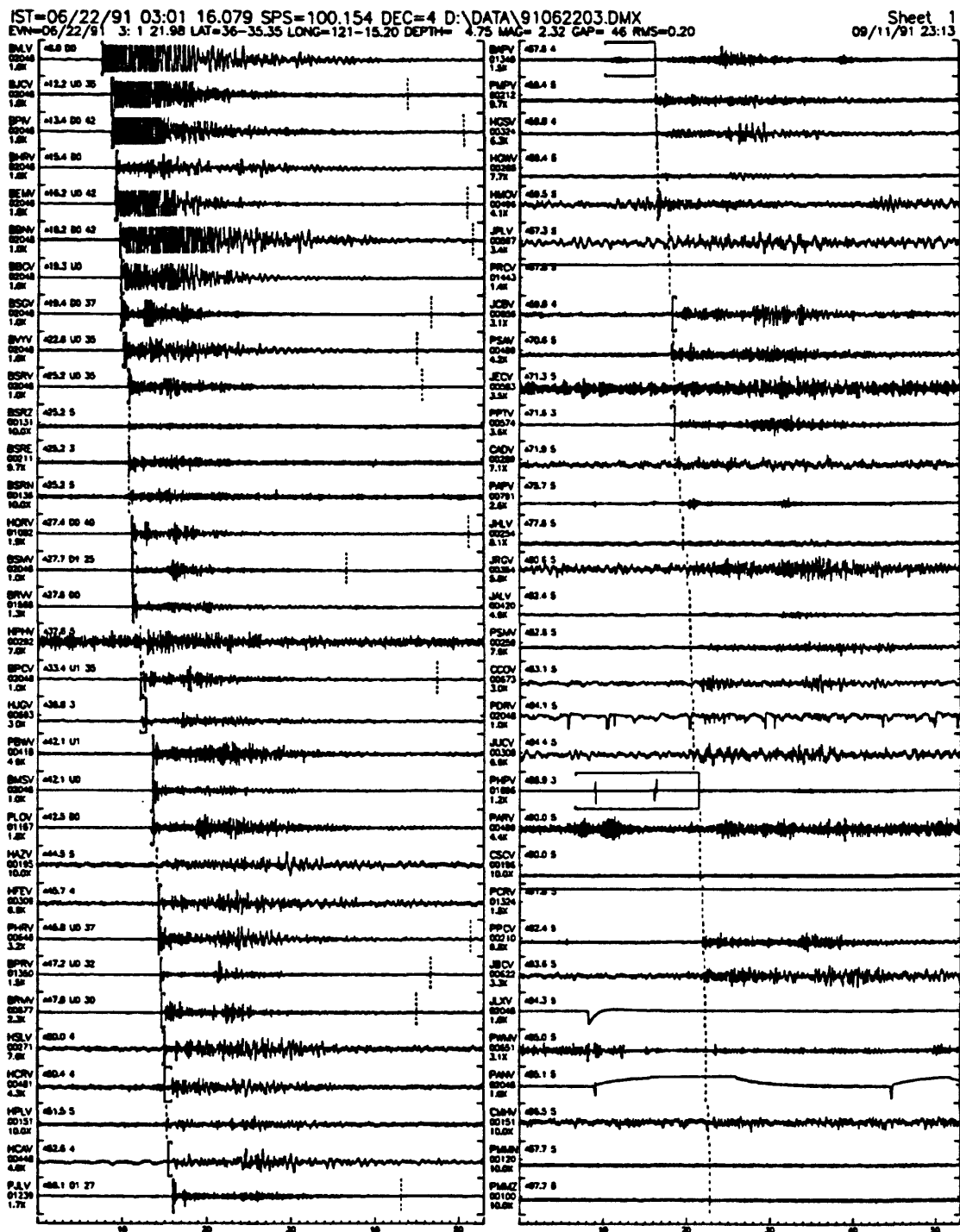


Figure 3. Example output from SUDSPLOT.

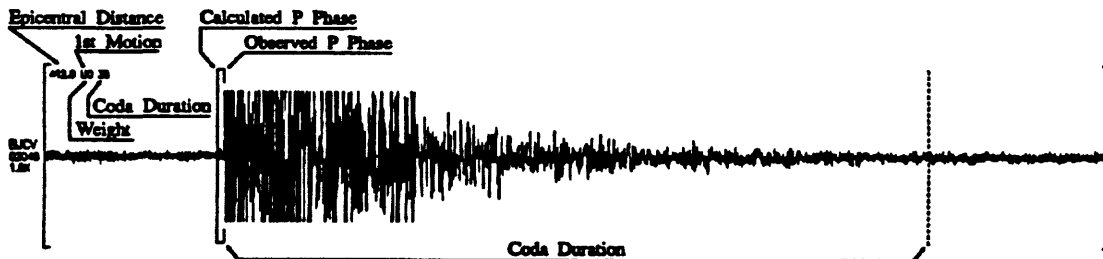
Figure 3 is an example of "raw" data. No source location has been performed. It is simply confirmation that an event did occur at a certain time. You could print this out and put it away in a binder.

However, once we have located the event we can make a plot that tells us a lot more. On the next page is a plot of an event which has been located using HYPO71PC. Notice that on this plot the stations are displayed ordered by epicentral distance. At the top of the page below the header is the event line from HYPO71PC. It includes the event origin time, epicenter coordinates, hypocentral depth, magnitude, gap, and RMS from the solution, and for each station there is now included a magnification value.

Most programs either plot traces with a constant scale factor for all the traces or else they autoscale each trace to fit inside the plotting area allocated for that trace. In SUDSPLOT we use a slightly different idea. You can specify the maximum scaling that you want to occur. For example you can specify a maximum magnification of 10X. Then SUDSPLOT will try to autoscale each trace as long as the autoscaling will not result in a magnification greater than 10X. That keeps you from taking "dead" traces and magnifying them to the point where they look like signals and confuse the issue. The top portion of figure five shows one trace enlarged to show the detail. The observed P phase shown on the plot is the phase pick made by SUDSPICK. The calculated phase pick is the one predicted by the HYPO71 solution. Notice on figure four that some of the traces are marked by a dashed line. This dashed line indicates that SUDSPICK did not pick an arrival time for that trace. The location of the dashed line is set according to where the HYPO71 solution indicates that the arrival should have occurred. Every channel showing a dashed line has a weight of five.



SUDSPLOT



Marking of phase arrivals and coda duration when using the 1A option.

SUDSPLOT Version 1.32 - R.Banfill

Usage: SUDSPLOT [options] inputfile [options] <+
options:

```
/Ddriver  Driver filespec. (C:\SUDES\SUDSPLOT.DRV$)
/Pport    Output port. (LPT1:)
/A        Plot phase arrivals on each trace. (OFF)
/An       Start/end plot n seconds before/after earliest arrival. (OFF)
/Wn       Include stations with pick weight <= n. (5)
/B        Remove baseline. Subtract avg value from each sample. (OFF)
/Bn       Remove baseline. Subtract avg of first n samples. (200)
/Tn       Plot n traces per page. 1<n<=64. (64)
/T*       Auto traces per page.
/Mn       Maximum amplitude scaling magnification. (1X)
/X        Decimate, windowing. (OFF)
/Xn       Decimate, simple, plot every n'th sample. 1<n. (OFF)
/Jn       Jump, begin plotting n seconds into trace. (0)
/In       Length, plot n seconds of trace. (ALL)
/Slist    Station list filespec. (HYPO71PC.PRT)
```

inputfile SUDS format waveform file. (SUDSPLOT.DMX)

() indicates default values.

Arguments may be placed in any order, and are not case sensitive.

Figure 5. Expanded SUDSPLOT trace and help screen.

It takes about three minutes to produce one of these plots. The time required is not really a function of your computer's speed. Rather, it is the time required to send all the information for the plot through the parallel port to the HP laserjet printer. The whole bottleneck is the speed at which the parallel port operates. Our latest version of SUDSPICK supports PostScript so you may be able to get a little faster printing with this version if you have a fast PostScript printer.

The lower half of Figure 5 shows the help screen from SUDSPLOT which is basically a list of the possible options. To produce plots which show the results of processing the data with HYPO71 use the /A option. This tells SUDSPLOT to look for a .PRT file which is the output from HYPO71. The default filename is HYPO71PC.PRT. If you pass the /A switch a number (n) it tells SUDSPICK to only plot from n seconds before the first p onset to n seconds after the last. This is in effect, a close up view of the early part of the event. The /Wn switch can be used to exclude traces with pick weights greater than n. Figure 6 shows the effect of using /A2 and /W2 on the event shown in figure 4. Since only stations with pick weights of two or less have been included in this plot, the original 128 traces have been reduced to only 21 traces. In the expanded view, it is much easier to see just how precise the automatic picks really were.

The /T option controls how many traces are printed per page. The default value is 64. If you specify T* the program examines the SUDSfile and considers what other options you have specified. From that information it determines an optimum number of traces to display per page. If that is greater than 32 it prints them side by side. Otherwise, each trace is a page width long.

IST=06/22/91 03:01 16.079 SPS=100.154 DEC=1 D:\DATA\91062203.DMX
 EVN=06/22/91 3: 1 21.98 LAT=36-35.35 LONG=121-15.20 DEPTH= 4.75 MAG= 2.32 GAP= 46 RMS=0.20

Sheet 1
 09/11/91 23:23

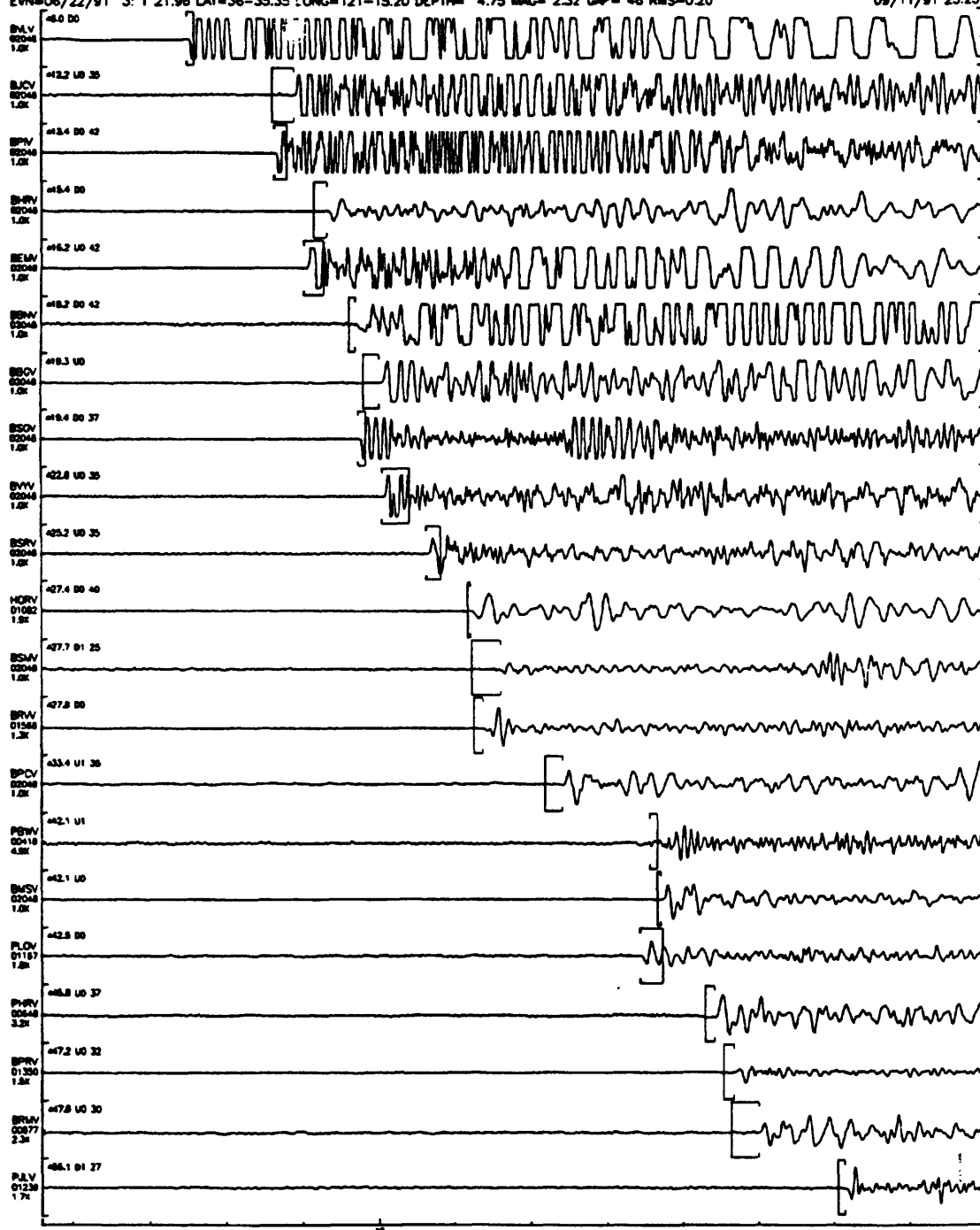


Figure 6. Expanded traces.

The /X option controls decimation. The whole reason for decimation is to decrease the time required to produce a plot. If you have many long traces with hundreds of sample points, decimation can make a dramatic difference in the amount of time required to produce the plot. You can perform simple decimation by just passing a number with the /X switch. Then SUDSPLOT will plot every nth sample.

The problem with that kind of decimation is that you may distort the character of the waveform. Instead, you may decimate using the /X option with no number passed. In this case, the program looks at the trace and sees how many samples there are to plot and how many pixels are available. From this it determines what the best decimation factor would be. Say, for example, it determines that factor to be 3. Then the program looks at a window of twice that number of samples and finds the minimum and maximum sample values. It then just plots those two points. Although not quite as fast as simple decimation, this technique preserves the peaks in the data. I recommend using just /X.

The /Jn option is used to start plotting n seconds into the trace, and the /Ln option controls how many seconds of trace are plotted after the start of plotting. The /Slist option would be primarily useful if you are using some program other than HYPO71PC to produce a .PRT file. The default is to use the file produced by HYPO71PC, but you can use this option to change that to something else.

Since SUDSPLOT is controlled by command line parameters, it is very much batch oriented. It also returns exit codes. If there is any kind of error it exits and passes a code to the batch file so that a "civilized" shut down can be made.

Two other switches I have not mentioned so far are /Ddriver and /Pport. The driver

switch controls which device driver is used for output. You can choose drivers for Hercules, EGA, VGA, laserjet, etc. You can specify the port with the /P option. So to send output to a VGA monitor you would include on the command line /Dvga /Pcon.

QMAP

This brings us to the QMAP program which Willie wrote years ago. It has been the basis for an untold number of mapping programs. I was given the original punch card FORTRAN code which needed to be modified for our purposes. Since I share Willie's aversion to large plotters with all their attendant problems with pens, paper registration etc. I decided that we needed to do something different. We needed to be able to produce maps, big maps, using a laser printer. We also wanted the process to be relatively painless. What we decided to do was use his code as the core of a program which would produce a large map by tiling it out over several sheets of paper and use registration marks to align the individual sheets. As an example of the speed of the program we wrote, I produced a large (15 8½ x 11 sheets) map last night in eleven minutes from the time I started the program to the time the map was complete. Figure 7 is a scaled down version of the map.

There is a legend produced with the map. We let you move the legend to wherever you want on the map. This is really the only enhancement to the original program. You move the legend by moving a boundingbox cursor on the screen to the point where you want the legend to be placed. The reason I made the legend mobile is that if I put it in one spot, sure enough, it would end up right on top of something you want to look at. You can define your classes and your symbol types for the legend. You also can specify a set of line

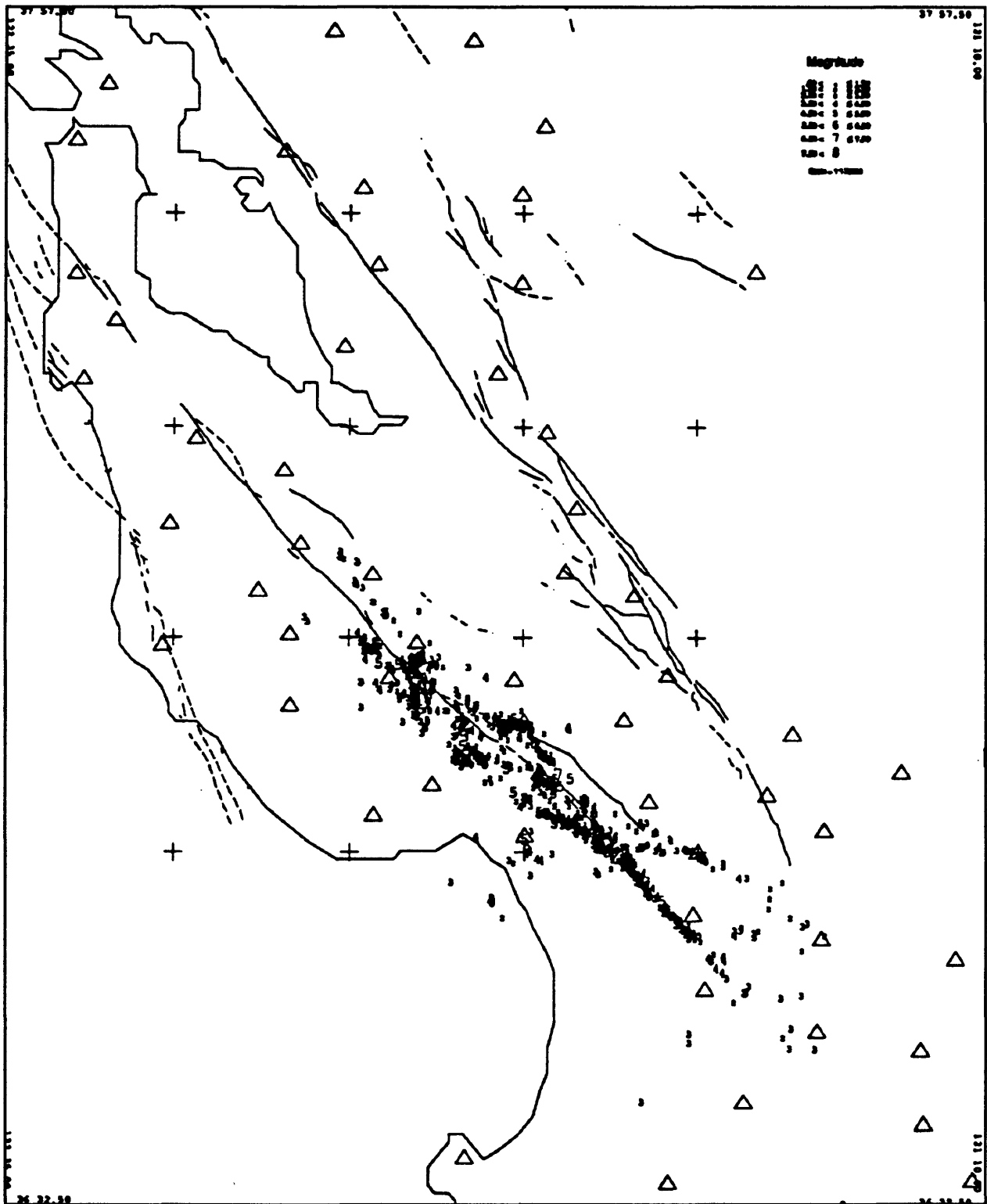


Figure 7. Loma Prieta aftershock sequence - October 1989.

segments which can be used to define such things as faults, coastlines, etc. Also, you can put

a title on the map.

There are two versions of this program. One generates a PostScript file and one generates an HP graphics language (HPGL-2) file. This could be used to drive a large HP plotter or else could be converted using a utility program to some other format, say to drive a Houston Instruments plotter. That program (called QMAPHP) also generates an HPGL-2 file which will tile the map on a laserjet III printer. The old laserjet II printers do not have the HPGL-2 interpreter so you are out of luck with those printers. If you have a laserjet III or better and you don't have a PostScript cartridge you can still print the maps but it will be extremely slow compared to printing with the PostScript cartridge. For instance, I printed this map on the laserjet III (connected to a 486 machine) and it took 42 minutes. With PostScript, it took about 4 minutes. I recommend PostScript.

Usage of QMAP is described fully in the PC-QMAP users guide, but I want to give sort of an overview of the manual here. The program uses an input file much like the other programs that started out during punch card days. Just think of the input file as a stack of punch cards. The input file is made up of six sections. There is the control section which lets you control things like the size of the map, its scale and so on. The legend information constitutes the second part of the input file. The third section contains the station data, basically the same information as in the header file for HYPO71. The pointdata section contains the point data which comes from the HYPO71.PUN files. Basically, those are just epicenter locations. The last section contains line data which could be anything. This is in the form of X Y pairs delimited by blank lines to move the pen before a move. For some of the data you must provide FORTRAN edit descriptors.

When you run the program you will see the map being printed sideways on the screen. It uses the entire screen. You will see a grid of fine gray lines which show you where the edges of the sheets of paper will be. Since the screen is being used for graphic display, all the program's text output goes into a log file. Basically, the log file is an annotated copy of the input file. Any points off the map are marked and if there are any errors, the error number is listed so you can look up the cause. In other words, the error trapping is pretty primitive by current standards.

10. MATHEMATICS FOR EARTHQUAKE LOCATION

by

W. H. K. Lee

U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

It used to be said that every equation put on the board put at least ten people to sleep. Unfortunately today we are going to deal with the mathematics of earthquake location (Figure 1), so I hope that you don't fall asleep too quickly. The earthquake location problem is essentially a problem in non-linear optimization. So my plan is to first spend some time discussing the nature of the earthquake location problem, some methods of optimization, and then show how these

optimization methods can be applied to the earthquake location problem. I will discuss two classes of optimization procedures. The first is called the method of steepest descent, and the other is called Newton's method. Both of these methods are based on a linearization of the problem by means of a Taylor's series expansion. Next, I will show how to apply these

Mathematics for Earthquake Location

W. H. K. Lee

- o **Nonlinear Optimization**
 - o Problem and definitions
 - o Methods in optimization
 - o Taylor expansion
 - o Method of steepest descent
 - o Newton's method
- o **Nonlinear Least Squares**
 - o Residual vector and objective function
 - o Jacobian matrix and Hessian matrix
 - o Gauss-Newton method
- o **Earthquake Location**
 - o Some simple methods
 - o Geiger's method
 - o Implementation of Geiger's method

Figure 1. Mathematics for earthquake location.

methods in the non-linear least squares procedure. I will also talk in some detail about the Gauss-Newton method, which is actually used in earthquake location under the name of Geiger's method (Geiger, 1912). Finally, I will talk about some simple methods of earthquake location, Geiger's method, and implementation of Geiger's method.

THE EARTHQUAKE LOCATION PROBLEM

Before plunging into the mathematics, I will try to give a simple illustration of the problem. Suppose we have a number of stations and an earthquake occurs. Then there will be arrival times at each station. The problem is: if we know those arrival times, can we use them to tell where the earthquake occurred? Can we determine when the earthquake occurred and how big it was? Those are the questions that a seismic network is designed to answer.

First of all you must realize that if you don't have good data, you can not get a good

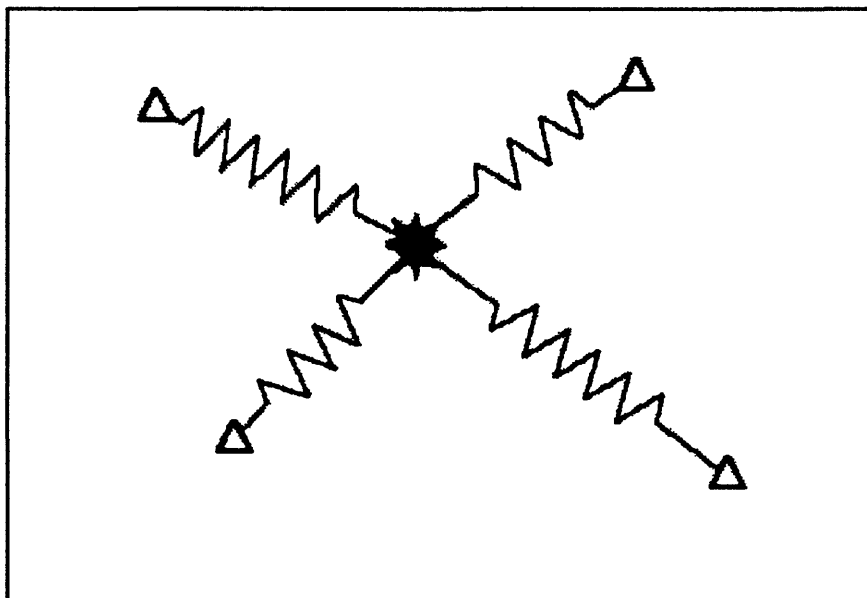


Figure 2. Spring analogy to hypocenter location problem.

location. A lot of people refuse to believe that and get themselves into trouble. Think of it this way: If you have no money in the bank, you can not withdraw money. It is the same thing in seismology. If you don't have good data, you can not locate the earthquake. To make the situation easier to visualize, consider a two-dimensional analogy consisting of a set of springs arranged as in Figure 2 (See p. 208). Here, the triangles represent stations, and the common point of the springs represents the location of the earthquake. In this analogy, the stiffness of the springs corresponds to how well you know the velocities and arrival times. With the arrangement of springs as shown in the figure and with arrival times and velocities well constrained (i.e., very stiff springs), there is very little choice as to where to put the common point (earthquake location). Displacement very far from the equilibrium position will result in a very strong restoring force.

On the other hand, if all the stations are on one side of the earthquake location as in

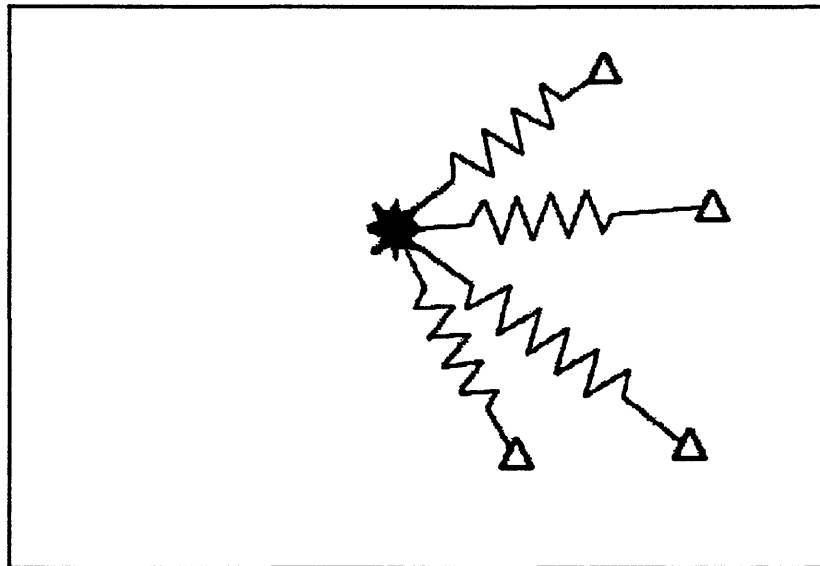


Figure 3. Spring analogy with all stations on one side of hypocenter.

Figure 3 (See p. 209), there is much greater latitude for spring motion in some directions without generating a strong resisting force. In this second case, what can you do? The best solution is to use other arrivals such as S-wave arrivals. Use of S-wave arrivals is equivalent to strongly constraining the length of the spring, making it much more difficult to move the common point of the springs.

In a real earthquake location problem, things are more difficult. The problem is four dimensional and all the stations are on the surface. This make it difficult to constrain the depth with P arrivals alone. What it all boils down to is: in earthquake location, the more stations you have, the better off you are, and you need to have potential source regions surrounded with stations if you want to get well constrained source locations. If the earthquake is outside the network, then you must use other arrival times. Many people ask why we only use first arrival times in our routine source locations. The problem is that unless you have three-component data, it is not that easy to pick other phases, especially with automated phase pickers. Also, if you have a very complicated earth model, it is not that easy to figure out what is the minimum travel-time path for secondary arrivals.

NON-LINEAR OPTIMIZATION

Now we will continue with the mathematics (Figure 4). Our basic problem in non-linear optimization is that we want to minimize some function:

$$\psi = F(x_1, x_2, \dots, x_m) \quad (1)$$

Where F is called the objective function. In the case of earthquake location, F is the sum of residuals squared. But, in general, F is an arbitrary function of the m parameters. The

question is, How do we go about this procedure?

An obvious approach is to use a search method in which we simply try many possible vectors, using an m -dimensional grid to generate the vectors. This is a brute force method, and it has the problem of being relatively slow, since so many functional evaluations must be made. Also, if a grid is used to generate the trial vectors, then the accuracy will be no better than half the grid spacing. Therefore, it is possible that the global minimum of the objective function lies between grid points. In one dimension, the search method is a reasonable approach. But, as we increase the number of dimensions, the required number of functional evaluations increases as a power function of the number of

Methods in optimization:

- (1) Search methods: function only
- (2) Derivative methods: use derivatives
 - o Steepest descent: 1st derivative
 - o Newton: 1st and 2nd derivatives

Figure 4. Methods in optimization.

Taylor Expansion

$$F(\mathbf{x} + \delta\mathbf{x}) \approx F(\mathbf{x}) + \mathbf{g}^T \delta\mathbf{x} + \frac{1}{2} \delta\mathbf{x}^T \mathbf{H} \delta\mathbf{x} + \dots$$

where \mathbf{g} is the gradient vector:

$$\mathbf{g} \equiv \nabla F(\mathbf{x}) = \begin{pmatrix} \partial F / \partial x_1 \\ \partial F / \partial x_2 \\ \vdots \\ \partial F / \partial x_n \end{pmatrix}$$

and \mathbf{H} is the Hessian matrix:

$$\mathbf{H} \equiv \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2^2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1} & \frac{\partial^2 F}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_n^2} \end{pmatrix}$$

Figure 5. Taylor expansion.

dimensions. This makes the search method look unattractive. Nevertheless, as computers increase in speed, the search method may come into use in the future.

A faster approach to minimizing the objective function is the derivative method. The derivative method uses the information from the function's derivatives to increase the convergence speed of the algorithm. There are two derivative methods in common use. One uses only the first derivative, and is called the method of steepest descent. The other uses both first and second derivatives, and is called Newton's method. The rationale behind the derivative methods is to use the information in the function's derivatives to move the current guess towards the minimum in the most direct manner. A physical analogy is that the quickest way to move from the top of a hill to the bottom is to move perpendicular to the contour lines of the hill, that is, follow the gradient of the topography.

To simplify the notation, define a parameter vector \mathbf{x} such that:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_m \end{pmatrix} \quad (2)$$

This vector may be due to our initial guess, and it is probably not correct, so we will want to adjust the vector. Therefore we introduce the adjustment vector:

$$\delta \mathbf{x} = \begin{pmatrix} \delta x_1 \\ \delta x_2 \\ \cdot \\ \cdot \\ \delta x_m \end{pmatrix} \quad (3)$$

The adjustment vector is added to the parameter vector and a new residual is calculated. If

the residual is acceptably small, then the adjusted parameter vector is accepted as the solution. Otherwise, the process is iterated.

The means of obtaining the adjustment vector is based on a Taylor expansion of the function. Figure 5 shows the first few terms of a Taylor expansion for a vector function $F(\mathbf{x} + \delta\mathbf{x})$ expressed in vector notation. In that expansion, \mathbf{g} is called the gradient vector, and \mathbf{H} is called the Hessian matrix. These two quantities are analogous to the first and second derivatives from the one-dimensional Taylor series.

The steepest descent method

only uses the gradient vector term, while Newton's method uses both the gradient vector and the Hessian matrix. When we approximate a function using a Taylor series, the more terms we use, the more accurate the approximation, so we would expect that for a given number of iterations, Newton's method would give a better approximation to the function. However, calculation of the second derivative can sometimes introduce significant computational error, so Newton's method can sometimes backfire.

As shown in Figure 6, the first correction term used in the steepest descent method is

Method of Steepest Descent

Use only the first-correction term:

$$\delta\Psi = \mathbf{g}^T \delta\mathbf{x} = \|\mathbf{g}\| \|\delta\mathbf{x}\| \cos\theta$$

where θ is the angle between the vectors \mathbf{g} and $\delta\mathbf{x}$, and $\|\cdot\|$ denotes the vector norm or length.

For any given $\|\mathbf{g}\|$ and $\|\delta\mathbf{x}\|$, $\delta\Psi$ depends on $\cos\theta$ and takes the maximum negative value if $\theta = \pi$.

Therefore, the maximum reduction in Ψ is:

$$\delta\mathbf{x} = \lambda \left(\frac{-\mathbf{g}}{\|\mathbf{g}\|} \right)$$

Figure 6. Method of steepest descent.

just the dot product of the gradient vector and the adjustment vector. The correction term takes on its maximum negative value when the angle between the two vectors is π . So, essentially, the optimum value of the adjustment vector is given by the minus value of the gradient vector divided by the norm of the gradient vector. The λ multiplier is used to control the step size to reduce the chance of overshooting the minimum.

Newton's method uses both the first and second order terms. Figure 7

shows the first and second order correction terms written out in summation form. In order to find the adjustment vector which produces the greatest movement toward the minimum, we differentiate the expression for the correction terms with respect to δx_k , and set the resulting expression equal to zero. That gives the expression shown at the bottom of Figure 7. In this method, we need to be able to determine the inverse of the Hessian matrix (H^{-1}). Then, by multiplying H^{-1} by the gradient vector (g), we get the optimum value for the adjustment vector. Although there are some practical difficulties in implementing this procedure, the mathematics is relatively straight-forward.

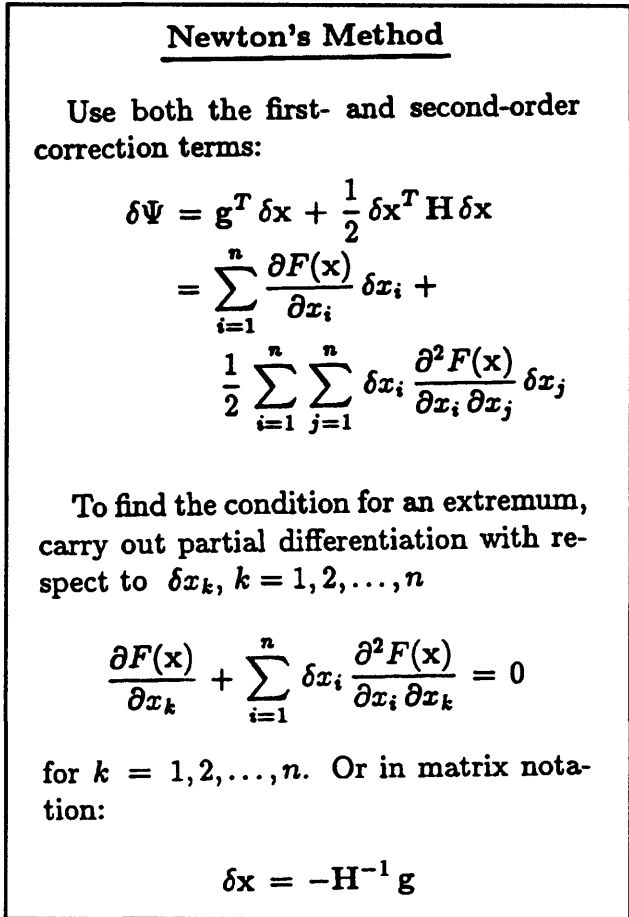


Figure 7. Newton's method.

NONLINEAR LEAST SQUARES

Usually, there are more observations than unknowns. In this case, we say that the system is "over-determined." Rather than just use a subset of the observations to solve the problem, the traditional approach has been to use the method of non-linear least squares.

The first step in this method is to define a residual vector. This is simply a vector containing the differences between the observed values and the predicted values of the travel times for all stations. For some

reason, Geiger reversed the sign of this quantity, and the practice has become entrenched in the seismological community. Of course, it makes no practical difference in the solution of the problem, but if you have learned about least squares from an applied mathematics outlook, the sign convention may look strange to you.

In seismology the theoretical times we use in computing the residual are computed using some model of the earth's velocity structure. Even if we can make our data fit the model perfectly, that still does not mean that we have found the exact location of the

Nonlinear Least Squares (1)

Residual vector:

$$\mathbf{r} = \begin{pmatrix} r_1(\mathbf{x}) \\ r_2(\mathbf{x}) \\ \vdots \\ r_m(\mathbf{x}) \end{pmatrix}$$

Objective function:

$$F(\mathbf{x}) = \sum_{k=1}^m [r_k(\mathbf{x})]^2 = \mathbf{r}^T \mathbf{r}$$

Gradient vector:

$$\frac{\partial F(\mathbf{x})}{\partial x_i} = \sum_{k=1}^m 2 r_k \frac{\partial r_k}{\partial x_i}$$

for $i = 1, 2, \dots, n$.

Figure 8. Nonlinear least squares (1).

earthquake in the real earth. It only means that we have located the earthquake with respect to our model of the earth. This may or may not be close to the true hypocenter of the earthquake.

The objective function is essentially the sum of the squared residuals, and the task is to minimize the objective function. This is exactly analogous to the two-dimensional problem of fitting a straight line to a set of data points by minimizing the L_2 norm.

As shown in Figure 9, the residual associated with each data point for some trial

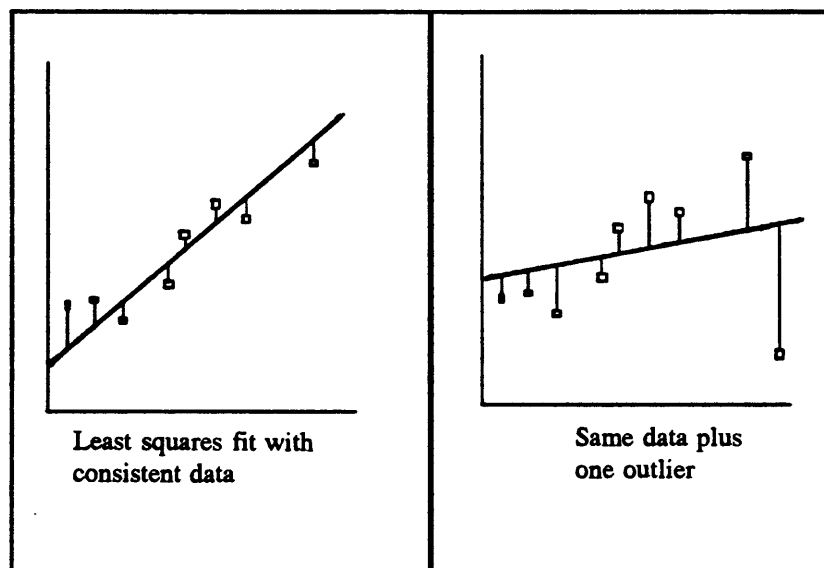


Figure 9. Least squares fit of straight line to data.

straight line is just the distance between the line and the data point. A measure of the misfit between the line and the data is the sum of the squared residuals. This measure works well in fitting the line to the data, as long as the data fit the model reasonably well. However, you can get into trouble if the data contains an outlier point, as in the right hand part of Figure 9. In that case, the least squares fit of the line to the data results in a line which fits most of the

data poorly. The outlier point "contaminated" the solution. The least squares process gave an unusually high weight to the "bad" data point, and this skewed the fit towards that point. A better measure of misfit for data which contains a few outliers is the L_1 norm. Unfortunately, minimization using the L_1 norm is not a straight-forward procedure, so least squares is still commonly used.

As shown in Figure 8, the gradient vector is calculated simply by taking partial derivatives of the objective function with respect to the x_i 's and summing. Calculation of the Hessian matrix is a little bit messier, but still, straight-forward (See Figure 10). To simplify notation we define the "Jacobian" matrix (A), which is a matrix of partial derivatives with the ij th element equal to the partial of the i th residual with respect to the j th model parameter. In terms of the Jacobian matrix, the gradient vector is

defined as: $g = 2A^T r$. Using this notation, the Hessian matrix can be approximately expressed as: $H \approx 2A^T A$. This approximation is valid as long as the residuals (r_k) are, on average, quite small. In this case, the second term of the defining equation for the Hessian

<u>Nonlinear Least Squares (2)</u>	
Hessian matrix:	
$\frac{\partial^2 F(x)}{\partial x_i \partial x_j} = 2 \sum_{k=1}^m \frac{\partial r_k}{\partial x_i} \frac{\partial r_k}{\partial x_j} + 2 \sum_{k=1}^m r_k \frac{\partial^2 r_k}{\partial x_i \partial x_j}$	
for $i = 1, 2, \dots, n$, and $j = 1, 2, \dots, n$.	
Define Jacobian matrix A as:	
$A \equiv \begin{pmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \dots & \frac{\partial r_1}{\partial x_n} \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \dots & \frac{\partial r_2}{\partial x_n} \\ & & \vdots & \\ \frac{\partial r_m}{\partial x_1} & \frac{\partial r_m}{\partial x_2} & \dots & \frac{\partial r_m}{\partial x_n} \end{pmatrix}$	
Gradient vector: $g = 2A^T r$	
Hessian matrix: $H \approx 2A^T A$	
Adjustment vector: $\delta x = -[A^T A]^{-1} A^T r$	

Figure 10. Nonlinear least squares (2).

matrix becomes much smaller than the first term and can be neglected (See first equation in Figure 10). This condition will be fulfilled as long as the model fits the data reasonably well, and there are no seriously bad observations. Otherwise, the approximation is invalid and the method breaks down. In the notation just introduced and with this approximation to the Hessian matrix, the adjustment vector is expressed as: $\delta x = -[A^T A]^{-1} A^T r$. This approach to solving for a non-linear problem was developed over 200 years ago, and is still in common use today. It works well as long as you abide by the restriction that the residuals should be small. Also, you have to have some way of calculating all the derivatives, or you can't solve the problem.

SOLVING THE EARTHQUAKE LOCATION PROBLEM

There are some simple methods of locating earthquakes. In some situations, it is possible to get an approximate location of the earthquake without resorting to a computer. If you have a dense seismic network, you can approximate the earthquake epicenter as the station with the earliest arrival. As long as the earthquake originated inside your network, this will give you an estimate with a probable error of the same order as the station spacing.

Another approach is to contour the P-wave arrival times. This is an extension of the earliest station method and if your network is dense enough, this may give a fairly good estimate of the epicenter location. Even if the earthquake occurs outside the network, the curvature of the contours should give you a rough idea of where the earthquake originated.

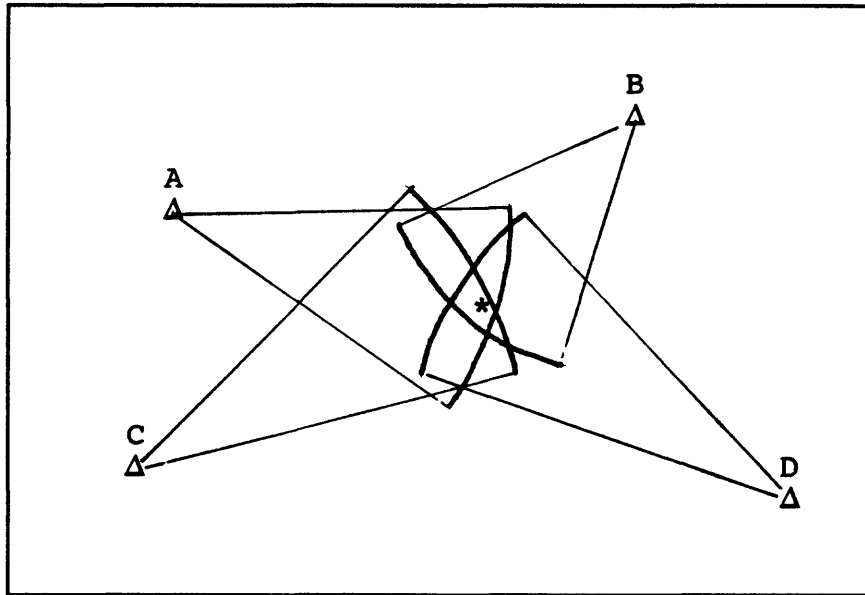


Figure 11. Locating epicenter using S-P times.

If you just have a few stations, you can't use either of the above methods effectively. But as long as you have S-P times, you can still get a pretty good approximation of the epicenter. Use the formula: $\Delta \approx 7.5 (S-P)$. Here Δ is in kilometers and S-P is in seconds. The 7.5 is the result of using some average values for P-wave and S-wave velocities, so if you know the seismic velocities in your area well, you might modify that formula slightly. After determining the epicentral distance using the above formula for each station in your network, just swing arcs on a map with each arc centered on the station and with epicentral distance equal to the arc length. As shown in Figure 11, the epicenter of the earthquake lies in the region where the arcs overlap. This method was commonly used in the past as a quick means of locating earthquakes.

The earthquake location problem can also be solved using the least squares method. Then, the problem is expressed in the following manner. You have a set of m stations with

coordinates (x_k, y_k, z_k) , $k=1,2,\dots, m$. You also have a set of arrival times T_k , $k=1,2,\dots, m$. The problem is to find the "best" origin time (T_0), and hypocenter (x_0, y_0, z_0) . We will assume some velocity model, usually a set of layers over a half-space. Of course, given this assumption, the solution can only be used for local earthquakes, since at distances of over a few hundred kilometers, the sphericity of the earth becomes a significant factor. Also, the assumption that the earth can be adequately modeled by a stack of constant velocity layers may not always lead to acceptable results. In some regions, lateral heterogeneities in the velocity structure can become quite significant, and a simple layered model will give inadequate results. However, we tend to use these simple models whenever we can, since they reduce the computational complexity significantly.

We express the model vector that we are solving for as:

$$\chi = (t, x, y, z)^T \quad (4)$$

To begin the solution process, we assume a trial vector:

$$\chi^* = (t^*, x^*, y^*, z^*)^T \quad (5)$$

We must be able to calculate the theoretical travel time from a trial hypocenter to a station using our velocity model. Then the theoretical arrival time at the k th station is:

$$t_k(\chi^*) = T_k(\chi^*) + t^* \quad (6)$$

where t^* is the trial origin time. Now, we calculate the residual at the k th station as:

$$\begin{aligned} r_k(\chi^*) &= \tau_k - t_k(\chi^*) \\ &= \tau_k - T_k(\chi^*) - t^* \quad \text{for } k = 1, 2, \dots, m \end{aligned} \quad (7)$$

As I mentioned earlier, this definition of the residual follows the convention established by

Geiger, and is opposite in sign to the convention used in most of mathematics. Following the method outlined in the last section on nonlinear least squares, we define the objective function as:

$$F(\chi^*) = \sum_{k=1}^m [r_k(\chi^*)]^2 \quad (8)$$

As before, we define the residual vector:

$$\mathbf{r} = [r_1(\chi^*), r_2(\chi^*), \dots, r_m(\chi^*)]^T \quad (9)$$

The adjustment vector becomes:

$$\delta\chi = (\delta t, \delta x, \delta y, \delta z)^T \quad (10)$$

So, at each iteration, we must solve the equation:

$$\mathbf{A}^T \mathbf{A} \delta\chi = -\mathbf{A}^T \mathbf{r} \quad (11)$$

Here \mathbf{A} is the Jacobian matrix and it has the value:

$$\mathbf{A} = \begin{pmatrix} \frac{\partial r_1}{\partial t} & \frac{\partial r_1}{\partial x} & \frac{\partial r_1}{\partial y} & \frac{\partial r_1}{\partial z} \\ \frac{\partial r_2}{\partial t} & \frac{\partial r_2}{\partial x} & \frac{\partial r_2}{\partial y} & \frac{\partial r_2}{\partial z} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial r_m}{\partial t} & \frac{\partial r_m}{\partial x} & \frac{\partial r_m}{\partial y} & \frac{\partial r_m}{\partial z} \end{pmatrix} \quad (12)$$

Using the arrival time residuals defined in Equation (7) and evaluating the partial derivatives at χ^* , the Jacobian matrix becomes:

$$\mathbf{A} = - \begin{pmatrix} 1 & \frac{\partial T_1}{\partial x} & \frac{\partial T_1}{\partial y} & \frac{\partial T_1}{\partial z} \\ 1 & \frac{\partial T_2}{\partial x} & \frac{\partial T_2}{\partial y} & \frac{\partial T_2}{\partial z} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{\partial T_m}{\partial x} & \frac{\partial T_m}{\partial y} & \frac{\partial T_m}{\partial z} \end{pmatrix} \chi^* \quad (13)$$

Substituting Equation (13) into Equation (10) and carrying out the matrix operations results in a set of four simultaneous linear equations in four unknowns:

$$\mathbf{G} \delta \chi = \rho \quad (14)$$

where

$$\mathbf{G} = \begin{pmatrix} m & \sum a_i & \sum b_i & \sum c_i \\ \sum a_i & \sum a_i^2 & \sum a_i b_i & \sum a_i c_i \\ \sum b_i & \sum a_i b_i & \sum b_i^2 & \sum b_i c_i \\ \sum c_i & \sum a_i c_i & \sum b_i c_i & \sum c_i^2 \end{pmatrix} \quad (15)$$

and

$$\rho = \left(\sum r_k, \sum a_k r_k, \sum b_k r_k, \sum c_k r_k \right)^T \quad (16)$$

In Equations (15) and (16) the summations are to be taken over $k=1,2,\dots,m$. Also the terms a_k , b_k , and c_k are defined as:

$$a_k = \left. \frac{\partial T_k}{\partial x} \right|_{\chi^*}, \quad b_k = \left. \frac{\partial T_k}{\partial y} \right|_{\chi^*}, \quad c_k = \left. \frac{\partial T_k}{\partial z} \right|_{\chi^*} \quad (17)$$

Solution of Equation (14) gives the adjustment vector, which when added to the trial parameter vector, gives the updated estimate of the solution vector. This process of updating the solution vector continues until the change in the solution is less than some specified

amount, or until a certain number of iterations have been completed.

Table 1 Summary of Geiger's method.

Geiger's Method for Hypocenter Location

- 1) Guess a trial origin time t^* , and a trial hypocenter (x^*, y^*, z^*) .
 - 2) Compute the theoretical travel time and its spatial partial derivatives evaluated at (x^*, y^*, z^*) from the trial hypocenter to the k th station for all k .
 - 3) Compute matrix G and vector p
 - 4) Solve the system of four linear equations for the adjustments δt , δx , δy , δz .
 - 5) An improved origin time is then given by $t^* + \delta$, and an improved hypocenter by $(x^* + \delta x, y^* + \delta y, z^* + \delta z)$.
 - 6) Repeat steps 2 to 5 until some cutoff criteria are met. At that point, set $t_0 = t^*$, $x_0 = x^*$, $y_0 = y^*$, and $z_0 = z^*$ as the solution to the problem.
-

The solution method just outlined is often referred to as the "normal equations" approach. It works well if the data constrain the solution well, and if the data contain few, if any, bad errors. Otherwise, inversion of the matrix G may become unstable and the method will fail. There are a number of alternative approaches which are more computationally stable, but they will not be discussed here. However, there are many good references which treat the subject in great detail. Table 1 summarizes Geiger's method for hypocenter location.

Geiger's method is sensitive to the initial guess you use. If it is too far away from the global minimum, then the method may not converge to the global minimum. You can see why this is so by examining the case of one dimensional minimization. Figure 12 shows a function which has several minima within the range of interest. If you start the search for the minimum near x_1 , the adjustments will move the trial parameter vector downhill until the point (x_1, y_1) is reached, and that will be returned as the answer. The true global minimum at (x_2, y_2) will never be found. The situation is more complicated in four dimensions but still

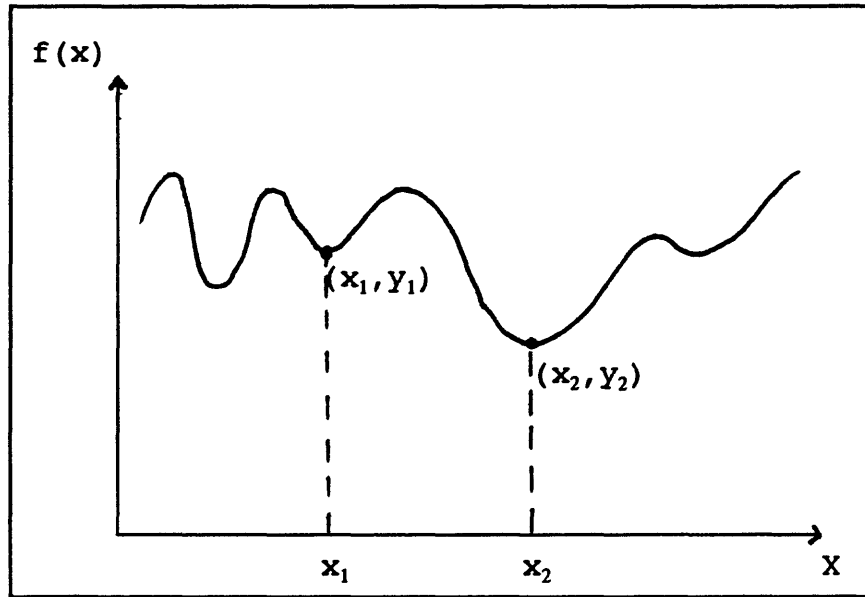


Figure 12. Function with several minima.

analogous. This tendency to get trapped in local minima is a common failure of nearly all non-linear optimization schemes, with the possible exception of the "simulated annealing" algorithms and grid search algorithms.

Even if you have a lot of observations, it is not always possible to get well constrained hypocenter locations. This is not really a failing of the optimization algorithm. Rather, it is due to an unfortunate geometry in the problem. If the earthquake occurs far outside the network it may be essentially un-locatable. Even if it occurs within the network, if you do not have a station closer to the epicenter than about one focal depth, you will not be able to constrain the focal depth very well. Sometimes, you may know in advance that this is the case. One approach to this difficulty is to reformulate the problem so that the focal depth is not one of the parameters to be solved for. Instead you may use some prior knowledge about earthquakes in the region you are studying to fix the focal depth. In

HYPO71 the program will determine if there is inadequate control over focal depth, and if so will automatically fix the focal depth in the solution it calculates. HYPOELLIPSE, HYPOINVERSE, and many other programs will always solve for the focal depth unless you instruct them not to.

SUGGESTED READING

Geiger, L. (1912). Probability method for the determination of earthquake epicenters from the arrival time only. Bull. St. Louis Univ., 8, 60-71.

Gill, P.E., W. Murray, and M.H. Wright (1981). "Practical Optimization". Academic Press, London, 401 pp.

Klein, F.W. (1978). Hypocenter location program--HYPOINVERSE, U.S. Geol. Surv. Open-File Report 78-694.

Lahr, J.C. (1979). HYPOELLIPSE: A computer program for determining local earthquake hypocentral parameters, magnitude, and first motion pattern. U.S. Geol. Surv. Open-File Report 79-431.

Lanczos, C. (1961). "Linear Differential Operators", Van Nostrand, London, 564 pp.

Lee, W.H.K., and J.C. Lahr (1975). HYPO71 (revised): A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes. U.S. Geol. Surv. Open-File Report 75-311, 116 pp.

Lee, W. H. K., and S. W. Stewart, (1981). "Principles and Applications of Microearthquake Networks." Academic Press, New York, 293 pp.

11. LOCAL EARTHQUAKE LOCATION PROGRAMS

by

J. C. Lahr

U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

An earthquake location program is designed to solve the following problem (See Figure 1). Given a set of arrival times from an earthquake, find the origin time and location of the earthquake and evaluate the limits of the error of the location. A secondary purpose of the program is to give feedback to the seismologist to aid in isolating and correcting reading errors. The program must be able to

summarize, for a set of locations, the overall quality of the solutions and the average residuals for each station. Finally, the program must be able to compute the magnitude from either amplitude or coda duration measurements.

The hypocenter location produced by the program is the one which "best" matches the observed arrival times with the times predicted by the earth model. In this sense, the hypocenter location produced by the program is not the same as the true hypocenter if the

Fundamental problem to be solved:

Given a set of arrival times from an earthquake, find the "best" origin time and location.

Secondary goals:

- o Give feedback to seismologist to aid in isolating and correcting reading errors.
- o Evaluate the limits of error of the "best" location.
- o Summarize for a set of locations the overall quality of the solutions and the average residuals for each station.
- o Compute the magnitude from either amplitude or coda duration measurements.

Figure 1 Fundamental problem to be solved.

velocity structure of the earth model differs significantly from the true velocity structure of the earth. The question arises, "What is meant by the "best" solution?" This is the solution that minimizes the root mean square (RMS) residual (Figure 2). That is, if we define R_i as the difference between the i th observation and the i th predicted arrival time, and if we define RMS as:

$$RMS = \sqrt{\frac{\sum_{i=1}^N R_i^2}{N-1}} \quad (1)$$

then the "best" solution is the one which minimizes RMS.

There are several ways in which the minimum RMS can be found (Figure 3). The simplest method might be to perform an exhaustive search. Of course, this implies that the solution space must be discretized in some manner to reduce the number of points to be searched to a finite value. One could grid the volume of space which is thought to enclose the hypocenter, and then evaluate RMS at every point of a grid within that space. The point which has the minimum value of RMS would then be returned as the solution. This method is relatively sure-fire as long as the volume of space does, in fact, enclose the solution, and as long as the grid spacing is fine enough for the required solution accuracy. The method is time consuming, however. If there are M grid points per dimension, then there must be M^3

The "BEST" solution has the lowest RMS

Given a proposed earthquake location and origin time, and a proposed earth velocity model, the arrival times to each station may be computed.

The *residual* of any phase is the difference between the observed and the computed arrival time.

Residual = Observed minus Computed time

Location programs attempt to minimize the Root Mean Square Residual (RMS) - the square root of the mean square residual.

$$RMS = \sqrt{\frac{\sum_{i=1}^N R_i^2}{N-1}}$$

Figure 2 The "best" solution has the lowest RMS.

evaluations of RMS to find the solution. However, as computers become faster, this "brute force" method or a combination of it with the method described next, becomes more feasible.

A much faster method is to start at a trial hypocenter, linearize the equations, and compute an improved solution. The improved solution

becomes the next trial hypocenter and the process is repeated until changes between trial hypocenters are small. The principle problem with this method is distinguishing a local minimum from a global minimum. That is, for many earthquake data sets, the RMS does not increase monotonically away from the global minimum. Rather, the RMS has one or more local minima in addition to the global minimum which we seek. If the first trial hypocenter location is within the "valley" of a local minimum, the solution will converge to that local minimum rather than the global minimum.

How is the minimum RMS found?

- o Exhaustive evaluation of RMS at all possible locations - there are algorithms that use this approach and as computers speed up this becomes more viable.
- o Start at a trial hypocenter, linearize the equations and compute an improved solution. Repeat process using new location as next trial location. Stop when changes are small.

The principal problem with this approach is distinguishing a local minimum from the very smallest (global) minimum.

Figure 3 How is the minimum RMS found?

LIMITATIONS ON ACCURACY

Apart from possibly converging to the wrong minimum, what limits earthquake location accuracy? There is always a certain amount of uncertainty in the observed arrival times and in the earth model used to compute the travel times. How these errors are mapped

into spatial location errors depends on the geometric relationship of the stations to the hypocenter and on the phases that are used.

Unless an earthquake is shallow and within a dense network, it is essential to have S-phase readings in addition to P-phase readings. The combination of P- and S-phases gives greater constraint on the origin time, and limits what could otherwise be an unbounded trade off between origin time and location.

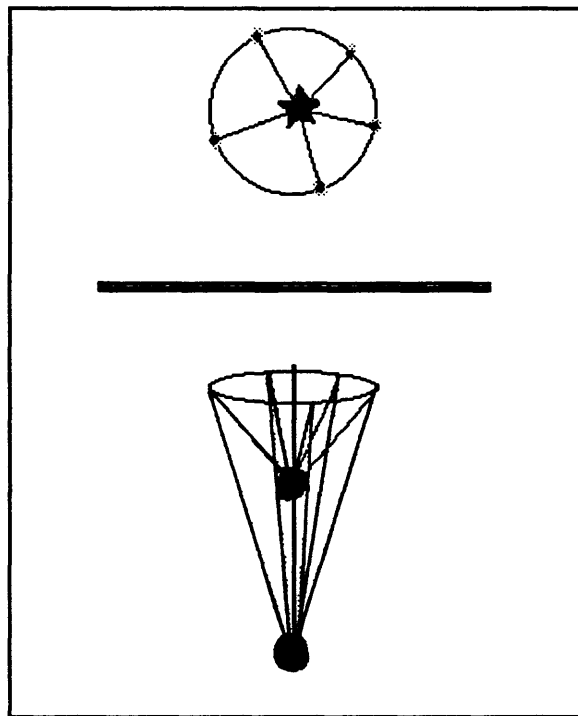


Figure 4 Trade-off between depth and origin time.

Consider an earthquake at the center of a circular array of stations (Figure 4). The P-phase will arrive simultaneously at all of the stations. This fixes the earthquake epicenter but places no constraint on depth. You can see why this is so by looking at a cross section of the problem. There is a trade-off between hypocenter depth and origin time. For any depth, all

the P- arrivals will arrive simultaneously. As depth is increased, the only effect is that the origin time must be earlier. Since there is no information about origin time, there is no control over depth. Just one S-phase reading, would constrain the path length between the hypocenter and that station. This would serve to fix the origin time and hence the depth.

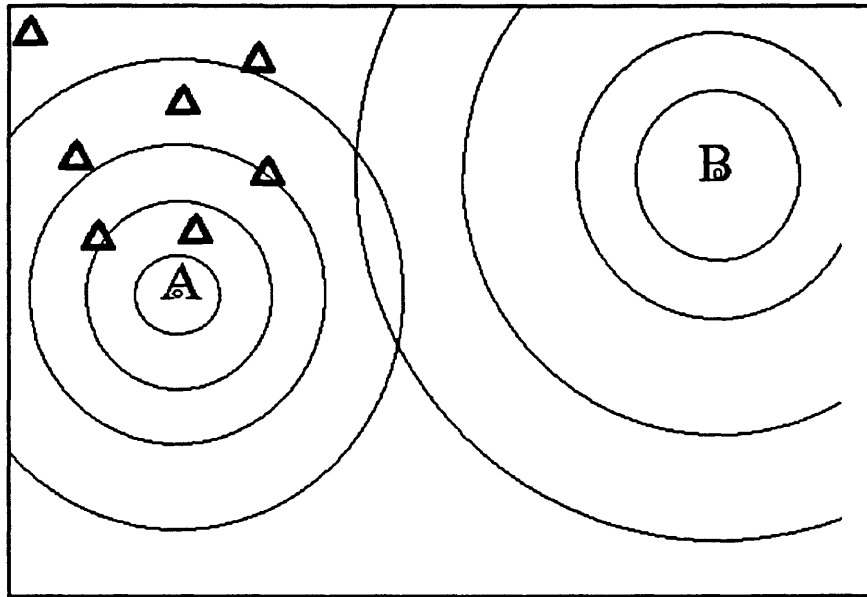


Figure 5 Wave front curvature variation with epicentral distance.

If an earthquake occurs outside the network and only P-wave arrivals are available, the information that is used to get the distance to the epicenter is the curvature of the P-waves as they pass through the net. If the epicenter is very close to the net, as for the source at A in Figure 5, then the curvature will be significant. Source B is further away and the curvature (as measured within the net) is less than for source A. If the source is farther away than about ten times the diameter of the net, then the P-waves will be essentially plane waves as they pass through the net. So, for all sources which are very far outside the net, the curvature is essentially the same. Therefore, as events get farther outside a network, S-

phases become more important to constrain the distance to the source.

If there are S-readings as well as P-readings, the S-readings have the effect of constraining the origin time. A simple case to consider is that of a halfspace with uniform velocity. If both the S-wave and the P-wave arrival times at a station are known, then the distance to the source from that station can be computed. That means you can draw a hemisphere centered on that station, of such a radius that the source must lie somewhere on the hemisphere. If this is done for 3 or more stations, the hypocenter will be fairly well defined by the intersections of the hemispheres. There still can be errors caused by reading errors, and errors in the velocity model, but the constraints on the hypocenter location will be much more stringent than if there were no S-wave readings.

The depth of hypocenters inside a network may be difficult to constrain, even apart from those near the center of a circular array. The problem is analogous to that of constraining the distance from the network to a distant earthquake. If the earthquake is shallow, relative to the station spacing, then the curvature is significant and the depth is well constrained. However, for deeper earthquakes it becomes progressively more difficult to constrain the hypocenter depth without S-wave readings.

Even with the additional constraints provided by S-wave readings, earthquakes become progressively harder to locate as they occur farther away from the net. A simple string experiment illustrated in Figure 6 shows the problem in two dimensions. Attach the ends of two strings that are about 2 m long at two separate locations about 0.5 meter apart. The points of attachment represent seismic stations. Slide a small paper tube about 5 cm long onto each string. The 5 cm represents the uncertainty in distance to each station due to

reading and model errors. With the tubes near the end of the strings align the strings so that the tubes overlap. Move the strings so that different portions of the tubes overlap, and note the spatial limits of the crossover points. Next, slide the tubes up much nearer the "stations," repeat the overlap process, and note the spatial limits of the crossover points. As shown in Figure 6, the crossover area defined when the tubes are close to the "stations" is smaller than the crossover area defined when the tubes are far away from the "stations." This crossover area represents the uncertainty in the solution due to

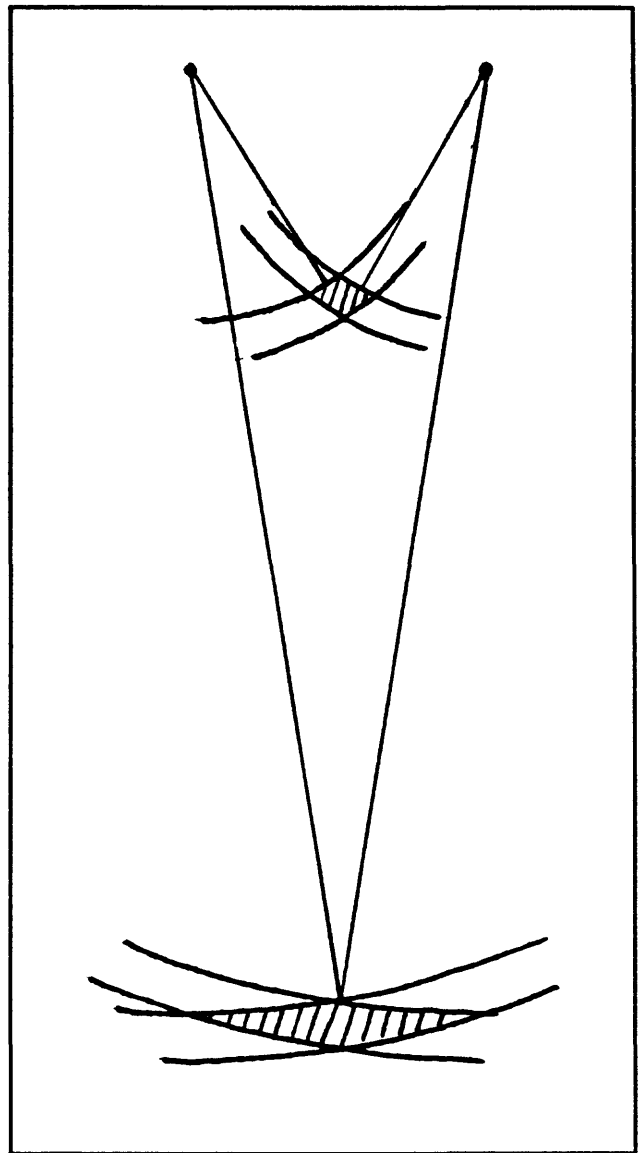


Figure 6 String experiment.

errors, and this experiment demonstrates that the uncertainty increases with distance from the network.

The previous observations lead to a rule of thumb you should keep in mind when looking at your own or anyone else's earthquake data. Always be suspicious of any

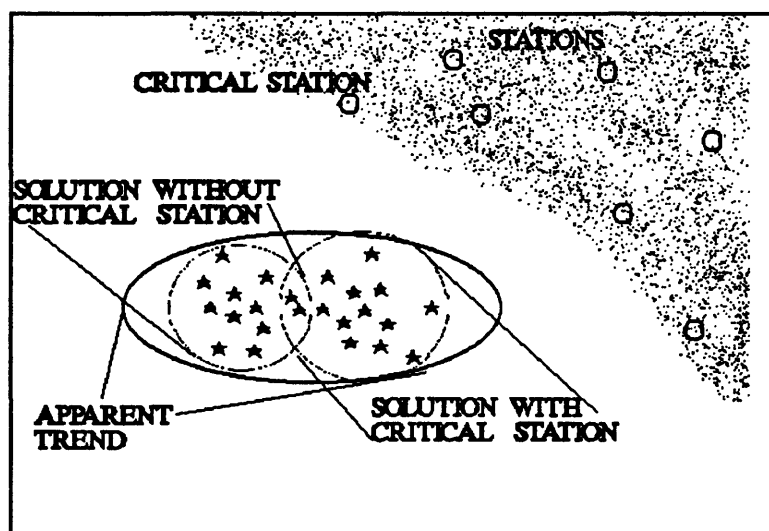


Figure 7 Station subset problem.

distribution of earthquakes, and not accept it at face value until thoroughly testing for problems which can arise while locating those earthquakes. There can be problems with multiple minima, and also problems with variations in the station subset used. If you have a network with lots of stations and you locate a group of earthquakes, that group may often have a spatial distribution with a distinct trend. This can lead you to believe that the distribution is due to some geological control such as a fault.

If the stations used in each location are reviewed, however, it may be found that one or more of the stations was used for only a fraction of the earthquakes. It may be useful then to plot all the earthquakes for which a given station was used, and then plot all the earthquakes for which the station was not used. The two data sets may produce two distinct clusters which, when combined, give the trend seen in the complete data set. What you have then, is not a trend. Rather, it is a more localized distribution whose computed location depends on the stations used.

This simple variation in station subset can actually become quite complex. For a

group of earthquakes which occur in some region, there will probably be some subset of the net which always records the events and another, quite variable, subset which may or may not record the events. The resulting complex variations in station subsets over time may combine in various ways to produce false trends in the earthquake distribution. To avoid drawing false conclusions about these trends, it is useful to calculate the spatial errors associated with each solution. If the orientation of the errors is parallel to the trend in the data, there may be a localized source, rather than a linear source.

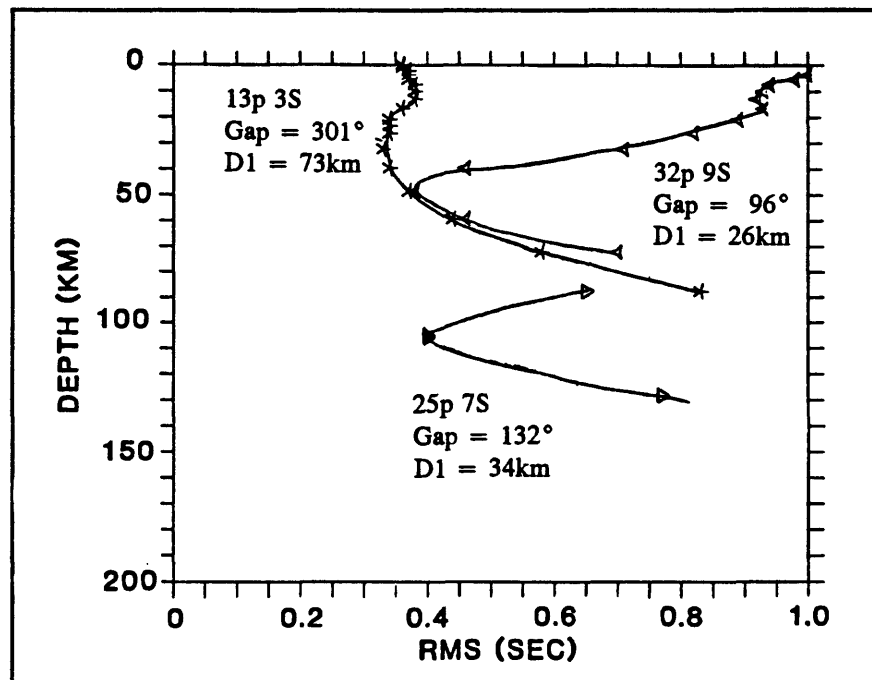


Figure 8 Multiple minima problem.

Figure 8 illustrates a multiple minima problem from a data set collected in Alaska. In the figure, the vertical axis is depth, increasing downward and the horizontal axis is root mean square residual. In some cases there is no problem. The bottom of the three curves shows a smooth decrease with depth to the minimum RMS and a smooth increase as depth

increases past the minimum point. Not only is there just one minimum, but the minimum is very narrow. This is a well constrained solution.

The curve marked "32P 9S" presents more of a problem. There is a minimum at a depth of about 13 kilometers followed by a small peak, and then another minimum at a depth of about 50 kilometers. In this case, if your location program starts with a trial location near the surface, it will converge at the shallow minimum, even though the minimum at 50 kilometers is actually the deepest or "global" minimum. The error bars will indicate a well constrained minimum, and only the RMS will indicate that there might be a problem with the solution. On the other hand, if you started the solution at a depth below 20 kilometers, the algorithm would converge at the global minimum of 50 kilometers. The RMS returned would be much lower than for the shallow solution, but it is possible that the error bars could be wider if the slope near the global minimum is less than the slope near the shallow minimum. The third case shown in Figure 8 has a minimum at the surface and another at a fairly shallow depth. The error bars returned for the 20-kilometer depth solution will be rather large since the slope is so small around that solution. One might note that the minimum near 20 kilometers is less well constrained than the minimum at the surface, and choose the surface focus as the best solution. This is a mistake. Always choose the solution with the lowest RMS.

Another problem with multiple minima that we sometimes have in our Alaska network occurs when the only stations available form an approximately linear array. In this case, there will be two minima which satisfy the data, as shown in Figure 9. The solution with the lower RMS should be selected.

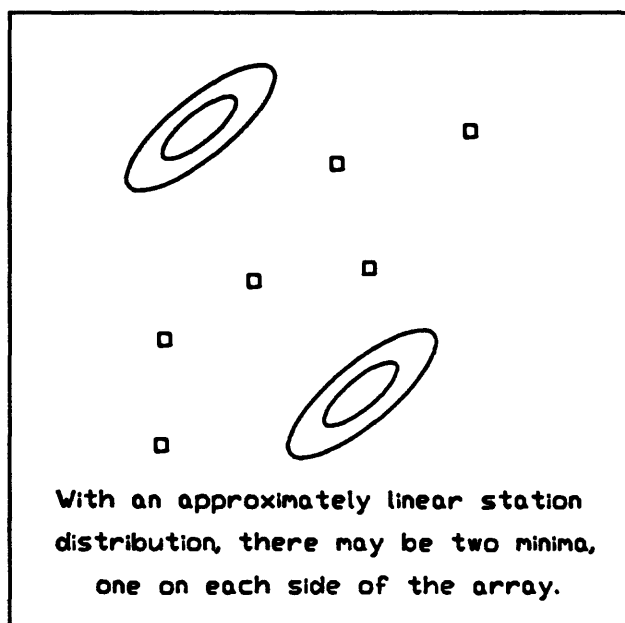


Figure 9 Multiple minima for linear station distribution.

Figure 10 shows a set of aftershocks for the 1984 Sutton, Alaska earthquake. The top panels show map views, and the bottom panels show cross sections through the earth. The mainshock epicenter was located north of the mapped fault zone, and we were trying to decide if it had actually occurred on a buried portion of the fault. In the cross section view in the left-hand lower panel, you can see that the fault zone does intersect the cluster of computed aftershock hypocenters. The question remained whether the apparent trend in the hypocenters was due to the actual distribution of shocks up and down the fault zone, or whether the trend was due to lack of control of the solutions which just happened to parallel the fault zone. The center two panels illustrate a "worst case" scenario in which the errors line up with the fault both in the strike direction and in the dip direction. If this were the case, we would not have been able to conclude that the aftershocks were distributed along the

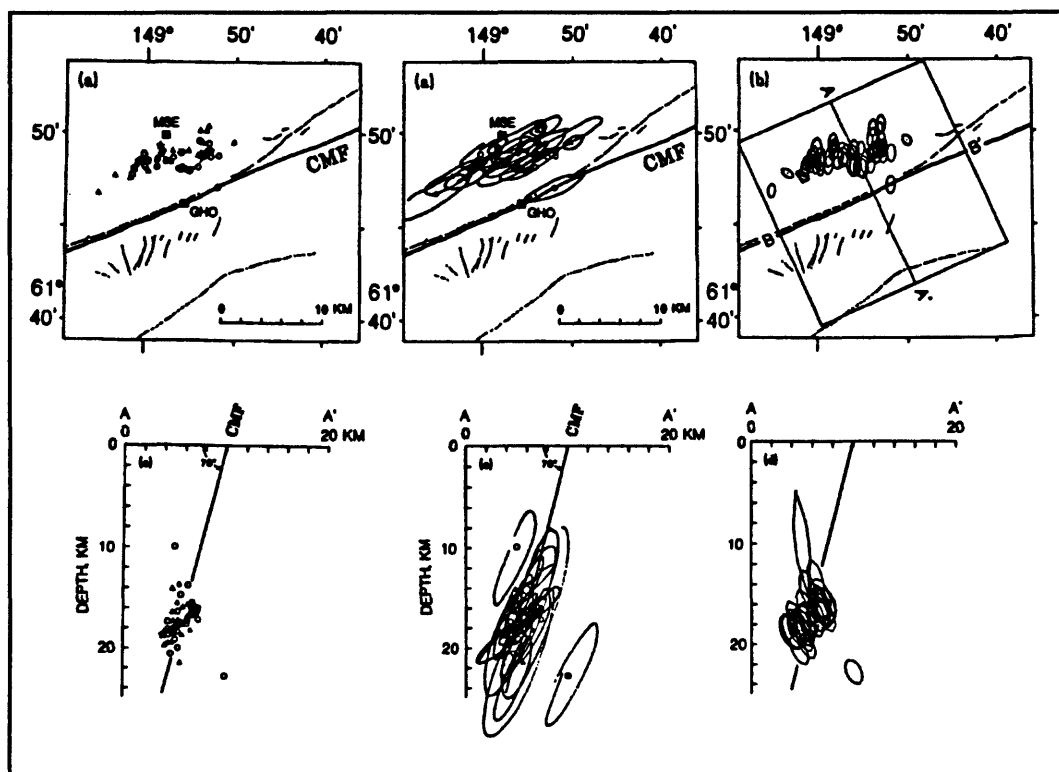


Figure 10 1984 Sutton, AK aftershocks.

fault. Luckily, that was not the case. The right-hand panels show the actual error bars associated with the solutions. The errors are roughly perpendicular both to the strike direction of the fault zone and to the dip of the fault. Therefore we concluded that the distribution along the trace of the fault in both map and cross section views reflected an actual trend in the earthquake hypocenter locations.

One approach to take in finding a true pattern in a set of earthquakes is called the "master event method" (Figure 11). Start with one earthquake that is exceptionally well located and has readings for every possible phase. It could be a main shock with very clear arrivals. Next adjust the station delays so that earthquake has zero RMS. Then, another earthquake with nearly the same pattern of arrivals will locate near the master event, even if

a different subset of stations is used.

Another approach is called the "average residual method" (Figure 11).

Here you compute the average residual at each station for the entire set of earthquakes. Then use the averages as station delays and recompute the locations. The process can be iterated a few times. This will reduce the relative errors for members of this set of

earthquakes. Neither of these methods can improve the absolute error in source location for earthquakes in the group, but they reduce the relative location errors so that a pattern that wouldn't have been apparent otherwise becomes clear.

I mentioned before that there could be scatter introduced into the solutions because of different combinations of stations being used in the solutions. An approach that is often useful in situations like this, is to find a subset of phase readings that were used for locating most of the events. Then relocate just those events that

Some approaches to take in looking for a true pattern in a set of earthquakes include:

Master Event Method

Derive station delays so that one event, the *master*, has zero RMS residual. Then locations of nearby events with respect to the master will be more precise.

Average Residual Method

Use the average residuals from a set of earthquakes as the station delays for these events. This can be iteratively repeated.

Figure 11 Some approaches in looking for earthquake patterns.

Common Phases Method

Scatter can be due to each event being located by a different combination of P and S readings.

Find a subset of readings that most of the events have in common. Relocate just the events that have all of these readings, using no other readings. This will eliminate one source of scatter and may reveal a clearer picture.

Figure 12 Common phases method.

have arrival times for all of these phase readings, using just these phase readings. Exclude all other events and phases. This can be a very effective means of eliminating scatter and may reveal a clearer picture of the true hypocentral distribution. The method is usually referred to as the "common phases method" (Figure 12).

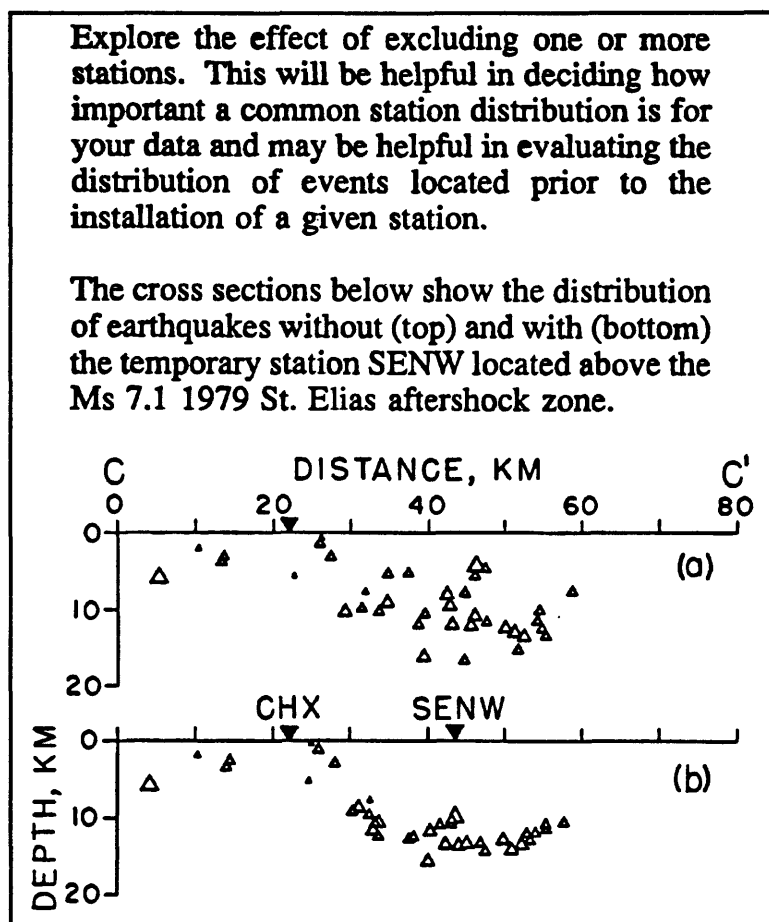


Figure 13 Effect of excluding stations.

One may be interested in finding the importance of a particular station to the stability of a set of hypocenter locations. Figure 13 shows two cross-section plots of some Alaskan aftershock data. The upper cross section shows a very scattered distribution of hypocenters. For a short period of time we added an extra station directly above the center of the cluster

of earthquakes. That station, SENW, is shown on the lower cross section. During the time that SENW was operating, the earthquakes we located had the pattern shown in the lower cross section. The added station greatly reduced the scatter in depth for events below the station.

EARTHQUAKE LOCATION PROGRAMS

There are many good location programs. I will discuss HYPO71 and HYPOELLIPSE since those are the ones I am most familiar with. The basic inputs to HYPO71 (Figure 14) are the velocity model, station coordinates, elevations, calibration information, control parameters, and, of course, the arrival times.

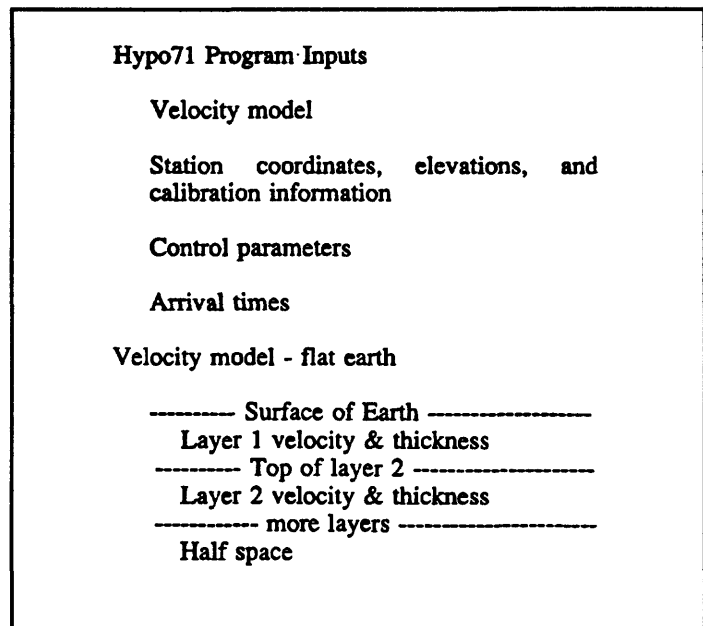


Figure 14 HYPO71 program inputs (1).

The velocity model that is used in HYPO71 consists of a series of layers over a half-space. The depth of the top and the velocity are specified for each layer and for the halfspace.

HYPO71 Program

The station list parameters used by HYPO71 include the station name, coordinates of the station, station elevation, P-delays, calibration parameters, and Magnitude corrections.

The station names used by HYPO71 may be up to 4 characters long. Elevations are not used by the program even though they are on the parameter list. The P-delay is specified on the station list and the S-delay is calculated from the P-delay as:

$$S\text{-delay} = P\text{-delay} \left(\frac{V_p}{V_s} \right)$$

The delay is added to the computed travel time and may be specified for two models. A

positive residual indicates that the observed time is larger than the computed time, and would be reduced by a positive delay.

The control parameters for HYPO71 include the first trial depth, the distance weighting limits and the V_p/V_s ratio. One might think that all the available data should be used, but for a large network such as Calnet, the best solution is often obtained by using only stations close to the earthquake. This is because errors in the computed travel time due to model errors will increase with distance. A small percentage in velocity, after a few hundred kilometers, can yield a substantial error in computed travel time. The purpose of distance weighting is to give greater weight to the close stations. The way it works is that for stations

Station List Parameters
Station Name (4 characters)
Coordinates
Elevation (not used in Hypo71)
P-delay (S-delay is $P\text{-delay} \cdot V_p/V_s$) Delay is <i>added</i> to computed travel time. May be specified for two models.
Calibration parameters
Magnitude corrections
Control Parameters
First trial depth
Distance weighting distance limits
V_p/V_s ratio

Figure 15 HYPO71 program inputs (2).

less than the lower distance limit (XNEAR), the weight given to the data is one. Beyond XNEAR the weight decreases linearly to zero at the outer limit (XFAR).

Other command switches recognized by HYPO71 include the check for missing data switch, the minimum number of first motions required for a focal mechanism plot, azimuthal weighting, first trial coordinates, and others listed in Figure 16.

Check for missing data switch
Min # of first motions for plot
Summary record output control
Printed output control - how much?
Print RMS computed at 10 nearby points?
Apply azimuthal weighting?
Sort stations by distance?
Start each EQ on new page?
First trial coordinates; if blank, use first station.

Figure 16 HYPO71 program inputs (3).

Additional control over HYPO71 is achieved through the use of the reset list (Figures 17, 18, and 19). This is an optional series of up to thirteen lines. Each line is used to reset one of 13 variables. If no reset list is present, HYPO71 will use a set of default values which are suitable for earthquakes recorded by Calnet.

Test(01) is the cutoff value for RMS value below which Jeffrey's weighting is not used. It should be set to a value approximately equal to the overall timing accuracy of P-arrivals in seconds. The purpose of Jeffrey's weighting is to catch large errors in arrival-time picks. For small networks without about 30 or more arrival times per earthquake Jeffrey's weighting will not work well and should not be used.

Test(02) limits

simultaneous changes in epicenter and depth. If the epicenter adjustment is greater than this amount, the adjustment is recalculated with focal depth fixed.

Test(03) is the critical F-value for the stepwise multiple regression. It should be set according to the number and quality of P- and S- arrivals. If test(03) is set to zero then simple

multiple regression is performed. If test(03) is set to 2 or greater then the iterative process may be terminated prematurely.

Test(04) is used to terminate the iterative location process. If the hypocentral adjustment is less than test(04) the solution is terminated.

Test(05) is used to control the maximum focal depth adjustment. If the adjustment (DZ) at any step is greater than test(05) then DZ is reset to $DZ/(K+1)$ where $K = DZ/\text{test}(05)$.

Test(07), test(08), and test(09) are the coefficients used for calculating duration magnitude (Lee et al. , 1972) according to the formula (using default values):

Test Variables [values used for a sparse network in Alaska are given in brackets.]

```
# Default
01 0.1 sec
    Below this RMS apply
    Jeffrey's Weighting. Not
    recommended for fewer than 30
    stations or after errors have
    been corrected.
    [Alaska, 100]

02 10 km
    Fix depth in any iteration
    with at least this much
    horizontal shift.
    [Alaska, 40 km]

03 2. Critical F-value to control
    which variables to include
    in the regression. Smaller
    number is less restrictive.
    [Alaska, 0.5]
```

Figure 17 HYPO71 program inputs (4).

$$\text{FMAG} = -0.87 + 2\log T + .0035D$$

Test(10) is used to control the maximum adjustment in latitude or longitude. If either adjustment exceeds test(10) at some iteration, then a new, reduced adjustment is calculated using test(10).

Test(11) is the maximum number of iterations in the solution process. Test(12) is used if the focal depth adjustment would put the hypocenter in the air. In that case, a new adjustment is calculated utilizing test(12).

Test(13) is used to control the radius of a sphere centered on the hypocenter location estimate. The RMS residual is calculated at ten points on the sphere to see if there is a nearby location with

lower RMS than the final location. This will not happen if the best location has been found.

05	5 km	On any iteration limit change in depth to this amount. For a computed step of $N \cdot \text{Test}(05)$ move only $\text{Test}(05)/(N+1)$. [Alaska, 35 km]
06	4.	Further F-value iteration control. [Alaska, 100.]
07	-.87	
08	2.0	
09	.0035	Duration magnitude (FMAG) coefficients. [Alaska, -1.15, 2.0, 0.0]
10	100.	Limit horizontal step. Similar to Test(05) algorithm. [Alaska, 35 km]

Figure 18 HYPO71 program inputs (5).

11	8.	Maximum number of iterations. [Alaska, 15]
12	0.5	If upward step would result in air quake, step only this fraction of the way to the surface.
13	1 km	For auxiliary RMS calculations, set radius equal to $\text{Test}(13) \cdot \text{Sqrt}(3)$. [Alaska, 10 km]

Figure 19 HYPO71 program inputs (6).

HYPOELLIPSE Program

I will briefly discuss the HYPOELLIPSE program. I mainly want to emphasize the differences between this program and HYPO71. HYPOELLIPSE is complex and has a lot of options (Figures 20, 21, and 23). Depending on your needs, this may or may not be an advantage. Compared to HYPO71, HYPOELLIPSE is not very stable.

HYPO71 has not changed in the

past 20 years, but HYPOELLIPSE probably changes every two or three months.

HYPOELLIPSE has lots of options that are useful for research and for routine processing of Alaska data, but having many options makes the program more difficult to learn.

Instead of just one velocity model, HYPOELLIPSE allows up to ten velocity models to be specified. The V_p/V_s ratio may be varied in each layer. If detailed crustal velocity information is available, this flexibility is quite helpful.

In addition to constant velocity layers over a half space, one can also specify a linear increase over a half space. Also, in addition to calculating travel times using a layered model, up to three different travel time tables may be used. Code for this feature was

HYPOELLIPSE PROGRAM

Disadvantages - Very Complex

This program has many options that were developed both to routinely locate earthquakes in Alaska and to study various location strategies. The options are very useful but it is often difficult to understand their operation without time consuming trials and in some cases referring directly to the source code.

ADDED FEATURES

Travel time calculations

Up to ten velocity models (five if V_p/V_s varies with depth) may be specified.

In addition to constant velocity layers, one can specify a linear increase over a half space.

Figure 20 HYPOELLIPSE program (1) .

borrowed from Fred Klein's

HYPOINVERSE program. All

distance calculations are based on a

spherical earth, so the program

could be used for more distant

earthquakes.

There are two different

ways of handling elevation

corrections. The first is to add a

station delay which is a function of

the station elevation. The way that I

prefer, in areas where there is a lot of

topographic relief, is to set the top of

the model to the maximum elevation of

interest (Figure 22). All the station

elevations, which are specified with

respect to sea level, are less than the

elevation of the top of the model so all

of the stations are, effectively, located

within the model. If an earthquake occurs above a station, the travel time is computed for a

ray traveling down to the station. There is no elevation correction per se. One disadvantage

of this approach is that it is possible to get hypocenter locations in the air. On the other

Three travel time tables may be used.

Distances are computed for a spherical earth, so flat-earth as well as spherical-earth tables may be used.

Elevation corrections may be made based on station elevations.

For areas of great topographic relief, such as near a volcano or within a mine, the stations may be embedded within the model.

A user modified subroutine allows different velocity models to be used in different geographic regions.

Up to ten different delay models may be specified for selected regions.

Figure 21 HYPOELLIPSE program (2).

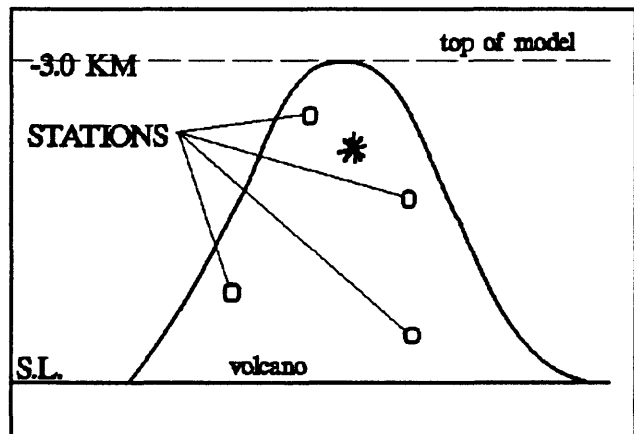


Figure 22 Stations embedded in model.

hand, such a location provides a clear indication of model or data problems that might otherwise go undetected.

In order to locate earthquakes throughout Alaska, one of three different models and starting depths is automatically selected depending on the initial location of the event. To implement this feature for a different region, one user-defined subroutine would need to be modified.

After the problem of multiple minima was recognized, a

new option, called the "global option," was added (Figure 23). The solution still may not be global, but this option increases the probability of finding the global solution. With the global option, the earthquake is started at three different depths and the local minimum corresponding to each is determined. Then the RMS values for the solutions are compared to see which has the minimum value. That solution is then returned as the final location. This approach has proved helpful in locating earthquakes with multiple minima such as were illustrated in Figure 8.

HYPOELLIPSE also has options which allows the variation of RMS with depth to be computed and to make repeated runs of either artificial or real data with added random

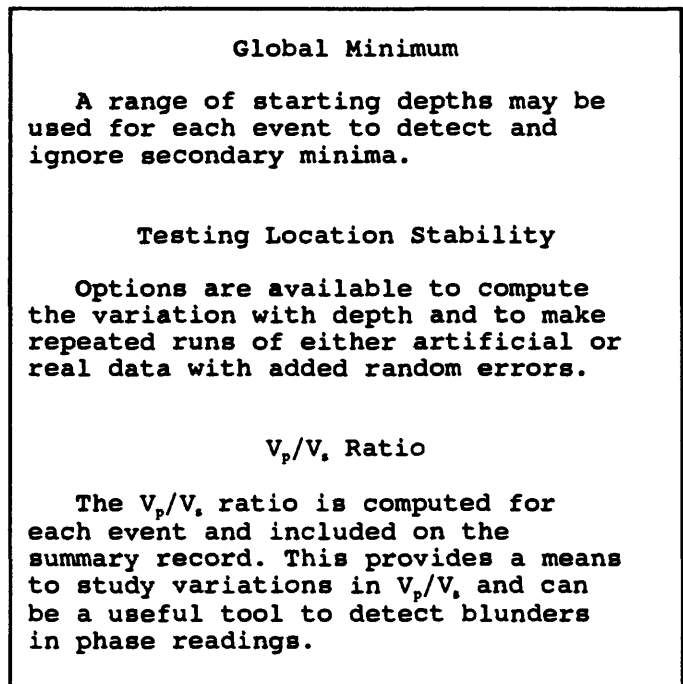


Figure 23 HYPOELLIPSE program (3).

errors. These options provide a means of testing location stability.

HYPOELLIPSE will optionally compute the V_p/V_s ratio for each earthquake processed. The computed V_p/V_s ratio may be used in part as a diagnostic tool. If the ratio is out of bounds, then a phase interpretation error is indicated. Spatial or temporal variations in V_p/V_s may also be studied.

In HYPOELLIPSE the station list contains a history of the station polarity, the station coordinates, the telemetry delay, and the gain as a function of time. Because of this, an arbitrarily long sequence of earthquakes may be processed even if station parameters have changed during the sequence. As time progresses, the program automatically updates station parameters from the information in the station list. We have found this to be a very useful feature. The station list allows independent P- and S- delays instead of having to use a fixed V_p/V_s ratio to compute S-delay, as in HYPO71.

Location error limits are specified by HYPOELLIPSE in terms of the error ellipsoid (Figure 24). This is the region in space in which there is 66% confidence that the earthquake occurred. Figure 24 illustrates an error ellipsoid and its projection into map view. The ellipsoid can also be

projected onto a vertical cross section. The ellipsoid shape is determined entirely by the velocity model and the geometric relationship of the stations to the hypocenter. The scaling

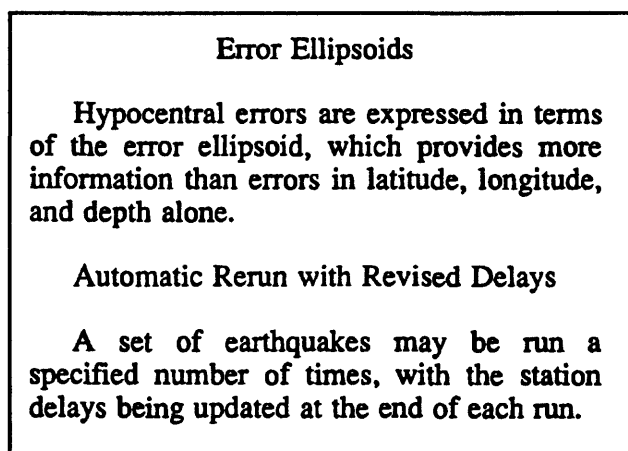


Figure 24 Error ellipsoids.

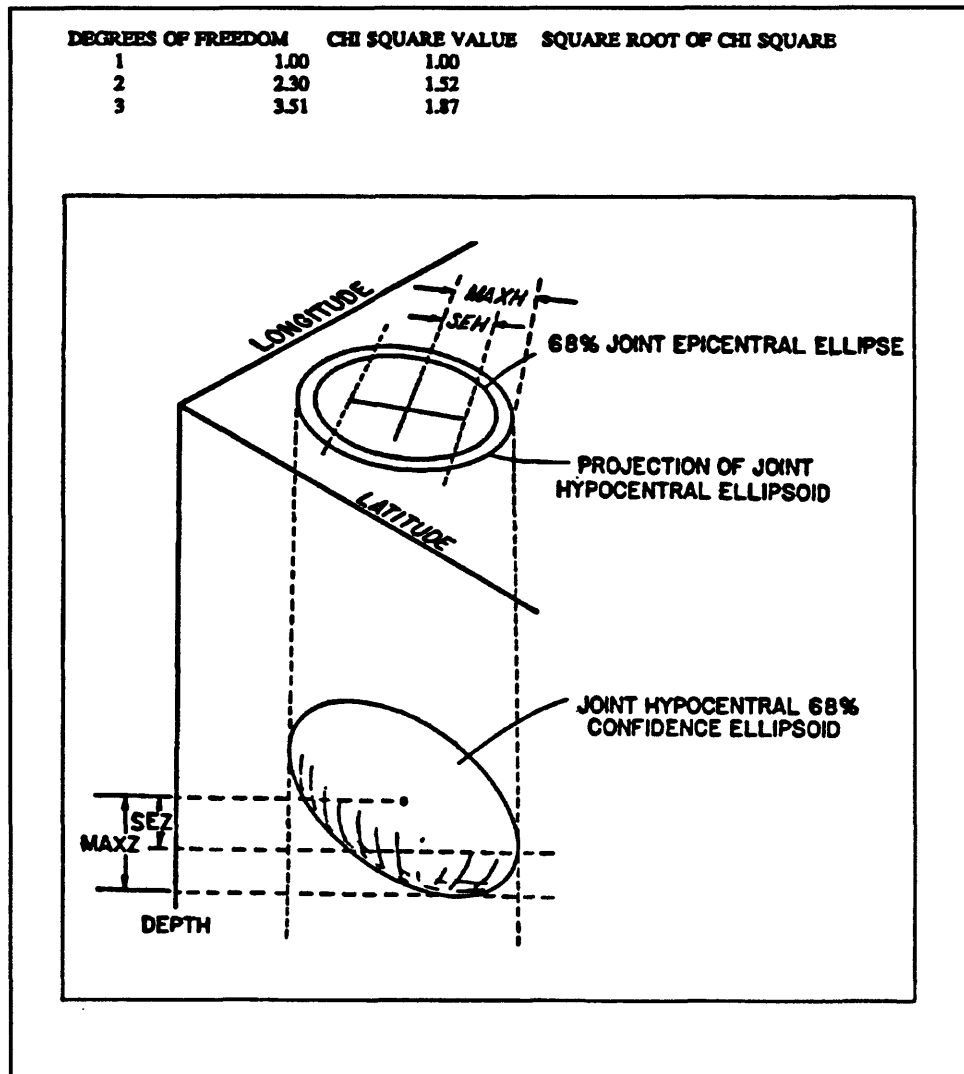


Figure 25 Error ellipsoid relationships.

of the ellipsoid is based on an estimate of the probable reading errors of the arrival times.

The error ellipsoid estimates the precision of the location but does not include any information on potential systematic errors that might bias the location.

REFERENCES

- Klein, F.W. (1978). Hypocenter location program--HYPOINVERSE. U.S. Geol. Surv. Open-File Report 78-694.
- Lahr, J.C. (1979). HYPOELLIPSE: A computer program for determining local earthquake hypocentral parameters, magnitude, and first motion pattern. U.S. Geol. Surv. Open-File Report 79-431.
- Lee, W.H.K., R.E. Bennett, and K.L. Meagher (1972). A method of estimating magnitude of local earthquakes from signal duration. U.S. Geol. Surv. Open-file Report. 28.
- Lee, W.H.K., and J.C. Lahr (1975). HYPO71 (revised): A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes. U.S. Geol. Surv. Open-File Report 75-311, 116 pp.
- Lee, W.H.K., and C.M. Valdes (1989). User manual for HYPO71PC, IASPEI Software Library , 1, 203-236.

12. COMPUTING TRAVEL TIME AND DERIVATIVES

by

W. H. K. Lee

U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

I will start by presenting an introduction to the problem: Why do we need to compute travel times and derivatives? Next, I will discuss the basic theory behind ray tracing and show how this can be applied to some simple one-dimensional velocity

Computing Travel Time And Derivatives

- 1) Introduction
- 2) Seismic Ray Tracing
- 3) Simple 1-D Velocity Models
- 4) Numerical Solution for 2-D and 3-D Velocity Models

Figure 1. Computing travel time and derivatives.

models. Finally, I will present some solutions for two-dimensional and three-dimensional velocity models. I am not going to discuss all the mathematical details of this problem. These details have been published, and if you have the mathematical background, you can read this material at your leisure. However, it is necessary to know why we need to do all this.

In Chapter 10, we discussed the earthquake location problem. Recall that it boils down to an iterative procedure of solving a set of linearized equations of the form:

$$\delta t + \left(\frac{\partial T_k}{\partial x} \right) \delta x + \left(\frac{\partial T_k}{\partial y} \right) \delta y + \left(\frac{\partial T_k}{\partial z} \right) \delta z = r_k = T^o_k - T^c_k \quad (1)$$

for $k = 1, 2, \dots, m$

We have made a guess at the earthquake location and need to improve the guess so that it

minimizes the sum of the residuals squared. The adjustment vector consists of four components, δt , δx , δy , δz , which are related to the residual (r_k) through the partial derivatives. For each observation, you can write one of these equations, and you just need to solve it by some means. In order to solve Equation (1), we need to calculate the theoretical travel time T_k^e and the spatial partial derivatives.

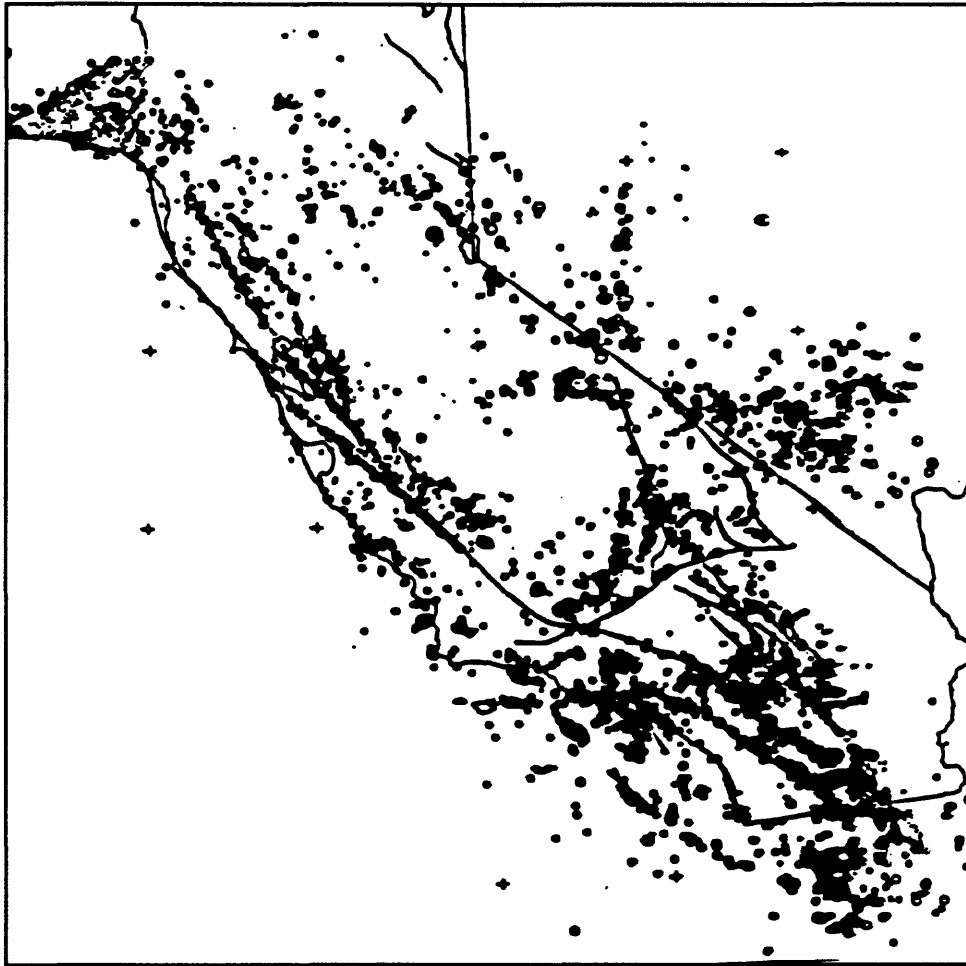


Figure 2. Distribution of earthquake epicenters in California for a two year period.

More than twenty years ago, J. P. Eaton (1969) solved this problem for the multiple layer case. If you look at his paper, you will see that the algebra is very messy. However, a

simple layered model is far from being the real earth.

This became apparent over the years. Figure 2 shows an epicenter map of California for the two year period 1980 - 1981. On the whole, it is not bad. As you can see, a great number of the epicenters plot right on the fault, right where we expect them to. But this is a very large scale map.

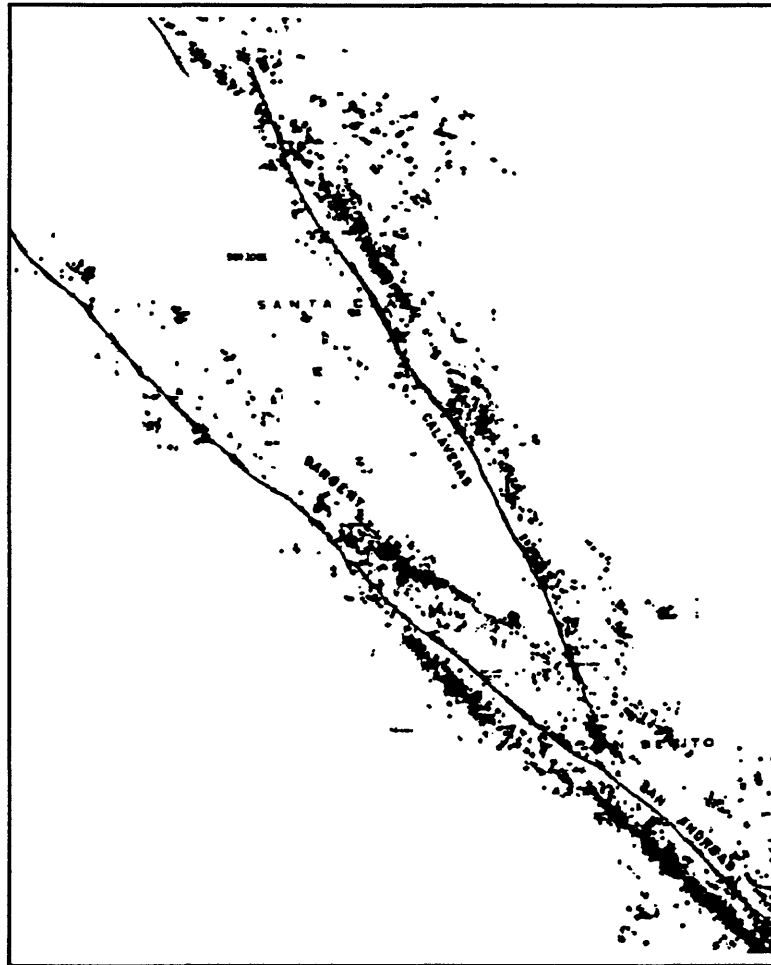


Figure 3. Distribution of earthquake epicenters in the San Francisco Bay area (1969 - 1971).

Figure 3 shows an epicenter map for the San Francisco Bay area. Notice, that the

earthquake epicenters, on average, are off the faults by a couple of kilometers. Of course, this makes the geologists unhappy! The reason for this behavior is that the velocity is not the

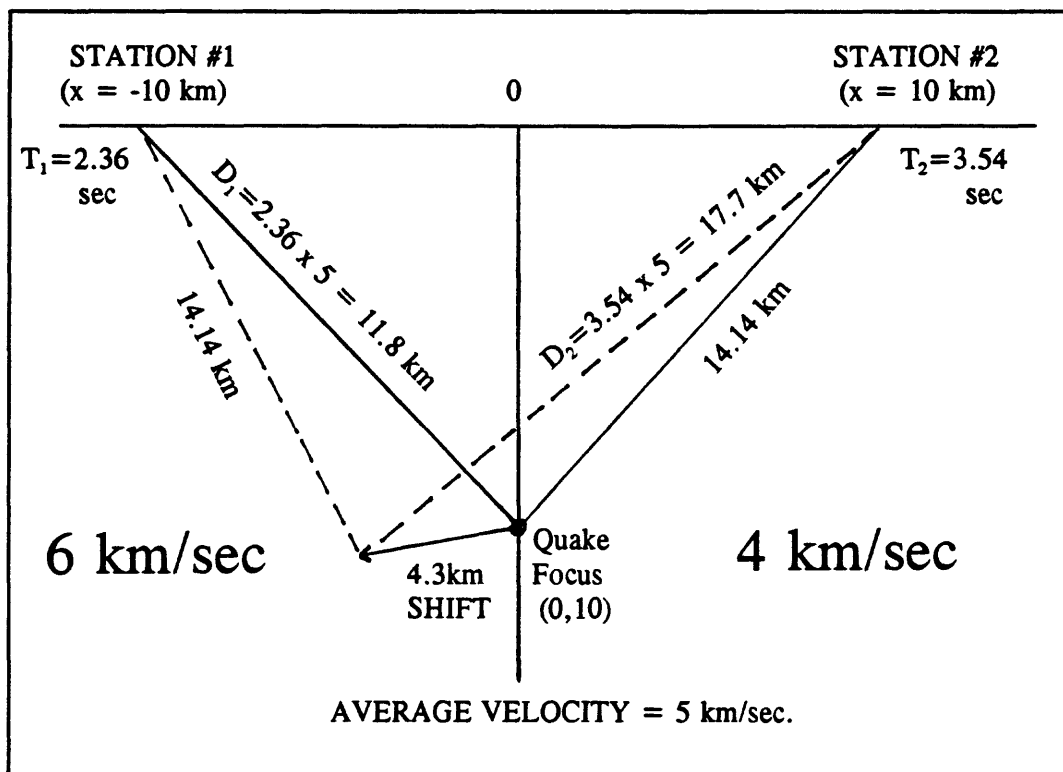


Figure 4. Velocity model to explain earthquake epicenter shift.

same on both sides of the fault. Figure 4 shows a hypothetical cross section through a fault. There are two stations on opposite sides of the fault. The fault is in the center of the diagram and is idealized as a vertical boundary with different velocities on opposite sides. The material on the left-hand side has a P-wave velocity of 6 kilometers per second, and the material on the right has a P-wave velocity of 4 kilometers per second. If an earthquake occurs at a depth of 10 kilometers as shown, the first arrivals at station 2 will take 3.54 seconds to arrive, but the first arrivals at station 1 will only require 2.36 seconds to travel from the focus to the station. If you didn't know the fault was there and you were using an

average P-wave velocity of 5 kilometers per second, then the travel time to station 2 would correspond to a (model) distance of 17.7 kilometers, while the travel time to station 1 would correspond to a distance of 11.8 kilometers. Therefore, in the constant velocity solution, the hypocenter would be at the intersection of two arcs of these lengths. This is the point shown in Figure 4, which is 4.3 kilometers to the left of the fault. In order to avoid this sort of problem, we must be able to handle more complex velocity models. This will involve seismic ray tracing.

Figure 5 shows the framework for the problem. There is a Cartesian coordinate system imposed on some isotropic medium. This system is used to describe the ray path extending from point A to point B. P is an arbitrary point on the ray path which has coordinates (x,y,z) , and which can be described by the position vector \mathbf{r} . We

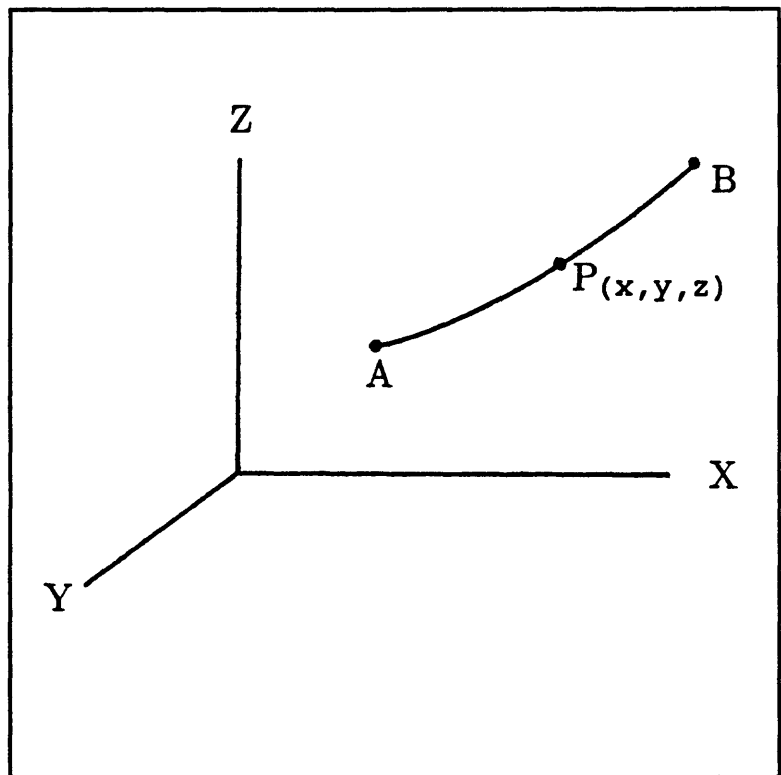


Figure 5. Framework for seismic ray tracing.

assume that velocity can be an arbitrary function of position expressed as $v(x,y,z)$. The mathematics is greatly simplified by introducing the "slowness", u , which is just $1/v$. The

gradient of the slowness is given by:

$$\nabla u = \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial u}{\partial z} \end{pmatrix} \quad (2)$$

Then the time T required to travel from point A to point B is given by:

$$T = \int_A^B u ds \quad (3)$$

where ds is an element along the ray path. For a minimum time path, Fermat's principle states that the time T has a stationary value. Using the calculus of variation, the minimum time path is given by the "ray equation":

$$\frac{d}{ds} \left[u \left(\frac{d\vec{r}}{ds} \right) \right] = \nabla u \quad (4)$$

Equation 4 is the ray equation expressed in vector notation. Expressed as components, the ray equation is:

$$\begin{aligned} \frac{d}{ds} \left(u \frac{dx}{ds} \right) &= \frac{\partial u}{\partial x} \\ \frac{d}{ds} \left(u \frac{dy}{ds} \right) &= \frac{\partial u}{\partial y} \\ \frac{d}{ds} \left(u \frac{dz}{ds} \right) &= \frac{\partial u}{\partial z} \end{aligned} \quad (5)$$

Equation 5 is a set of 3 second order ordinary differential equations. This type of equation has been studied for many years. Consequently, there are many "tricks" that have been devised to aid in the solution of these equations.

ONE-DIMENSIONAL CONTINUOUS VELOCITY MODEL

The simplest case is the constant velocity case. This class of problem can be solved by trigonometry. The next level of complexity is a model in which velocity is a function of z alone. This is the model which has seen the most use by seismologists, and is in wide spread use, even today. In this case, the ray equation becomes:

$$\frac{1}{v(z)} \frac{dx}{ds} = \text{const}, \quad \frac{d}{ds} \left(\frac{1}{v(z)} \frac{dz}{ds} \right) = -\frac{1}{v^2} \frac{dv}{dz} \quad (6)$$

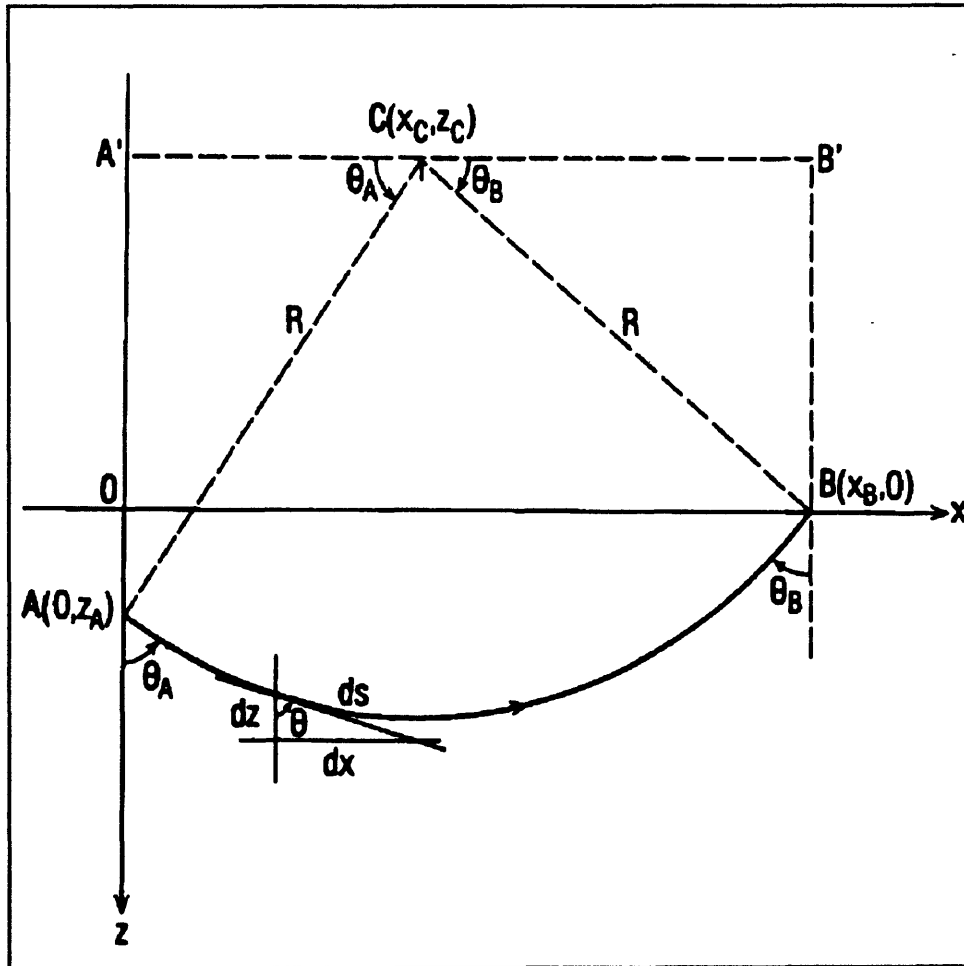


Figure 6. Geometry for velocity given by $v = v(z)$.

The geometry is shown in Figure 6. From Figure 6, the direction cosines are:

$$\begin{aligned}\cos \alpha &= \frac{dx}{ds} = \sin \theta \\ \cos \gamma &= \frac{dz}{ds} = \cos \theta\end{aligned}\tag{7}$$

Then, from Equations (6) and (7) we get "Snell's Law":

$$\frac{\sin \theta}{v(z)} = \text{const.} \equiv p\tag{8}$$

where p is called the ray parameter.

The second member of Equation (6) reduces to:

$$\frac{d\theta}{ds} = p \frac{dv}{dz}\tag{9}$$

From this we can solve for the travel time as a function of epicentral angle:

$$T = \int_A^B \frac{ds}{v(z)} = \int_A^B \frac{dz}{v(z) \cos \theta} = \int_{\theta_A}^{\theta_B} \frac{d\theta}{(dv/dz) \sin \theta}\tag{10}$$

Let's consider the linear case where the velocity is given by:

$$v(z) = g_0 + gz\tag{11}$$

This velocity distribution results in a ray path which is an arc of a circle as shown in Figure

6. The radius of this circle is given by:

$$R = \left(\frac{d\theta}{ds} \right)^{-1} = \frac{v}{g \sin \theta}\tag{12}$$

The center of this arc is given (in terms of the parameters from Figure 6) as:

$$x_c = \frac{x_B^2 - 2(g_0/g)z_A - z_A^2}{2x_B}, \quad z_c = \frac{-g_0}{g} \quad (13)$$

The travel time for this linear velocity case is:

$$T = \int_{\theta_A}^{\theta_B} \frac{d\theta}{g \sin \theta} = \frac{1}{g} \ln \left[\frac{\tan(\theta_B/2)}{\tan(\theta_A/2)} \right] \quad (14)$$

$$\theta_A = \tan^{-1} \left(\frac{z_A - z_c}{x_c} \right), \quad \theta_B = \tan^{-1} \left(\frac{-z_c}{x_B - x_c} \right)$$

Finally, the spatial derivatives of T at the source are:

$$\left. \frac{\partial T}{\partial x} \right|_A = \frac{-\sin \theta_A}{(g_0 + g z_A)} \quad (15)$$

$$\left. \frac{\partial T}{\partial z} \right|_A = \frac{-\cos \theta_A}{(g_0 + g z_A)}$$

It is possible to derive all these equations from first principles. However, the algebra required to do it that way is more involved than the algebra required to derive the equations from the general equations.

ONE-DIMENSIONAL MULTI-LAYER VELOCITY MODELS

The next case to consider is a layered earth model. The simplest version of this case is referred to as a layer over a half-space. This is shown in Figure 7. The travel time for this case can be determined using simple trigonometry and is given as:

$$T_r = \frac{\Delta}{v_2} + \frac{2h_1}{v_1} \cos \theta \quad (16)$$

The critical distance is the smallest offset at which a refracted wave will appear. It is:

$$\xi = 2h_1 \tan \theta \quad (17)$$

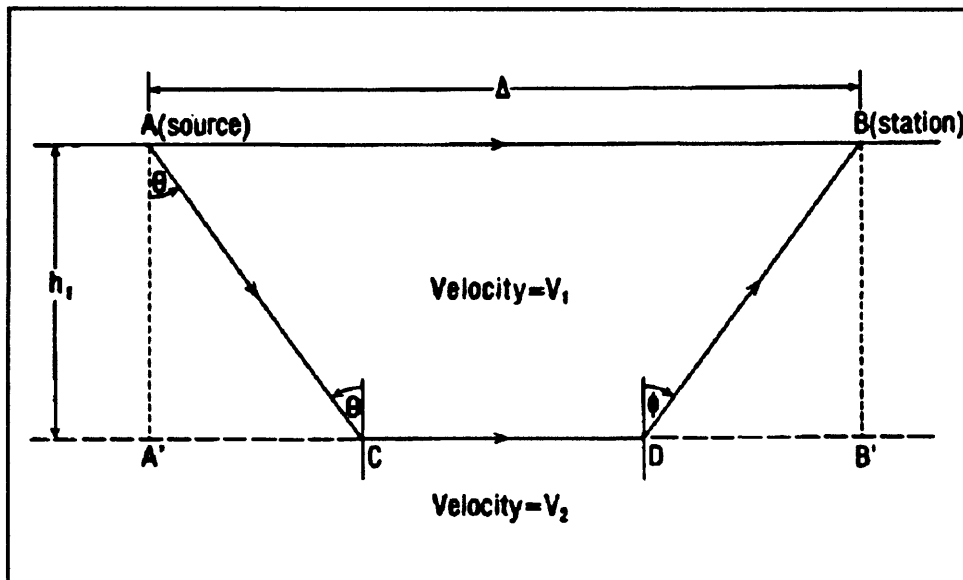


Figure 7. Layer over a half-space model.

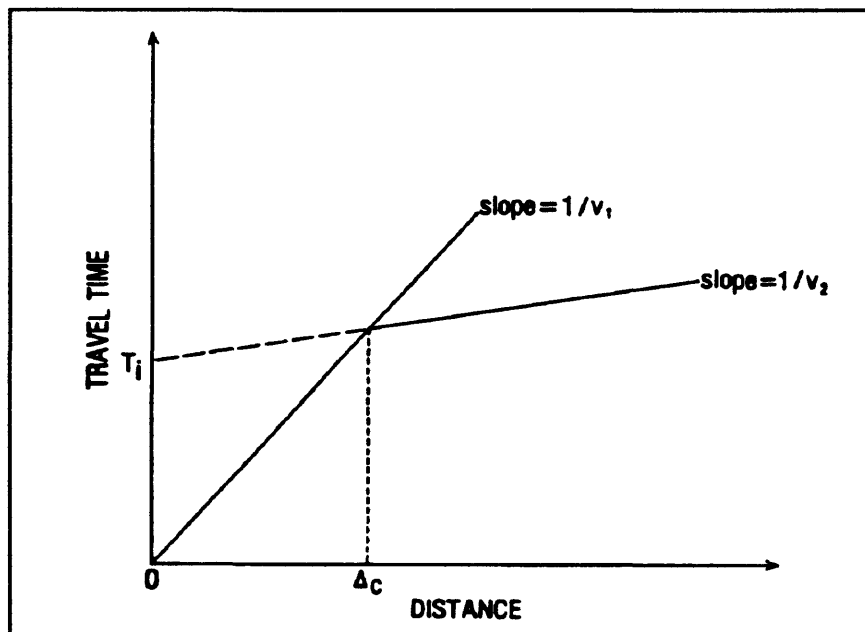


Figure 8. Travel time vs distance for layer over half-space model.

Figure 8 is a plot of travel time versus distance for this layer over a half-space model. The first branch of this curve is for the direct wave and has slope $1/v_1$, and the second branch, which is for the refracted wave, has slope $1/v_2$. The intersection of these two branches is the crossover distance at which the refracted arrival is observed before the direct arrival. The layer over a half-space model can be generalized to multiple layers as shown in Figure 9. The mathematics gets rather messy, but if you have the patience, you can work out the formulas for the travel time and the critical distance. They turn out to be:

$$T_k = \frac{\Delta}{v_k} + 2 \sum_{i=1}^{k-1} \frac{h_i (v_k^2 - v_i^2)^{\frac{1}{2}}}{v_i v_k}$$

$$\xi_k = 2 \sum_{i=1}^{k-1} \frac{h_i v_i}{(v_k^2 - v_i^2)^{\frac{1}{2}}}$$
(18)

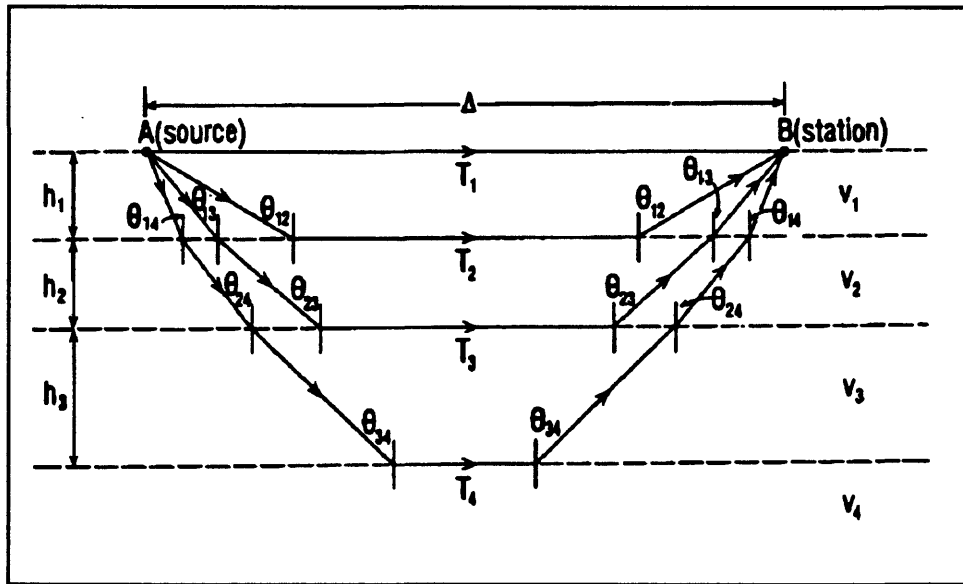


Figure 9. Multiple layer case.

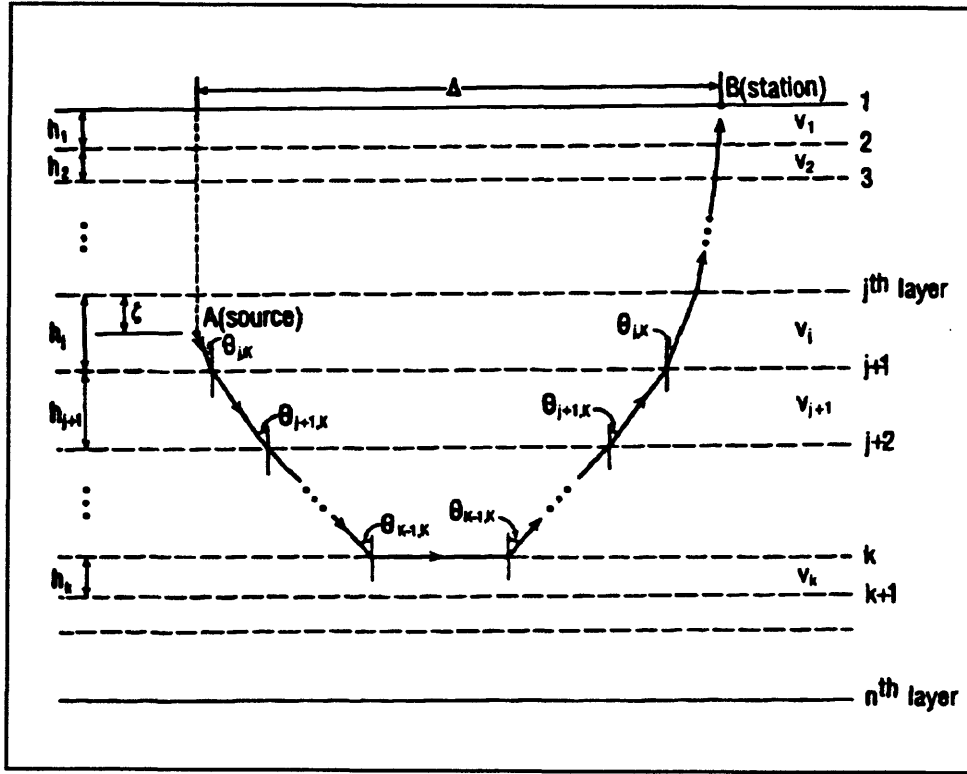


Figure 10. Source at depth.

The next step is to consider a source at depth in a multiple layer case. Figure 10 shows the geometry. This case was originally worked out by J.P. Eaton (1969). The equations get a little messier, but the problem is still basically one that can be solved by trigonometry. The travel time and critical distance are:

$$\begin{aligned}
 T_{jk} &= \frac{\Delta}{v_k} - \frac{\zeta \Omega_{kj}}{v_j v_k} + \sum_{i=1}^{j-1} \frac{h_i \Omega_{ki}}{v_i v_k} + 2 \sum_{i=j}^{k-1} \frac{h_i \Omega_{ki}}{v_i v_k} \\
 \xi_{jk} &= -\frac{\zeta v_j}{\Omega_{kj}} + \sum_{i=1}^{j-1} \frac{h_i v_i}{\Omega_{ki}} + 2 \sum_{i=j}^{k-1} \frac{h_i v_i}{\Omega_{ki}}
 \end{aligned} \tag{19}$$

$$\text{Where} \quad \Omega_{ki} = \sqrt{v_k^2 - v_i^2}$$

These last few models have been used to get expressions for the refracted wave travel times. However, we have not yet provided an expression for the direct arrivals. For an

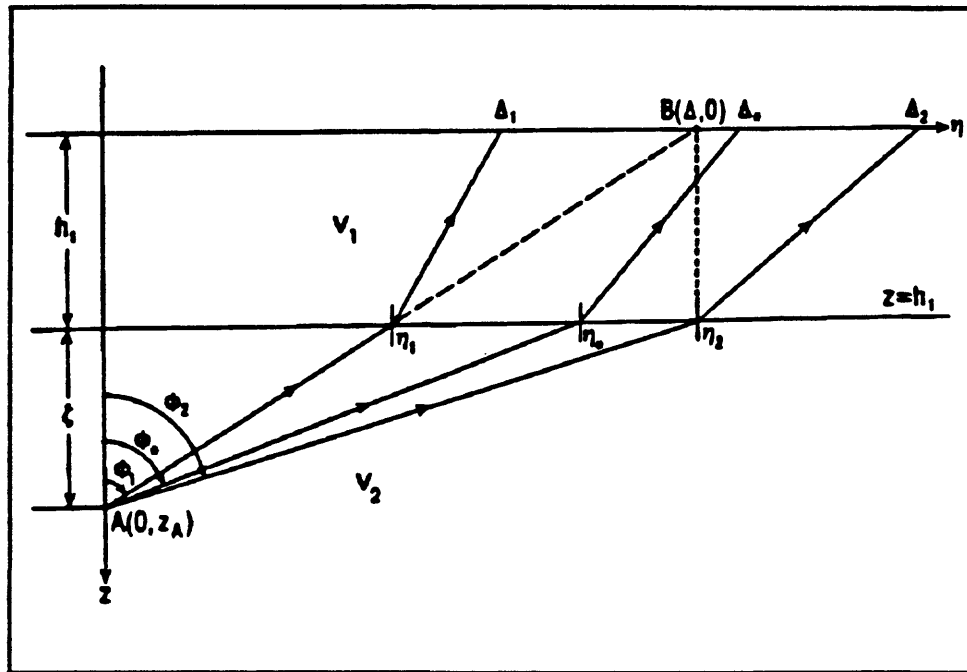


Figure 11. Travel path for a direct wave in a layered model.

earthquake which occurs in the first layer, the direct path is entirely in the first layer. But, if the earthquake occurs in the second or deeper layer, then you need to take into account the refraction of the direct wave, which occurs at each interface on the way to the surface. This problem was also solved by J. P. Eaton (1969). The approach is straight-forward. Figure 11 shows the geometry. First, two extreme take-off angles are defined such that the actual take-off angle must be between them. The corresponding epicentral distances bracket the distance from the station to the epicenter. Then, angles between these two extremes are tried using a bisection method, until one which provides a satisfactory match to the epicentral distance is determined. This process converges rather rapidly. As shown in Figure 11 the ϕ_1 initial guess is determined by picking a straight line path between the source and the receiver. Because of upward refraction at the layer boundary, that ray must emerge short of the

receiver. The other extreme angle is chosen by dropping a perpendicular from the receiver to the boundary layer. The line connecting the intersection point with the source defines the angle.

Once you have determined a ray path for a given n layer case, you can also compute the total travel time associated with that ray path. The spatial partial derivatives of the travel time evaluated at the source for the layered model are given in Equation (20).

$$\begin{aligned}\frac{\partial T_{jk}}{\partial x} \big|_A &= -\frac{(x_B - x_A)}{\Delta v_k} \\ \frac{\partial T_{jk}}{\partial y} \big|_A &= -\frac{(y_B - y_A)}{\Delta v_k} \\ \frac{\partial T_{jk}}{\partial z} \big|_A &= -\frac{(v_k^2 - v_j^2)^{1/2}}{(v_j v_k)}\end{aligned}\tag{20}$$

The primary reason for showing some of the mathematics here is just to illustrate the point that even for very simple models, the mathematics associated with seismic ray tracing can become quite complicated. There is a fair amount of symmetry in many of the formulations and this is an aid in calculation. For instance, the x and y equations from Equation 20 are interchangeable with each other. The travel time derivative expression for the direct wave (not shown here) also shows the same type of symmetry. This symmetry can also be helpful when you are trying to work out the details of a ray tracing problem. If you derive an expression for a ray path (in a flat layered model), and it does not show this type of symmetry, then you know you have probably done something wrong.

MORE COMPLEX MODELS

The equations for the last few models are all derivable using trigonometry and basic

principles because the models are quite simple. However, If you need to work with a two-dimensional or three-dimensional velocity model, you have to go back to the ray equation. Even then, there are only a few analytic solutions for special cases so you must solve the equation numerically in order to get ray paths and travel times. In order to solve this set of equations, it must first be reduced to a set of linear equations. We can do this because of a theorem which says that any 2nd order ODE can be reduced to two 1st order ODEs. The proof is straight-forward and is shown in Figure 12.

$$\begin{aligned}
 &\text{Given } y'' = f(x) \\
 &\text{where } y' \equiv \frac{dy}{dx}, \quad y'' = \frac{d^2y}{dx^2} \\
 &\text{choose } \omega_1 \equiv y \quad \omega_2 \equiv \omega'_1 = y' \\
 &\text{then } \omega'_2 = (y')' = y'' \\
 &\text{Hence, we have} \\
 &\quad \left[\begin{array}{l} \omega'_1 = \omega_2 \\ \omega'_2 = f(x) \end{array} \right]
 \end{aligned}$$

Figure 12 Reduction of a 2nd order ODE to two 1st order ODEs.

In practice, it is very important how you choose the parameters which define the problem. Otherwise, this process of reducing the equation into a set of first order equations can become very messy. Essentially, most people have done this by trial and error.

Eliseevnin (1965) proposed a solution to the ray equation which is outlined in Figure 13. Later on, this solution was used by Jacob (1970), and Julian (1970). The parameters used

by Eliseevnin were $x, y, z, \alpha, \beta, \gamma$, and t . The resulting solution is very messy, although workable.

Eliseevnin (1965)

Choice: $x, y, z, \alpha, \beta, \gamma$; $t = \text{time}$

Solve:

$$\begin{aligned} dx/dt &= v \cos\alpha \\ dy/dt &= v \cos\beta \\ dz/dt &= v \cos\gamma \\ d\alpha/dt &= \frac{\partial v}{\partial x} \sin\alpha - \frac{\partial v}{\partial y} \cot\alpha \cos\beta - \frac{\partial v}{\partial z} \cot\alpha \cos\gamma \\ d\beta/dt &= -\frac{\partial v}{\partial x} \cos\alpha \cot\beta + \frac{\partial v}{\partial y} \sin\beta - \frac{\partial v}{\partial z} \cot\beta \cos\gamma \\ d\gamma/dt &= -\frac{\partial v}{\partial x} \cos\alpha \cot\gamma - \frac{\partial v}{\partial y} \cos\beta \cot\gamma + \frac{\partial v}{\partial z} \sin\gamma \end{aligned}$$

Note: $\cos\alpha \equiv dx/ds$, $\cos\beta \equiv dy/ds$, $\cos\gamma \equiv dz/ds$

Figure 13. Eliseevnin's parameterization.

In 1971 Rob Wesson introduced a different parameterization shown in Figure 14. The resulting equations, although fewer in number, and free of trigonometric quantities, are still rather complex. One would like to have a simpler formulation.

Still another parameterization was introduced by Julian and Gubbins (1977). They defined the parameter q which is the element of path divided by the total ray path. The resulting equations are shown in Figure 15. The equations are algebraically simpler than the previous formulations. However, I think that they are still much more complicated than is desirable.

Wesson (1971)

Choice: $y = y(x)$, $z = z(x)$

Solve:

$$\frac{d}{dx} \left[\frac{y'}{v \sqrt{(1 + y'^2 + z'^2)}} \right] + \frac{\sqrt{(1 + y'^2 + z'^2)}}{v^2} \frac{\partial v}{\partial y} = 0$$

$$\frac{d}{dx} \left[\frac{z'}{v \sqrt{(1 + y'^2 + z'^2)}} \right] + \frac{\sqrt{(1 + y'^2 + z'^2)}}{v^2} \frac{\partial v}{\partial z} = 0$$

where $y' = dy/dx$, and $z' = dz/dx$.

Figure 14. Wesson's (1971) parameterization.

Julian and Gubbins (1977)

Choice: $q = s/S$

Solve:

$$-u_x (y'^2 + z'^2) + u_y x' y' + u_z x' z' + u x'' = 0$$

$$-u_y (x'^2 + z'^2) + u_x x' y' + u_z y' z' + u y'' = 0$$

$$x' x'' + y' y'' + z' z'' = 0$$

where s = element of arc length,

S = total path length, and

$'$ = differentiation with respect to q .

Figure 15. Parameterization by Julian and Gubbins (1977).

In the late 1970's I was working with a group of mathematicians. One day I proposed

that there must be a simpler representation of the ray equation than those that had been developed. One of the mathematicians, by the name of Victor Pereyra, did manage to devise

Pereyra, Lee, and Keller (1980)

Choice: $\omega_1 = x, \quad \omega_3 = y, \quad \omega_5 = z$
 $\omega_2 = x', \quad \omega_4 = y', \quad \omega_6 = z'$

Solve:

$$\begin{aligned}\omega'_1 &= \omega_2 \\ \omega'_2 &= v(u_x - G\omega_2) \\ \omega'_3 &= \omega_4 \\ \omega'_4 &= v(u_y - G\omega_4) \\ \omega'_5 &= \omega_6 \\ \omega'_6 &= v(u_z - G\omega_6)\end{aligned}$$

where s = element of arc length,
 $'$ = differentiation with respect to s .
 $u_x = \partial u / \partial x, \quad u_y = \partial u / \partial y, \quad u_z = \partial u / \partial z$
 $G = u_x \omega_2 + u_y \omega_4 + u_z \omega_6$

Figure 16. Parameterization by Pereyra, Lee, and Keller (1980).

a parameterization which resulted in a much simpler representation of the ray equation.

Although there are six equations in the solution, three of them have only one right-hand side term and the other three have very simple algebraic expressions for the right-hand side.

I was quite happy with this representation. I wrote some code based on it and everything worked fine. At that time I had a summer student by the name of Rob Comer. I showed him the representation that I was working with and it apparently interested him, so a

few days later he presented me with an even better representation. Rob never published the solution so actually, it was published in one of my papers. The representation is shown in Figure 17. The parameterization is similar to the last case, except that here the slowness has been introduced as one of the parameters. The resulting six equations are about as simple as possible. In each case, there is only a single multiplication or else a single calculation of the gradient. This computational simplicity results in significant savings in computing time when the ray tracing algorithm is at the heart of some iterative scheme such as hypocenter location. Although the set of equations is quite simple, it is the result of almost 20 years of work by a number of people.

Lee, Luk, and Mooney (1984) [due to R.P. Comer]

Choice: $\omega_1 = x, \quad \omega_3 = y, \quad \omega_5 = z$
 $\omega_2 = u \, dx/ds, \quad \omega_4 = u \, dy/ds, \quad \omega_6 = u \, dz/ds$

Solve:

$$\begin{aligned}\omega'_1 &= v \, \omega_2 \\ \omega'_2 &= \partial u / \partial x \\ \omega'_3 &= v \, \omega_4 \\ \omega'_4 &= \partial u / \partial y \\ \omega'_5 &= v \, \omega_6 \\ \omega'_6 &= \partial u / \partial z\end{aligned}$$

where s = element of arc length,
 $'$ = differentiation with respect to s .

Figure 17. Parameterization due to R. P. Comer.

It also occurred to us that we would like to be able to get the travel time directly

rather than obtaining it by integration along the ray path. Not only is the integration procedure time consuming, but it also tends to accumulate numerical errors to a significant degree. By adding some more equations to the problem, it is possible to get the travel time directly. Figure 18 shows an outline of the method.

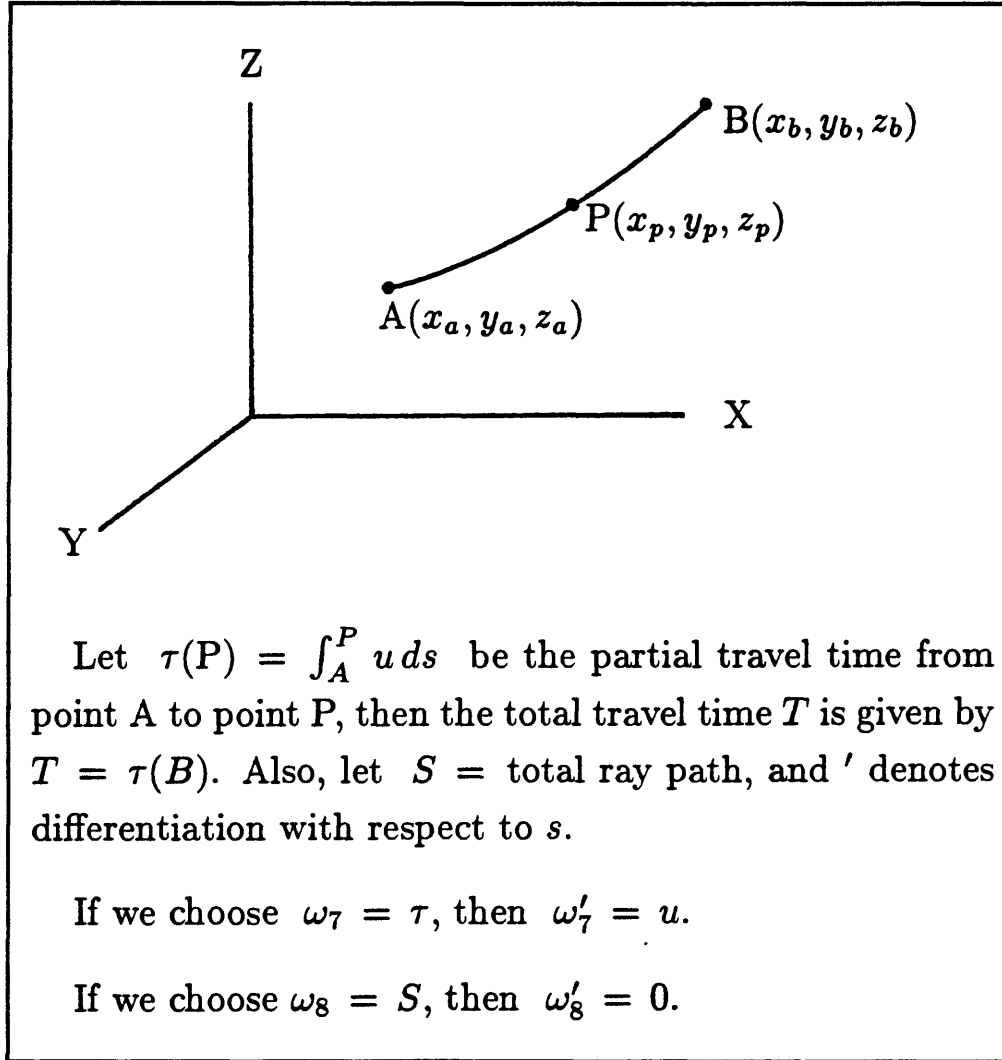


Figure 18. Calculating the travel time.

The integral in the figure gives the partial travel time from point A to point P along the ray path. Then the total travel time is given by $T(B)$ since B is also an arbitrary point on

the ray path. We introduce a new parameter $\omega_7 = T$. We want to normalize the ray so we define another new parameter $\omega_8 = S$, where S is the total path length. We normalize by letting $q = s/S$. The final set of equations is given in Figure 19.

Final Set of Seismic Ray Equations

Normalization:

Let $q = s/S$ and $' = d/dq$, for normalization.

Final set of equations to be solved are (for $q \in (0, 1)$):

$$\omega'_1 = \omega_8 v \omega_2$$

$$\omega'_2 = \omega_8 \partial u / \partial x$$

$$\omega'_3 = \omega_8 v \omega_4$$

$$\omega'_4 = \omega_8 \partial u / \partial y$$

$$\omega'_5 = \omega_8 v \omega_6$$

$$\omega'_6 = \omega_8 \partial u / \partial z$$

Solutions are:

$$\omega_1 = x, \omega_3 = y, \omega_5 = z, \omega_7 = \tau, \omega_8 = S$$

$$\omega_2 = u dx/ds, \omega_4 = u dy/ds, \omega_6 = u dz/ds$$

Boundary Conitions are:

$$\omega_1(0) = x_a, \omega_3(0) = y_a, \omega_5(0) = z_a$$

$$\omega_1(1) = x_b, \omega_3(1) = y_b, \omega_5(1) = z_b$$

$$\omega_7(0) = 0, \omega_2^2(0) + \omega_4^2(0) + \omega_6^2(0) = u^2(0)$$

Figure 19. Final set of equations for seismic ray tracing.

Because of the normalization, q is between zero and one. The solution of this set of

equations gives you all the parameters of the ray path. For any point on the ray, you know the coordinates, the travel time to that point, the direction cosines, and the total path length to that point. There are 8 boundary conditions for solving the set of 8 equations. These are also listed in Figure 19.

Of course, we do not solve this set of equations analytically. The solution is done using computer software. Figure 20 is a flow chart of the solution procedure. The driver program first sets up the model which will be used to solve the problem, and then coordinates the activity of the various routines which carry out the solution.

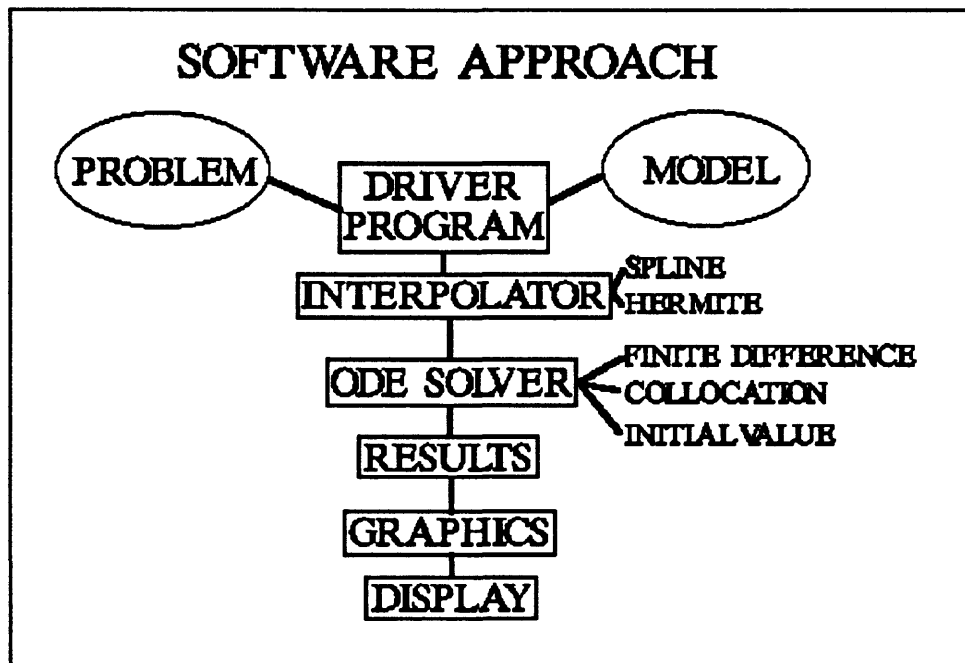


Figure 20. Procedure for solving ray equation.

One of the biggest problems in the solution process has to do with interpolation. We spent quite a bit of time working on the interpolation problem. We finally settled on using spline interpolation and Hermitian interpolation.

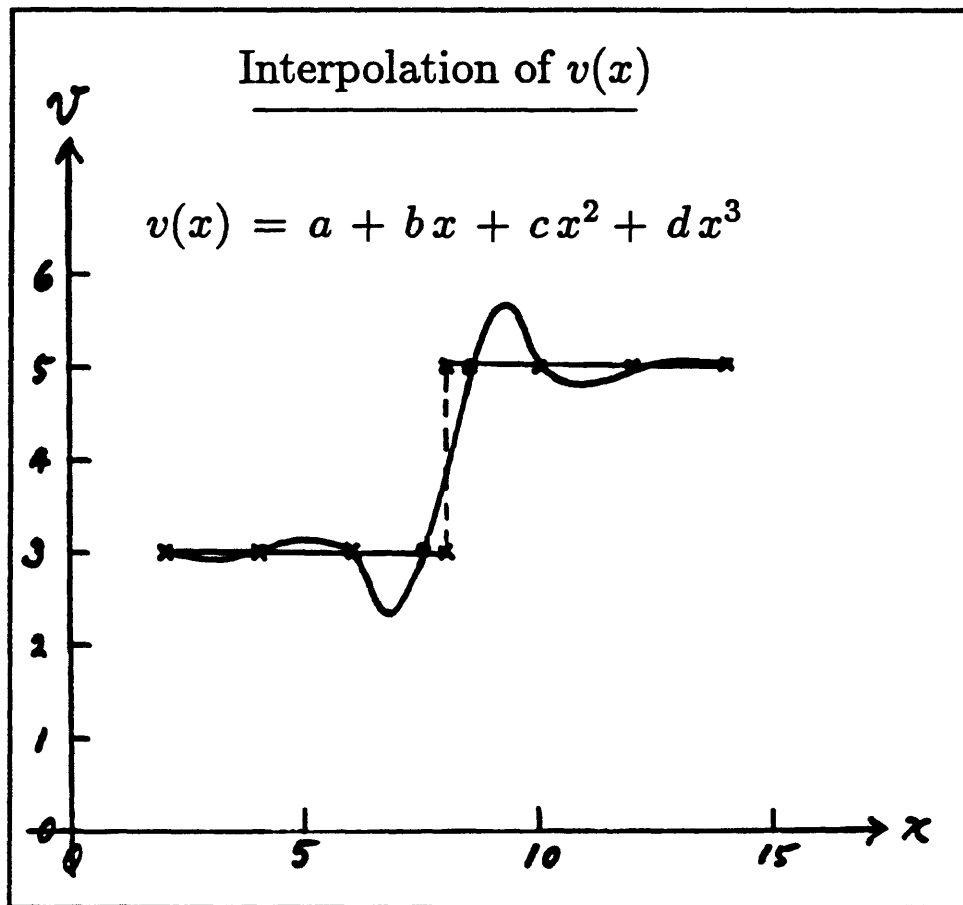


Figure 21. Interpolation problems.

Basically, there are four methods of solving these systems of differential equations. These are finite difference, collocation, initial value, and finite element. We use the first three of these in our work. The reason we use more than one solution method is so that we can compare answers from the different methods. If they are in reasonable agreement, then we can be much more confident about the answer.

I briefly mentioned that interpolation is a part of this ray tracing scheme. The reason that interpolation is important is that we need to be able to calculate the value of a function and its derivative at arbitrary points, even though the function is only sampled at discrete

points. Figure 21 illustrates the kind of problem that can occur when interpolation is carried out across a discontinuity of some sort. In this case, I am modeling a discontinuity in velocity such as might occur crossing a fault. The figure shows the result of fitting a cubic spline to the discontinuity data. As you can see, there is a significant amount of overshoot at the discontinuity. It turns out that the Hermitian interpolation scheme does not have this sort of behavior at discontinuities so we use this form of interpolation in velocity models with discontinuities.

APPLICATION OF RAY TRACING PROGRAM TO SOME SIMPLE MODELS

Figure 22 shows a schematic diagram of a simple fault model. This is a two-dimensional model. There are three horizontal layers with different velocities and a velocity discontinuity in all of the layers at the fault.

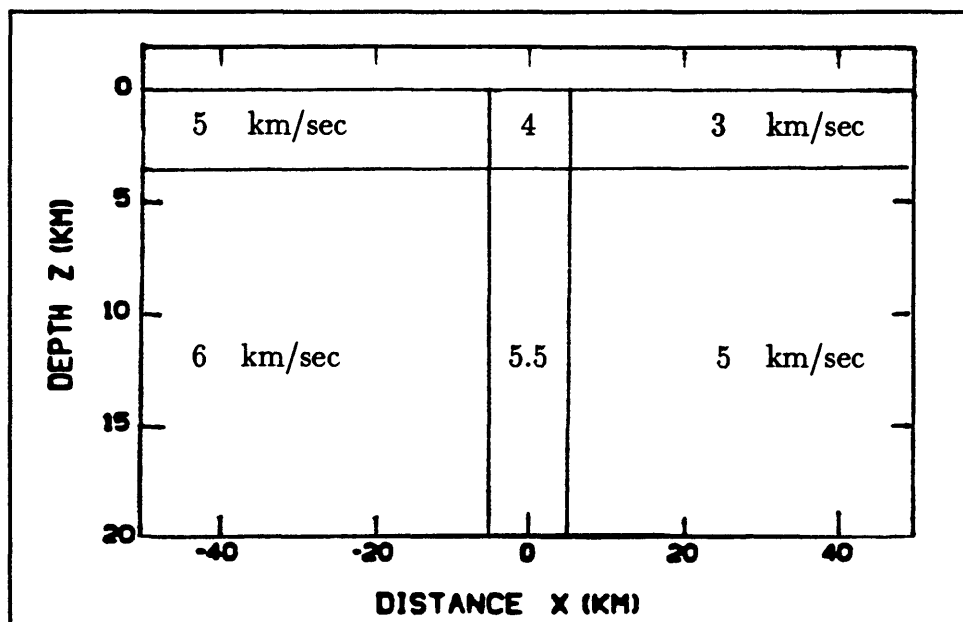


Figure 22. A simple fault model.

Figure 23 shows the ray paths for an earthquake occurring at a depth of 10 kilometers (about halfway through the model). There is a concentration of rays on the right hand side of the fault. This is due to a focusing effect associated with the velocity discontinuity. It is also possible to calculate the expected amplitude of seismic energy arriving at the surface for the event shown. A plot of this is given in Figure 24. As expected, it shows a maximum where the ray density is highest in Figure 23, right on the right-hand edge of the fault. There is also a peak in the amplitude east of the fault.

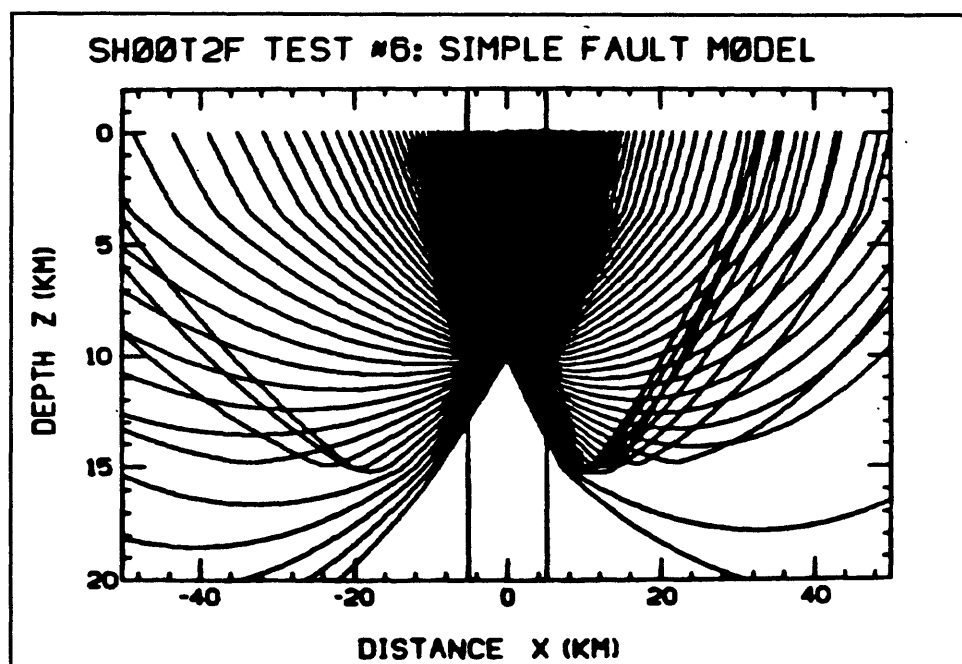


Figure 23. Ray paths for the simple fault model.

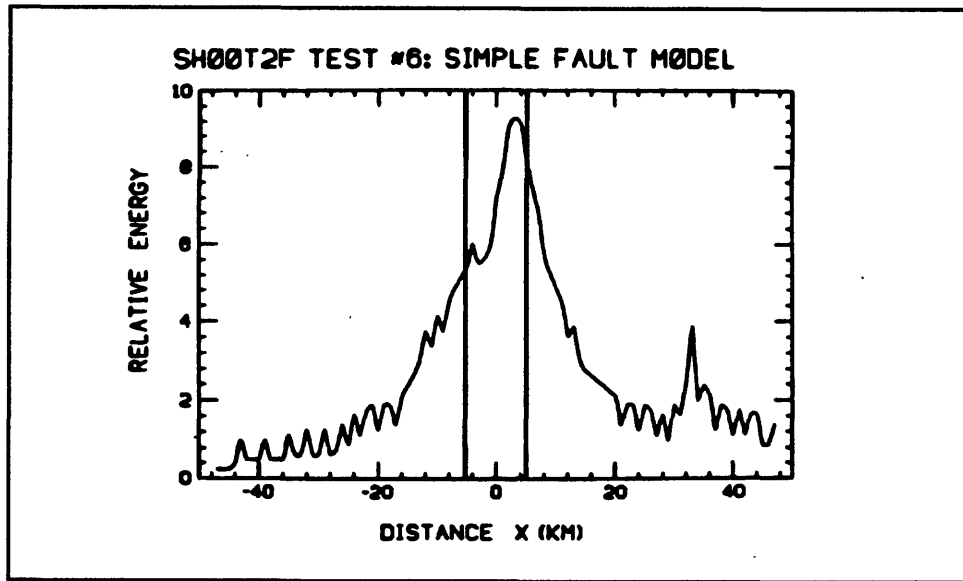


Figure 24. Expected amplitude of seismic waves.

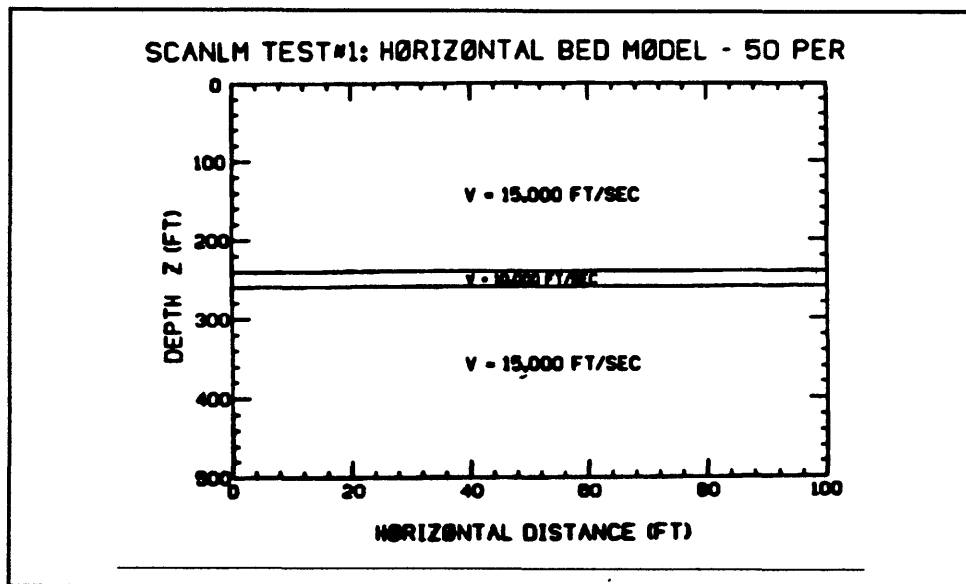


Figure 25. Model LVZ (horizontal layers).

Figure 25 shows a model created for what was essentially a cross-hole tomography experiment. This relates to a project undertaken by Los Alamos Laboratories. They drilled two holes about 100 feet apart in a region where they had very good control on the geology.

The researchers were interested to see if they could use information from seismic waves to infer the geologic structure between the two holes. I developed this model as part of some theoretical work I did in conjunction with that effort. The model consists of three horizontal layers, with the center layer a low velocity zone.

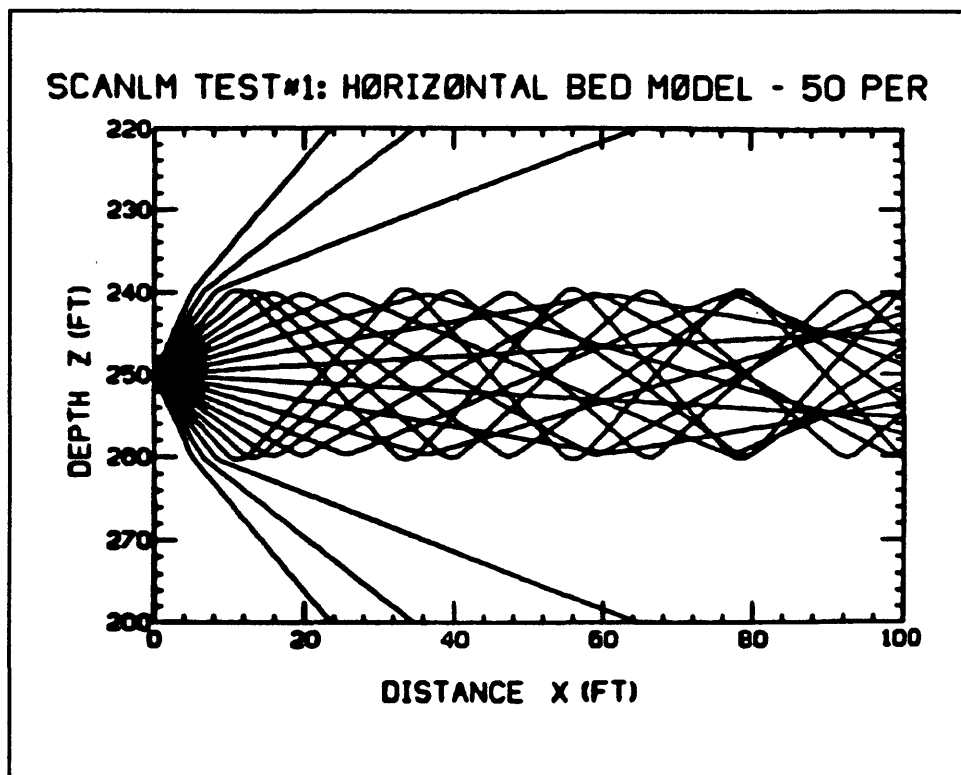


Figure 26. Ray paths for low velocity zone.

Figure 26 shows some ray paths for a small region about the low velocity zone. As expected, all rays which initially contact the interface at an angle greater than the critical angle are trapped inside the low velocity zone. It acts as a wave guide between the two bore holes. This example was really just a demonstration model. The actual geology in the region separating the two bore holes was much more complicated. The geology could be modeled as horizontal layers but there was a reverse fault dipping toward the east in the region between

the bore holes. A model of the actual geology is shown in Figure 27.

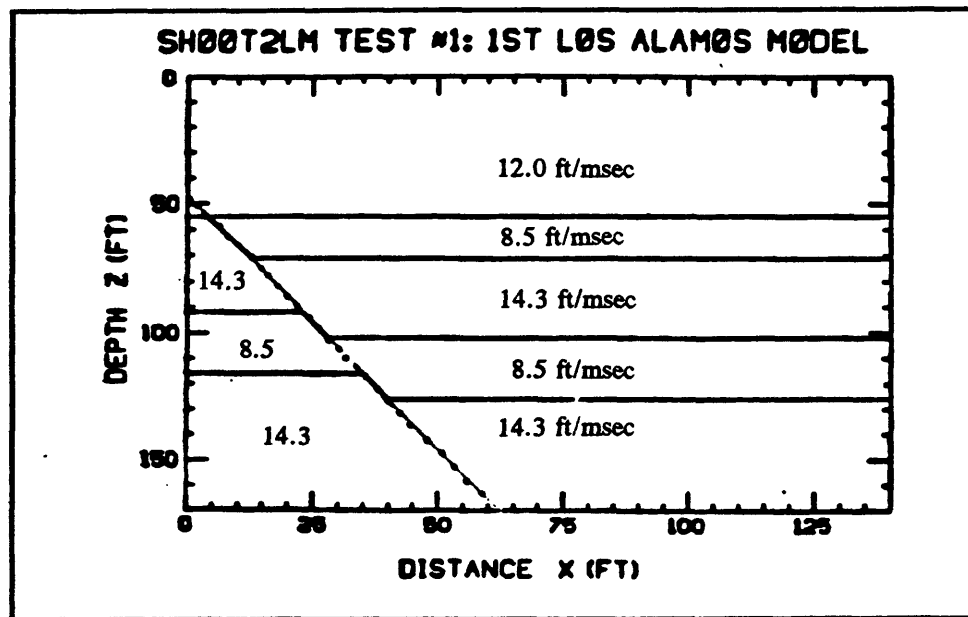


Figure 27. Los Alamos model.

The plan for the experiment was to use seismic sources located in the right-hand bore hole at a spacing of about 25 feet and have a similar array of sensors in the left-hand bore hole. The information collected by the sensors could then be used to try to reconstruct the geology of the region separating the two bore holes. Rather than demonstrate the complete results of this experiment, I will show the expected ray paths for a source located in the lower of the two low velocity zones. This particular set of ray paths is interesting because the low velocity zone is offset by the fault near the receiver bore hole. Since we already know that the low velocity zone (LVZ) acts as a wave guide, it is interesting to see the effect of, in effect, pinching and bending the wave guide.

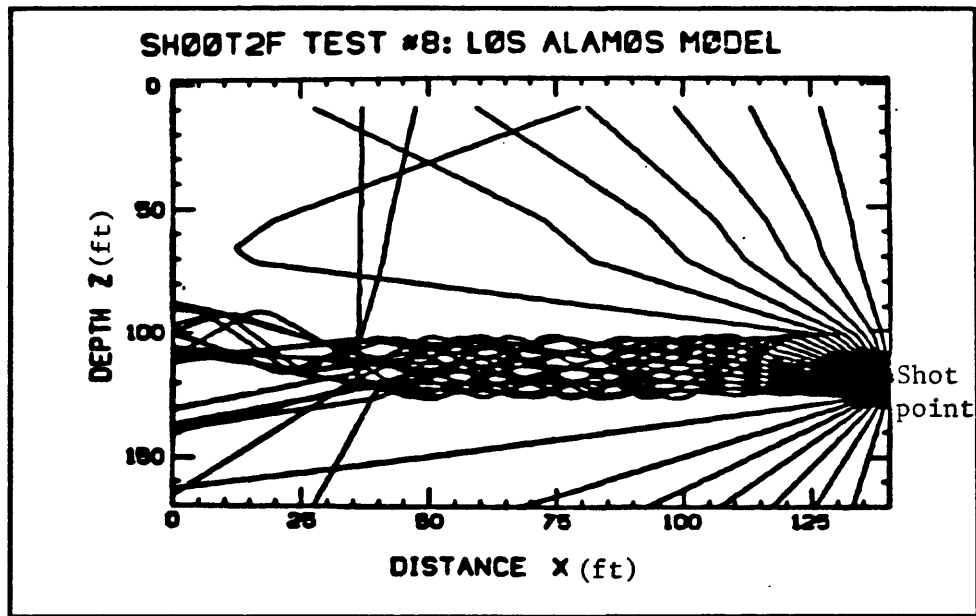


Figure 28. Ray paths through the lower part of the Los Alamos model.

As you can see, most of the energy does, in fact, continue to follow the offset LVZ across the fault although there is some leakage there. Also, some of the rays which left the source at less than the critical angle are reflected at the fault.

REFERENCES

- Eaton, J. P. (1969), HYPOLAYR, a computer program for determining hypocenters of local earthquakes in an earth consisting of uniform flat layers over a half-space. Geol. Surv. Open-File Rep. 155.
- Eliseevnin, V. A. (1965). Analysis of rays propagating in an inhomogeneous medium. Sov. Phys. Acoust. (Engl. Transl.), 10, 242-245.
- Jacob, K. H. (1970). Three dimensional ray tracing in a laterally heterogenous spherical earth. J. Geophys. Res. 75, 6675-6689.
- Julian, B. R. (1970). "Ray Tracing in Arbitrarily Heterogenous Media," Tech. Note 1970-45. Lincoln Lab., Massachusetts Institute of Technology, Cambridge.
- Julian, B. R., and Gubbins, D. (1977). Three-dimensional seismic ray tracing. J. Geophys., 43, 95-113.
- Lee, W. H. K., Luk, F., and W. D. Mooney (1984). Applications of seismic ray-tracing techniques to the study of earthquake focal regions in "Inverse Problems of Acoustic and Elastic Waves", Santoza, F., Pao, Y. H., Symes, W. W., and C. Holland (Editors) SIAM, Philadelphia.
- Lee, W. H. K., and S. W. Stewart, (1981). "Principles and Applications of Microearthquake Networks." Academic Press, New York.
- Pereyra, V., Lee, W. H. K., and H. B. Keller (1980). Solving two point seismic ray-tracing problems in a heterogenous medium. Part 1. A general adaptive finite difference method. Bull. Seismol. Soc. Am., 70, 79-99.
- Wesson, R. L. (1971). Travel-time inversion for laterally inhomogeneous crustal velocity models. Bull. Seismol. Soc. Am., 61, 729-746.

13. DEVELOPMENT OF EARTHQUAKE MAGNITUDE SCALES

by

J. P. Eaton

U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

Assigning magnitudes to earthquakes is something that anybody who deals with networks must cope with. The concept of magnitude was developed by Richter at the time of the development of the first practical local network, the one that was established in the late 1920's in southern California. The network was based on the newly developed Wood-Anderson seismographs. As soon as they began to work up the records from that local earthquake data set using this uniform set of instruments, they recognized that in order to talk about earthquakes, you had to have some way of characterizing their size.

The principal problem we have in characterizing earthquakes is the tremendous range in size of the earthquakes that we are interested in. Recorded up close, one can detect an event equivalent to about 50 pounds of dynamite. We would like to locate that and be able to say how big it is. On the other extreme, we have earthquakes with energy that far exceeds any man-made explosion. One problem is simply the large range in size as measured by any parameter like amplitude or duration. The other problem is setting up a scale that tries to accommodate this range in size. The issue is whether earthquakes remain similar enough from very small to very large so that any single parameter you might choose to measure size holds throughout that range. As it turned out, the measure developed by Richter is not very good when you go to very large earthquakes. But for the small earthquakes in the range from

about magnitude 1 to magnitude 4 or 5, some of the assumptions that Richter made, in order to get on with the work, seem to be borne out well enough.

My plan, in this talk, is to first present an outline of the topics I will discuss today, and then I will go through the historical development of the amplitude magnitude scale. Then I will touch very briefly on the origin and development of the duration magnitude scale, the one used in central California. Finally, I will cover, in summary form, a paper that I have been involved in writing for the last couple of years, which tries to tie together both amplitude and duration magnitude to show how they are related to one-another and also to set up procedures for calculating those magnitudes practically from the instruments that we are running in the northern California network. Things that we will touch on include the early developments due to Richter and limitations of the original magnitude scales. Our current problem is determining magnitudes from short-period electromagnetic seismographs. Finally, we will talk about how to implement a magnitude determination system.

EARLY DEVELOPMENTS

Richter initially set up the magnitude scale to describe earthquakes in southern California. Given the characteristics of the network that he had, and the Wood-Anderson instruments that he used, he covered the range of magnitudes from about 2.5 to about 6. The earthquakes he analyzed had epicentral distances of less than 600 kilometers. Magnitude turned out to be such a useful concept that his co-worker, Gutenberg became involved and helped Richter extend the concept of magnitude, empirically, to work with other earthquakes.

They first extended the method to teleseisms where the amplitude that they measured

was that of surface wave trains from shallow focus earthquakes. The maximum amplitude waves of this type commonly had periods of about 20 seconds. Then Gutenberg and Richter extended the method to body waves, the P-, S-, and multiply reflected or converted P- and S- waves, as recorded on a variety of instruments, for very large, shallow earthquakes. Finally,

EARLY DEVELOPMENTS

- Richter Defines M_L for Wood-Anderson Seismographs in Southern California (Magnitude Range $M_{2.5}$ to M_6 and Epicentral Distance less than 600 km)
- Gutenberg and Richter Develop Empirical Methods to Compute Magnitudes for Teleseisms from
 - 1) 20^{\pm} sec surface waves of shallow quakes
 - 2) body waves of shallow quakes
 - 3) body waves of deep quakes

Figure 1 Early developments.

they took up the problem of determining the magnitudes of deep-focus earthquakes. There, they based the magnitude estimate on measurements of the amplitude of body waves in a manner derived from the shallow-focus method. All of these extensions to the original local magnitude scale are empirical in nature.

LIMITATIONS OF ORIGINAL SCALES

It turns out that there are some problems that have been encountered with the magnitude scales. For large earthquakes, the body wave magnitude scale saturates. No matter how large the earthquake is, the magnitudes computed from body waves tend not to get much above 6 to 6.5. The surface wave scale is less affected by this problem but for very large earthquakes like the major subduction zone earthquakes, the surface wave scale saturates. This happens somewhere around magnitude 8. Typically we think of the largest earthquakes

as having magnitudes of about 8.3 to 8.5. But, when you look at the seismograms, it becomes clear that among that set of very large earthquakes, there are some, like the Chilean earthquake in 1960 and the Alaskan earthquake in 1964 that are very much larger. It turns out that the limitation is in the instrument recording the earthquake. A significant amount of

the energy of these very large earthquakes is at periods which are much longer than the instruments are capable of responding to. As new, longer period instruments have come into use, it has become possible to assign a better measure of the size of these very large earthquakes using the moment magnitude scale.

Another problem with the original procedure as applied to local earthquakes has to do with the gain of the Wood-Anderson instruments. These instruments have a relatively low gain, are available only in horizontal component form, and are not designed for use with telemetered systems. The low magnification means that you must work with earthquakes larger than about 2.5. Finally, M_L is defined in terms of the Wood-Anderson records which can't be manipulated very well. There probably are fewer than 20 of those instruments existing in the state of California compared to about 500 electromagnetic instruments. Clearly, we have to liberate ourselves from some of these limitations.

LIMITATIONS OF ORIGINAL SCALES

- Short Period Body Wave Teleseism Magnitudes Saturated for $M > 6$
- Long Period ($\approx 20\text{sec}$) Surface Wave Teleseism Magnitudes Saturated for $M > 8$
- Wood-Anderson Gain too Low ($V_0 \approx 2000$) to Record Local Earthquakes With $M < 2.5$ Effectively
- M_L Defined in Terms of Wood-Anderson Records

Figure 2 Limitations of original scales.

CALCULATION OF LOCAL EARTHQUAKE MAGNITUDES

There are two approaches to computing magnitudes for a network using high-gain short-period electromagnetic seismographs. One is to continue using amplitude-based magnitude estimates and the other is to compute magnitudes based on duration.

There are two approaches to computing amplitude magnitudes. One is to use the modern instruments and with the knowledge of their response and the response of the Wood-

Anderson instruments, completely transform the records to the equivalent synthetic Wood-Anderson seismogram. Then compute magnitudes using the synthetic seismogram. The other method, which is the one that I have pursued, is to take advantage of the fact that, except for the gain, the response of the electromagnetic instruments is quite similar to that of the Wood-Anderson instrument. The principal difference is a 6 dB per octave increase in magnification with increasing frequency. In this approach, we use the records from our instruments directly without elaborate transformation to compute magnitude in approximately the same way as if we were using Wood-Anderson seismographs.

The other basic approach to computing magnitudes takes advantage of the fact that as

CALCULATION OF LOCAL EARTHQUAKE MAGNITUDES FROM HIGH-GAIN SHORT-PERIOD ELECTROMAGNETIC SEISMOGRAPHS

- Amplitude Magnitudes
 M_l from Synthetic Wood-Anderson
MX
- Duration Magnitudes
FMAG (Lee)
MF

Figure 3 Calculating local earthquake magnitudes.

earthquakes get larger, the duration of the records from the onset of the P-phase until the signal drops below some threshold also increases with the size of the earthquake. A variety of people working around the world worked out the systematics of this and developed a scheme of determining magnitudes based on duration. This is a secondary scale since it is derived in such a way as to match the M_L scale for a certain range of earthquakes.

IMPLEMENTATION

In terms of implementing these procedures, the first task is to work out the relation between the amplitude of the zero magnitude earthquake and distance. The second task is to work out the duration of the zero magnitude earthquake versus distance in order to implement the duration magnitude.

These need to be done using data from

the region where you need to compute magnitudes. Once you have worked out those relationships, it turns out that there are a number of other corrections that have to be added to what you calculate from your records. One correction is for the different site responses. Richter recognized this problem and worked out a set of site corrections with values up to about $\pm (.2 \text{ to } .3)$ units.

Richter was dealing only with Wood-Anderson seismographs. They were all alike and

IMPLEMENTATION

- Determination of $\log A_0$ vs R and $\log T_0$ vs D Equations from the Northern California Data Set
- Determination of MX and MF Site Corrections, Sensitivity Corrections, and Component Corrections
- Dependence of MX and MF Site Residuals on Rock Type

Figure 4 Implementation of magnitude scales.

all had the same magnification. We use a variety of different instruments with different magnifications so we have to worry about corrections which compensate for different instrumental sensitivities. In the case of amplitude magnitude, that is simple since we know the gain of each of our instruments and we know the gain of the Wood-Anderson instrument. However, for duration magnitude, the problem is not so simple because the implementation of this magnitude measure involves setting a threshold which determines the end of an event coda. Using instruments of different sensitivity results in a different duration, hence different magnitude.

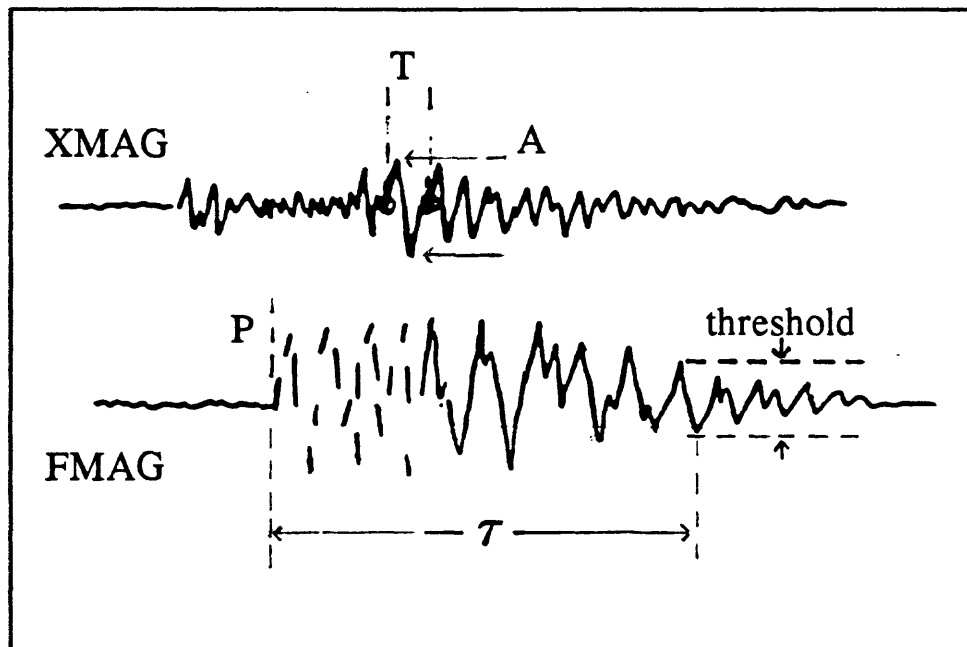


Figure 5 Measurements for magnitude determination.

Local magnitude was defined by Richter on the basis of Wood-Anderson records which were only for the horizontal component. In the modern networks, most of the instruments are vertical component. This is because the vertical component usually has the sharpest P-wave onset, which is required for locating earthquakes. But for a given

earthquake, the maximum amplitude recorded on the vertical component will be different than on a horizontal component so we must have a correction factor for the component used.

As Richter defined magnitude, there are three critical elements involved in the measurement. One is that the amplitude measured must be the maximum recorded for that event. This was the peak to trough distance divided by two as shown at (A) in Figure 5. When we use modern electromagnetic instruments to measure amplitude, we also must measure the period T in order to be able to convert the measurement to an equivalent Wood-Anderson amplitude.

The second critical element is the specification that the recording be made on a standard Wood-Anderson horizontal component seismograph with free period of 0.8 second, damping equal to 0.8 of critical, and measured static magnification of 2100.

The third critical element involves the standard recording distance of 100 kilometers. This definition is independent of what practical steps you must use to measure magnitude when the standard conditions are not met. To permit his definition to be implemented, Richter worked out an empirical relation between measured amplitude and distance. This is a vital tool which you must have in order to be able to calculate magnitude.

In the duration magnitude approach (FMAG in Figure 5), much of the record is clipped so you wait until the record comes back on scale. Then you measure the time from the first onset of the P-wave until the signal amplitude drops below the threshold. Then that time (T) is used to compute the magnitude. The reason we adopted this method for our network was that the deconvoluted records that we used for so long only rarely would allow you to actually see the maximum amplitude.

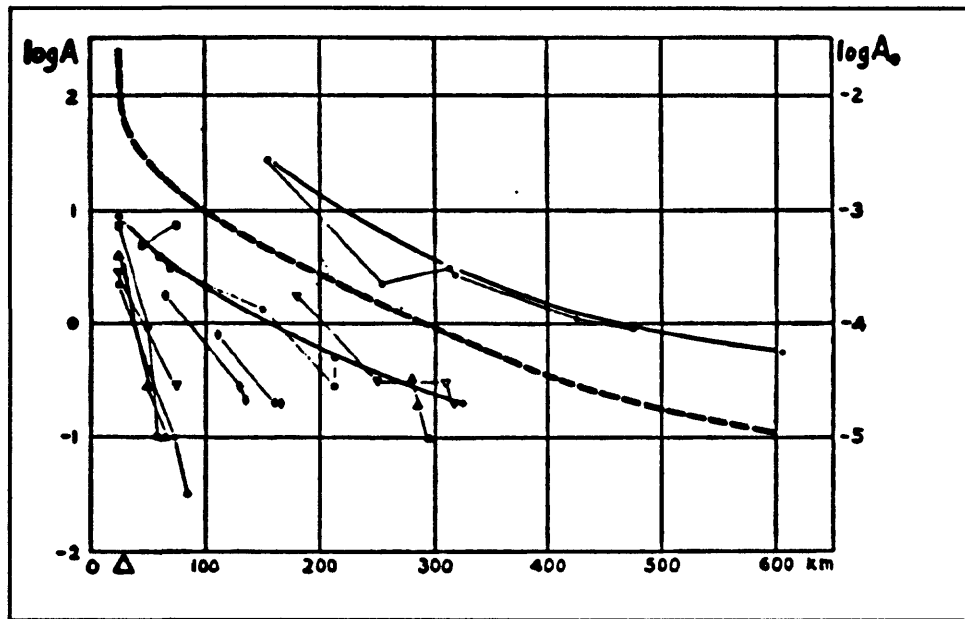


Figure 6 Origin of the magnitude scale.

Figure 6 is an illustration from Richter's original paper. It is a plot of log amplitude versus distance for a group of individual earthquakes. In this plot, each earthquake is represented by a series of measurements connected by lines. What Richter noticed, was that for a given distance, the slopes were always constant. This meant that you could construct a general reference curve which was always parallel to the slope of the lines connecting observations. In Richter's definition, the standard earthquake had an amplitude of one micron at a distance of 100 kilometers.

Figure 7 is a plot of log amplitude versus log distance. In this plot, the curve breaks into two parts. Above 200 kilometers, the slope of the line is -3.0 and at distances less than 200 kilometers, the slope of the line is -1.6. Actually, there are systematic variations about the two lines, but they still represent a pretty good fit. In a modern application where you would like to be able to compute magnitudes automatically, it is necessary to have some

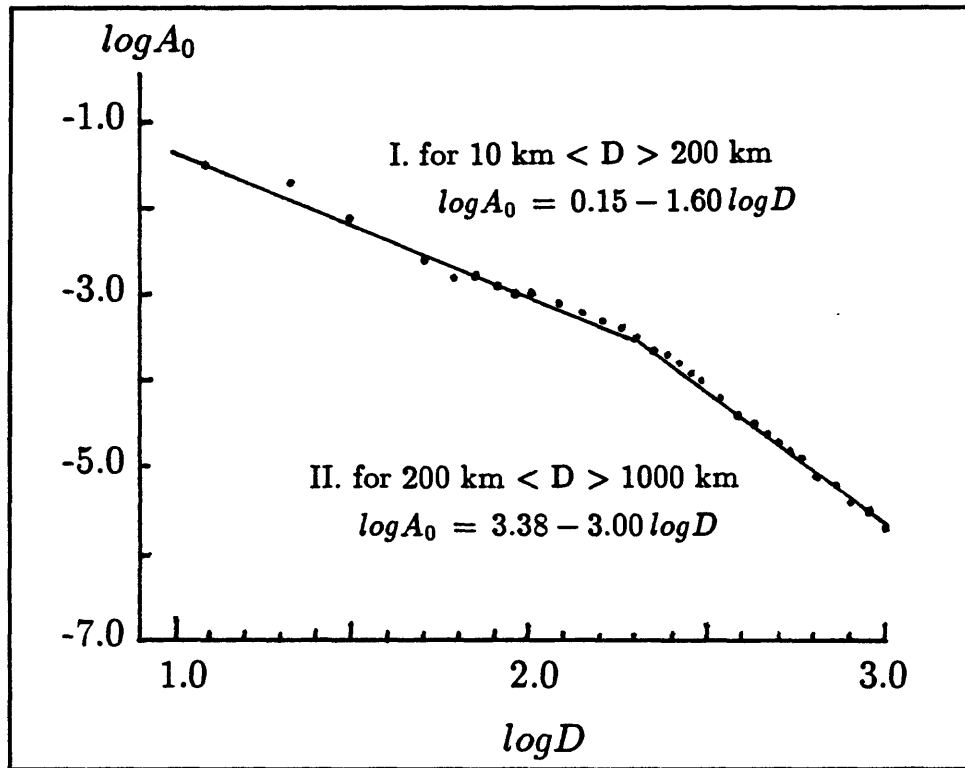


Figure 7 $\log A_0$ vs $\log D$ (Richter: $\log A_0$ vs $\log \Delta$, with $h = 7 \text{ km}$)

means of doing these computations numerically. For many years we calculated equivalent Wood-Anderson magnitudes using the linear approximation shown in Figure 7.

Figure 8 shows a variety of earthquakes recorded by our network. The top trace is of a very small (M 1.5) earthquake recorded only about 10 kilometers away. The second trace is from a bigger earthquake from the same site but at an epicentral distance of 112 kilometers. Both of these traces are recorded on vertical component instruments. The third and fourth traces are from a magnitude 2.8 earthquake. This is the magnitude range where the original Wood-Anderson recordings began to be usable. The third trace is the vertical component and the bottom trace is the north-south component. Note the difference in

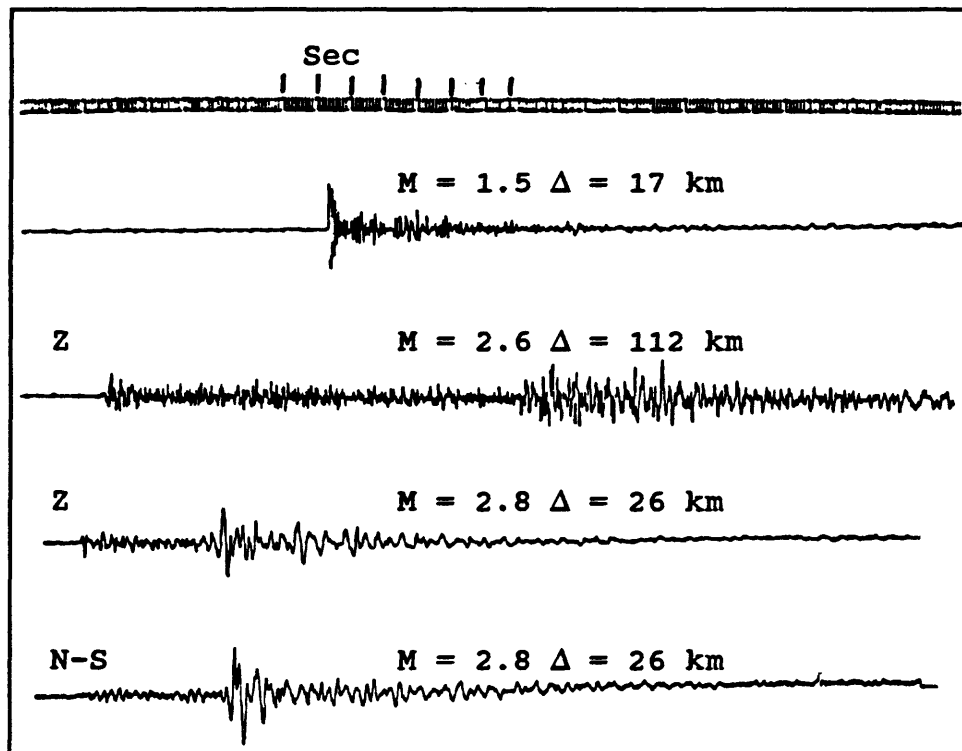


Figure 8 Earthquakes recorded by Calnet.

character between the two records. The P-wave onset is much sharper on the vertical component and the maximum amplitude, which occurs during the S-phase is substantially larger on the horizontal than it is on the vertical. For a magnitude 2.8 earthquake, the total length of the record is less than 30 seconds.

In extrapolating magnitude scales to large events, they used records such as the one shown in Figure 9. This is a record taken by Galitsin. His recorders ran backwards so the record is read from the upper right-hand corner to the left side, and so on. The dots are one minute apart, so the records in Figure 8 would fit between one pair of dots. Clearly, this is a much bigger event and it takes a major leap of faith to attempt to use the same scheme that was used for local events on these very large events. For these large amplitude teleseisms,

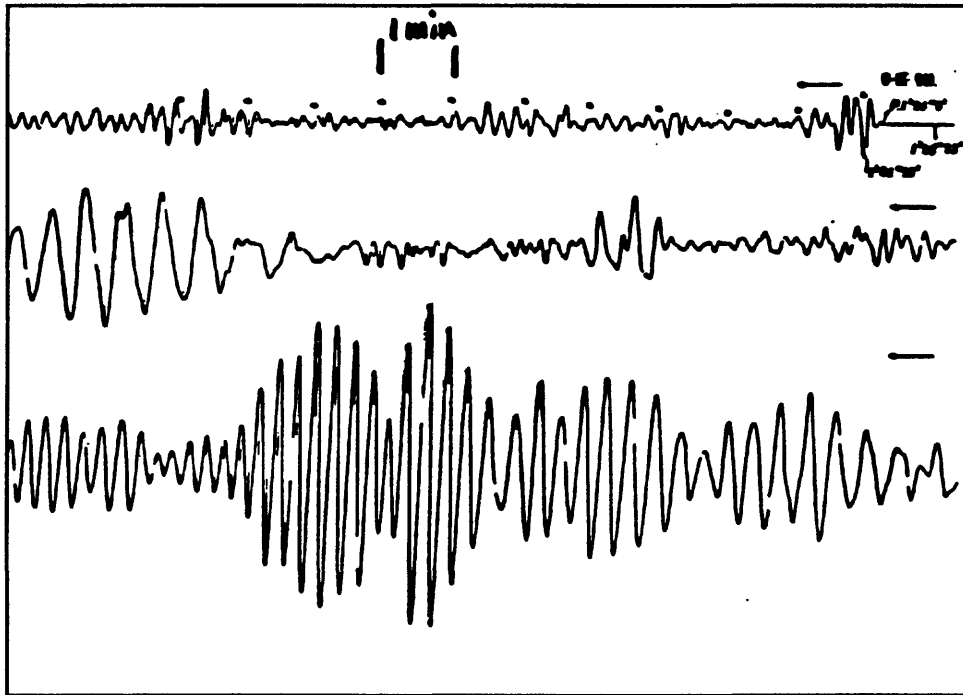


Figure 9 Teleseism recorded by Galitsin.

the maximum amplitude occurs in the surface wave train, usually at a period of about 20 seconds. Since the Wood-Andersons don't have good sensitivity at these periods, it was necessary to use other instruments with different magnification curves. The data taken this way was simply reduced to an equivalent Wood-Anderson record for the Magnitude calculation.

For years this was pretty much a static procedure. Gutenberg and Richter had worked out the method, and everyone followed it. The procedure is summarized in the nomogram shown in Figure 10. One simply connects the trace amplitude (on the right) and the epicentral distance (on the left) with a straight line and then reads off the magnitude where the line intersects the center scale. This was a practical way that magnitudes were computed for local earthquakes and large shallow earthquakes for many years.

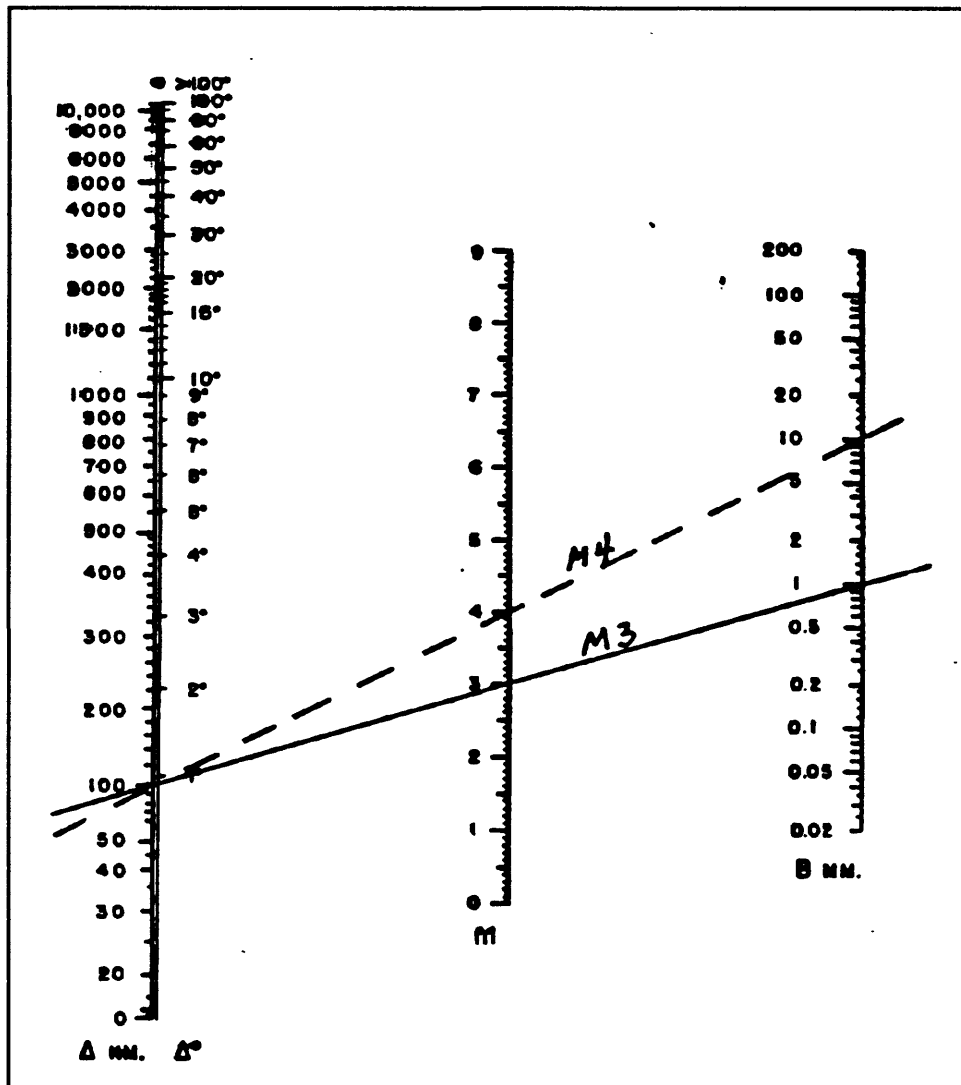


Figure 10 Nomogram for determining earthquake magnitudes.

When we began to develop the dense microearthquake networks, we started dealing with earthquakes down to about magnitude one and we started to see records like that shown in Figure 11. This is a local event with a magnitude somewhere around one. On many of these records, the trace simply disappears when the event starts and doesn't show up again until the amplitude becomes small. From a record like this, there is no possibility of

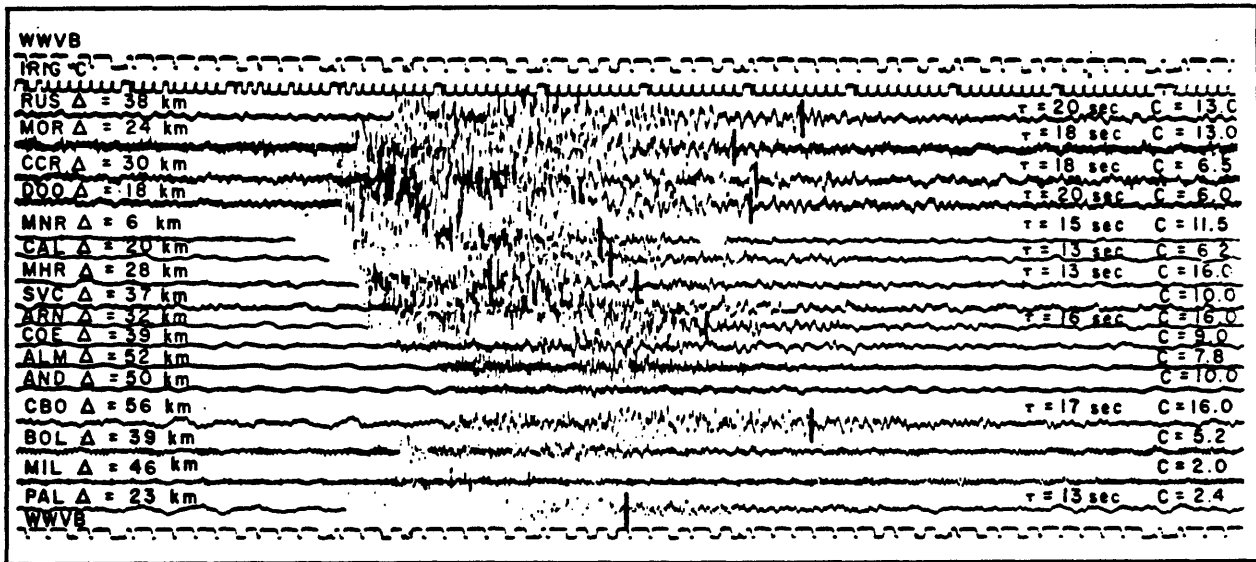


Figure 11 Develocorder record for local event.

measuring the maximum amplitude so we were forced to adopt the duration magnitude method.

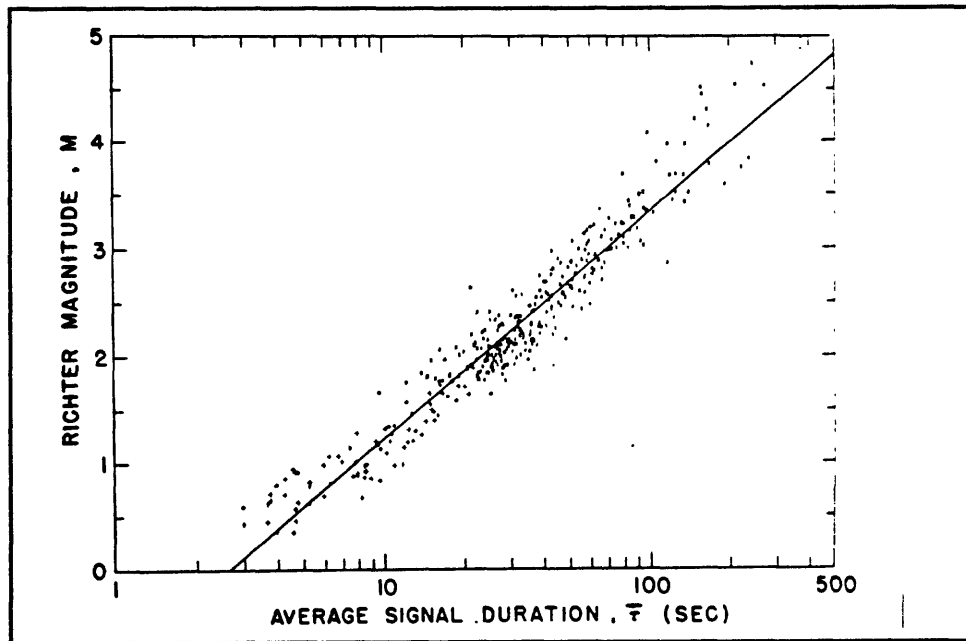


Figure 12 Magnitude versus duration.

Figure 12 shows a plot of Richter magnitude versus signal duration for a group of

earthquakes. As you can see, the data cluster reasonably well around a straight line. This empirical observation forms the basis for the duration magnitude method. The relation, as determined by Lee and others (1972) is:

$$\hat{M} = -0.87 + 2.0 \log T + 0.0035 D \quad (1)$$

where T is the average duration of the event as recorded by several stations and D is the average distance of those stations.

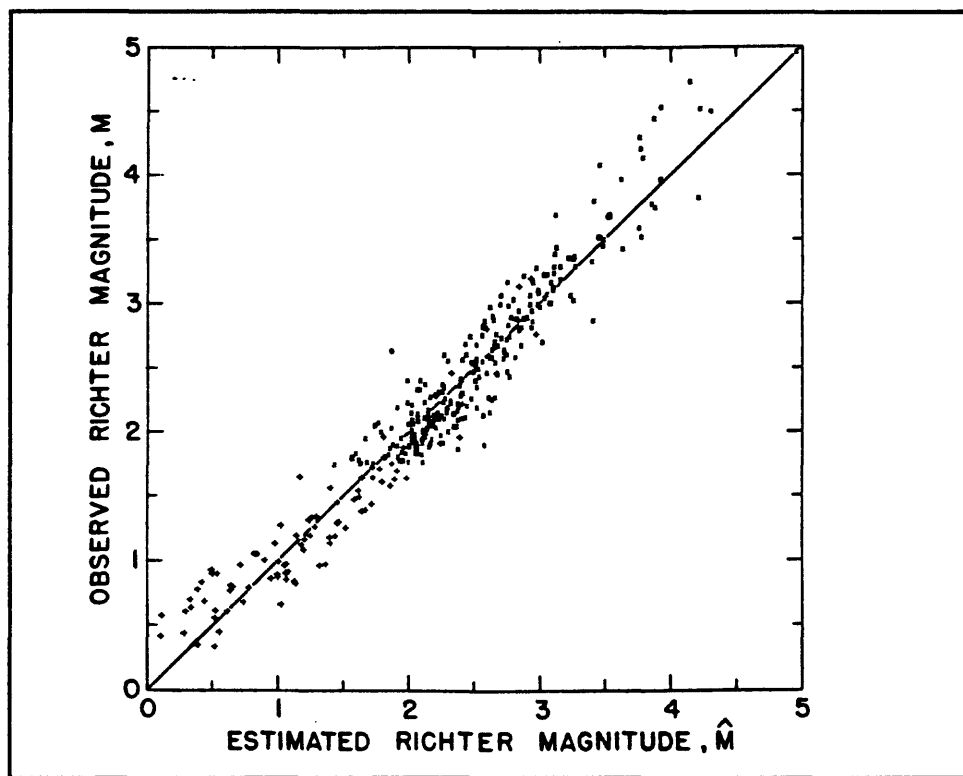


Figure 13 Comparison of Richter magnitude and duration magnitude.

Figure 13 shows a comparison of magnitudes computed using the standard amplitude method to magnitudes computed using the duration measure. Through at least the range from about magnitude 1.5 to magnitude 3.5 there is a reasonably good correspondence between the

two methods. It turns out that for the larger earthquakes, this method tends to underestimate the magnitude. In the process of trying to resolve this problem, Carl Johnson, in southern California, who was working up a method for analyzing earthquakes entirely by computer, did a study of the decay of trace amplitude with time. He characterized the decay in the manner shown in Figure 14. The trace was divided into 5 second intervals, and for each

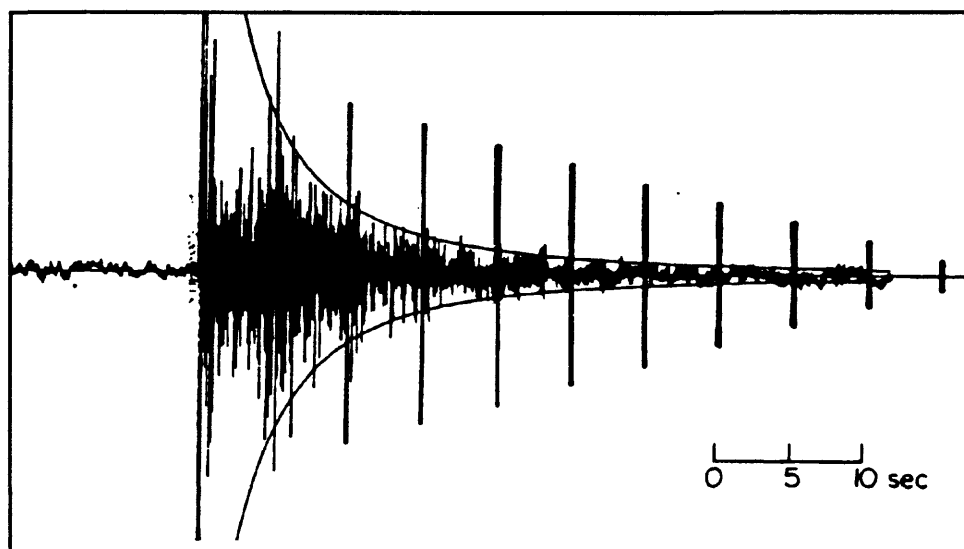


Figure 14 Magnitude estimates over 5 second intervals.

interval, he calculated the average absolute value in each interval and plotted that as a measure of how the amplitudes were decaying. Each 5 second interval produces a quasi-independent magnitude estimate for the event.

Figure 15 is a plot of $\log B$ (from Figure 14) versus \log duration where duration is measured from the P-wave onset. The plot indicates that there is a linear relationship of the form:

$$\log B = \log B_0 - m \log T \quad (2)$$

This relation turns out to be important in handling different sensitivities and different

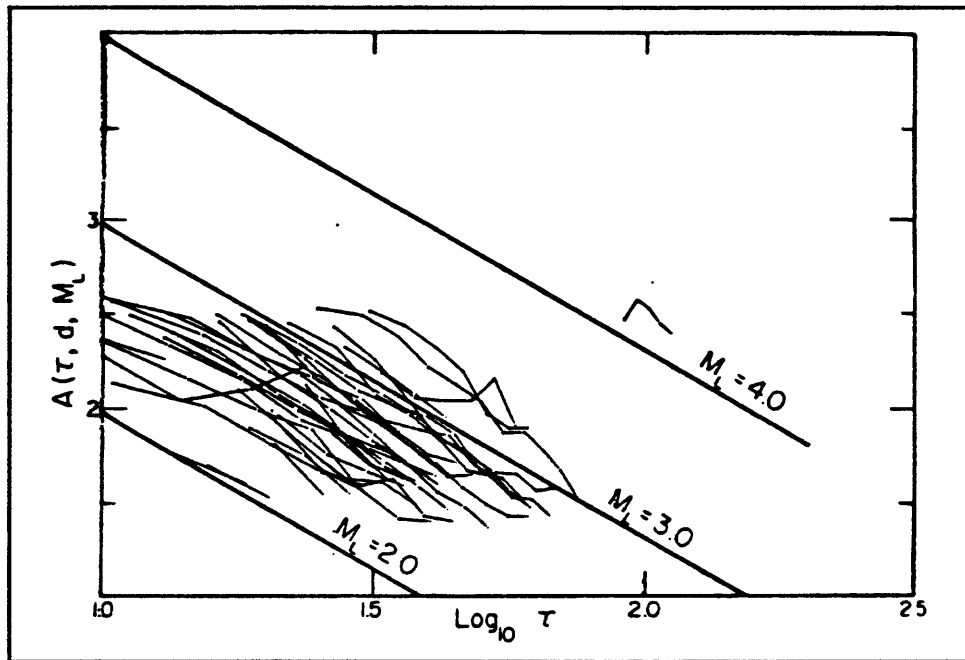


Figure 15 Log B vs log T.

thresholds in measuring duration for duration magnitudes.

Then there is the issue of the difference in response between the Wood-Anderson instrument and the instruments used in our network. The lower part of Figure 16 shows a plot of log magnification versus log frequency for the Wood-Anderson and for one of the USGS instruments operating at fairly high sensitivity. All the curves in this figure are asymptotic approximations to the actual response curves. The Wood-Anderson has flat response at frequencies above its natural frequency of 1.25 Hz. The USGS system has a response proportional to ground velocity above its natural frequency of 1.0 Hz; so, with regard to displacement, its response increases at 6 dB per octave. At higher frequencies, our system cuts off in the discriminator used to recover the telemetered data. The cutoff point is somewhere around 25 to 30 Hz. In practice, the Wood-Anderson records can't be read at

these high frequencies since recording is on paper.

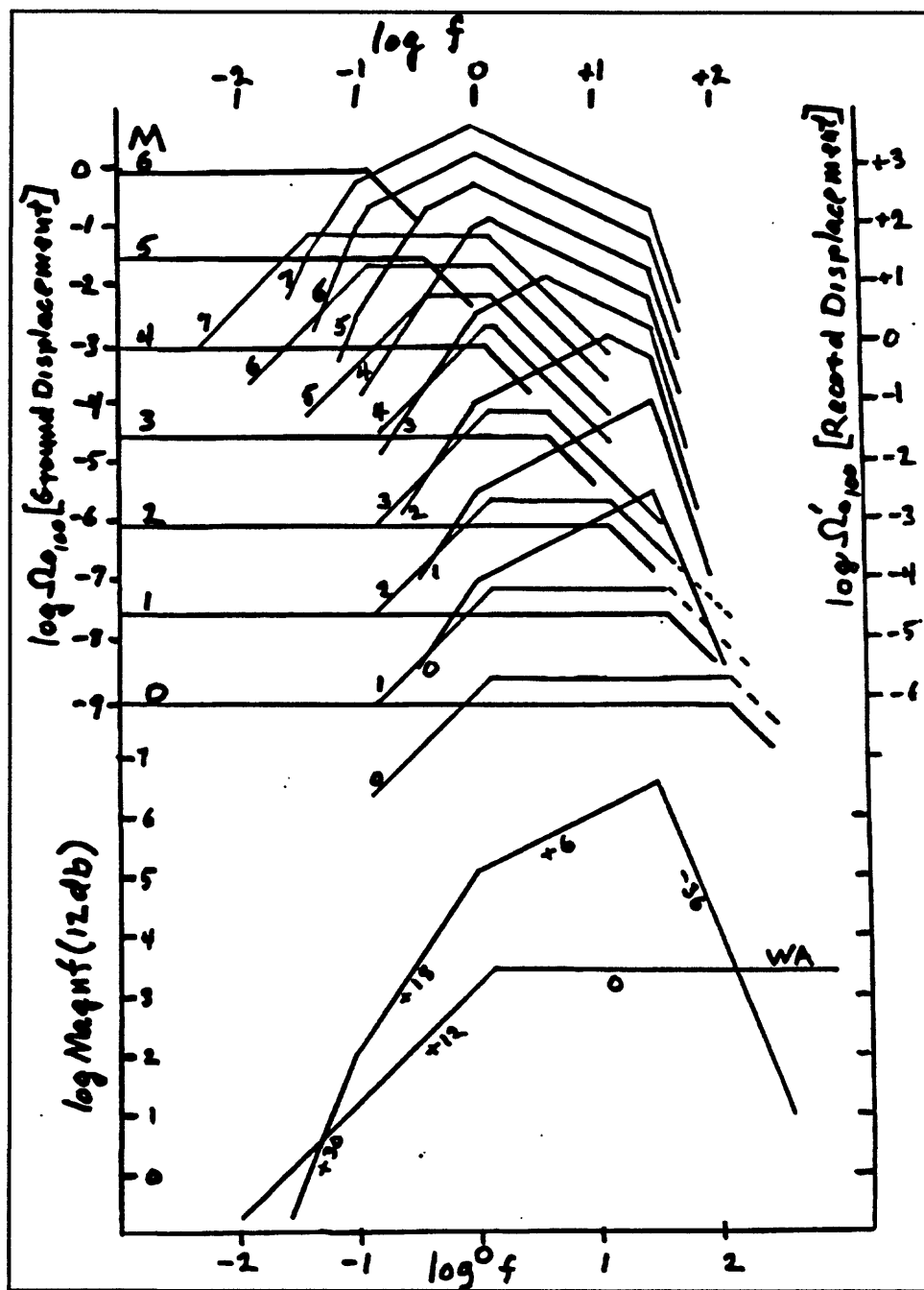


Figure 16 (upper) Displacement spectra vs magnitude and (lower) instrument response from two seismometers.

It is interesting to see what the record of an earthquake is going to look like on these two systems. To do that, I have taken the Brune (1970) spectra of earthquakes of various magnitudes and converted that into the spectra which we would expect to record by adding the response spectra of the instruments to the far-field spectra predicted by the Brune model. The results are shown in the top part of Figure 16. As magnitude increases, the corner frequency decreases so, for the Wood-Anderson instrument, the flat part of the response spectrum diminishes with increasing magnitude. For our system, something analogous happens. However, there is no flat part on the response spectrum, so the measured spectrum is distorted at all magnitudes.

Figure 17 shows what the seismograms would look like (in the frequency domain). It was produced by registering the two sets of curves so that the maxima correspond. The Wood-Anderson performs better at low frequencies for all magnitudes. However, at the higher frequencies, the electromagnetic system has higher output than the Wood-Anderson system. Since the small events have more of their energy concentrated at high frequencies, the electromagnetic seismometers are better suited for recording those small events.

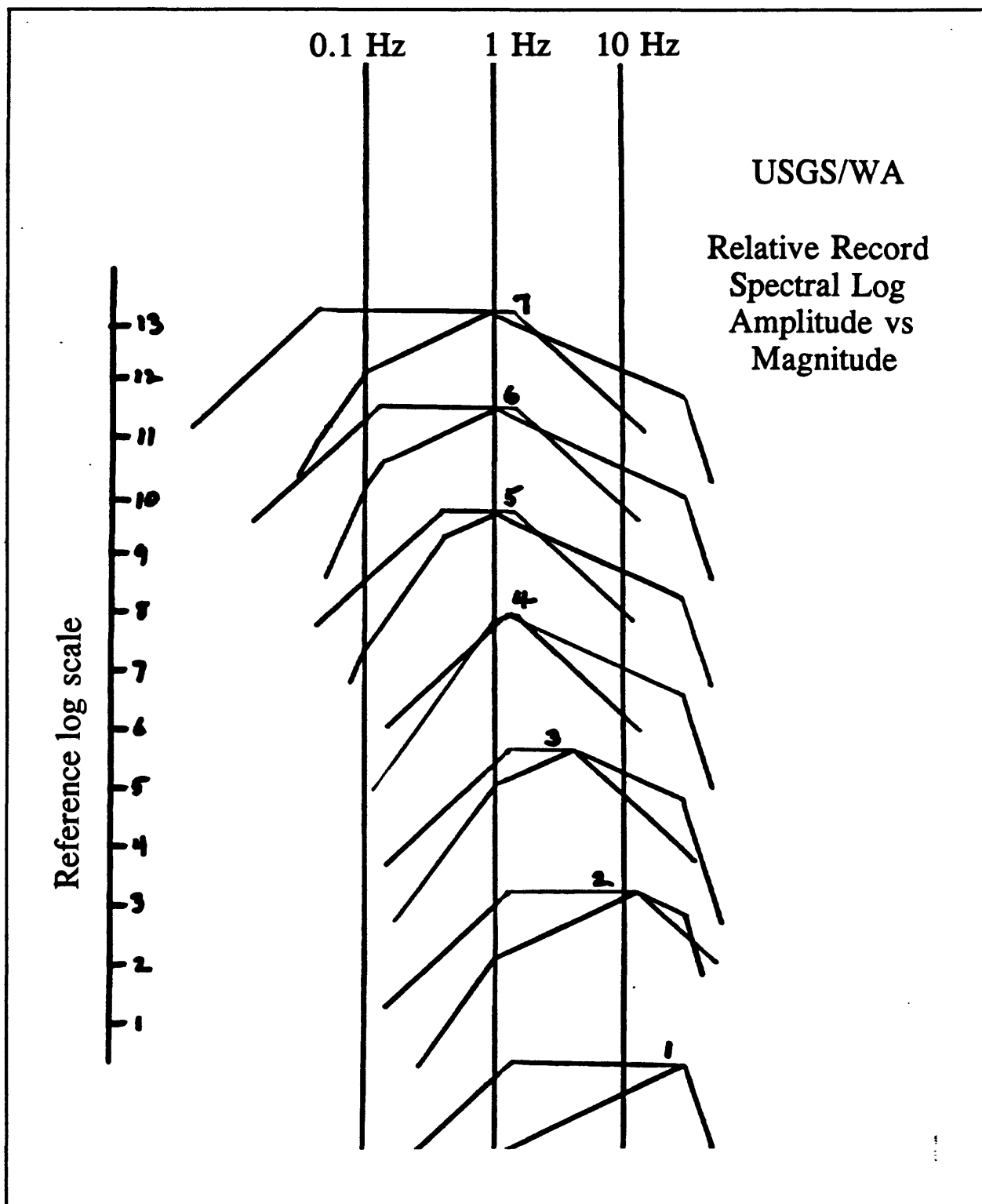


Figure 17 Combined displacement spectra and instrument response spectra.

$$\mathbf{XMAG} = \log(A) + \mathbf{X(D,h)} + \mathbf{SX(T,attn)} + \mathbf{CX(cmp)} + \mathbf{SITX}$$

$$\mathbf{FMAG} = c \cdot \log(\tau) + \mathbf{F(D)} + \mathbf{SF(attn)} + \mathbf{CF(cmp)} + \mathbf{SITF}$$

A = max amplitude (Peak to trough)

T = period associated with A

τ = record duration (measured from P)

c = const. relating FMAG to XMAG

X(D,h) = dependence of A on epicentral distance & h

F(D) = dependence of τ on D

SX(T,attn) = XMAG sensitivity correction

SF(attn) = FMAG sensitivity correction

CX(cmp) = XMAG component correction

CF(cmp) = FMAG component correction

SITX = XMAG site correction

Figure 18 Equations for the two magnitudes.

Figure 18 shows the complete form of the two magnitude equations as we use them today. XMAG is the amplitude magnitude. It contains the log of the amplitude, as Richter would have had, a term which accounts for decrease in amplitude with distance, which was used by Richter, and a site correction, which was also used by Richter. We have two additional terms. The SX term accounts for the sensitivity of the instrument and the CX term is used to compensate for which component is used. The duration magnitude (FMAG) equation is entirely analogous. Log duration substitutes for log amplitude (and has a constant multiplier). The F(D) term represents the decrease in zero magnitude earthquake duration on a standard vertical USGS instrument. There is also a site correction, as in the amplitude magnitude equation. SF is a sensitivity correction analogous to SX and CF is an instrument

component correction term, analogous to CX. What is needed is to establish coefficients for the two equations.

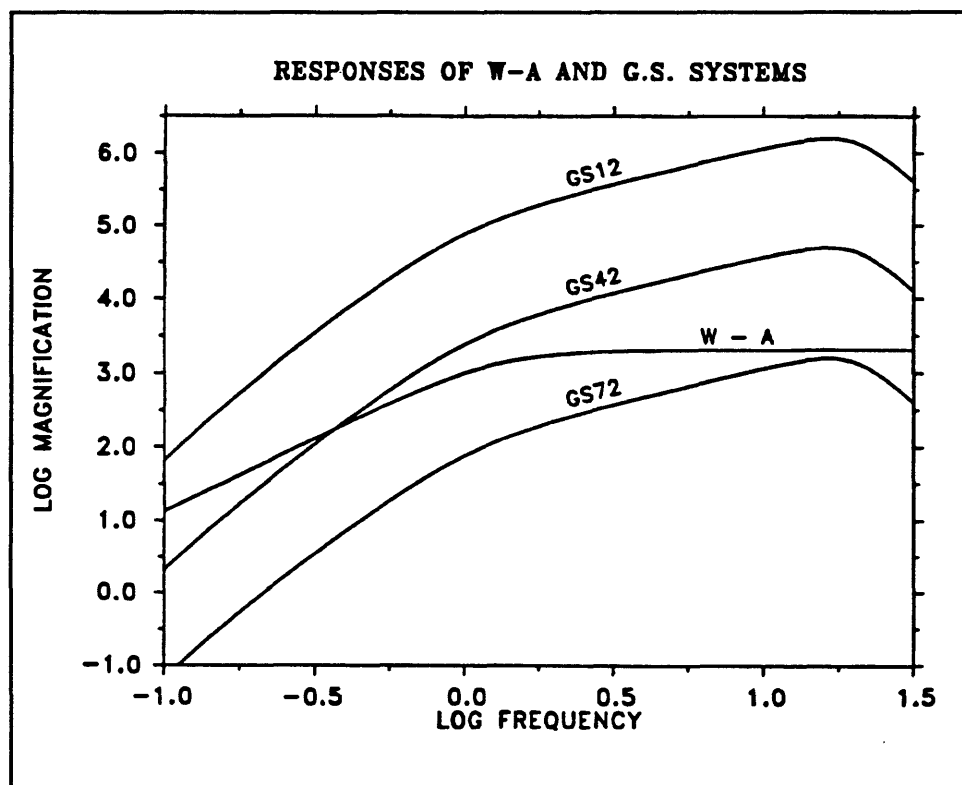


Figure 19 Frequency response of Wood-Anderson and G.S. Systems.

Figure 19 shows actual response curves for a Wood-Anderson seismograph and for some USGS instruments. In our system the responses are characterized by the attenuation relative to the maximum amplification of 92 dB so the GS12 has an overall gain of 80 dB. Most of our high gain stations run at either 12 dB or 18 dB attenuation. In addition to these high gain stations which characteristically, have been set out to run against background noise, we have had some intermediate gain stations which operate at 42 dB. Recently, we have also added a single 3-component station in this building, which goes directly into the

A/D convertor. It runs at 72 dB attenuation from the maximum gain. This system will record a magnitude 5 earthquake on scale if it occurs 50 or 60 kilometers away. So for our system, the sensitivity correction in the magnitude equation must be able to span this range of 60 dB for it to be of any use to us.

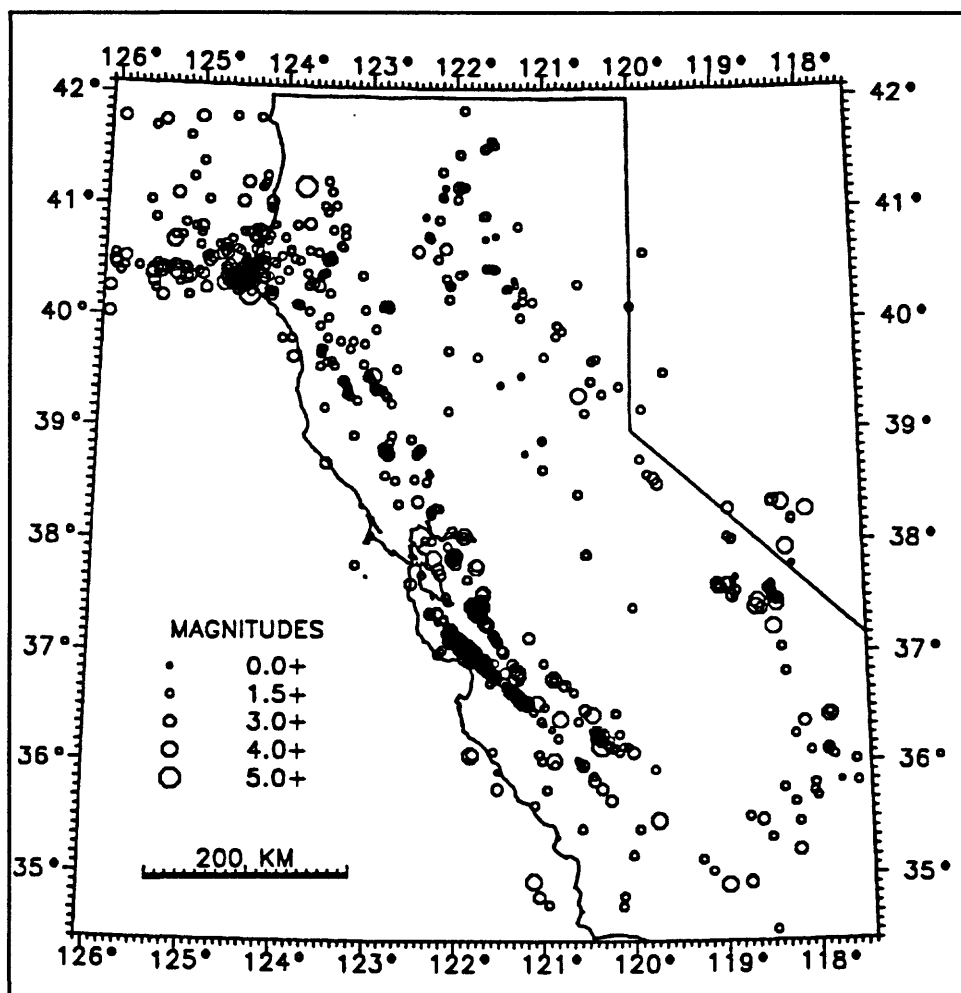


Figure 20 Events used in magnitude equation calibration.

Figure 20 is a plot of the data set, drawn from the northern network which I have used for calibrating the magnitude equations. There are about 1270 earthquakes distributed

throughout the region in this data set. They are mostly larger events although the small events are well represented.

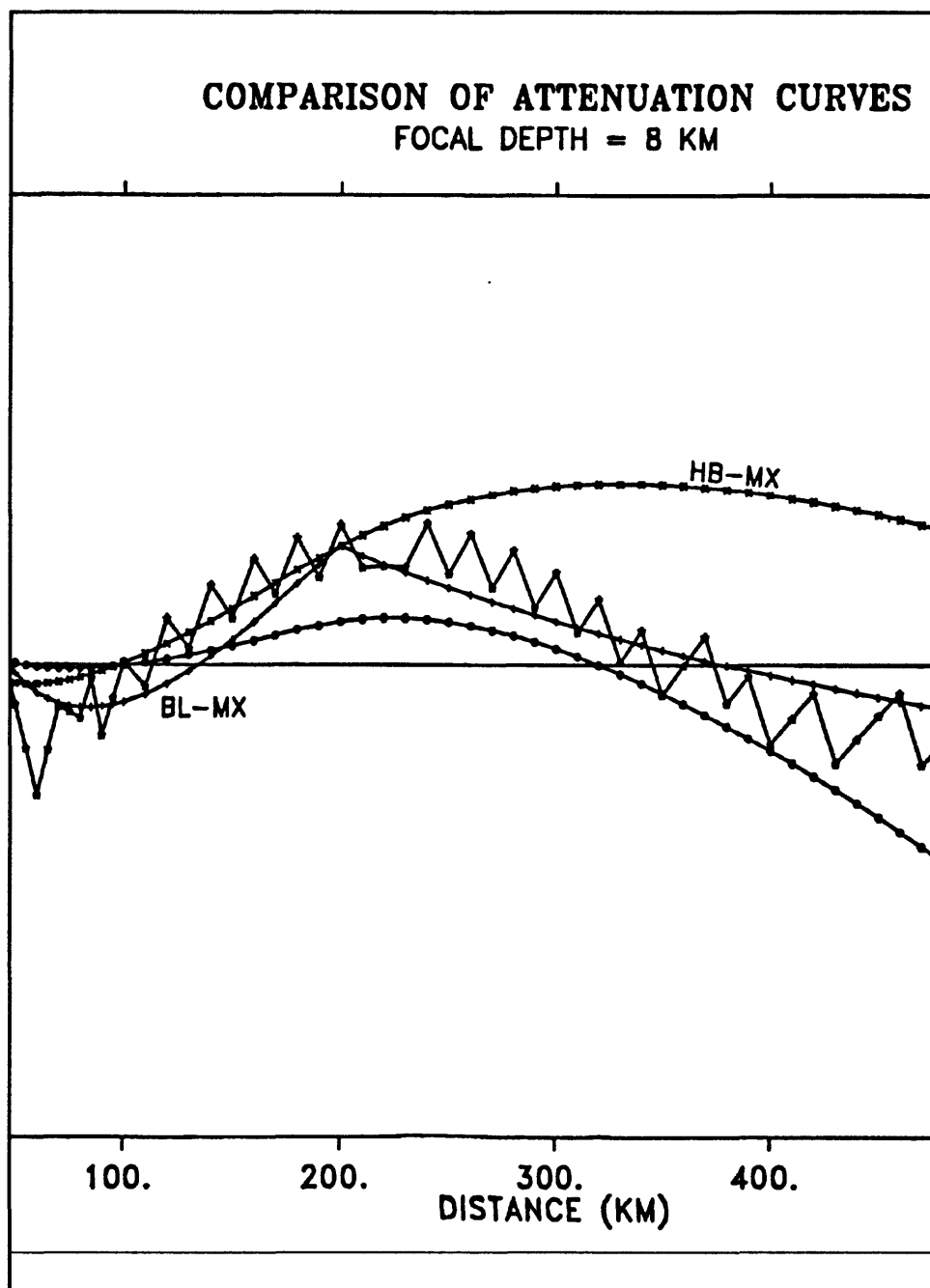


Figure 21 Comparison of attenuation curves.

Figure 21 shows a comparison of attenuation curves developed by different people using different data sets. The comparison is made by taking the difference between the attenuation curve being compared, and an attenuation relation I developed using the northern California data set. The sawtooth curve represents the difference between the original Gutenberg-Richter attenuation curve and my northern California attenuation relation. The curve labeled BJ-MX is a comparison of an attenuation relation due to Bakun and Joyner (1984) with mine. Their relation was developed using real Wood-Anderson records and synthetic Wood-Anderson records derived from our electromagnetic instruments. Their data set cut off at about 400 kilometers, so their attenuation relation is really applicable only for distances out to 400 kilometers. Notice that in the range from zero to 400 kilometers, their attenuation curve matches the Gutenberg-Richter curve quite well. The Bakun and Joyner curve is based on the assumption that all of the attenuation is due to geometric spreading and to absorption along the path. That doesn't account for the fact that the seismogram gets more complicated with distance. Hutton and Boore recently have done the same analysis on a southern California data set. They use the same equations as Bakun and Joyner but with different parameters. They found that they got better fit in southern California with higher geometric spreading and lower absorption. But, when I compare the southern California data set with the northern California data set, it appears that there is a real difference.

There are a variety of equations for duration magnitude. Those turn out to be a little more troublesome to compare since they all involve a distance. In the comparison illustrated in Figure 22, I fixed the distance at 100 kilometers. In the figure, duration magnitude is plotted versus log duration at 100 kilometers. Lee's (1972) curve is plotted using the □

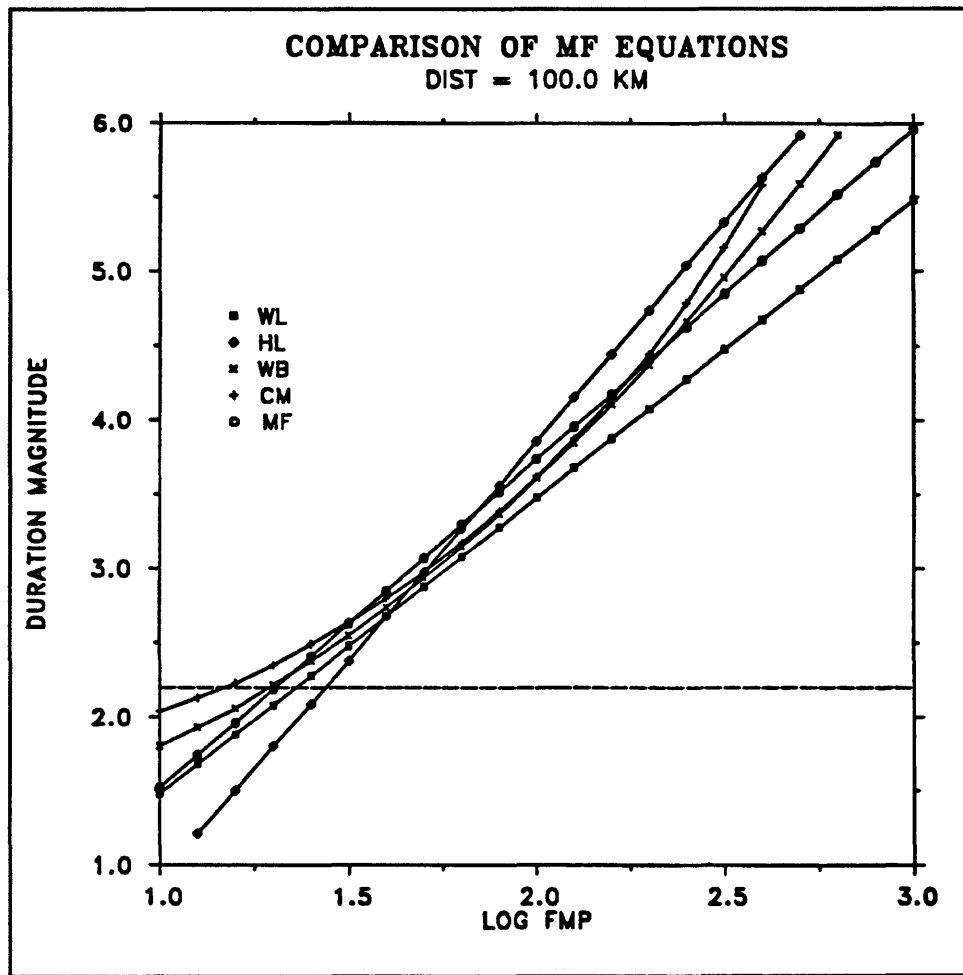


Figure 22 Comparison of MF equations.

symbols. My curve is the one plotted with ○ symbols. The fit between all these curves is reasonably good in the range magnitude 3 to magnitude 4, but it deteriorates quickly outside that range.

The fitted amplitude magnitude and duration magnitude equations are given in table 1. The amplitude magnitude is the one denoted MX_{ij} and the duration magnitude is the one denoted MF_{ij} . My curve uses the Bakun-Joyner attenuation formulation for distances less than 185 kilometers, but beyond that a simple linear form fits better. Also, a different correction

Table 1. Fitted Amplitude and Duration Magnitude Equations

The "station" magnitudes,	
MX _{i,j} and MF _{i,j} are calculated by the equations:	
MX _{i,j} =log(A _{i,j} /2)-logA ₀ +log(SWA/SGS _i) _{21,j} +XC _i +XS _i	
MF _{i,j} =2.22(logτ _{i,j} -logτ ₀)+log(CAL15/CAL _i)+FC _i +FS _i +HF(h) _j	
where,	
-logA ₀ =0.955+0.82logR _{i,j} +0.00405R _{i,j}	D _{i,j} < 185.3 km
=3.57+2.55(logR _{i,j} -log185.3)	D _{i,j} > 185.3 km
δ(-logA ₀)=0.09sin(0.07(D _{i,j} -25.0))	if D _{i,j} < 70.0 km
and	
-2.22logτ ₀ =-0.81+0.0011D _{i,j} +0.005(D _{i,j} -40.)	D _{i,j} < 40.0 km
=-0.81+0.0011D _{i,j}	40.0 km < D _{i,j} < 350.0 km
=-0.81+0.0011D _{i,j} +0.0006(D _{i,j} -350.0)	D _{i,j} > 350.0 km
HF(h) _j =0.014(h _j -10.0)	if h _j > 10 km
Further definitions	
MX _j =median(MX _{i,j})	event MX
MF _j =median(MF _{i,j})	event MF
RX _{i,j} =MX _{i,j} -MX _j	station MX residual
RF _{i,j} =MF _{i,j} -MF _j	station MF residual
DXF _j =MX _j -MF _j	difference MX _j -MF _j for event j
ARX _i =SUM(RX _{i,j})/N	average MX residual at site i
ARF _i =SUM(RF _{i,j})/N	average MF residual at site i
ADXF=SUM(DXF _j)/N	average (MX-MF) over entire set of events

is used for distances less than 70 kilometers. Similarly, the duration magnitude correction term is broken into three different formulations based on distance. There is also a slight depth dependence in the duration magnitude equation. For a given event, there will be many different stations which record it and which will go into the magnitude calculation. I use the median value of all the magnitude estimates to avoid biasing the estimate due to the

inevitable bad station or two that creep into the estimate.

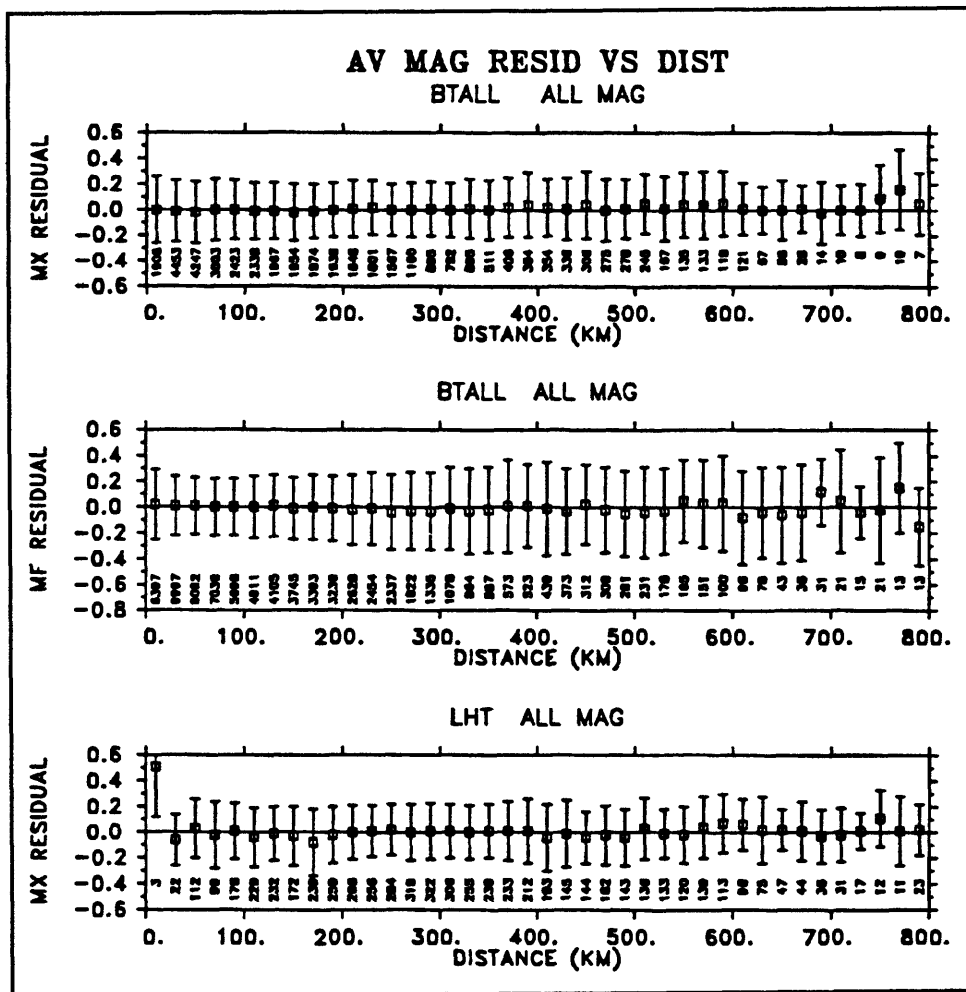


Figure 23 AV magnitude residual versus distance.

For a given event, each station has a residual which is its station magnitude minus the event magnitude. The fit of these equations was done by minimizing these residuals. The top plot in Figure 23 is the average station magnitude residual for all the earthquakes where the data has been broken up into 20 kilometer bins and the averages over all those bins have been computed. The bars represent the standard deviations of the station residuals. The center plot is essentially the same thing except that it is for duration magnitudes. In both

cases, the residuals show no systematic dependence on distance and the fluctuation about zero is very small. The third plot in Figure 23 is from an auxiliary data set which consists of earthquake records read by hand from paper records. Only amplitude information is available for this data set, so there have been no MF residuals calculated.

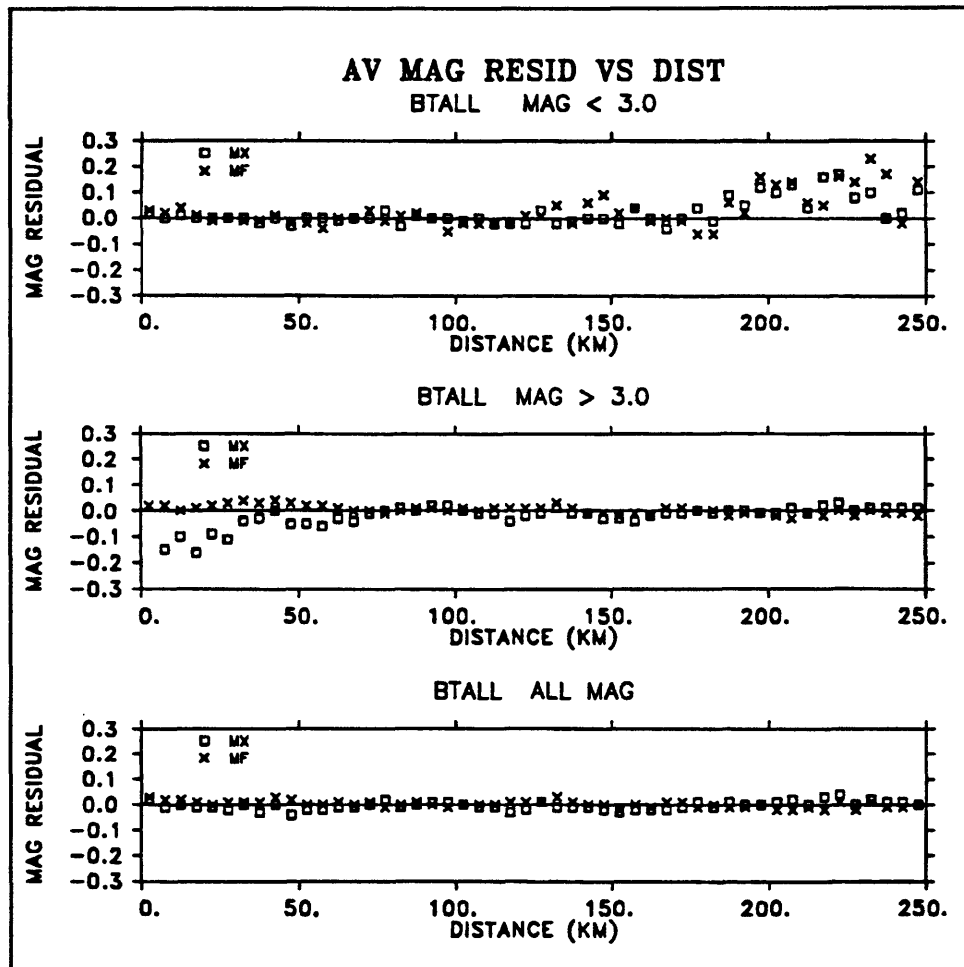


Figure 24 Magnitude residuals at small distances.

Figure 24 shows a little bit closer comparison made for shorter distances. In these analyses, the data was grouped into 5 kilometer bins out to a distance of 250 kilometers. The data set was broken into two parts for this analysis. One part consisted of earthquakes with

magnitude smaller than 3 and the other part consisted of earthquakes with magnitudes greater than 3. For the small magnitude subset, shown in the top plot, the residuals are uniformly small until a distance of about 150 kilometers. Then the residuals become positive. This is an artifact which is due to the effect of noise on low amplitude records. Also, there is quite a scatter in station response, so when you get out to the range where most stations don't see the event, the few stations that do are recording with an abnormally large response.

The subset of larger earthquakes is plotted in the second frame. For these larger earthquakes, things are well behaved at the larger distances and the duration magnitudes are well behaved at small distances, but at short distances there are problems with the amplitude magnitudes. The reason is that at small distances, the only stations which record the event without clipping either have abnormally low response or are at a site which received a low amplitude signal. Panel three shows the residuals for the complete data set on this distance scale. As you can see, the average residual is just a few hundredth's of a magnitude unit.

In Figure 25 the residuals between the amplitude magnitude and duration magnitude are plotted as a function of magnitude. The data were binned into $\frac{1}{2}$ magnitude unit intervals and averaged. The data set has been broken into two parts for this analysis. The first subset shown in the top plot had at least 10 measures of amplitude and 10 measurements of duration for each event. The second subset, shown in the center plot, fails that test, but has at least 5 of each measurement. The third plot is from the combined data sets. It shows that over the range from about magnitude .5 to about magnitude 5.5 the average residual is about zero. That is, the two scales track each other quite well over this range. The observations that fall off the curve have very few observations so they are not statistically significant.

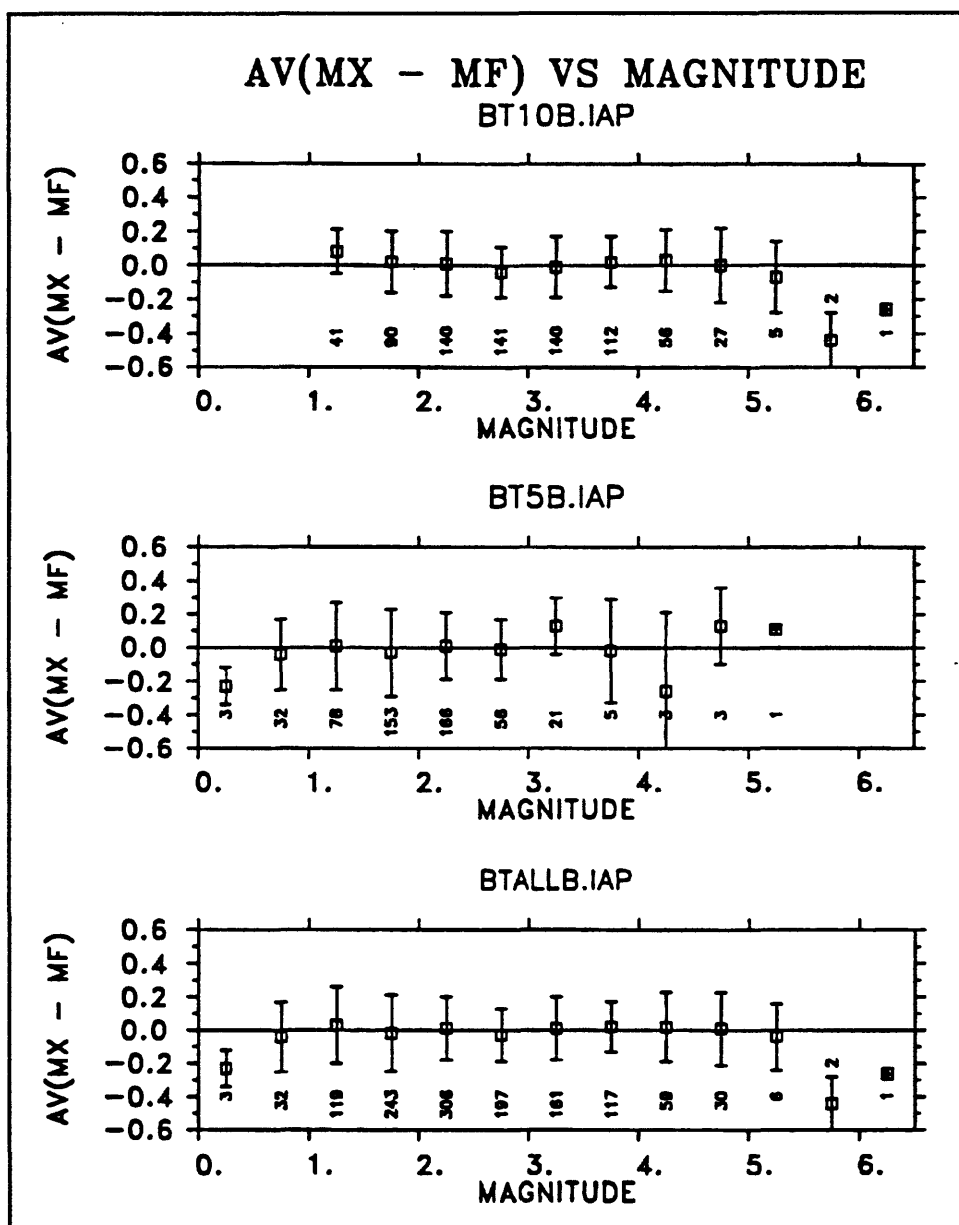


Figure 25 AV(MX - MF) vs magnitude.

If you plot the differences between the two scales on the map, you can see that they are systematic. Over certain regions, the residuals are consistently small while in some areas they are consistently large. I think this is telling us something about the earth, not about the procedure. You can see these differences when you look at the seismograms of events

generated in different regions. The character of the seismograms tends to change from region to region.

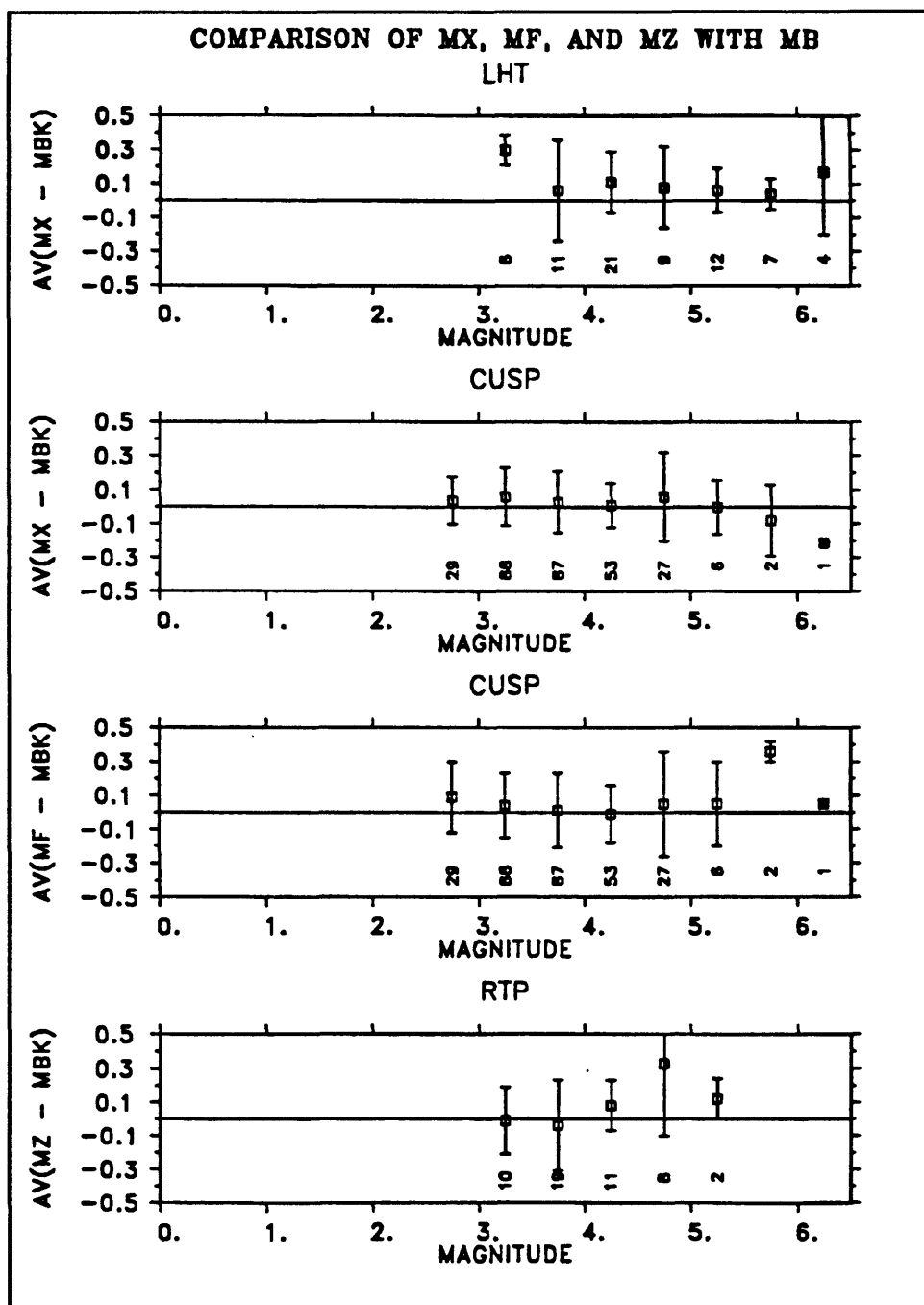


Figure 26 Comparison of MX, MF, and MZ with MB.

Figure 26 is a comparison of the various magnitudes computed by these methods with the comparable magnitude determined by UC Berkeley using the Wood-Anderson records they have available. There are several problems with this comparison. There are only four stations with Wood-Anderson seismographs installed. Only one of those stations has been there since the site corrections were originally determined by Gutenberg. Site corrections for the other instruments have been chosen to be consistent with that one station. Also, the Berkeley magnitudes are computed on the basis of the original Richter nomogram which departs systematically from what appears to be the best fit for the northern California data. Also, the number of events in this comparison is small since there weren't that many cases of overlap.

Looking at the second plot in Figure 26, we see that the average difference between the magnitudes is 0.04 for magnitudes between 2.5 and 5.5. The result is essentially the same for the comparison between duration magnitude and Richter magnitude as shown in the third plot. The real time processor magnitude comparison, shown in the fourth plot matches reasonably well in the magnitude range from about 3 to 4. It looks like a little adjustment in the coefficients could improve the fit.

Figure 27 shows one final comparison. I have constructed a $\log A_0$ versus \log distance table for northern California. In the first frame I have compared it with the equation that we use to compute MX. The fit is very good out to at least 600 kilometers.

If you compare this A_0 versus distance table with the one used by Richter, as shown in the second frame, you can see that there are some systematic differences. However, they are only on the order of a tenth magnitude unit except for very short distances. The Bakun-

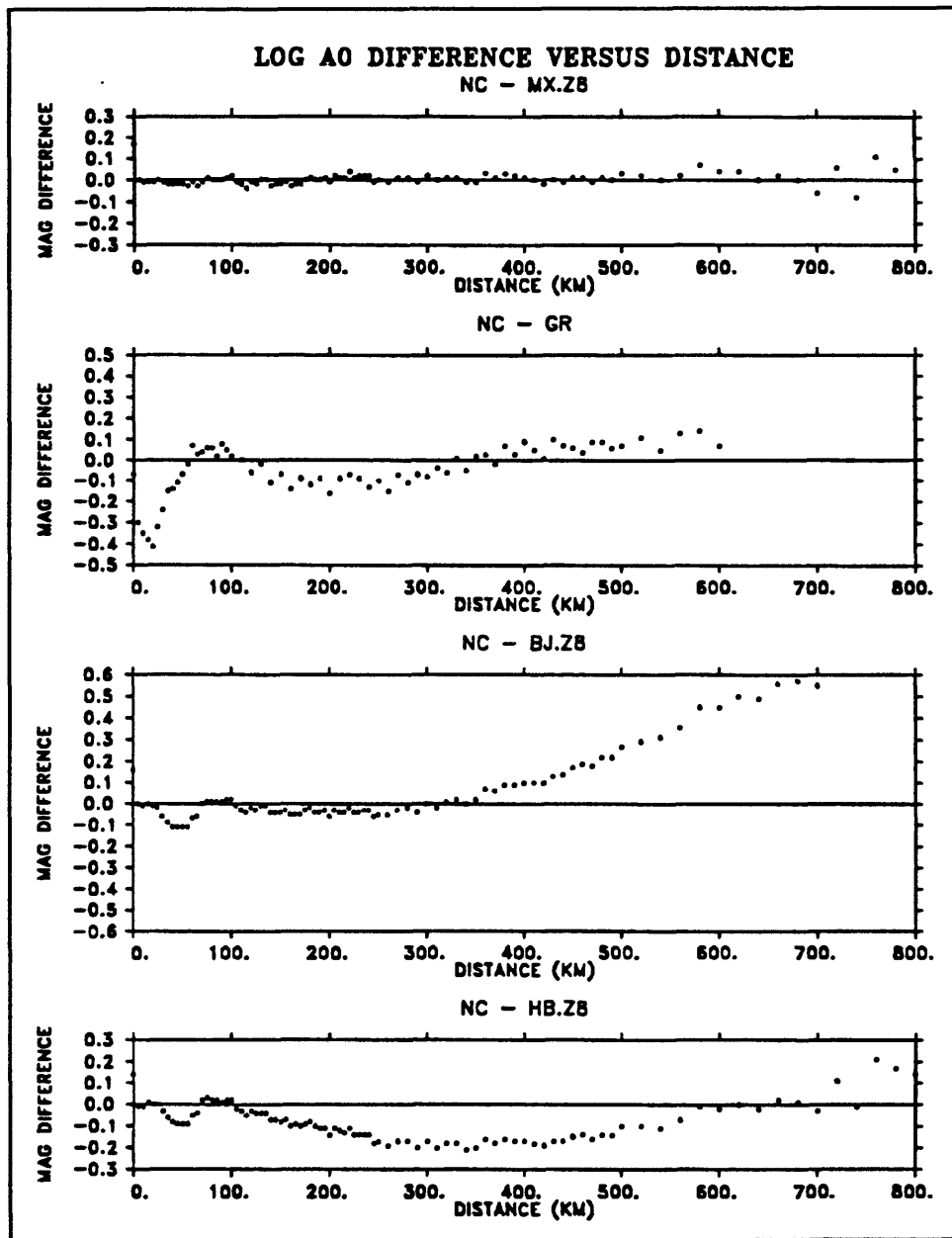


Figure 27 Log A_0 difference versus distance.

Joyner relation fits our data as well as theirs out to about 400 kilometers, but the fit falls apart beyond 400 Km since they had no data to constrain their model at large distances. For the southern California data, the relation by Hutton and Boore has a somewhat poorer fit.

This is not too surprising since the data shown is from northern California and the relation was developed from southern California data.

REFERENCES

- Bakun, W. H., and W. B. Joyner (1984). The M_L scale in central California. Bull. Seism. Soc. Am. **74**, 1827-1843.
- Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. J. Geophys. Res. **75**, 5997-5009.
- Eaton, J. P. (1992). Determination of amplitude and duration magnitudes and site residuals from short-period seismographs in northern California. Bull. Seism. Soc. Am. **82**, in press.
- Eaton, J. P. (1992). Determination of amplitude and duration magnitudes and site residuals from short-period seismographs in northern California. Geol. Surv. Open-File Rep. 92-0086.
- Gutenberg, B., and C. F. Richter (1942). Earthquake magnitude, intensity, energy, and acceleration. Bull. Seism. Soc. Am. **32**, 163-191.
- Hutton, L. K., and D. M. Boore 1987). The M_L scale in southern California. Bull. Seism. Soc. Am. **77**, 2074-2094.
- Lee, W. H. K., Bennett, R. E. , and Meagher, K. L. (1972). A method of estimating magnitudes of local earthquakes from signal duration. Geol. Surv. Open-File Rep. 28.
- Richter, C. F. (1935). An instrumental earthquake magnitude scale. Bull. Seism. Soc. Am. **25**, 1-32.

14. CODA WAVES FOR MAGNITUDE AND Q

by

W. H. K. Lee
U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

In this discussion of coda waves I intend to present only a brief overview of the subject. There are literally hundreds of papers devoted to the study of coda waves, and if you have an interest in this subject, you will have to spend some time reading the literature.

I will start by presenting a list of some useful papers. This selection of references is obviously subject to personal preference and bias.

However, I think the papers listed here are among the most important ones dealing with the subject of using coda waves to estimate magnitude and attenuation.

Chronology and References (1)

- 1935. M_L : Richter, BSSA, 25, 1-32.
- 1945. M_S, m_b : Gutenberg, BSSA, 35, 3-12, 57-69, 117-130.
- 1966. M_O : Aki, BERI, 44, 73-88.
- 1969. *Coda*: Aki, JGR, 74, 615-631.
- 1970. Ω_o, f_o, ϵ : Brune, JGR, 75, 4997-5009.
- 1972. Ω_o, f_o, ϵ (How to): Hanks and Wyss, BSSA, 62, 561-589.
- 1972. M_D : Lee, USGS Open-file Report.
- 1975. *Coda*: Aki and Chouet, JGR, 80, 3322-3342.

Figure 1 Chronology and references (1).

MAGNITUDE SCALES AND CODA DURATION

Everyone is probably aware of the classic paper by Richter (1935) introducing the local magnitude scale, M_L . His work was widely accepted, and almost ten years later, Gutenberg, (1945 a,b,c) introduced M_s and m_b . Aki (1966) was apparently the first to define seismic moment M_0 , even though others had discussed the concept. Also, in 1969, he did some original work on coda waves which was published in Aki (1969). In 1970, there was another very important paper by Brune (1970), in which he introduced a model of the seismic source which was quantified by the parameters Ω_0 , f_0 , and ϵ . Two years later, Hanks and Wyss published a paper which told how to determine these parameters from seismic data (Hanks and Wyss, 1972). In 1972, I and my associates (Lee et al. 1972) introduced the M_D scale (duration magnitude). Also, in 1975, Aki and Chouet did further studies on the coda (Aki and Chouet, 1975).

At the same time, Bob Herrmann was applying the coda duration for magnitude and moment estimation (Herrmann, 1975). In 1977, Kanamori, introduced the moment magnitude scale, M_w (Kanamori, 1977). In 1987, there was a review by Herraiz and Espinosa (1987) summarizing the results of coda wave studies that had been done over the last ten years. Also, that same year there was a very important paper by Frankel and Wennerberg (1987) on the coda model. In 1988, Wu and Aki (1988) edited the book, "Scattering and Attenuation of Seismic Waves." In 1989, Carlos Valdes and David Novelo-Casanova published the program QCODA in volume one of the IASPEI Software Library (Valdes and Novelo-Casanova, 1989).

What do we know about coda waves? First, we note that the amplitude of coda waves

decays very slowly. Even after all the phases have arrived, the seismograph still receives energy. For example, after the great Chilean earthquake, the whole earth shook for about four days. What are these coda waves, and why are they so persistent?

A second observation people have made is that coda duration is not very sensitive to epicentral distance. A station 50 kilometers away from an earthquake will ring for about the same length of time as a station 100 kilometers away.

Another interesting thing that we observe, at least here in California, is that coda duration is not very sensitive to instrument gain, if you set the gain according to the background

noise level. If the amplitude of the background noise is high, we lower the instrument gain. If the amplitude of the background noise is low, we increase the instrument gain. Therefore, the way we set instrument gain compensates for the site effects, and no matter what the

Chronology and References (2)

- 1975. τ for M , M_0 : Herrmann, BSSA, 65, 899-913.
- 1977. M_W : Kanamori, JGR, 82, 2981-2987.
- 1987. Review: Herraiz and Espinosa, PAGEOPH, 125, 499-577.
- 1987. Coda model: Frankel & Wennerberg, BSSA, 77, 1223-1251.
- 1988. "Book": Wu and Aki (Eds), "Scattering and Attenuation of Seismic Waves", PAGEOPH, 128, Nos. 1/2.
- 1989. QCODA: Valdes & Novelo-Casanova, IASPEI Software Library, 1, 237-255.

Figure 2 Chronology and references (2).

Empirical Observations about Coda Waves

- o Amplitude of coda waves decays slowly
- o Coda duration not sensitive to epicentral distance
- o Coda duration not sensitive to instrument gain

Figure 3 Empirical observations about coda waves.

instrument gain is, the coda wave amplitudes are adjusted to the background noise.

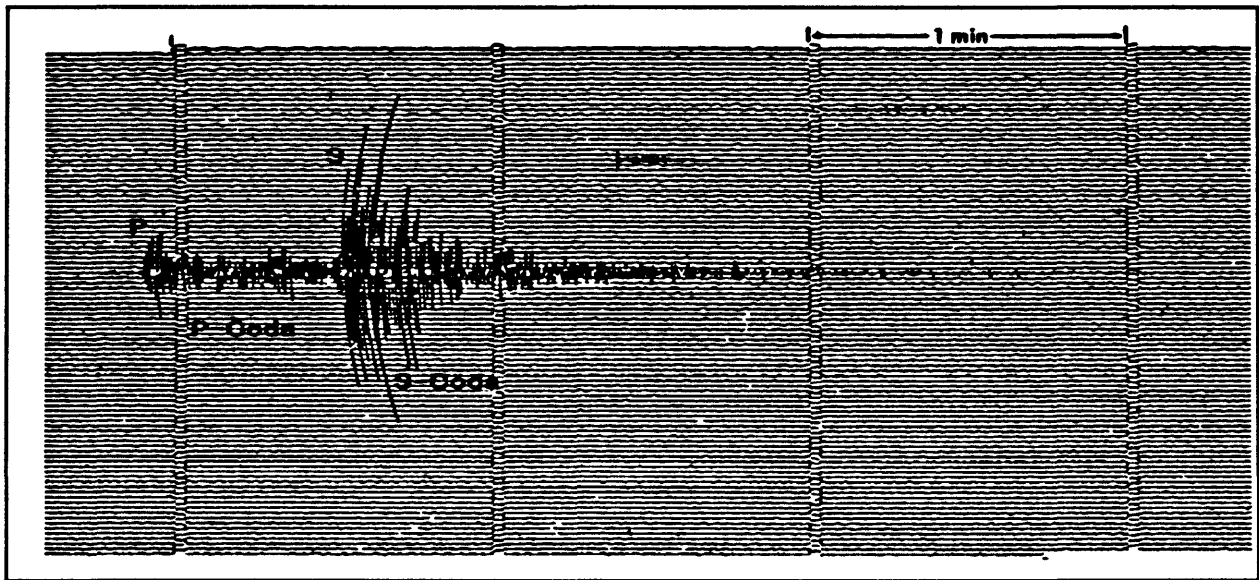


Figure 4 Seismogram from $M_L = 3.7$ earthquake recorded at station GUD (Madrid, Spain).

Figure 4 is taken from a review paper by Herraiz and Espinosa (1987). You can see the P-wave arrival and the S-wave arrival and maybe some surface waves later on. The P-coda and the S-coda are the decaying wave trains that follow the respective phase onsets. You can see that the coda length is many times greater than the length of the phase which generates it.

Figure 5 is taken from Herrmann (1975) and is a seismogram from an earthquake which occurred in Illinois in 1965. Again, you can see that most of the record consists of coda waves, although there are some phase arrivals visible near the start of the record. The solid envelope lines are the result of fitting the data to a power law function. In this case, the exponent has the value -1.5.

Figures 6 and 7 illustrate a phenomenon that we observe for earthquakes in northern

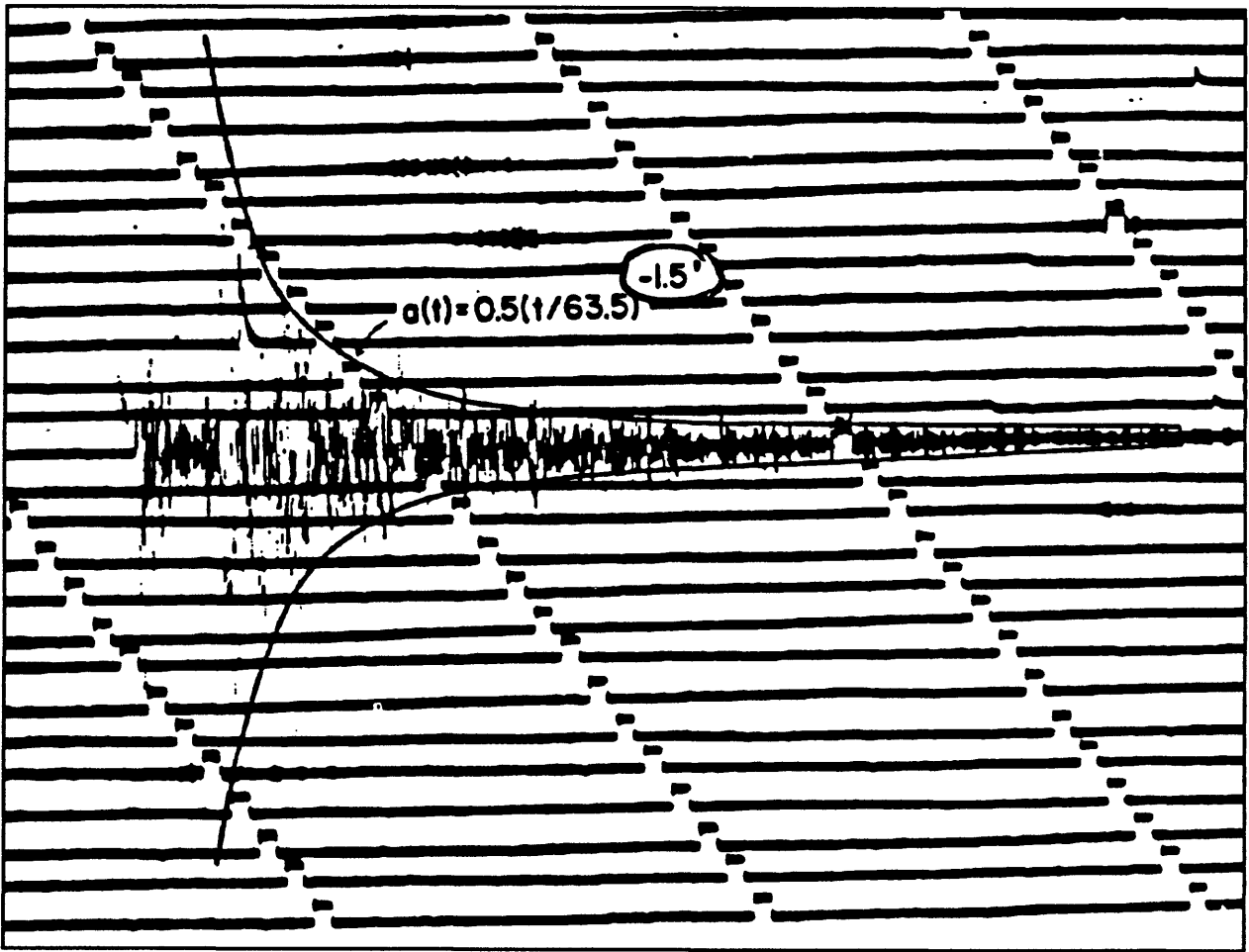


Figure 5 Seismogram of the Tamms, Illinois earthquake.

California. The figures show a 32-channel record section for an earthquake recorded by Calnet. The channels are arranged in order of increasing epicentral distance, with the closest station less than 21 kilometers away and the most distant station 78 kilometers away from the source. Notice that for all these channels, the coda duration is about the same. This is typical behavior and leads us to conclude that coda duration is relatively insensitive to epicentral distance. The stations shown in these figures also cover a wide range of azimuths yet there is no systematic dependence on azimuth angle either. This insensitivity to azimuth is also typical of coda duration for earthquakes recorded by Calnet.

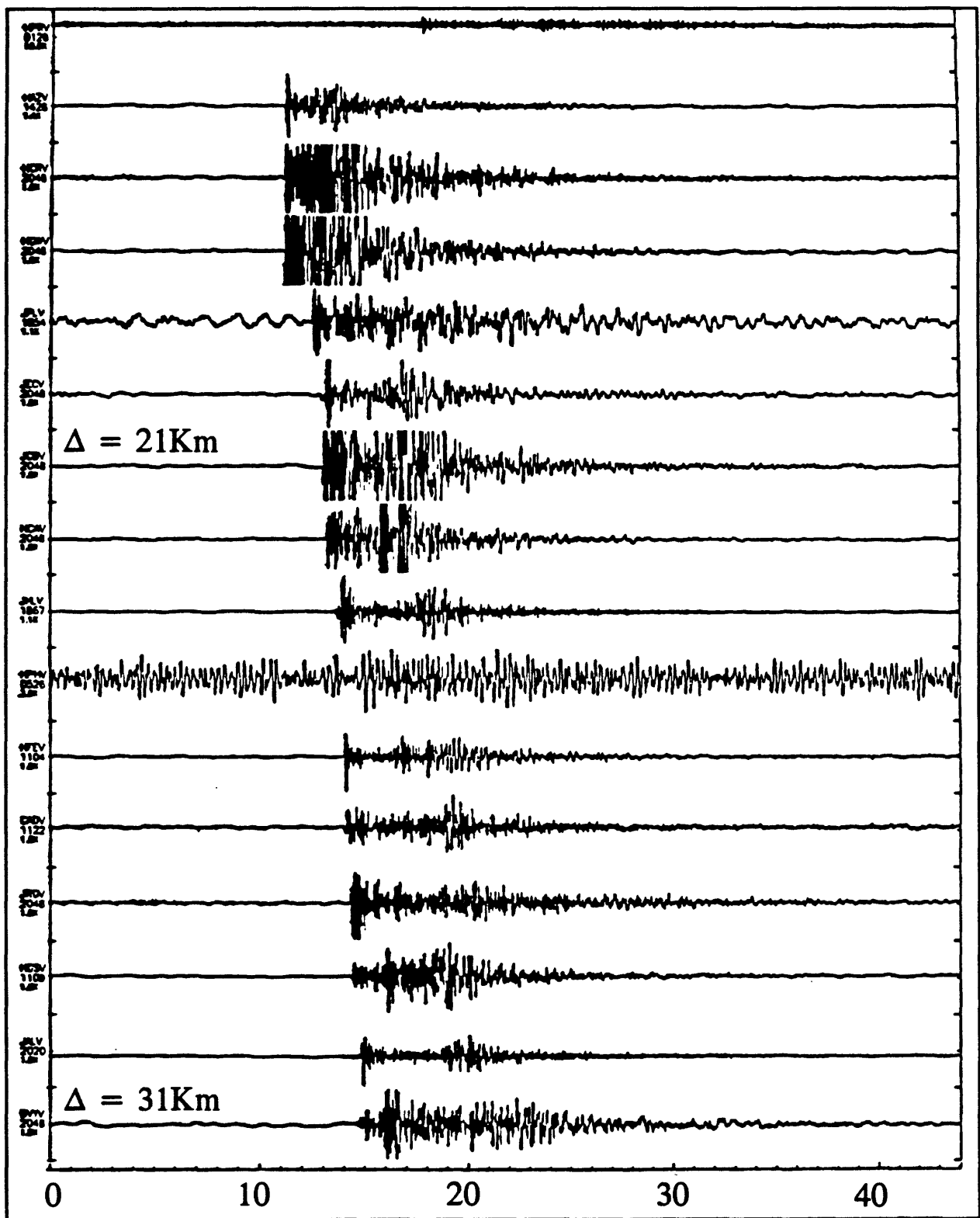


Figure 6 Record collected by Calnet (1).

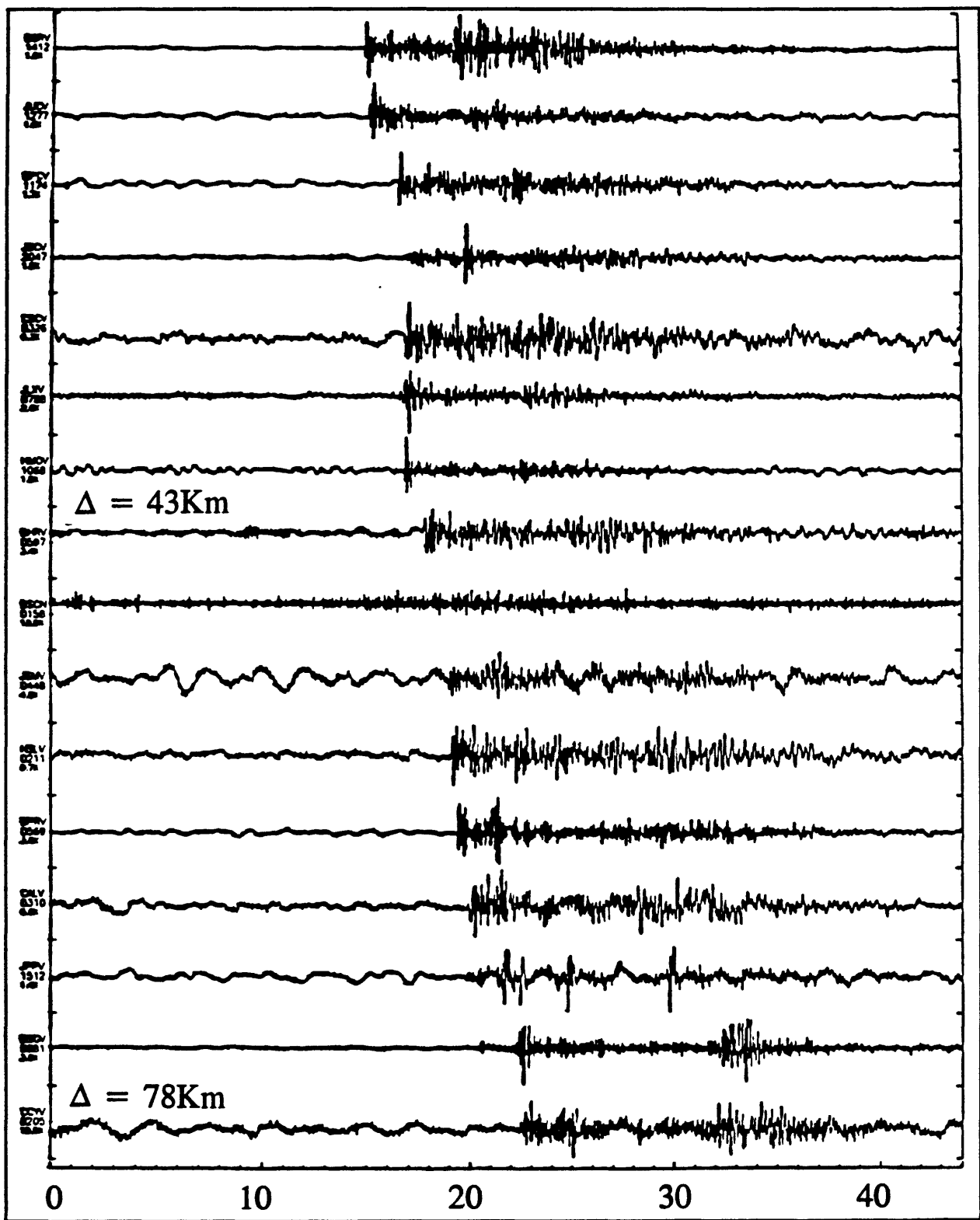


Figure 7 Continuation of Calnet record from Figure 6.

This is very interesting behavior and a number of people have tried to quantify it. In 1975, Herrmann (1975) proposed the very simple coda amplitude expression:

$$A(t) = C t^{-\alpha} \quad (1)$$

Here C is some measure of the strength of the source and $t^{-\alpha}$ controls the decay rate for the coda wave train. Since local magnitude is given by $\log(A(t))$ plus some correction terms, it follows that magnitude can be expressed in terms of duration as:

$$M_L \propto \alpha \log \tau \quad (2)$$

where τ is the duration.

Aki and Chouet (1975) introduced a formula that includes attenuation of the coda waves in the rock mass. Their formula for coda amplitude is:

$$A(t) = C t^{-\alpha} e^{\left(\frac{-\omega t}{2Q_c}\right)} \quad (3)$$

and the corresponding formula for magnitude is:

$$M_L \propto \alpha \log \tau + \beta \tau \quad (4)$$

Finally, Frankel and Wennerberg (1987) proposed the formulas:

$$A(t) = C t^{-1.5} e^{\left(\frac{-\omega t}{2Q_I}\right)} \sqrt{1 - e^{\left(\frac{-\omega t}{Q_S}\right)}} \quad (5)$$

and

$$M_L \propto 1.5 \log \tau + \beta \tau + \dots \quad (6)$$

Notice that they explicitly used the α value of 1.5 that Herrmann (1975) had observed. They

introduced two factors in their expression for $A(t)$. The first factor which they denoted Q_i is called the intrinsic Q , and the other new factor is Q_s which is the Q due to scattering.

Although the amplitude expression is somewhat complicated, the expression for magnitude is nearly of the same form as previous magnitude expressions.

Before continuing with the discussion on how to use duration to compute magnitude, let's go back and look at the original Richter local magnitude definition. I will be brief, since the topic has been covered in some detail in Chapter 13.

Richter's scale is based on using Wood-Anderson seismographs, to measure the amplitude of ground displacement and then taking the base 10 logarithm of the amplitude as a function of distance from the source. The basic formula is:

Richter Local Magnitude

- Based on Wood-Anderson seismograph
- Uses \log_{10} , and selects $A_o(\Delta)$
- Defined by: $M_L = \log A - \log A_o(\Delta)$

Richter's scale depends on:

$$A_m(\Delta) = S \cdot F(\Delta)$$

where

$A_m(\Delta)$: maximum amplitude

S : source strength

$F(\Delta)$: path effect

Taking logarithm:

$$\log A_m(\Delta) = \log S + \log F(\Delta)$$

Equate $M_L = \log S$, then:

$$M_L = \log A_m(\Delta) - \log F(\Delta)$$

Figure 8 Richter local magnitude.

$$M_L = \log A - \log A_o(\Delta) \quad (7)$$

Richter's formula for magnitude is dependent on two factors, a source factor and the path effect on the propagating energy. Symbolically:

$$A_m(\Delta) = S \cdot F(\Delta) \quad (8)$$

where $A_m(\Delta)$ is the maximum amplitude at some distance, S is the source strength, and $F(\Delta)$

is the path effect. Taking logarithms of both sides of Equation (8) gives:

$$\log A_m(\Delta) = \log S + \log F(\Delta) \quad (9)$$

Richter noted that the magnitude of an earthquake should be independent of the effects of the propagation path. That is, we want to estimate the earthquake size at the source. This is the motivation for defining the magnitude as $M_L = \log S$, which gives the magnitude as:

$$M_L = \log A_m(\Delta) - \log F(\Delta) \quad (10)$$

As Jerry Eaton mentioned, when we first set up the network more than 20 years ago, we were recording data on a strip of film. The records looked like the one in Figure 9.

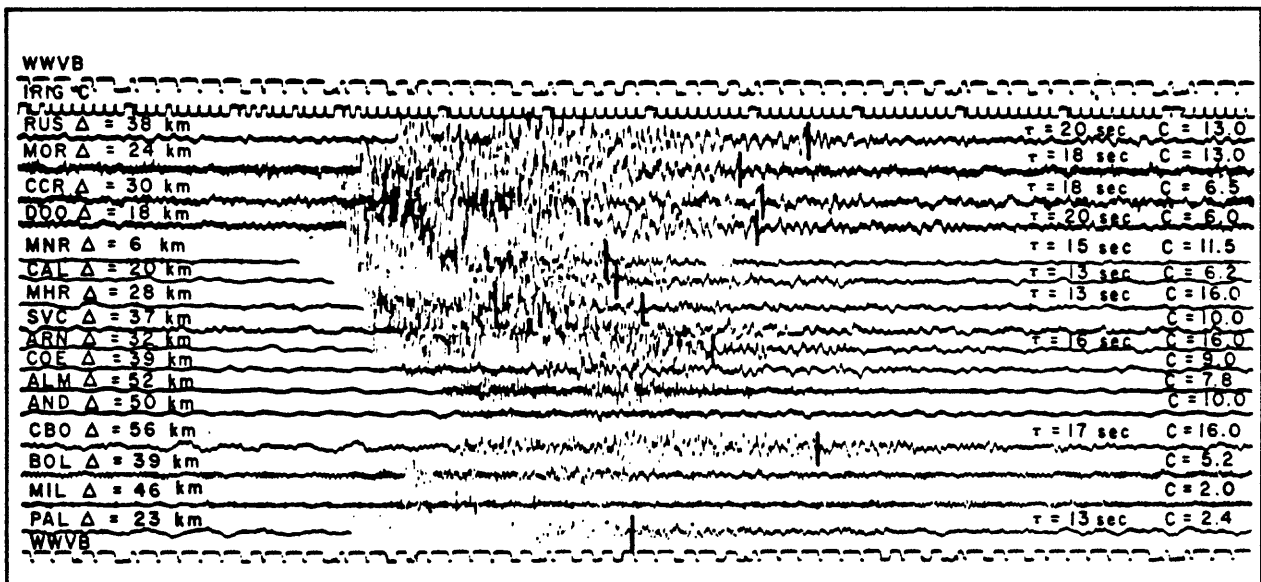


Figure 9 Typical develocorder record.

As you can see, it is nearly impossible to measure amplitudes with records like this. Also, we did not have very good calibration on our instruments when we first started doing this. Since the magnitude measurement depends on the measured amplitude of the trace, if amplitudes are incorrect due to poorly calibrated instruments, then magnitudes will also be incorrect.

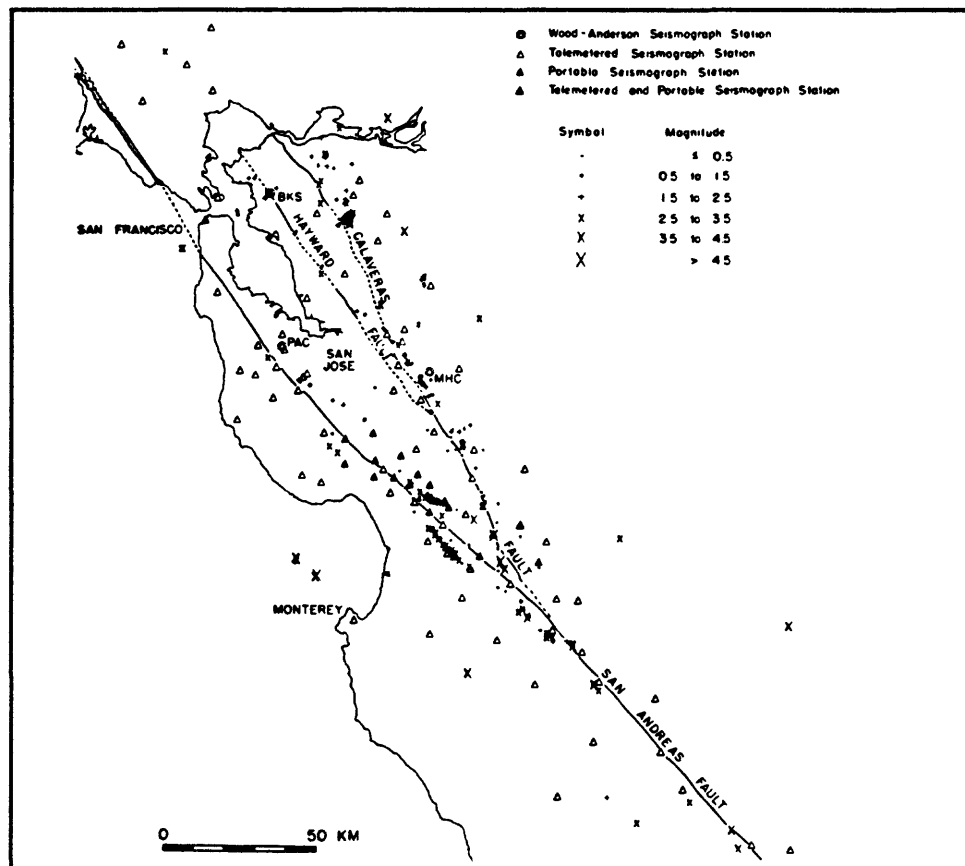


Figure 10 Location of Wood-Anderson seismographs.

At that time, I was processing the data and was completely frustrated by our inability to produce reliable magnitude estimates. However, a colleague (Bob Hamilton) told me about a paper written by a Hungarian seismologist by the name of Bisztricsany (1958), in which magnitude was determined as a function of duration of the surface waves. I got the paper, but was not able to read it, since it was written in Hungarian. Further searching uncovered a paper by a Soviet seismologist (Solov'ev, 1965), in which magnitude was estimated from the total duration of the seismic event. Tsumura (1967) also published a method for determining magnitude by measurements of total event duration.

Knowledge of their work encouraged me to attempt a correlation between event

duration and Richter magnitude for some California earthquakes. There were two Wood-Anderson stations being operated by Berkeley, and after a great deal of trouble, I was able to resurrect a third instrument and install it at Stanford. This gave me the three stations circled on the map in Figure 10 to use for Richter magnitude determination. My plan was simply to

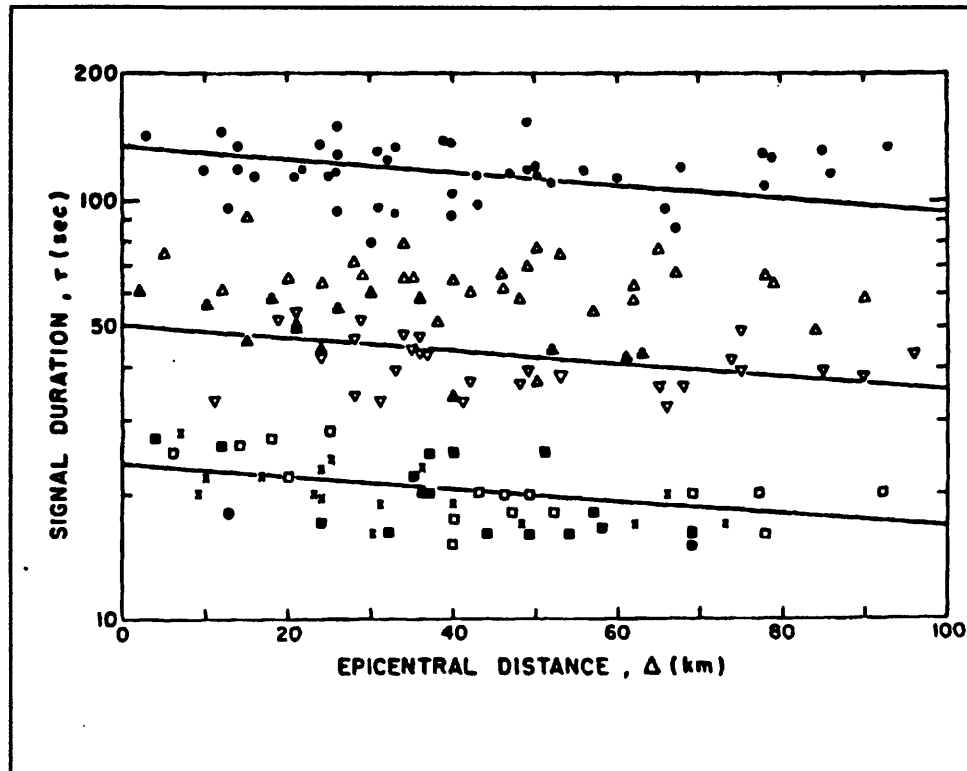


Figure 11 Dependence of duration on distance.

collect a number of records with this little network of Wood-Anderson seismographs, and compute magnitudes for all those events. At the same time, the events would be recorded by Calnet, and I could measure the durations off the develocorder recorder records. Then I hoped to find a simple relation between the durations and the measured magnitudes.

One of my concerns in this investigation was whether there would be a significant dependence of duration on distance. The plot in Figure 11 shows the dependence I found for

a set of 8 earthquakes. In this plot, each earthquake is represented by a distinct symbol. The lines represent best fit lines to the data. As you can see, there is only a slight dependence of duration on epicentral distance.

The next thing I did was compare the Richter magnitude with the average signal

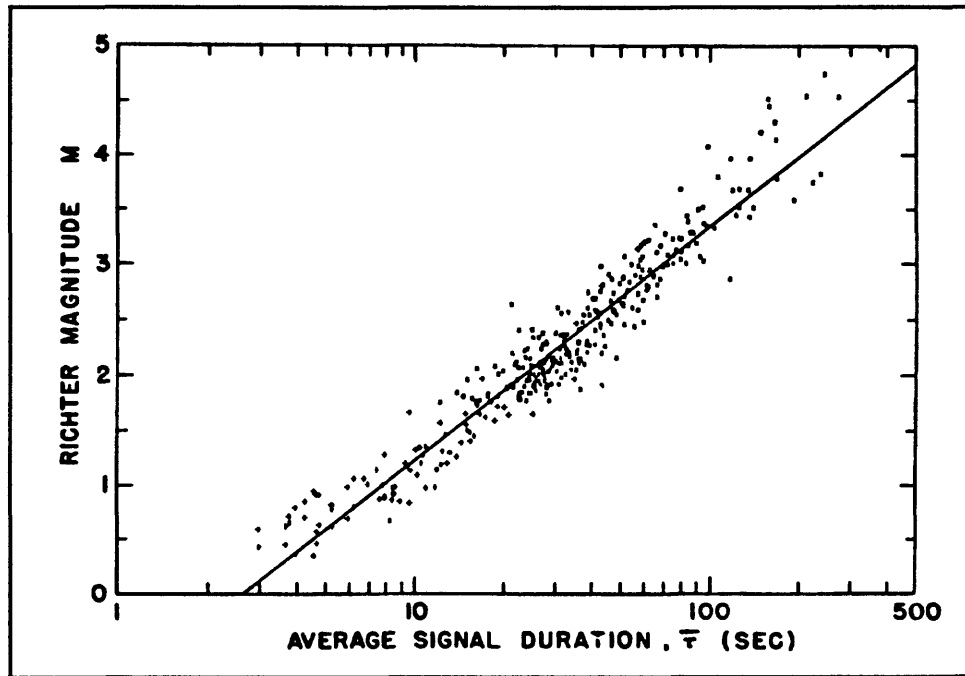


Figure 12 Richter magnitude vs signal duration.

duration. As Figure 12 shows, there is a definite linear dependence. The fit of a line to the data is quite good from about magnitude 1 to about magnitude 3 or 4. From about magnitude 3.5 up, I really didn't have enough data to constrain the fit very well.

From this study, (Lee et al. (1972) developed the duration magnitude equation as:

$$m_D = -0.87 + 2 \log \tau + 0.0035 \Delta \quad (11)$$

As you can see, the correction factor for distance is only about 0.3 units per 100 kilometers.

As I stated in the report of this study, one of the prime reasons for using this magnitude

formulation is its convenience. There is no theoretical basis for the formula. It is just an empirical relation, which has been shown to work for a large number of earthquakes in central California.

You would think that something this simple would be widely accepted, but that was not the case. The report from which this material is drawn was never published, because it was rejected for publication. To this day, I still don't understand why. I can only guess that since it was a new concept, which was at variance with the established way of doing things, the reviewers didn't want to be troubled by it. Possibly, the lack of a theoretical basis was also a problem for the reviewers. However, if you have a better mousetrap, people will use it! There is just no other simple way to estimate magnitudes for small earthquakes. So there are actually a large number of people using this formula or a similar one to estimate magnitude from signal or coda durations.

CODA WAVES AND ATTENUATION

There is still no satisfactory model to explain the generation of coda waves. There are essentially two schools of thought. One school, led by Aki and his students, argues that coda waves are generated by simple back-scattering. The argument is roughly like this. Suppose you have a seismic source and a receiver separated by some distance. The seismic waves propagate from the source radially, and after a certain length of time, some waves reach the receiver. However, if there are "discontinuities" in the volume of space surrounding the source and receiver, some of the seismic waves will be scattered off the "discontinuities". These scattered waves will arrive at the receiver some time after the direct wave. In a

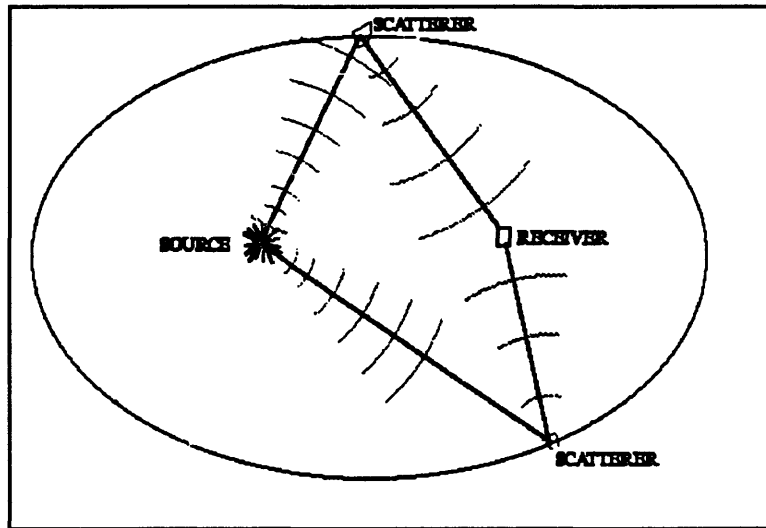


Figure 13 Back-scattering model.

constant velocity space, the energy which reaches the receiver at a given instant is from the scatterers which all lie on an ellipsoid surface with the source and receiver as foci. As time goes on, the scattered waves will sample larger and larger volumes of space.

Frankel and Wennerberg (1987) have a much more complicated model. I don't have time to go over it here in any detail, but I will make a few brief comments. They argue that there are two types of attenuation, intrinsic attenuation and scattered attenuation. The intrinsic attenuation is a property of the rock itself, and the scattered attenuation is due to propagation effects. The attenuation we measure is some combination of those two types. There is a big debate going on between those two schools of thought.

In either case, study of coda waves can give us useful information that would be hard to get with the direct wave alone. This is because the coda waves sample a much larger volume of rock than do the direct waves. As I mentioned before, there can be focusing and de-focusing effects along certain paths in the earth. If you measure just the amplitude of the

direct waves in order to compute magnitude, your measurement may be biased by such an effect. This is especially true if you only have one or two stations. On the other hand, the coda waves will be less subject to such effects, since they sample a much larger volume of the earth on their way to the seismometer.

Since coda waves sample such a large volume of the earth, we can also use coda waves to measure the attenuation of the earth. Aki and Chouet (1975) proposed the following equation for the amplitude of the coda wave.

$$A(\omega, t) = C(\omega) t^{-\alpha} e^{\left(\frac{-\omega t}{2Q_c}\right)} \quad (12)$$

In this formulation, the amplitude of the coda wave depends on the source function $C(\omega)$.

The source function does not depend on time. The $t^{-\alpha}$ term accounts for geometric spreading.

The exponential term accounts for the attenuation which is not due to geometric spreading.

What we would like to do is find the value for Q_c as a function of frequency by analyzing seismograms. This can be formulated as a linear least squares problem if we rearrange terms and take natural logarithms of both sides of Equation (12) to get:

$$\ln(At^\alpha) = \ln C - \frac{\omega}{2Q_c} t \quad (13)$$

Since, for every t you can measure At^α , you have the information to do a least squares fit of $\ln(At^\alpha)$ versus t . The slope parameter (b) which you get is equal to $\omega/2Q_c$, from which:

$$Q_c(\omega) = \frac{\omega}{2b} \quad (14)$$

There are several ways to carry out this Q estimate in practice. What Aki and I did a

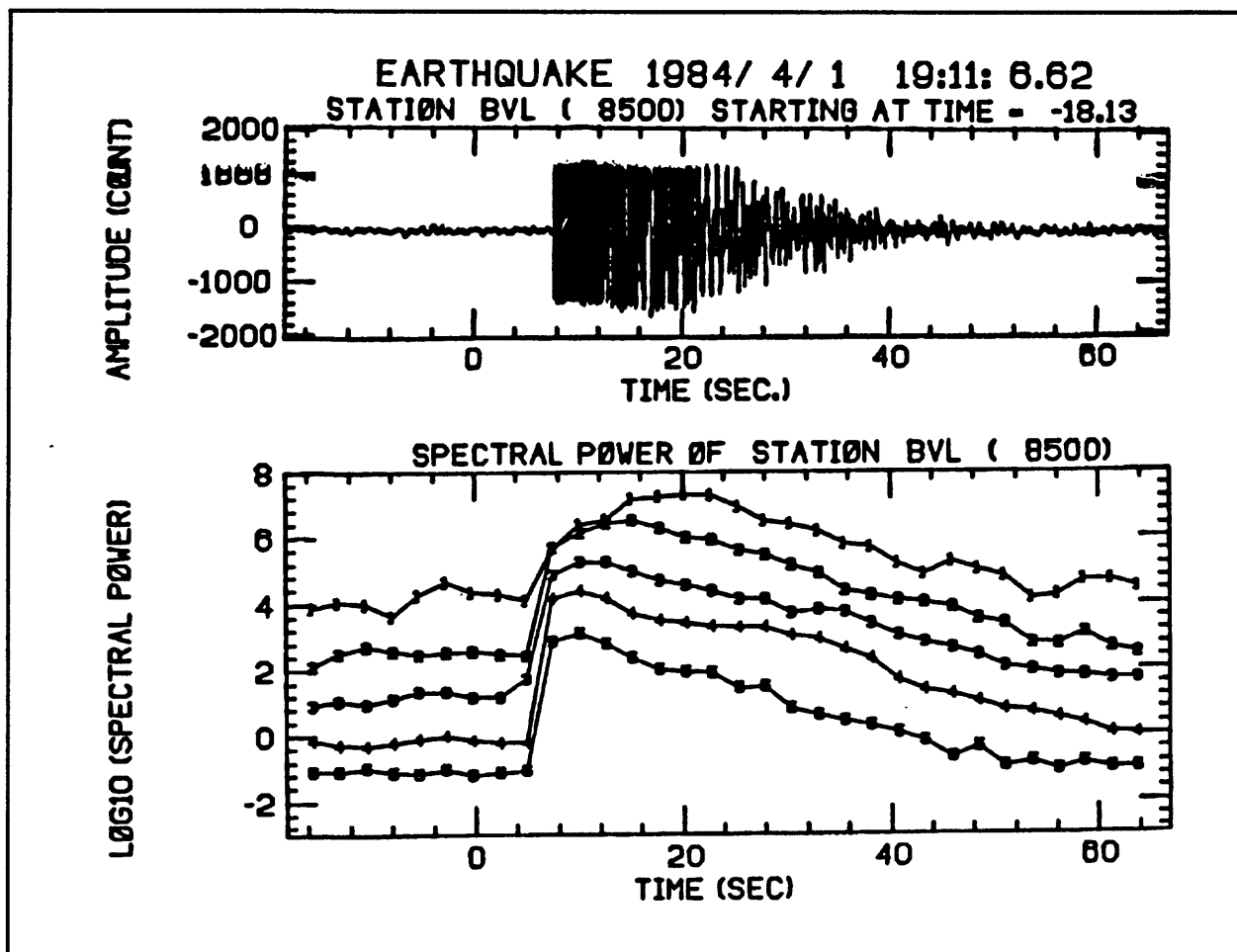


Figure 14 Seismogram and resulting spectral power.

few years back was to use a moving window across the seismogram (Lee et al. 1986). Figure 14 shows a seismogram and a plot of spectral power as a function of time for that seismogram. The spectral power plot is actually for 5 different frequency bands. These are 1-2 Hz, 2-4 Hz, 4-8 Hz, 8-16 Hz, and 16-32 Hz from the top down. We didn't attempt anything past 32 Hertz because of the bandwidth of the telemetry system. It was a pretty straight-forward matter to fit straight lines to the decay portion of the spectral power plots. The only complication is dealing with the effect of the direct S-wave energy. You can either start far enough down the seismogram that the direct S-waves are substantially gone, or you

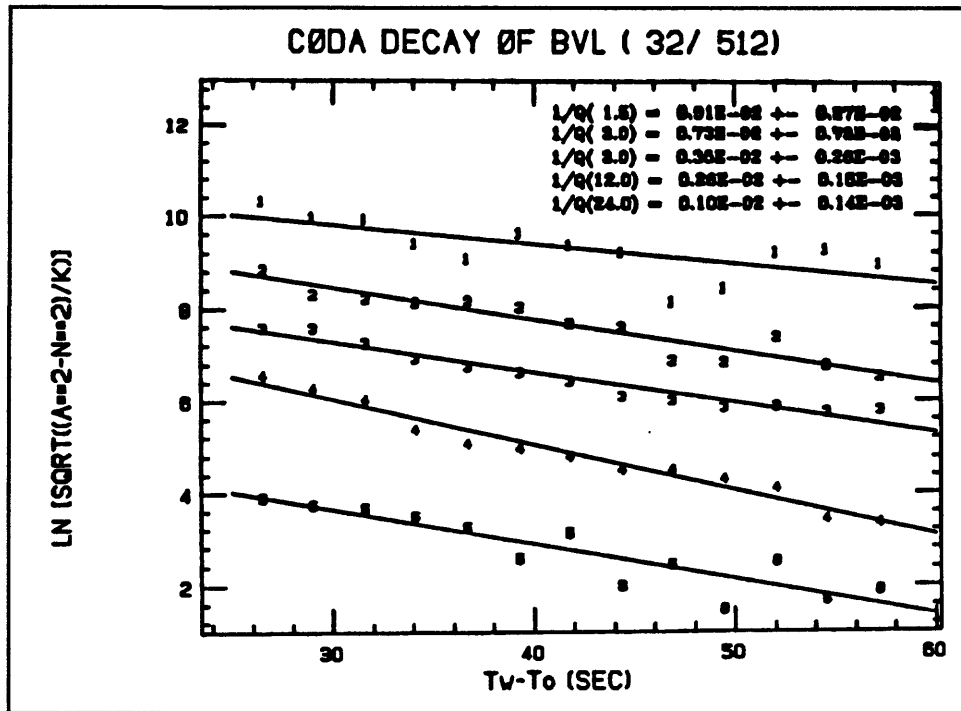


Figure 15 Straight line fit to spectral data.

can use the K-factor due to Sato (1977) and start at the S-wave onset. Figure 15 shows the result of fitting the data shown in Figure 14. Each of the 5 estimates is considered to be the value of Q_c for the center frequency of its band. That is, for the band 1-2 Hz, the estimate is considered to be for the frequency 1.5 Hz.

This has been a quick survey of a complex subject. All I really wanted to get across in this chapter, was the concept that there is a large part of the seismogram not due to energy which arrived by a direct path or by reflection off a major boundary in the earth. We have some theories which explain the existence and behavior of coda waves, but the subject is still very much, an active research area. The behavior of coda waves is analogous to that of echoes in a large canyon. A very loud shout may echo for several seconds, while a whisper is gone almost immediately. Similarly, the largest earthquakes have coda durations of days,

while small local events have coda durations of seconds. This behavior forms the basis for an empirical method of estimating magnitude, which can provide reasonable estimates even when the signals are clipped or when much of the earthquake energy is outside the response of the seismometers. We can also use coda waves to make estimates of the attenuating properties of the rock in a region which includes both the source and the receiver. The methods are simple and provide estimates of Q_c as a function of frequency.

REFERENCES

- Aki, K. (1966). Generation and propagation of G waves from the Niigata earthquake of June 16, 1964. Part 2. Estimation of earthquake moment, released energy, and stress-strain drop from the G-wave spectrum. *Bull. Earthquake Res. Inst. Univ. Tokyo* **44**, 73-88.
- Aki, K. (1969). Analysis of seismic coda of local earthquakes as scattered waves. *J. Geophys. Res.* **74**, 615-631.
- Aki, K., and Chouet, B. (1975). Origin of coda waves: Source, attenuation, and scattering effects. *J. Geophys. Res.* **80**, 3322-3342.
- Bisztricsany, E. (1958). A new method for the determination of the magnitude of earthquakes. *Geofiz. Kozl.* **7**, 69-96 (in Hungarian with English abstract).
- Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *J. Geophys. Res.* **75**, 4997-5009.
- Frankel, A. and Wennerberg, L. (1987). Energy-flux model of seismic coda: separation of scattering and intrinsic attenuation. *Bull. Seismol. Soc. Am.* **77**, 1223-1251.
- Gutenberg, B. (1945a). Amplitudes of surface waves and magnitudes of shallow earthquakes. *Bull. Seismol. Soc. Am.* **35**, 3-12.
- Gutenberg, B. (1945b). Amplitudes of P, PP, and S and magnitude of shallow earthquakes. *Bull. Seismol. Soc. Am.* **35**, 57-69.
- Gutenberg, B. (1945c). Magnitude determination for deep focus earthquakes. *Bull. Seismol. Soc. Am.* **35**, 117-130.

- Hanks, T. C., and Wyss, M. (1972). The use of body wave spectra in the determination of seismic source parameters. *Bull. Seismol. Soc. Am.* **62**, 561-589.
- Herraiz, M., and Espinoza, A.F. (1987). Scattering and attenuation of high-frequency seismic waves: development of the theory of coda waves. *Pure Appl. Geophys.*, **125**, 499-577.
- Herrmann, R. B. (1975). The use of duration as a measure of seismic moment and magnitude. *Bull. Seismol. Soc. Am.* **65**, 899-913.
- Kanamori, H. (1977). The energy release in great earthquakes. *J. Geophys. Res.* **82**, 2981-2987.
- Lee, W. H. K., Bennett, R. E., and Meagher, K. L. (1972). A method of estimating magnitude of local earthquakes from signal duration. *Geol. Surv. Open-File Rep.* 28.
- Lee, W.H.K., Aki, K., Chouet, B., Johnson, P., Marks, S., Newberry, J.T., Ryall, A.S., Stewart, S.W., and D.M. Tottingham (1986). A preliminary study of coda Q in California and Nevada. *Bull. Seismol. Soc. Am.* **76**, 1143-1150.
- Richter, C. F. (1935). An instrumental earthquake magnitude scale. *Bull. Seismol. Soc. Am.* **25**, 1-32.
- Sato, H. (1977). Energy propagation including scattering effects, single isotropic approximation, *J. Phys. Earth.*, **25**, 27-41.
- Solov'ev, S. L. (1965). Seismicity of Sakhalin. *Bull. Earthquake Res. Inst., Univ. Tokyo* **43**, 95-102.
- Tsumura, K. (1967). Determination of earthquake magnitude from total duration of oscillation. *Bull. Earthquake Res. Inst., Univ. Tokyo* **45**, 7-18.
- Valdes, C.M., and Novelo-Casanova, D.A. (1989). User manual for QCODA. *IASPEI Software Library*, **1**, 237-255.
- Wu, R.S., and Aki, K. (Editors) (1988). "Scattering and Attenuation of Seismic Waves", *Pure Appl. Geophys.*, **128**, nos 1/2.

15. PCEQ and QCODA

by

C. M. Valdes
Institute of Geophysics, National University of Mexico
Mexico City, Mexico

INTRODUCTION

Collecting seismograms is only the first step in studying an earthquake. Once you have acquired some seismograms you need to process the data. In the past, I often had to process seismograms on smoked paper. It is very difficult to work with these analog seismograms. They are wonderful tools and have some advantages over digital seismograms but at the same time, it is very difficult to keep up with the processing when the seismicity is at a very high level.

Many years ago, I worked with the HYPO71 earthquake location program, first in Mexico and then in Wisconsin. It was a very difficult program to use. Our computer had only 100 Kbytes of memory and the program required 135 Kbytes, so we had to wait until night to run the program because otherwise, we couldn't fit the program into the main core memory of the computer.

When I met Willie Lee, he had an AT computer sitting on his desk and he wanted to be able to run HYPO71 on his AT. I didn't know what to think! It didn't seem possible to run such a large program on such a small computer, but he told me that his computer had 512 Kbytes of memory; more than enough to hold the program. It seemed to me that this type of computer would be able to help many people doing seismology.

What I want to do today is give an overview of the PCEQ and QCODA programs. This is intended to be a rather general review since they are both published in one of the IASPEI volumes (Lee, 1989). What I would like to do now is put more emphasis on what kind of research can you do with those programs.

DIGITAL SEISMOGRAMS VS ANALOG SEISMOGRAMS

Using digital seismograms allows you to do much more quantitative analysis than is possible with the analog seismograms. What I want to talk about first is the difference between an analog and a digital seismogram. Of course, most of you probably know the difference between the two. If you have worked with analog seismograms before, you know that one of their advantages is that you can immediately get a feel for the general level of seismicity just by looking at an analog seismogram. You go out in the field and remove one of the papers, and you can tell right away how many earthquakes there were, and whether they were large or small (based on coda duration). You can also tell immediately if the instrument didn't work at all for whatever reason. That is an advantage that I always liked, and I would like a system that could produce both a paper record and a digital seismogram.

Also, with analog seismograms, we know how long it takes for the storage media to fill up. For instance, we know that with a particular recorder, in five days the pen will be exactly at the end of the paper. Of course, if there are a lot of earthquakes in that time period it is going to be very hard to read, in detail, all the characteristics of the waveforms. Also, once we have recorded at a particular scale, we can not change the magnification easily.

When I was in Mexico we had a special room in which we would project the seismograms on a large screen. Then we would manually pick points off the projected image and punch the information. That was a way to digitize an accelerogram, but I remember that it would take us at least two hours to digitize one minute of the trace. Those were very low frequency seismograms, so you can see that it would be almost impossible to do anything with the relatively high frequency seismograms recorded by a local network.

Another problem with the analog seismograms was the limited number of parameters that we could extract from them. We could clearly distinguish the break of the P-arrival and we could also distinguish the S-arrival. We could measure the amplitudes, assuming that the dynamic range of the instrument had not been exceeded and the peaks clipped. We could see the polarity of the first motion and invert that to get a focal mechanism. We could also measure the coda duration of the seismogram. However, it was very difficult to tell anything about the frequency except in a very qualitative way. This was a shame, since there is a lot of information about the rupture and about the medium contained in the waveforms. Essentially, we throw away 98% of our information by keeping only those few parameters I mentioned.

A digital seismogram is just a stream of numbers that represent the waveform. The ground motion is converted into a voltage and the level of the voltage is sampled periodically. Each sample is then converted into a binary representation of the voltage at that instant in time. One problem with the digital data is that it requires some processing in order to show the seismogram. This is necessary since the digital seismogram is only numbers, and nobody can just look at the stream of numbers and easily tell where the various arrivals are.

This is one of the few shortcomings of the digital method. However, it is possible to just hook a printer or some similar device to the computer and get a hard copy seismogram automatically. I am not sure if anyone has done this, but it certainly is possible.

Storage capacity is also a concern related to the digital method. We currently use magnetic media for most of our storage. In the early days that was a very expensive type of medium, but today it is relatively cheap. You can see that at a sample rate of 100 samples per second, a one minute long event will require 6000 samples to be stored for each channel. Without high capacity storage devices there would be a significant problem. Today the storage cost of magnetic media is actually cheaper than the cost of printing the data onto paper records, so it is less expensive to save all our records digitally. Even so, one must be certain that the size of the storage medium is great enough that it will not be saturated while the computer is running unattended. Recent increases in hard disk capacity are making this much less of a problem than it was in the not-so-distant past.

A major advantage of digital recording is that many more types of analysis are possible with digital data than with analog data. Before the advent of digital recording, we were restricted to estimates of hypocenter, magnitude, and first motion. There was really very little information that could be extracted which could lead to a better understanding of the earthquake mechanism or possibly be applied to the prediction of earthquakes. With digital data, the time scale may be altered at will. This allows us to see the ground motion in great detail and to easily pick first motion time and polarity. Filtering of the trace is also possible with digital seismograms. It is a straight-forward matter to convert the time domain information into frequency domain data, and see how the earthquake motion was distributed

by frequency. Thus we can extract many more parameters from the seismogram than was possible with the analog records. With digital data we are moving into a more quantitative type of analysis. We can estimate various parameters related to the rupture process as well as estimate properties of the earth through which the seismic waves have passed.

PROCESSING DIGITAL SEISMOGRAMS USING PCEQ

Because there are so many possibilities that come from recording digital data, the need arose to develop a convenient way to carry out the processing. Otherwise, we would soon be inundated with more seismograms than we could process, and instead of storing paper records we would just be storing large quantities of unprocessed digital seismograms. PCEQ is a program that I developed to facilitate convenient processing of digital seismograms. I tried to make the program as interactive as possible, and to minimize the amount of typing required to interact with the computer. I realized that if much typing was required, it was going to be very painful to process a lot of seismograms. PCEQ was designed to process waveform files in the format produced by XDETECT or MDETECT. However, it is a relatively simple matter to convert any storage format to another as long as the specifications for each format are known, so PCEQ can actually be used to process earthquake waveforms recorded by many different systems. The program is designed to run on a PC, PC/XT, or PC/AT. The computer must have 640 Kbytes of memory installed. Although the program will run on a PC or PC/XT, it will be very slow on those machines. The system must be equipped with a mouse, and the mouse driver must be loaded before the program is started. The reason for the mouse is that the seismogram is spatial in nature, and

a mouse makes it easy to move from one part of the seismogram to another. Willie Lee always suggested that there should be an alternative way of interacting with the PCEQ program via the "arrow" keys, but that has not been done at this time.

I will describe a few of the characteristics of the PCEQ program. There are other programs that do similar things, and if you can operate PCEQ, then probably you can also use these other programs. Once the program is started and a file is loaded, you are presented

- PCEQ typical display

IRIG	PHPM	PBWM	PUCM	PHRM	PAGM	PTAM	PLOM	PANM	PARM
PSTM	PHOM	PSTM	PPFM	PPTM	VEGM				

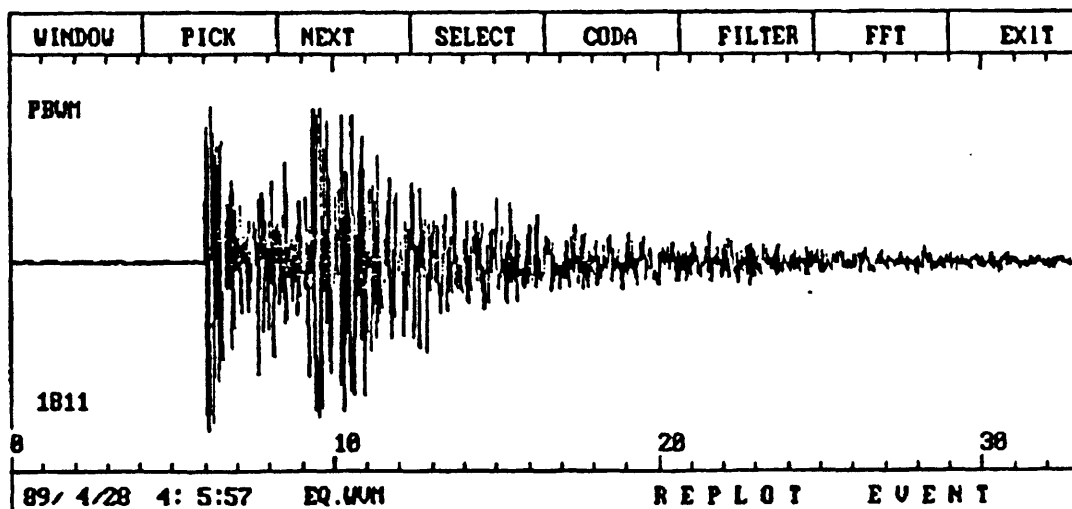


Figure 1 Typical PCEQ display.

with a screen that looks something like Figure 1. At the bottom of the screen is the window showing the complete event for one channel. Along the bottom, the time is plotted in seconds. On the left hand side of the window are listed the station identifier and the

maximum amplitude of the signal in digital counts. Just above the seismogram are listed a set of analysis options, which may be selected by pointing with the mouse and clicking. Most of them are self-explanatory, but I will discuss a few of them here. At the top of the screen are listed the names of the stations in the station list. Here, only 16 are shown. If you have more, the empty space on the screen will fill with station names. It is possible to either scroll through the station list using the "NEXT" option, or to select randomly from the station list by clicking on the appropriate station name.

"WINDOW" is the parameter which allows you to select a portion of the seismogram and enlarge it. Figure 2 shows a two second interval from the last seismogram selected by using the "WINDOW" option.

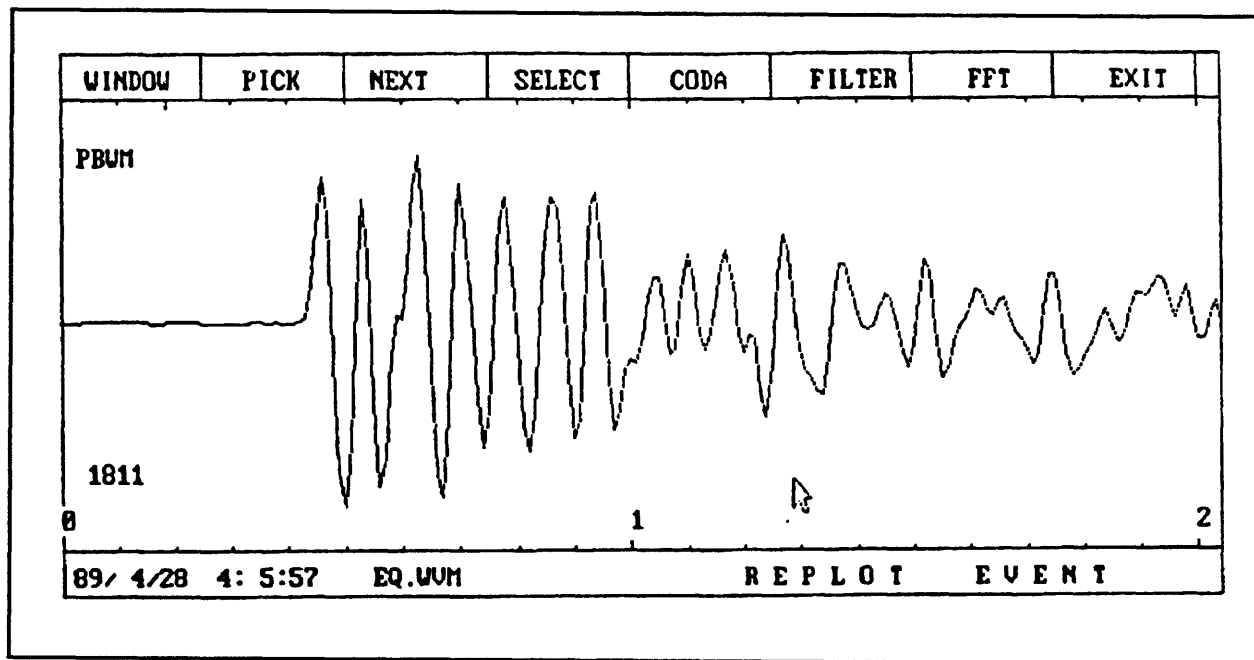


Figure 2 Expanded seismogram using "Window" option.

It is very hard to tell exactly what is happening within this interval when looking at the original trace, but on the windowed trace all the wavelets are quite distinct. It is very

easy to tell the first motion direction on the two second trace.

After selecting this two second window, we would probably choose the "PICK" option next. "PICK" allows you to pick either P-arrivals or S-arrivals and describe them. This information is written in the HYPO71 format to an output file.

HYPO71 is an earthquake program that was written by Willie Lee and John Lahr (1975). The format of the input and output files of HYPO71 is quite cumbersome, and people have complained about it for years. Even so, the program is a "standard", so PCEQ supports the HYPO71 format when writing its pick information. Once you select the "PICK" option, the menu items at the top of the window are replaced with options relevant to the picking process. This is shown in Figure 3.

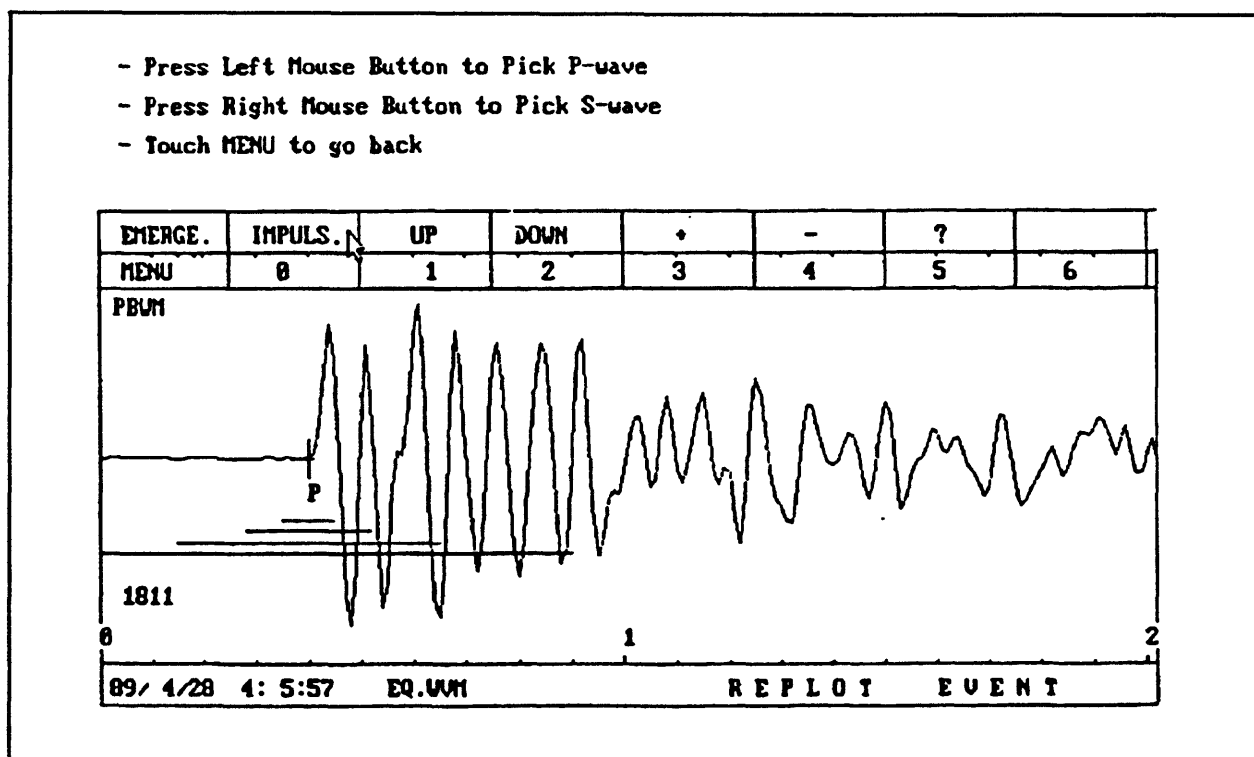


Figure 3 P-wave arrival time picking using the "Pick" option.

The actual pick position is produced by moving the mouse pointer to the place on the

waveform where you think the pick should be, and then pressing either the left button (P-wave) or right button (S-wave). A symbol appears centered on your pick location. If you decide to change the pick, just repeat the process. In that case the new symbol will write on top of the old one. The HYPO71 pick parameters are selected by moving the mouse pointer to the relevant box, and then clicking the left mouse button. You can mark the arrival as being either emergent or impulsive. Also various polarity descriptions recognized by HYPO71 can be chosen. The numbers in the second row of boxes represent the weight HYPO71 is to place on that pick when computing the hypocenter. Here 0 represents the highest weight. To return to the main menu select "MENU".

The "CODA" option on the main menu allows the user to select two points on the seismogram. One should be at the beginning and one at the point where we think the coda has died down to the background level. This information is used by the HYPO71 program to determine coda duration magnitudes (Lee et al., 1972).

Another useful selection is the "FILTER" option. This option allows the user to select an adjustable band-pass filter, apply it to the seismogram and display the result. This could be used, for example, to eliminate high frequency components so that the general character of the waveform is more distinct. Figure 4 shows the seismogram from Figure 1 after application of a 1 - 2 Hz band-pass filter. Comparing the two seismograms, we can see that the peaks of the low frequency energy do not correspond to the peaks of the unfiltered seismogram. That sort of information can be very useful. Often, we can not easily see where the S-wave first appears because the S-wave may be obscured by the P-wave coda. By viewing the seismogram using different filter combinations, it may be possible to see the S-

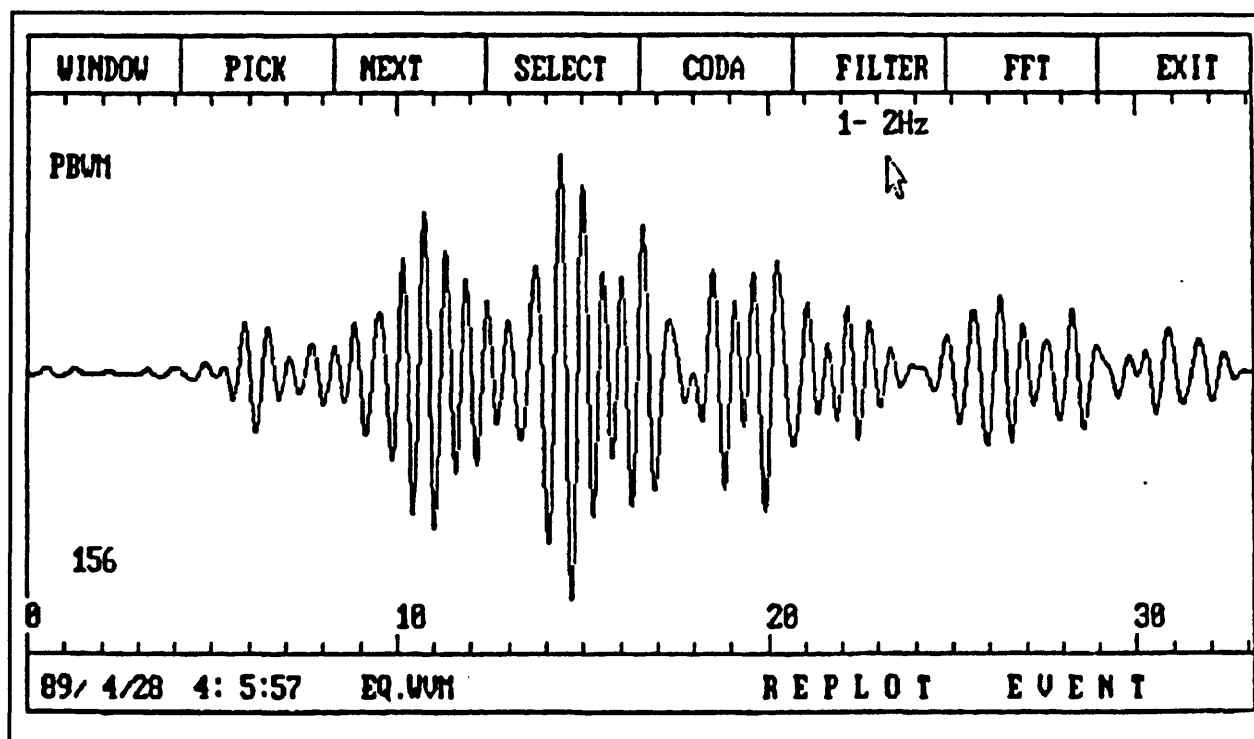


Figure 4 Bandpass-filtered seismogram.

wave arrival. This is one of the advantages of a digital seismogram. It is easy to do this kind of manipulation.

The next menu selection is called "FFT". Of course, this is an abbreviation for the fast Fourier transform. When this option is selected, whatever portion of the seismogram is in the current window will be transformed into the frequency domain. Figure 5 shows how the screen looks after this option is selected.

The top trace is a copy of the signal which is being transformed, and the bottom part of the display is the resulting amplitude spectrum. The reason for performing this kind of analysis has to do with modeling the earthquake source. In 1970 Brune developed a model for how the far field spectrum should behave for an ideal double-couple earthquake source. In Brune's model there are two parameters which describe the far field spectrum. These are the

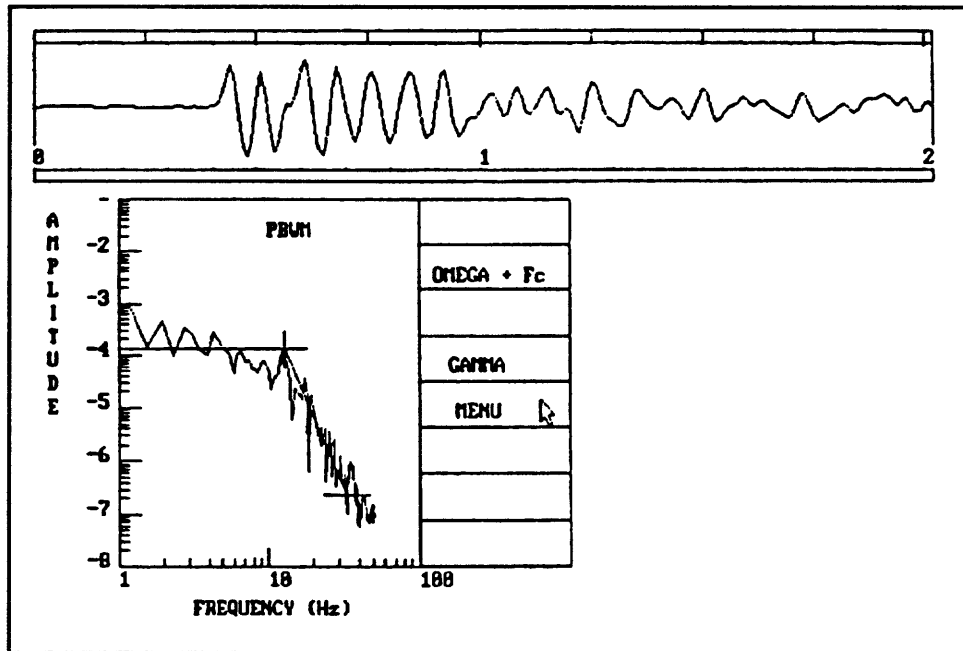


Figure 5 Seismogram and its power spectrum.

level of the low frequency flat part of the spectrum, and the corner frequency at which the spectrum begins to decay. By measuring these parameters, you can fit your observed data to a Brune model of the source, and in that way develop more information about the nature of the source rupture process. This is a type of analysis which is not practical at all for analog seismograms.

To make these measurements on your digital seismograms, use the mouse to select the "OMEGA + Fc" OPTION. Then move the mouse pointer to the height where you think the low frequency level is and the point where you think the corner frequency is, and click the left button. A line will then be drawn on the screen at the level you have chosen to the corner frequency point. If you do not like that location, just repeat the process. When you are happy with the position of the low frequency level, then select "GAMMA". This will

allow you to select the slope of the decay portion by moving the mouse pointer to the base of the decay portion of the plot, and pressing the left mouse button. Keep in mind that the scales on the spectral plot are logarithmic so relatively small changes in position of the low frequency level can result in rather major changes to computed estimates of the source parameters.

Brune's theory was originally developed for the shear wave so we should be looking at the S-wave. However, people have modified his formula so that it can be applied to P-waves and, as you can see, in this example the P-wave portion of an event is being analyzed. In any case, once we select a low frequency level and a corner frequency, the information is written into a file. No further processing is done within the PCEQ program.

How would we use these parameters in practice? Brune's formulation relates the seismic moment of the earthquake to these measured parameters. The seismic moment is a much more robust way of estimating the size of an earthquake than the coda duration, because it doesn't depend on the path through which the waves propagated. Of course the source must be reasonably modeled as a double couple for this estimate to have much meaning. The formula for the seismic moment M_0 is

$$M_0 = \frac{4\pi\beta^3 R\Omega_0}{2R_{\theta\phi}} \quad (1)$$

Here Ω_0 is the low frequency level, β is the shear wave velocity, R is the hypocentral distance, and $R_{\theta\phi}$ is the radiation pattern of the seismic waves. In Brune's model, the rupture is assumed to take place over a circular area. For this model it is possible to calculate the radius of the rupture zone. It is given by:

$$r = \frac{2.34\beta}{2\pi f_c} \quad (2)$$

Here r is the source radius, and f_c is the corner frequency.

One more parameter which we can calculate according to the Brune model is the stress drop. This is given by

$$\Delta\sigma = \frac{7M_0}{16r^3} \quad (3)$$

Here $\Delta\sigma$ is the stress drop. When you look carefully at these equations you can see how easily errors can be produced. f_c for instance, could easily be mis-located from 10 Hz to 15 Hz. That will change r significantly, and when we get to the stress drop calculation, r is cubed so there would be a large change in the computed stress drop. In this model, the assumption is that all the stress is released at once when the fault moves. That is probably not close to what usually happens. Brune's model also does not take into account how the high frequency waves in the seismogram will appear. Even so, this sort of analysis can give you some idea of what is going on in the source region of the earthquake under study.

There is a very good paper by Archuleta and others (1982) called "Source Parameters of the 1980 Mammoth Lakes, California Earthquake Sequence." That is one of the best references for a description of how to use source parameters. Figure 6 is a plot taken from that paper.

Figure 6 is a rather complicated plot, but it provides us with a lot of information. Along the bottom of the plot is the source radius in kilometers. The left-hand vertical axis is seismic moment (M_0), the top horizontal axis is corner frequency (f_c), and the right-hand

vertical axis is low frequency level(Ω_0). The diagonal lines are lines of constant stress drop. These are essentially the parameters we obtain from the preceding analysis. The large cross-hatch is the set of error bars. You can see that they are very large, at least in part because of the uncertainty in the waveforms.

If you have only analog seismograms you will only be able to get magnitudes. However, many people have developed relations between seismic moment and the various magnitudes in common use. These

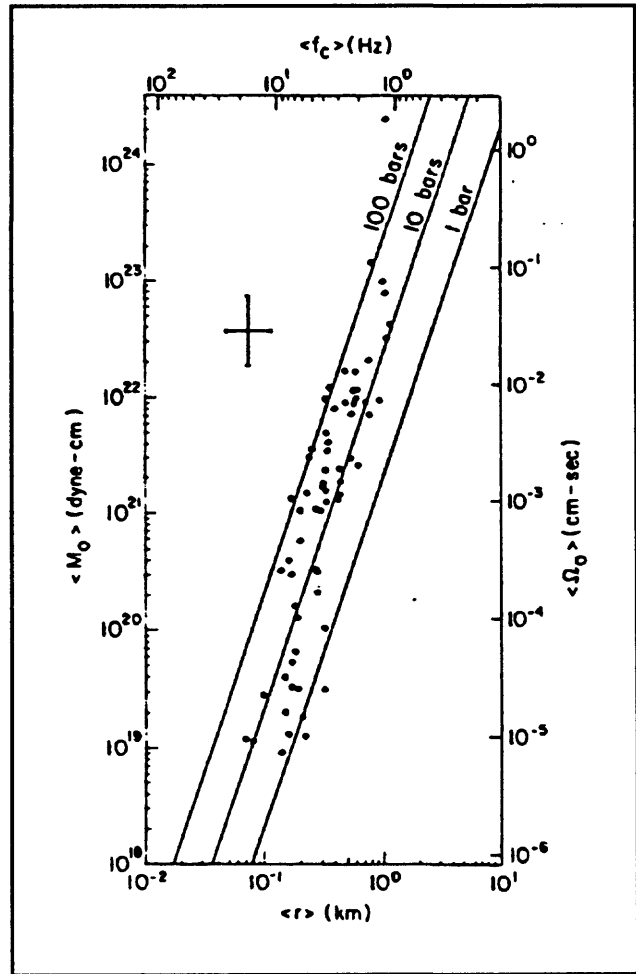


Figure 6 Relation between source parameters.

relations are useful for relating data sets obtained with analog seismograms to those obtained with digital systems. A seismic moment of about 10^{19} to 10^{20} will be equivalent to a magnitude of about 1 to 1.5. A large earthquake might have a moment of 10^{25} to 10^{27} . As you can see from the figure, earthquakes with higher moments seem to have higher stress drops than do earthquakes with lower moments. That is, earthquakes with moments greater than about 2×10^{20} have stress drops between 10 bars and 100 bars, while earthquakes with smaller moments have stress drops mostly below 10 bars.

When we are working with a digital seismogram, we need an instrument response correction. Each instrument has its own characteristic response curve, and we have to correct for that or else we will bias our seismograms. Typically, the shape of the response curves are very simple. The program will allow you to make this correction. It will ask you for the name of the file containing the instrument correction data. The format is very simple, and is explained in the manual (Valdes, 1969).

The input files for the program are typically specified with a .WVM extension, and are produced by the MDETECT or XDETECT program. Otherwise, you may create a waveform file from digital data you have recorded with another system. The output files produced by the PCEQ program are PHASE.OUT and SPEC.OUT. PHASE.OUT contains the arrival times and other pick information as well as the coda duration of the event. The information is written in HYPO71 format, so that it is relatively easy to perform a hypocenter location with the data from this program. SPEC.OUT contains the spectral parameters of the waveforms.

The program was written in FORTRAN with a few routines in C. It was also written in a modular form. So you could easily write your own routines for some analysis you wish to perform and link them into this program as subroutines. The graphics used in this program are from the Metawindow subroutine library. The Metawindow library supports a wide range of display adapters so it was relatively easy to make this program work on PCs with different graphics hardware. We wanted to make this program as portable as possible.

PROCESSING SEISMOGRAMS WITH QCODA

The quality factor Q obtained from coda waves Q_c has been an important tool in seismic attenuation studies and recently in earthquake prediction. The QCODA program is designed to allow you to easily measure Q_c using digital seismograms collected with a PC system.

QCODA could have been added as another window within the PCEQ program if there was more memory available under DOS. However, it could not be implemented that way. Essentially, this program uses the main core of PCEQ. It looks very similar, but you notice that a lot of the windows are not there. For instance there is no picker window, no spectrum window, and no filter window. QCODA is also an interactive program, although there is very little for the user to do once a few parameters have been selected. The purpose of the program is to calculate the seismic wave attenuation based on the coda waves. The coda waves sample a much larger volume of rock than do either the direct P or S waves, so coda waves can provide a better picture of the attenuating properties of the rock mass.

The part of QCODA which actually performs the analysis was written some years ago by W.H.K. Lee (Lee et al., 1986). What I did was give the program an easy to use interface. Essentially, things operate as they do in PCEQ. The program estimates coda Q using several different models. It uses the back scattering model of Aki and Chouet (1975), the single scattering model of Sato (1977), and the single scattering model in the frequency domain due to Phillips and Aki (1986). The reason we chose to use all three methods is that coda Q is such a difficult thing to measure that we can have more confidence in the measure if several different methods produce similar estimates. Typically, the beginning of the S-

wave coda should begin at about twice the travel time of the S-wave. So the algorithm by Aki and Chouet starts analysis at twice the S-wave travel time. When Lee et al., 1986 coded the algorithm, he included a correction factor due to Sato (1977), that will allow you to use the entire S-wave in the analysis. The advantage of doing that is that sometimes the decay of the S-wave is quite rapid, and if you wait until twice the S-wave arrival time, it has already decayed to close to the noise level. If that happens, we can not extract much information from the waveform. We have to go back farther in the trace to be able to analyze some of the small events. Two of the coda Q routines are modified to do this analysis early in the trace so actually, there are five distinct analysis routines included in QCODA. All of them are measuring the same parameter in slightly different ways.

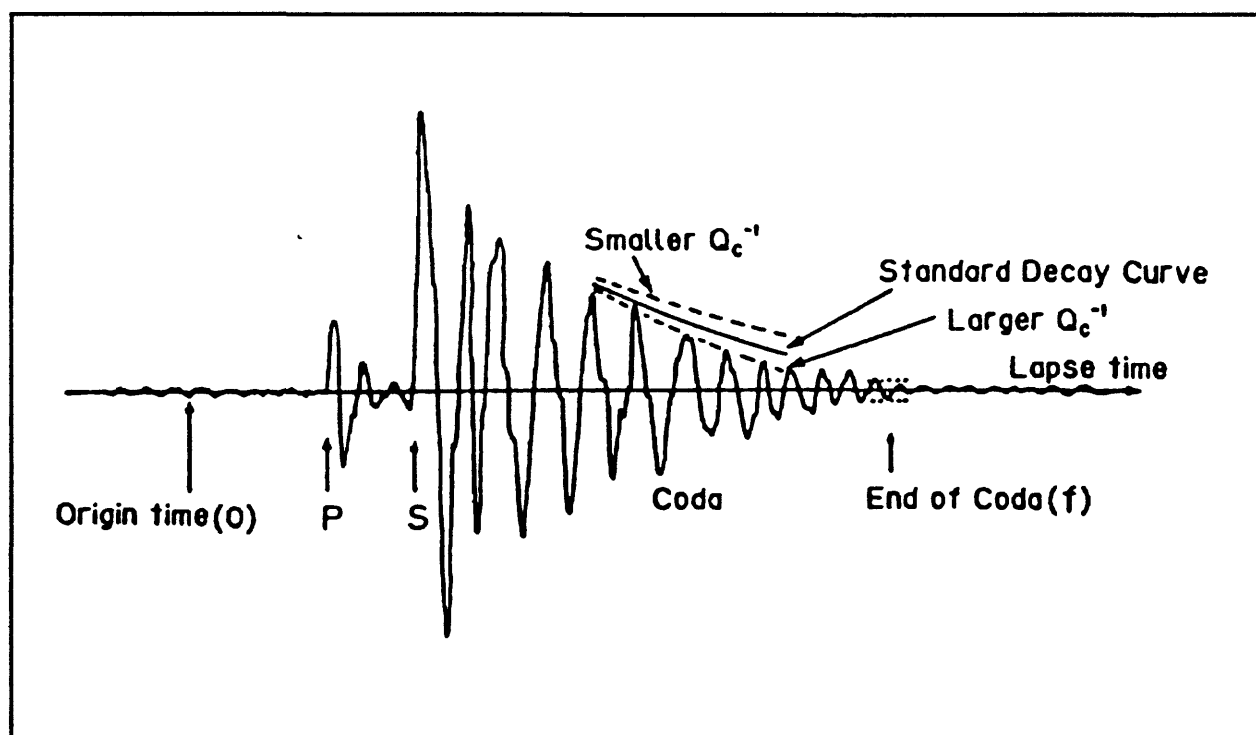


Figure 7 Effect of attenuation on seismogram.

Just looking at a seismogram you can get some idea of what the Q of the rock mass is. As shown in Figure 7, if the seismogram "rings" for a long time then you know that the attenuation is relatively low. On the other hand, a seismogram which dies out very quickly indicates that the attenuation is relatively high. So you might wonder why you can't estimate Q by some method, which essentially measures how long it takes for the signal to die out. The reason is that Q is a frequency dependent parameter, and the coda of a seismic wave is a mixture of frequencies. The very high frequencies have short wavelength, and these short wavelengths are going to be scattered very differently than the long wavelength waves from the low frequency components of the signal. In order to get an unambiguous estimate of Q , you need to see how it behaves as a function of frequency.

In our analysis routines we use fairly tight frequency bands. The center frequencies of the bands are 1.5, 3.0, 6.0, 12.0, and 24 Hz. When we look at the higher frequency waves we are going to be seeing the effects of small scatterers. Conversely, the low frequency waves will show the effects of large scatterers.

I want to spend a little time here showing you a few examples of the things people have done using coda Q . One of the things people have been doing with Q is seeing how it changes with time in a particular region. There is a belief among some investigators that as an area comes under higher and higher stress, the attenuation properties of the rock mass, particularly for high frequency components will change significantly. If that is so, we should be able to see those changes reflected in significantly different values for high frequency Q estimates.

Figure 8 shows the results of a study by Lee et al. (1986) of the Round Valley

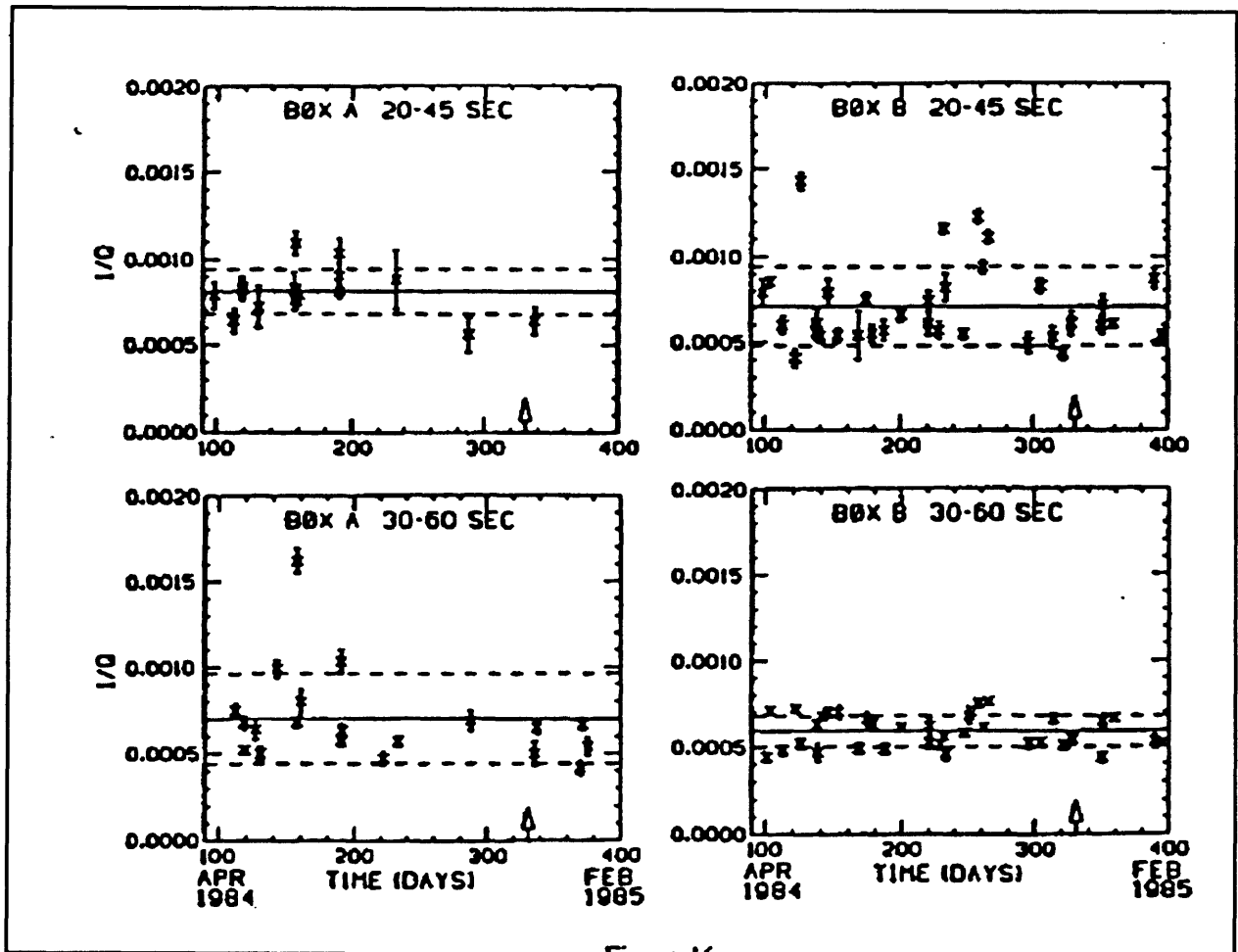


Figure 8 Coda Q Analysis from Round Valley earthquake.

earthquake. The figure shows values of Q_c obtained at two different times within the coda of the earthquake. The top row shows the results for a window from 25 seconds to 45 seconds into the event and the lower row shows the analysis results for a window from 30 to 60 seconds into the event. The analyses on the left hand side were performed on signals collected from within the aftershock region and the right hand side results were collected outside the aftershock region. The plots are of attenuation versus time. As you can see from the error bars on the graphs, the potential size of the error is of the same order as the spread in the estimates. Q is not an easy thing to measure!

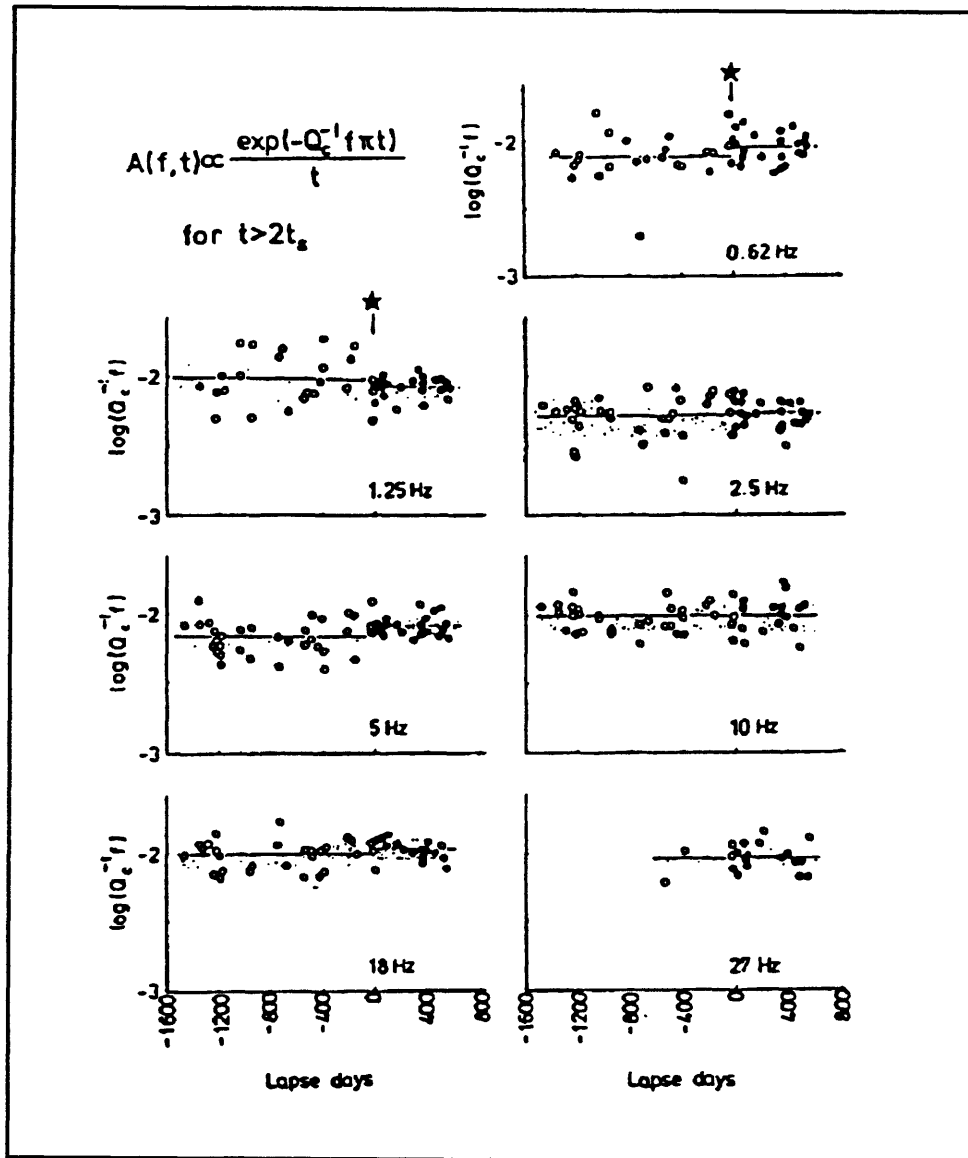


Figure 9 Temporal variations of Q_c by frequency band.

Figure 9 shows temporal variations of $\log(f/Q_c)$ for seven frequency bands before and after an earthquake in central Asia from a study by Sato et al. (1988). The frequency bands that they used were centered at 0.6, 1.0, 2.0, 5.0, 10.0, 15.0, and 27.0 Hz. The fore shocks are open circles and the after shocks are solid. The horizontal lines running through the data

are centered on the mean of that group. The shaded areas on the plot define the region within one standard deviation of the mean. The star near the top of the plot indicates the time of the earthquake. Their results are that there is no significant change in coda Q before and after the earthquake. However, it looks like for some frequencies there may be a slight change in the mean after the main shock.

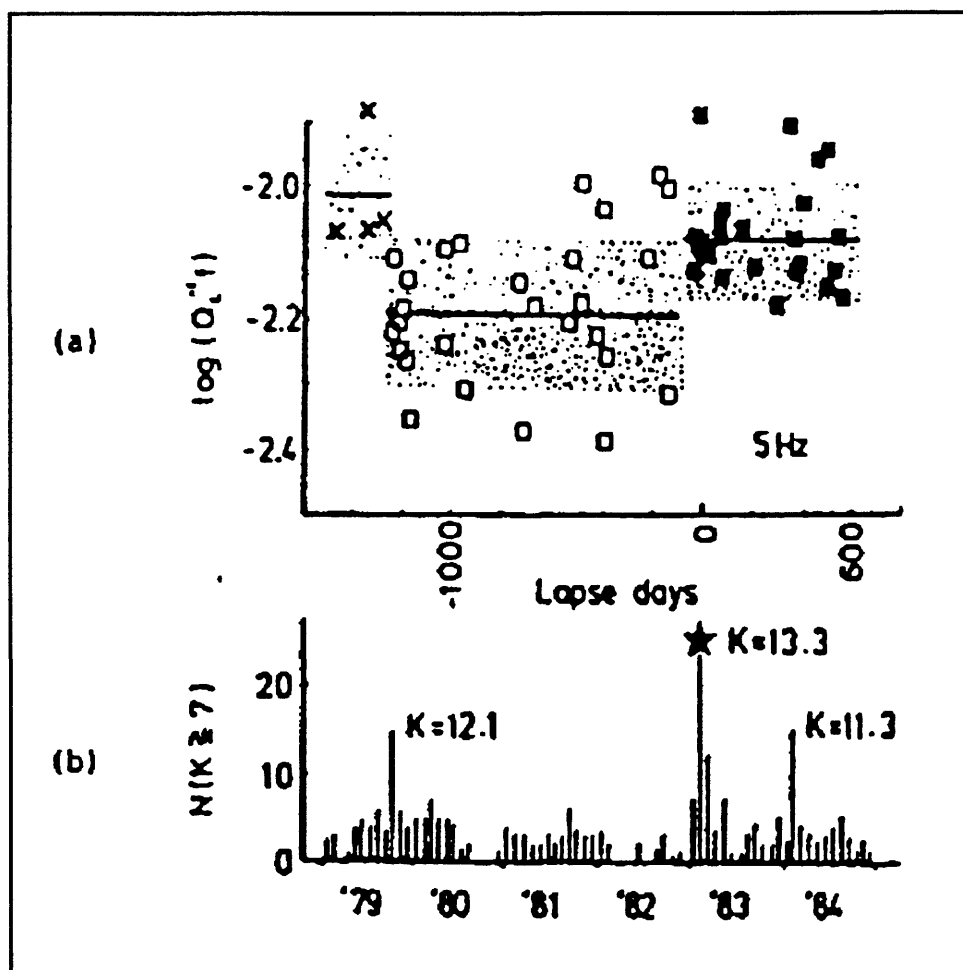


Figure 10 Detail of the 5.0 Hz band from Figure 9.

Figure 10 shows detail from Figure 9 for the 5.0 Hz band. Again, the mean and standard deviation are plotted. Although a formal statistical test might not show a temporal

dependence of Q_c , there certainly appears to be a sustained drop in $\log(f/Q_c)$ in the months preceding the mainshock followed by a return to the levels that existed prior to Sept. of 1979.

When you start QCODA it asks you for the travel time for the S-wave. That is the only information that you have to provide it with. Then the program will start producing estimates of the coda values for the different frequencies.

There are two output files for QCODA. The first is called QCODA.OUT. it provides all the information related to coda Q in a format that explains the calculated values and their errors. QCODA.SUM gives the same information written in a much more compact way. This is so that the output information may be easily input into some other program without having to strip out all the comments. There is one line per coda estimate in this file.

REFERENCES

- Aki, K., and Chouet, B. (1975). Origin of coda waves: Source, attenuation, and scattering effects. *J. Geophys. Res.* **80**, 3322-3342.
- Archuleta, R. P., Cranswick, E., Muller, C., and P. Spudich, (1982). Source parameters of the 1980 Mammoth Lakes, California, earthquake sequence, *J. Geophys. Res.* **87**, 4595-4607.
- Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *J. Geophys. Res.* **75**, 4997-5009.
- Lee, W. H. K., Aki, K., Chouet, B., Johnson, P., Marks, S., Newberry, J. T., Ryall, A. S., Stewart, S. W., and D. M. Tottingham, (1986). A preliminary study of coda Q in California and Nevada, *Bull. Seism. Soc. Am.* **76**, 1143-1150.
- Lee, W. H. K., Bennett, R. E., and Meagher, K. L. (1972). A method of estimating magnitude of local earthquakes from signal duration. *Geol. Surv. Open-File Rep.* 28.
- Lee, W. H. K., and J. C. Lahr, (1975). HYPO71 (revised): A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes. *Geol. Surv. Open-File Rep.* 75-311, 1-116.
- Sato, H. (1977). Energy propagation including scattering effects, single isotropic approximation, *J. Phys. Earth.*, **25**, 27-41.
- Sato, H. (1988). Temporal change in scattering and attenuation associated with the earthquake occurrence--a review of recent studies of coda waves. *Pure and Appl. Geophys.* **126**, 465-497.
- Valdes, C.M. (1989) User manual for PCEQ, IASPEI Software Library, 1, 175-201.
- Valdes, C.M. and D.A. Novelo-Casanova (1989) User manual for QCODA, IASPEI Software Library, 1, 237-255.

16. MATHEMATICS FOR WAVEFORM ANALYSIS

by

A. Lomax
Seismographic Station
University of California, Berkeley, CA 94720

INTRODUCTION

What use are seismograms? What do we do with seismograms? At least in my work, there are two basic ways in which they are used. One is visualization and one is processing. Visualization is a more "human" processing technique in which we look at graphical output and sometimes numbers, interpreting data, looking for correlations and anomalous features. It is on the whole, an intuitive process. For this sort of analysis, it is very useful to have the interactive, visual tools which in the last few years have become very powerful. Another aspect of using seismograms is actually processing them; sometimes with very complex algorithms and sometimes with simple manipulations. In practice, anything you do with seismograms is actually a combination of these two. Visualization requires processing in order to view the seismogram in the way you want. Any processing involves decisions which may be based in some way on things you have seen in the seismograms. They are not really completely separate.

OPERATIONS ON SEISMOGRAMS

A seismogram is a two or more dimensional object that you may want to view in many different ways in order to find out what kind of information it contains. Viewing a

seismogram involves transformation and display in graphical form. With that display you take measurements and make observations. It is a circular process. Conclusions you draw after viewing the seismogram may lead you to make additional transformations to the data which can lead to further transformations and so on. Hopefully, this all leads to something interesting in the end.

Very basic transformations which may be made on a seismogram are shifting and time scaling. Mathematically these operations take the form:

$$A'(t') = cA(t) + a; \quad t' = t + t_0 \quad (1)$$

Of course, in practice, you wouldn't change the actual data of the seismogram. There would be no sense in that. But the data being sent to the display may be shifted in this way.

Generally you start with a time series which is in the form of an amplitude as a function of time. For real data this takes the form of a group of numbers representing the amplitudes, and a group of numbers representing the discrete times corresponding to those amplitude numbers. Mean removal is a process that frequently must be applied to this set of numbers prior to more complex processing. Mathematically, this looks like:

$$A'(t) = A(t) - \frac{1}{N} \sum_{j=1}^N A(t_j) \quad (2)$$

Another basic transformation that is useful when you have multi-component data is component rotation. This allows you to get out of the fixed reference frame in which the seismometers are mounted. Typically a three component seismometer is set up with one component vertical, one component north-south, and one component east-west. By mathematically rotating the components you have, in effect, changed the way the

seismometer was set up in the field. The transformation is not identical, but, if there is no noise or other problems, it is practically identical. In two component form, the component rotation formula is:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3)$$

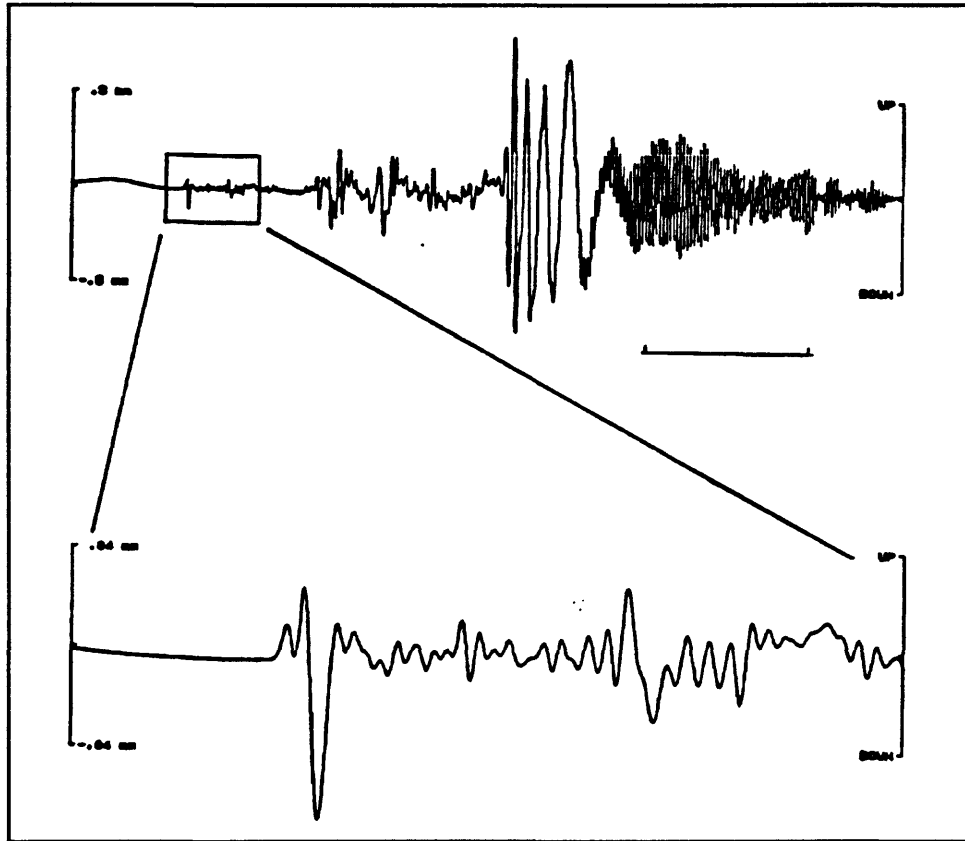


Figure 1 Scaling and shifting a seismogram.

Figure 1 shows the effect of scaling and shifting a seismogram. The top trace is a complete teleseismic record that is about an hour long. I might be interested in examining the initial part of the seismogram shown by a rectangular box. The lower portion of the figure shows a scaled and shifted version of the data inside the box. As you can see, the vertical

scale has been expanded and the horizontal scale has been expanded. Also, the record is shifted by several minutes. Intuitively, in the sense of viewing, I call this a "zoom." As I mentioned before, the mathematical operations are performed only on the data being sent to the display, not to the original data.

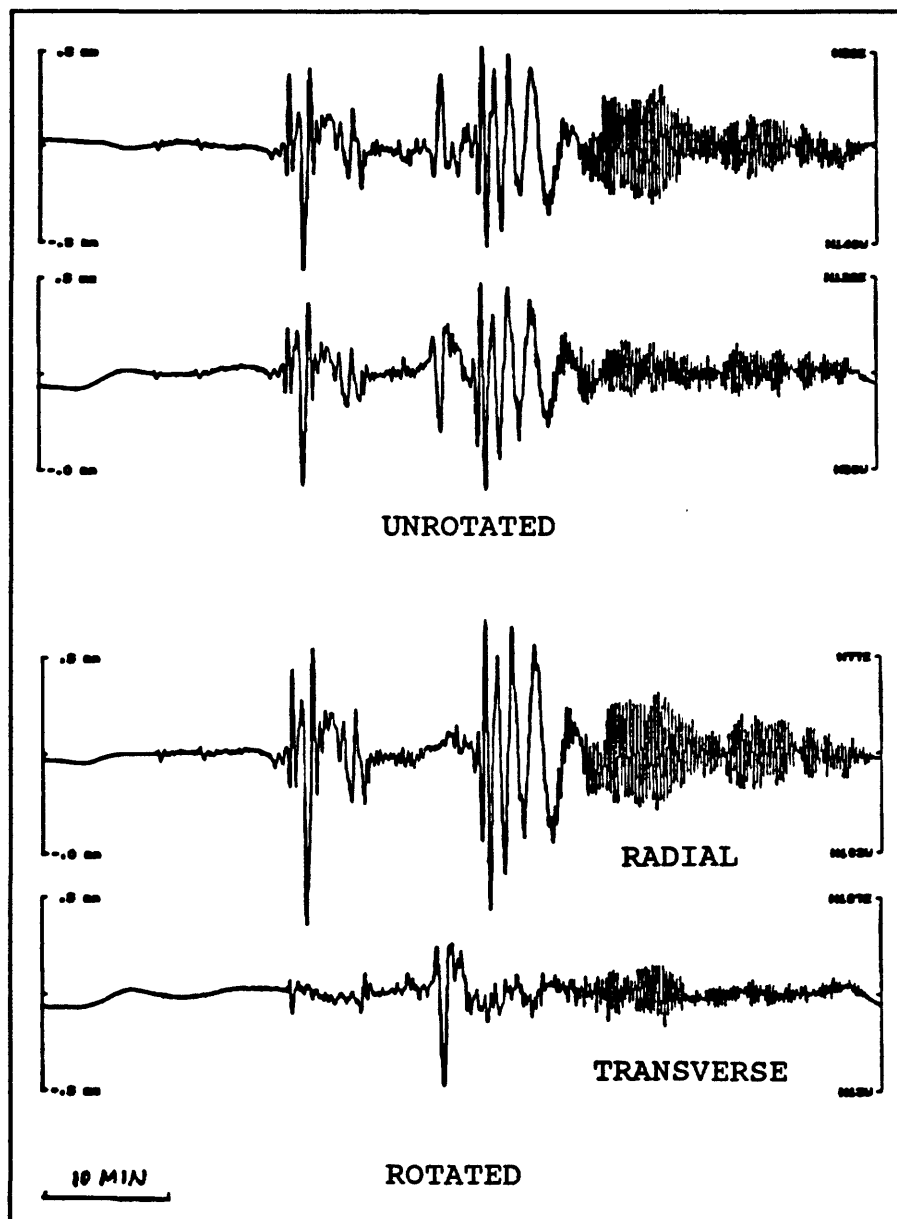


Figure 2 Unrotated and rotated horizontal component Seismograms.

Figure 2 shows an example of component rotation. This is the same teleseismic record as Figure 1, but showing the two horizontal components. The components are oriented corresponding to the instrument mounting directions. All of the phase arrivals show up on both components. There is nothing particularly unique. However, for most earthquakes, particularly teleseismic events, you may want to rotate components so that the horizontal components are oriented with respect to the direction to the earthquake. The bottom part of the figure shows the same two seismograms rotated in this fashion.

Now the top component is showing ground motion towards and away from the epicenter and the lower seismogram shows motion transverse to the path. Now you can clearly identify the S-wave, Love-wave, and Rayleigh waves on the two seismograms. So in the early stages of analysis, just by rotating the coordinates with respect to the epicenter direction, you can see a lot more information; particularly the identification of phases or polarization.

OTHER TRANSFORMATIONS

A couple of other transformations (Figure 3) that are commonly used are integration and differentiation of the time series. Integration typically accentuates the longer periods. This tends to emphasize the surface waves. Differentiation emphasizes the higher frequency body waves in the seismogram. Something I will talk about quite a bit is the Fourier transform. This transform takes a time series, and converts it into a new series in frequency. It breaks the time series into its spectral components, so the information is seen as a function of frequency instead of time.

Figure 4 shows another aspect of integration and differentiation of seismograms. The center trace is a recording of ground velocity as a function of time. By integrating the series we get the displacement time series, as shown on the top trace. On the other hand, we can differentiate the series, and get the acceleration, as shown on the bottom trace. Also, you

Other Transformations

Integration, Differentiation

Integration brings out longer periods
(surface waves).
Differentiation emphasizes higher frequencies
(body waves).

Fourier Transform

$$X(t) \Leftrightarrow \bar{X}(\omega)$$

Breaks time series into spectral components.
Characterizes time series as a function of
frequency instead of time.

Figure 3 Other transformations.

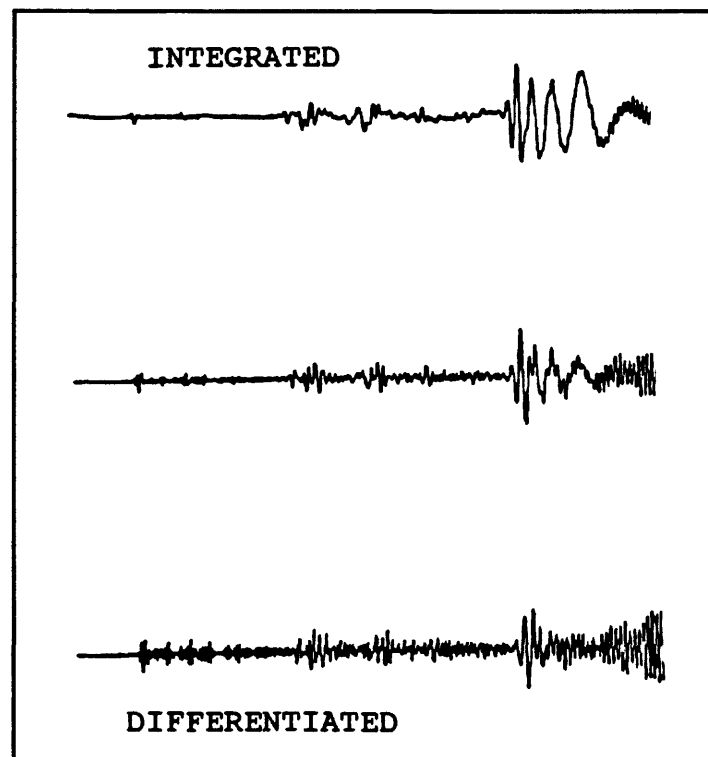


Figure 4 Integration and differentiation applied to seismogram.

can see on these traces that the integrated series shows enhanced low frequencies and the differentiated series shows enhanced high frequencies. A good example I have seen lately are the strong motion records from the Loma Prieta earthquake. Strong motion records are usually acceleration. Someone at Berkeley has done a very good job of integrating a set of these records into displacement records. You can get a lot of numerical problems when doing something like this, but in this case the results are quite good. What you end up with are a set of records that look much more like an earthquake to a seismologist. The acceleration time series is much more complicated and difficult to interpret than the displacement seismogram.

SEISMOGRAM DISPLAY

You may want to view time series in different ways for different purposes. The basic display is one component at one station for a single event. It will be in the form of amplitude versus time. If you have a multiple-component station like a three-component broadband instrument, you can display those together. There is a relationship between traces so it is an advantage to be able to view them together. There are other cases where you may have many stations for a given event or many events for a given station, and you construct a seismic section.

Figure 5 shows the same teleseismic event that was displayed earlier, but it is a complete three component set and also rotated into the coordinate system of the epicenter and station. The top trace is the vertical component, the middle is the horizontal component away from the epicenter and the bottom is the horizontal transverse component. When I work with

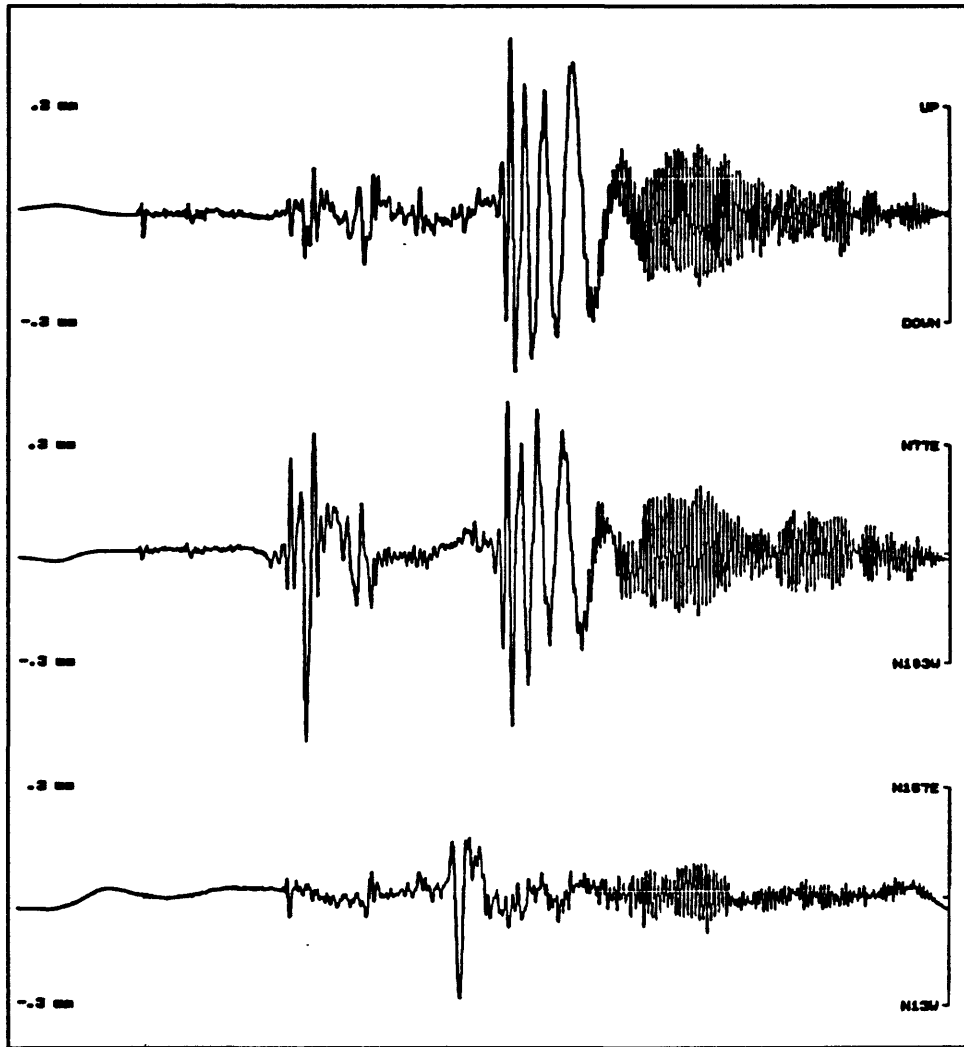


Figure 5 3-component teleseismic record.

three-component data, I always make sure that the amplitude scale on all three components is always the same. This way you can readily compare the amplitudes of the different components. However, many programs have a default setting, such that the traces are all scaled to a given height. I think that is something to be careful about, because you are losing a lot of information that is readily available otherwise. Little subtleties like this can sometimes skew your analysis, because of something you didn't see in the seismograms.

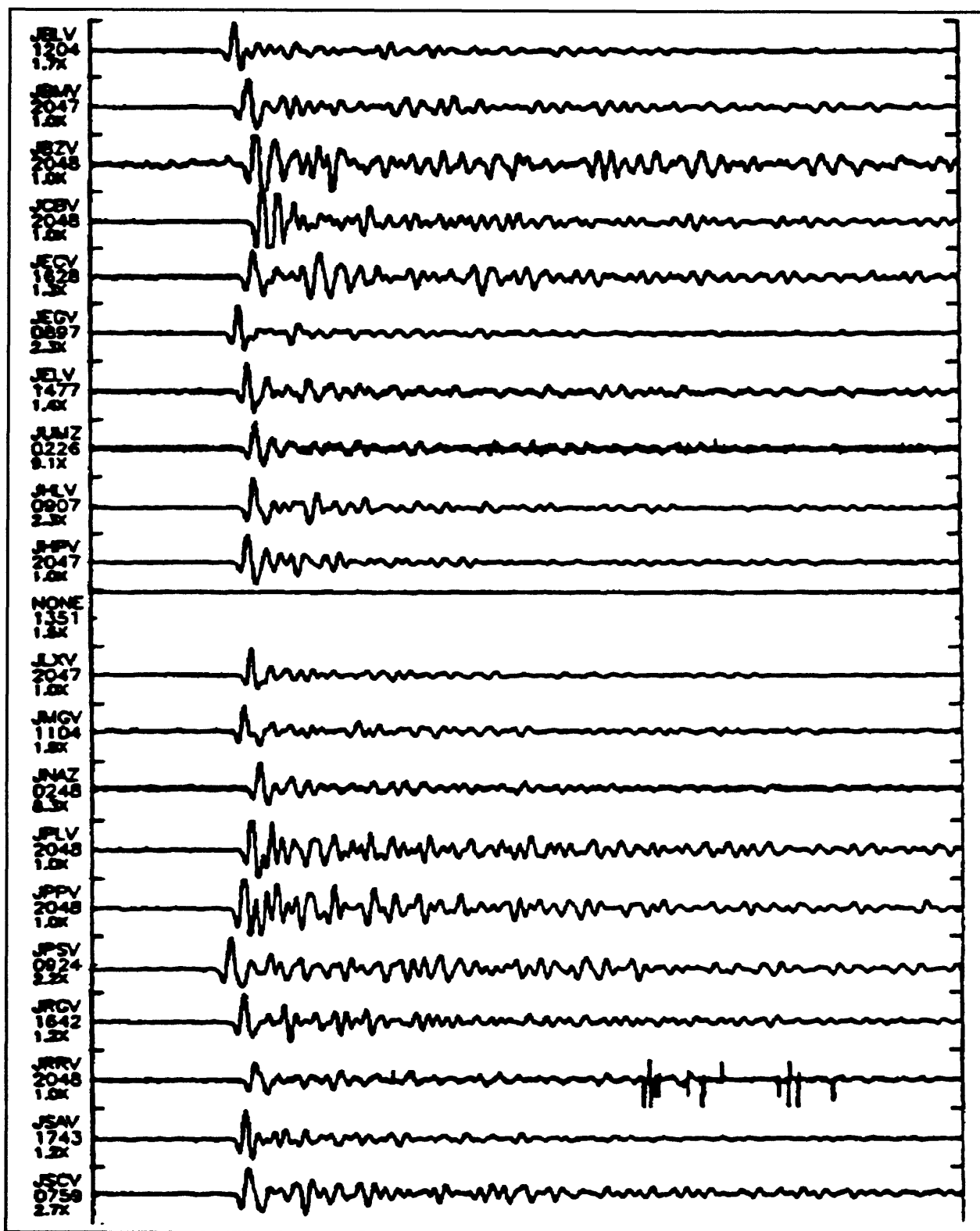


Figure 6 Example of a seismic section.

Figure 5 is an example of a seismic section. Usually seismic sections have the traces ordered in some manner. Here they are ordered so that they all start at the same time. This makes it rather easy to make correlations by eye of waveforms.

COORDINATE SYSTEMS

There is also the whole question of coordinate systems used for seismogram display. Most of what we have looked at so far is displayed in Cartesian coordinates. This coordinate system emphasizes amplitude of motion along fixed directions as a function of time, and is equivalent to the instrument output and to analog records generated by our instruments. You have to remember that the coordinate system is something imposed on the data. In fact, as I showed earlier, you may need to rotate out of a particular set of coordinates in order to see some of the information in the data. That is one of the key limitations of Cartesian coordinates. They favor a particular direction. However, this coordinate system is useful for phase picking and if properly oriented, for distinguishing P, SV, and SH motion.

A very different method of display is with polar coordinates. In a polar coordinate system, you are looking at a vector in space and considering its magnitude and the angles it makes in space with a horizontal plane, and with a vertical plane. The display in polar coordinates gets a little bit away from the directionality problems associated with Cartesian coordinates, though you still have some fixed directions that the angles are measured from. It is a little bit more useful for doing things like phase type identification and looking for off-axis effects like multi-pathing.

There are other fairly different displays. One example is time versus frequency which

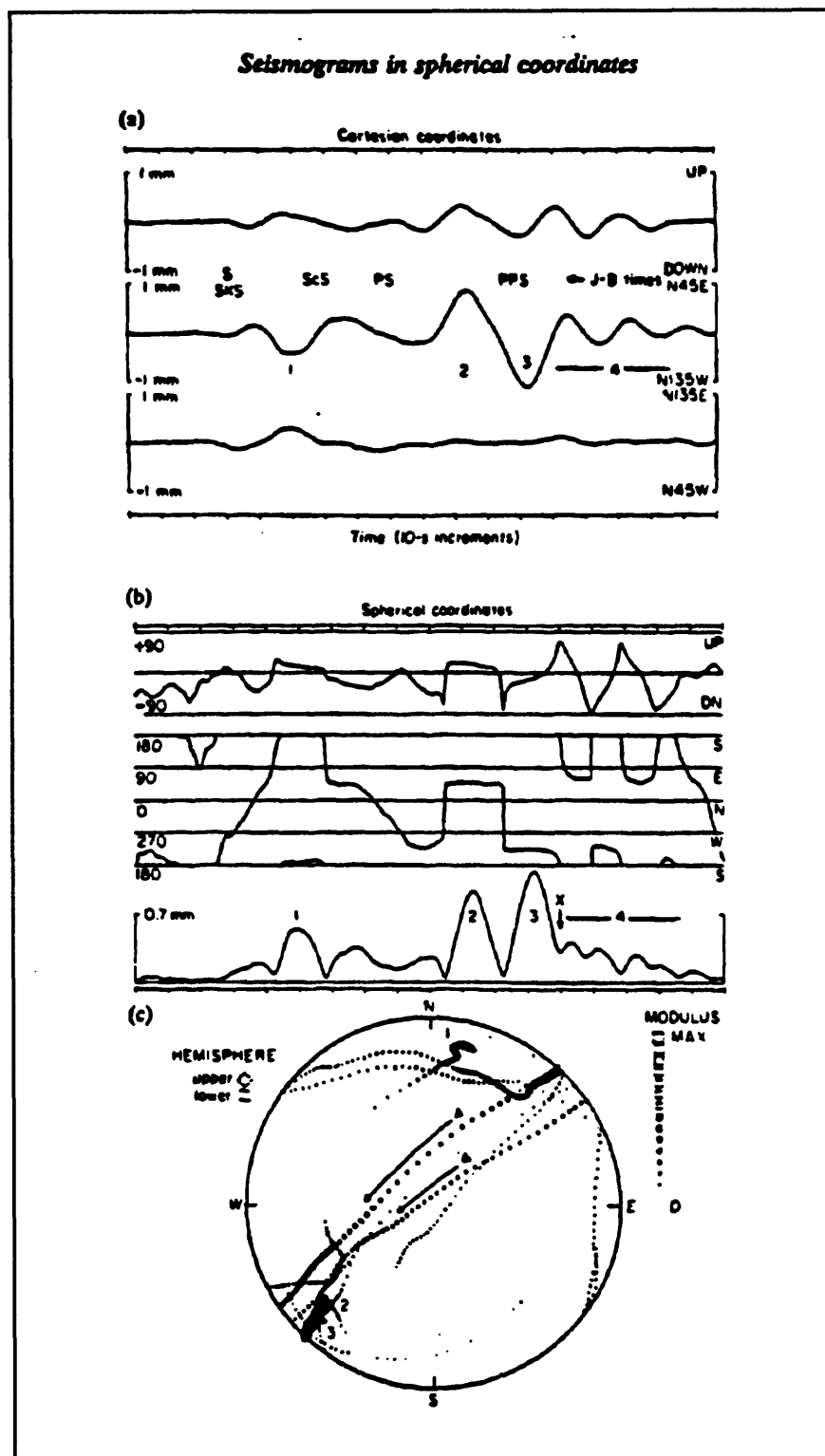


Figure 7 Seismograms in different coordinate systems.

is seen quite often. Here, instead of emphasizing amplitude as a function of time, you emphasize amplitude as a function of frequency at a given time. This is used for group velocity determination, looking at spectral content of phases, distinguishing signals from noise, and has seen quite a bit of use related to nuclear test ban verification.

Figure 7 [(a), (b), and (c)] is an example of a seismogram displayed in different coordinates. The top three traces [(a)] are the three components displayed in rectangular coordinates. The center set of traces [(b)] shows the seismogram plotted in spherical coordinates. The first of these traces is the angle from the horizontal, the second is the azimuthal angle, and the third is the amplitude of the motion. It is rather difficult to get used to using this kind of seismogram. However, the arrival of certain phases, which are difficult to see on the normal seismograms, have distinct character in spherical coordinates. There is a phase arrival in the last third of these seismograms. It is not distinguishable on the first set of traces but it shows up in spherical coordinates as the transition from a mostly "square wave" to a group of triangular pulses. This can be carried even further. The information can be plotted on a stereonet as shown in [(c)] of Figure 7. Here, each of the points represents one sample. The stereonet plot gives you information about the polarization of the waves which is a key factor in determining differences between certain phases. On this plot, the body waves all plot more or less as one point while the Rayleigh wave sweeps across the plot. As you can see this transformation brings out significantly different features in the data.

Figure 8 is an example of a time-frequency display. Here, the time axis is vertical and the horizontal axis is period. In this plot, the seismogram is decomposed into its different frequencies and those are plotted vertically. You can see that the high amplitude waves are

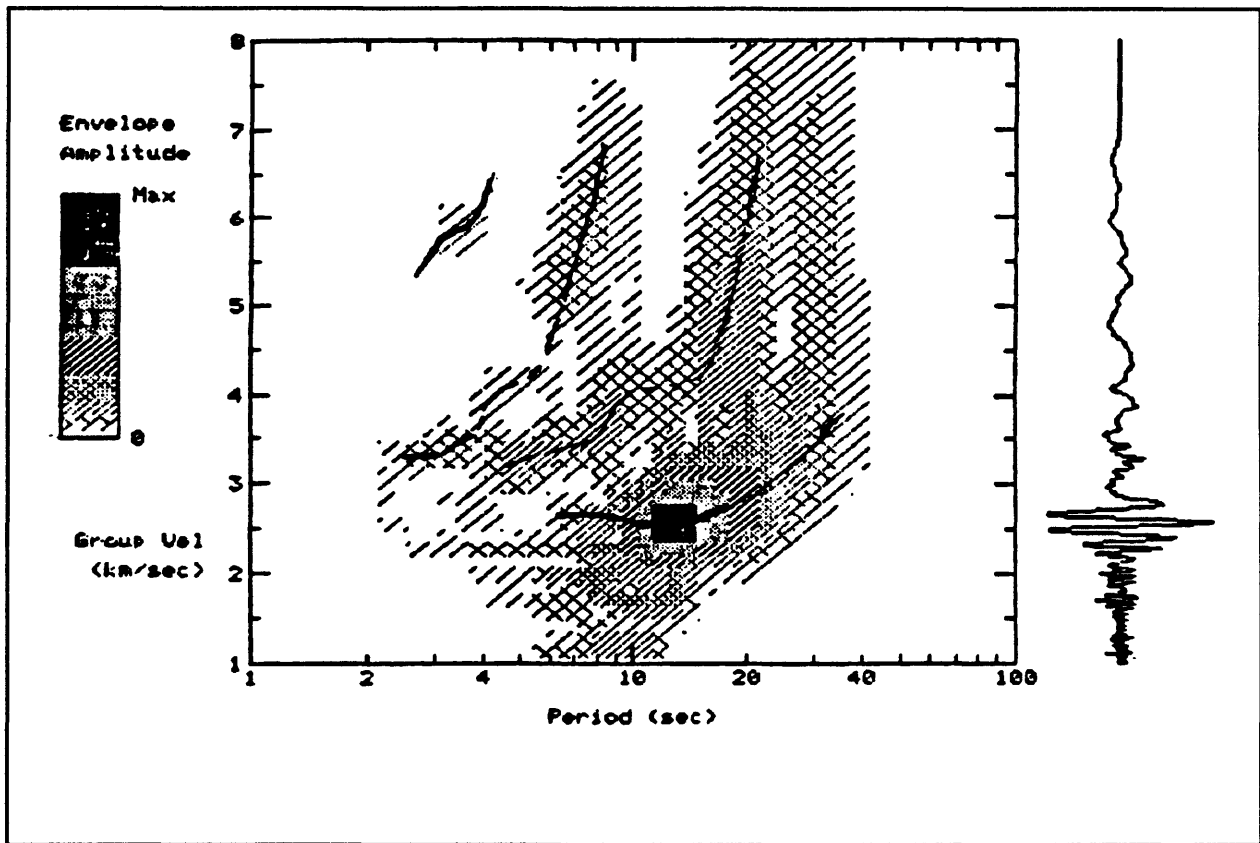


Figure 8 Time-frequency display.

mostly composed of 15 second period components. This plot is used to determine group velocity dispersion curves.

THE FOURIER TRANSFORM

One of the fundamental transformations used in seismogram processing is the Fourier transform. Figure 9 shows the mathematical definition of the Fourier transform for a continuous time function. We are actually dealing with discrete time sequences so the transform is given as a sum instead of an integral, but the discrete time transform is in other ways analogous to the definition given here. In the definition, $f(t)$ is the original time

series, and it is converted through the integral into a function of frequency $F(\omega)$. Basically, this transformation compares the time series with sinusoids of every frequency and phase over all time to see how much energy the times series has at that frequency and phase. So the result of the transformation is a function which gives the amplitude of the series as a function of frequency.

The inverse Fourier transform, which

is very symmetric to the forward transform, is used to recover the original time series from its frequency domain expression. It is standard practice to refer to the original time series as residing in the "time domain", and the transformed series as residing in the "frequency domain." One of the main reasons for the popularity of the Fourier transform is that such computationally intensive operations as filtering, convolution, and correlation can be done very efficiently in the frequency domain.

Every time domain function has a unique representation in the frequency domain and vice-versa. Thus, every time domain function and its frequency domain representation constitute what is called a "Fourier transform pair." Figure 10 shows a few of the common Fourier transform pairs. Here, the left side shows the time domain representation, and the right side shows the frequency domain representation. However, because of the symmetry of

Fourier Transform

The *Fourier Transform* $F(\omega)$ of the function $f(t)$ is defined as

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

Compares the function $f(t)$ with the sinusoid $\cos(\omega t) + i \sin(\omega t)$ over all time t for each frequency ω . The original function $f(t)$ can be expressed as

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega$$

Figure 9 The Fourier transform.

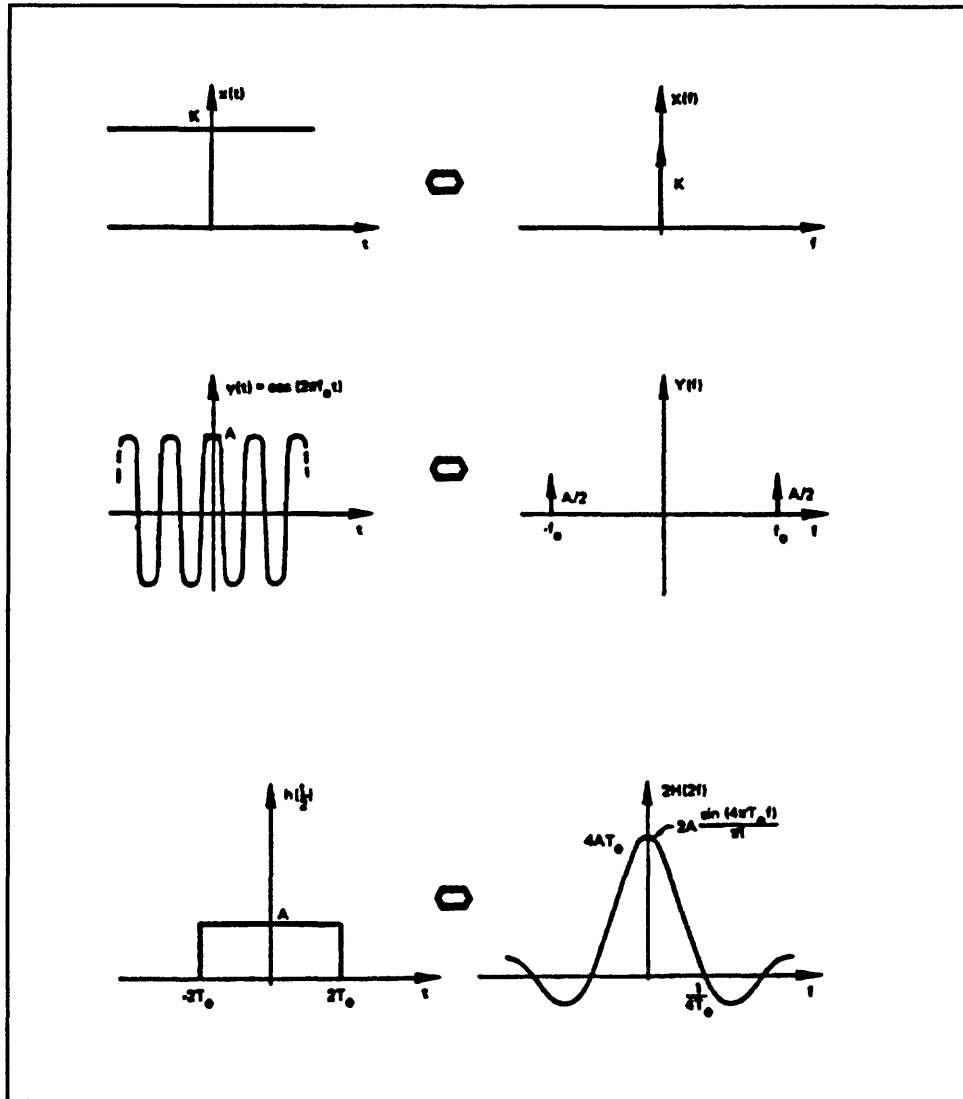


Figure 10 Some Fourier transform pairs (from Brigham, 1974).

the Fourier transform, by merely replacing all "t's" with "f's" and vice-versa, an additional set of three transform pairs are also given by this diagram. That is, not only does $\{x(t) = K\}$ transform into $\{X(f) = K\delta(f)\}$, but $\{x(t) = K\delta(t)\}$ transforms into $\{X(f) = K\}$. I mentioned earlier that the transform works by comparing the function to sinusoids of all frequencies and phases. The second transform pair shown in Figure 10 is of a pure sinusoid of frequency f_0

which transforms into spikes at frequencies of $\pm f_0$. This is the expected result.

An important transform pair, particularly for seismology, is the "boxcar to sinc" transform pair. This is shown as the third pair in Figure 10. One reason for the importance of this pair is that any time series you use in practice is of finite duration. This corresponds to multiplying in the time domain with a boxcar function. This, in turn, corresponds to convolution in the frequency domain with the sinc function.

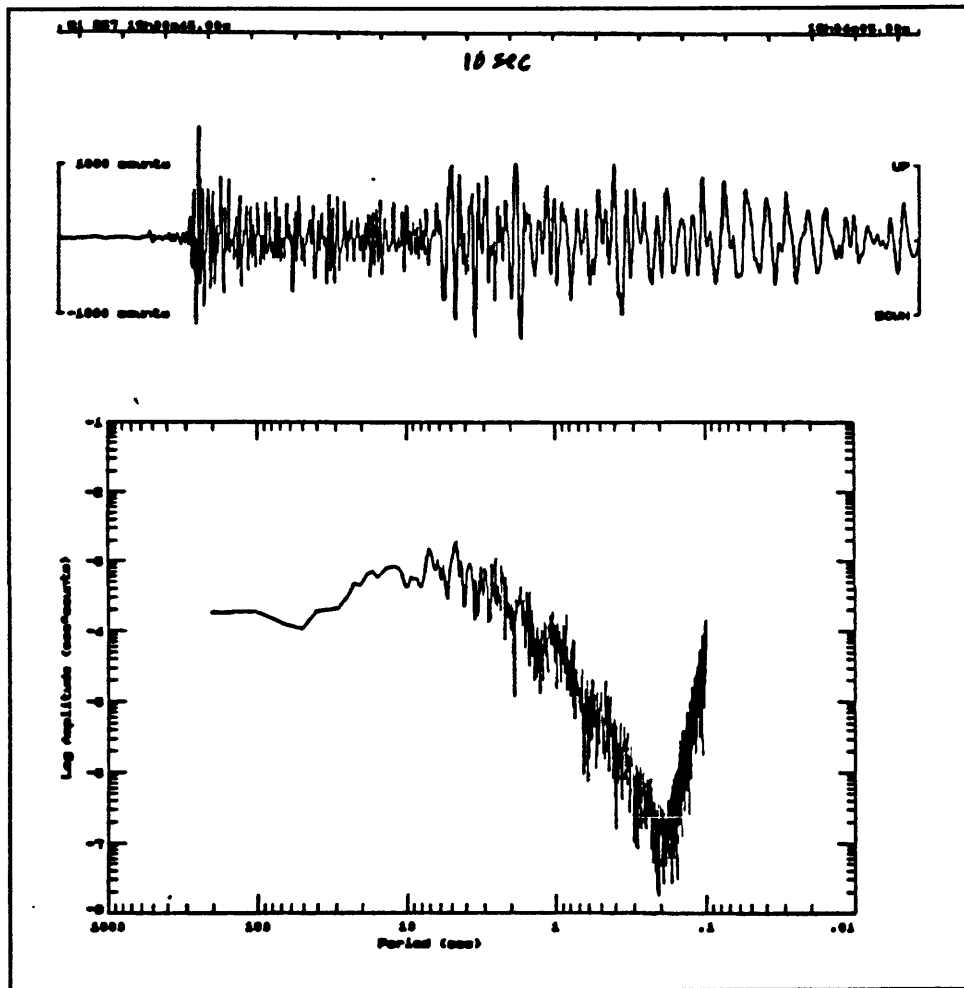


Figure 11 A seismogram and its transform.

Figure 11 shows a seismogram and the resulting spectrum produced by means of the

Fourier transform. The seismogram was produced by a nuclear event and was recorded at Berkeley. The amplitude spectrum is displayed with logarithmic axes for both amplitude and period (1/f). With just a glance at the amplitude spectrum, you can tell which frequencies had the highest representation in the seismogram. In this seismogram, the dominant frequencies seem to lie between about 2 seconds to about 10 seconds. I think that most of that energy is from the latter part of the wave which, as you can see, is quite coherent. Probably, if I had only transformed the first part of the seismogram, I would have found the dominant period to be around 1 second.

PROCESSING METHODS

I want to talk a little bit more about processing methods and, specifically, about filtering of a time series and about convolution and correlation. All of these techniques can be done in either the time domain or in the frequency domain. However, since the advent of a means of doing Fourier transforms very efficiently (the fast Fourier transform or FFT), it is usually much more efficient to do these operations in the frequency domain. Filtering can be thought of as a kind of restriction of a function or a limiting of

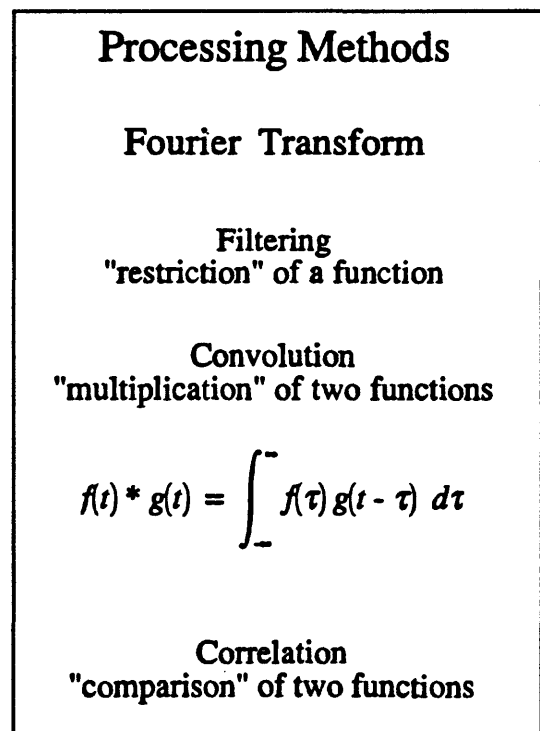


Figure 12 Processing methods.

information. Convolution can be thought of as a "multiplication" of two functions.

Correlation is related to convolution, but may be thought of as a comparison between the two functions. We might use correlation to compare a recorded seismogram to a synthetic seismogram produced from some model we are testing.

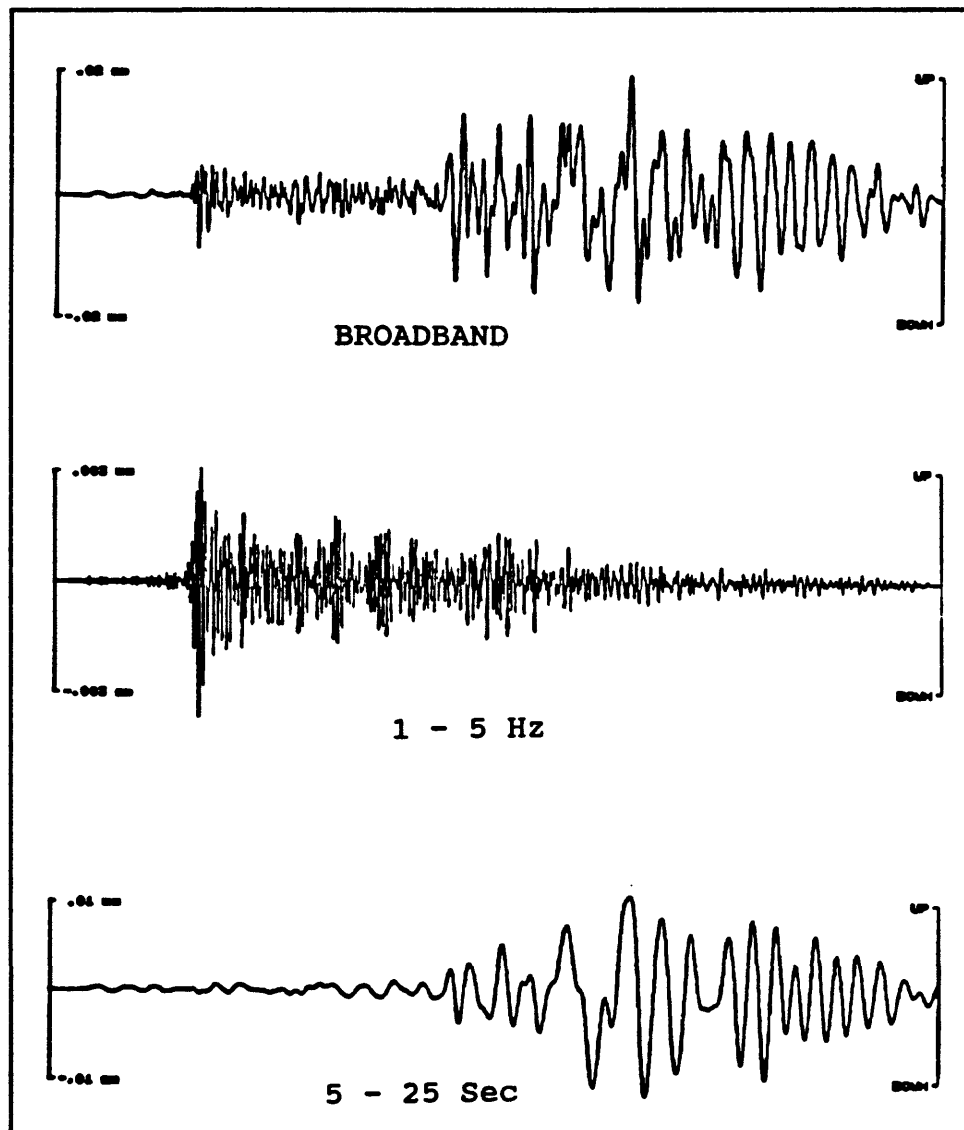


Figure 13 Filtering a seismic trace.

Figure 13 shows an example of filtering applied to a seismogram. The top trace is the

vertical-displacement component from a broadband seismometer recording of the nuclear test blast I showed earlier. The center trace shows the same seismogram filtered so that the energy is restricted to be between 1 to 5 Hertz. That is the higher frequency end of the recording. Most of the P-wave energy is between 1 to 5 Hertz, so by filtering in this manner, we emphasize the P-wave in this seismogram. The bottom trace shows the seismogram filtered at long periods. This causes the P-wave to be de-emphasized and the surface wave train to be emphasized. A generalization of this result is that you can emphasize surface waves by low-pass filtering seismograms and emphasize body waves by high-pass filtering.

CONVOLUTION

Figure 14 is an attempt to show schematically what happens when you convolve two functions. At the top of the figure are two functions $h(t)$ and $x(t)$ which are to be convolved. The first step in convolving these two functions is to fold one of these two functions around the vertical axis. In this case, $h(t)$ has been folded. Then, for every displacement (t) we get the area under the curve given by the multiplication of the folded and shifted version of $h(t)$ with $x(t)$. The value of that integral for that displacement t is the value of the convolution for that displacement (t) .

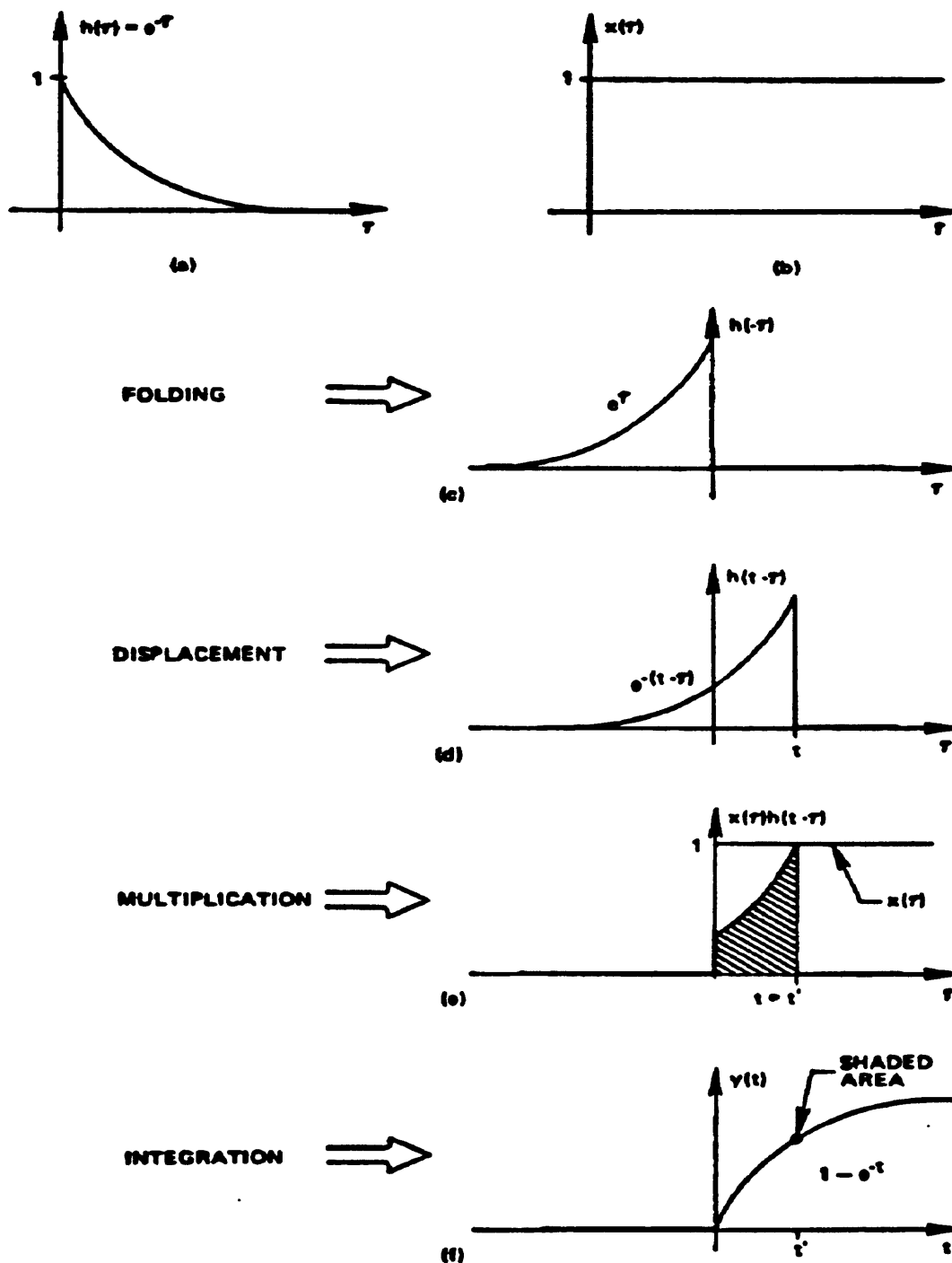


Figure 4-4. Convolution procedure: folding, displacement, multiplication, and integration.

Figure 14 Pictorial view of convolution (from Brigham, 1974).

Figure 15 shows an application of convolution. In seismology one often deals with

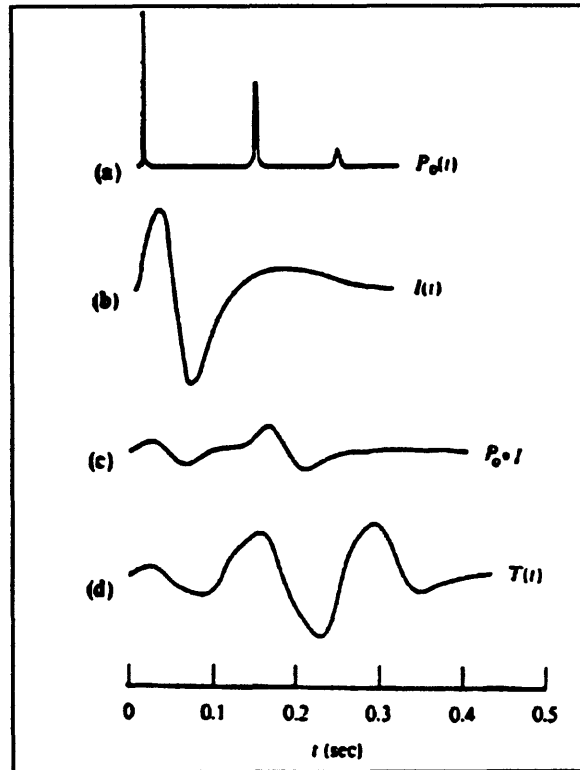


Figure 15 Convolution application.

either a synthetic, or otherwise known source function, for an event under study. In the case of Figure 15, we are dealing with a seismic event caused by exploding a charge of TNT under water. The source function for this event is given by part (a) of the figure. However, whenever, we measure a seismic signal, our measurements are affected both by the instruments we use to measure the signal and by the effects of the media through which the signal must pass. Part (b) of Figure 15 shows the response characteristics of a recording system to a spike input. This is often referred to as the "impulse response" of the system. Under assumptions of linearity and time-invariance, we can expect the response of the

recording system to the source function to be given by the convolution of the source function with the impulse response of the recording system. This is shown in part (c) of Figure 15. Part (d) of the figure shows the result of adding in some of the effects due to transmission through a medium.

As I have mentioned before, the various processing steps can be applied in either the frequency domain or in the time domain. But because of the generally greater efficiency of frequency domain processing, most processing is done in that domain.

Figure 16 shows the overall process in schematic form. First, the seismogram $f(t)$ is transformed into the frequency domain via the Fourier transform. This

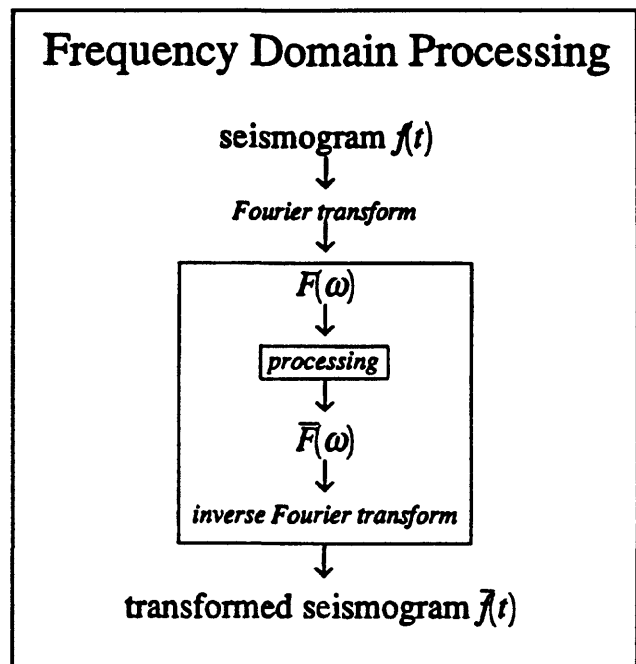


Figure 16 Frequency domain processing.

results in the series $F(\omega)$. A number of manipulations may then be performed on $F(\omega)$, and the resulting modified series can then be inverse transformed back to the time domain.

OTHER PROCESSING METHODS

I want to briefly mention a few other processing methods which are often used in seismology. Earlier I referred to polarization analysis in the context of plotting seismograms in different coordinate systems. Actually, there are some generalized methods for

determining the polarization of a group of waves. These can be used to determine particle motion and to identify various phases. This finds application in automated monitoring of seismic activity including nuclear test ban verification.

Seismic arrays consist of seismometers arranged in ordered fashion. Because of the

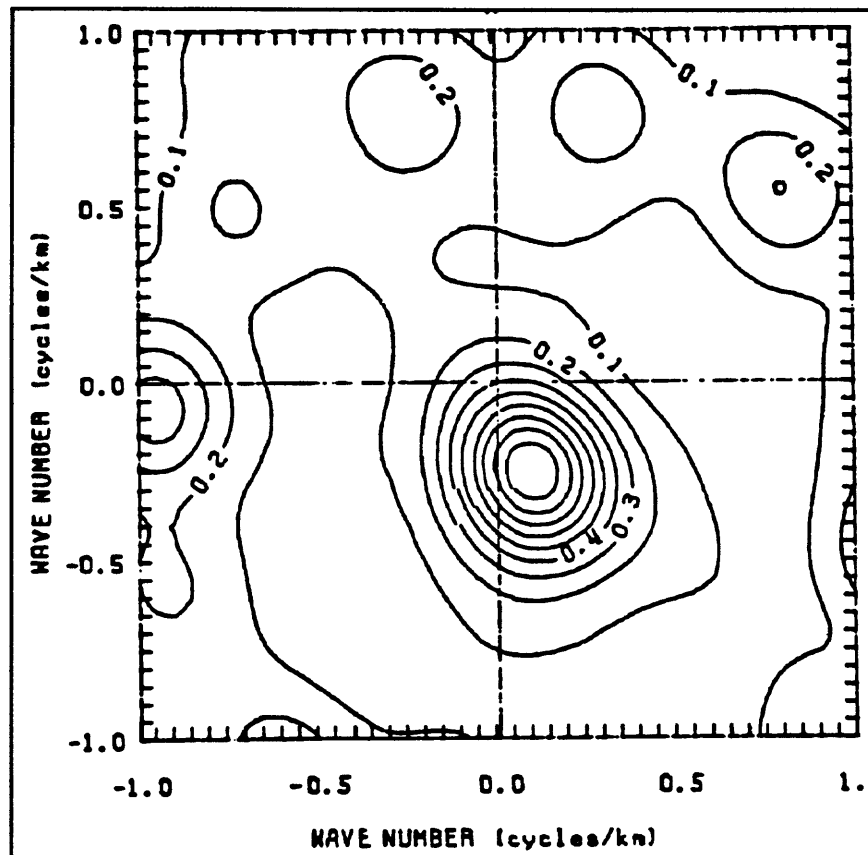


Figure 17 Array analysis.

spatial order, it is possible to do Fourier transforms both in time and space and thus determine characteristics of the wave field impinging on the array. This sort of processing is generally referred to as "array analysis." Figure 18 shows an example of array analysis performed on an array of 36 stations in Taiwan. What this gives you is the strength and direction that a significant portion of the arrivals had. You can also determine the apparent

velocity of a group of waves which impinge on the array. Array analysis is well suited to the study of scattered waves, those which did not come directly from the epicenter.

A very fundamental form of analysis in seismology is comparison of seismograms to

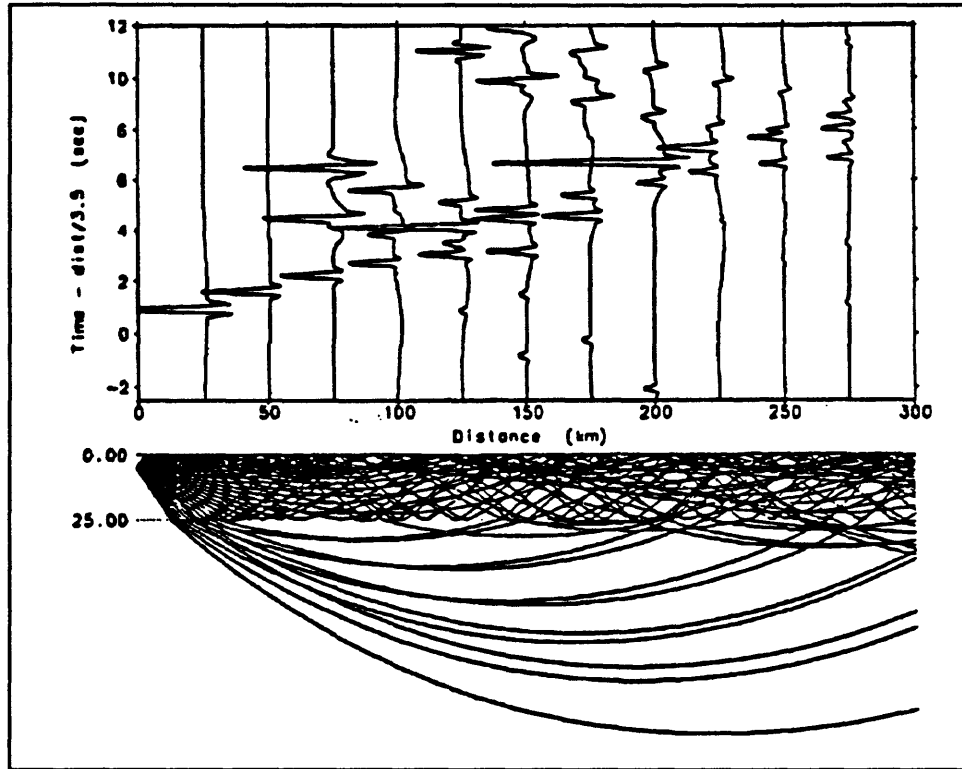


Figure 18 Computation of synthetic seismograms.

synthetic seismograms. By this type of analysis, it is possible to fine tune the earth model we are using. This sort of analysis involves quite a bit of mathematics and physics. Figure 15 shows a series of synthetic seismograms calculated for various distances from an idealized source and convolved with a simple earth model.

REFERENCES

Brigham, E.O. (1974). "The Fast Fourier Transform", Prentice-Hall, Inc., Englewood Cliffs, NJ.

17. THE SEISGRAM PROGRAM

by

A. Lomax
Seismographic Station
University of California, Berkeley, CA 94720

The SEISGRAM program which is distributed by IASPEI is an interactive graphics program for the display and basic analysis of 3-component digital seismograms (Lomax, 1991). In this respect, it is quite different from programs which are oriented toward batch processing and handling of large quantities of data. SEISGRAM can handle 1-component seismograms, but it was developed around 3-component, broadband seismograms, so it has been optimized for that purpose. I am working on another version, which is more geared toward larger numbers of single-component data, so it will be significantly more useful for network data.

The SeisGram Program

SeisGram is interactive graphics software for the display and basic analysis of 3 component, digital seismograms.

Program requirements:

An IBM compatible PC running PC DOS version 3.1 or later.

At least 640 k of RAM memory.

A hard disk and a graphics adapter (EGA, VGA or better).

An 80386 CPU and matching co-processor or 80486 CPU are recommended.

Figure 1 The SEISGRAM program.

EQUIPMENT REQUIREMENTS

The program runs on almost any IBM compatible PC running the later versions of DOS (Version 3.1 or later). The current version of the program takes a lot of memory. The 640 Kbytes of memory which DOS provides to programs is quite inadequate, when you are

attempting to do signal processing on long streams of data. This version of the program takes nearly all the available memory for its data arrays. Consequently, the user needs to be a little bit aware of how memory is being managed on his system in order to be able to run SEISGRAM successfully. I plan for later versions of SEISGRAM to either allocate memory on an as-required basis, or else to swap unused memory to disk to free up space when it is needed. A hard disk is a necessity and, of course, you need some sort of graphics adapter. Although the program will run on any of the older CPU's, it is advantageous to have a machine with either an 80386 or 80486 CPU. It is also very important to have a math co-processor since the program is so numerically intensive. Some testing I did indicates that the program performs floating point operations about ten to thirteen times faster with a math co-processor than without.

PROGRAM OPERATION

When you start SEISGRAM, the first thing you see is a screen like that in Figure 2. At this point, four basic options are available to you as shown under the "Select" menu. The bottom line is a status line which gives you information on the data groups available and read, and on the memory available. Basically, at this point in the program you either read a data file into memory using the "Newgroup" option, display some data already loaded, initialize the program, or else exit the program. This is the non-graphical part of the program.

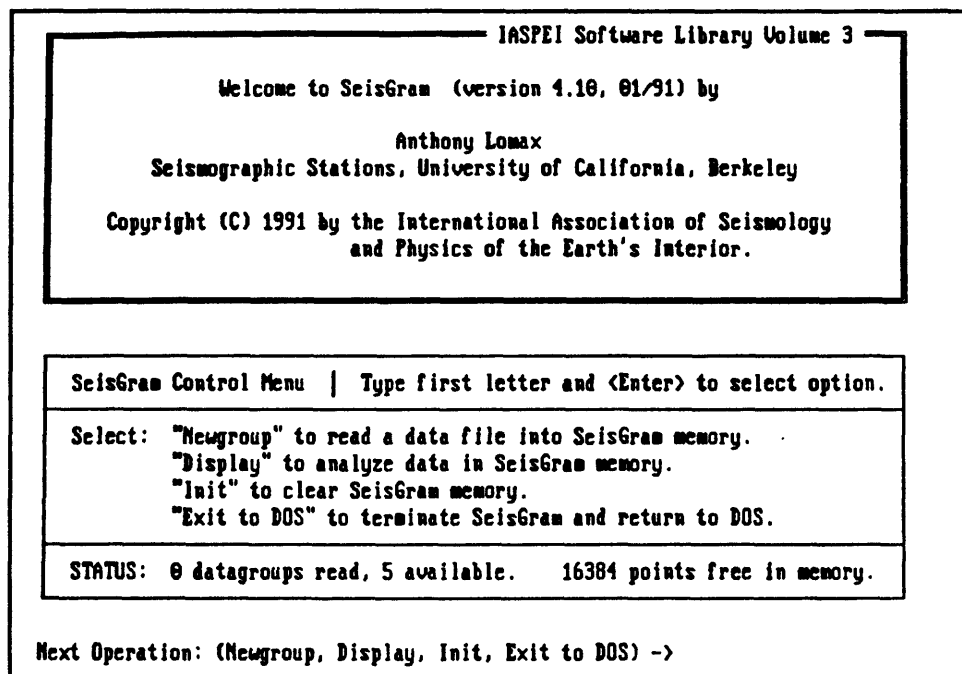


Figure 2 Opening screen control menu.

SEISGRAM works with distinct "display groups" consisting of from one to three traces from a single station. With the current version of SEISGRAM you may have up to five display groups in memory at a time. When more than one display group is loaded, the screen is divided so that the

display groups are clearly separated from each other. At any time, one of the display groups is active. Most of the operations which SEISGRAM performs will apply only to the active display group. Although, in general, up to five display groups can be loaded at any time,

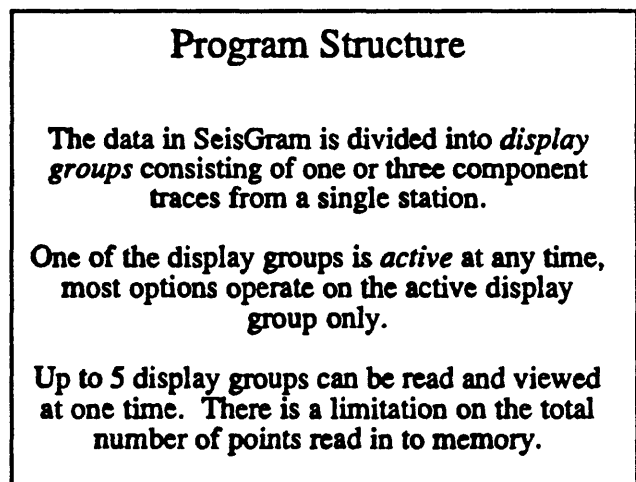


Figure 3 Program structure.

lack of available memory can limit this number to something less than 5.

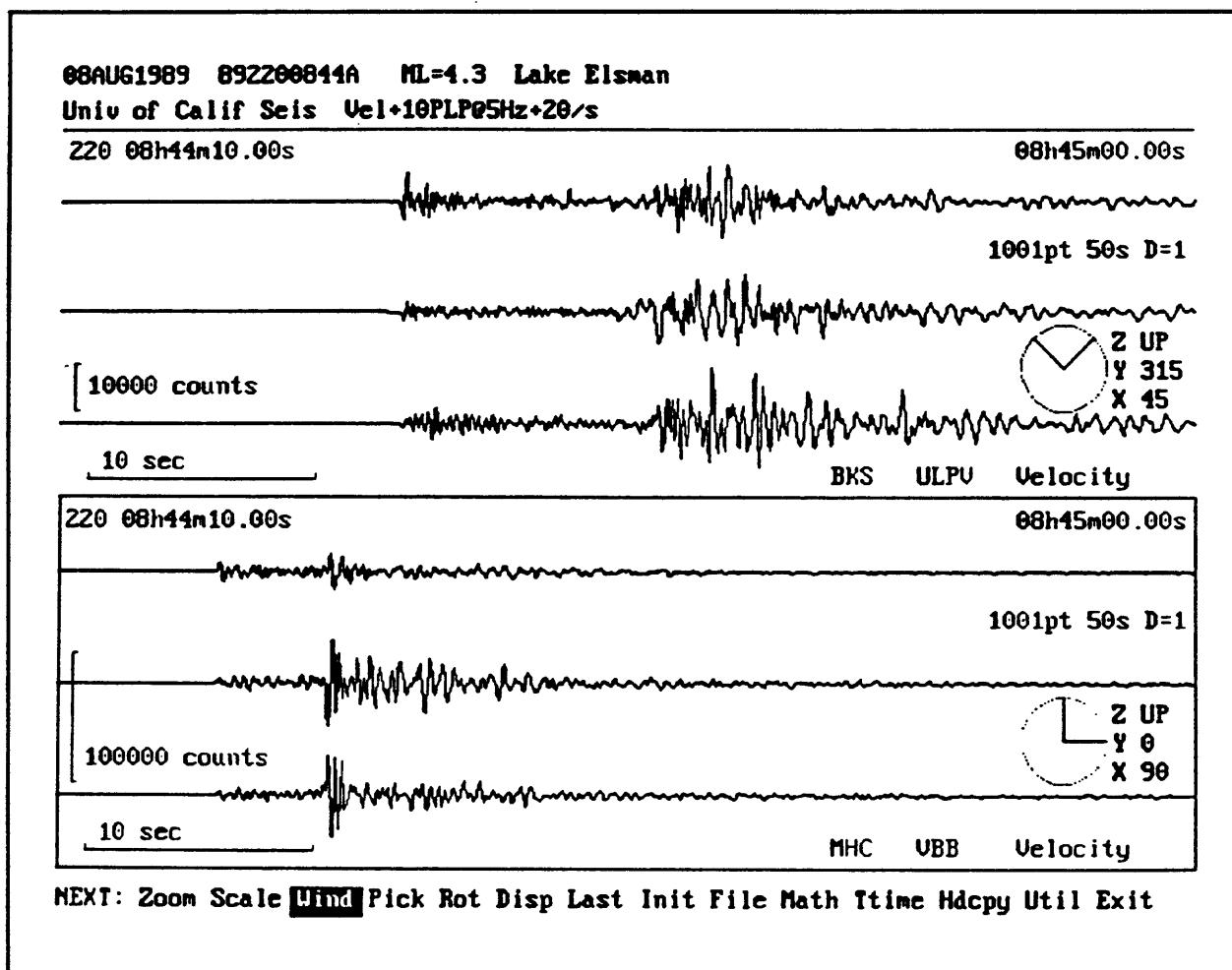


Figure 4 Basic display with two groups shown.

After you have read in the display groups, the screen will look something like Figure 4. In this case two display groups have been loaded. The upper three traces are from station BKS and the lower three traces are from station MHC. The box around the lower set of traces shows that it is the active display group. The header at the top of the screen refers to the active display group, and consists of data information, comments, network information, and velocity information. Within each display group, the information in the upper left-hand

corner is the Julian date of the event and the start time of the window. The upper right corner of the display group shows the stop time for the window. The number of points within the window is also listed along with the corresponding time length in seconds and any decimation applied to the time series. The amplification in digital counts and a time scale are displayed near the lower left corner of the display window. Finally, the station identifier, instrument type, and type of record are displayed in the lower right hand corner of the display window. The traces are always displayed so that the top trace is vertical and the bottom two traces are horizontal components. The compass symbol between the bottom two traces shows the orientation of the two horizontal traces. The color of the "hands" on the compass corresponds to the color of the trace to which they refer. Along the bottom of the screen is one line giving a list of options. The highlight on the option line can be moved using either the mouse or the arrow keys on the numeric keypad. By pressing the "enter" key, the currently highlighted option is executed. Alternatively, the option may be activated by pressing a letter key corresponding to the option name.

Figure 5 shows the same screen, but with only one active display group, hence only one set of traces displayed. On a VGA display, I have found that up to about 3 groups can be displayed simultaneously without significant loss of resolution.

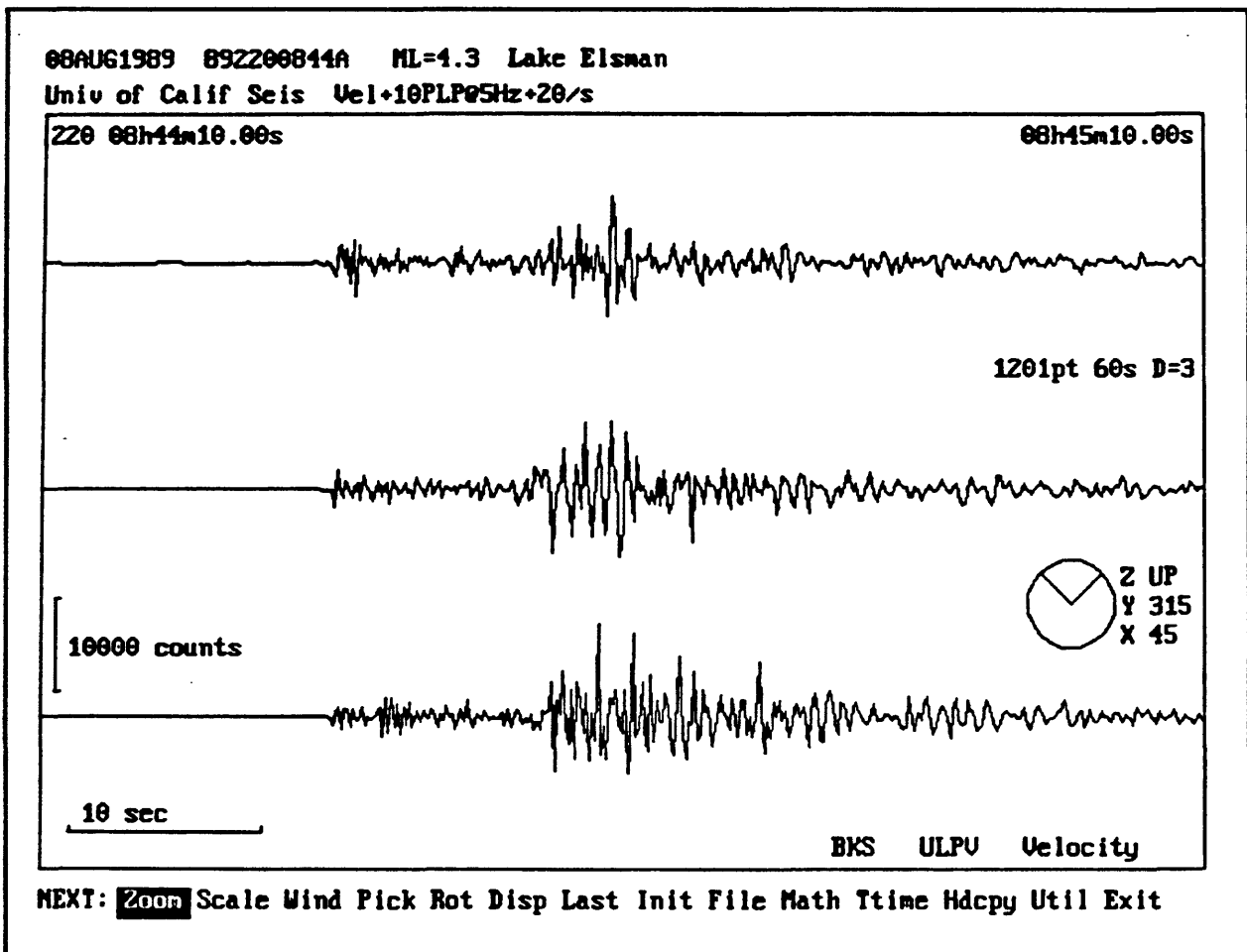


Figure 5 Single group being processed.

PROGRAM OPTIONS

The basic viewing options available on the main screen are "ZOOM", "SCALE", "WINDOW", "ROTATE", and "DISPLAY." ZOOM, SCALE, and WINDOW all involve displaying a segment of the seismogram which is currently active. ROTATE is used to transform the horizontal components to a specified orientation in space; perhaps radial and transverse to the path of the seismic wave. DISPLAY is used to superimpose the three

traces. For certain operations, it is nice to be able to do this.

Basic analysis options are "PICK", and "TRAVEL TIMES." PICK is for general time-based amplitude picks and for differential amplitude picks. TRAVEL TIMES allows you to overlay a set of predicted arrival times over a seismogram to assist in identifying phase arrivals. At the moment it only works for teleseismic phases.

The "MATH" option allows you to perform several types of processing. Using this option you can perform such basic operations as mean removal and time domain integration. Also, a large selection of frequency domain options are available under the MATH option.

Finally, there are some general options available on the main menu. These are "FILE", "HARDCOPY", and "UTILITIES." FILE is used for writing output files. HARDCOPY is used to send output to a laser printer or graphics printer. UTILITIES includes a number of services for such things as removing traces, editing text and numbers, and managing memory.

Figure 7 shows the effect of using ZOOM on a set of traces. ZOOM is a very quick way of narrowing the display to just a small segment of the seismogram. After selecting the

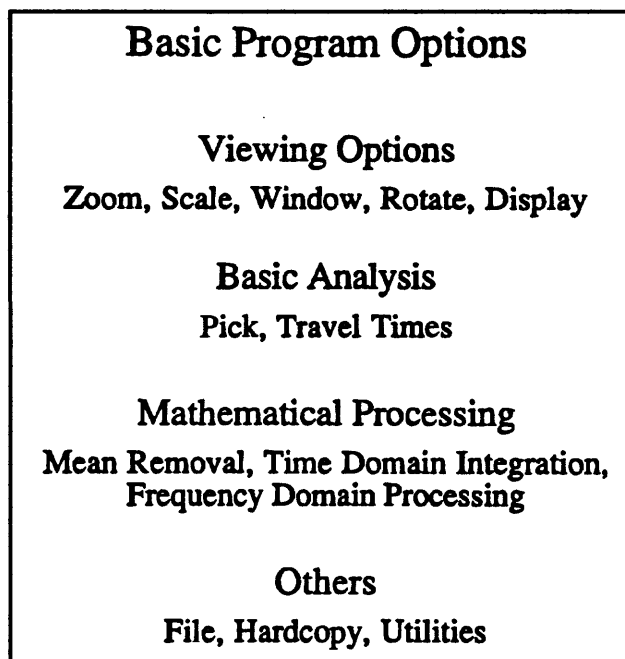
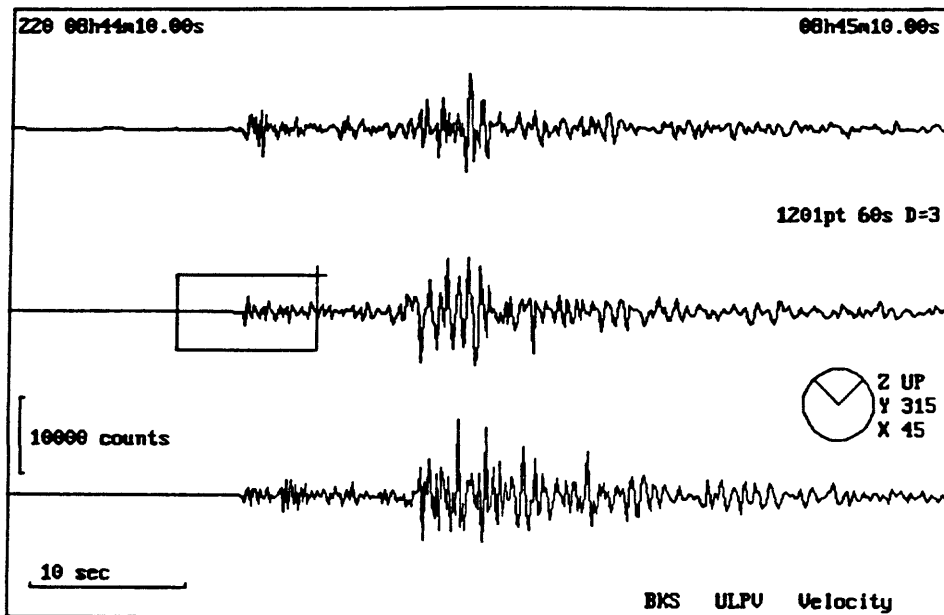


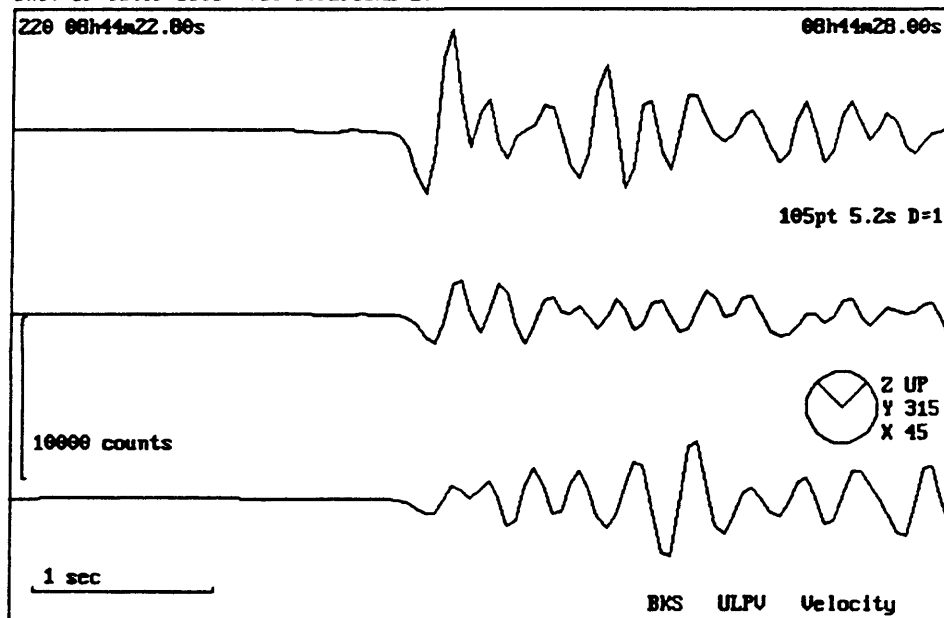
Figure 6 Basic program options.

Use middle trace for zoom window.



Move locator, press <Enter> at two diagonal corners of zoom window, Q to quit.

00AUG1989 052200044A ML=4.3 Lake Elswan
Univ of Calif Seis Vel=10PLP05Hz+20/s



NEXT: Zoom Scale Wind Pick Rot **Disp** Last Init File Math Ttime Hdcpy Util Exit

Figure 7 Using the zoom feature.

option, you use the mouse to select two corners of a box which defines the new display window. Then either by clicking one of the mouse buttons or pressing the Enter key, the lower window fills with the selected portion of the seismogram. When you are displaying three components, you always select on the middle component to define the zoom window. All the traces in one display group share a common amplitude scaling so no matter how small a section of the seismogram you are viewing, it is always possible to see the relative amplitudes of the different components.

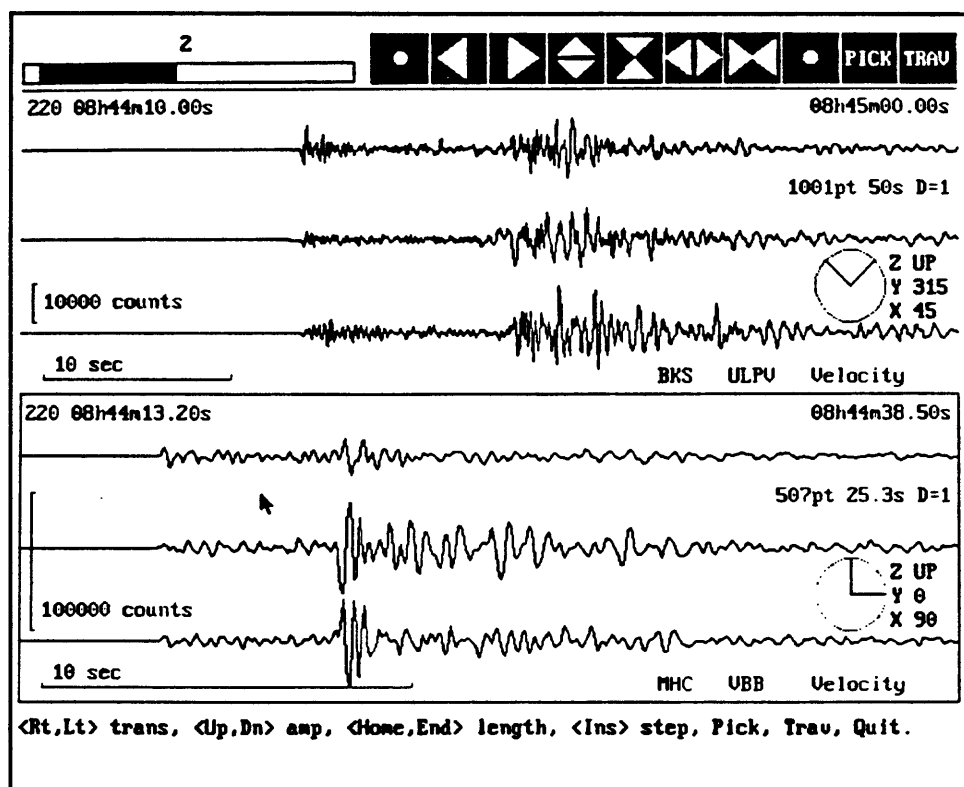


Figure 8 Using the window option.

Another way to view a segment of a seismogram is to use the WINDOW option (See Figure 8). This method is not quite as quick as ZOOM but it is more flexible. In the WINDOW mode, you select options either by clicking on the symbols at the top of the

screen, or by pressing the letter corresponding to the desired option and then pressing the Enter key. For instance, if you click on the symbol which looks like two arrows pointing together, the window is squeezed in. So, the seismogram is expanded to show more detail. In each case, the arrow symbolizes the action which will be performed on the window, not action taken on the seismogram. The dot symbols control the step size for all window adjustments. PICK and TRAV are ways to get to the PICK menu without going back to the main menu. The scale bar in the upper left symbolizes the position in the seismogram occupied by the present window. The shaded part is what is being displayed, and the unshaded part is not in view currently. You can switch back and forth between display groups either by moving the mouse pointer to the desired display group and clicking, or else by using the PgUp and PgDn keys on the numeric keypad.

Figure 9 shows a case in which a very short window has been defined around the onset of the P-wave. This trace would have been derived from the full seismogram by repeatedly clicking on the ►◄ symbol and using either ► or ◄ to move to the appropriate part of the seismogram. The advantage of using WINDOW for this sort of operation is that you can proceed in small steps. This allows you to see the change in the seismogram evolve slowly so that you are better able to judge when you have selected the best window to work with. Also, by using either the left or right arrows, you can use WINDOW to slowly scan along the seismogram.

The final option for re-scaling is called SCALE. SCALE is a purely numerical method in which you select a start time and a length. SCALE is the least intuitive and least interactive means of scaling the windows but it does have the advantage that you can specify

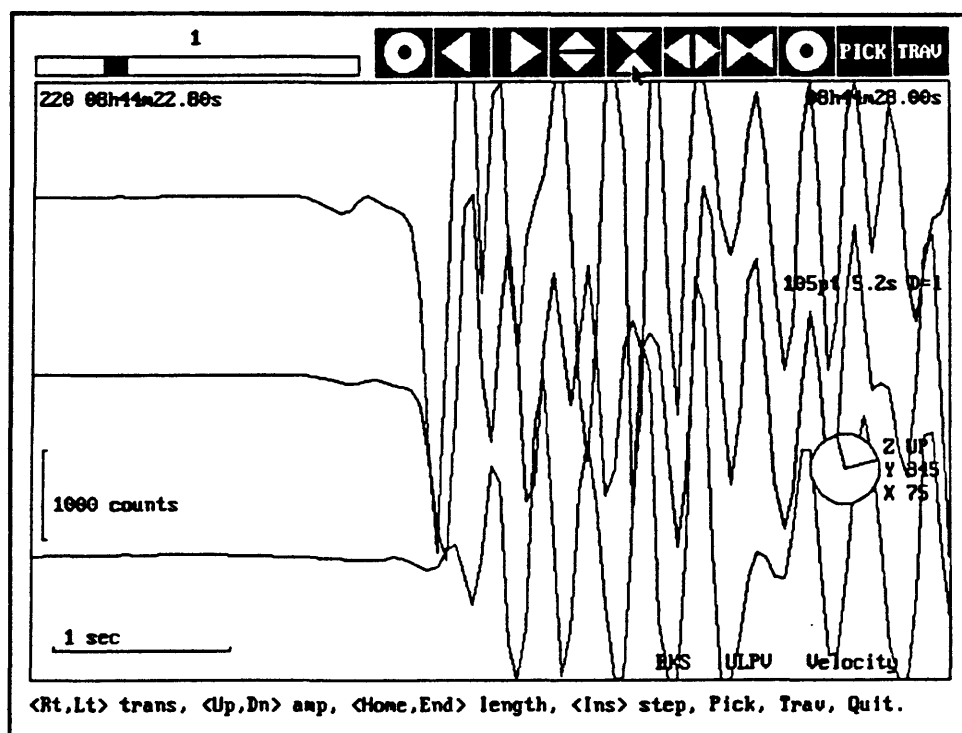


Figure 9 Short segment of seismogram selected using the window option.

the window parameters very precisely.

SEISGRAM also has a rotation option. In Figure 10 three traces have been overlaid using the display option. It is useful to be in this mode when you are doing rotations since it makes it easier to see correlations between the horizontal components as you rotate them.

One interesting application of ROTATE is to rotate the traces until the P-wave doesn't show on one of the horizontal components. When the seismograms have been rotated into this position, then you have the seismometer rotated into the natural coordinates for that earthquake. That is, one horizontal component is aligned in the radial direction along the source-receiver path, and the other component is aligned in the transverse direction. Combining this with S-P times, you can get a crude estimate of the epicenter location with

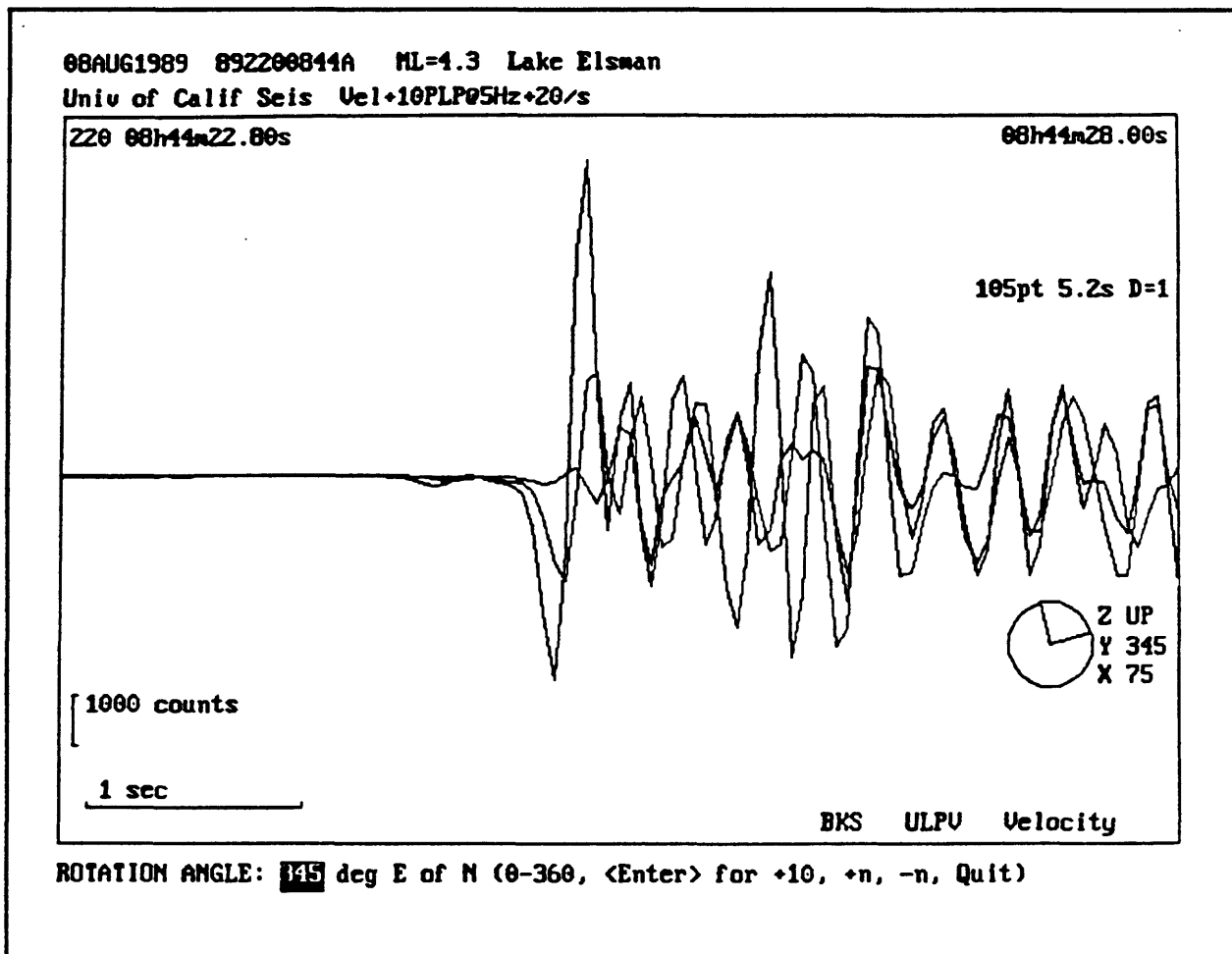


Figure 10 Use of the rotate option.

just one station. Perhaps more important, by rotating the seismograms into the natural coordinates for the earthquake, you will be in a better position to analyze the rest of the seismogram scientifically.

Within PICK, there is a display of the traces and you can move the cursor along a trace. As you do, the display in the upper left corner continually updates with the time and amplitude at the cursor. You can select which trace the cursor is following. There is a whole sub-menu under PICK. Using the PHASE option you can pick phases and label them, or you can remove a phase pick that you made with DELETE. With A-T you can make an

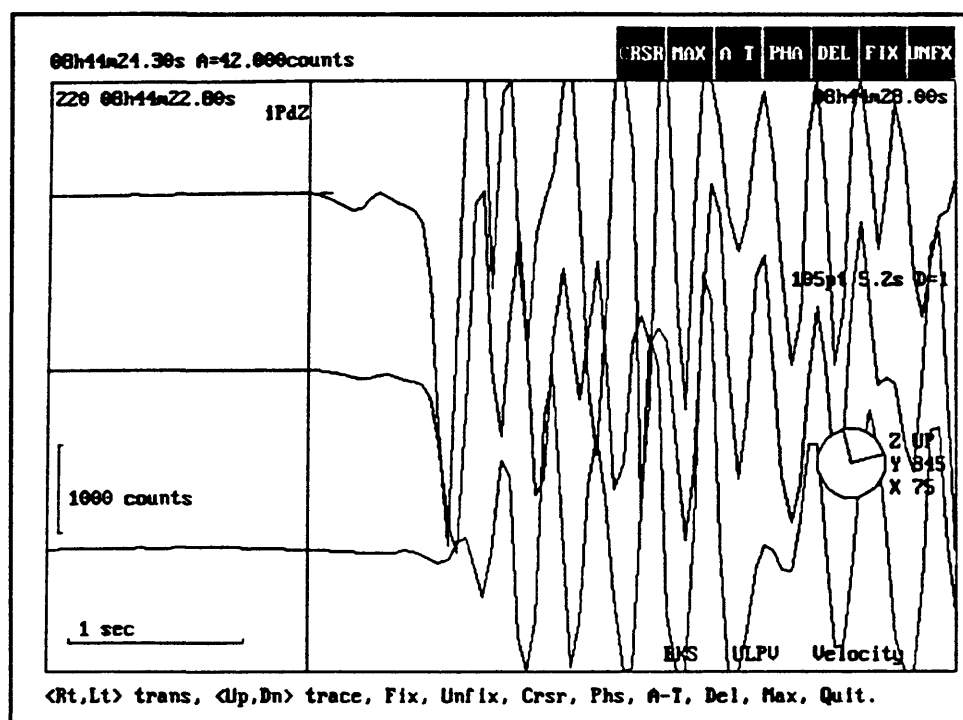


Figure 11 Using the pick option.

amplitude-period pick for teleseismic body waves. This is useful for observatory work. If you click on MAX, the cursor will jump to the maximum value. In the FIX mode, a marker is set at some point along the trace and as you move the cursor, the differential time is calculated and displayed on the screen.

Figure 12 shows an example of differential picking. FIX was selected when the cursor was at the little X on the bottom trace. The FIX time and amplitude are displayed on the top line in the upper left corner of the figure. After fixing the point, the cursor was moved along the trace to the onset point of the S-wave. The time and amplitude of that point are displayed on the next line down at the upper left. These are followed by the differential time and differential amplitude between the two points. In this case, the differential time is the S-P time at that station. Being able to easily determine differential times and amplitudes is a

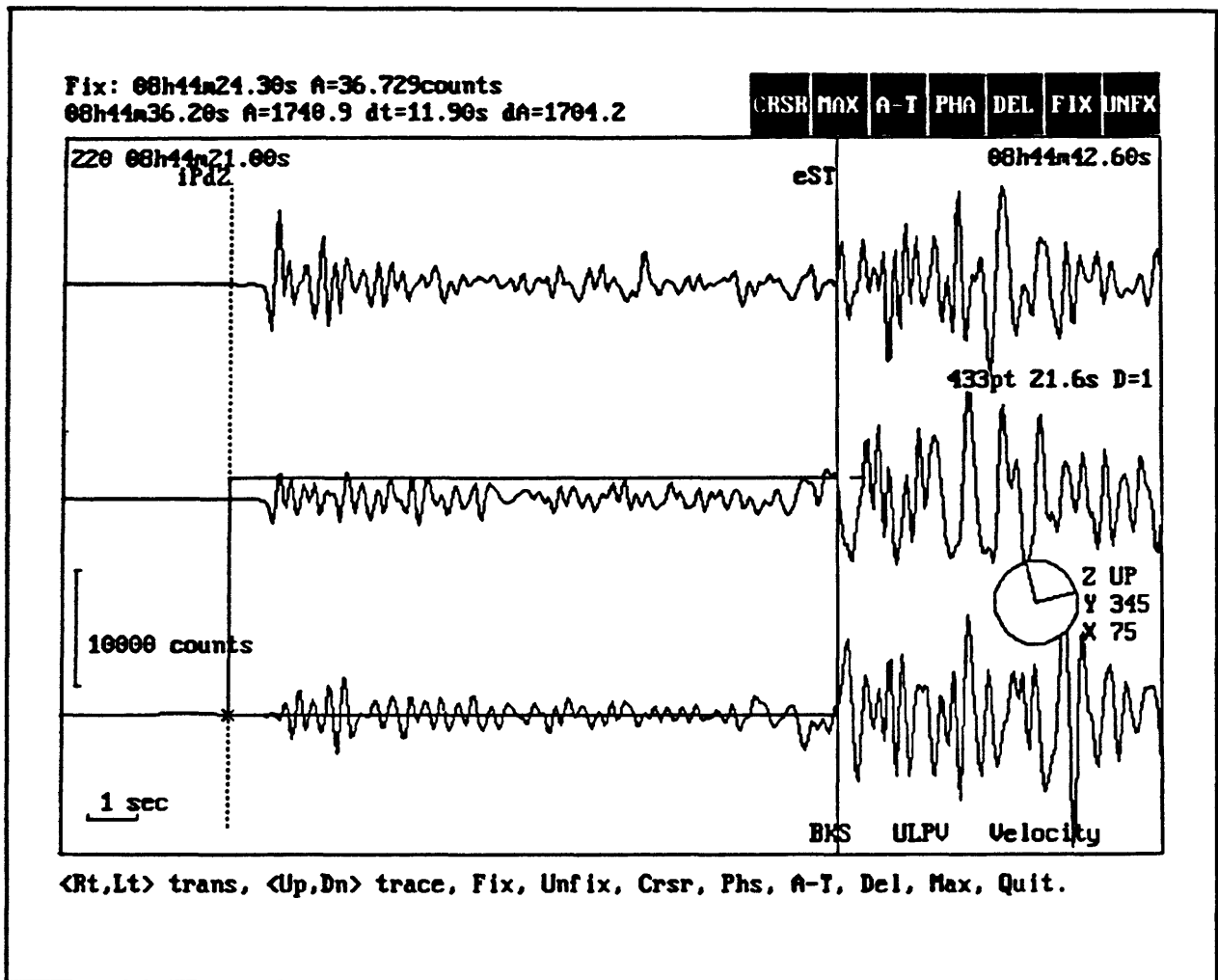
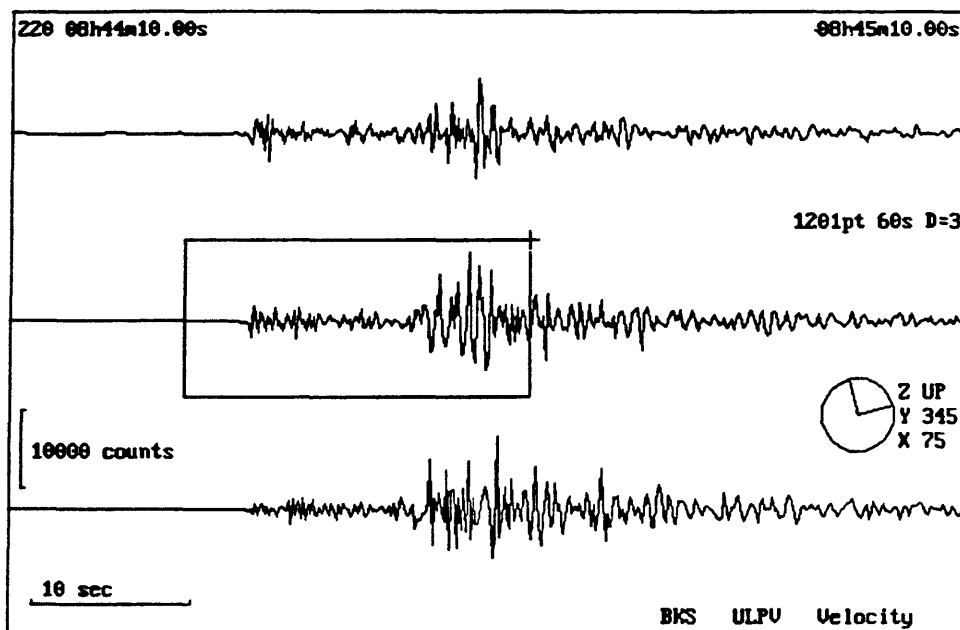


Figure 12 Using the differential pick option.

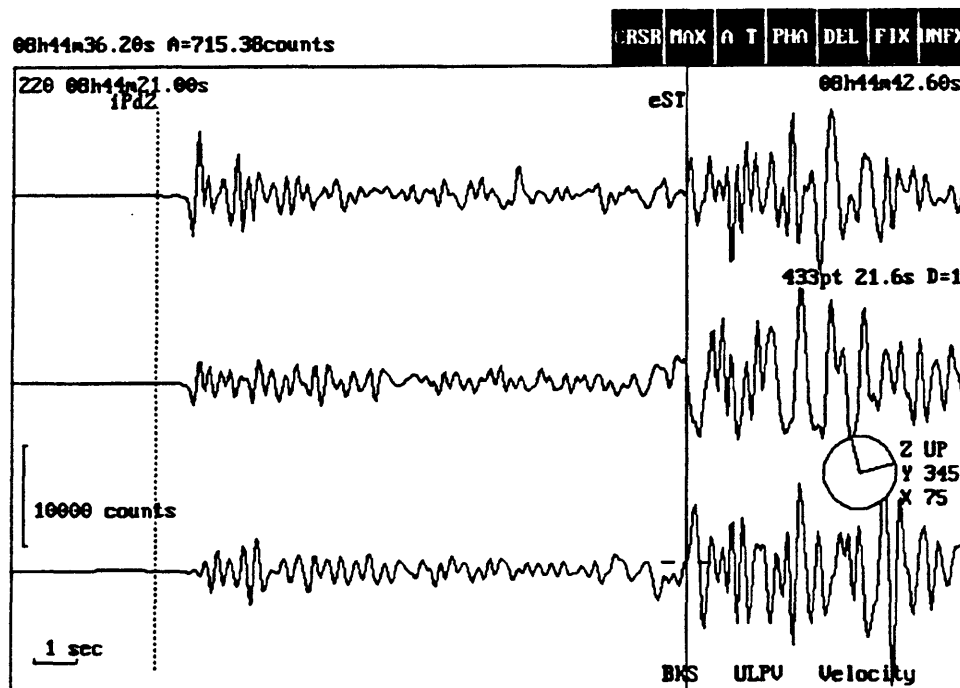
powerful tool in routine analysis of seismograms.

Figure 13 illustrates the sequence you might use to pick both the P and S phase onsets on a 3-component seismogram. In the upper window, ZOOM is being used to select the portion of the seismogram which contains both phase onsets. After selecting the PICK option and picking both times, the screen appears as in the lower half of Figure 13. At the Berkeley Seismographic Station we base our locations, typically on P times alone. However, since we started using broadband seismometers, we have been able to pick S times on the horizontal

Use middle trace for zoom window.



Move locator, press <Enter> at two diagonal corners of zoom window, Q to quit.



<Rt,Lt> trans, <Up,Dn> trace, Fix, Unfix, Crsr, Phs, A-T, Del, Max, Quit.

Figure 13 Picking P and S onsets with the pick option.

components quite well using this program, and consequently have improved our locations significantly. While the P time can be picked fairly well by automatic methods, the S time is quite a bit more difficult, since the S phase often arrives while there is still signal from prior phases. This means that human intervention is still required to pick S-wave arrivals accurately.

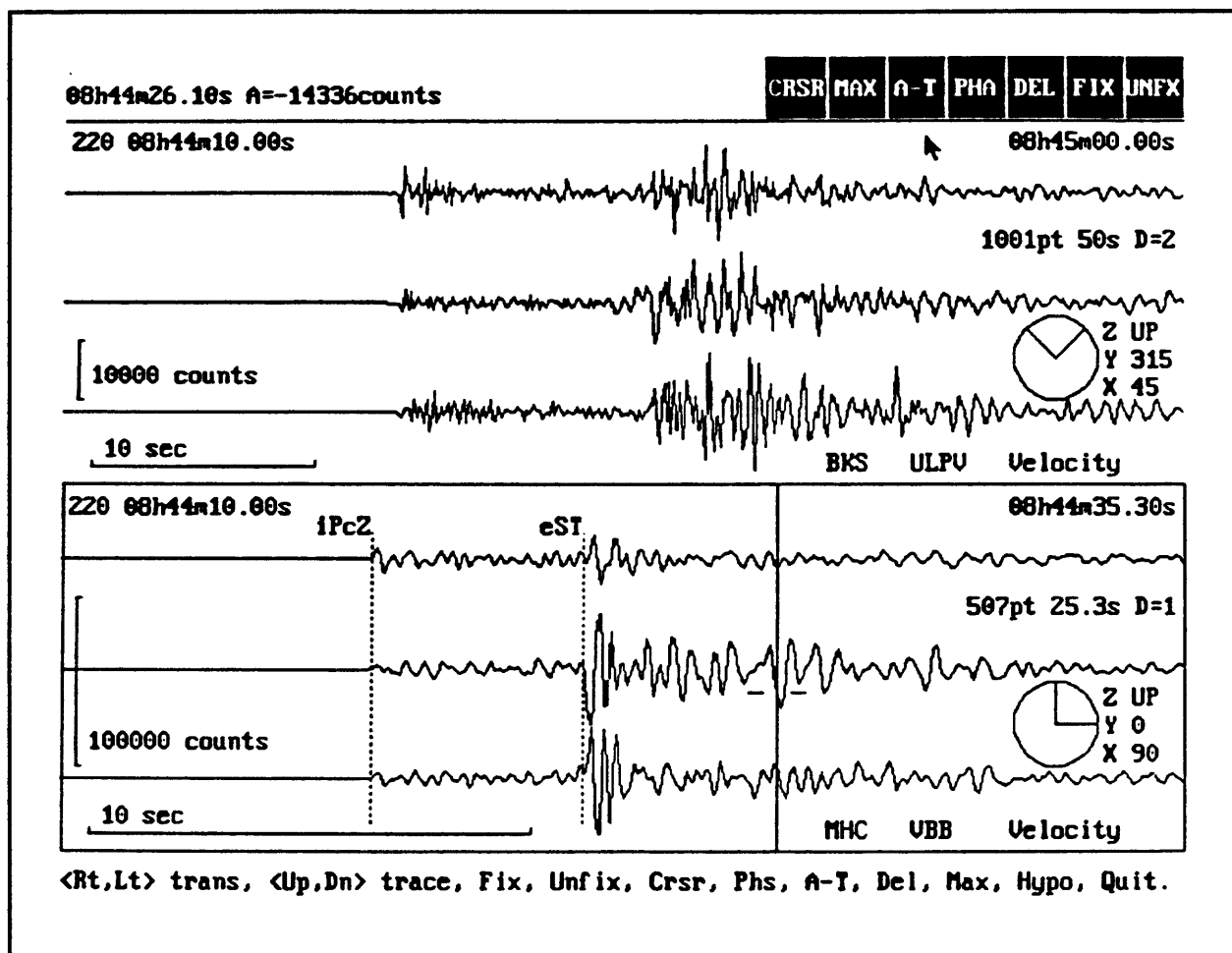


Figure 14 Using pick with multiple display groups.

Figure 14 shows the PICK process when there are two display groups loaded. The lower group is active so the pick cursor is in the lower half of the screen. The phases already have been picked. If the phases had been picked in the upper group while it was active, then

the picks would still show up even when that group was not active as in this figure. In fact, you can change the display group while operating under the PICK menu. When you do so, the PICK cursor always moves to the currently active group.

If you have the same station, instrument, and time window displayed in more than one display group and you make a pick in the active group, the program will also display the pick in the corresponding inactive group window. That allows you to select a small section of a long record, pick one or more arrivals within that section, and have the picks displayed on the complete record. This can be very helpful when analyzing teleseismic records.

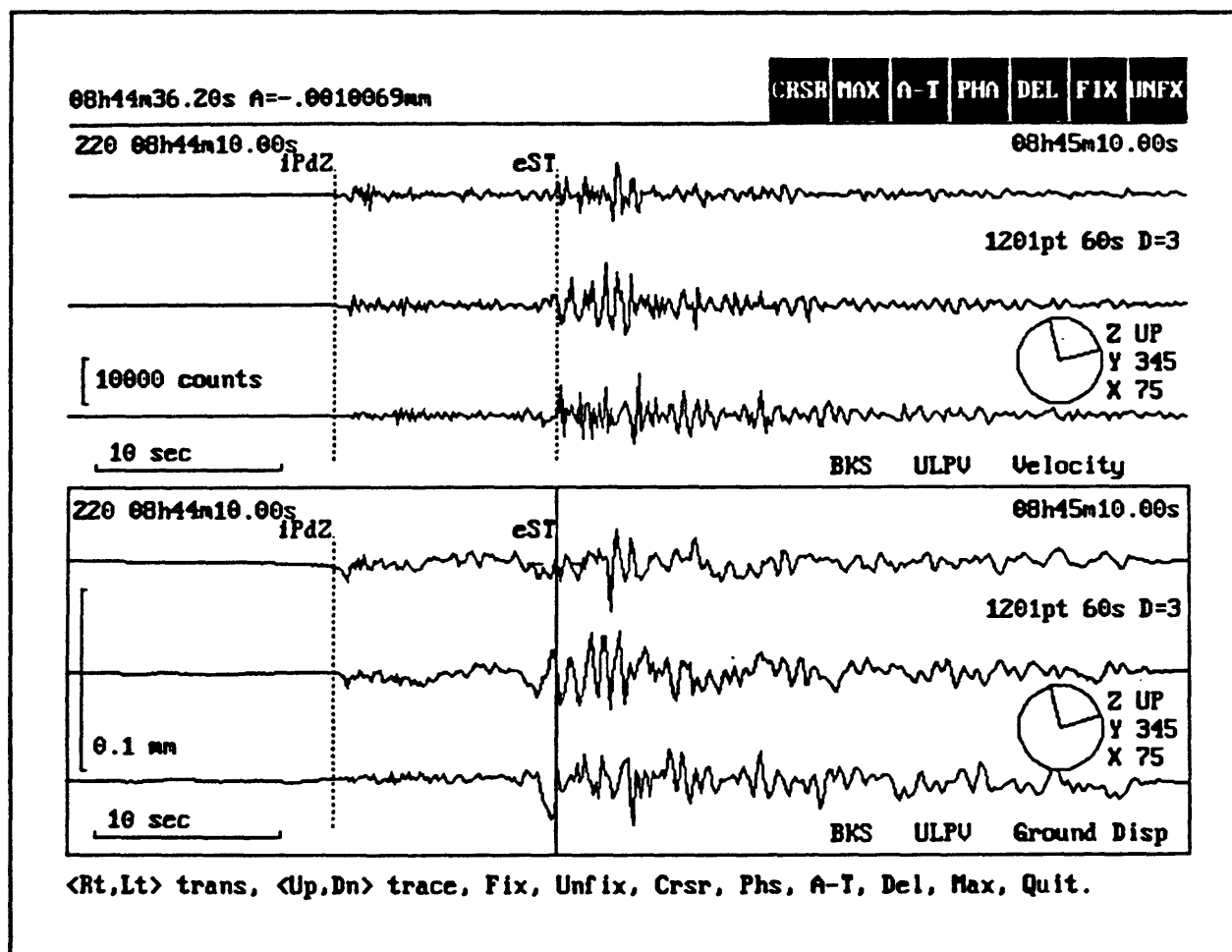


Figure 15 Instrument deconvolution.

Figure 15 shows a case where the same record is being displayed in two different windows but in one of the windows, the record has been processed to remove the instrument response. This type of processing is referred to as **deconvolution**. Earlier I mentioned that the seismogram was the result of convolving the source function with the instrument response and the impulse response of the earth. In the deconvolution process, we attempt to remove the effects of the instrument, so that we can see the signal as it appeared without the effects of the instrument modifying it. To perform this deconvolution operation in SEISGRAM you must have a file set up and available to the program, so that the program knows the characteristics of the instrument. Then you select instrument deconvolution under the MATH menu. SEISGRAM will take the original instrument output, which, in this case, is velocity, and remove the instrument effects. It also integrates the output so that the display in the lower window is of ground displacement. As you can see, there can be some significant differences in the way records look when the instrument effects have been removed. Of course, much of the difference is due to the integration. As I mentioned earlier, integration tends to emphasize the longer period features of the seismogram, and the lower record certainly does show much higher amplitude long period features than the upper record. However, we know that our instrument response falls off for low frequencies, so part of the lack of long period features on the upper record is probably due to poor instrument response at those low frequencies.

Another effect of removing the instrument response has been to sharpen the P-wave arrivals. On the uncorrected record, the P-wave arrivals are smoothed, but on the corrected record, the P-wave arrivals are sharp, one-sided pulses.

However, one of the most interesting effects of removing the instrument response is the change in the apparent arrival time of the S-phase. On these records, the S-wave arrival time was picked prior to removing the instrument response. But notice that on the deconvolved seismograms, a long period part of the S-wave arrives nearly a second earlier. Here, removing the instrument response results in a significantly different picture of the earthquake.

This also shows that if you look at an earthquake using broadband seismometers, you may see things differently than if you only use short period instruments. In fact, the larger the earthquake, the more important it is to recover the long period information, since a greater and greater portion of the energy is at low frequencies as the earthquakes get larger.

SEISGRAM does Fourier transforms. It has to in order to perform much of the math which is used to analyze seismograms. However, it also has the option to calculate and display an amplitude spectrum. Figure 16 is a display of a velocity amplitude spectrum produced by SEISGRAM. A lot of the same information that is displayed on the time series screen is also shown on the amplitude spectrum display. You can choose the horizontal scale to be either frequency or period. Also, there are various options for scaling of the axes. If you wish, SEISGRAM will produce a hard copy of the spectrum. The spectrum display is always from one display group.

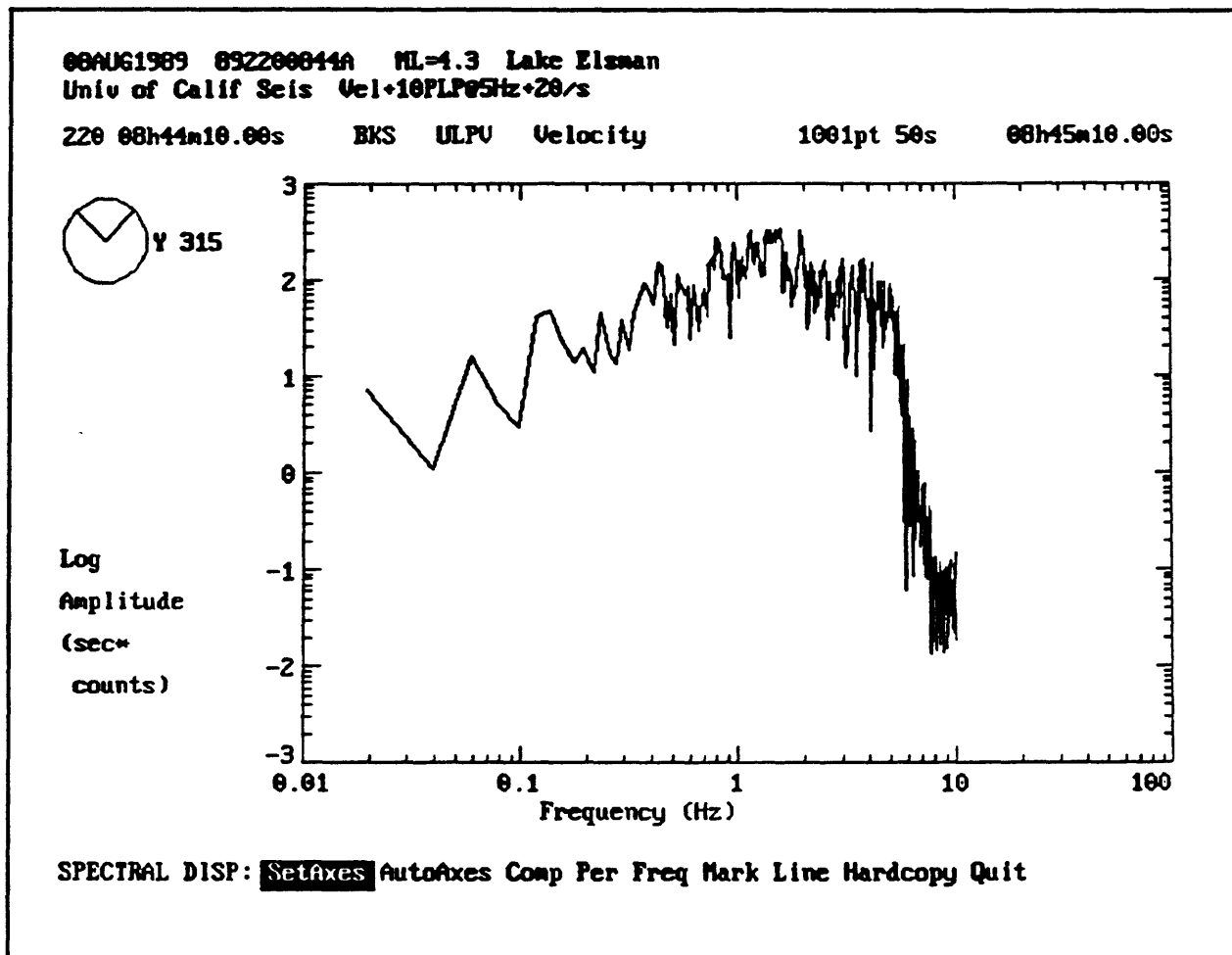


Figure 16 Velocity amplitude spectrum produced by seisgram.

SEISGRAM has an interactive option which allows you to superimpose Jeffrey's-Bullen model travel times on a seismogram. You can adjust the epicentral distance and the depth by clicking on the icons at the top of the screen. As you do, the various phase arrivals will move to the new locations predicted by the model. By adjusting the origin time, the various phase arrivals will move synchronously across the seismogram. At Berkeley, this has provided a very effective means of locating teleseisms. The Berkeley network covers northern California. Its dimensions are about 400 miles long by 300 miles wide. Relative to an earthquake source, 6000 miles away, the network is a point, so use of travel times to

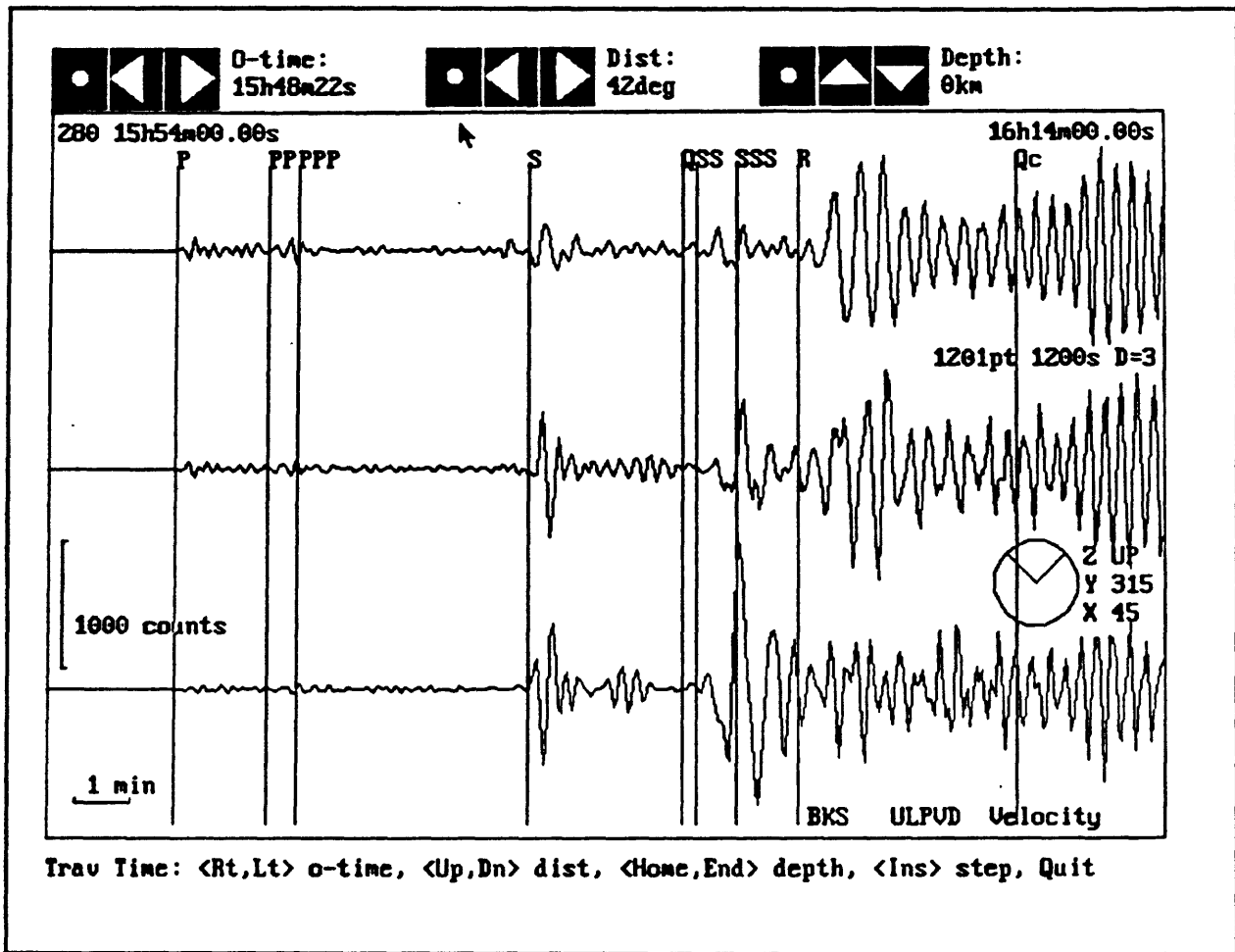


Figure 17 J-B travel times superimposed on seismograms.

locate teleseisms doesn't really work. Instead, we first rotate the seismometers to null out one component of the P-wave. This gives the azimuth to the source. Then we use the travel time option to get the best match of predicted phase arrivals to actual arrivals. Using this technique, we can locate an earthquake across the Pacific ocean with an accuracy of about 300 to 500 kilometers.

It is worth pointing out that a typical microearthquake location is performed using only the P-wave arrival and possibly the S-wave arrival. Here, we are, in effect, using many arrivals on each channel to locate an earthquake. This shows that there is much more

information contained in a seismogram than is currently being used on a routine basis. Hopefully, in the future, people will start incorporating more and more of the information contained in seismograms into routine analysis of small earthquakes. I think one very important application that may become much more routine is to estimate the moment and focal mechanism from 3-component broadband recordings of earthquakes. This will give us a better idea of the kinematics of the many earthquakes we are observing on a daily basis.

REFERENCES

Lomax, A. (1991). Users manual for SeisGram. IASPEI Software Library, 3, 13-80.

18. FOCAL MECHANISM: THEORY AND HISTORY

by

P. A. Reasenber
U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

The computer program described in this chapter was developed in 1983 - 1985 after the Coalinga earthquake sequence in California. We recorded over 2000 earthquakes for which we could calculate focal mechanisms. There was obviously a need for some automated way of doing that. In this chapter I will cover some of the theory behind calculating focal mechanisms. I also intend to at least touch on all of the subjects which I think are relevant to the procedure. However, this is not intended to be a rigorous development of focal mechanism theory.

As a framework for this discussion, Figure 1 shows the processes which occur within the earth and the means by which we recover information about those processes. Usually, for an earthquake to occur there must be some pre-existing faults and a stress regime which favors slip

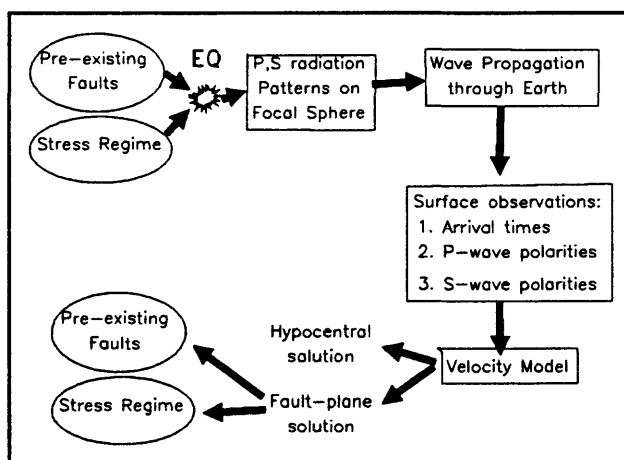


Figure 1 Schematic of the inversion process.

on those faults. When an earthquake occurs, P and S waves are radiated away from the source in some pattern, and the propagation of these waves is controlled by the velocity

structure of the earth. Eventually, the waves arrive at a seismic net where we can record them and observe such things as arrival time, P-wave polarity and S-wave polarity. Using a velocity model, we can calculate a hypocentral solution and a fault plane solution, and from the fault plane solution, we are often able to determine the orientation of the fault and the stress regime that produced the earthquake.

You can't talk about earthquakes without going back to Reid's (1910) theory of elastic rebound which is the most commonly accepted explanation of what happens during an earthquake. According to this theory, there is a region of the earth which is being elastically strained by some stress. At some point, the stress

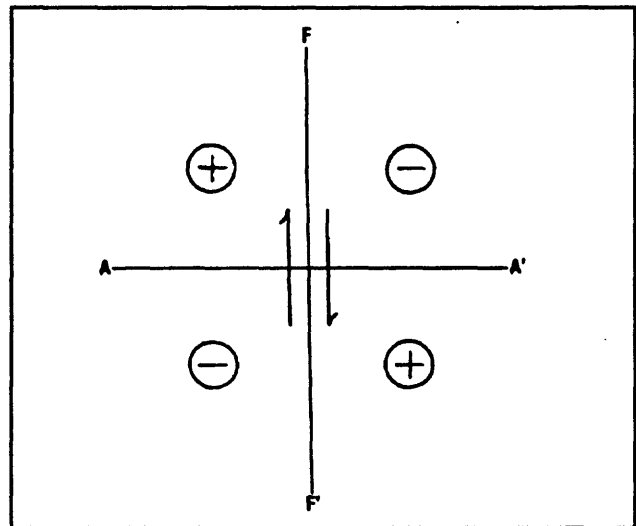


Figure 2 Compressions and dilatations along right-lateral fault.

exceeds the strength of the rock (or frictional strength of the fault in the case of an earthquake on a preexisting fault) and the elastic strain which had been developing is suddenly released through slippage on some portion of a plane. That motion causes elastic waves to be radiated through the earth surrounding the region of slip. Figure 2 shows an idealized model of a vertical fault oriented along the line F-F' on which right-lateral slip has occurred. It is fairly intuitive to see that in quadrants marked with a plus symbol the motion will result in compression while in quadrants marked with a minus sign, the motion results in dilatation. The waves generated by this earthquake will carry some information about the

sense of slip since they were produced subject to this systematic orientation of compressions and dilatations. That is, waves that depart from a compressional quadrant will have a compressional first motion, while waves that depart from a dilatational quadrant will have a dilatational first motion. This is the basic information that we use to reconstruct the orientation and sense of motion on the fault plane for any earthquake we analyze.

STRESS REGIME AND RADIATION PATTERNS

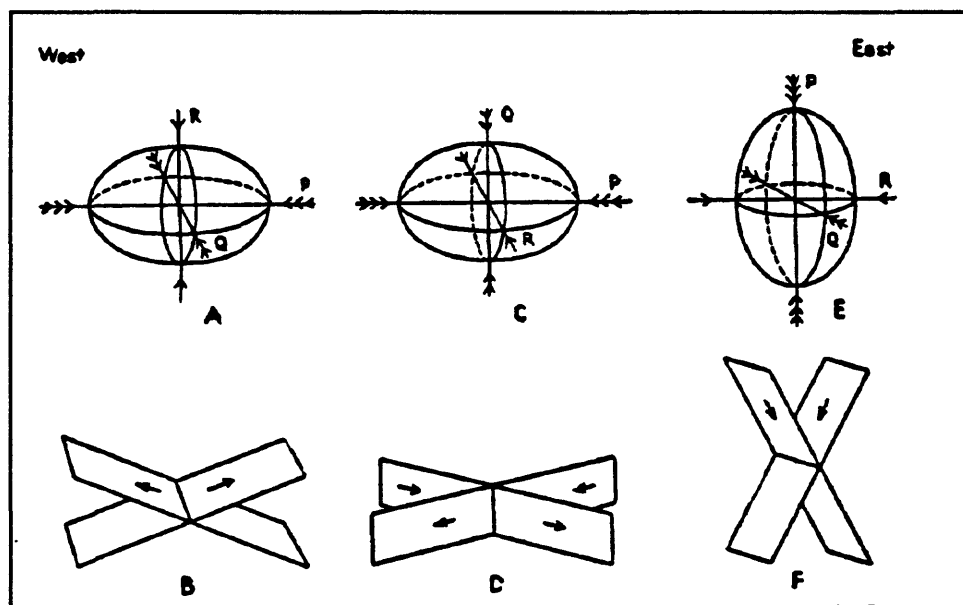


Figure 3 Stress ellipsoid and fault type.

Let's look in more detail at what happens as the stresses interact with some fault. At any point in the earth there is a stress field which is usually described by a stress tensor. It is possible to resolve the stress tensor into 3 principal components, orthogonal to each other, which can be considered as axes of a **stress ellipsoid**. These axes are commonly referred to as the greatest principal stress axis (P in Figure 3), the intermediate principal stress axis (Q

in Figure 3), and the least principal stress axis (R in Figure 3). When the least principal stress is vertical, as in case A of Figure 3, then reverse faulting as shown in B will occur. If the intermediate principal stress is vertical (case C) then strike-slip faulting occurs (case D). Finally, if the greatest principal stress is vertical (case E) then normal faulting can occur (case F). So, we see that there is a definite relation between the orientation and slip occurring on faults during an earthquake, and the stress regime within which the earthquake occurred. When we determine the fault plane orientation and slip direction for an earthquake, we can, in effect, go backward, and estimate the stress regime for that region of the crust.

Figure 4 shows a summary of what the P-wave radiation pattern looks like. In this figure, the fault plane is considered to be the $\Theta=90^\circ$ plane. The length of a vector from the origin to some point on the radiation pattern represents the relative amplitude of the P-wave at that orientation relative to the fault plane.

The radiation pattern shown in the upper part of Figure 4 is actually just a two dimensional slice through a three dimensional pattern. If we could surround the earthquake source with a sphere and mark points on the sphere according to whether compressional or dilatational pulses had propagated through the point, we would end up with a quadrantal "beach ball" pattern on the sphere. Since three dimensional objects don't lend themselves well to manipulation on paper, we work with certain projections or maps of the sphere. The lower half of Figure 4 is such a projection. Notice that the amplitude of the pulses decreases towards the **nodal planes** (the planes separating compressions from dilatations). In the figure this is shown by decreased size of the polarity symbols (pluses and minuses). This directly corresponds to the reduced vector length near the nodal planes shown in part A of Figure 4.

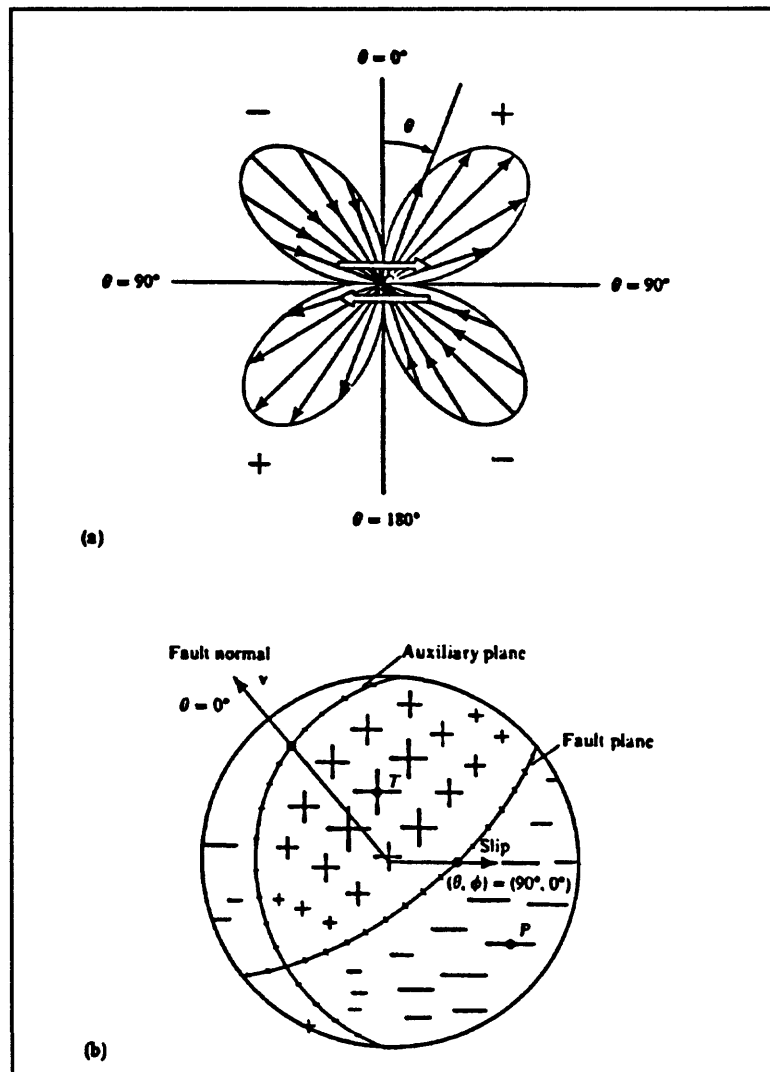


Figure 4 P-wave radiation pattern.

S-waves are radiated with a different pattern, as shown in the top half of Figure 5.

For S-waves, the maximum amplitude occurs on the fault plane and perpendicular to the fault plane. On the focal sphere projection, the maximum amplitude of the displacement occurs on the fault planes at the slip vectors and decreases from there.

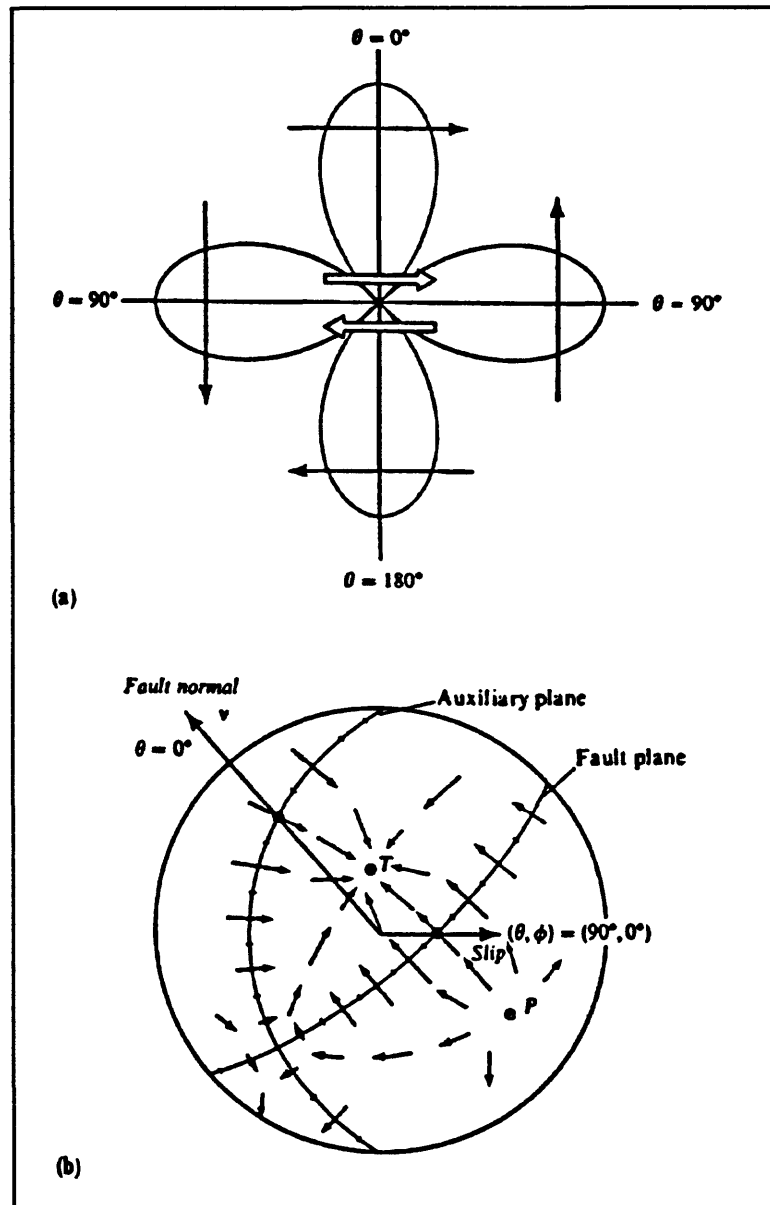


Figure 5 S-wave radiation pattern.

We need some kind of coordinate system in which to describe all the displacements produced by the earthquake. Figure 6 which is taken from Aki and Richards (1980) shows the focal sphere and orientation angles defined relative to geographic north. It shows one ray leaving the source region. The orientation of the ray is given by its azimuthal angle ϕ and

the take-off angle i_t . These are the angles that define the ray with respect to the earth coordinates.

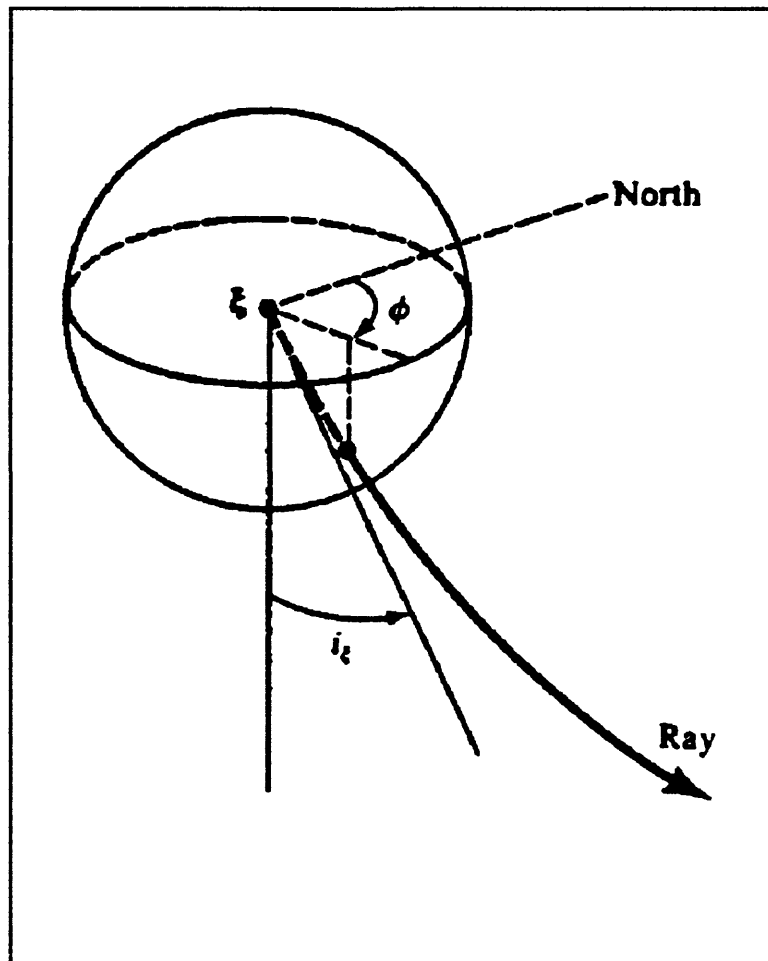


Figure 6 Focal sphere with coordinate system (from Aki and Richards, 1980)

Early in the history of seismology, it was noticed that there was a systematic orientation to the polarization of P-waves for a given earthquake. In 1934, Honda compiled the set of observations shown in Figure 7 for an earthquake in Japan. In this diagram, each seismograph location is indicated with a circle and the vertical polarization of P-wave first arrivals is indicated with a "u" for up and a "d" for down. Since these rays are refracted to

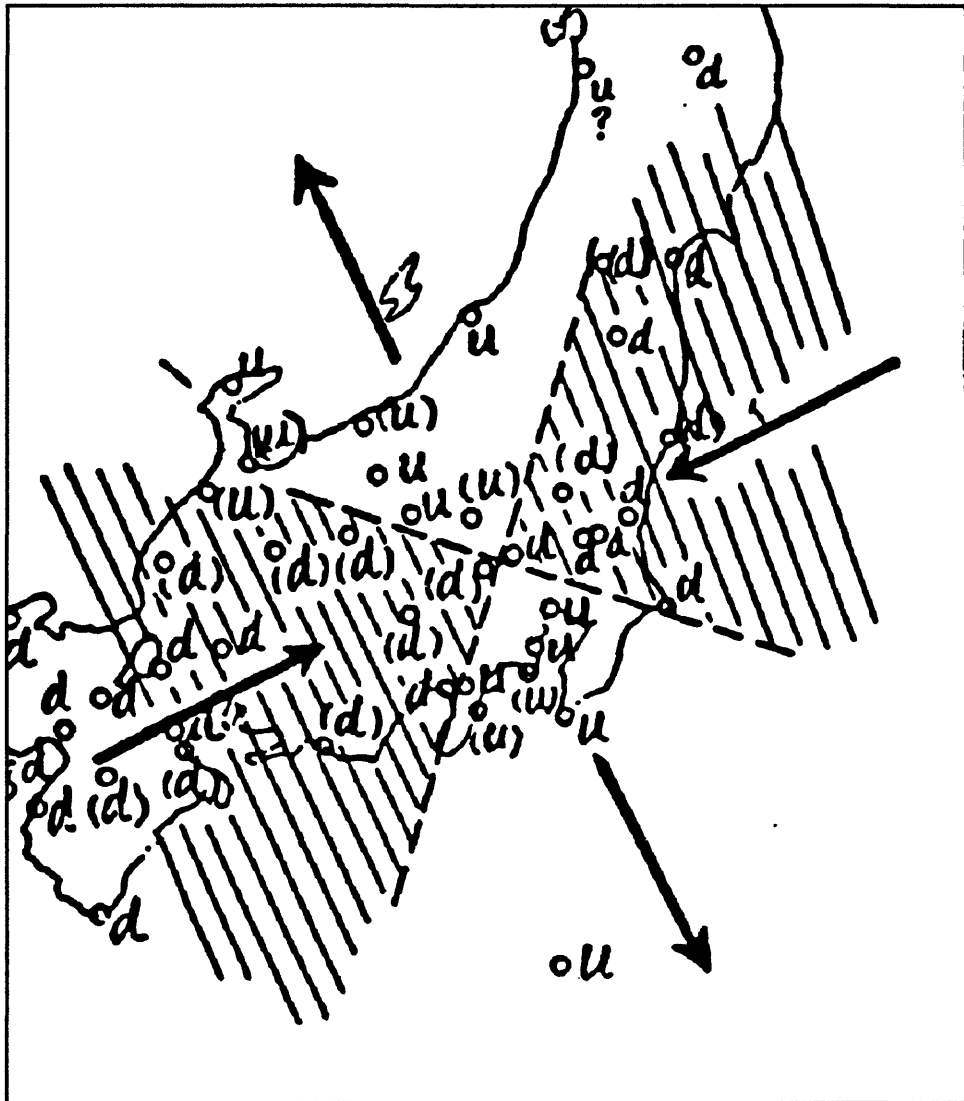


Figure 7 Distributions of compressions and dilatations for earthquake of Sept. 21, 1931 [Honda, 1934].

near vertical incidence, compressional first motion appears as up motion, and dilatational motion appears as down at the surface. As Honda's diagram shows, there is a clear pattern, with the ups and downs separated into four quadrants. Honda inferred that the direction of maximum compression must lie in the dilatational quadrants, and the direction of minimum compression must lie in the compressional quadrants. This is shown by the large inward and

outward pointing arrows in each of the quadrants. Although Honda's analysis worked for this earthquake, whose fault is vertical and slip is horizontal, in more general circumstances we will need to account for refraction of the seismic waves and for faults whose orientation and slip are oblique to the ground surface.

Figure 8 shows a schematic diagram of an arbitrarily oriented fault with arbitrarily oriented slip. The wedge shaped block of rock shown in the figure underlies the fault plane

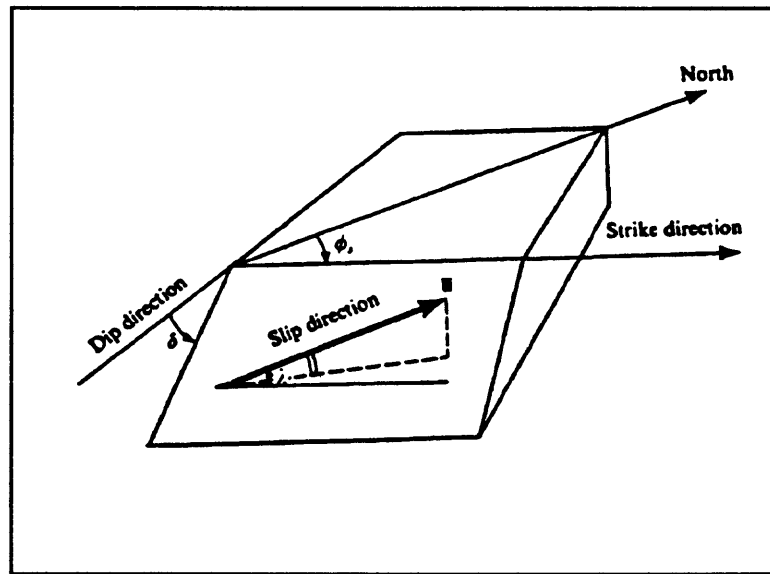


Figure 8 Fault plane parameters.

and is referred to as the foot wall block. The rock above the fault plane (which is not shown in this view) is called the hanging wall. The strike of the fault is the azimuthal orientation of a line formed by the intersection of the fault plane and a horizontal plane. This angle is symbolized as ϕ_s . The dip of the fault plane is the angle measured (perpendicular to the strike) from the horizontal to the fault plane, and is denoted as δ . The slip vector (\mathbf{u}) gives the sense of motion of the hanging wall with respect to the foot wall. It is defined by an angle (λ) measured in the fault plane, from the strike direction to the slip direction.

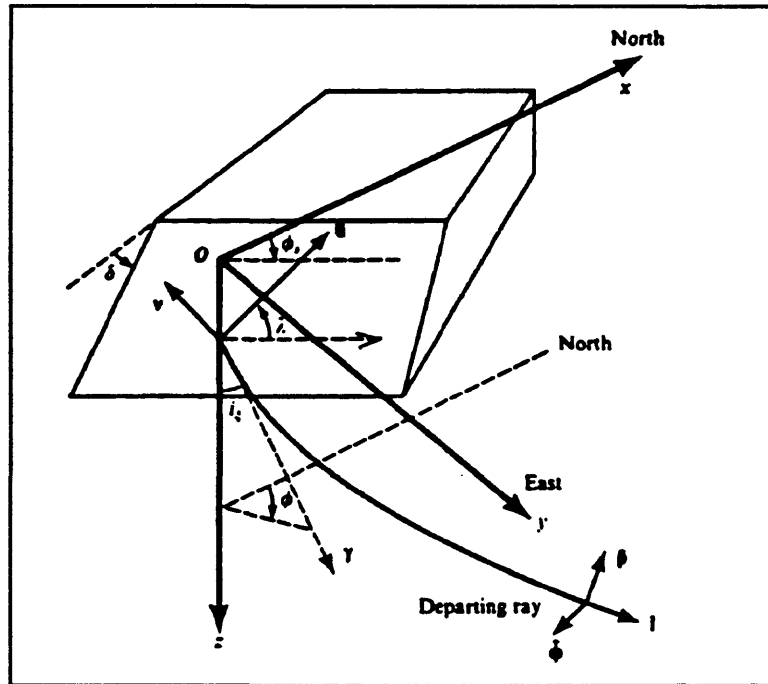


Figure 9 Cartesian coordinates for fault plane.

Figure 9 shows the same diagram but with the axes of a Cartesian coordinate system added, and also with a ray leaving the source region, as in Figure 6. Using the relations shown in this diagram we can move easily from source coordinates to earth coordinates, as required in the analysis of the focal mechanism. Specifically, the relations illustrated in Figure 9 can be used to develop a set of equations which define the radiation patterns of P, SV, and SH waves in terms of the angles shown in Figure 9. These equations are part of the means of relating an observed P-wave pulse to a specific point on the focal sphere. Figure 10 shows the equations as presented by Aki and Richards (1980).

$$\begin{aligned}
\mathcal{F}^P &= 2(\gamma \cdot \mathbf{v})(\gamma \cdot \dot{\mathbf{u}})/\dot{u} \\
&= \cos \lambda \sin \delta \sin^2 i_c \sin 2(\phi - \phi_s) - \cos \lambda \cos \delta \sin 2i_c \cos(\phi - \phi_s) \\
&\quad + \sin \lambda \sin 2\delta(\cos^2 i_c - \sin^2 i_c \sin^2(\phi - \phi_s)) \\
&\quad + \sin \lambda \cos 2\delta \sin 2i_c \sin(\phi - \phi_s), \tag{4.84} \\
\mathcal{F}^{SV} &= [(\gamma \cdot \mathbf{v})(\dot{\mathbf{u}} \cdot \hat{\mathbf{p}}) + (\gamma \cdot \dot{\mathbf{u}})(\mathbf{v} \cdot \hat{\mathbf{p}})]/\dot{u} \\
&= \sin \lambda \cos 2\delta \cos 2i_c \sin(\phi - \phi_s) - \cos \lambda \cos \delta \cos 2i_c \cos(\phi - \phi_s) \\
&\quad + \frac{1}{2} \cos \lambda \sin \delta \sin 2i_c \sin 2(\phi - \phi_s) \\
&\quad - \frac{1}{2} \sin \lambda \sin 2\delta \sin 2i_c (1 + \sin^2(\phi - \phi_s)), \tag{4.85} \\
\mathcal{F}^{SH} &= [(\gamma \cdot \mathbf{v})(\dot{\mathbf{u}} \cdot \hat{\boldsymbol{\phi}}) + (\gamma \cdot \dot{\mathbf{u}})(\mathbf{v} \cdot \hat{\boldsymbol{\phi}})]/\dot{u} \\
&= \cos \lambda \cos \delta \cos i_c \sin(\phi - \phi_s) + \cos \lambda \sin \delta \sin i_c \cos 2(\phi - \phi_s) \\
&\quad + \sin \lambda \cos 2\delta \cos i_c \cos(\phi - \phi_s) \\
&\quad - \frac{1}{2} \sin \lambda \sin 2\delta \sin i_c \sin 2(\phi - \phi_s). \tag{4.86}
\end{aligned}$$

Figure 10 Equations describing radiation patterns (from Aki and Richards, 1980).

SPHERICAL PROJECTIONS

Because the geometry is such a challenge, people have developed graphical methods to facilitate the description of fault orientations and positions of rays within the earth. One method, which I will explain, employs an equal area projection. Another method employs a stereographic projection. Both methods provide a means of plotting the rays and fault planes on a flat piece of paper.

Figure 11(a) shows the lower half of the focal sphere with a line OP projecting from the center of the sphere to the point P on the surface of the sphere. Figure 11(c) is a map view of the focal sphere showing the point P. Evidently, line OP strikes S 42°E and plunges 40°. The position of point P in Figure 11(c) is given by the following transformation. If the spherical coordinates of the point P are (R, α, β) where α is the azimuthal angle and β is the

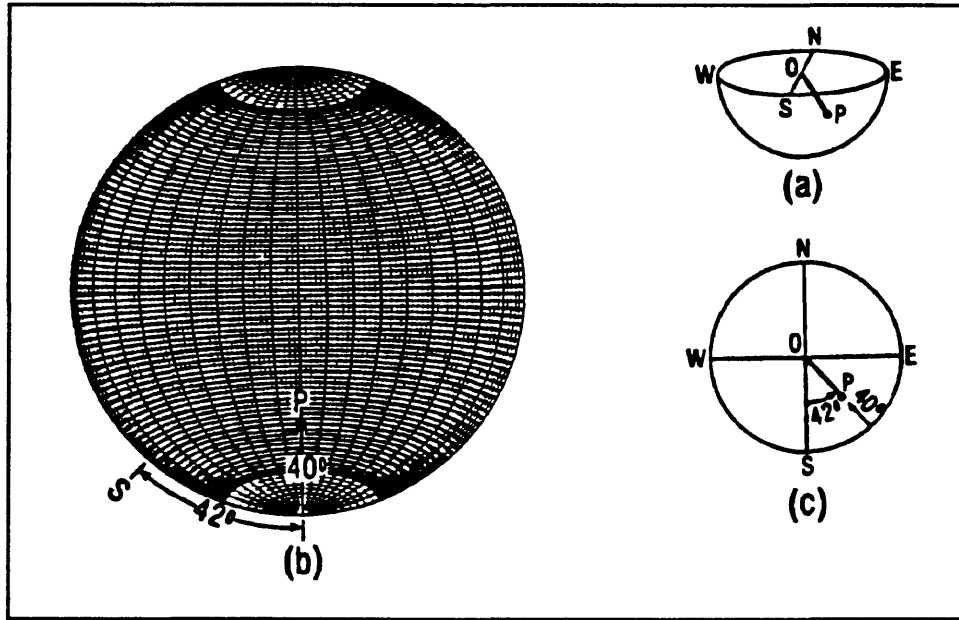


Figure 11 Equal area plot of a line.

take-off angle, then the equal angle transformation to plane polar coordinates (r, θ) is

$$r = 2R \sin(\beta/2) \quad , \quad \theta = \alpha \quad (1)$$

The grid of lines shown in Figure 11(b), which is called a stereonet, is produced by taking points at regular angle intervals on the focal sphere and transforming them in this manner, then plotting the transformed points. The meridians on the stereonet are the projections of great circles on the focal sphere. The major divisions (darker lines) are in ten degree increments. To plot the point P using the equal area net, first lay a sheet of tracing paper over the net and mark the north pole on the tracing paper. Then count divisions along the perimeter of the net until you reach the azimuthal angle of P. Mark that spot on the perimeter and then rotate the tracing paper until the mark is aligned with the east-west line on the equal area net. Finally count in from the perimeter of the net until the plunge angle has been reached, and mark the spot. That is the projection of point P (line OP).

The equal area projection maps only the lower hemisphere onto a plane and ignores the upper hemisphere. With a clever trick, however, we can also plot rays piercing the upper hemisphere, without any disadvantage. To do so, we recognize the fact that the radiation patterns from a double couple source are point symmetric (P-waves) or anti-point symmetric (S-waves) with respect to the origin. This means for P-waves, every upgoing ray is matched by a downgoing one of the same polarity on the opposite side of the focal sphere. Accordingly, upgoing rays (α, β) plotted on the lower hemisphere at the point ($\alpha + 180^\circ, 180^\circ - \beta$), are the ray's azimuth and take-off angles, respectively.

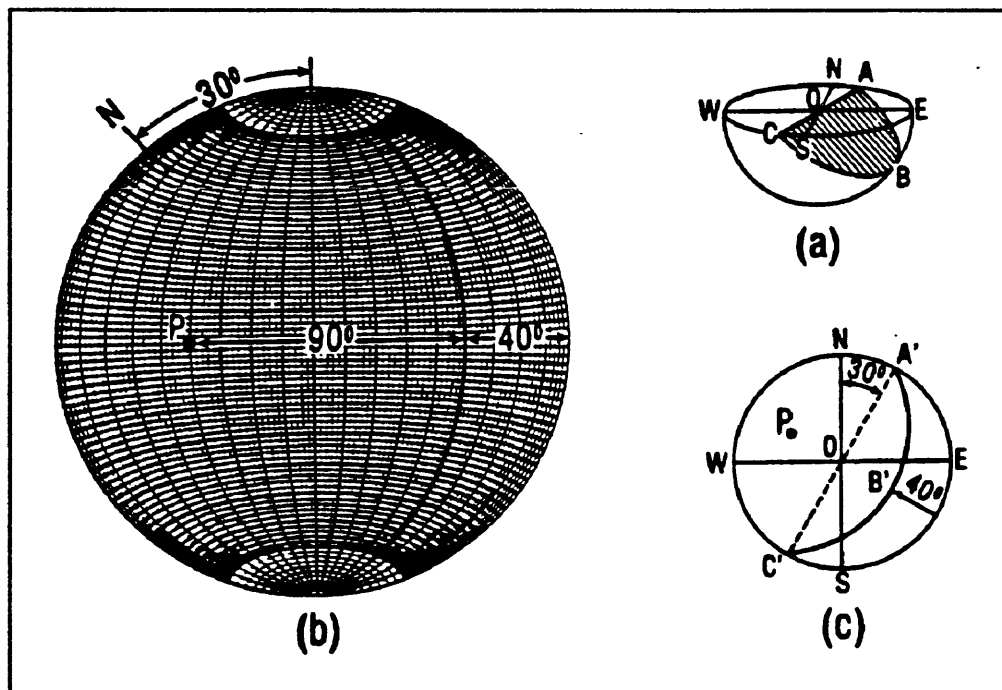


Figure 12 Equal area plot of a plane.

Since a plane is a collection of points, it follows that we can also plot planes using an equal area net. In fact, the great circles on the equal area net each represent a plane striking north-south with a dip between 0° and 90° , either to the east or west. We make use of this

fact to plot a plane of arbitrary orientation. First mark the strike direction of the plane on the tracing paper at the perimeter of the net. Then rotate that mark until it lines up with one of the poles. Move in from the perimeter until you reach the great circle with the same dip as the plane you are plotting, and trace that great circle onto the paper. When the tracing paper is re-oriented with respect to north, the line traced on the paper is the equal-area projection of the plane. This process is illustrated in Figure 12.

FAULT PLANE SOLUTIONS

Figure 13 is an illustration of a fault plane solution for a microearthquake near Santa Barbara, California. The two planes shown on the projection are orthogonal to each other and (almost completely) separate the compressions and dilatations into adjacent quadrants.

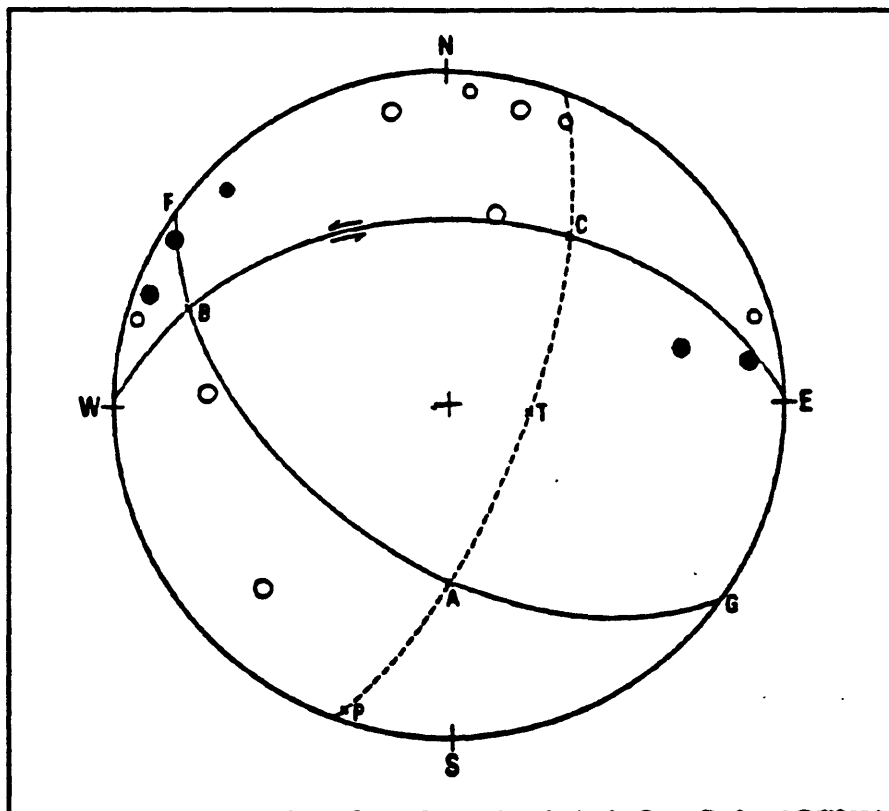


Figure 13 Fault plane solution.

The east-west striking plane has been identified as the fault plane and the other plane as the auxiliary plane. The least principal stress axis is marked as T and the greatest principal stress axis as P. The slip vector is marked as C and it indicates that slip was partly left lateral and partly reverse.

With the knowledge that a double-couple source will always produce a radiation pattern with compressions and dilatations in adjacent quadrants, the only variation being the orientation of the nodal planes that separate the quadrants, we should, in theory, always be able to find a fault plane solution of the form shown in Figure 13. The remaining question is how to go about doing it in practice.

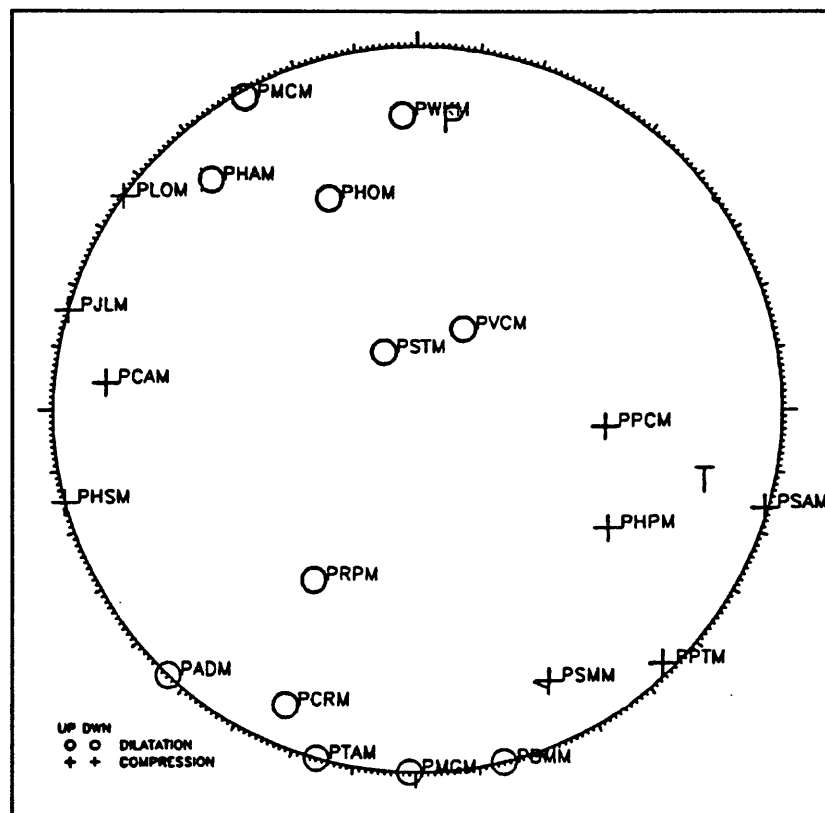


Figure 14 Equal area projection of first motions from earthquake near Parkfield, California.

Figure 14 shows a computer generated plot of compressions and dilatations plotted on an equal area projection of the focal sphere. The compressions are shown as crosses and the dilatations as open circles. Rays which were up-going on the focal sphere have been reflected through the origin and plotted on the lower hemisphere as heavy symbols. So, for instance, the ray to station PPCM actually was upgoing and toward the west, but has been plotted as a downgoing ray east of the origin.

The old fashioned method for reconstructing the focal mechanism of the earthquake would be to simply stare at the plot for a while until you can visualize two orthogonal planes (in equal area projection) that can separate most or all of the dilatations from the compressions. There are some graphical aids for doing this, but the procedure is essentially one of inspired guessing and trial and error. Of course, you can't just draw these planes freehand. You have to use an equal area net or some other aid to insure that you maintain the orthogonality of the planes as you adjust them. The process is fairly time consuming and this was the motivation for producing a computer program to automate it.

One such program--FPFIT--is described in the next chapter by Dave Oppenheimer. Dave will cover the methods used in that program, and some of the practical aspects of its use. I would just like to comment here on some of the program's features and limitations as they relate to the history and theory we have discussed so far.

There are two basic sources of error in the data which have to be accounted for. One is a blunder, in which the polarity of the first motion is simply reversed. This might result when some component of the network is repaired or replaced and wires get crossed. Another more subtle problem is that of lateral refraction of rays. As I mentioned earlier, we assumed

that the azimuthal direction does not change from the source to the receiver. In fact, there is often some horizontal refraction, and if the receiver is located near a nodal plane, the refraction can result in the arrival being plotted in the wrong quadrant on the focal sphere. McNally and McEvilly (1977) showed an example of that problem for an earthquake in California. They found that horizontal refraction across the San Andreas Fault could cause an arrival to be deflected from about 5° to 30° degrees in a horizontal direction.

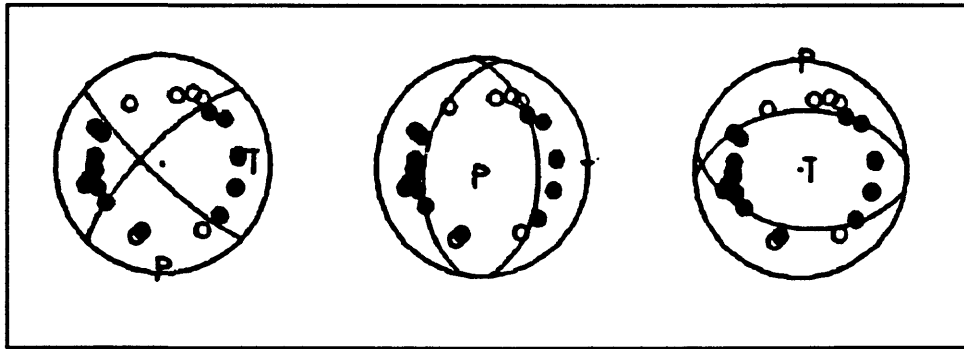


Figure 15 Non-unique Solutions to polarity data.

Yet another problem that arises is that of solutions being non-uniquely determined. Figure 15 shows the data from one earthquake, fit with three totally different solutions, which all fit the data equally well. When we search for solutions visually/manually, we often bring to the problem some prejudice about what the solution should look like. For instance, we may know that the earthquake occurred on the San Andreas fault, or in some subduction zone whose geometry we think we pretty well understand. Amazingly, these are the solutions we most often find, even if others are possible. If an automated means of finding solutions is used, that only reports the "best" fit to the data, there is a chance that the correct solution may not be reported (if it is not the best fit solution). The FPFIT program, by contrast, presents a suite of well fit models so that the analyst can choose the appropriate one based on

both the data fit and a knowledge of the regional geology and tectonics.

Figure 15 also points out the difficulties that can arise if you try to do fault plane solutions without enough stations or with an inadequate distribution of stations. Notice that all of the stations are at approximately the same distance from the origin in Figure 15.

Suppose, for instance, that there had been several compressions near the origin. That would immediately discredit the center solution and either the first or third solution, depending on the exact orientation of the compressions. Alternatively, a mixture of compressions and dilatations near the origin would tend to discredit both the second and third interpretations.

Still another potential problem in finding fault plane solutions is poor control over the hypocenter. This is particularly critical for stations which would plot near the origin since shifts in the hypocenter will affect the locations of those stations most strongly. In fact, unless the solution quality is very good, the stations that plot near the center of the diagram should be relied on least, even though they appear to exercise significant control over the solution.

Suggested Reading

- Aki, K., and Richards, P. (1980). "Methods of Quantitative Seismology." Freeman, San Francisco, California.
- Brillinger, D. R., Udias, A., and B.A. Bolt, (1980). A probability model for regional focal mechanism solutions. Bulletin of the Seismological Society of America **70**, No. 1, 149-170.
- Honda, H. (1962). Earthquake mechanism and seismic waves. J. Phys. Earth **10**, 1-97.
- Kasahara, K., (1963). Computer program for a fault-plane solution. Bulletin of the Seismological Society of America **53**, No. 1, 1-13.
- Lee, W. H. K., and S. W. Stewart, (1981) "Principles and Applications of Microearthquake Networks." Academic Press, New York, New York.
- McNally, K. C. and T. V. McEvilly, (1977). Velocity contrast across the San Andreas fault in central California: small-scale variations from P-wave nodal plane distortion. Bulletin of the Seismological Society of America **67**, No. 6, 1565-1576.
- Reid, H. F. (1910), "The California Earthquake of April 18, 1906, Vol. 2: The Mechanics of the Earthquake." Carnegie Inst., Washington, D. C.

19. THE FPFIT PROGRAM

by

D. H. Oppenheimer
U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

I am going to start about where Paul left off, although I will cover a little bit of the material he discussed. First I am going to talk about why this program was written and how it works.

Then I am going to discuss how you

use this program. You may have never seen this program before and may not use it again,

but it is nice to know what you can accomplish with a program like this one. Finally, I want to talk about some practical results and show you a few of the ways things can go wrong and lead you into bizarre interpretations.

This program started out from work that Eberhart-Phillips and Reasenbergs did on the Coalinga earthquake. They had collected a large amount of aftershock data using Calnet supplemented with some analog 5-day recorders. They had available some software written by Rex Allen which would automatically pick the P-wave onsets and determine first motions. Using this software, they were able to generate a tremendous amount of first-motion data rather quickly. To process this volume of data, they decided to write a program which would determine the focal mechanisms of these event. The program was limited to producing

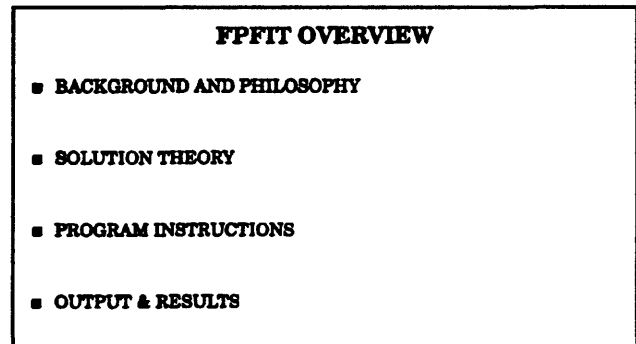


Figure 1 FPFIT overview.

BACKGROUND AND PHILOSOPHY

- **Eberhart-Phillips and Reasenber study of the Coalinga Earthquake**
 1. 2000+ earthquakes recorded by NCSN, field instruments.
 2. First-motions determined by computer
 3. Main shock fault plane not defined by aftershocks
 4. Constrained solution based on main shock mechanism
- **Reasenber and Oppenheimer Open File 85-739**
 1. Enhanced method for general solution
 2. Added confidence estimates (Mark Matthews)
 3. Documented and published code
- **Revised version 1988**
 1. Fixed a few bugs, improved user interface
- **Philosophy**
 1. Modern networks produce copious amounts of first-motion data-impossible to generate focal mechanisms by hand.
 2. Graphical analysis results in subjective bias.
 3. Robust in the presence of incorrect first-motions
 4. Produce feedback to detect reversed stations, lateral refraction effects.
 5. Ease of use

Figure 2 Background and philosophy.

solutions whose rake was equal to that of the main shock, which was reverse. It was the first time I had seen anyone process that much first motion data. One of the most interesting things about this project was that the Coalinga aftershock sequence did not define the fault plane. There was a raging controversy for about two years regarding which of the two nodal planes was actually the fault plane. You certainly couldn't tell from the locations of the aftershocks, so there was hope that having a large set of focal mechanisms for the aftershocks would shed some light on the problem. I had been working with seismicity in the

Geysers area and saw this program as a means to help with the analysis of seismicity in that region. However, the program would have to be generalized to deal with an arbitrary focal mechanism before it would be useful to me, so Paul and I collaborated to produce the first version of FPFIT, which was published in a USGS Open File Report. With the help of Mark Matthews, who was a graduate student in statistics at Stanford, we were also able to produce confidence statistics for the fault plane solutions produced by the program. In 1988 I worked some more on the program, fixing a few bugs and improving the user interface.

It used to be that a relatively small number of earthquakes were instrumentally recorded, and fault plane solutions were only done on very large earthquakes. Under those circumstances it was feasible to do fault plane solutions manually. However, modern earthquake networks produce so much data that the analysis has to be automated, or it will never get done. Also, we wanted to remove some of the bias that inevitably creeps into analyses that are done manually. We also wanted our program to be robust. There have been other programs produced for this sort of analysis, but in the presence of erroneous data they tended to become unstable. We also wanted the program to produce some feedback so that we could be aware of how good or bad our polarity data was.

FPFIT SOLUTION THEORY

There are certain assumptions built into this program. The first one is that we are going to consider only double-couple solutions. We use the P-wave radiation pattern for a double-couple source from Aki and Richards (1980). The program considers only P-wave first motions although, in principle, the code could be extended to include S-wave first

motions and P/SV ratios
also. When we wrote the
program, Calnet had only a
few horizontal component
stations, so there was not

that much good S-wave data

available. Also, S-wave first motion is much harder to detect reliably by automatic means.

We also assumed that first motion uncertainty is reflected in the arrival time weighting scheme (0 = very confident) to (4 = no confidence). Finally, we assume that there is no error in the take-off angles. That is, the velocity model is exact, and the hypocenter location is perfect.

The solution method of FPFIT is based on a grid search. We consider every possible source orientation. This approach was taken because of the non-linearity of the problem. Although this approach is very inefficient compared to something like least-squares, it takes only about 15 seconds to get a solution with the program running on a SPARC station and, maybe, twice as long on a PC. You can plow through a lot of data in a hurry. Figure 4 shows the basic problem geometry. The foot wall is shown in the figure. The slip vector is the direction of motion of the hanging wall. The rake angle is the angle measured in the fault plane between the strike vector and the slip vector. The resulting angle convention is shown in the figure. The strike angle is given by the intersection of the fault plane and a horizontal plane, referenced to north. The dip angle is the angle between the horizontal and the fault plane, measured perpendicular to the strike.

SOLUTION THEORY

■ Assumptions

- 1. Double-couple solution only**
- 2. P-first motions**
- 3. Arrival time weighting reflects first-motion uncertainty.**
- 4. No error in take-off angles (i.e., velocity model and resulting earthquake location are accurate)**

Figure 3 FPFIT solution theory.

■ **Method**

1. Grid search over all possible source models $[M^i]$, where M^i is a double couple mechanism defined by strike, dip, and rake.

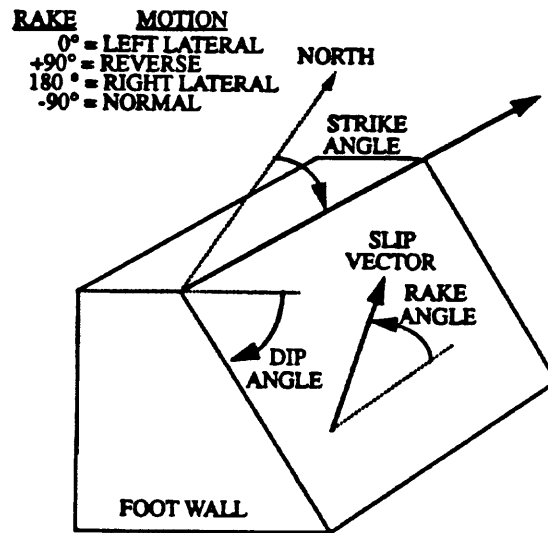


Figure 4 Fault plane geometry for problem.

Figure 5 shows the misfit function used to evaluate the various models. F^i is the misfit value for the model being tested. The program keeps track of every one of them, and when all combinations have been tried, the lowest value is chosen to be the solution. There are K stations per earthquake. For the model, we sum over the K observations the absolute value of the difference between the observed polarity and the predicted polarity. The polarities have assigned values of either 0.5 or -0.5, so if they are the same, the absolute value is zero. Otherwise it has the value 1. Each observation is weighted by its reliability W_o^k and (optionally) by a weight based on its distance from a nodal plane. The reason the second weighting factor is optional is that there are two schools of thought on the relative

2. Calculate a one-norm "misfit function" F_i for each earthquake by comparing the observed polarity at the k^{th} station with the polarity predicted for each M_i .

$$F^i = \frac{\sum_k (|P_{obs}^k - P_{pred}^{i,k}| \cdot W_o^k \cdot W_{pred}^{i,k})}{\sum_k (W_{obs}^k \cdot W_{pred}^{i,k})}$$

where P is the observed and predicted first motion polarity (+0.5 for compression, -0.5 for dilatation).

The term W_{obs}^k is the observed weight of the reading at the k^{th} station, and the term

$$W_{pred}^{i,k} = \sqrt{A(i,k)}$$

where $A(i,k)$ is the predicted P-wave radiation amplitude for the i^{th} source model at the k^{th} station.

3. Keep track of all F^i and choose the source model with the minimum score (F_{min}). In case of tie score, choose the solution for which the denominator of Equation 1 is maximized. This criteria maximizes the distance on the focal sphere between the observations and the nodal planes.

Figure 5 Solution method (Continued).

importance of data near a nodal plane. Some people think near-nodal data is very important since it tightly constrains the nodal plane. Others think that because of the possibility of lateral refraction and because such data tends to be low amplitude, that it should be given the least weight. The program lets you choose which approach you want to use. The practical effect of using the optional weight is to change the uncertainty estimate of the solution.

4. Specification of one plane and its rake is sufficient to represent both the slip and auxiliary plane. Therefore, FPFIT attaches no special importance to the plane used to represent the model.
5. The misfit function is determined by a two-stage three-dimensional grid search. The first (coarse) stage uses 20° increments in each of the three parameters (strike, dip, and rake). The coarse search identifies the solution corresponding to F_{min} and, when they exist, multiple solutions corresponding to significant relative minima in misfit. The second (fine) stage searches the 90% confidence region of the coarse search region with 5° increments in strike, dip, and 10° in rake.

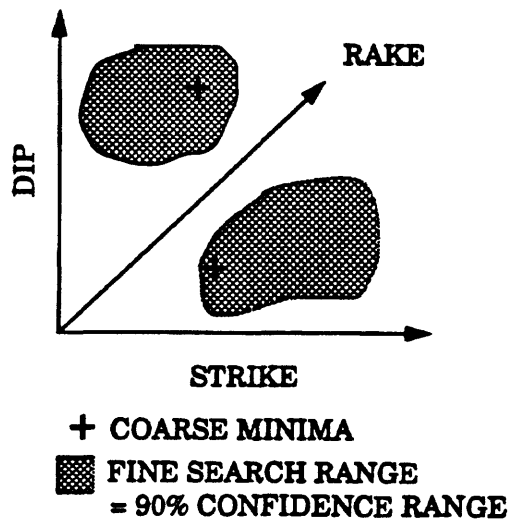


Figure 6 Solution method (Continued).

It is quite possible that a whole region of the focal sphere may fit the data perfectly. In that case, out of all the models with a perfect score, the one in which the denominator of the misfit equation is maximized will be chosen as the solution. This criteria maximizes the distance on the focal sphere between the observations and the nodal planes. You may or may not be happy with that result, but at least it is an objective criterion in that case. However, if

you have knowledge about the geology of the region which leads you to prefer some solution which is not quite what FPFIT produced, your judgement should take precedence.

Because of the symmetry of the problem, and because of the orthogonality of the nodal planes, the number of possible orientations which must be searched is only half that which would be required if there were no symmetry in the problem. By specifying the orientation of one plane and the rake, the other plane and its rake are completely specified. The program doesn't know which is the fault plane and which is the auxiliary plane. It is up to the analyst to make that determination. So, the approach taken in FPFIT is to calculate the solution for one plane and generate the other plane afterwards.

To speed things up, we do a coarse search first. The misfit function is determined by a two-stage three-dimensional grid search. The increment size for the first stage is 20° . The program determines the coarse solution F_{\min} or multiple minima if they exist. Figure 6 shows a hypothetical situation in which two minima have been generated after the first grid search. The program then computes the 90% confidence range for the minima. Then a fine grid search (grid interval = 5°) is conducted within the 90% confidence regions. In this manner, a relatively high resolution is achieved without spending an excessive amount of time searching regions where no solution is possible. The program knows when the various regions containing minima are discrete because the confidence regions do not overlap.

Before talking more about handling multiple minima, I want to discuss how the confidence ranges are determined. The way in which the confidence limits are determined affects both computational time and the interpretation of the reliability estimates produced by the program. The program makes the assumption that there is a priori knowledge of the

6. Confidence range is determined from *a priori* knowledge of the variance of the data. To do this we take advantage of the fact that the data are provided with assigned "quality codes" (0, 1, etc.) that allow us to treat the misfit function F as a sum of binomial processes.

$$\text{Var}(F) = \frac{n}{\left(\sum_{j=1}^N n_j \cdot \bar{w}_j \right)^2}$$

$$\sigma_F = \sqrt{\text{Var}(F)}$$

where n = number of observations,

N = the number of data classes represented by the observations

n_j = the number of observations in data class j

$$\bar{w}_j = \frac{1}{\sqrt{r_j(1-r_j)}}$$

and

r_j = the error (discrepancy) rate for class j

Figure 7 Determination of confidence range.

distribution of errors in the first-motion data. The assumption is that it is a binomial process where you have "zero-weight" readings, "one-weight" readings, "two-weight" readings, and so on. Readings with a weight of zero are assumed to be correct about 95% of the time. You wouldn't want to use any readings with a reliability of less than 50% since that doesn't add anything to the solution. As you can see from Figure 7, the variance of the misfit is given by

the number of observations divided by a squared sum of weights for a given data class times the number of observations in that class. You can see that if the discrepancy rate for some class (r_j) is zero then the weight for that class (\hat{W}_j) will become infinite, resulting in a variance of zero.

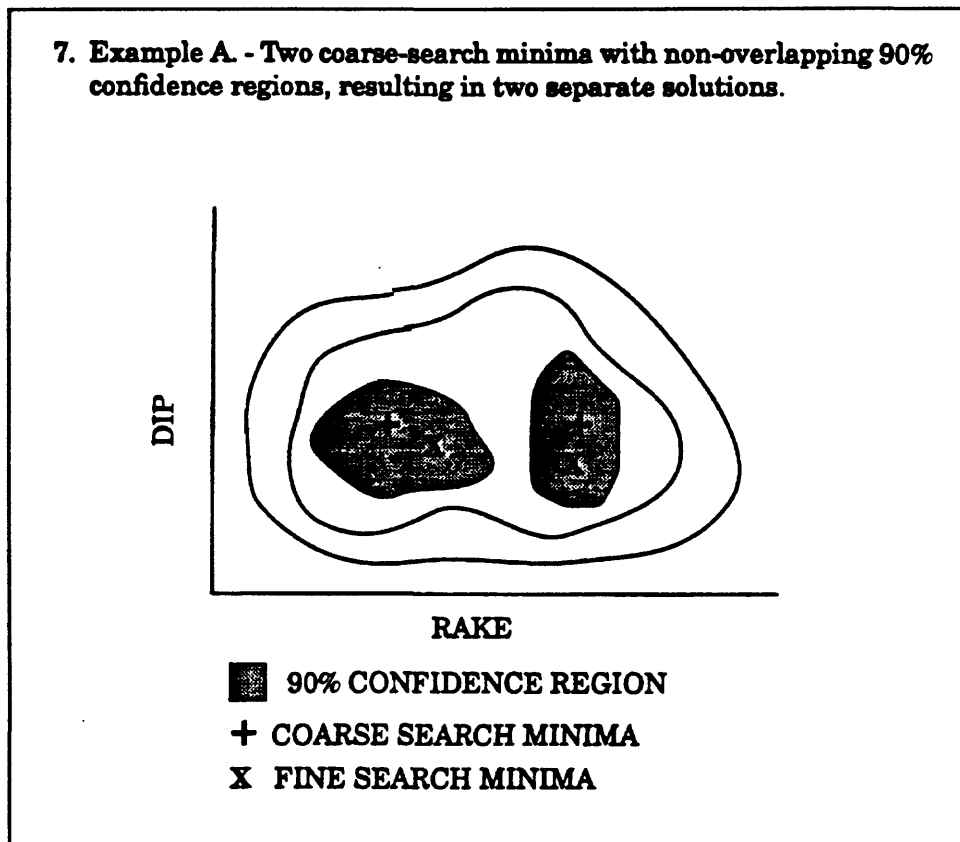


Figure 8 Handling multiple minima (case A).

You will probably not be able to provide estimates for the discrepancy rates for each class when you begin using FPFIT. However, FPFIT has a set of built-in discrepancy rates which you can use when you first start using the program. Since FPFIT provides information on discrepancies by class as part of each solution, after doing enough fault plane solutions, you should be able to determine an empirical estimate of the discrepancy rate for each of the

data classes used in your network. Eventually, you should replace the built-in discrepancy rates with your empirically determined rates. If you use the wrong discrepancy rates, you will probably still get the right solution. However, the 90% confidence regions that result from using your discrepancy rates may be so large that the program will spend an inordinate amount of time on the fine grid search.

Figure 8 illustrates a situation where the program has gone through its coarse search and has resolved two minima with non-overlapping 90% confidence regions. In the figure the "+" symbols represent the coarse search minima and the "X" symbols represent the fine search minima. In this case, the program will report two distinct solutions and those solutions will have parameters as specified by the "+" solutions.

If you used different values for the discrepancy rates, it is possible that the two confidence regions would overlap. In that case the situation might look similar to Case B in Figure 9. Here, the coarse search has produced two minima but instead of having the two non-overlapping confidence regions of Figure 8, their confidence regions overlap to form one large confidence region. The result is a single solution shown at point "X". This is one possible effect of changing the discrepancy rates. Of course, it is also possible to go the other direction and have the program report large numbers of similar solutions. The moral here is that if you use values which are similar to the default ones, the program will work quite well. But if you change the discrepancy rates radically, you may get unacceptable results from the program. So I recommend only making large changes in the discrepancy rates if you have very good reason to do so.

Example B. - Two minima, but secondary minima is within 90% confidence range of first, resulting in a single solution.

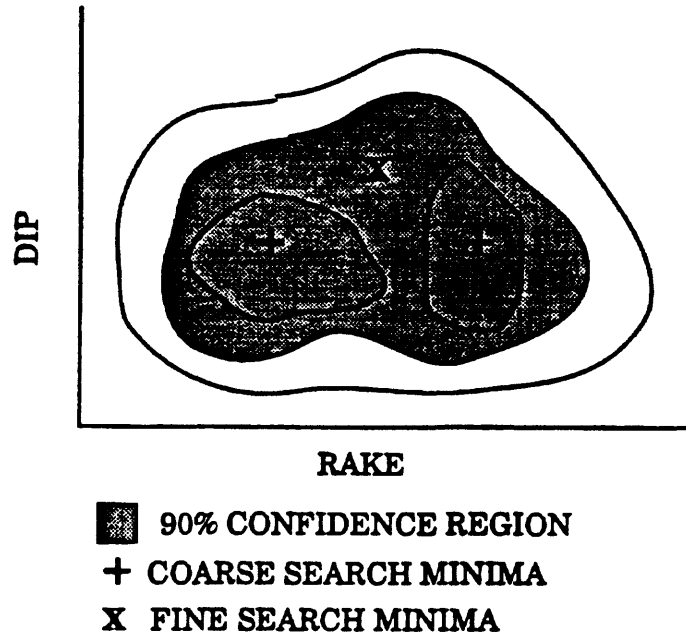


Figure 9 Multiple minima (case B).

The program may sometimes tell you that it can't converge. What this means is that the problem is so poorly determined, that after the coarse search found the minimum "+" in Figure 8, the fine search moved the minimum to the edge of the 90% confidence region. The program will actually produce a solution but it will mark it with a "C" indicating lack of convergence. The program expected to find the minimum within the 90% confidence region but instead the solution appears to be moving outside that region. Just as with a bad hypocenter solution, the confidence region is getting stretched out in space. The data is just too poor to constrain the solution very much.

Example C. - Fine search minima occurs on edge of search range, resulting in a "lack of convergence"

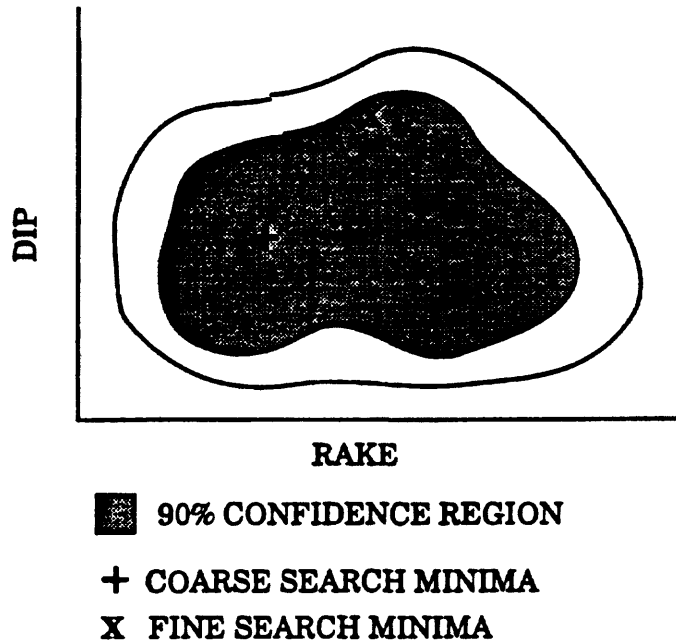


Figure 10 Lack of convergence.

Reporting confidence bounds on the solution is slightly tricky since the confidence regions can be so irregular. We would like to be able to report things like:

" fault plane strike = $30^\circ \pm 5^\circ$." Our solution is to approximate the confidence bounds on a parameter by projecting from the solution to the edge of the confidence region in the direction of the parameter and using the length of that line segment as the $\frac{1}{2}$ range of that parameter. So, for instance, in Figure 11, the confidence range for the rake is centered at the calculated value and varies by \pm the length of the heavy horizontal line.

Example D. - Two dimensional portrayal of reported 90% confidence region. Since FPFIT conducts a grid search, the confidence region is an irregular shape. The shape is best portrayed by the confidence range of the P- and T-axes shown at the lower right of the focal mechanism plot. An approximation of this region given in the summary line, as shown below.

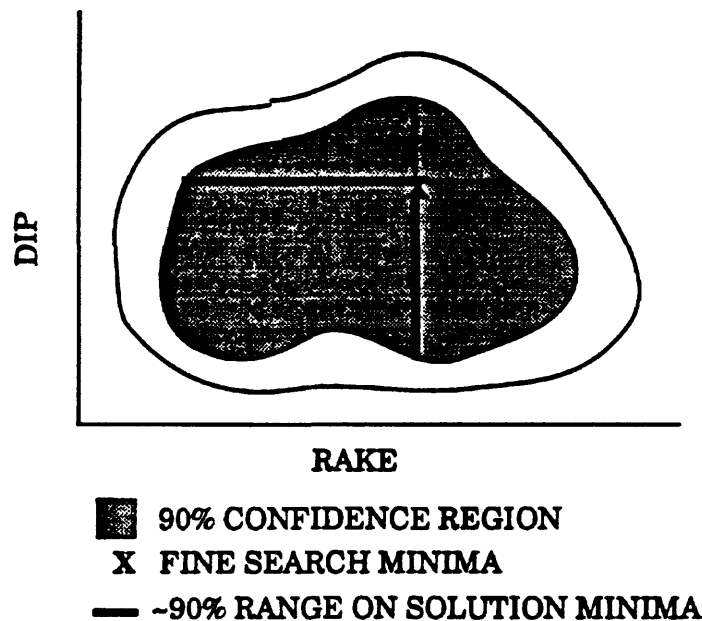


Figure 11 2-D portrayal of 90% confidence region.

PROGRAM OPERATION

Figure 12 shows an opening screen from FPFIT. The screen is essentially, a list of commands . Much of this is different from the description given in the Open File Report, since I have made modifications since that time.

Most of the commands are fairly obvious. For instance, "hyp" is used to set the file name of the input file, "for" tells FPFIT what the format of

the input file is, "out" is used to set the name of the output file for reports, and so on. Most people will want to use the "pol" command and create a file for input into the plotting program. I don't recommend using the "fit" option since it prints out the full fit function for the earthquake and is quite voluminous. The "ttl" option is used if you want to insert your own title. Normally you don't have to worry about anything else since the program has defaults built in. After any necessary i/o commands have been executed, you can do several things. You may compute a fault plane solution by issuing the "fps" command. To inspect the parameters, use the "sho" command. If you do this on a routine basis, you may want to

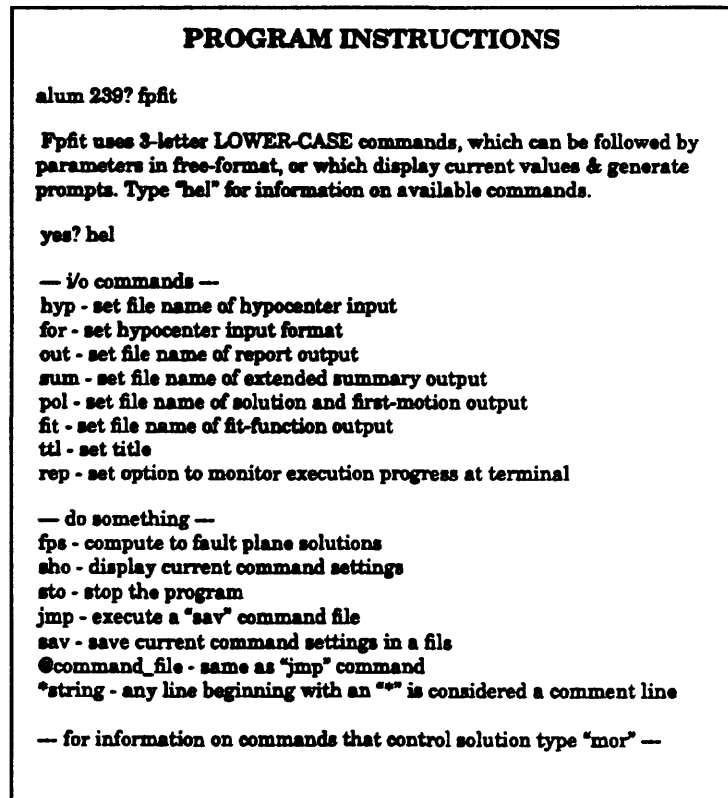


Figure 12 Program instructions.

save all the command settings by executing the "sav" command. I think this is all pretty straight-forward.

You can also constrain solutions using some of the commands shown in Figure 13. For instance, you can choose to weight the data by the P-wave radiation function by means of the "amp" option. You can have the program report only the first solution instead of multiples by

```
yes? mor

-- solution control --
amp - set option to weight data by p-radiation amplitude function
bst - set option to search for multiple solutions
cmp - set option to generate a composite solution
mag - set minimum acceptable event magnitude
obs - set minimum # of p first-motions per event (ignored for composites)
dis - set maximum permitted distance
res - set maximum permitted p-residual

ain - set permitted angle-of-incidence range
kil - set names of stations to ignore
rev - set names of stations with reverse polarities
hnd - assign first-motion error rates to p-wt codes of hand-picked data
mac - assign first-motion error rates to p-wt codes of machine-picked data
dir - set restricted downdip azimuth search range
dip - set restricted dip angle search range
rak - set restricted rake angle search range
fin - set option to restrict fine search range to coarse search range
yes?
```

Figure 13 Extended FPFIT commands.

means of the "bst" command. If you have several earthquakes in a file, you can instruct FPFIT to do a composite solution by means of the "cmp" option. One option you may use frequently is the "kil" option. By means of this option, you can instruct the program to ignore the data from a specified station. You might need to do this if you have a station which is getting refracted arrivals which are contaminating the solution. The "hnd" and "mac" commands are what you use to assign the error rate which is used in determining confidence intervals. The reason we broke this out into two categories is that machine picked arrivals are usually not as reliable as hand picked arrivals. If you have some information about the geometry of the fault plane, you might choose to restrict the search range using the "dir", "dip", or "rak" commands.

Figure 14 is a summary of the four kinds of output files produced by FPFIT. The

report file, which I will discuss shortly, gives you a lot of information which will help you evaluate your data. The summary file looks a lot like a HYPO71 summary card, and is used as input to two plotting programs; QPLOT and

PTPLOT. There is the "fit" file which I mentioned earlier with all the information about the misfit function. Finally, there is the "pol" file which contains the fault plane solution, the 90% confidence region, and the station polarities. This is the required input file for the plotting programs FPPAGE and FPPLLOT.

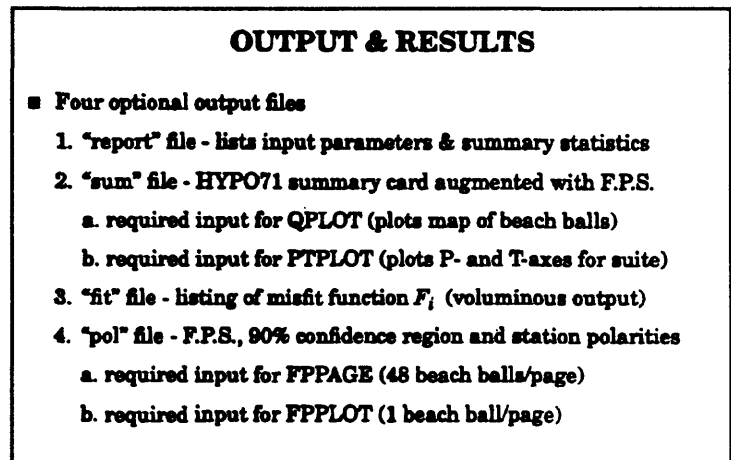


Figure 14 Output and results.

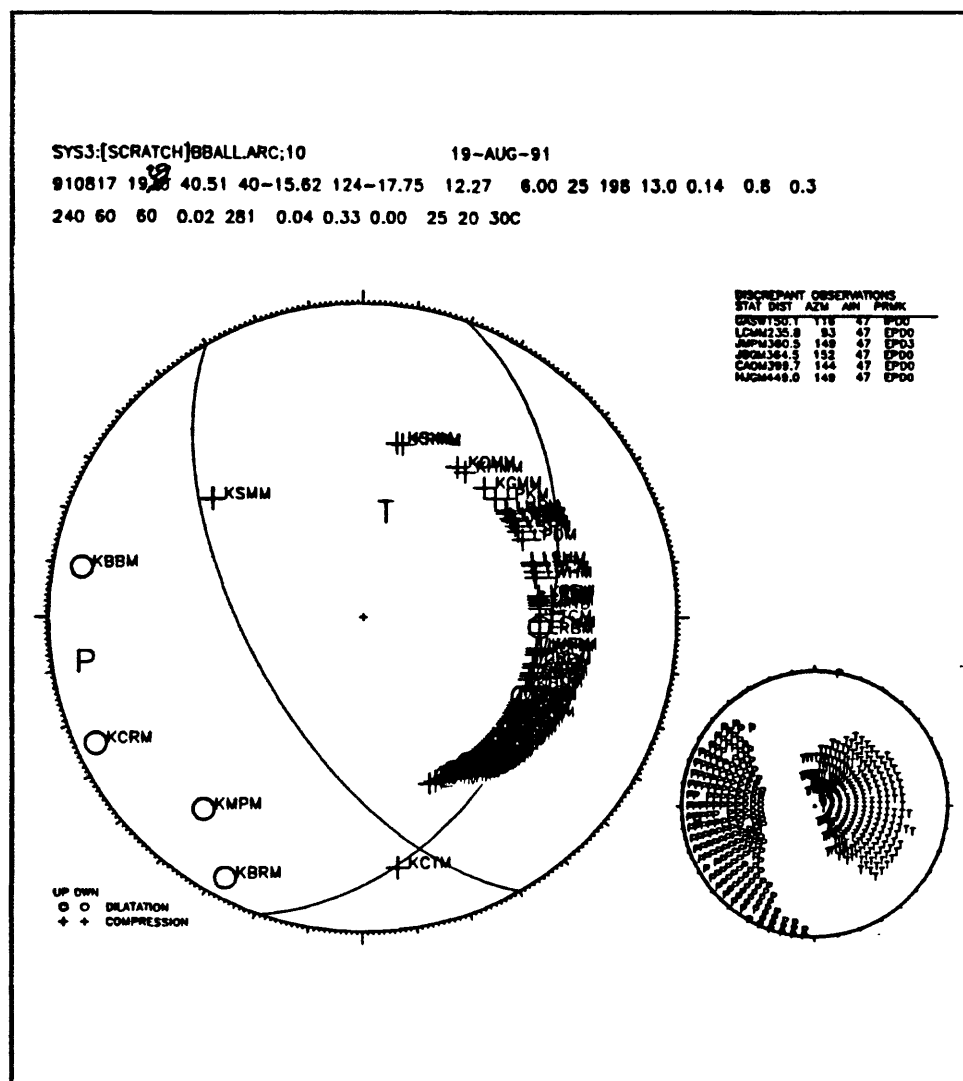


Figure 15 Sample output.

Figure 15 shows an example of output which is from a magnitude 6 earthquake which occurred in August, 1991 near Cape Mendocino. The top data line (after the title) is a duplicate of the HYPO71 summary card. It gives the origin time, latitude, longitude, depth, magnitude, number of stations, azimuthal gap, distance from nearest station, and RMS. The second line contains the fault plane solution. This example is for only one plane. The first four parameters are: the azimuth of the down-dip direction of the fault plane, the dip of the

plane, the rake angle, and the misfit value. The table labeled "discrepant observations" provides a listing of the data which does not fit the solution. This includes the station name, distance in kilometers, the azimuth, take-off angle, and the P remarks.

On the stereonet plot, the bold-faced symbols are for rays which would have plotted on the upper hemisphere but which the program has reflected through the origin. You should be very wary when you see up-going rays which plot near the middle of the "beach ball." Just a slight change in the hypocenter location can move them around significantly, so the fault plane solution may depend in an unstable manner on the accuracy of the hypocenter location. The curved "band" of station locations is due to the fact that those arrivals are all P_n (refracted off Moho). The small plot of P and T axes shows the 90% confidence region for the solution. It gives you a sort of graphical representation of the uncertainty in the solution. Basically, this is a thrust fault solution, but the fault plane solution has a very large uncertainty. Numerically, the uncertainty is shown as the last three parameters of the fault plane solution line. In this case, the uncertainty in strike is 25° , in dip the uncertainty is 20° , and in rake, the uncertainty is 30° . The "C" following the rake uncertainty indicates that the solution did not converge.

Some other numbers displayed in the fault plane solution line are the number of machine picked first motion readings (0.00), the number of first motion readings (281) and the 90% confidence region above the minimum. This is a change made about 6 months ago. It came out of a request to look at the Loma Prieta data. What we wanted to test was whether a constrained solution would still fall within the 90% confidence interval of the unconstrained solution.

```

parameter settings for fpfir
-----
ttl 1 '/usr/openssl/arc/9104r1.arc'          Pri Sep 13 14:23:58 1991'
hup '/usr/openssl/arc/9104r1.arc'
out '/usr/openssl/ep/9104.rap'
run '/usr/openssl/ep/9104.epa'
pol '/usr/openssl/ep/Pol/9104.pol'
jmp '9104.inp'
for 4
mag 0.
obs 15
dis 0.1000E+06
res 100.0
sin 0.      100.0
amp 1
bet 0
fin 1
rep 1
amp 0
had 0.1790E-01 0.5160E-01 0.1774 0.2165
mac 0.3954      1.000      1.000      1.000
dir 90.00      250.0      20.00      5.000
dip 10.00      90.00      20.00      5.000
rak -100.0     160.0      20.00      10.00
rev 1 'CHBB'      0      0
rev 2 'FRIB'      0      0
rev 3 'HIMB'      0      0
rev 4 'PASC'      0      0
rev 5 'SINC'      0      0
rev 6 'CHLO'      0      0
rev 7 'DBOG'      0      0
rev 8 'EXSO'      0      0
rev 9 'LACO'      0      0
rev 10 'MILO'     0      0
rev 11 'MOCO'     0      0
rev 12 'CVLL'     0 001204
rev 13 'ABJM'     0      0
rev 14 'ARRM'     0 000602
rev 15 'BSCN'     0      0
rev 16 'CBMM'     0      0
rev 17 'LCMM'     0      0
rev 18 'MFVM'     0 091107
rev 19 'PPRM'     0      0
rev 20 'PSMM'     0 090321
rev 21 'ADNR'     0 041105
rev 22 'ADNR' 000309      0
rev 23 'ANTR'     0 041105
rev 24 'ANTR' 000309      0

SUMMARY OF STATIONS HAVING POLARITIES IN DISCREPANCY WITH BEST FIT SOLUTION (* DENOTES REVERSED STATION)
STATION      DISCREPANCIES  AGREEMENTS      TOTAL  WEIGHTED ERROR RATE  TOTAL ERROR CONTRIBUTION
BAVM          15           04           99          0.057          8.0005
BSCN          19           76           95          0.070          8.0006

```

Figure 16 Parameter settings.

Figure 16 is an example of a report file generated by FPFIT. It echoes all your input commands so that if things don't look right you can trace where the error originated. Then it lists the station name, how many times it produced a discrepancy, how many times it was in agreement, total number of times used, weighted error rate and total error contribution. To be useful, these statistics should be the result of processing a fairly large suite of data. In the case of this sample, a whole month of data was processed to produce this summary.

BENK	16	51	67	0.088	0.0004
BCCM	12	97	109	0.053	0.0003
BCCB	5	12	17	0.110	0.0001
BCKR	3	63	66	0.007	0.0000
BCKH	7	42	49	0.019	0.0001
BENK	16	93	109	0.024	0.0002
BENK	17	95	112	0.065	0.0006
*BENK	6	94	100	0.029	0.0003
*BENK	2	91	93	0.009	0.0001
TOTAL	2029	11987	14016		
SUMMARY OF HAND-PICKED DATA WITH RESPECT TO BEST FIT SOLUTIONS					
QUAL	DISCREPANCIES	AGREEMENTS	TOTAL	WEIGHTED	ERROR RATE
0	64	1878	1934	0.0227	**
1	390	2968	3358	0.1000	**
2	887	2248	3135	0.2649	**
3	151	221	372	0.3395	**
TOTAL	1492	7307	8799		
SUMMARY OF MACHINE-PICKED DATA WITH RESPECT TO BEST FIT SOLUTIONS					
QUAL	DISCREPANCIES	AGREEMENTS	TOTAL	WEIGHTED	ERROR RATE
0	452	3724	4176	0.0144	**
1	81	940	1029	0.0596	**
2	4	8	12	0.1372	**
3	0	0	0	0.0000	
TOTAL	537	4680	5217		
DISTRIBUTION OF SOLUTION MISFIT SCORES					
MISFIT SCORE	NUM	CUM NUM			
0.000 - 0.025	225	225			
0.025 - 0.050	76	301			
0.050 - 0.075	74	375			
0.075 - 0.100	43	418			
0.100 - 0.125	16	434			
0.125 - 0.150	9	443			
0.150 - 0.175	9	452			
0.175 - 0.200	6	458			
0.200 - 0.225	1	459			
0.225 - 0.250	0	459			
0.250 - 0.275	0	459			
0.275 - 0.300	0	459			
0.300 - 0.325	1	460			
0.325 - 0.350	0	460			
0.350 - 0.375	0	460			
0.375 - 0.400	0	460			
0.400 - 0.425	0	460			
0.425 - 0.450	0	460			
0.450 - 0.475	0	460			
0.475 - 0.500	0	460			

Figure 17 Continuation of summary report.

Figure 17 is a continuation of the summary report in Figure 16. By looking at this report you may be able to tell if a station is reversed. For instance, if the number of discrepancies for some station was 93 and the number of conforming observations was 14, then you could be pretty sure that station was reversed. If you do have a station which is reversed, you might want to look at the last column which shows how much that station

affected the total error in that suite of data. After listing the station statistics, the report lists the error rates by data class. In the case shown, the empirical weighted error rates are rather far from the initial assumed values, so following the error rate is a double asterisk to highlight this discrepancy. This indicates that you may want to use a value which is closer to what is listed in the report for future analyses. Next is a listing of the distribution of the misfit scores, and distributions of the dip ranges, strike ranges, and rake ranges.

INTERPRETING THE OUTPUT OF FPFIT

Now I want to spend some time discussing how to interpret the output of FPFIT.

Figure 18 is a map view of the magnitude 6 earthquake whose fault plane solution was

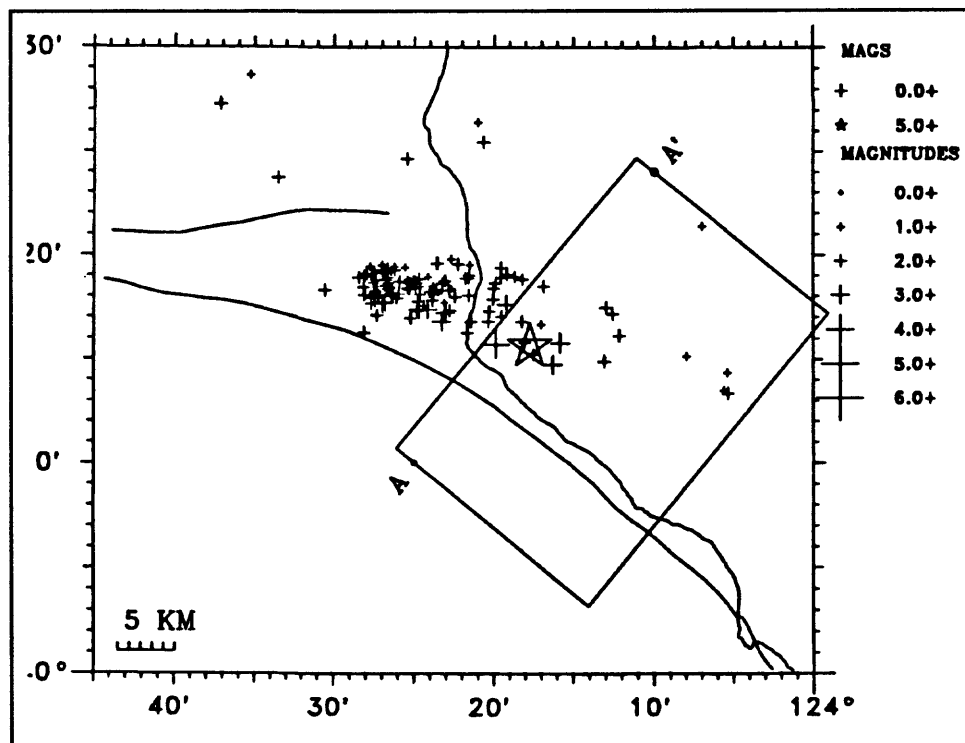


Figure 18 Epicenter map for Cape Mendocino earthquake.

shown in Figure 15, along with some aftershocks. The star marks the epicenter. Recall that the fault plane solution was for a thrust fault striking to the northwest. This doesn't seem to fit the pattern of aftershocks shown in the map.

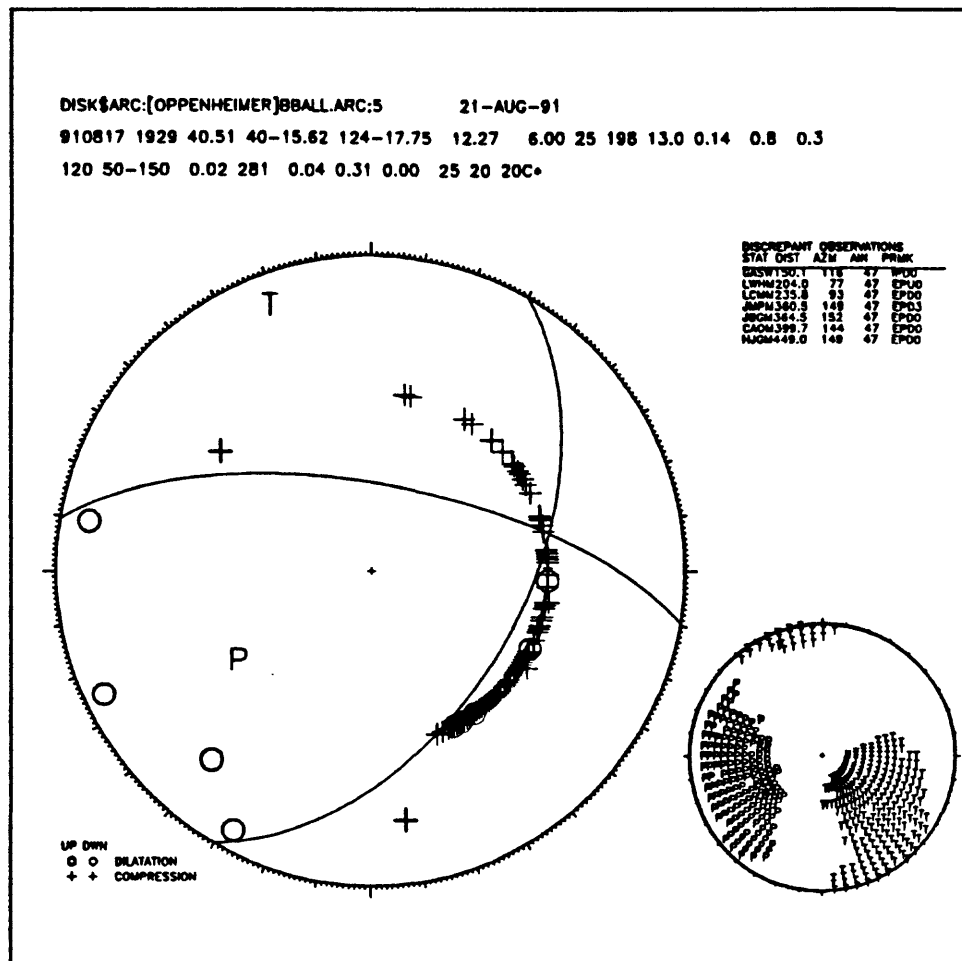


Figure 19 Alternate solution for Cape Mendocino earthquake.

However, the program also produced a fault plane solution shown in Figure 19. If you thought that the trend in aftershocks was delineating the slip plane, then you would conclude that the plane striking nearly east-west in Figure 19 was the fault plane, and the

motion was left-lateral with a component of normal dip-slip. However, we know that aftershocks occurred on the Mendocino transform fault, and its slip should be right-lateral. So, this fault plane solution is inconsistent with prior knowledge we have of the area.

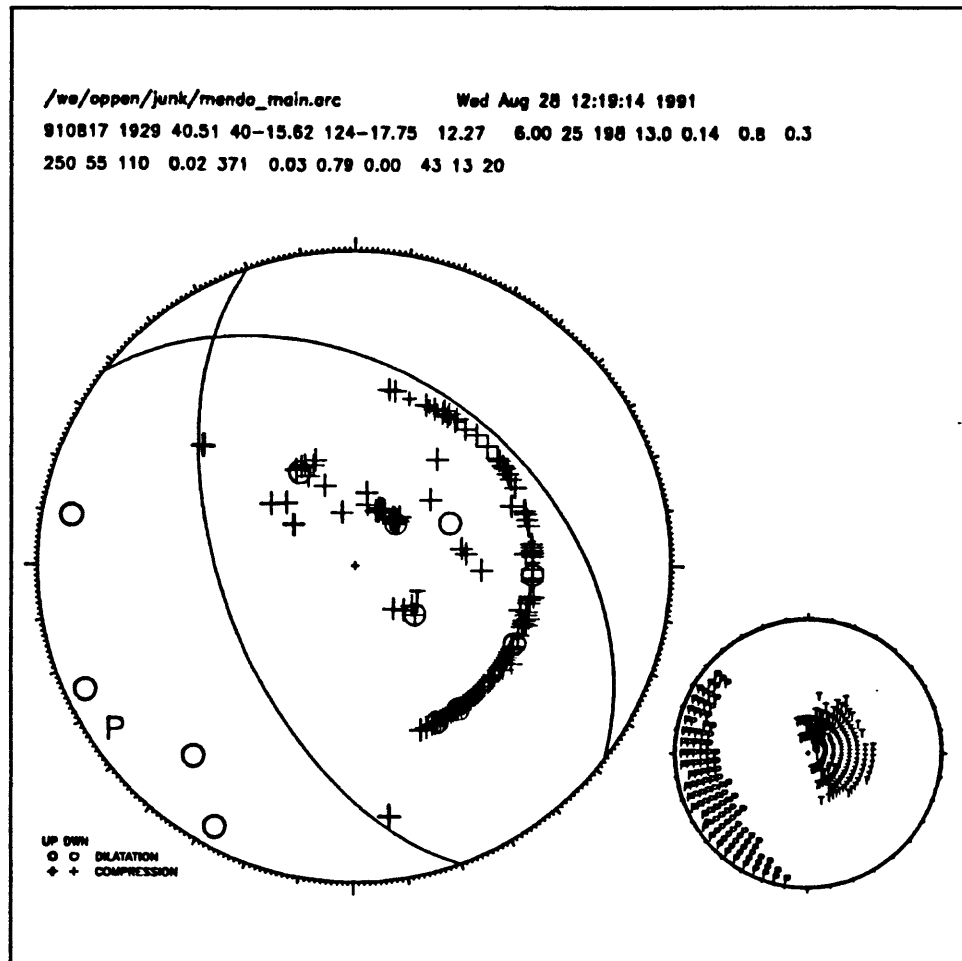


Figure 20 Fault plane solution using world-wide data.

To resolve the problem, I redid the fault plane solution with world-wide data added to the Calnet data. This solution is shown in Figure 20. As you can see, the additional data clearly constrains the mechanism to be a thrust fault. So why doesn't the fault plane line up with the trend of the aftershocks? To answer this question, consider the plot of background

seismicity plotted on the same map as shown in Figure 18. You can see that in the period August, 1974 through June, 1991 the area with the apparent aftershocks of Figure 18 is a rather active region. Apparently the earthquake triggered activity on this persistent feature, even though the earthquake did not occur on a fault plane paralleling the feature.

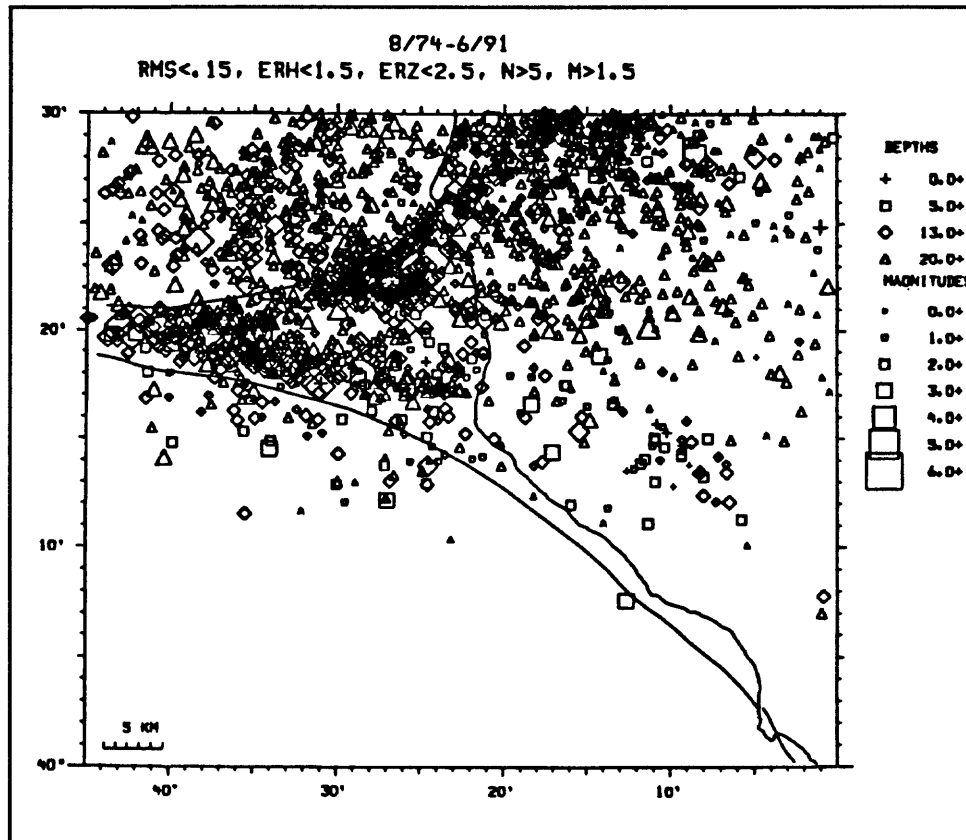


Figure 21 Background seismicity of Cape Mendocino area (8/74 - 6/91).

The focal mechanism plots shown so far were all produced with the program FPPLLOT. However, I mentioned that there is another program called FPPAGE.

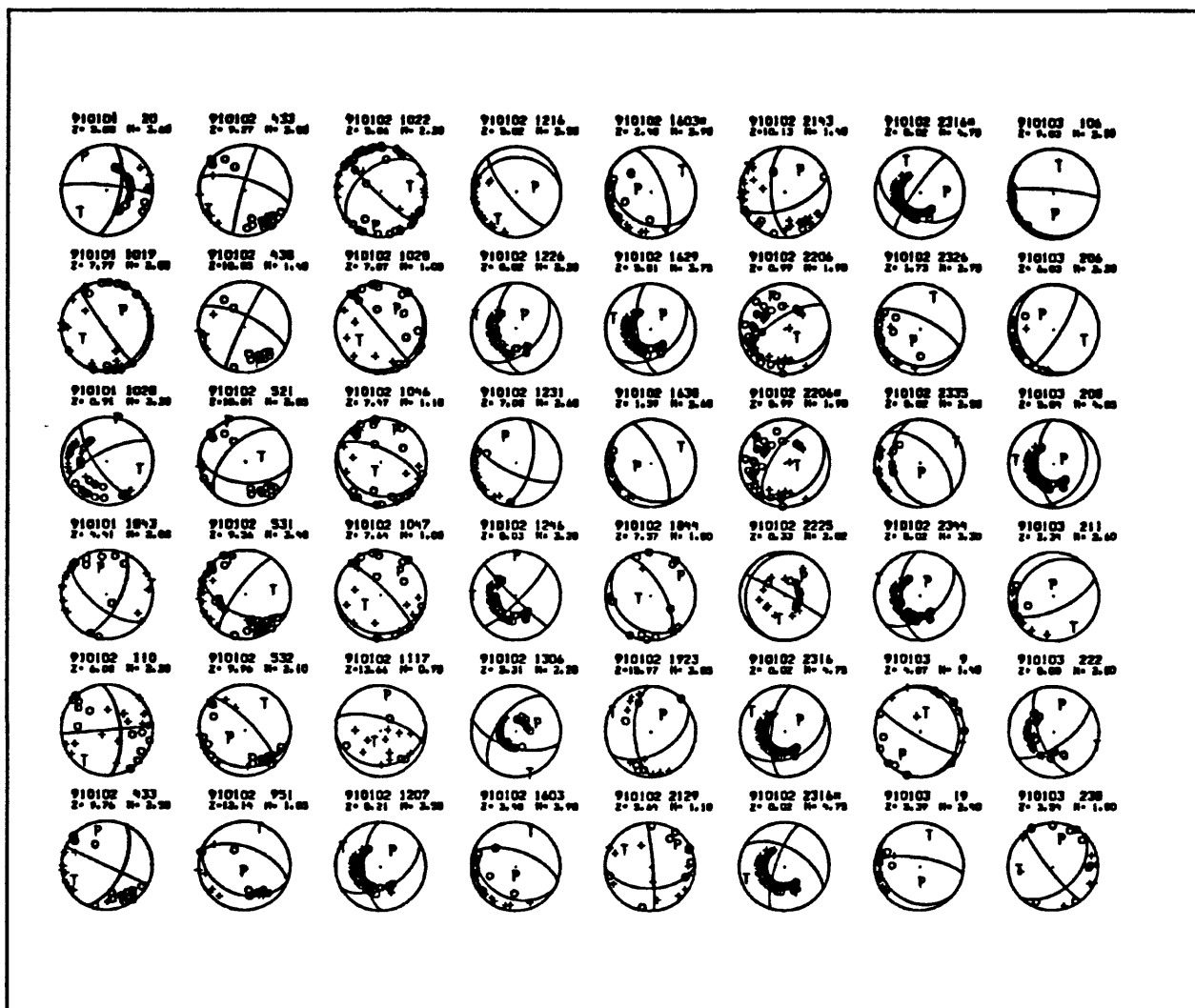


Figure 22 Output from FPPAGE.

Figure 22 shows the first page generated by FPPAGE from a monthly compilation of focal mechanisms for earthquakes recorded by the Calnet. Some of the plots here illustrate the kinds of meaningless results that you have to watch out for. The middle two plots on the first line show hypocenter parameters that include a depth of near zero. This is a tip-off that the hypocenter location is no good. I am not sure what is going on here. It could be an earthquake swarm off the coast. What ever the cause might be, the hypocenter location is

is going to have to be improved before it makes any sense to try to get a focal mechanism for this event. The power of the program is that you don't have to look at just one earthquake. If you observe long enough, you can start to put data together and then make some reasonable interpretations. One earthquake certainly does not provide enough information for you to gain any understanding of what is going on in some region.

Another example is provided by the study of the Morgan Hill earthquake. Figure 23 shows the main shock and aftershock sequence of the Morgan Hill earthquake. We ran all

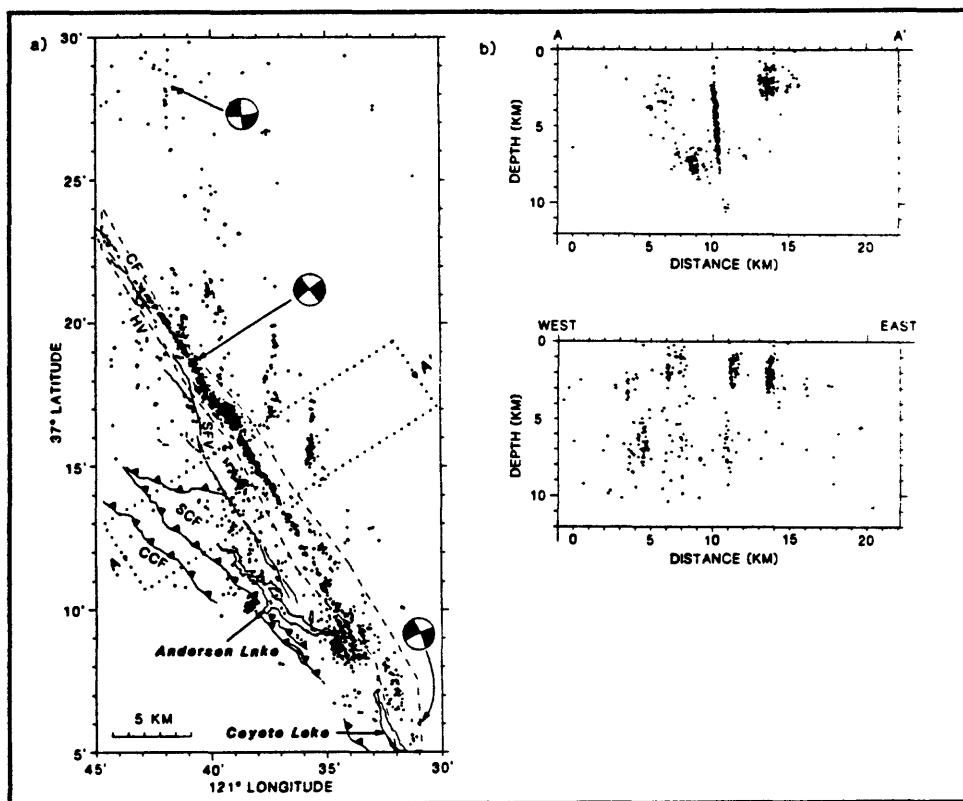


Figure 23 Aftershocks of the Morgan Hill earthquake.

this data through the program and got the set of representative mechanisms shown in Figure 24.

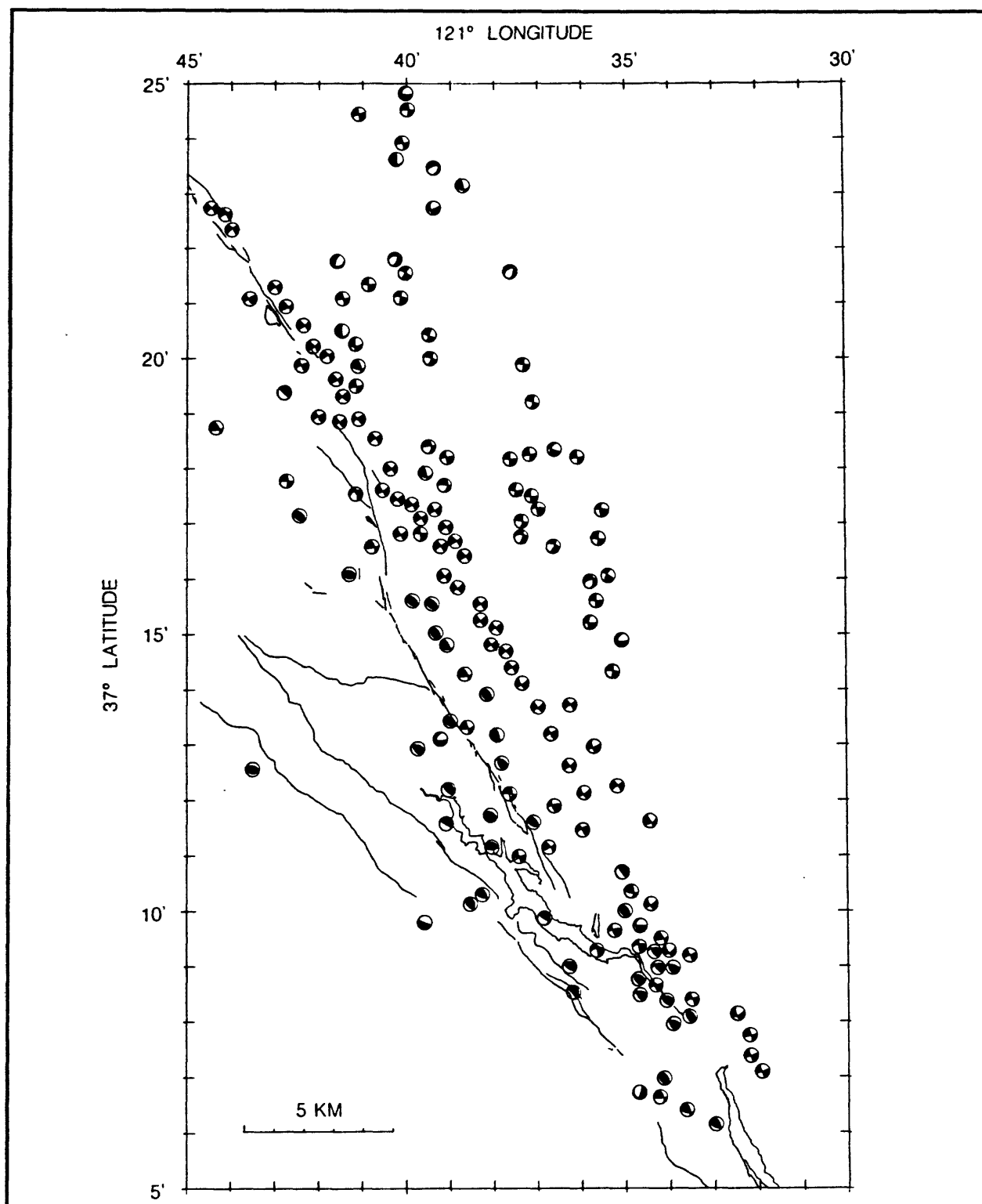


Figure 24 Fault plane solutions for 157 aftershocks of the 1984 Morgan Hill earthquake.

This is where we learned something that I don't think was apparent to anybody at the time. As expected along the Calaveras fault, the motion was predominantly right-lateral strike-slip with fault planes trending the same as the Calaveras fault plane. However, north-east of the Calaveras fault the motion was predominantly north-south right-lateral strike-slip and south-west of the Calaveras, the motion was primarily thrusting with the planes parallel to the San Andreas fault. This information provides some important constraints on the state of stress along the San Andreas fault. If you had only looked at the main shock you would have concluded that the P-axis was north-south and the T-axis was east-west. But if you look at everything off the fault, you reach the opposite conclusion. This analysis made us realize that it was important to look at all the shocks in some region when attempting to determine the state of stress there.

I want to give you a warning about lateral refraction. Figure 25 shows the fault plane solution for the Morgan Hill main shock. The set of readings which lie on the inner circle are P_n arrivals. The two stations CCO and CMH are clearly discrepant. These two stations are very close to the fault (less than 9 kilometers). The energy which reached these two stations was refracted across the fault plane. If all the P_n data had not been available (as would be the case for a small earthquake) the solution would have been influenced very strongly by these discrepant points.

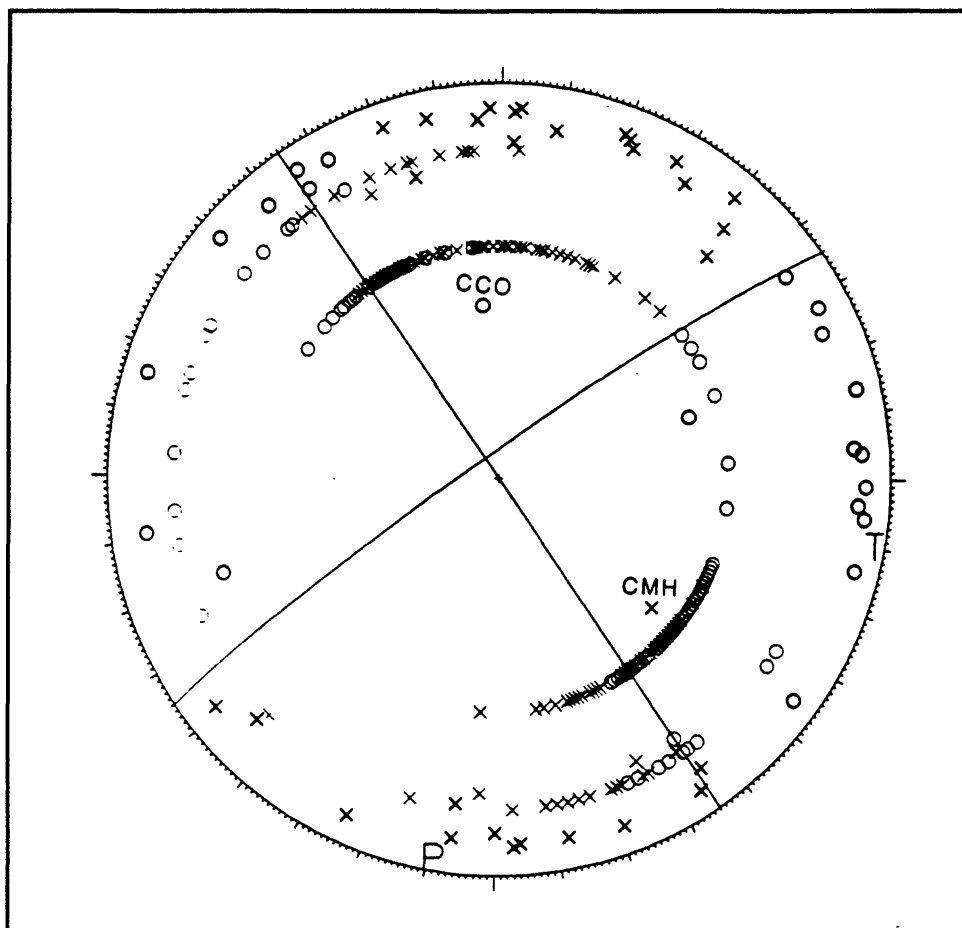


Figure 25 Fault plane solution for Morgan Hill main shock.

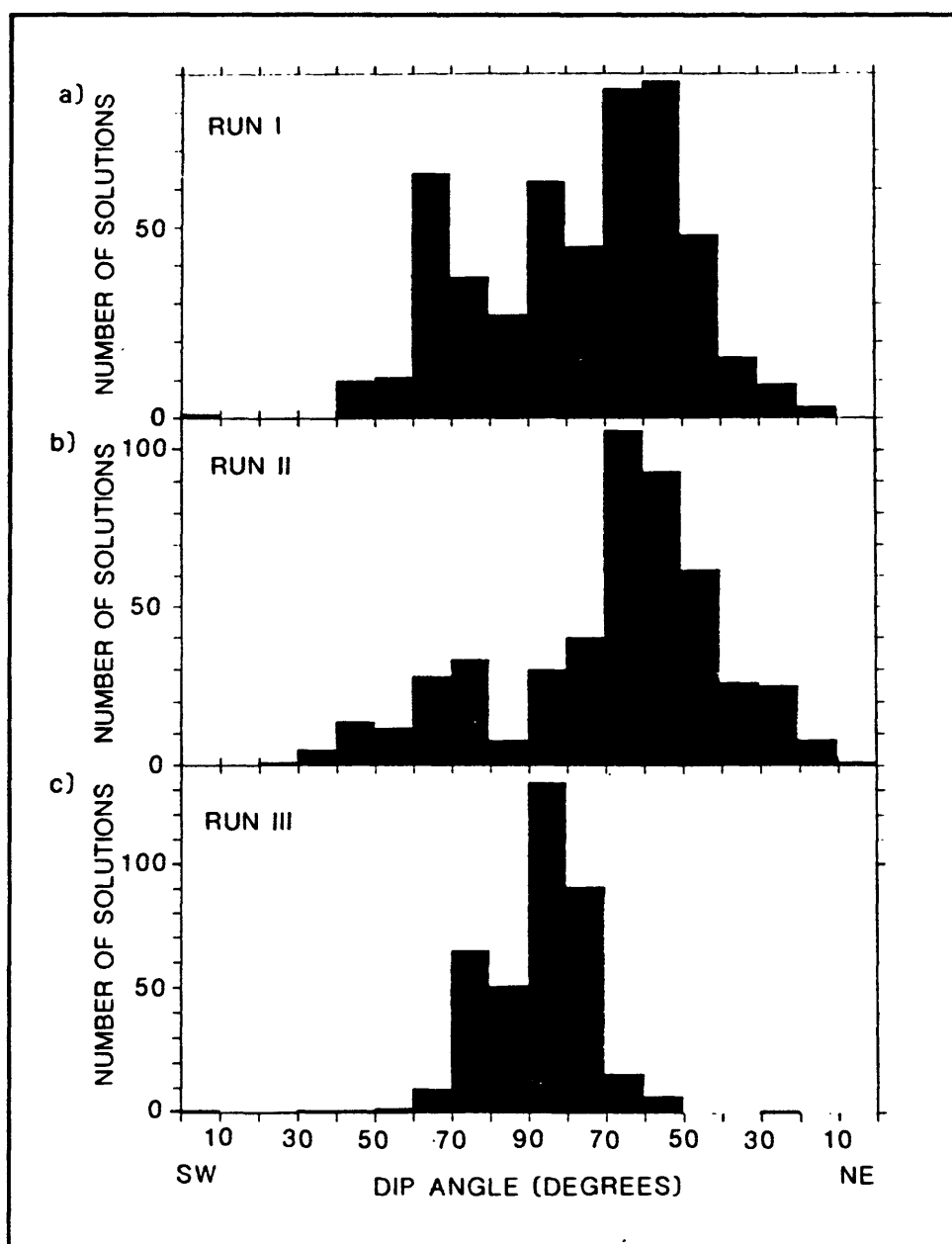


Figure 26 Histogram of dip angles for aftershock solutions.

Figure 26 shows histograms of the dip angles for fault plane solutions for the Morgan Hill aftershock series. These histograms were obtained by analysis of the report file provided by FPFIT. Because there was so much lateral refraction, the average dip was about 60° to

the north east. Yet, the aftershock sequence paints a picture of nearly vertical fault planes as shown in Figure 24. This inconsistency is due to lateral refraction. To counter this problem, we first turned on the feature which weights observations according to their distance from the fault plane. Even this was not enough and eventually, we had to throw out some of the observations that were seriously contaminated by lateral refraction.

You can also get a sense of how good your data is by plotting some of the data provided in the report file. Figure 27 shows the misfit data and the 90% confidence intervals for strike, dip and rake of the solutions. You can see that most of the solutions had a very

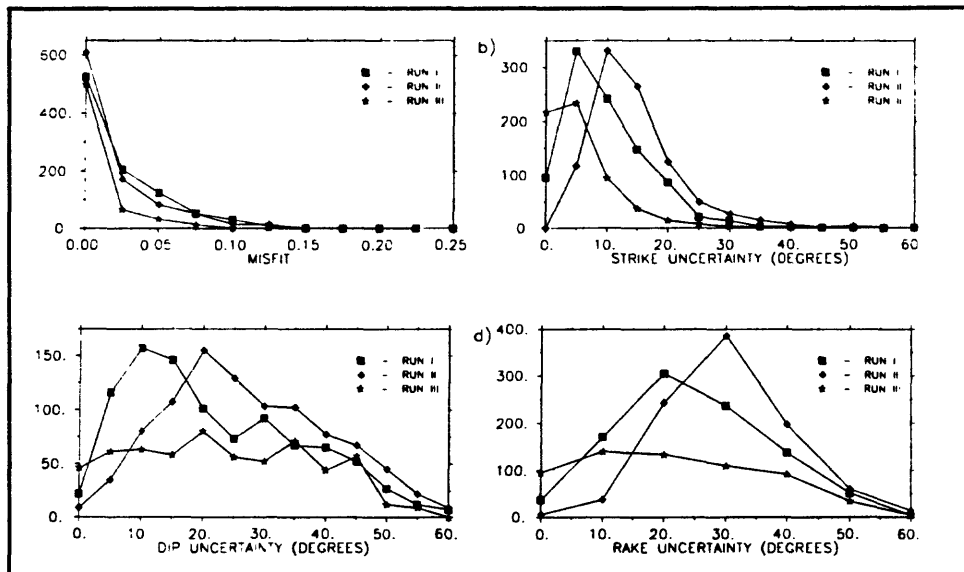


Figure 27 Distribution of misfit scores and 90% confidence estimates for solution parameters.

low misfit value. However, the strike uncertainty is around 15° and the dip and rake uncertainties are even worse because of the lateral refraction problem. This shows that when you process a large enough data set with FPFIT, you can get a characterization of how good your data are.

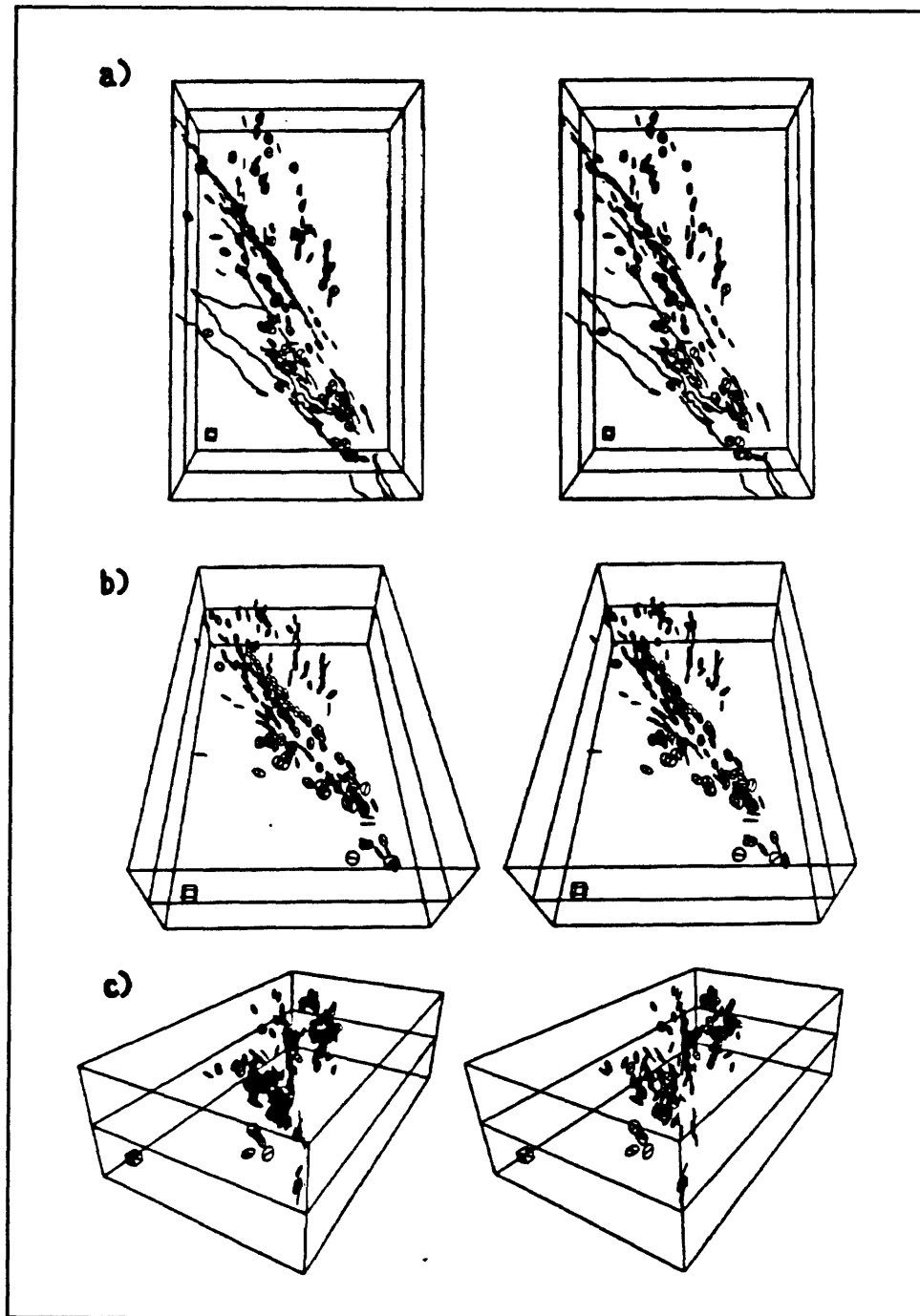


Figure 28 Three-dimensional plots of fault-plane solutions.

There is software available which can take the output of FPFIT and plot the

mechanism on a three-dimensional plot. Figure 28 is a set of views of the Morgan Hill data processed by such a program. On these plots, the different mechanisms are quite evident. You can see the events on the Calaveras fault clearly differentiated from the others not only by their epicentral alignment, but also by the alignment of their mechanisms. Similarly, the events north of the Calaveras all show a north-south strike and have vertical fault planes. This sort of three-dimensional visualization may help you to understand the regional stress field.

You can easily look at plots of P and T axes. There is a program called PTPLOT which was used to produce the plots shown in Figure 29. The program also allows you to plot confidence ellipses and find the center of the distribution. In this figure we have divided the earthquakes into groups with consistent P and T axes. There are also programs available to invert the data for stress.

Figure 30 shows some representative fault plane solutions for a region north of the Geysers. This plot looks like a real mess even though the solutions are particularly well determined. If you had only looked at one earthquake out of this ensemble you could arrive at a conclusion about the regional state of stress that is totally at odds with most of the other data. When you look at the whole ensemble, you see that the situation is complicated. This kind of data can arise in a region which is in a state of uniaxial compression or uniaxial tension. In that case, the orientations of the planes are free to rotate around the axis of maximum compressive stress. At first glance, the effect may seem to be "random" orientations, but careful analysis should reveal the underlying stress regime. The power of the FPFIT program is that it enables you to process the large amounts of fault plane data

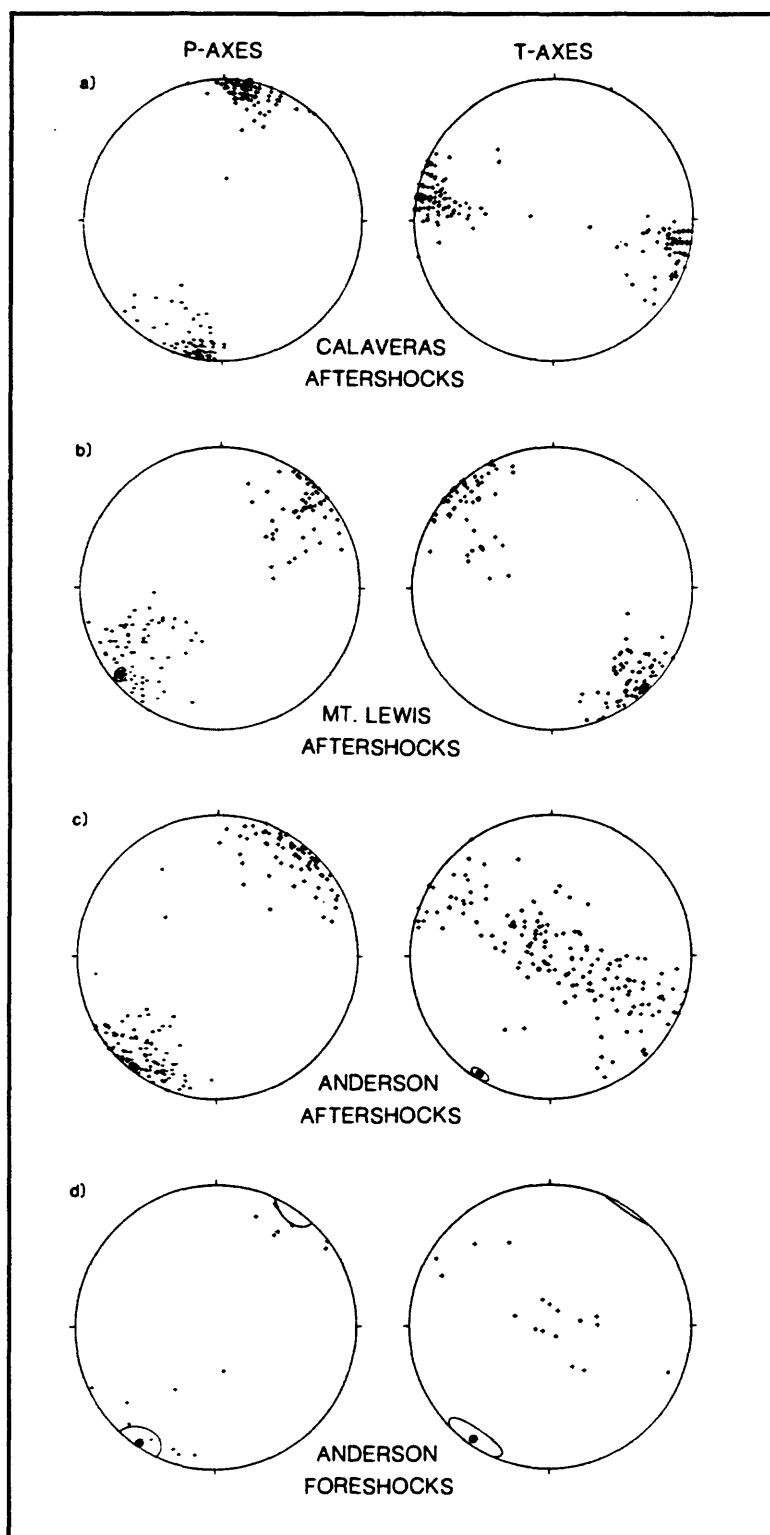


Figure 29 Distribution of P and T axes for Morgan Hill earthquakes.

which are necessary in order to unravel the details of a complicated stress regime.

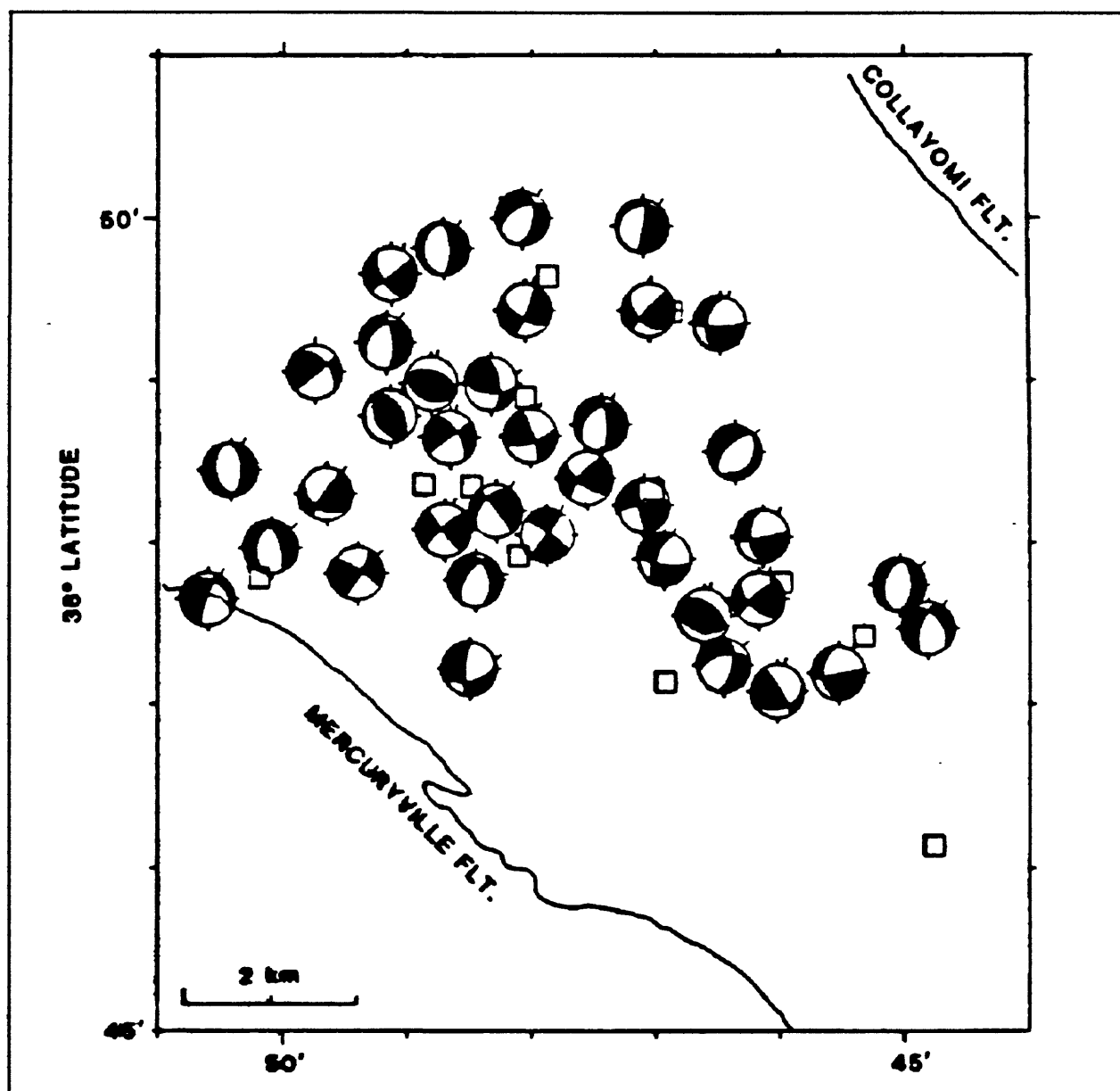


Figure 30 Representative fault plane solutions.

SUGGESTED READING

Aki, K., and P. Richards, (1980). "Methods of Quantitative Seismology." Freeman, San Francisco, California.

Oppenheimer, D.H., Reasenber, P.A., and R.W. Simpson, (1988). Fault plane solutions for the 1984 Morgan Hill, California, Earthquake Sequence: Evidence for the state of stress on the Calaveras Fault,. Geophys. Res., **93**, no. B8, 9007-9026.

Reasenber, P. and D. Oppenheimer, (1985). "FPFIT, FPLOT, and FPPAGE: FORTRAN computer programs for calculating and displaying earthquake fault-plane solutions. U. S. Geological Survey Open-File Report No. 85-739.

20. MATHEMATICS FOR SEISMIC TOMOGRAPHY

by

W. H. K. Lee
U. S. Geological Survey, Menlo Park, CA 94025

INTRODUCTION

In this chapter, I will provide a brief introduction to Seismic Tomography. The plan is to first spend some time talking about the development of medical tomography, since that was the first successful application of tomography. Then I will compare the field of medical tomography to the field of seismic tomography. In the course of this discussion, I will be talking about the central-slice theorem, the fundamental theorem on which the discipline is developed. Finally, I will discuss a commonly used technique of seismic tomography, or, simultaneous inversion, as it has often been called. This is the process of determining hypocenter parameters and earth structure simultaneously.

"Tomo" is from a Greek work that means "slice." In tomography, we examine a series of two-dimensional slices from a three-dimensional object, and use that information to infer the internal structure of the three-dimensional object. The beginning of

tomography can be traced back to the discovery of x-rays by W. Röntgen in 1895. You have probably seen the famous x-ray picture of his hand, which is in many physics

<u>Tomography</u>		
Tomo:	slice	
1895:	W. Röntgen	- x rays
1917:	J. Radon	- mathematics
1963:	A. Cormack	- method
1971:	G. Hounsfield	- scanner

Figure 1 History of tomography.

textbooks. Basically, the x-ray technique works because there is a large contrast between the absorption of x-rays by bone and the absorption of x-rays by flesh, so the film becomes exposed faster in those regions where the x-rays do not pass through bone. The medical field has found the x-ray technique to be highly useful, and it is used routinely today. However, we would like to be able to examine the soft tissues of the body to search for abnormalities within those tissues. The x-ray technique is not capable of doing this.

The mathematical foundation of medical tomography was created well before there was any serious effort to develop such a technique. In 1917 an Austrian mathematician by the name of Johann Radon proved the central slice theorem. For years his result was only of theoretical interest. However, in 1963, Allen Cormack proposed a technique for imaging soft tissue which was based on the central slice theorem. By 1971, the method became operational through the efforts of a British engineer by the name of Geoffrey Hounsfield. He used an x-ray scanner coupled with a computer in his system. Since then, the method has become refined and is widely used in hospitals around the world. It is now possible to image the entire human body with a resolution of about 1 millimeter.

The Nobel prize is not given for mathematics, so of the three men who contributed so much to the method, all but Radon received a Nobel prize. This is somewhat ironic, since the mathematical foundation that Radon laid has also had a great impact in the field of radio astronomy. In fact, during the 1970's two British astronomers received the Nobel prize essentially by applying the mathematics worked out by Radon to image the universe.

MEDICAL TOMOGRAPHY

My discussion of medical tomography is based on the paper that won Cormack his Nobel prize. The idea is really very simple (Cormack, 1963). There is an x-ray source of intensity I_0 and

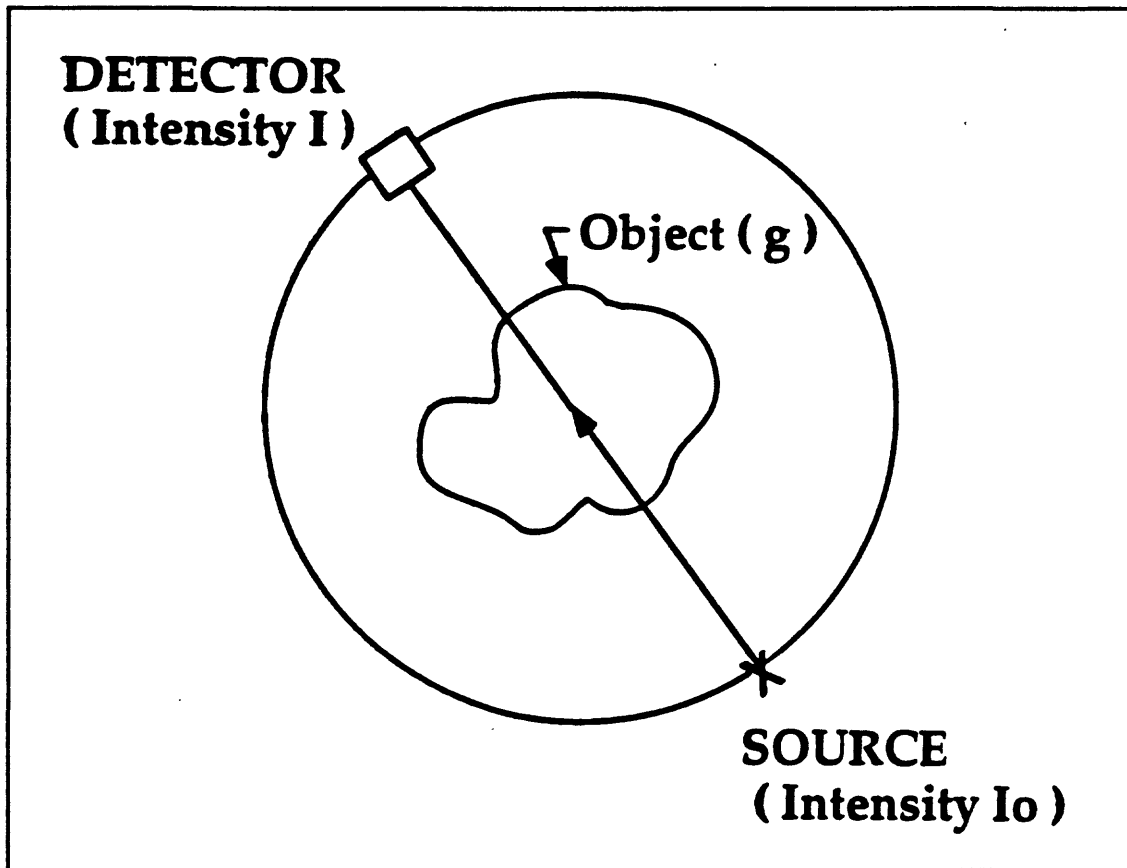


Figure 2 Schematic of the tomography experiment.

an object that is to be imaged which has a linear absorption coefficient $g(r)$. There is also a detector which is sensitive to the amplitude of the received x-ray signal. In medical tomography, the source and detector are rotated around the body being scanned so that a multiple set of paths through the body are sampled. At each orientation of the source-scanner system, the x-rays travel in essentially a straight line through the object and are received by

the detector with intensity I . For any position of the source-detector pair, the intensity of the received signal is given by:

$$I = I_0 e^{-\int_L g(s) ds} \quad (1)$$

The line integral in the exponent is taken over the path L between the source and the receiver and ds is a small increment along the path L . Essentially, the attenuation of the x-rays is controlled by the value of the line integral. Now solve for the line integral to get:

$$f_L = \ln\left(\frac{I_0}{I}\right) = \int_L g(s) ds \quad (2)$$

Since we can measure I_0 and I , we can compute f_L . What we need to be able to do is determine $g(s)$ from measured values of f_L . Cormack observed that the three-dimensional problem could be reduced to this two-dimensional problem by simply treating the body as a sequence of two-dimensional slices. The question was, could $g(s)$ be determined by repeated measurement of f_L at different orientations? It turns out that $g(s)$ can be reconstructed as shown by the central slice theorem.

THE CENTRAL SLICE THEOREM

Consider some object $g(r)$ where r is a position vector defined on an $(x - y)$ coordinate system. $g(r)$ is to be scanned from a number of different directions. For that coordinate system, every scan of the object will be at some angle ϕ from the x axis; and in the scanning reference frame, the object is defined relative to a rotated set of coordinate axes, $(x' - y')$. After the x-rays pass through the object, they are received by the detector and

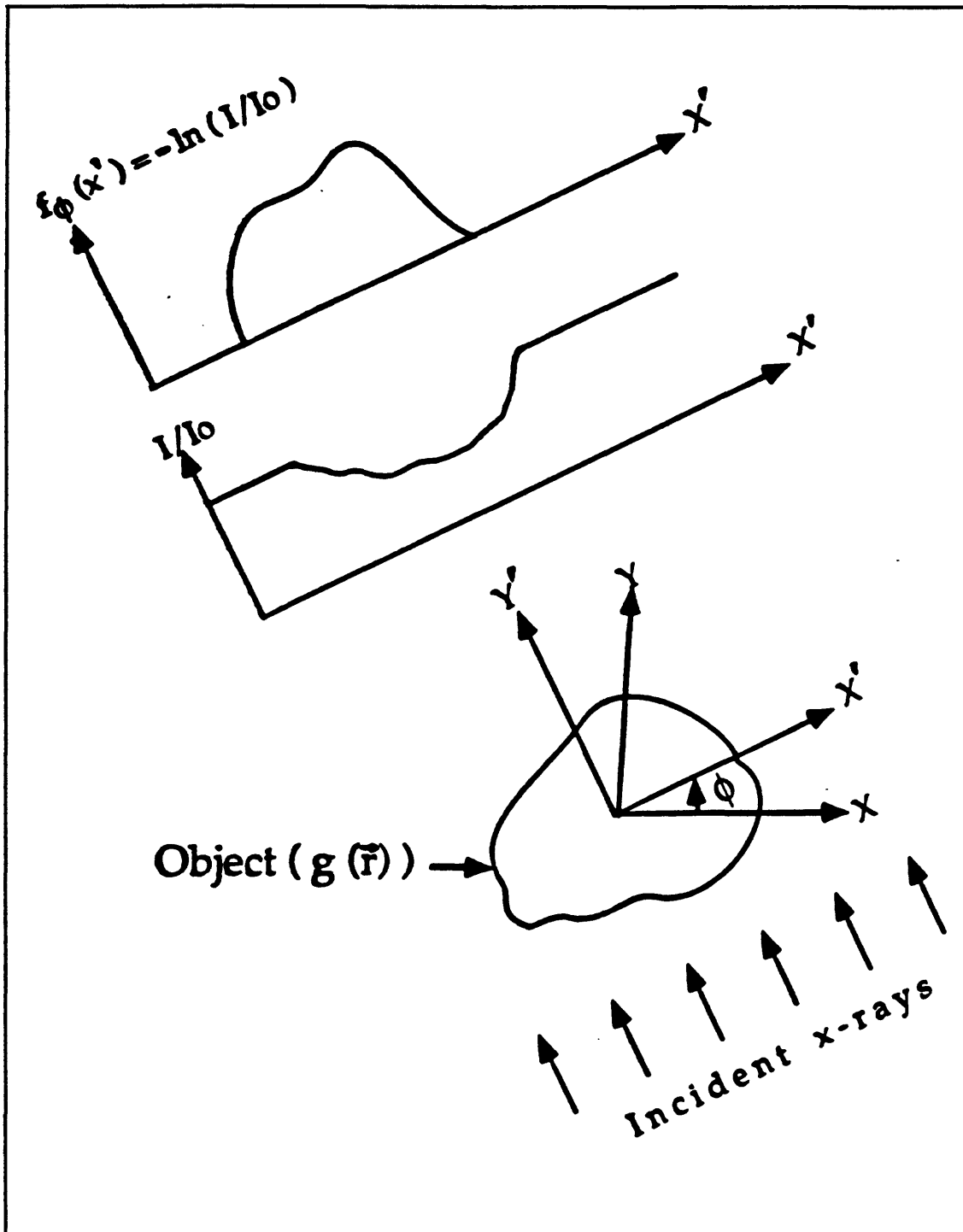


Figure 3 The central slice theorem.

from the measured intensities and the known source intensity, a quantity I/I_0 is computed.

This quantity is a function of the rotated coordinate x' . Another quantity which is equivalent to f_L of Equation (2), but which is in terms of the rotation angle is

$$f_{\phi}(x') = \ln\left(\frac{I}{I_0}\right) = \int_{-\infty}^{+\infty} g(\vec{r}) dy' \quad (3)$$

This is the function shown in the top part of the sketch in Figure 3, and it is referred to as the "projection function." Now define a new function F_1 : the one-dimensional Fourier transform of the projection function. As you know, the Fourier transform allows you to transform a function of time or space into a function of frequency. For this transformation, the spatial variables are x' and y' . The corresponding spatial-frequency variables will be defined as ξ' and η' . In terms of these variables, the transform is:

$$F_1[f_{\phi}(x')] = \int_{-\infty}^{+\infty} f_{\phi}(x') e^{i2\pi\xi'x'} dx' \quad (4)$$

Now substitute the expression for $f_{\phi}(x')$ from Equation (3) to get:

$$F_1[f_{\phi}(x')] = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} g(\vec{r}) dy' \right) e^{i2\pi\xi'x'} dx' \quad (5)$$

But Equation (5) can also be expressed in the form:

$$F_1[f_{\phi}(x')] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\vec{r}) e^{i2\pi(\xi'x' + \eta'y')} dx' \bigg|_{\eta'=0} \quad (6)$$

Which is just the two-dimensional Fourier transform evaluated at $\eta'=0$. Since the Fourier transform is invertible, it is possible to invert the intensity data and arrive at the object function $g(\mathbf{r})$. This is the basis of medical tomography.

COMPARISON OF MEDICAL TOMOGRAPHY TO SEISMIC TOMOGRAPHY

Table 1 Comparison of Medical and Seismic Tomography.

	Radiology	Seismology
Equation:	$f_L = \int_L g(s)ds$	$T = \int_\Gamma u(s)ds$
	wheres s is an element along L	wheres s is an element along Γ
Unknown:	$g(s)$ = absorption coefficient	$u(s)$ = slowness = 1/ velocity
Ray path:	L is straight line	Γ is usually not a straight line
Sources:	at known locations	usually unknown and uncontrollable
Detectors:	many and controllable	few and not optimally located
Funding:	well funded	poorly funded

Medical tomography and seismic tomography share some similarities, but they also have important differences. In radiology, we measure the intensity of received x-rays, which is a function of the absorption coefficient of the travel path. In seismology, we measure a travel time which is a function of the slowness along the raypath. In medical tomography, the path (L) is a straight line. However, in seismology, the raypath is almost never a straight line. Also, in medical tomography, the source and detector locations are known exactly at all times. In seismic tomography, the source locations are unknowns which must be solved for,

while inverting for the velocity structure. A final difference is that a much larger amount of money has been spent on medical tomography than has been spent on seismic tomography. When you throw a tremendous amount of money and talent at some problem, you can probably come up with a solution in a hurry.

Actually, the developments in medical tomography have had little to do with seismic tomography. The people working on seismic tomography were unaware of the advances being made in the medical field. One reason was that the problem was attacked at about the same time in both fields, and since scientists tend not to talk across disciplines, and the subject was too new to have been reported in the popular press, the seismologists working on the problem were just not aware of the existence of related research in medical tomography. However, by the late 1970's some seismologists started reading some of the medical literature and adopted some of the techniques used in medical tomography.

During the early years of the development of seismic tomography, the technique was called "simultaneous inversion." The idea was to get some information on the velocity structure of the earth at the same time that source parameters were being determined. The work I am going to discuss was done using a dense seismic network and only P-wave arrivals.

In the mid-1970's it occurred to a number of seismologists that since a seismic network may have hundreds of stations, and since only 4 stations are required to do a hypocenter inversion, all that redundant information might be put to use to invert for the velocity structure of the earth simultaneously with the hypocenter inversion. Crosson (1976 a,b) worked out the problem for the one-dimensional case. At the same time, Aki and Lee

(1976) published the results of a three-dimensional inversion. A complete description of their method can also be found in chapter 6 of Lee and Stewart (1981).

At about the same time, Aki, Christoffersen and Husebye (1977) worked out the three-dimensional inversion method for velocity parameters using teleseismic events. For the global arrival-time data, Dziewonski, Hager and O'Connell (1977) used the ISC data set to determine the spherical harmonic coefficients of velocity perturbations in the Earth's mantle.

Seismic tomography is not limited to imaging the Earth or part of the Earth using body-wave arrival-time data, it has been extended to include surface waves and waveforms as well (e.g., Geller, 1991; Iyer and Hirahara, 1993). Although inverting for earth's velocity structure is most common, other parameters can also be imaged. For example, spatial variations in seismic attenuation from observed S to P wave amplitude ratios was studied by Ho-Liu, Kanamori and Clayton (1988), and slip velocity by tomographic source imaging of small earthquakes was carried out by Frankel et al. (1986).

In the following discussion, I will concentrate on only one-type of seismic tomography, namely the simultaneous inversion of earthquake hypocentral parameters and velocity structure from local earthquakes. This approach is commonly used in seismology, and the basic idea can be extended for inverting for other parameters.

SIMULTANEOUS INVERSION

The simultaneous inversion problem can be posed in the following manner. There is an array of (m) stations indexed $k=1,2,\dots,m$. The array is used to record (n) earthquakes indexed $j=1,2,\dots,n$. The hypocenter parameters for the jth earthquake are $(t_j^\circ, x_j^\circ, y_j^\circ, z_j^\circ)$. There

are n earthquakes so the complete set of hypocenter parameters is $(t_1^\circ, x_1^\circ, y_1^\circ, z_1^\circ, t_2^\circ, x_2^\circ, y_2^\circ, z_2^\circ, \dots, t_n^\circ, x_n^\circ, y_n^\circ, z_n^\circ)$. There are a total of $4n$ unknown parameters in this vector.

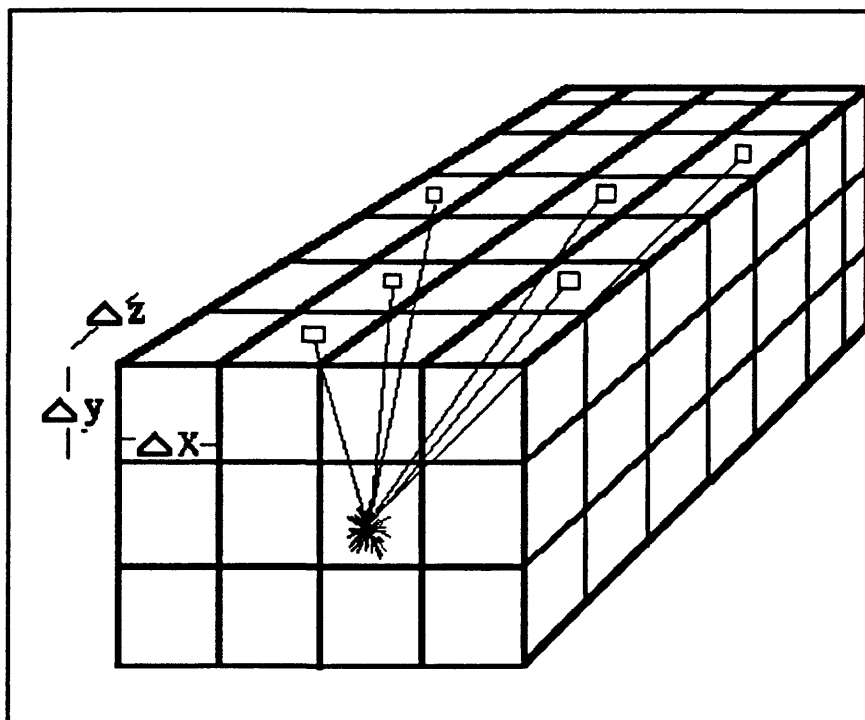


Figure 4 Grid system imposed on problem.

To parameterize the velocity structure, we superimpose a three-dimensional grid on the region of the crust containing the earthquake sources and the seismic stations. Figure 4 shows a schematic sketch of the system. Each of the cells within the grid system is assumed to have constant velocity but the velocity can vary from cell to cell. In practice, the dimensions of each block can be different. For example, around a fault, we would probably want to have a finer grid than at great distance from the fault. However, for the purposes of this discussion, assume the cells are all equi-dimensional. Assume that there are (L) blocks indexed $l=1,2,\dots,L$. The velocity parameters of the problem are indexed (v_1, v_2, \dots, v_L) . The complete set of unknowns for the problem consists of the set of hypocenter parameters

and the set of velocities, so the total number of unknowns for the problem is $4n + L$.

It is mathematically convenient to put all the unknowns in one parameter vector which can be expressed as:

$$\zeta = (t_1^\circ, x_1^\circ, y_1^\circ, z_1^\circ, \dots, t_n^\circ, x_n^\circ, y_n^\circ, z_n^\circ, v_1, \dots, v_L)^T \quad (7)$$

We need to define a trial parameter vector which will be given as:

$$\zeta^* = (t_1^*, x_1^*, y_1^*, z_1^*, \dots, t_n^*, x_n^*, y_n^*, z_n^*, v_1^*, \dots, v_L^*)^T \quad (8)$$

The trial vector is essentially a guess which will be improved during the solution process.

For any trial vector, the residual associated with the j th earthquake and the k th station is:

$$r_{jk}(\zeta^*) = T_{jk} - T_{jk}(\zeta^*) - t_j^* \quad (9)$$

Here T_{jk} is the measured arrival time for the j th earthquake at the k th station. $T_{jk}(\zeta^*)$ is the theoretical travel time, and t_j^* is the theoretical origin time. There are (mn) of these residuals and in order to have a solution, we must have (mn) much larger than $(4n+L)$. Given these conditions, the problem is to minimize:

$$F(\zeta^*) = \sum_{j=1}^n \sum_{k=1}^m [r_{jk}(\zeta^*)]^2 \quad (10)$$

Equation (10) is just the sum of the squared residuals over all the stations and all the earthquakes for (mn) observations.

This is a generalization of the earthquake location problem. Recall that in the earthquake location problem, we are given a set of m stations with m corresponding arrival times, and we want to find the "best" origin time and hypocenter location, assuming a velocity model. We make a starting guess for the desired parameters, and through an

iterative procedure, we improve the guess until it meets some stopping criterion. In simultaneous inversion, we do this but we also solve for the velocity model at the same time. Just as in the earthquake location, we define a residual vector:

$$\mathbf{r} = (r_{11}, r_{12}, \dots, r_{1m} \quad \dots \quad r_{n1}, r_{n2}, \dots, r_{nm})^T \quad (11)$$

Then we define the adjustment vector:

$$\delta \zeta = (\delta t_1, \delta x_1, \delta y_1, \delta z_1, \dots, \delta t_n, \delta x_n, \delta y_n, \delta z_n, \delta v_1, \dots, \delta v_L)^T \quad (12)$$

At each step of the process we must solve a set of simultaneous equations for the adjustment vector in order to improve the estimate of the parameter vector. In matrix notation, the system of equations is:

$$\mathbf{B} \delta \zeta = -\mathbf{r} \quad (13)$$

Here \mathbf{B} is the Jacobian matrix generalized to include the n earthquakes and L velocity parameters, i.e.,

$$\mathbf{B} = \left(\begin{array}{cccc|c} \mathbf{A}_1 & 0 & 0 & \dots & 0 \\ 0 & \mathbf{A}_2 & 0 & \dots & 0 \\ & & \vdots & & \\ 0 & 0 & 0 & \dots & \mathbf{A}_n \end{array} \right) \mathbf{C} \quad (14)$$

Here the 0 's are $(m \text{ by } 4)$ matrices with zero elements and the \mathbf{A}_j 's are $(m \text{ by } 4)$ Jacobian matrices of the form:

$$\mathbf{A}_j = - \begin{pmatrix} 1 & \frac{\partial T_{j1}}{\partial x} & \frac{\partial T_{j1}}{\partial y} & \frac{\partial T_{j1}}{\partial z} \\ 1 & \frac{\partial T_{j2}}{\partial x} & \frac{\partial T_{j2}}{\partial y} & \frac{\partial T_{j2}}{\partial z} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{\partial T_{jm}}{\partial x} & \frac{\partial T_{jm}}{\partial y} & \frac{\partial T_{jm}}{\partial z} \end{pmatrix} \bigg|_{\zeta^*} \quad (15)$$

The **C** element of matrix **B** is an (mn by L) matrix where each row of the matrix has L elements and describes the sampling of the velocity blocks of a particular raypath. If Γ_k is the raypath connecting the kth station to the jth earthquake hypocenter, then the elements of the ith row [$i = k + (j-1)m$] of **C** are given by:

$$C_{il} = -\Pi_{jkl} \left(\frac{\partial T_{jk}}{\partial v_l} \right) \bigg|_{\zeta^*} \quad \text{for } l = 1, 2, \dots, L \quad (16)$$

where Π_{jkl} is defined by

$$\Pi_{jkl} = \begin{cases} 1 & \text{if } l\text{th block penetrated} \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

It is advantageous to work with slowness rather than with the velocity, so we define the slowness in the lth block is defined as $u_l = 1/v_l$. From this, we have:

$$\delta u_l = -\frac{\delta v_l}{v_l^2} \quad (18)$$

To approximate the travel time derivative with respect to velocity to first-order accuracy, we make the approximation:

$$\frac{\partial T_{jk}}{\partial u_l} = \frac{\partial T_{jkl}}{\partial u_l} \quad (19)$$

Where $T_{jk l}$ is the travel time for the ray Γ_{jk} in the l th block. Also, $T_{jk l} = u_l S_{jk l}$, where $S_{jk l}$ is the length of the raypath Γ_{jk} in the l th block. Assuming that the dependence of $S_{jk l}$ on u_l is of second order,

$$\frac{\partial T_{jk l}}{\partial u_l} = S_{jk l} = v_l T_{jk l} \quad (20)$$

Thus Equation (16) becomes:

$$C_{j l} = \prod_{j k l} \frac{T_{jk l}(\zeta^*)}{v_l^*} \quad \text{for } l = 1, 2, \dots, L \quad (21)$$

Since each ray path samples only a small number of blocks, most elements of matrix **C** are zero. Therefore, the matrix **B** is sparse, with most of its elements being zero. Equation (13) is a set of (mn) equations written in matrix form which may be rewritten as:

$$\delta t_j + \frac{\partial T_{jk}}{\partial x} \Big|_{\zeta^*} \delta x_j + \frac{\partial T_{jk}}{\partial y} \Big|_{\zeta^*} \delta y_j + \frac{\partial T_{jk}}{\partial z} \Big|_{\zeta^*} \delta z_j - \sum_{l=1}^L \frac{\prod_{j k l} T_{jk l}}{v_l^*} \delta v_l = r_{jk}(\zeta^*)$$

for $k = 1, 2, \dots, m, \quad j = 1, 2, \dots, n$

(22)

Equation (22) is a generalization of the equation used for hypocenter determination. The only difference is that the simultaneous inversion equation has the summation term which is an approximation to the sums of partial derivatives of the travel times with respect to velocity in each block.

To summarize the solution, first we guess a trial parameter vector. Essentially, one must have some estimates of where all the earthquakes are and have some initial approximation for the velocity. Next, calculate the theoretical travel times and the spatial derivatives for the trial vector (all stations and all earthquakes). This information is used to

construct the **B** matrix. Also, compute the residual vector, which is essentially the difference between the observed travel times and the theoretical travel times. Then solve a system of (n by m) linear equations. This gives the adjustment vector, which is added to the trial vector to produce an updated trial vector. This is repeated until the solution converges to some pre-defined limit or is terminated after a pre-

Numerical Solution of the Simultaneous Inversion Problem

- 1) Guess a trial parameter vector ζ'' as given by equation 8
- 2) Compute the theoretical travel time T_{jk} and its spatial partial derivatives
- 3) Compute matrix **B** as given by equations 14, 15, and 16 and compute the **r** vector as given by equation 9.
- 4) Solve the system of linear equations as given by equation 22 for the adjustment vector $\delta\zeta$ using the normal equations approach or generalized inversion.
- 5) replace the trial parameter vector ζ'' with $(\zeta'' + \delta\zeta)$
- 6) repeat steps 2 through 5 as necessary.

defined number of iterations. The process is analogous to hypocenter location, but instead of doing it for just one earthquake, use many earthquakes, and instead of using an assumed velocity model, solve for that as well.

The central idea of seismic tomography is not that complicated. Yet, it took nearly two decades to develop the method to its current state. It all started with the realization that local networks provided far more data than was required to locate earthquakes. From that point, the method has been refined, and now there are many mathematical variations in the implementation of this simultaneous inversion algorithm. However, all of them share the essential theoretical formulation of the problem that I have outlined here.

It was not my intention to try to cover all details of the process in this brief discussion of seismic tomography. Instead, I just hope to have given you a basic

understanding of how the method works. If you have an interest in working with seismic tomography, there are many excellent references available as given in the next section on "Suggested Reading".

SUGGESTED READING

Because this chapter is intended as an introduction, I would like to recommend further reading for the interested reader. Chapter 5 of Lee and Stewart (1981) deals with generalized inversion and nonlinear optimization and refers to classic texts, such as Lanczos (1956, 1961), Forsythe, Malcolm and Moler (1977), and Gill, Murray and Wright (1981). For readers not familiar with these mathematical topics, a careful study of the above mentioned references and more recent texts (e.g., Menke, 1984; Tarantola, 1987) will be of great help. A collection of classic papers in inversion of geophysical data may be found in Lines (1988).

For seismic tomography, a general review by Thurber and Aki (1987) is a good start. In addition, Claerbout (1985), Nolet (1987), and Iyer and Hirahara (1993) treat this subject in some detail. For medical tomography, Swindell and Barrett (1977) is a good introduction. In addition, there are classic papers by Cormack (1963) and by Hounsfield (1973), and a good text book by Herman (1980).

Aki, K., A. Christoffersen, and E. S. Husebye (1977). Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.*, **82**, 277-296.

Aki, K., and W. H. K. Lee (1976). Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes, Part 1. A homogeneous initial model. *J. Geophys. Res.*, **81**, 4381-4399.

- Claerbout, J. F. (1985). "Imaging the Earth's Interior", Blackwell Scientific Publications, Oxford, 398 pp.
- Cormack, A. M. (1963). Representation of a function by its line integrals, with some radiological applications, *J. Appl. Phys.*, **34**, 2722-2727.
- Crosson, R. S. (1976a). Crustal structure modeling of earthquake data, Part 1. Simultaneous least squares estimation of hypocenter and velocity parameters, *J. Geophys. Res.*, **81**, 3036-3046.
- Crosson, R. S. (1976b). Crustal structure modeling of earthquake data, Part 2. Velocity structures of the Puget Sound region, Washington, *J. Geophys. Res.*, **81**, 3036-3046.
- Dziewonski, A. M., B. H. Hager, and R. J. O'Connell (1977). Large-scale heterogeneities in the lower mantle, *J. Geophys. Res.*, **82**, 239-255.
- Forsythe, G. E., M. A. Malcolm, and C. B. Moler (1977). "Computer Methods for Mathematical Computations", Prentice-Hall, Englewood Cliffs.
- Frankel, A., J. Fletcher, F. Vernon, L. Haar, J. Berger, T. Hanks, and J. Brune (1986). Rupture characteristics and tomographic source imaging of earthquakes near Anza, southern California, *J. Geophys. Res.*, **91**, 12633-12650.
- Geller, R. J. (1991). Supercomputers in seismology: determining 3-D Earth structure, *SIAM News*, **24**, no. 5.
- Gill, P. E., W. Murray, and M. H. Wright (1981). "Practical Optimization", Academic Press, London, 401 pp.
- Herman, G. T. (1980). "Image Reconstruction from Projections", Academic Press, New York, 316 pp.
- Ho-Liu, P., H. Kanamori, and R. W. Clayton (1988). Applications of attenuation tomography to Imperial Valley and Coso-Indian Wells region, southern California, *J. Geophys. Res.*, **93**, 10501-10520.
- Houndsfield, G. N. (1973). Computerized transverse axial scanning tomography: Part I. Description of the system, *Br. J. Radiol.*, **46**, 1016-1022.
- Iyer, H. M., and K. Hirahara, (Editors) (1992). "Seismic Tomography: Theory and Practice", in press.
- Lanczos, C. (1956). "Applied Analysis", Prentice-Hall, Englewood Cliffs, 539 pp.

- Lanczos, C. (1961). "Linear Differential Operators", Van Nostrand, London, 564 pp.**
- Lee, W. H. K., and S. W. Stewart, (1981). "Principles and Applications of Microearthquake Networks", Advances in Geophysics, Supplement 2, Academic Press, New York, 293 pp.**
- Lee, W. H. K. and V. Pereyra (1992). Mathematical introduction to seismic tomography. In press.**
- Lines, L. R. (Editor), (1988). "Inversion of Geophysical Data", Geophysics Reprint Series, No. 9, Society of Exploration Geophysicists, 543 pp.**
- Menke, W. (1984). "Geophysical Data Analysis: Discrete Inverse Theory", Academic Press, Orlando, 260 pp.**
- Nolet, G. (Editor), (1987). "Seismic Tomography", D. Reidel Publishing Co., Dordrecht, 386 pp.**
- Swindell, W., and H. H. Barrett, (1977). Computerized tomography: taking sectional x rays, Phys. Today, 32-41.**
- Tarantola, A. (1987). "Inverse Problem Theory", Elsevier, Amsterdam, 613 pp.**
- Thurber, C. H., and K. Aki, (1987). Three-dimensional seismic imaging, Ann. Rev. Earth Planet. Sci., 15, 115-139.**

21. APPLICATIONS OF SEISMIC TOMOGRAPHY

by

H. M. Iyer
U. S. Geological Survey, Menlo Park, CA 94025

WHAT IS TOMOGRAPHY

Tomography is a technique for producing a three-dimensional image of an object by repeatedly sampling, in different orientations, some property of the object such as absorption of x-rays in the medical case, or slowness (reciprocal of velocity) in the seismic case. The technique was first developed in the medical field where it was used to produce images of tumors and other abnormalities in soft tissue.

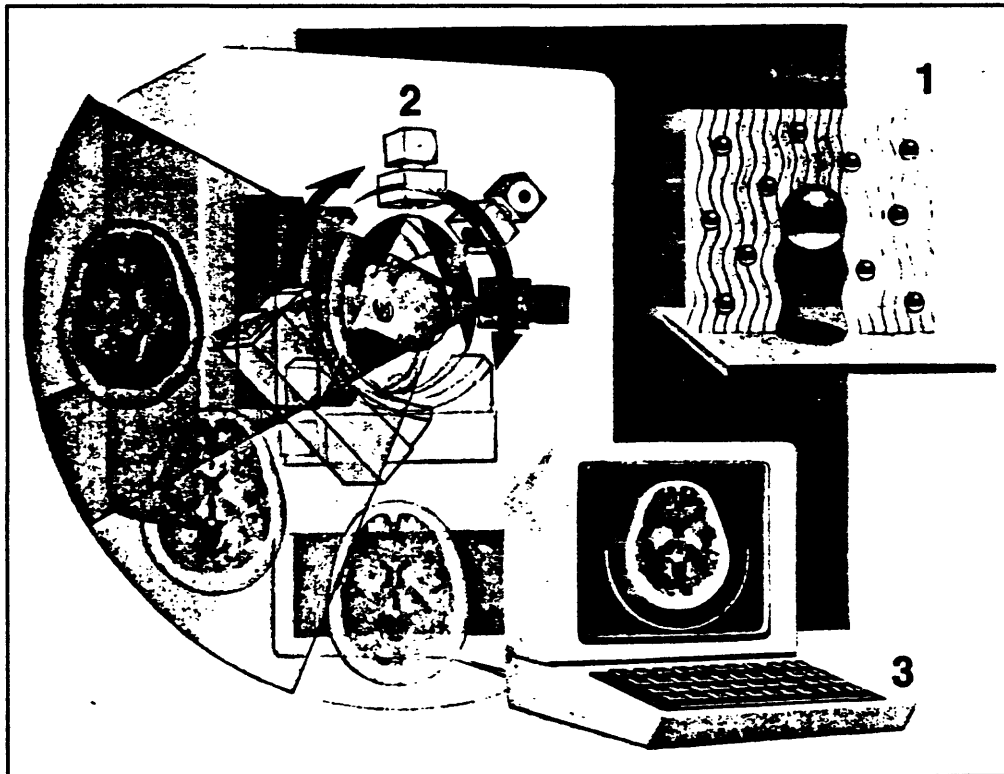


Figure 1 Illustration of medical tomography.

Figure 1 (National Geographic Magazine, January 1987) is a schematic illustration of the medical tomography procedure. Adaptation of medical tomography to seismic imaging has provided a powerful new tool in seismology.

The money and time that has been invested in medical tomography has resulted in systems which can produce really spectacular three-dimensional images of the interior of the human body. We have not come anywhere close to that in seismic tomography. Partly, this is a matter of funding and partly it is due to the different nature of the two processes. Figure 2 shows a schematic comparison of the two techniques. In medical tomography (CT Scan) the

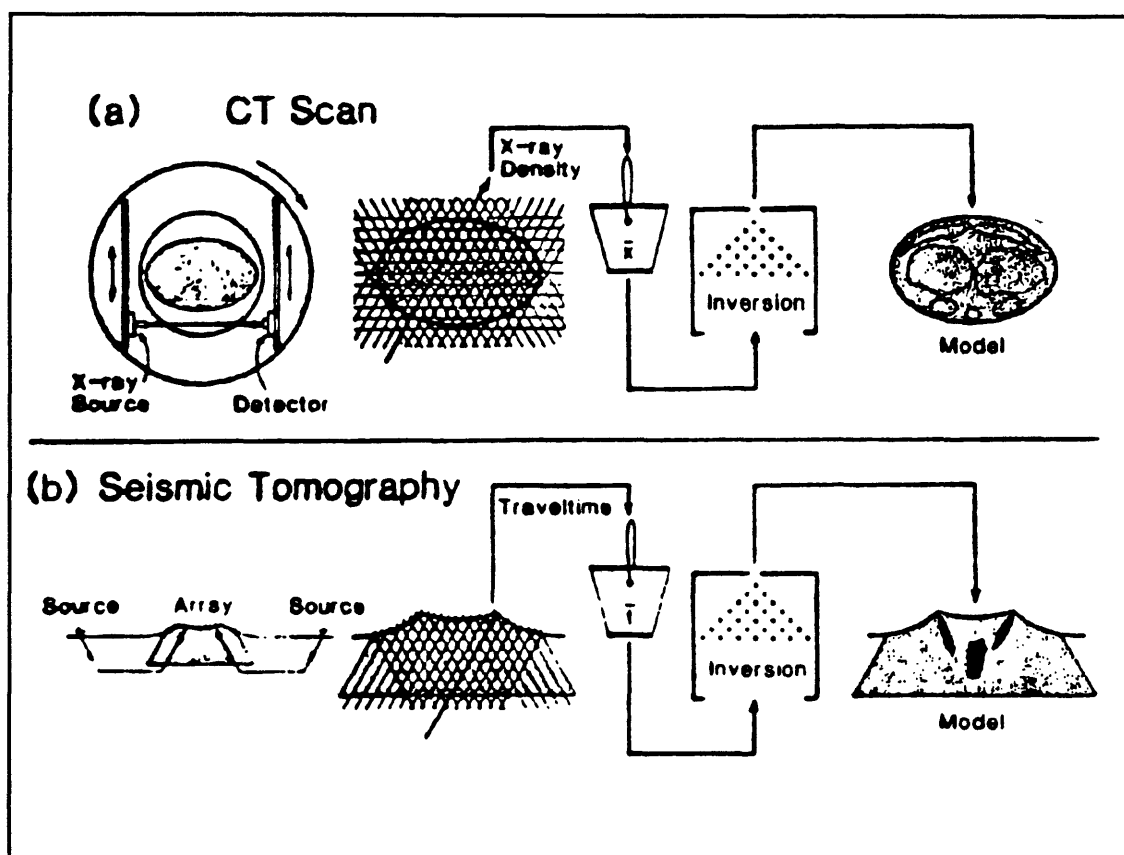


Figure 2 Comparison of medical and seismic tomography.

source-receiver pairs are at known positions. In seismic tomography, although we know exactly where the receivers are, the location of the source is one of the unknown quantities

except when explosions are used as sources, as shown later. Also, in medical tomography, the source-receiver pairs are rotated around the object being scanned so that the object is sampled from all possible orientations (within the limits of sampling). Conversely, in seismic tomography, although we have some control over where we place receivers, the sources tend to exist only in certain locations. Thus the volume we wish to image is incompletely sampled.

Around the late 1970's when Lee and Aki were working on local tomography, Aki was also working with a group in Norway on a means of using seismic waves from teleseisms to invert for a model of the earth's velocity structure on a larger scale. Also, a group from Harvard was working on the problem of global seismic tomography.

The object of global tomography is to use body waves and surface waves to image global-tectonic features such as oceanic rifts, subduction zones, and so on. The resolution of the technique is on the order of 500 to 1000 kilometers. Soon, very similar techniques were applied on a regional scale to view large-scale continental and oceanic heterogeneities. Despite the limitations of the seismic tomographic method, it has been able to produce some spectacular images of the earth's velocity structure. Figure 3 is a tomographic map of global shear wave velocity perturbations. The top panel shows velocity perturbations at a depth of 150 kilometers. The white lines are plate boundaries. The bottom panel is a cross section (see middle panel) through the upper mantle from the Moho at a depth of 26 kilometers to a depth of 670 kilometers along the meridian 110°W . All the continents have high velocities, while almost all the low velocity regions follow the oceanic ridges. The cross section through Antarctica, North America, and Asia, shows clearly that the high velocity regions under the

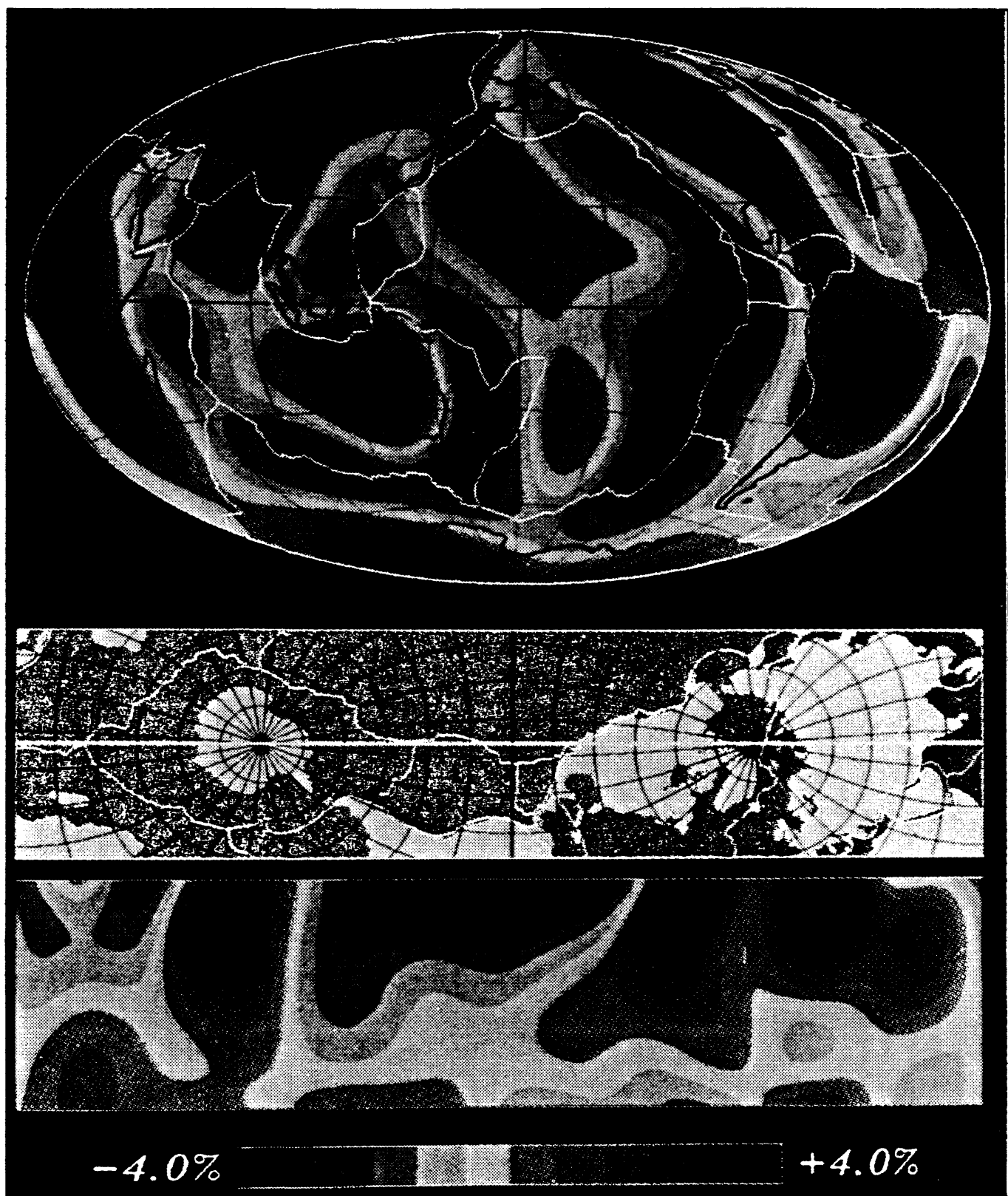


Figure 3 Relative perturbations in upper mantle shear velocities
(Taken from Anderson and Dziewonski, 1984).

continent extend to great depths into the mantle. This is the kind of information provided by global tomographic studies.

TOMOGRAPHY EXPERIMENTS AT THE USGS

Interesting and exciting though these studies are, I will not spend any more time on this aspect of tomography. Rather, I want to talk about some work we at the USGS have done on a more regional scale using the technique which Aki and his Norwegian colleagues developed.

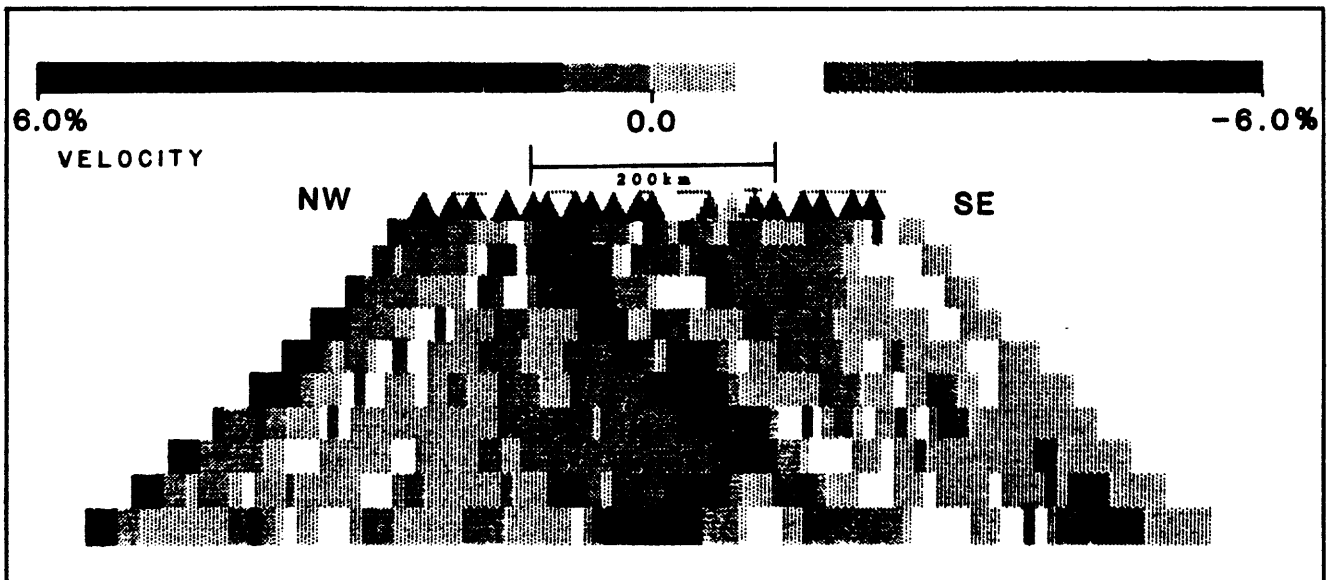


Figure 4 Tomographic cross section under Crater Lake (Taken from Harris et al., 1991).

Figure 4 shows one of the tomographic models we made of the Juan de Fuca subduction zone near Crater Lake in Oregon. We operated a 300 kilometer long linear seismic array along the coast of Oregon and this tomographic model is one of the results of our modeling. Normally, one images a structure like a subducting slab by mapping the seismicity associated with the

structure. In this experiment, we produced the image by tomographic means.

Another powerful tool in seismic tomography, first developed by Aki and Lee, is the use of seismic waves from local earthquakes. Seismology has advanced to the point where we are no longer content to simply locate earthquakes using simple one-dimensional velocity models. The size and density of seismometer arrays and the processing power of computers have increased so much that we are now able to extract a great deal of information from the events that we monitor. We are able to invert for the velocity structure of the seismic zones we monitor, and relocate the earthquakes with that information. When we can see the earthquake locations superimposed on velocity models derived from tomographic modeling, we can begin to see how tectonics and structure influence the generation of earthquakes on some fault. Tomographic models of 3-d crustal structure using local earthquakes can be useful low cost by-products of any PC-based seismic network operation which might be deployed for monitoring fault zones, volcanoes, geothermal areas, or dams.

Why should we do teleseismic tomography? The difference between local earthquake tomography and teleseismic tomography is that in local earthquake tomography we are using the earthquakes within the model volume and we are using fairly high frequency waves so we resolve structures with dimensions on the order of a kilometer or so. In teleseismic tomography, we are using low frequency waves from distant sources, and our spatial resolution is much lower than that obtainable with local scale tomography. However, teleseismic tomography is able to give us information about the regional features within which our local model is embedded, and therefore by combining the two analyses we can gain a better understanding of the regional tectonics than would be possible using only one or

the other of the two techniques. Since there is now a teleseismic trigger which operates on a PC system it is reasonable to expect that one can do this sort of combined analysis using a relatively inexpensive PC-based system.

There is another high resolution tomography which is different from the methods I have discussed so far in that it does not rely on any natural source of seismic energy. One sets up a dense seismic array surrounding the region of interest and then uses explosions as the seismic sources. In principle, the technique can yield spatial resolutions of the order of tens of meters. However it is a very expensive procedure. At the USGS we have used this procedure twice in the study of magma chambers of volcanoes. These two high-resolution imaging experiments provided us with a spatial resolution of about one Kilometer. However, there is no reason that the experiment cannot be scaled down so as to look at very fine structures such as potential engineering or archeological excavations. Then it becomes much cheaper.

In the remainder of my talk I am going to concentrate on the above three topics. These are local earthquake tomography, teleseismic tomography, and high resolution tomography. The mathematics of tomography has already been covered in the previous chapter. What I want to do is discuss the practical results we get from doing seismic tomography experiments. It is good to understand the mathematics, but now there are a number of programs available which can do the inversion for you.

LOCAL EARTHQUAKE TOMOGRAPHY

The expert in local-earthquake tomography at the USGS is Donna Eberhart-Phillips.

In fact, she should be giving this talk! Unfortunately, she is at Pasadena today. So, I will talk a little about some work she did using local earthquake tomography.

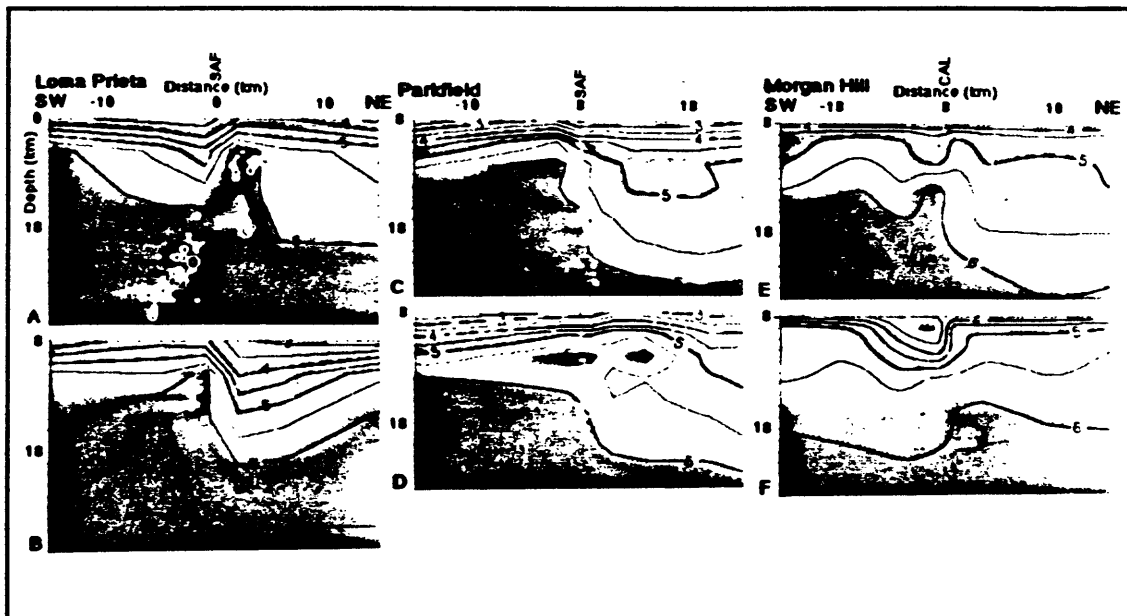


Figure 5 Velocity anomalies along three earthquake rupture zones (Taken from Michael and Eberhart-Phillips, 1991).

Figure 5 is an illustration of the kind of results one can get using local earthquake tomography. The left two frames are cross-sections taken from southwest to northeast over a distance of 30 kilometers at three earthquake source areas. The white patches are locations of aftershocks of the Loma Prieta earthquake. Superimposed on the aftershock map is a set of velocity contours derived from tomographic analysis. These show a high velocity anomaly in the region where the earthquake has occurred. What is it? Is it an asperity? It would not be possible to detect the presence of this feature without the tomographic analysis. At both Parkfield (middle) and Morgan Hill (right) similar velocity anomalies associated with the earthquake rupture region have been imaged. I don't want to go into any more detail on what these velocity anomalies may mean. You can read the paper if you are interested. The point I

want to emphasize is that tomography was the means by which these features were discovered. This is the kind of information that is going to enhance our detailed understanding of seismic zones. Figure 6 is an illustration of the type of ray path coverage used to illuminate a source region which is to be modeled. Obviously, the region where the

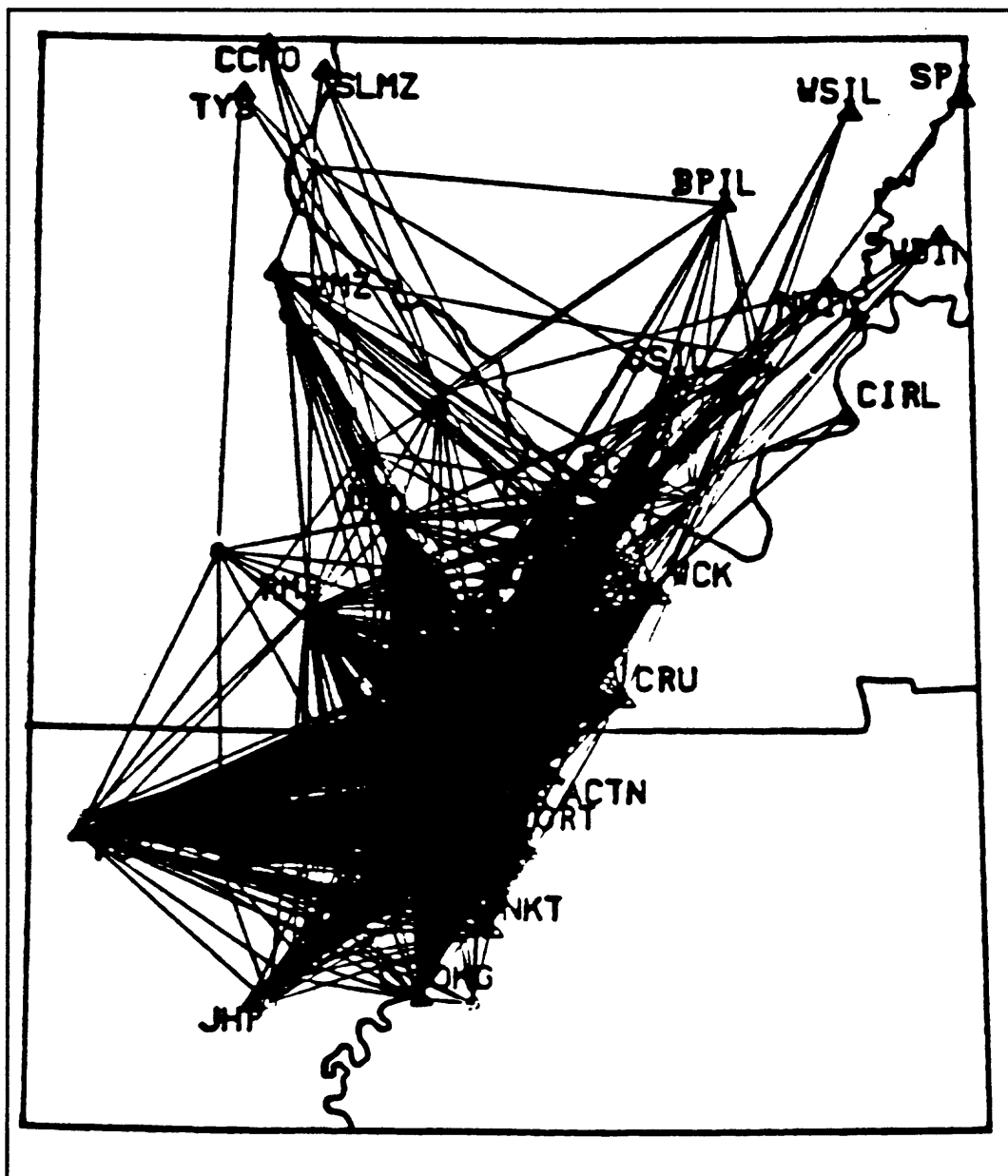


Figure 6 Ray path coverage for typical tomography experiment.

ray coverage is the most dense is the region in which you can expect the most stable inversion.

What I want to do now is talk about the tomographic modeling Donna did using aftershocks from the 1983 Coalinga earthquake. I will not attempt to show you all the details of the inversion. Rather, I just want to present a "cook book" approach to seismic tomography. If you have a seismic network and if the seismicity is high in the area, for example an aftershock sequence, and you collect a lot of data from this region, what type of information you can get and how do you go about processing the data to obtain the information?

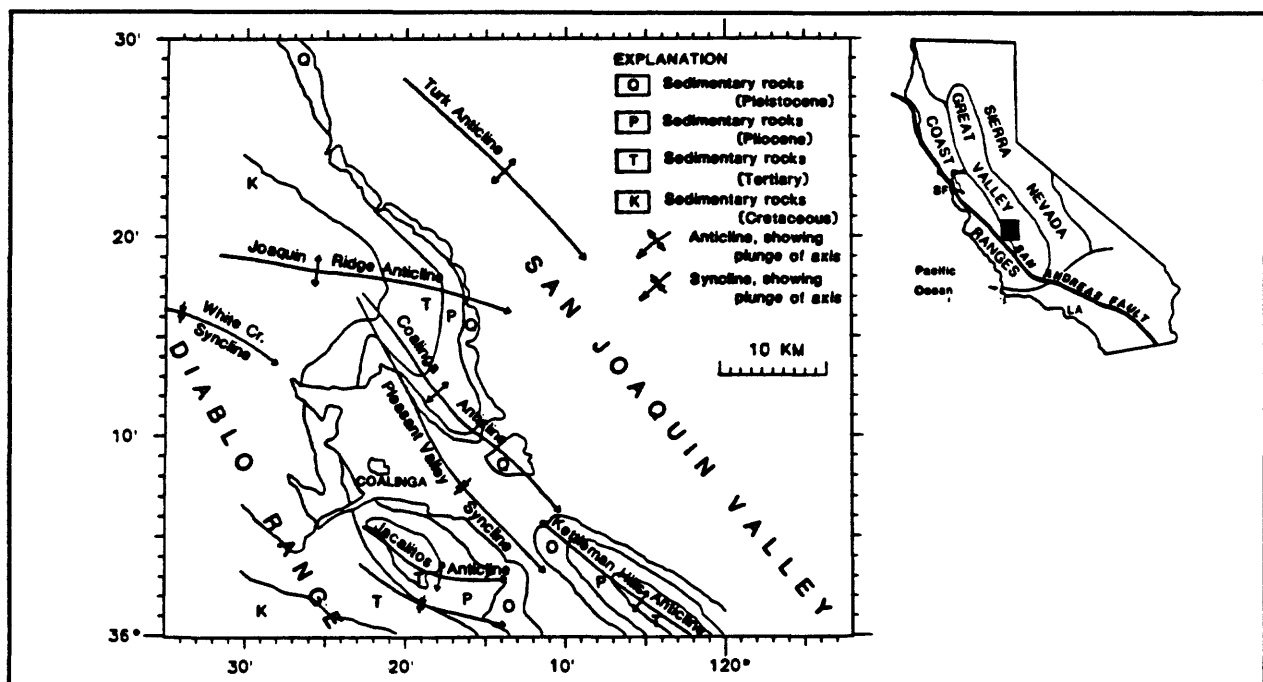


Figure 7 Base map of California with Coalinga detail (Taken from Eberhart-Phillips, 1990).

Figure 7 shows in detail the Coalinga area. The Coalinga earthquake of 1983 was not a strike-slip type earthquake. Rather, it was a thrust-faulting event. It occurred in a thrusting,

anticlinal type of environment. There were lots of aftershocks and two years after the earthquake had occurred, there was an earthquake in the Kettleman Hills anticline just to the south-east of the source region for the 1983 earthquake. In her study, Donna modeled the aftershock region between the Coalinga anticline and the Kettleman Hills anticline.

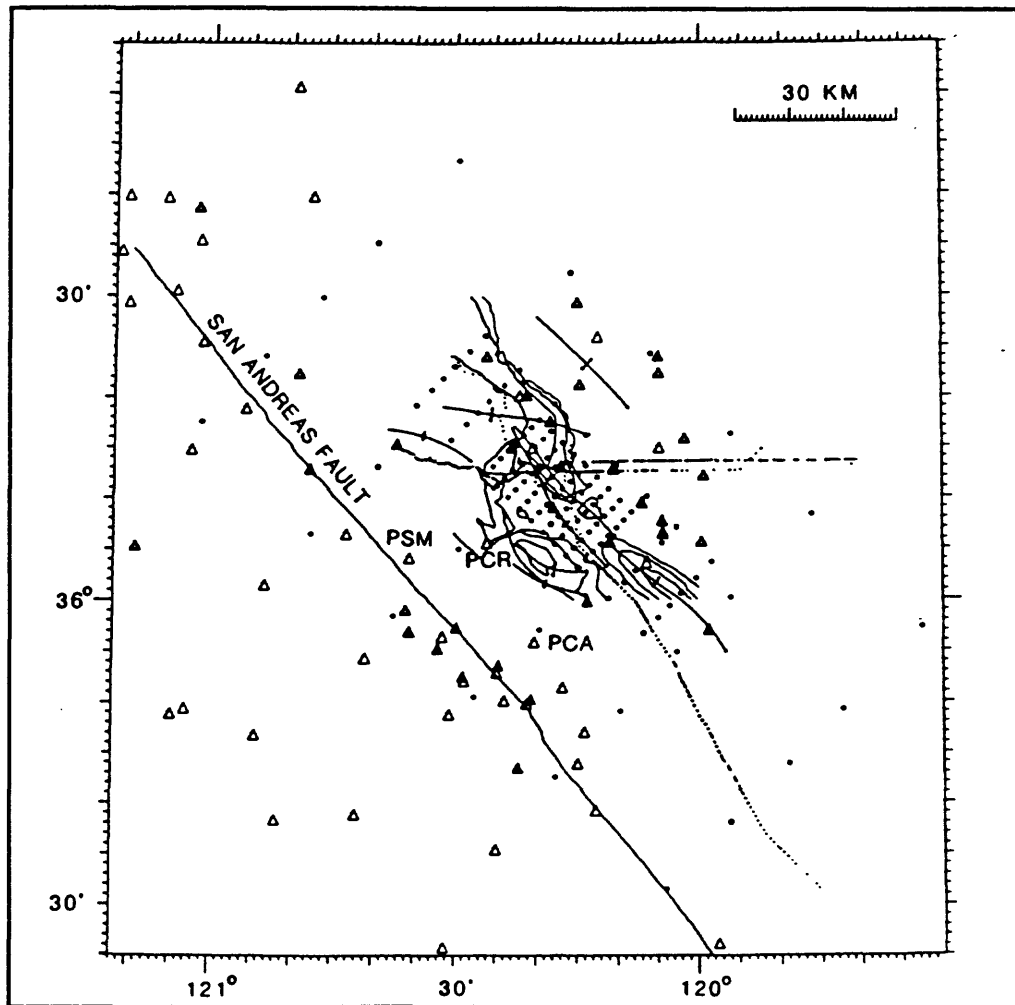


Figure 8 Distribution of seismographs shown on geologic base map (Taken from Eberhart-Phillips, 1990).

Figure 8 shows a map of stations superimposed on a geologic base map for the Coalinga experiment. The open triangles represent Calnet stations used in the experiment and

closed triangles are temporary three-component seismometers deployed for the experiment. The closely spaced dots represent grid points of the model and the rows of coarsely spaced dots running roughly SE-NW are points on the refraction array. The model was divided into a number of small cells with vertices at the grid points and each cell has individually variable velocity. During the inversion, ray paths were initially assumed to be straight and as cell

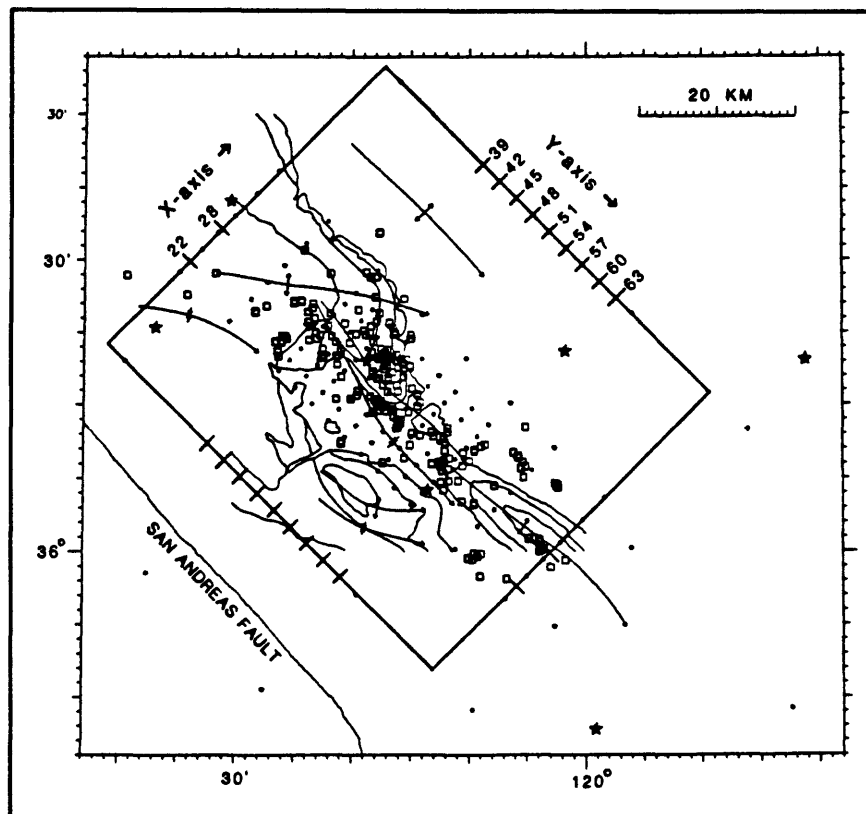


Figure 9 Earthquake data used in inversion
(Taken from Eberhart-Phillips, 1990).

velocities were adjusted to reduce residuals, the ray paths could become quite complex. The program used in this experiment enabled fixing the velocity of any cell. In this way information, either from geological observation or, perhaps, from the refraction survey could be incorporated in the model. However, Donna found that starting with a complicated model

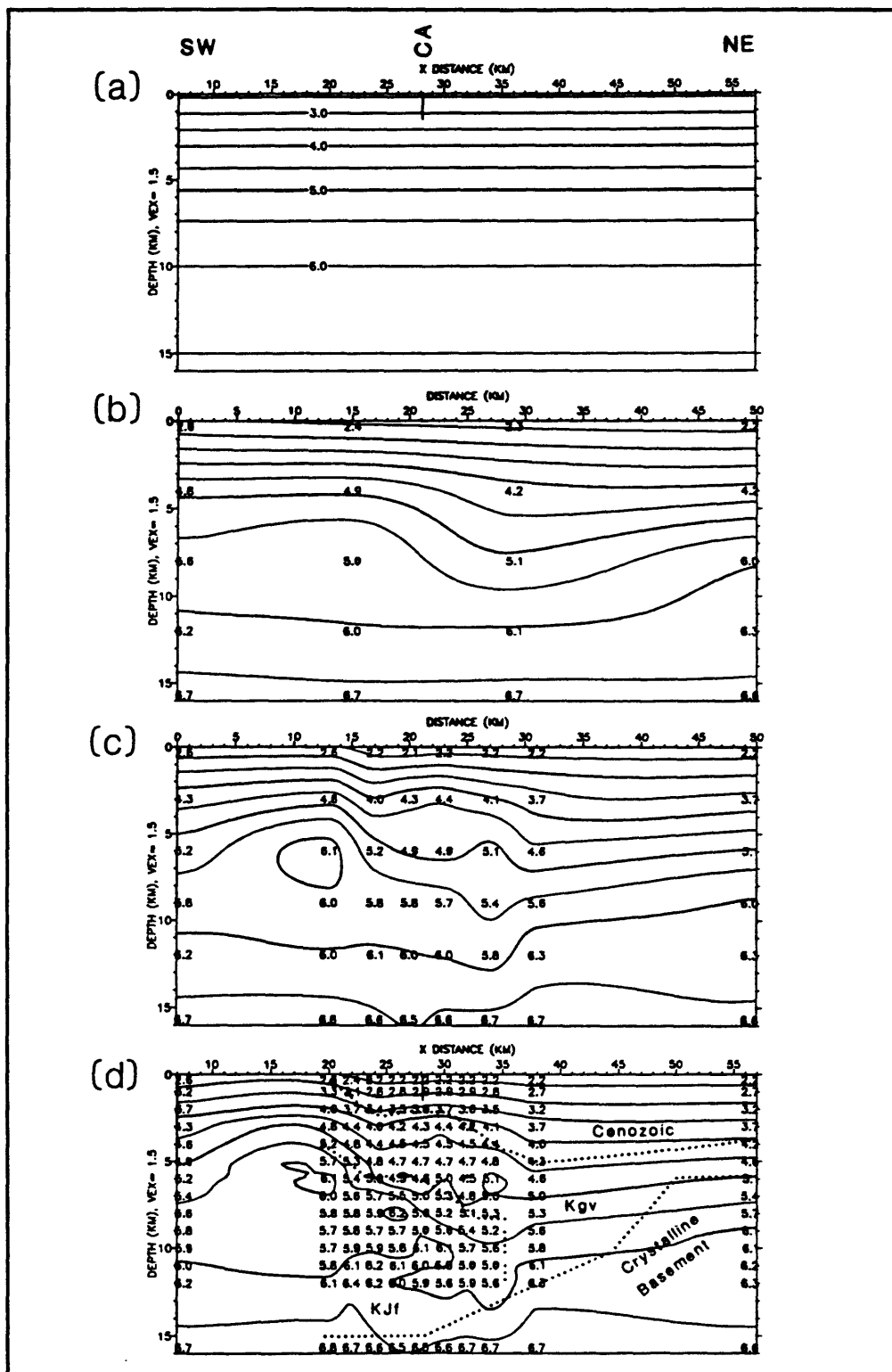


Figure 10 Results of inversion with simple starting model (Taken from Eberhart-Phillips, 1990).

actually produced poorer results than starting with a simple model.

Figure 9 shows the earthquake data used in the inversion. This figure is an enlarged view of Figure 8. Here, the open squares are earthquakes and the stars are refraction shots. The rectangular box shows the orientation of the coordinate system used for the model.

Figure 10 Shows a series of velocity cross sections perpendicular to the Coalinga

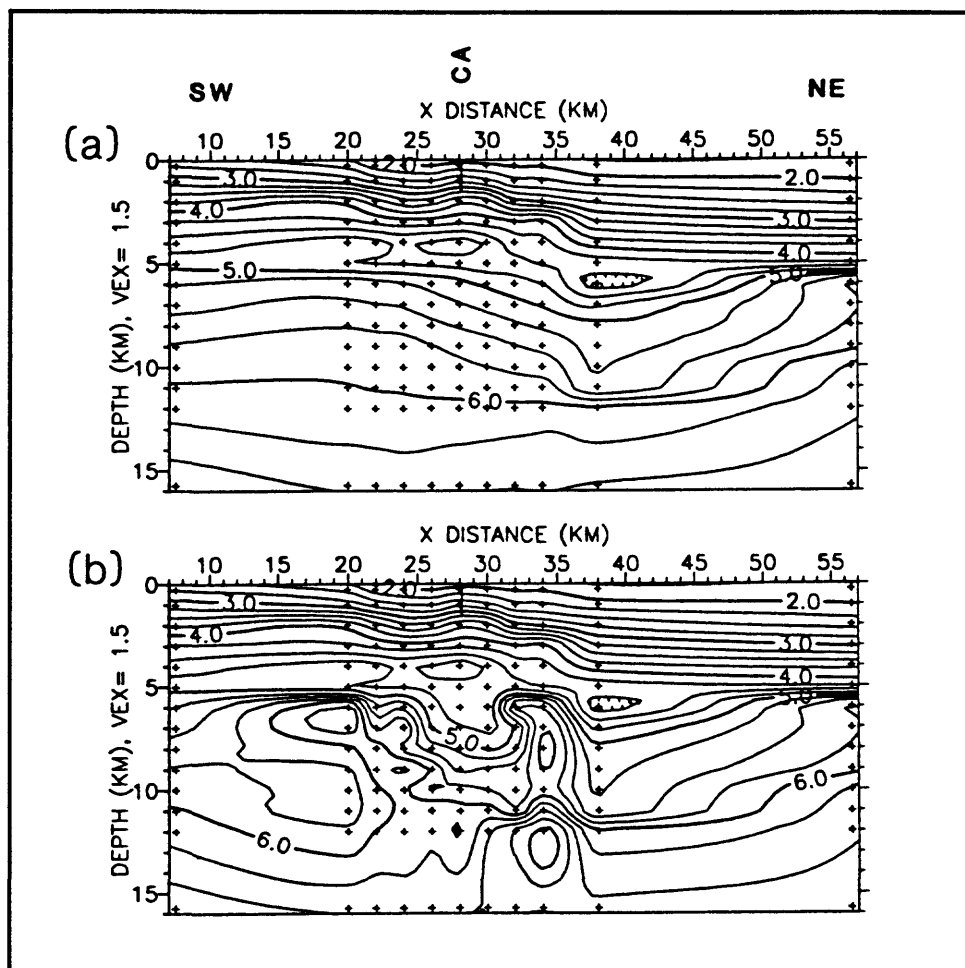


Figure 11 Inversion solution using refraction model as initial model (From Eberhart-Phillips, 1990).

anticline obtained at different stages of the inversion process. In part (a) a simple one-

dimensional velocity model is being used. The solution obtained is essentially a HYPO71 solution. Models (b) through (c) were obtained using progressively finer grid spacings. With each decrease in grid size the residual decreases and the model structure increases in complexity. Later, I will discuss the believability of these structures.

The results shown in Figure 10 were produced using a simple one-dimensional velocity model as the starting point. It is interesting to see the effect of using a more complex model as the starting point for inversion. Figure 11 shows an initial velocity model (a) produced using the seismic refraction data, and the resulting inversion solution is shown in (b). As you can see, the initial model is rather complex. The velocity solution variance associated with this model was twice as much as that associated with the solution with the simple starting model. In that sense, the inversion might be considered a success. However, the geologic structure suggested by this model is unrealistically complex and Donna decided to reject the model despite its lower residual. Starting the inversion process with a complex model actually seems to de-stabilize the inversion process. Simple initial assumptions seem to result in a more stable inversion process.

No matter what you try to do in seismology, if you have only P-waves to work with, you don't get very far. It is much better to have both P-waves and S-waves to work with. When both P-wave and S-wave velocities are modeled in some region, it becomes possible to deduce some of the material properties of the region. One can get an idea what sort of rock is in the region and whether it is relatively intact or whether it is extensively fractured. Figure 12 shows the results of some work Donna did in which she used both P-wave and S-wave velocities to produce models of the same region. You all know how hard it is to work

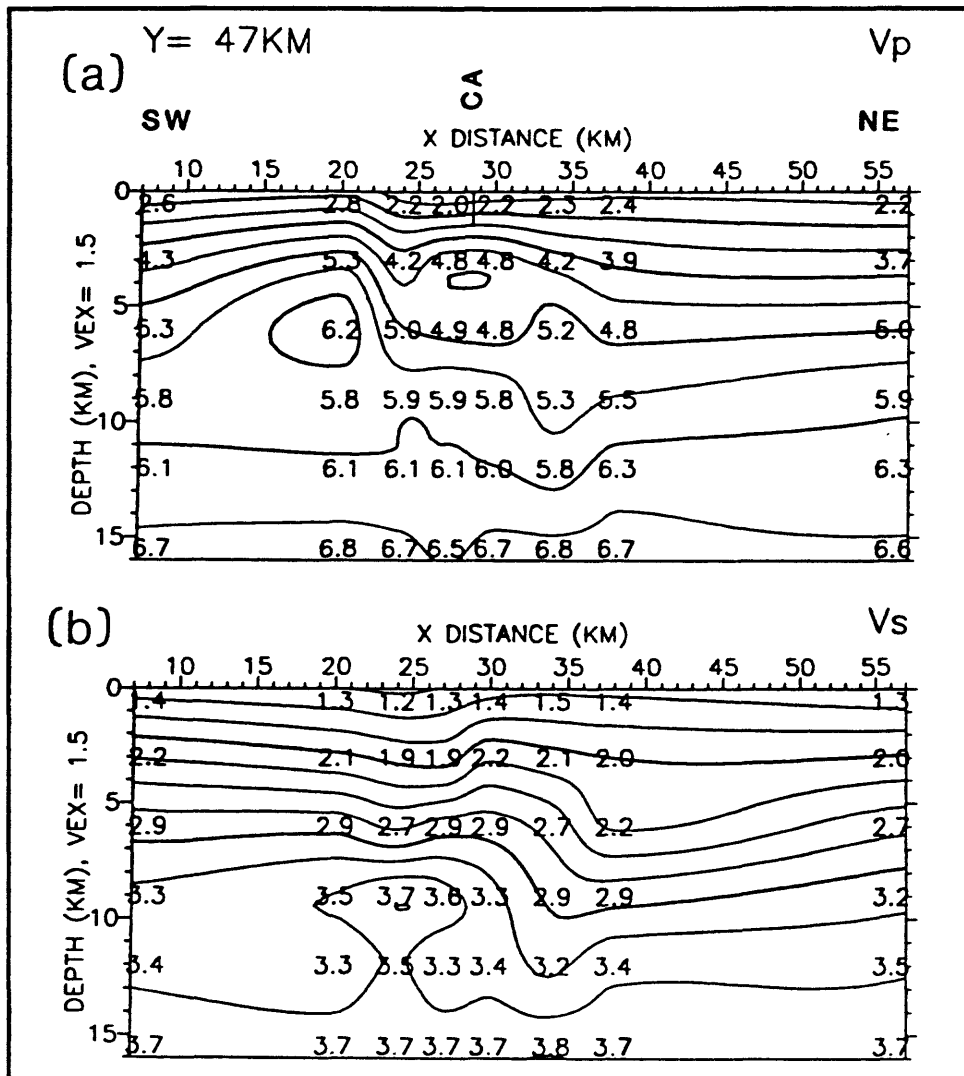


Figure 12 Cross section across Coalinga anticline using both P and S velocities in inversion (Taken from Eberhart-Phillips, 1990).

with S-waves because of the difficulty of picking the S-wave onset reliably. That was a problem here. In all cases where the S-wave onset was felt to be ambiguous, the observation was not used. Consequently, there are fewer observations and the model is less well constrained. The way the inversion program works is that at any grid point where the solution does not converge, the initial value of velocity is returned as the final answer. If

there are many grid points with this problem, a lot of false curvature shows up in the velocity cross sections. This was a problem with the S-wave models. In Figure 12, the upper panel shows the result of P-wave modeling and the lower panel shows the result of S-wave modeling. There is considerable difference between the models within the finely gridded portion of the model.

When we do tomographic inversion, the problem we have to live with is that we are taking a time measurement for some ray and distributing that measurement to different parts of the ray. There is bound to be some kind of smearing in this process. The smearing is usually along the ray path and is much more obvious in teleseisms when we know what direction the rays are coming from than is the case working with local earthquakes. One of the ways the inversion program tries to control the smearing is through a damping factor. If the solutions generated with the program are very unstable, you can sometimes improve things by adjusting the value of the damping factor.

One of the things we need to know about any model we produce by tomographic inversion is how believable are the points in the model. The model residual gives some information, but we can get a much better idea by looking at the resolution matrix. Figure 13 shows plots of parts of the resolution matrix for four different points. In these plots, the tighter the contours, the less smearing there is from block to block. The top 4 frames ((a) through (d)) show the results for the P-wave inversion and the lower 4 frames show the results for the same 4 points from the S-wave inversion. Although the P-wave results are acceptable for all the points, the 2 points shown in the left hand frames are badly smeared in the S-wave inversion.

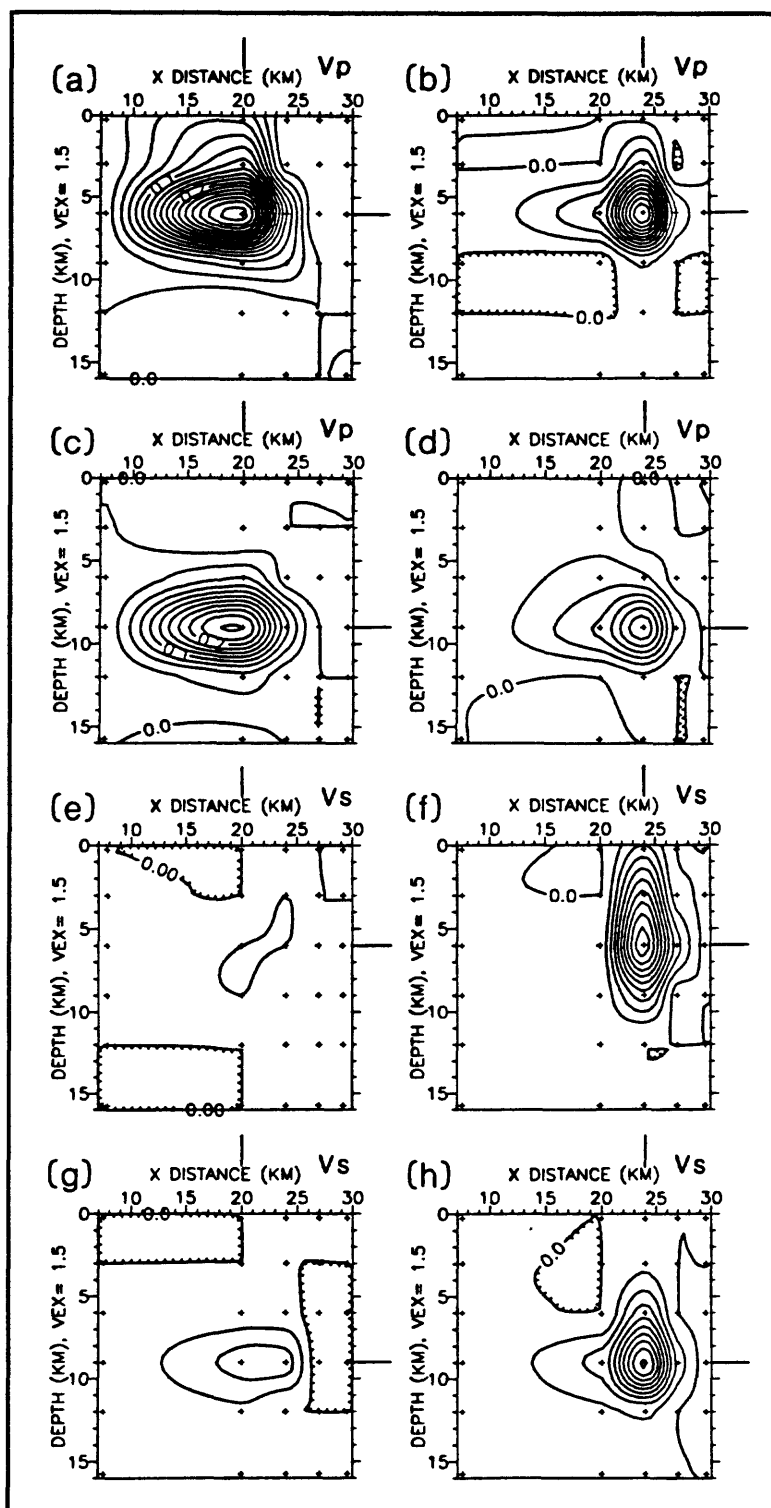


Figure 13 Plots of portions of the resolution matrix (Taken from Eberhart-Phillips, 1990).

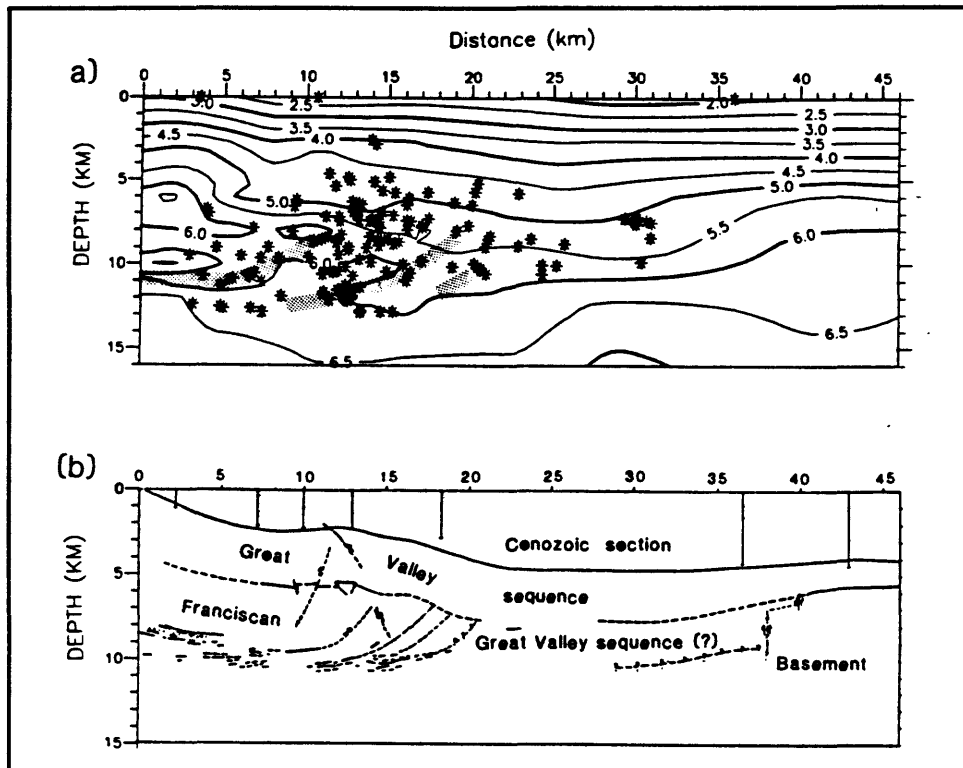


Figure 14 Comparison of inversion solution with reflection record (Taken from Eberhart-Phillips, 1990).

Part (a) of Figure 14 shows a cross section through the aftershock zone with the velocity solution and the aftershock locations superimposed and part (b) shows a reflection record for the same cross section. The main feature I want to point out here is the close correspondence between the region of structural complexity on the left hand side of the figure, and the region in which the tomographic model shows a highly complex velocity structure. These also match up with the aftershock zone quite well.

A similar study was also performed in the Loma Prieta aftershock region. Figure 15 shows an index map of the region. The mainshock is indicated by the large octagon, seismometers by the open triangles, and aftershocks used in the inversion by +’s. The large

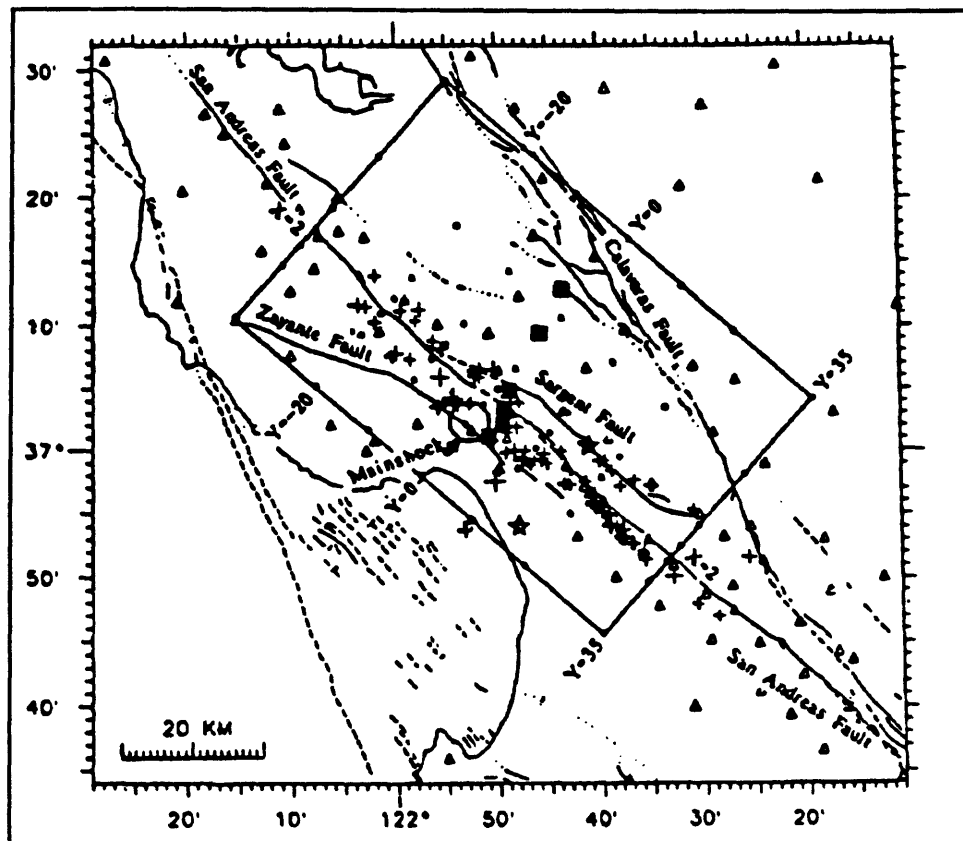


Figure 15 Index map of Loma Prieta epicentral region (Taken from Eberhart-Phillips et al., 1990).

rectangle is the outline of the region for which a velocity model was produced by the inversion.

Figure 16 shows the initial model used for the 3-D velocity inversion and three slices through the final model at the points $Y=-20$, $Y=0$, and $Y=35$. These points are marked on the index map of Figure 15. Notice the pronounced velocity anomaly between the San Andreas and Sargent faults in part (c) of the figure.

Figure 17 shows the results of a resistivity survey done in the same location by one of our colleagues. The high resistivity section of this plot very nearly coincides with the high

velocity anomaly from the tomographic model.

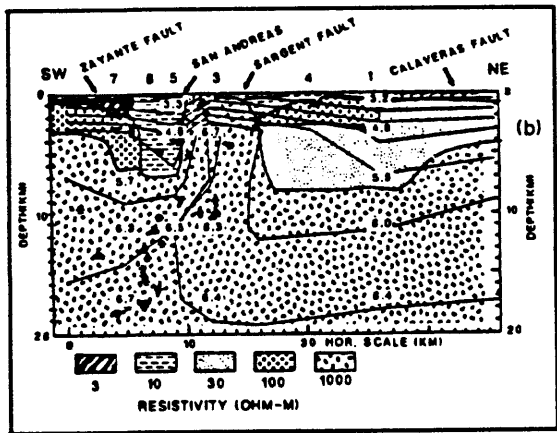


Figure 17 Resistivity profile in Loma Prieta source region.

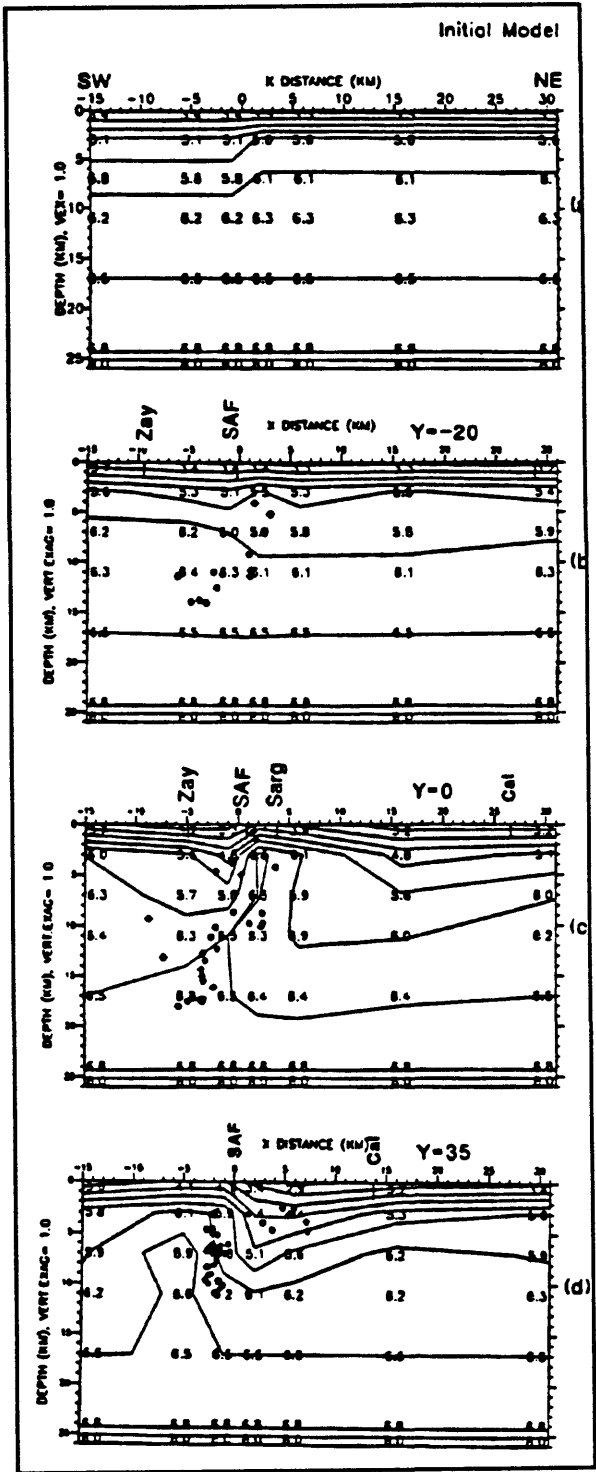


Figure 16 Initial model and three slices through model.

TELESEISMIC TOMOGRAPHY

So far we have been talking about situations in which the seismic sources originate within the model volume. In teleseismic tomography, the earthquakes originate at great distances from the model volume. The rays travel through the mantle and probably through the core as well. The seismic rays come into our model region obliquely and essentially parallel to each other.

Usually, the signal to noise ratio for teleseismic signals is much less than in local earthquake signals and so our task is complicated by the difficulty of accurately picking phase arrivals. However, we have some very nice interactive programs which allow us to manipulate the signals in ways which make it easier to make accurate phase picks. In our work we typically use some standard earth model to predict what the velocity of the waves should be immediately underneath the model. Then we solve for the velocity perturbations in the model which give the measured arrival times. This is somewhat easier than trying to solve for absolute velocity throughout the model.

I want to quickly give you a flavor of teleseismic tomography using, as an example, a recent study of the Long Valley caldera. Figure 18 shows a geologic map and station index map for the study area. In 1976, I and a co-worker did some very simple teleseismic experiments which led us to believe that there might be a magma chamber under the caldera. Then, in the late 1970's some very large earthquakes started happening in the Long Valley area and since we already knew there was a magma chamber in the region, we were afraid that there might be a volcanic eruption there. Although an eruption never took place, since that time, it has become a very popular place to investigate. In Figure 18, the region

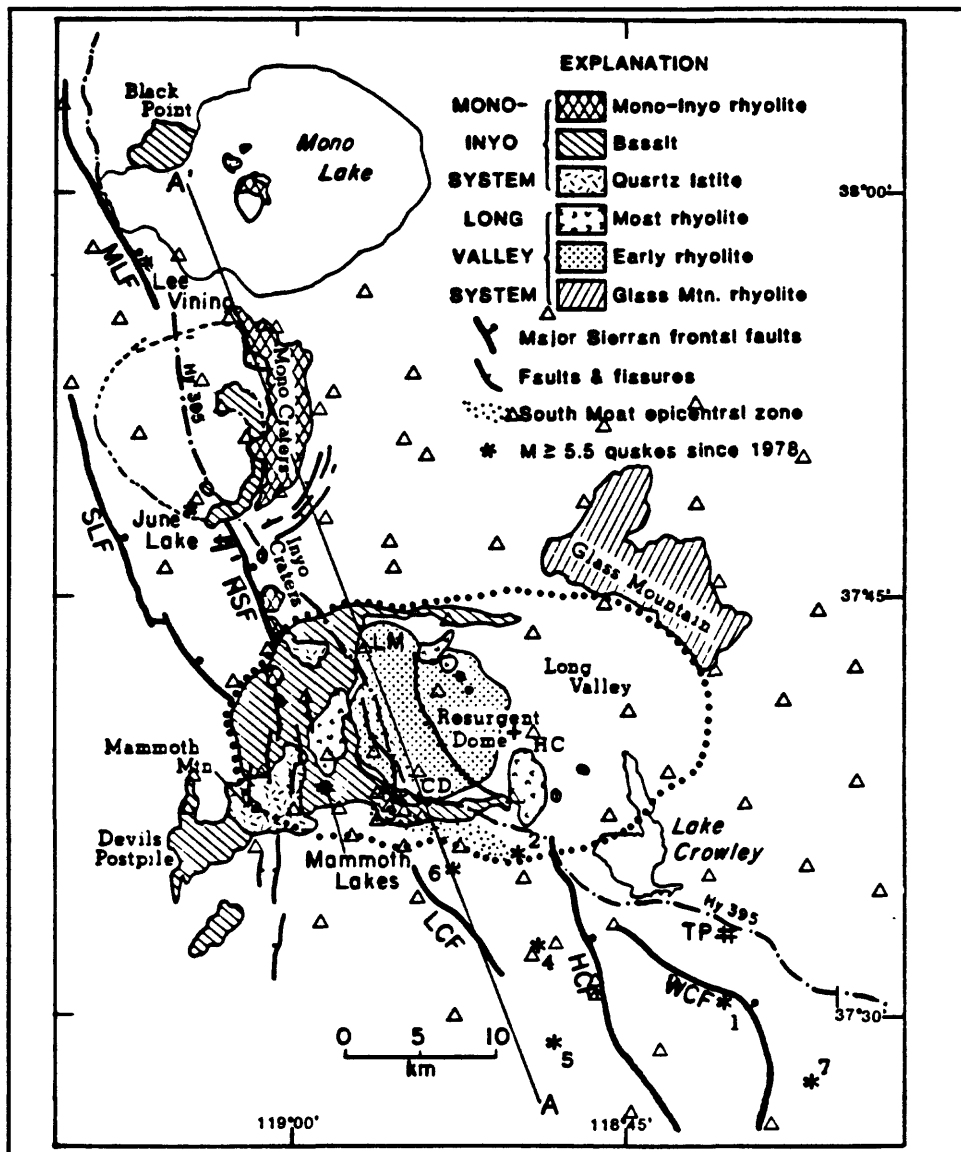


Figure 18 Geologic index map of the Long Valley region (From Dawson et al., 1990).

encircled by dots is the Long Valley caldera. It is about 25 kilometers in diameter and is the remnant of a much larger volcano which erupted violently about 700,000 years ago. The explosion created the caldera and sometime after, a resurgent dome was emplaced inside the caldera. Just to the northwest of the caldera is Mono Craters, a younger volcanic feature.

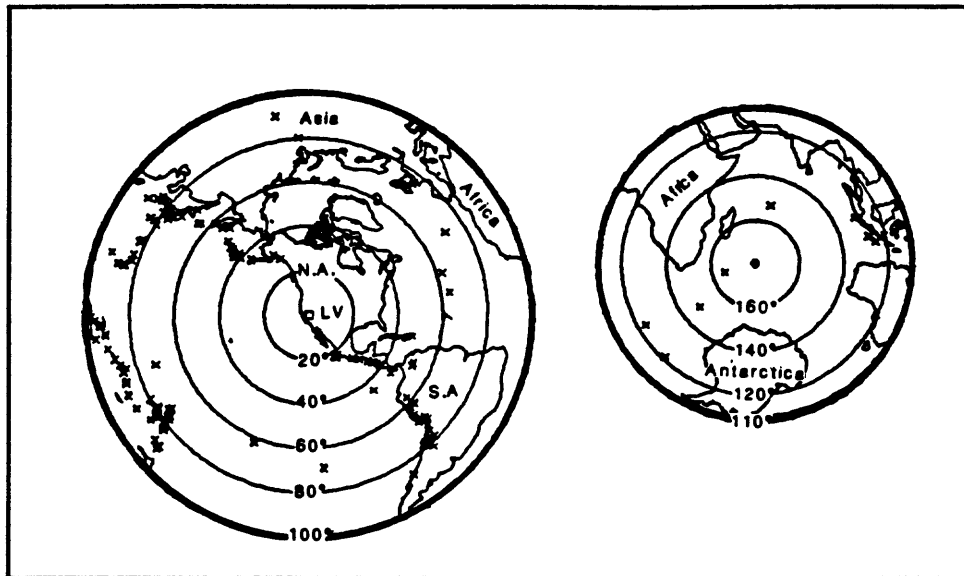


Figure 19 Location of teleseisms used in Long Valley experiment.

Figure 19 shows the location of teleseisms used in the Long Valley experiment. In the figure, Long Valley is represented by the square in the center of the left-hand azimuthal equidistant map. Earthquake epicenters are indicated by crosses. The right-hand map is an antipodal view of the epicenter map and shows events which arrived as PKIKP phases.

Figure 20 (left panels) show the velocity anomalies that were revealed by the tomographic modeling. The first layer is shown at the top, and shows velocity anomalies of up to -5.4. This low velocity region coincides with the Long Valley caldera. A similar low velocity region shows up in the Mono Craters region as well. Figure 20 (left panels) also show the velocity structure for the second and third layers which together extend from a depth of 8.02 kilometers to a depth of 28 kilometers. Note the pronounced, low velocity region which coincides with the Long Valley caldera. The magnitude of the low velocity

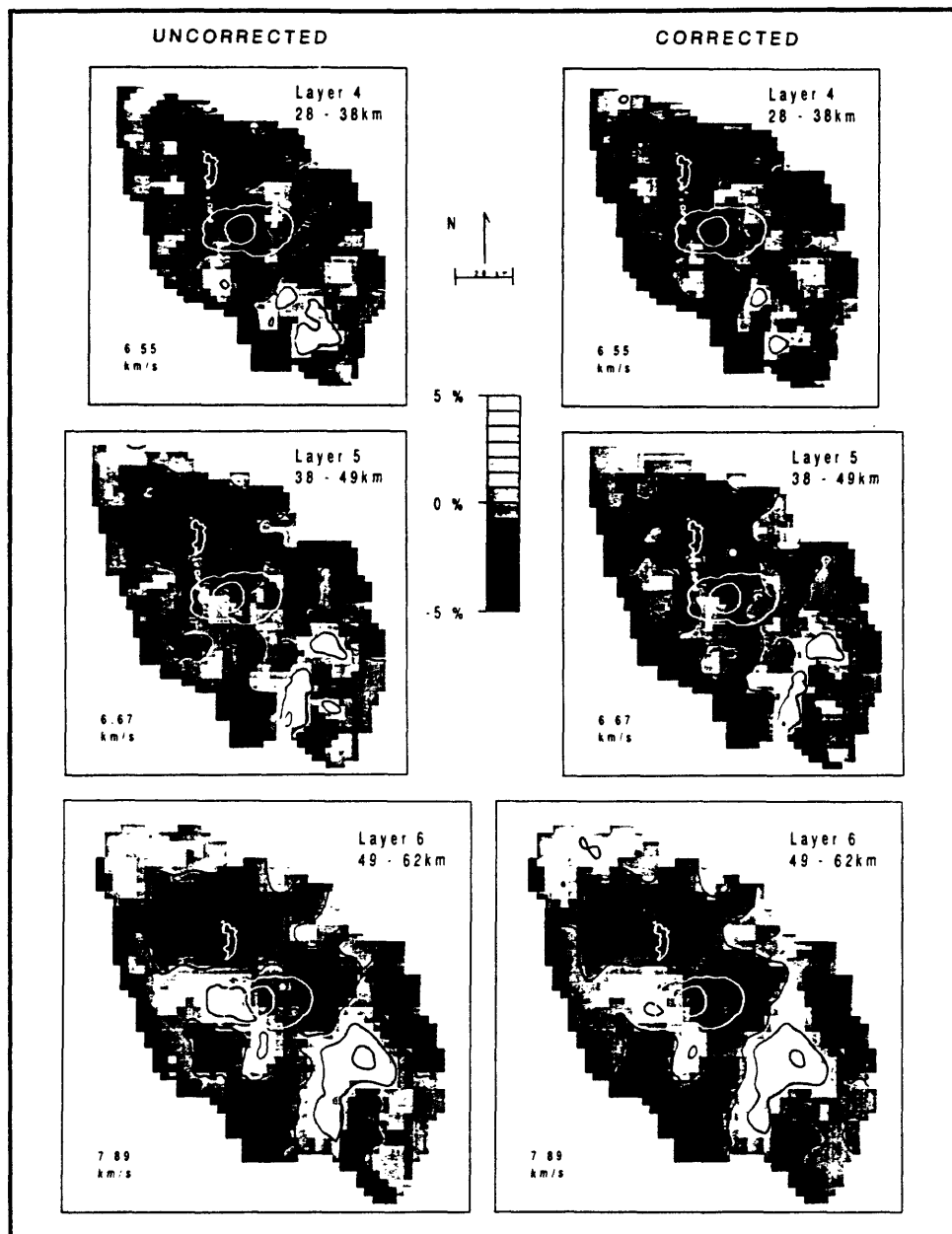


Figure 20 Long Valley and Mono Craters velocity anomalies. Left panels: uncorrected for surface layer. Right panels: corrected for surface layers (From Dawson et al., 1990).

anomaly underneath the Mono Craters area is smaller at this depth than it is near the surface. However, it is still quite evident in this plot. In the next layer down, the Mono

Craters velocity anomaly nearly disappears and the Long Valley anomaly weakens although it is still quite evident.

The preliminary results of this study shown above made some geologists very uncomfortable. The Large low velocity zone under the Long Valley caldera implies that there is a very large magma chamber under the caldera. They did not expect to see a magma chamber as large as the caldera. The general understanding, supported by many previous studies, was that the magma chamber should have been confined to the western half of the caldera. We discovered that there was a problem with our method. Seismic waves traveling through the low velocity surface layers at an angle were "smearing" the apparent low velocity over an area that was significantly larger than the true area of the low velocity region associated with the magma chamber. This is something that one has to be careful of in any tomographic study. To correct the problem, we used information from seismic refraction studies in the region to make a model of the surface velocity over the magma chamber. Using this model, we were able to correct the tomographic model and the result was a much more realistic size for the low velocity zone associated with the caldera. These "corrected" models are shown in the right panels of Figure 20.

HIGH RESOLUTION TOMOGRAPHY -- NEHT

In High Resolution tomography, the energy source is neither local nor teleseismic earthquake activity. Rather, we use explosive charges as the seismic energy source. Figure 21 shows a location map for a recent high resolution tomography experiment. Both Newberry volcano in Oregon, and Medicine Lake volcano in California are considered to be potential

geothermal energy sources. However, to estimate the feasibility of their energy potential, it was necessary to get some idea of the depth to the magma chamber and the size of the magma chamber for each volcano. To do so, we decided to employ a high resolution

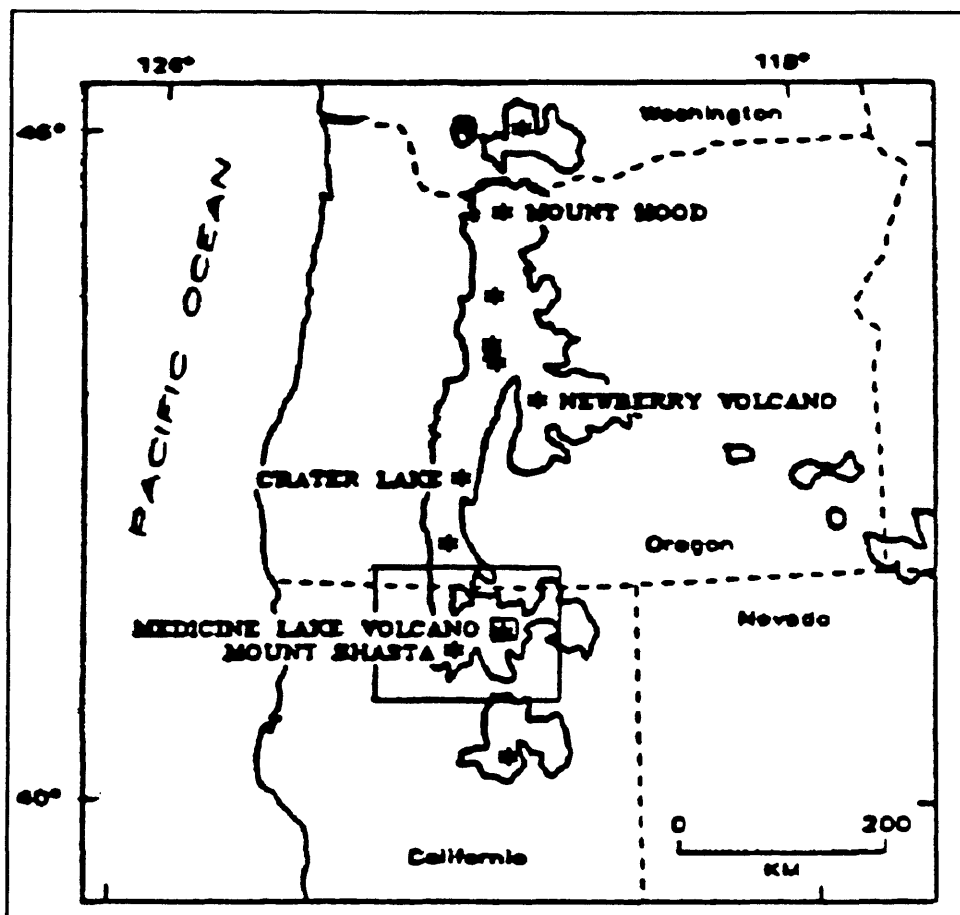


Figure 21 Location map For Medicine Lake volcano (From Evans and Zucca, 1988).

tomographic technique referred to as NeHT.

Figure 22 shows a map of the Medicine Lake region. The numbered stars on the map represent the locations of source explosions for the experiment. These were very large explosions, each one produced by several thousand pounds of explosive. The seismic array is

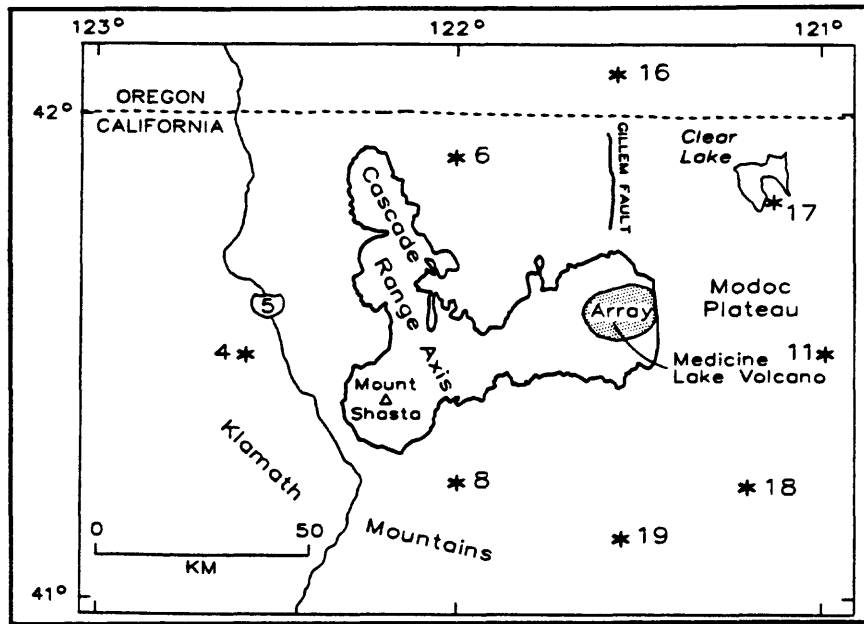


Figure 22 Map of Medicine Lake region (From Evans and Zucca, 1988).

shown as the stippled region surrounding the volcano.

Figure 23 shows a closeup view of the region around the volcano. The caldera is enclosed by the dashed curve and the seismograph locations are indicated by the dots. Medicine Lake is indicated by the symbol ML. John Evans, who was the leader of this study, used almost the same procedure to invert the data as was used in the teleseismic tomography experiment. However, in addition to computing velocity perturbations throughout the study region, in collaboration with Jay Zucca, John Evans also computed differential attenuation using the spectral information in the signals.

Figure 24 and Figure 25 show the velocity perturbation and the differential attenuation for two layers near the surface. Where there is both low velocity and high attenuation, there is probably a magma chamber. Regions which have high velocity and high attenuation are

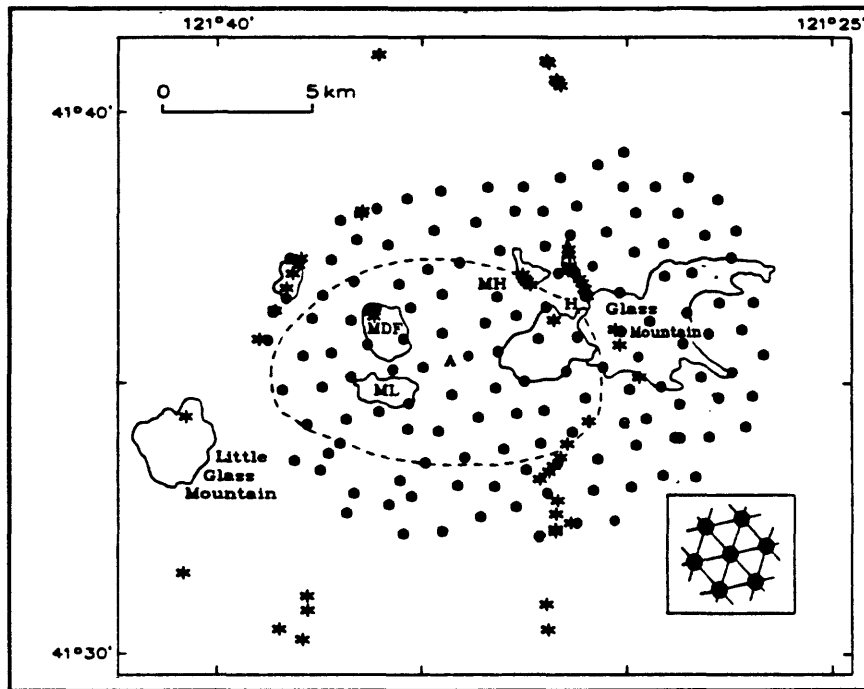


Figure 23 Medicine Lake volcano array (From Evans and Zucca, 1988).

probably good prospects for geothermal development. That situation can be caused by boiling-hot water. This experiment was expensive but the exciting thing is that we have been able to get detailed information about the subsurface environment using seismic tomography. There are people all over the world doing many types of investigations which involve tomography. I think that over the next 10 or 20 years, anyone who is working in seismology cannot help getting involved in seismic tomography.

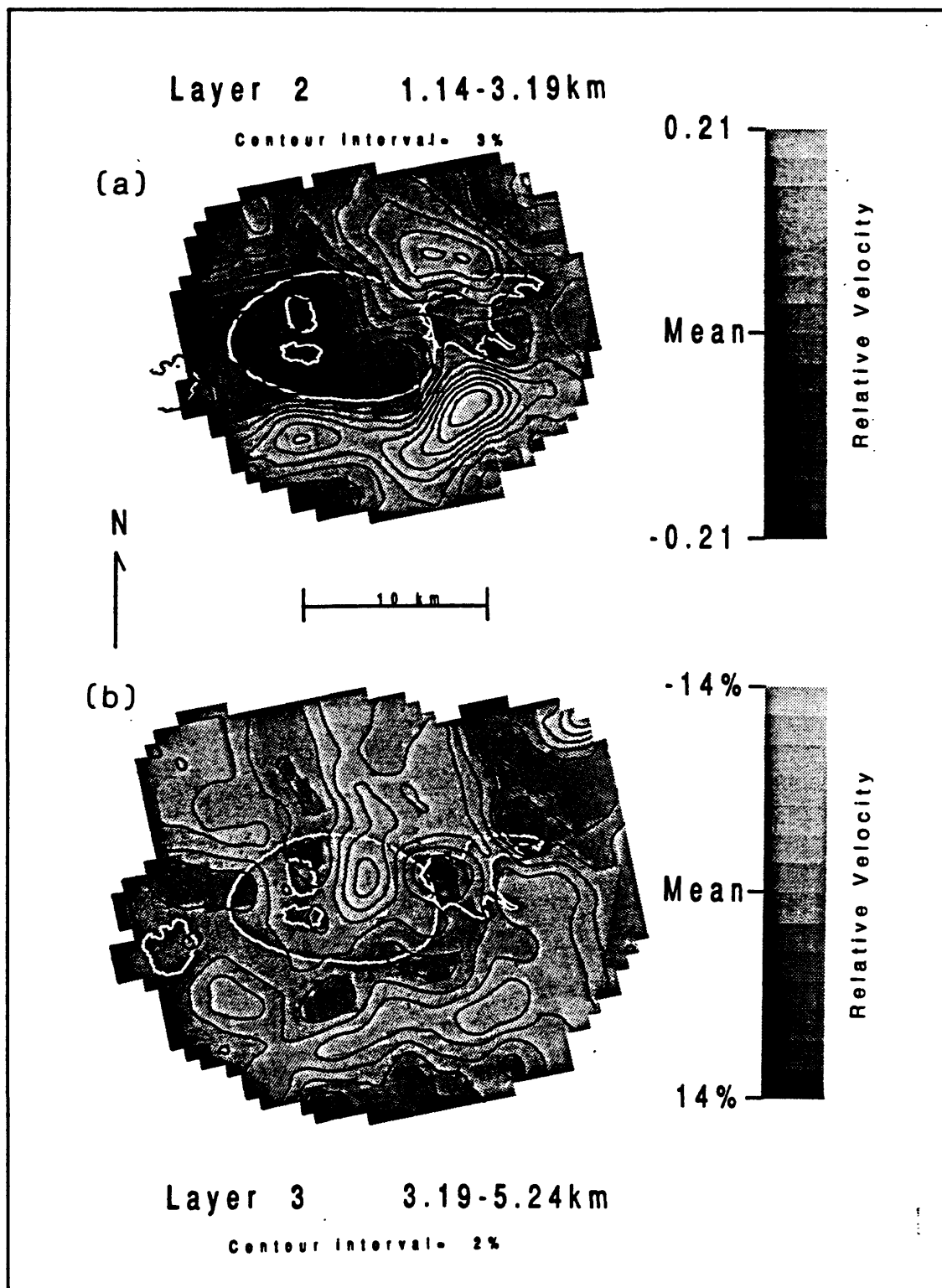


Figure 24 Velocity perturbations: layers 2 & 3 (From Evans and Zucca, 1988).

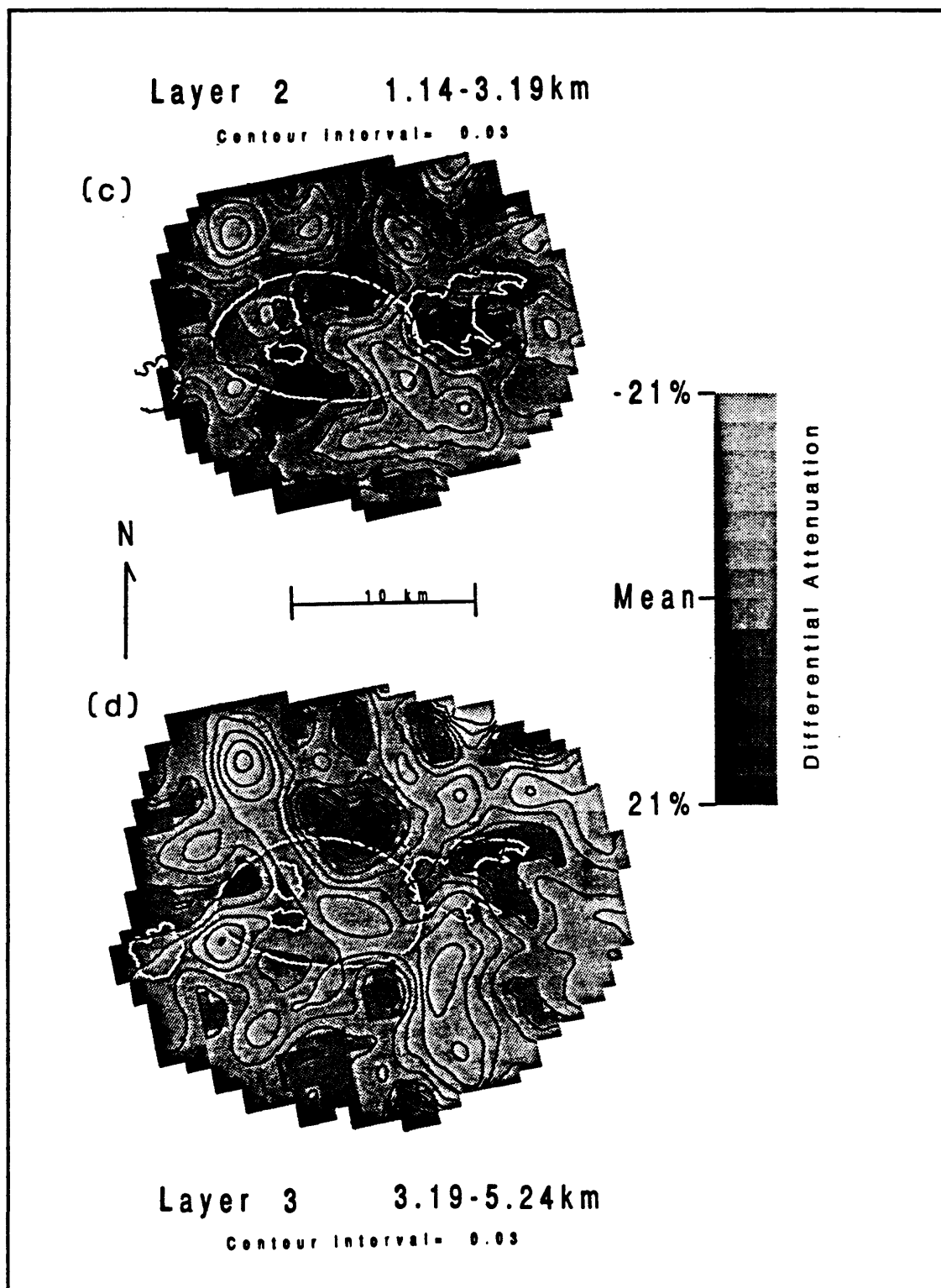


Figure 25 Differential attenuation: Layers 2 & 3 (From Evans and Zucca, 1988).

SUGGESTED READING

- Aki, K., Christoffersson, A., and E. S. Husebye (1976). Three-dimensional seismic structure of the lithosphere under Montana LASA: Seismological Society of America Bulletin, **66**, 501-524.
- Aki, K., Christoffersson, A., and E. S. Husebye (1977). Determination of the three-dimensional seismic structure of the lithosphere. Journ. Geophys. Res., **82**, no. 2, 277-296.
- Aki, K. and W.H.K. Lee (1976). Determination of three dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes. Part I. A homogeneous initial earth model. Journal of Geophysical Research, **88**, 4381-4399.
- Anderson, D.L., and A.M. Dziewonski (1984). Seismic tomography, Scientific American, **251**, 60-68.
- Dawson, P. B., Evans, J. R., and H. M. Iyer (1990). Teleseismic tomography of the compressional wave velocity structure beneath the Long Valley region, California. Journ. Geophys. Res., **95**, No. B7, 11,021-11,050.
- Dziewonski, A.M., and D.L. Anderson (1984). Seismic tomography of the earth's interior, American Scientist, **72**, 483.
- Eberhart-Phillips, D. (1989). Active faulting and deformation of the Coalinga anticline as interpreted from three-dimensional velocity structure and seismicity, Journ. Geophys. Res. **94**, 15565-15363.
- Eberhart-Phillips, D. (1990). Three-dimensional P and S velocity structure in the Coalinga Region, California, Journ. Geophys. Res. **95**, 15343-15363.
- Eberhart-Phillips, D., Labson, V. F., Stanley, W. D., Michael, A. J., and B. D. Rodriguez (1990). Preliminary velocity and resistivity models of the Loma Prieta earthquake region. Geophysical Research Letters, **17** No. 8, 1235-1238.
- Ellsworth, W. L., and R. Y. Koyanagi (1977). Three-dimensional crust and mantle structure of Kilauea volcano, Hawaii. Journ. Geophys. Res., **82**, No. 33, 5379-5394.
- Evans, J.R. and J.J. Zucca (1988). Active high-resolution seismic tomography of compressional wave velocity and attenuation structure at Medicine Lake Volcano, Northern California Cascade Range. Journ. Geophys. Res. **93**, 15016-15036.

- Harris, R.A., Iyer, H.M., and P.B. Dawson (1991). Imaging the Juan de Fuca Plate Beneath Southern Oregon using teleseismic P wave residuals. *Journ. Geophys. Res.* **96**, 19,879-19,889.
- Iyer, H.M. (1988). Seismological detection and delineation of magma chambers beneath intraplate volcanic centers in western U. S. A. in "modeling of Volcanic Processes" (Chi-Yu King and Roberto Scarpa Eds.) Friedr. Vieweg & Sohn, Braunschweig/Weisbaden.
- Iyer, H.M. (1989) Seismic Tomography. In: James D.E. (ed) *Encyclopedia of Geophysics*, 1133-1151, Van Nostrand Reinhold & Co.
- Michael, A.J. and D. Eberhart-Phillips (1991). Relations among fault behavior, subsurface geology, and three-dimensional velocity models, *Science* **253** no. 9, 651-654.
- Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California. *Journ. Geophys. Res.* **88**, No. B10, 8226-8236.
- Thurber, C., and Aki, K. (1987). Three-dimensional seismic imaging, *Ann. Rev. Earth and Planet. Sci.*, **15**, 115-139.

APPENDIX

Implementing a PC-Based Seismic System: Questions & Answers

DISCLAIMER

<p>This document was prepared by W. H. K. Lee to express his personal views. <i>Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by any organizations.</i></p>
--

1. Why this document?

I prepared this document because many people have asked me how to implement a PC-based seismic system. *My answers are obviously based on my own experience, and my recommendations are of necessity subjective.*

2. What is a seismic network?

A seismic network consists a number of individual seismometers in the field and a recording and processing center. Telemetry may be analog or digital and signals are transmitted from the field either by hardwire, radios, satellites, telephones, etc. For a general description, please see Lee and Stewart (1981).

3. What is a PC-based seismic system?

A PC-based seismic system is a computer system configured for acquisition, processing, and analysis of data from a seismic network. In the past the computer used was usually a mini-computer with specialized hardware and software (costing hundreds of thousands of dollars, and taking several weeks to install). Because personal computers are now inexpensive, a minimum PC-based seismic system can be implemented for as little as \$5,000 and requires less than 1 hour to install. If you already have a suitable IBM-compatible PC/AT, 286 PC, or 386 PC, then the cost can be as little as \$2,000.

A PC-based seismic system expects multiple channels of analog signals (either hardwired or telemetered) as inputs. The analog signals are usually a few volts in amplitude (peak-to-peak), and will be digitized at a prescribed sampling rate. The system records the digital data either continuously or by event trigger.

4. How to build and operate a seismic network?

At present, several commercial companies offer many types of seismic networks at various price levels (usually about \$100,000 for a network of a few stations). If money is no object and you don't want to get your hands dirty, buying a turnkey seismic network or even

turning the whole operation over to a commercial company is your best bet. On the other hand, if your budget is limited and you wish to do it yourself as much as possible, then your best bet is first to study the literature and visit some established seismic networks. You can then design and build your own seismic network using whatever commercially available parts that are useful to you. The same reasoning applies to the data acquisition and processing part of a seismic network.

5. What is the PC-Quake system?

In 1987, I was conducting a study of explosive sources in a quarry. There was no commercial seismic system that I could afford to do my job. Consequently, I had to use whatever components were available (mostly on loan from my colleagues) and implemented a simple system that worked much better than expected. Because no direct funding was available from the U. S. Geological Survey (USGS), I had to spend quite a bit of my own time and money, and I also had to persuade several non-USGS friends to help out.

Like most scientists, I want to share the fruit of my labor with anyone who is interested. Consequently, all work done on USGS time was released to the public in Lee et al. (1988), Ellis (1989), and Tottingham and Lee (1989). In addition, a volume (Lee, 1989) called "Toolbox for Seismic Data Acquisition, Processing, and Analysis" was published as IASPEI Software Library Volume 1 under the auspices of the IASPEI Working Group on Personal Computers in collaboration with the Seismological Society of America (SSA). This volume contains contributions from USGS and non-USGS scientists. For lack of a better name, I called this PC-based seismic system the "PC-Quake system" in Lee (1989).

In essence, Lee and Stewart (1981) and Lee (1989) contain information on how to build a seismic network and seismic data acquisition and processing system. Neither USGS, IASPEI, SSA nor Lee is selling a seismic network or a data acquisition and processing system.

6. What are the characteristics of the PC-Quake system?

The PC-Quake system as published in Lee (1989) is a general-purpose seismic system, with the following characteristics:

- (1) Uses IBM-compatible 286, 386sx, 386, or 486 PCs with an 8-MHz AT bus.
- (2) Uses analog-to-digital boards (DT 282x series) made by Data Translation, Inc.
- (3) Digitizes 16 analog channels (upgradable to 256 channels) at user selectable rate.
- (4) Digitizes up to a few thousand samples per second per channel.
- (5) Displays digitized data continuously and in real time.
- (6) Saves digitized data continuously or by event triggering.
- (7) Automatically picks P-arrivals and locates events.
- (8) Offline analysis includes manual picking, filtering, FFT, & coda Q.
- (9) Channel to digitize, or display, or trigger is user selectable.

7. Who needs a PC-based seismic system?

Any seismologist who runs a seismic network or collects seismic data in the field should consider a PC-based seismic system because it is the most cost-effective way to do "digital seismology". Even if one had a computerized seismic system, a PC-based seismic system could serve as an inexpensive backup system. It can also be used to teach students about observational seismology.

Because the PC-Quake system displays the digitized data in real time, it replaces the costly Develocorder film system. Furthermore, any digital data recorded by this system is available for further analysis in a few tens of seconds, instead of minutes or hours as in many other computerized systems.

8. Who is using the PC-Quake system?

At present, about 60 PC-based seismic systems are in operation around the world by about 30 different groups. After about 3 years since the system was published in Lee (1989), it has become the most popular computerized seismic system due to its low cost and ease of implementation.

Current users include: U.S. Geological Survey, U.S. Bureau of Mines, University of Idaho, Montana Bureau of Mines, California Department of Water Resources, University of Alaska, Stanford University, University of Southern California, Arizona State University, Memphis State University, Los Alamos National Laboratory, and institutions in Chile, Colombia, Costa Rica, Ecuador, Hong Kong, Indonesia, Italy, Japan, Jordan, Mexico, New Zealand, Philippines, Puerto Rico, Spain, Taiwan, Turkey, USSR, and Venezuela.

9. What is IASPEI?

IASPEI is the acronym for the International Association of Seismology and Physics of the Earth's Interior, an international scientific organization to promote the study of seismology and earth's interior. In 1987 a Working Group on Personal Computers was established to promote the exchange and sharing of information on the use of personal computers for geophysical research.

10. What is IASPEI Software Library?

Under the auspices of the IASPEI PC Working Group, a series of volumes (including software) are now being published in collaboration with the Seismological Society of America (SSA). This series is being edited by W. H. K. Lee and Frank Scherbaum, with an editorial advisory board chaired by H. Kanamori. IASPEI software is written and contributed by volunteers.

Each volume is devoted to a major topic in seismology. For example, Volume 1 is a toolbox for seismic data acquisition, processing, and analysis, and contains 6 executable programs with detailed documentation. The executable volume is available for \$250 through the Seismological Society of America, 201, Plaza Professional Building, El Cerrito, CA 94530, USA (Phone 415-525-5474). In addition, source code is published in a source code package that is sold separately for \$250.

Some complimentary copies of the IASPEI Software Library volumes are available to any one who needs the software but can not afford them or lack the U.S. currencies. Members of Seismological Society of America or of Seismological Society of Japan can order the software volumes at 50% and 40% discount, respectively, if they are paying with personal funds.

11. What are the hardware requirements for the PC-Quake system?

A detailed discussion on the hardware requirements has been published in IASPEI Software Library Volume 1. Critical components are:

- (1) PC -- Any 100% IBM-compatible 286, 386sx, 386, or 486 PC. It must have the industry standard 8-MHz AT bus and Intel DMA controller. A minimum of 2 megabytes of RAM (of which 1 or more Megabytes are extended memory), a math co-processor (except for 486 PC), a hard disk, a floppy disk drive, and Hercules, or EGA, or VGA graphics adapter and corresponding monitor.
- (2) Hard Disk -- At least 30 Megabytes capacity. This will hold about 100 earthquakes (16 channels, digitizing at 100 samples per second per channel, and saving a event for 100 seconds). A larger capacity hard disk is needed if you have more than 16 channels.
- (3) A/D Board -- Must be DT2821, or DT2824, or DT2827 made by Data Translation. We recommend the 16-channel, 12-bit DT2824 for most seismic applications. The DT2821 board is identical to DT2824, but with the D/A components that we don't use. The DT2827 is a 4-channel, 16-bit A/D board.
- (4) Multiplexer -- A multiplexer designed by Jim Ellis of U. S. Geological Survey (Ellis, 1989) will upgrade the system to handle up to 128 channels. Two improved multiplexers (up to 256 channels with a modified 12-bit DT2824 A/D board, and up to 64 channels with a modified 16-bit DT2827 A/D board) are now available from Quesig Electronics. IASPEI software has been revised to support Ellis' and Quesig Electronics' multiplexers.
- (5) LAN -- I recommend using the LANtastic kit by Artisoft to link the online PC with the offline PC. This way, digitized data can be accessed from the offline PC without interrupting the online PC.
- (6) MASS STORAGE -- Originally, I recommend using the IBM 3363 optical WORM drive to archive digitized data on 200-megabyte cartridges. This is an inexpensive and quick method to archive and retrieve data. However, IBM does not seem to be willing to sell it. At present, several removable mass storages are available, including erasable optical cartridges, DAT tapes, etc. A very cost effective and common mass storage is the QIC-80 tape drives, using 1/4 inch cartridge tapes (up to 250 MB). The drive is about \$400 and the tape is about \$30 each.
- (7) MOUSE -- a mouse is required for data analysis. I recommend the Microsoft serial mouse.
- (8) Printer -- I recommend a HP LaserJet III laser printer with Postscript cartridge. A

dot-matrix printer (such as an IBM ProPrinter or an Epson FX series printer) is too slow for plotting seismic traces.

12. What are the software requirements for the PC-Quake system?

The necessary software for the PC-Quake system (16 channels) has been published in IASPEI Software Library Volume 1 to run under either IBM PC-DOS or Microsoft MS-DOS. I recommend using either version 3.3 or 4.01. For hard disk partitions larger than 40 megabytes, you must use Version 4.01. In addition, I recommend the following utilities: (1) PCTOOLS by Central Point Software for disk optimization, (2) XTREE by Executive Systems for file management, (3) PCKWIK by Multisoft Corp. for disk caching and print spooling, and (4) KEDIT by Mansfield Software as a general text editor.

For users who need to modify the published IASPEI software, a source code package is available. However, Microsoft compilers and two other commercial software packages are also required.

An update of IASPEI Software Library Volume 1 (to expand data acquisition to 256 channels and for EGA/VGA color monitors) is available, if the user of Volume 1 returns the Registration Form with a request for the update.

13. What brands of PC can be used?

Because the online data acquisition is a continuous operation, a high quality PC is required. Although many brands of PC claim to be 100% IBM compatible, this is often NOT the case. I have tested many PC brands during the past few years, and some did NOT work with the IASPEI software. The following PCs have been tested successfully with the IASPEI software:

- (1) IBM PC/AT, and Compaq Portable II (these older PCs are no longer in production).
- (2) Compaq DeskPro 286, 386s, and 386.
- (3) Everex Step 286, 386sx, 386, and 486.
- (4) AST Premium 286, 386, and 486.
- (5) NEC Powermate 386sx portable.
- (6) Toshiba T3200sx portable.

Although the better-known name brands cost more, they are often more reliable and much easier to obtain repair service worldwide. I recommend using an IBM PC/AT (if you can find one) or Compaq DeskPro PCs for laboratory environments, and using the NEC Powermate portable for field operations.

14. Can an IBM PS/2 be used?

At present, you can use an IBM PS/2 only for offline analysis.

15. How can one upgrade to more than 16 channels?

In 1989, Jim Ellis designed a multiplexer to allow a Data Translation A/D board to digitize up to 128 channels (USGS Open-file Report 89-191), and anyone can copy Ellis design (later modified by John Rogers and Grey Jensen), as it is in the public domain. This is a do-it-yourself task.

Two improved multiplexers (up to 256 channels with a modified 12-bit DT2824 A/D board, and up to 64 channels with a modified 16-bit DT2827 A/D board) are now commercially available from Quesig Electronics.

16. What are the updates on IASPEI Software Library Volume 1?

- (1) 256 Channel -- as explained in Question #15, the revised software can handle up to 128 channels in conjunction with Ellis' multiplexer (Ellis, 1989), or up to 256 channels in conjunction with a Quesig Electronics multiplexer.
- (2) SUN-compatible data -- the revised IASPEI data acquisition software uses Peter Ward's SUDS data format (Ward, 1989; Banfill, 1992), which is compatible with many analysis programs now running on SUN workstations.
- (3) Other updates -- PCEQ and QCODA have been revised to read SUDS format also. A program is available to automatically correct the time if IRIG-E time code is digitized along with the seismic signals. A program to convert SUDS data into an ASCII file is also available so that the digitized data can be exported to almost any computer.
- (4) New programs -- several new programs are now being written by Small Systems Support of Utah. These include SudsPick, a program to perform automatic P-pickings and visual corrections, SudsPlot, a program to plot seismic traces on a HP LaserJet printer, Squeeze, a program to condense seismic data for archiving purposes, and QMap, a program to plot hypocenter data on a map. These programs are now available as updates to IASPEI Software Library Volumes 1 and 2.

17. What is the shopping list for a PC-Quake system?

The PC-Quake system as published in Lee (1989) is a do-it-yourself system. It can be easily tailored to one's particular application. Because a 386 PC now costs only a few hundred dollars more than a 286 PC, we recommend using a 386 PC. A 486 PC is not necessary for data acquisition unless you need very high sampling rates (thousands of samples per channel per second instead of hundreds of samples per channel per second). I also recommend buying a top-quality PC, because it will be more reliable and hence far cheaper and less of a headache in the long run.

How you implement a PC-Quake system depends on your needs, your budget, and how much you wish to do yourself.

The standard system design calls for two identical IBM-compatible 286 or 386 PCs, with a 12-bit, 16-channel analog-to-digital board made by Data Translation. One PC is used for on-line data acquisition and the other PC is used for off-line data processing and analysis. These two PCs are linked by a high-speed local area network (LANtastic kit by

Artisoft), and an optical WORM drive (IBM 3363) is used to archive the digitized data on 200-megabyte removable cartridges. The off-line PC also serves as a backup to the on-line unit. A multiplexer is required to increase the number of digitized channels to a maximum of 256 channels.

There is a trade-off between the speed of a PC, the number of digitized channels, and the digitization rate. For example, an 8-MHz 286 PC is more than adequate for a 16-channel system digitizing at 100 samples per second per channel. Although one can use a 12-MHz 286 PC to digitize 128 channels at the same sampling rate, I recommend a more powerful PC (such as a 16-MHz 386sx PC, or a 20 or 25-MHz 386 PC).

For budgeting purposes, I suggest 5 configurations: (A) Minimum single PC system (16 channels), (B) Standard dual PC system (16-64 channels), (C) Medium-duty system (64-128 channels), (D) Heavy-duty system (256 channels) and (E) Field system (see Appendices A through E). Please note the prices listed are approximate and are based on recent prices in the U.S. These prices do NOT include sales taxes (7.25% if applicable) and shipping charges. You may obtain current prices from Symmetric Research.

18. Where to buy?

You can buy the individual pieces separately from various stores and manufacturers. You can then assemble the hardware, install the software, and test the system yourself. Several groups have done it this way, and it usually takes a few days to order and a few days to assemble and test. If you have had experience in working with PCs and have ordered everything correctly, then you should not have any problems. On the other hand, if a problem arises and you ordered the pieces from different stores, then it may take some time to straighten things out.

Alternatively, you can order every thing you need from Symmetric Research, 15 Central Way, Suite #9, Kirkland, WA 98033. Telephone: 206-828-6560; FAX: 206-827-3721.

I recommend that you compare prices, consider different alternatives, and decide what is best for you. The following is a list of potential suppliers:

- (1) Computer system and software: your favorite computer store(s).
- (2) A/D Board: Data Translation, Inc., 100 Locke Drive, Marlboro, MA 01752; Telephone: 508-481-3700.
- (3) UPS: Best Power Technology, P. O. Box 280, Necedah, WI 54646; Telephone: 800-356-5794.
- (4) LAN: Artisoft Inc., 3550 N. First Ave., #330, Tucson. AZ 85719; Telephone: 602-293-6363.
- (5) IASPEI Software: Seismological Society of America, 201 Plaza Professional Building, El Cerrito, CA 94530; Telephone 415-525-5474.
- (6) Multiplexers: Quesig Electronics, 1863 Channing Ave., Palo Alto, CA 94303; Telephone: 415-326-0501 (evenings only); FAX: 415-328-3293.
- (7) Amplifiers: Kinemetrics/Systems, 222 Vista Ave., Pasadena, CA 91107; Telephone 818-795-2220.

19. What are users' comments?

Excerpts from some users' comments are:

"We are very happy and busy with your system, since June '89 (diskettes are piling up!) ... We have [written] some routines ... which could be useful to others." -- *H. Meyer, Colombia.*

"I'm writing to express my sincere gratitude for the time and effort that the IASPEI Working Group on Personal Computers, has spent developing Volume 1 ... The PC Quake system worked perfectly from the beginning and has operated reliably since August, 1989. ... The availability of the low-cost digital recording and analysis package through your organization is the only way this small network was able to enter the digital age" -- *M. Stickney, USA.*

"The PCEQ package is working extremely well on our data. It is very pleasant to use and proving popular with our analysts." -- *T. Turbitt, UK.*

"I express our thanks for your great effort to develop and distribute the IASPEI Software Library. When I first [saw] the Volume 1 at Prof. Y. Fukao's office last year, I found it just for us! Our institute IISEE is a training institute of seismology and earthquake engineering for researchers in developing countries ... The first ten packages were given to the participants of special seminar course held in last December. ... all the participants were pleased very much." -- *H. Inoue, Japan.*

20. Is PC-Quake system a "finished" product?

The answer is **no**, because a system can not be designed to meet every possible need from the users. I believe that the PC-Quake system is a good start, but it is up to the users to modify the system to meet their particular needs. Recently, a cooperative program between the U.S. Geological Survey and the Taiwan Central Weather Bureau has been established to jointly develop an advanced earthquake monitoring system with early warning capability. Because the new system will be based on the PC-Quake system, more rapid improvement and new development will be made in the next few years.

21. How long it will take to learn the PC-Quake system?

Depending on your experience with PCs, it may take from one to a few days to learn how to use the system.

22. Why it takes two PCs? Can't you multi-task?

As explained in Answer #17, the online PC is dedicated to digitizing and recording seismic data. The reason is that we use a single-task operating system. Multi-tasking systems are far more complex, and we have not tried this approach. Because PCs are inexpensive, we believe that the dual PC design is cost effective and provides a backup for online data acquisition.

23. 12-bit or 16-bit?

The updated IASPEI Volume 1 software supports data acquisition at either 12-bit or 16-bit (dynamic range of either 72 dB or 96 dB). Quesig Electronics has a multiplexer for either 12-bit or 16-bit A/D board at about the same price. Therefore, the question of 12-bit or 16-bit boils down to this: "are you willing to spend an additional \$1,300 for a 16-bit A/D board (which supports only 4 channels), and an additional \$1,500 for a multiplexer (which you most likely will need)?"

If you need between 16 and 64 channels, then I would recommend you use a 16-bit system; otherwise, it is more cost effective to use a 12-bit system with lower dynamic range. Because the traditional analog telemetered network has a dynamic range of about 50 dB, it is more than adequate to use a 12-bit A/D board.

24. What are the electric power requirements?

A typical PC normally requires about 400 watts of electrical power, depending on the peripherals attached. Therefore one should plan for about 800 watts for a dual PC-based seismic system. Most PCs nowadays can operate either on 110 volt and 60 Hz, or 220 volt and 50 Hz. For field operations, it is possible to power a portable PC (which requires less power than a desktop PC) via a UPS (uninterruptable power system) using large batteries. In countries where electric power is not steady, it is a good practice to use a UPS.

25. What is the trend in seismic data acquisition and analysis?

I believe that all seismic data acquisition will eventually be performed by PCs, and proprietary or complicated seismic systems will disappear. The reason is simple: PCs are inexpensive and a PC-based seismic system can be easily implemented and maintained.

With funding from the USGS Volcano Programs, two portable seismic arrays are being implemented now: (1) a 64-channel, 16-bit cabled array, and (2) a 48-channel (sixteen 3-component stations), 16-bit, digital telemetered array. Both of these arrays use the PC-Quake system for data acquisition, processing, and analysis (Lee, 1992).

An extensive earthquake monitoring program is now being implemented by the Central Weather Bureau in Taiwan: (1) 600 free-field digital accelerographs are being deployed; these accelerographs are compatible with the IASPEI software, (2) up to 80 buildings and structures are being instrumented with a 32- or 64-channel, 16-bit cabled array, and (3) a prototype earthquake warning system is being developed in cooperation with USGS. Because items (2) and (3) are based on the PC-Quake system, more rapid improvement and new development will be made in the next few years.

The current trend in seismology is to use high-speed PC or SUN workstations for computation and analysis. To facilitate the use of digital seismic data on SUN workstations, we adopted a SUN-compatible data format, and data collected by our PC-based system can be easily ported to SUN workstations (or any other computer).

26. How much money should one budget?

The costs of hardware and software vary from country to country. In general, the United States has perhaps the lowest prices. You should compare prices among various alternatives. If a price is too low, then you may not get the "real" product. Some computer stores substitute inferior parts in a name-brand product and sell it cheaply. Some do not provide any services or technical help. My advice is to buy only from authorized dealers with a good reputation. It may cost you a little more, but can save you from a lot of headaches.

If you plan to buy a system in the U.S., you should budget an additional \$ 3,000 for travel expenses, and \$ 2,000 for shipping and insurance. If you need software modifications, it will require someone on your staff who can program in C, Fortran, and Macro Assembler. Alternatively, you should budget enough money for consulting (typical rate in the U.S. is \$50 to \$100 per hour).

In general, \$23,000 is a good figure for budgeting purposes in buying a standard dual PC-based seismic system (see Appendix B) in the U.S., including funds for travelling and shipping.

References

- Banfill, R., (1992). SUDS: Seismic Unified Data System, version 1.31, Small Systems Support, P. O. Box 205, Big Water, Utah 84741, USA.
- Ellis, J. O., (1989). Expanding the input multiplexer for the Data Translation, Inc. Model DT2821 analog-to-digital converter, *U.S. Geological Survey Open-File Report 89-191*.
- Lee, W. H. K. (Compiler), (1992). Descriptions of seismic array components: Part 1. DIGGER, DIGIREC, and MULTIPLEXER, *U.S. Geological Survey Open-File Report 92-430*, 276 pp.
- Lee, W. H. K. (Editor), (1989). "Toolbox for Seismic Data Acquisition, Processing, and Analysis", *IASPEI Software Library Volume 1*, Seism. Soc. Am., El Cerrito, CA.
- Lee, W. H. K., and S. W. Stewart, (1981). "Principles and Applications of Microearthquake Networks", *Advances in Geophysics, Supplement 2*, Academic Press, New York, 293 pp.
- Lee, W. H. K., D. M. Tottingham, and J. O. Ellis, (1988). A PC-based seismic data acquisition and processing system, *U.S. Geol. Surv. Open-File Report 88-751*, 31 pp.
- Tottingham, D. M., and W. H. K. Lee, (1989). XDETECT: A fast seismic data acquisition and processing program, *U.S. Geol. Surv. Open-File Report 89-205*, 8 pp.

Appendix A. Minimum Single PC System (16 Channels)

The following list contains all the necessary hardware and software parts for a minimum single PC system (which can be upgraded to more than 16 channels by adding a Quesig Electronics' multiplexer). Each item is described and followed by the quantity needed and the approximate sale price in the U.S.

(1) Compaq DeskPro 386s PC (20 MHz) with 2 MB RAM, 40 MB hard disk, 1.44 MB floppy drive, serial/parallel port, clock and calendar, and enhanced keyboard	1	\$ 2,000
(2) Intel 80387sx math coprocessor (20 MHz)	1	200
(3) Compaq VGA color monitor (14")	1	500
(4) IBM PC-DOS version 4.01	1	100
(5) Data Translation DT2824-PGH analog-to-digital board	1	1,300
(6) Data Translation DT707 screw terminal panel	1	200
(7) Microsoft serial mouse & mouse pad	1	100
(8) High-quality power surge protector	1	50
(9) IBM ProPrinter and printer cable	1	250
(10) IASPEI Software Library Volume 1	1	250
(11) PCTOOLS by Central Point	1	100
		<hr/>
Total :		\$ 5,050

Appendix B. Standard Dual PC System (16-64 Channels)

The following list contains all the necessary hardware and software parts of a dual PC system for a small to medium size seismic network. Each item is described and followed by the quantity needed and the approximate sale price in the U.S. *Add about \$3,000 to upgrade the following 12-bit system to a 16-bit system (i.e., replace Items (6) and (7) by two DT 2827 A/D boards and specify Item (19) for two 16-bit multiplexers).*

(1) Compaq DeskPro 386 PC (25 MHz) with 4 MB RAM, 100 MB hard disk, 1.44 MB floppy drive, serial/parallel port, clock and calendar, and enhanced keyboard	2	\$ 5,000
(2) Intel 80387 math coprocessor (25 MHz)	2	400
(3) Compaq VGA color monitor (14")	2	1,000
(4) 5.25-inch 1.2 MB floppy disk drive	2	200
(5) IBM PC-DOS version 4.01	2	200
(6) Data Translation DT2824-PGH analog-to-digital board	2	2,600
(7) Data Translation DT707 screw terminal panel	2	400
(8) LANtastic hardware & software by Artisoft	1	400
(9) Microsoft serial mouse & mouse pad	1	100
(10) High-quality power surge protector	2	100
(11) HP LaserJet III laser printer with 3 MB RAM, etc.	1	2,000
(12) QIC-80 internal backup tape drive	1	400
(13) Miscellaneous supplies	1	500
(14) IASPEI Software Library Volume 1	1	250
(15) PCTOOLS by Central Point	1	100
(16) Quesig Electronics' multiplexer for 64 channels	2	3,000
(17) System integration, testing and training	1	1,000
Total :		\$ 17,650

Appendix C. Medium-duty System (64-128 Channels)

The following list contains all the necessary hardware and software parts of a dual PC-based system for a medium to large size seismic network. Software for modifying the IASPEI programs is also included. Each item is described and followed by the quantity needed and the approximate sale price in the U.S.

(1) Compaq DeskPro 386/33 PC (33 MHz) with 4 MB RAM, 200 MB hard disk, 1.2 MB floppy drive, serial/parallel port, clock and calendar, and enhanced keyboard	2	\$ 6,000
(2) Intel 80387 math coprocessor (33 MHz)	2	400
(3) Compaq VGA color monitor (14")	2	1,000
(4) 5.25-inch 1.2 MB floppy disk drive	2	200
(5) IBM PC-DOS version 4.01	2	200
(6) Data Translation DT2824-PGH analog-to-digital board	2	2,600
(7) LANtastic hardware & software by Artisoft	1	500
(8) Microsoft serial mouse & mouse pad	1	100
(9) High-quality power surge protector	2	100
(10) HP Laserjet III laser printer with 3 MB RAM, etc.	1	2,000
(11) QIC-80 internal backup tape drive	2	800
(12) Miscellaneous supplies	1	500
(13) IASPEI Software Library Volumes 1, 2, 3, & 5	1	1,000
(14) IASPEI Software Library Volume 1 Source Code Package	1	250
(15) PCTOOLS by Central Point	1	100
(16) KEDIT by Mansfield Software	1	150
(17) XTREE EASY by Executive Systems	1	50
(18) PCKWIK Power Pak by Multisoft	1	100
(19) Microsoft compilers, ATLAB, and METAWINDOW software	1	2,000
(20) Quesig Electronics' multiplexer for 128 channels	2	3,500
(21) System integration, testing and training	1	1,000
Total :		<hr/> \$ 22,550

Appendix D. Heavy-duty System (256 Channels)

The following list contains all the necessary hardware and software parts of a dual PC-based system for a large size seismic network. Software for modifying the IASPEI programs is also included. Each item is described and followed by the quantity needed and the approximate sale price in the U.S.

(1) Compaq DeskPro 386/33 PC (33 MHz) with 4 MB RAM, 300 MB hard disk and controller, 1.2 MB & 1.44 MB floppy drives, serial/parallel port, clock and calendar, and enhanced keyboard	2	\$ 10,000
(2) Intel 80387 math coprocessor (33 MHz)	2	400
(3) Compaq VGA color monitor (14")	1	500
(4) Orchid Pro-designer IIs Super-VGA adapter and color monitor (16") .	1	1,000
(5) IBM PC-DOS version 4.01	2	200
(6) Data Translation DT2824-PGH analog-to-digital board	2	2,600
(7) LANtastic hardware & software by Artisoft	1	400
(8) Microsoft serial mouse & mouse pad	1	100
(9) High-quality power surge protector	2	100
(10) HP Laserjet III laser printer, with 3 MB RAM, Postscript cartridge ..	1	2,500
(11) QIC-80 internal backup tape drive	2	800
(12) Miscellaneous supplies	1	500
(13) IASPEI Software Library Volumes 1, 2, 3, & 5	1	1,000
(14) IASPEI Software Library Volume 1 Source Code Package	1	250
(15) PCTOOLS by Central Point	1	100
(16) KEDIT by Mansfield Software	1	150
(17) XTREE by Executive Systems	1	50
(18) PCKWIK Power Pak by Multisoft	1	100
(19) Microsoft compilers, ATLAB, and METAWINDOW software	1	2,000
(20) Quesig Electronics' multiplexer for 128 channels	2	3,500
(21) System integration, testing and training	1	1,000
<hr/>		
Total :		\$ 27,250

Appendix E. Field System (16 Channels)

The following list contains all the necessary hardware and software parts of a dual portable PC-based system for field deployment. This system is designed to collect seismic data in the field for several hours per day and it is powered by batteries. Although this system is configured for 16 channels, it can be easily upgraded to more than 16 channels by adding a Quesig Electronics' multiplexer. Each item is described and followed by the quantity needed and the approximate sale price in the U.S. *Add about \$6,000 to upgrade the following 12-bit system to a 16-bit system (i.e., replace Items (4) and (5) by two DT 2827 A/D boards and add two Quesig Electronics' 16-bit multiplexers).*

(1) NEC 386sx Powermate PC (16 MHz) with 4 MB RAM, 40 MB hard disk, 1.44 MB floppy drive, serial/parallel port, clock and calendar, gas plasma VGA screen, and enhanced keyboard	2	\$ 6,000
(2) Intel 80387sx math coprocessor (16 MHz)	2	400
(3) IBM PC-DOS version 4.01	2	200
(4) Data Translation DT2824-PGH analog-to-digital board	2	2,600
(5) Data Translation DT707 screw terminal panel	2	400
(6) LANtastic hardware & software by Artisoft	1	400
(7) Microsoft serial mouse & mouse pad	1	100
(8) High-quality power surge protector	2	100
(9) UPS by Best Technology (AC power via batteries)	2	3,000
(10) IBM ProPrinter and printer cable	1	250
(11) 16-channel low-noise seismic amplifier	1	2,000
(12) Miscellaneous supplies	1	500
(13) IASPEI Software Library Volume 1	1	250
(14) PCTOOLS by Central Point	1	100
(15) KEDIT by Mansfield Software	1	150
(16) XTREE by Executive Systems	1	50
(17) PCKWIK Power Pak by Multisoft	1	100
(18) System integration, testing and training	1	1,000
<hr/> Total :		\$ 17,600

NAME INDEX

Authors' names of Chapters in this report and the inclusive page numbers of these Chapters are indexed in **bold type**, page numbers of citations in the text are in ordinary type, and those of bibliographical references listed at the ends of the chapters (including joint authors) are in *italics*.

Names associated with known instrument, equations, etc., are not entered here but in the Subject Index.

Aki, K., 316-318, 323, 329, 331, *334-335*,
351-352, *358*, 410-411, 414-415, *423*,
426, *460*, 468-469, *476*, *478*, 481, 483-
484, *510-511*

Allen, R. V., 11, *50*, 133, *136*, *164*, 181
189, *424*

Allen, S. S., 130, *136*, *164*

Anderson, D. L., 482, *510*

Archuleta, R. P., 348, *358*

Bakun, W. H., 305, *315*

Banfill, R., 124, *136*, 172, 180-181, 183-
184, *189*, **190-206**, 517, *521*

Barrett, H. H., 476, *478*

Bennett, R. E., *250*, *315*, *335*, *358*

Bennioff, H., 83

Berger, J., *477*

Bisztricsany, E., 326, *334*

Bolt, B. A., 2, *50*, *423*

Boore, D. M., 305, 314, *315*

Brigham, E. O., 373, 378, *382*

Brillinger, D. R., *423*

Brodine, R. C., *164*

Brune, J., 299, *315-317*, *334*, 345, *358*,
477

Chouet, B., 316-317, 323, 331, *334-335*,
351-352, *358*

Christoffersen, A., 469, *476*, *510*

Claerbout, J. F., *476-477*

Clayton, R. W., 469, *477*

Cockerham, R., *51*

Comer, R., 268-269

Cormack, A., 461-464, *476-477*

Cranswick, E., *358*

Criley, E., 4, *50*

Crosson, R. S., 468, *477*

Dawson, P. B., 501, 503, *510-511*

Dziewonski, A. M., 469, *477*, 482, *510*

Eaton, J. P., 1-52, 55, 58, 64, 66, 71,
79-99, 130-131, 165, 252, 262-263,
280, **281-315**, 325

Eberhart-Phillips, D., 424-425, 485-486,
488-494, 496-498, *510-511*

Eliseevnin, V. A., 265-266, *280*

Ellis, J. O., *50*, 513, 515, 517, *521*

Ellsworth, W. L., 41, *510*

Espinosa, A. F., 317-319, *335*

Euler, L., 171

Evans, J. R., 130-131, *136*, 152-164,
505-509, *510*

Fletcher, J., *477*

Foley, J., *50*

Forsythe, G. E., *476-477*

Frankel, A., 317-318, 323, 330, *334*, 469,
477

Fukao, Y., 519

Galitsin, B., 80-81, 291-292

Gauss, K. F., 171

Geiger, L., 208, 215, 221, *225*

Geller, R. J., 469, *477*

Gill, P. E., 225, 476, 477
Griggs, D. T., 51
Gubbins, D., 266-267, 280
Gutenberg, B., 282-283, 292, 313, 315-317, 334

Haar, L., 477
Hager, B. H., 469, 477
Hanks, T., 316-317, 335, 477
Hamilton, R. M., 326
Harris, R. A., 483, 511
Healy, J. H., 3, 51
Herman, G. T., 476-477
Herraiz, M., 317-319, 335
Herrmann, R. B., 317-319, 323, 335
Hill, D. P., 17, 41, 51
Hirahara, K., 469, 476-477
Ho-Liu, P., 469, 477
Holland, C., 280
Honda, H., 411-412, 423
Houndsfield, G. N., 461-462, 476-477
Husebye, E. S., 469, 476, 510
Hutton, L. K., 41, 305, 314, 315

Inoue, H., 519
Iyer, H. M., 469, 476-477, 479-511

Jackson, W., 6
Jacob, K. H., 265-266, 280
James, D. E., 511
Jensen, G., 517
Johnson, C. E., 13, 15, 51, 296
Johnson, J., 188
Johnson, P., 335, 358
Jones, L. M., 17, 51
Joyner, W. B., 305, 315
Julian, B. R., 265-267, 280

Kanamori, H., 317-318, 335, 469, 477, 514
Kasahara, K., 423
Keller, H. B., 268, 280
King, C. Y., 511

Klein, F. W., 3, 26, 48, 52, 225, 246, 250
Koyanagi, R. Y., 3, 52, 510

Labson, V. F., 510
Lahr, J. C., 121, 133, 137, 143-144, 182, 225, 226-250, 343, 358
Lahr, N., 143
Lanczos, C., 225, 476-478
Lee, W. H. K., 51, 53-78, 115-137, 138-151, 165-189, 194, 203, 207-225, 243, 250, 251-280, 285, 295, 305, 315, 316-335, 337, 341, 343-344, 351-353, 358, 423, 461-478, 481, 484, 510, 512-526
Lester, R., 51
Lines, L. R., 476, 478
Litehiser, J. 51
Liu, C. C., 129
Lomax, A., 186-187, 359-382, 383-404
Louderback, G. D., 2, 51-52
Luk, F., 269, 280

McEvilly, T. V., 421, 423
McNally, K. C., 421, 423
Malcolm, M. A., 476-477
Marks, S., 335, 358
Matthews, M., 425-426
Meagher, K. L., 250, 315, 335, 358
Menke, W., 476, 478
Meyer, H., 519
Michael, A. J., 486, 510-511
Moler, C. B., 269, 280
Moler, C. B., 476-477
Mooney, W. D., 269, 280
Muller, C., 358
Murdock, J. N., 51
Murray, W., 225, 476-477

Newberry, J. T., 335, 358
Nolet, G., 476, 478
Novelo-Casanova, D., 185-186, 317-318, 335, 358

O'Connell, R. J., 469, 477
O'Neill, M. E., 51
Oppenheimer, D. H., 26, 48, 51-52,
420, 424-460

Pakiser, L. C., 51
Pao, Y. H., 280
Parker, D., 184-185
Pereyra, V., 268, 280, 478
Phillips, W. S., 351
Press, F., 3, 51-52

Radon, J., 461-462
Raleigh, C. B., 51
Reasenber, P. A., 405-423, 424-426,
460
Reid, H. F., 406, 423
Richards, P., 410-411, 414-415, 423, 426,
460
Richter, C. F., 281-283, 286-290, 292,
301, 313, 315-317, 324-325, 327-328,
335
Rodriguez, B. D., 510
Roentgen, W., 461
Rogers, J. A., 130, 137, 151, 517
Rubey, W. W., 51
Ryall, A. S., 335, 358

Santoza, F., 280
Sato, H., 333, 335, 351, 355, 358
Scarpa, R., 511
Scherbaum, F., 187-188, 514
Simpson, R. W., 460
Solov'ev, S. L., 326, 335
Spudich, P., 358
Stanley, W. D., 510
Stewart, S. W., 78, 130-131, 137, 225,
280, 335, 358, 423, 469, 476, 478,
512-513, 521
Stickney, M., 519
Swindell, W., 476, 478
Symes, W. W., 280

Tarantola, A., 476, 478
Tocher, D., 2
Thurber, C. H., 476, 478, 511
Tottingham, D. M. 130, 137-138, 151,
174-178, 335, 358, 513, 521
Turbitt, T., 519
Tsai, V., 527, 530
Tsumura, K., 326, 335

Udias, A., 423

Valdes, C. M., 121, 137, 181-182,
185-186, 250, 317-318, 335, 336-358
Van Schaack, J., 65, 100-114, 136
Vernon, F., 477

Walter, A., 48
Wallace, R. W., 51
Ward, P. L., 122-124, 137, 148, 172, 178-
179, 189, 517
Wennerberg, L., 317-318, 323, 330, 334
Wesson, R., 266-267, 280
Wright, M. H., 225, 476-477
Wu, R. S., 317-318, 335
Wyss, M., 316-317, 335

Zucca, J. J., 505-509, 510

SUBJECT INDEX

ACROSPIN, 172-173, 184-185
 A/D conversion, *see* analog to digital conversion
 Adjustment vector, 212-214, 218, 221, 222, 252, 472, 475
 Analog to digital conversion, 67-69, 126-129
 Apple Macintosh computer, 118
 ATLAB, 127-128
 Attenuation, 329-334, 351-357, 506
 Attenuation curves, 304-305
 Average residual method, 238

 Back-scattering of coda waves, 329, 351
 Bear Valley, CA experiment, 6-7, 16
 Benioff seismograph, 83
 Brune model of seismic source, 345-348

 Calaveras fault, CA, 9, 18, 452, 457
 California
 Dept. of Transportation (CalTran), 77
 Seismic networks, 3, 10-17, 22, 31, 42, 20, 22, 31, 38, 42, 102-104, 109, 281
 Seismicity maps, 12, 14, 18, 19
 Caltech Earthquake Detection and Recording (CEDAR), 15
 Central slice theorem, 464-466
 Coda Q, 329-334, 353-356
 Coda wave, **316-335, 353-356**
 Amplitude formulas, 323, 331
 Attenuation, 329-334, 351-357
 Chronology and references, 317-318
 Duration, 317-318, 320, 334, 337-338, 347
 Magnitude, 293-297, 305-314, 317-329
 Q_c , 331-334; *see also* Coda Q
 Common phases method, 238-239
 Component rotation, 360-363, 393-394
 Convolution, 372, 375-380, 400
 Correlation, 375-376

CUSP, 15, 43-45, 48, 121-122

Data

 Acquisition, *see* Data acquisition
 Analysis, *see* Data analysis/processing
 Format, 121-124
 Processing, *see* Data analysis/processing
 Data acquisition, **115-137**
 Analog to digital conversion, 126-129
 Computer hardware for, 116-118
 Computer software for, 124-125
 Desirable features, 125-126
 Display of data, 129-130
 General issues, 115-116
 Operation system for, 118-120
 Trigger algorithm, 130-132
 Data analysis/processing, **165-189, 190-206**
 Data flow, 168-169
 DEMUX, 176-177, 190
 Earthquake location, 179-183
 EXTRACT, 177-178
 FIXTIME, 174-175, 190
 HYPO2, 182
 HYPO71, 182-183
 HYPO71PC, 190, 197
 IASPEI Software for, 170-174
 Network data, 165-167
 PCEQ2, 181-182
 PICK1, 180-181
 PITSA, 185, 187-188
 QCODA2, 185-186
 QMAP, 183-185, 203-206
 SEISGRAM, 185-186, 383-404
 ST2ASC, 178-179
 STDESCR, 178-179
 SUDSMAN, 191
 SUDSPICK, 190, 191-194
 SUDSPLOT, 183-184, 195-203
 Waveform processing, 174-179, 185-188

- XPLAY, 175-176
- Data Translation, Inc., 126-129, 140-141
- Deconvolution, 400-401
- DEMUX, 176-177
- Derivative methods in optimization, 212-214
- Develco Discriminator, 100
- Develocorder, 2, 6, 47, 294, 325
- Disk Operating System, 60, 119-120, 140-142, 383
- DOS, *see* Disk Operating System
- Drum recorder, 55
- Duration, *see* Coda wave; Magnitude scale

- Earthquake location, **207-225, 226-250**
 - Accuracy, 228-240
 - Approximate, 218-219
 - Geiger's method, 219-224
 - HYPO71, 240-244
 - HYPO71PC, 182
 - HYPOELLOPSE, 245-249
 - Mathematics for, 207-225
 - Nonlinear least squares, 215-218
 - Nonlinear optimization, 210-214
 - Problem discussed, 208-210
 - Programs for, 226-250
 - Solution methods, 218-225
- Equations of motion; *see* Seismometers
- EXTRACT, 177-178
- Event detection, 33-34; *see also*
 - Trigger algorithm
- Event location, 133, 179-182; *see also* Earthquake location

- Fast Fourier transform (FFT), *see*
 - Fourier transform
- Fault plane solutions, *see* Focal mechanism
- Fermat's principle, 256
- Filtering, 372, 375-377
- FIXTIME, 71, 174-175
- Focal depths, 32-33, 229-230, 234-235
- Focal mechanism, **405-423, 424-460**
 - Fault plane solution, 418-422, 451, 453, 456-459
 - Cape Mendocino region, 441-448
 - FPFIT program, 424-460
 - Radiation pattern, 407-415
 - Spherical projection, 415-418
 - Stress regime, 407-408
 - Theory and history, 405-423
- Fourier transform, 345, 363-364, 371-375, 380, 401, 402, 466
- FPFIT, **424-460**
 - Background & philosophy, 424-426
 - Commands, 438-439
 - Confidence range, 431-437
 - Interpreting output, 445-459
 - Program operation, 438-445
 - Output & results, 439-445
 - Solution theory, 426-437
- FPPAGE, 448-449
- FPLOT, 448
- Frequency domain processing, 380

- Gauss-Newton method, 207-208
- Geiger's method, 207-208, 223
- Gradient vector, 211, 213-215, 217

- Hermitian interpolation, 272, 274
- Hessian matrix, 207, 211, 213-214, 217-218
- HYPO2, 182; *see also* HYPO71PC
- HYPO71, 182-183, 197, 240-244, 343-344
- HYPO71PC,
 - Program Inputs, 240-244
- HYPOELLIPSE, 245-249
 - Error ellipsoid, 248-249

- IASPEI, *see* International Association of Seismology and Physics of the Earth's Interior
- IASPEI Software, 170-172, 514
- International Association of Seismology and Physics of the Earth's Interior, 57, 73-75, 124, 170, 514
- IBM-compatible PC *see* Personal computer

- International Business Machines (IBM), 58-63
- International Seismological Observational Period, 174
- Inversion (teleseismic), 490-499
- IRIG-E, 175

- Jacobian matrix, 217, 221, 472
- Jeffrey's-Bullen model, 402-403
- Jeffrey's weighting, 242-243

- K-Factor, 333

- L-4 seismometer, *see* seismometers
- L-22 seismometer, *see* seismometers
- Lamont AH Format, 122
- Least squares (linear) method, 216-217, 219-222
- Location, *see* Earthquake location
- Loma Prieta earthquake, 204, 497-499
- Long Valley Caldera, CA, 500-504

- Macintosh (Mac) computer *see* Apple Macintosh computer
- Magnitude scale, 281-315
 - Amplitude, 285-293
 - Attenuation curves, 289, 304-305
 - Calculation of, 285, 288, 292-293
 - Duration, 293-297, 305-314, 317-329
 - Early developments, 282-283
 - Equations, 301, 307
 - F_{MAG}, 301; *see also* MF
 - Instrument response, 297-300, 302-303
 - Limitations of, 283-284
 - m_b , 317
 - M_D , *see* Duration
 - MF, 305-314
 - M_L , *see* Local scale
 - M_S , 317
 - M_W , 317
 - MX, 306-307
 - Nomogram, 292-293
 - Richter scale, 289, 324,
 - X_{MAG}, 301; *see also* MX

- Mark Products, 97
- Master event method, 237-238
- MDETECT, 138
- Mean removal, 360
- Medicine Lake Volcano, CA, 504-507
- Metawindow, 350
- Mini-computers, 56, 69, 117
- Model Vector, 220
- Mono Craters, CA, 501-504
- Morgan Hill, CA,
 - Earthquake, 450-454, 457-458
 - Velocity anomalies, 486
- Mt. Pinatubo, Philippines, 77
- Multiplexer 129
 - Ellis, 517
 - Quesig Electronics, 517-518, 525

- NCSN; *see* Northern California Seismic Network
- Networks, *see* seismic networks
- Newton's method, 211-214
- Nonlinear least squares method, 215-218
- Nonlinear optimization, 210-214
- Northern California Seismic Network, 11-12, 25-31, 38-40, 42

- Objective function, 215-216, 221
- OS/2 Operating System, 118-119

- P-picking, 180, 191-194, 394-399
- Parameter vector, 212-213, 222-223, 471-472, 474
- Parkfield, CA, 8, 16, 32, 35, 104-105, 113, 486
- Parkfield-Cholame, CA, Earthquake, 4-5, 32, 419
- PC *see* Personal computer
- PC-based seismic systems, *see* Seismic systems
- PCEQ, 171, 180-182, 185, 336-338, 341, 343, 347, 350-351
- PCEQ2, 181-182; *see also* PCEQ
- PC-Quake system, 513-526
 - 12-bit or 16-bit, 520

- Budget for, 521
- Field system, 526
- Hardware requirements, 515
- Heavy-duty system, 525
- Medium-duty system, 524
- Minimum single PC system, 522
- Power requirements, 520
- Shopping list for, 517-518
- Software requirements, 516
- Standard dual PC system, 523
- Suppliers for, 518
- Usable PCs for, 516
- User comments, 519
- Personal computer
 - IBM-compatible, 58-60, 62, 135
 - PC-based seismic systems, 53, 55-57, 59, 76, 101, 120, 168, 170, 190, 484-485, 512-526
- PICK1, 180-181
- PITSA, 185, 187-188
- Plotting
 - Maps, 190
 - Seismograms, 190
 - Waveform, 195
- PTPLOT, 440, 457
- QCODA, 171, 185-186, 317, 336-337, 351-352, 357
- QCODA2, 185-186; *see also* QCODA
- QMAP, 172, 183-185, 191, 203-206
- QPLOT, 440
- Quesig Electronics, 518
 - Multiplexer, 75, 517-518, 523-525
- Radiation pattern, 417, 419
 - Equation describing, 415
 - P-wave, 409, 414, 426
 - S-wave, 410
- Ray Equation, 256-257, 265, 268, 271-272
- Ray tracing, *see* Seismic ray tracing
- Real-time display of data, 129-130
- Real-time processor, 11
- Reid's theory of elastic rebound, 406
- Residual vector, 215, 221, 472
- Richter magnitude, *see* Magnitude scale
- Richter Nomogram, 292-293
- Root mean square (RMS) residual, 227-228
- Round Valley, CA, 353-354
- RTP, *see* Real-time processor
- S-picking, 193
- San Andreas fault, CA, 19, 36
- SCSN *see* Southern California Seismic
- SEISGRAM, 383-404
 - Display group, 385-387, 391, 398-399, 401
 - Equipment requirements, 383-384
 - Fourier transform, 401-402
 - Math menu, 400
 - Program operation, 384-388
 - Program options, 388-404
 - Pick, 394-399
 - Rotate, 393-394
 - Scale, 392-393
 - Travel time, 402-404
 - Window, 391-392
 - Zoom, 389-390
 - Program structure, 385
- Seismic data acquisition, *see* Data acquisition
- Seismic moment (M_0), 317, 347-348
- Seismic network, 2-3, 17, 512
 - Alaska, 234-237, 239-240, 247
 - Augmentation, 38
 - California *see* California
 - Cluster, 4, 27
 - Configuration, 31-32
 - Current status of (California), 25-31
 - Design and Implementation of, 20-25
 - Environmental, 11, 13, 15
 - Further development of (CA), 42-49
 - History of, 2-19
 - Hybrid, 24
 - Instrument Specifications, 105-114
 - Local earthquake, 4
 - PC-Based; *see* Personal Computer
 - Principal functions of (CA), 31-41

- Regional, 11, 13, 21, 25, 31-32, 42, 47
- Requirements, 34
- Seismic ray tracing, 251, 255, 264, 271, 273-274
- Seismic sensors, 53-55, 63, 65, 76, 105-107, 278
- Seismic systems, PC-based, 53-78
- Seismic Tomography, 461-478, 479-511
 - Applications of, 479-511
 - Central slice theorem, 464-466
 - Coalinga experiment, 488-497
 - Comparison with medical, 467-469
 - Experiments at USGS, 483-485
 - High resolution, 504-509
 - History of, 461-462
 - Local earthquake, 469-476, 485-499
 - Mathematics for, 461-478
 - Medical, 463-464
 - Simultaneous inversion, 469-476
 - Suggested readings, 476
 - Teleseismic, 500-504
- Seismograms
 - Analog, 337-338, 349
 - Archiving of, 47, 49
 - Broadband, 383
 - Calnet, 321-322, 325
 - Digital, 337-338, 340, 350-351, 383
 - Display, 365
 - Processing, *see* Data analysis/processing
 - Recovery of, 48
 - Seismic section, 365, 367-368
 - Synthetic, 376, 382
- Seismological Society of America, 57, 70 73-75, 124, 170
- Seismometers
 - Broadband, 64, 377, 396, 401
 - Calibration, 89-92
 - Development of, 80-85
 - Electromagnetic, 299
 - Equations of motion, 86-89
 - L-4, 84, 90, 95-96, 102, 106-107
 - L-22, 54, 102, 105
 - Long-period, 64
 - Short-period, 64, 154, 282
 - Theory & practice, 79-99
- SGRAM2, 186; *see also* SEISGRAM
- Shifting, 360-361
- Signal conditioning, 54, 103, 106
- Simple fault model, ray paths, 275-276
- Simulated annealing algorithm, 224
- Simultaneous inversion, 461, 468-469, 472, 475
- Single-scattering, 351
- Snell's Law, 258
- Southern California Seismic Network, 25-31, 36, 39-40, 42, 44-45, 48-49
- Spherical projections
 - Equal area projection, 415, 417, 420
 - Sterographic, 415
- SPIN, 172, 184
- Spline interpolation, 272, 274
- ST2ASC, 171, 174, 178-179
- STDESCR, 174, 178-179
- Steepest descent method, 207, 211-213
- Steronet, 370, 416, 442
- STPARSE, 138-139, 149
- SUDS, 122-124, 139, 148, 172, 174-178, 182, 186, 200
- SUDSMAN, 191
- SUDSPICK, 169, 171, 180, 190-194, 200
- SUDSPLOT, 172, 183, 190-191, 195-203
- SUN computer, 117
- Sutton, AK, earthquake, 236-237
- Taylor expansion, 207, 211, 213
- TDETECT, 152-164
 - Adjusting for performance, 160-164
 - Flowchart, 155
 - Input file, 160-164
 - Logic of, 154-159
 - Parameters, 160-163, 171
 - Program, 71, 131, 138-139
 - Trigger algorithm, 154, 156, 158-159, 161-164
- Telemetry, 3, 21-27, 31, 43-47, 53-54, 56, 65-67, 112-113, 333
 - Analog, 68, 73-74

- Basic techniques for, 100-114
- Cable, 66
- Digital, 75
- Microwave, 24, 26-27, 43, 45-46, 104
- Radio, 25-26, 43, 45-46, 54, 66-68, 76
 - 103-106, 109-111, 113
- Satellite, 66-67, 102-103
- Telephone, 54, 66, 68, 101, 103-104, 111
- Teleseisms, 152-154
- Time scaling, 360-361
- Time series
 - Differentiation of, 363-364
 - Integration of, 363-364, 400
- Tomography
 - Medical, 461, 463, 466-468, 476, 480-481
 - Seismic, *see* Seismic Tomography
- Travel time and derivatives, 251-280
 - Complex model, 264-275
 - One-dimensional model
 - Continuous velocity, 257-259
 - Multi-layer velocity, 259-264
 - Ray equation, 265-272
 - Ray tracing applications, 274-279
 - Three-dimensional model,
 - see* Complex model
 - Two-dimensional model,
 - see* Complex model
- Trigger algorithm, 130-132, 147, 168
- UNIX Operating System, 118
- VCO, *see* Voltage Controlled Oscillator
- Velocity model,
 - Complex (2-D, 3-D), 264-265
 - One-dimensional continuous, 257-259, 493
 - One-dimensional multi-layer, 259-264
- Voicegram, 153-154
- Voltage Controlled Oscillator, 106, 108-109, 113
- Wood-Anderson seismograph, 83, 285, 292, 297, 299, 302
- XDETECT, 138-151
 - AUTOEXEC.BAT, 142-143
 - CONFIG.SYS, 140, 143
 - Control parameters, 144-145
 - Daily log, 150
 - STATION.@@@, 149
 - Station definitions, 147-148
 - Trigger parameters, 146-147
- XPLAY, 175-176
- XPLAYV, *see* XPLAY
- Waveform analysis, 185-188, 359-382