2

Received by [illegible]

MAR 1 0 1987

LBL-22280

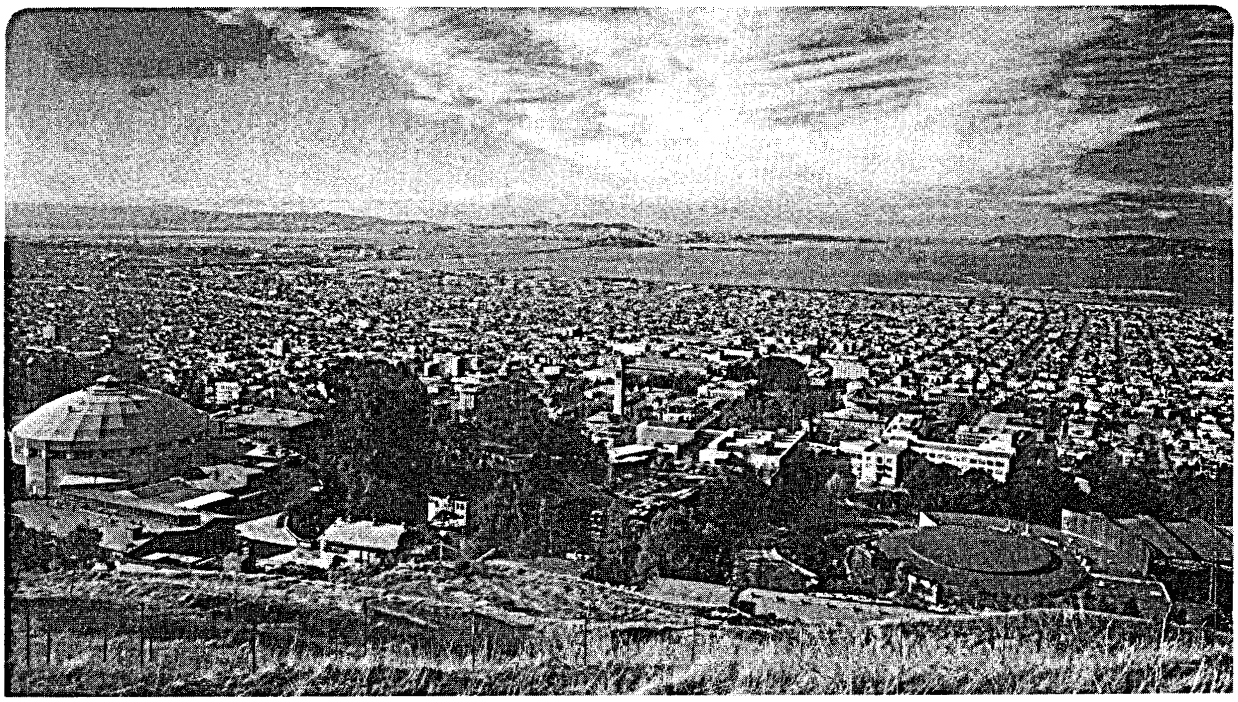
Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

EARTH SCIENCES DIVISION

SEISMIC VELOCITY STRUCTURE AND MICROEARTHQUAKE SOURCE PROPERTIES AT THE GEYSERS, CALIFORNIA, GEOTHERMAL AREA

D.R. O'Connell
(Ph.D. Thesis)

December 1986



Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098

DISCLAIMER

Seismic Velocity Structure and Microearthquake Source Properties at The Geysers, California, Geothermal Area

Copyright © 1986

LBL--22280
DE87 006170

Daniel Robert O'Connell

Department of Geology and Geophysics
The University of California at Berkeley
Berkeley, California 94720

and

Center for Computational Seismology
Earth Sciences Division
Lawrence Berkeley Laboratory
The University of California
Berkeley, California 94720

December 1986

This research was supported by Contract F19628-85-K-0025 of the Advanced Research Projects Agency of the Department of Defense and monitored by the Air Force Geophysics Laboratory and by The Director, Office of Basic Energy Sciences, Division of Engineering, Mathematics, and Geosciences, of the U. S. Department of Energy under contract DE-AC03-76SF00098.

The United States Department of Energy has the right to use this thesis for any purpose whatsoever including the right to reproduce all or any part thereof.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

The Government reserves for itself and others acting on its behalf a royalty free, nonexclusive, irrevocable, world-wide license for governmental purposes to publish, distribute, translate, duplicate, exhibit, and perform any such data copyrighted by the contractor.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Seismic Velocity Structure and Microearthquake Source Properties at The Geysers, California, Geothermal Area

Daniel Robert O'Connell

ABSTRACT

The method of progressive hypocenter-velocity inversion has been extended to incorporate S-wave arrival time data and to estimate S-wave velocities in addition to P-wave velocities. Synthetic tests demonstrate that the joint use of P and S-wave arrival time data has the following advantages over the use of P-wave data alone: (1) P-wave velocity and slowness gradient structure are more accurately estimated; (2) hypocenter mislocation errors are substantially reduced, especially hypocentral depth; (3) convergence of progressive inversions to local minima is more detectable using RMS data misfits of P and S-wave data; (4) velocity model and hypocenter estimates are more accurately determined when station corrections are used; (5) errors in linearized resolution and error estimated are reduced; and (6) complete elastic properties are estimated providing greater constraints for geologic interpretation of velocity structure. Adding S-wave data to progressive inversion does not completely eliminate hypocenter-velocity tradeoffs, but they are substantially reduced.

Results of a P and S-wave progressive hypocenter-velocity inversion at The Geysers show that the top of the steam reservoir is clearly defined by a large decrease of V_p/V_s at the condensation zone-production zone contact. The depth interval of maximum steam production coincides with minimum observed V_p/V_s, and V_p/V_s increases below the shallow primary production zone suggesting that reservoir rock becomes more fluid saturated.

The moment tensor inversion method was applied to three microearthquakes at The Geysers. Estimated principal stress orientations were comparable to those estimated using P-wave first motions as constraints. Well constrained principal stress orientations were obtained for one event for which the 17 P-first motions could not distinguish between normal-slip and strike-slip mechanisms. The moment tensor estimates of principal stress orientations were obtained using far fewer stations than required for first-motion focal mechanism solutions. The three focal mechanisms obtained here support the hypothesis that focal mechanisms are a function of depth at The Geysers.

Progressive inversion as developed here and the moment tensor inversion method provide a complete approach for determining earthquake locations, P and S-wave velocity structure, and earthquake source mechanisms.

In memory of my father

Daniel Edward O'Connell (1919-1981),

a man who loved life.

Acknowledgements

I am indebted to Dr. Lane. R. Johnson for his encouragement, support, and patience while allowing me to pursue my own course. His advice and example have been invaluable. I thank Dr. Thomas V. McEvilly for enthusiastic discussions and valuable advice and Dr. T. N. Narisimhan for his assistance.

My original interest in geophysics was sparked by Dr. Charles G. Bufe and I am grateful for the start he gave me and his continued support. I would like to thank Dave Oppenheimer for generously providing the resources and encouragement for the feasibility study that provided the groundwork that led to this dissertation as well as the USGS data used here and his continued interest and advice. I thank Dr. Paul Okubo for many helpful discussions.

I want to thank my fellow graduate students for their enthusiastic and open discussion of ideas. Although I cannot name everyone I would like to especially thank Jonathan Scheiner, Patty Murtha, John Aidun, Dave Tralli, Don Vasco, Phil Cummins, Jim Nelson, Michael Leonard, and Fred Eastwood for many interesting and helpful discussions.

All computations were carried out at the Center for Computational Seismology (CCS) of the Lawrence Berkeley Laboratory. I want to thank all the people who provided assistance at CCS and the U. C. Seismographic Station. The assistance of Dr. E. L. Majer was most appreciated. The outstanding facilities and personnel made this dissertation possible.

I feel especially blessed by my parents dedication to providing educational opportunities. It has been the continued encouragement and steadfast support of my family that made it possible to pursue my dreams.

Finally, I want to thank my wife Holly L. O. Huyck for her love and encouragement. In so many ways we have shared a common path and I feel blessed to have done so. I look forward to striving for knowledge and a more loving world with her.

Table of Contents

Chapter 1

Introduction

A wealth of information about geologic structure and ongoing tectonic processes are contained in seismic recordings of microearthquakes. Microearthquake locations can delineated the positions of active faults, microearthquake source mechanisms constrain the type and orientation of faulting, and estimates of seismic velocity structure provide information about geologic structure. Estimation of P and S-wave velocity structure provide strong constraints on material properties of geologic structures.

The problems of estimating microearthquake locations, source mechanisms, and seismic velocity structure are coupled. In order to estimate seismic source properties, accurate estimates of earthquake locations and material properties are required to calculate the complete medium response (Green functions) between sources and receivers. Accurate estimates of earthquake location require that the assumed seismic velocity structure is close to the truth. A recently developed approach to solving the coupled hypocenter-velocity inversion problem in the context of local earthquake data is progressive inversion. Pavlis and Booker (1980) and Pavlis (1982) developed the method of progressive inversion to locate earthquakes and estimate seismic velocity structure which explicitly accounts for the coupling between earthquake locations and assumed velocity structure. Progressive inversion was developed and tested for the case of P-wave arrival time data only. Since complete characterization of the seismic properties of a medium requires that S-wave velocity structure be known, the method of progressive inversion is extended to include S-wave arrival time data in Chapter 2.

Progressive inversion requires a linearization of a nonlinear problem. Pavlis and Booker (1983) conducted synthetic tests to determine the significance of nonlinearities on the problem of estimating seismic velocity structure. In Chapter 3, synthetic tests of progressive inversion are done to determine the effects of adding S-wave information to the problem. The effects of errors in starting velocity models on estimates of hypocenters, station corrections, and velocity structure are also investigated.

A long standing question is whether The Geysers geothermal field has a distinctive seismic signature. The results of Majer and McEvilly (1979) and Majer et al. (1986) suggest that the ratio of P-wave velocity to S-wave velocity, V_p/V_s, may be anomalously low at The Geysers. In Chapter 4 a progressive inversion with P and S-wave arrival time data from microearthquakes there is done to determine microearthquake locations, P and S-wave velocity structure, and V_p/V_s at The Geysers. Estimated microearthquake locations and V_p/V_s are compared to the locations of zones of steam production.

Studies of microearthquake source mechanisms have traditionally relied on P-wave first motion methods to constrain focal mechanisms and principal stress orientations associated with earthquake sources. These approaches require a large number of P-wave first motions to have the potential to constrain seismic source mechanisms. Oppenheimer (1986) found that P-wave first motion focal mechanism estimates at The Geysers were highly ambiguous for many earthquakes even though numerous P-wave first motions were available for all events considered. The moment tensor inversion approach developed by Stump and Johnson (1977) provides an alternative means to characterize seismic source mechanisms with a smaller number of data. Their method has not been previously used to estimate seismic moment tensors of microearthquakes.

Results of a progressive inversion, as developed in Chapter 2 and as applied to The Geysers in Chapter 4, provide the information required for reliably estimating seismic source properties; progressive inversion results can be used to generate the best estimate of Green functions. In Chapter 5, Stump and Johnson's (1977) method is used in conjunction with the progressive inversion results of Chapter 4, to estimate seismic moment tensors for three microearthquakes at The Geysers. The results are compared to those obtained using a P-wave first motion approach.

In addition to providing insights into geologic processes at The Geysers, this dissertation is intended to provide a complete approach for the utilization of three-component seismic recordings of microearthquakes. Taken together, the method of progressive inversion developed in Chapter 2 using P and S-wave data, and Stump and Johnson's (1977) method of moment inversion, provide a comprehensive approach to utilize microearthquake data such that the most complete and accurate information is obtained: microearthquake locations, P and S-wave velocity structure, and seismic source properties. This information in turn provides powerful constraints on ongoing tectonic processes and on material properties of geologic structures.

Chapter 2

Progressive Inversion with S-waves

2.1. Introduction

Two of the oldest outstanding problems in seismology are determining the hypocenter of an earthquake from seismic arrival-time data and determining seismic velocity structure from seismic arrival-time data. Often, these problems have been considered separately. It is clear however, that they are in fact coupled. Determination of an earthquake hypocenter from arrival-time data requires knowledge of seismic velocity structure. Much of the Earth's velocity structure has been deduced using earthquakes as sources. It is advantageous to solve these two problems together in a systematic approach. A recently developed approach to solving the coupled hypocenter-velocity problem in the context of local earthquake data is progressive inversion.

Progressive inversion is an approach developed by Pavlis and Booker, (1980) and Pavlis (1982) to determine hypocentral parameters, station corrections, and velocity structure using local event arrival times. The goals of progressive inversion are to improve hypocenter determination and to estimate velocity structure. The method was developed and tested using P-wave arrival time data only (Pavlis and Booker, 1980, Pavlis, 1982, and Pavlis and Booker, 1983a). The goal of this Chapter is to outline how to extend the method of progressive inversion to include S-wave arrival-time data. In Chapter 3, the effects of adding S-wave data to progressive inversion will be investigated using synthetic data.

There are several reasons to utilize S-wave information. Information contained in S-wave arrival-time data can help provide better constraints on event origin time and depth than P-wave data alone. In some cases, S-wave data will also help stabilize the process of progressive inversion for hypocenters and velocity structure. Thus, S-wave data can constitute a valuable addition to the joint problem of determining hypocenters and velocity structure.

Knowledge of S-velocities when combined with P-velocities provides tighter constraints on rock properties than does P-velocities alone. P-wave velocity, V_p, can be correlated with various rock types. But, interpretation of V_p in terms rock type can be highly ambiguous. In contrast, the ratio of P-velocity to S-velocity, V_p/V_s, has been found to help discriminate lithologies (Tatham, 1982). Also, V_p/V_s has been found to correlate with fracture density (Moos and Zoback, 1983), can be diagnostic of porosity variations (Domenico, 1984), and is sensitive to fluid and gas concentration (Tatham and Stoffa, 1976).

Another reason to use S-arrival time data is that S-waveforms contain information about the seismic source. One means of extracting seismic source properties is to use three-component waveform data in a moment tensor inversion (Stump and Johnson, 1977). However, in order to use three-component data to obtain moment tensor solutions, P and S-velocity structure is needed to calculate complete Green functions. Progressive inversion which includes S-arrival-time data provides a means to estimate both the P and S-velocity models required to calculate Green functions.

To summarize, motivations for including S-wave data are to improve hypocenter and velocity determination, obtain better constraints on rock properties, and to facilitate calculation of Green functions for moment tensor inversions. In this Chapter we outline how to extend the method of progressive inversion to include S-wave data. We begin with a review of the problem of earthquake location. Next, we will review progressive inversion as developed by Pavlis and Booker (1980), Pavlis (1982), and Pavlis and Booker (1983a). Then extensions and modifications to include S-wave data will be presented. Some aspects of our development differ from those of Pavlis and Booker. These differences will be pointed out and explained. Discussion of practical problems such as determination of S-wave arrival times is deferred to Chapter 4 which deals with real data.

2.2. Single-Event Location

Determination of earthquake location from seismic-wave arrival time data is, in general, a nonlinear problem. The problem can be stated as follows; we wish to minimize the objective function $F(\mathbf{x})$ given by

$$F(\mathbf{x}) = \mathbf{r}^{\mathrm{T}}\mathbf{r} \tag{2.1}$$

where $\mathbf{x}$ is the space-time vector of dimension ≤ 4 which describes a point seismic source, and $\mathbf{r}_i = \mathbf{T}^{obs_i} - \mathbf{T}^{cal_i}(\mathbf{x})$ is the residual vector of observed minus calculated travel times for $i = 1, m$ readings. The most common approach to the solution of (2.1) is Geiger's (1910) method of iterative least squares. Applying Geiger's method to local earthquakes requires a Taylor expansion of (2.1) about an initial estimate, $\mathbf{x}_0$, which gives

$$F(\mathbf{x}_0 + \delta\mathbf{x}) = F(\mathbf{x}_0) + \mathbf{g}^{\mathrm{T}}\delta\mathbf{x} + \frac{1}{2}\delta\mathbf{x}^{\mathrm{T}}\mathbf{H}\delta x + \cdots \tag{2.2}$$

where $\delta\mathbf{x}$ are the adjustments to the initial hypocenter, $\mathbf{g}$ is the gradient vector of $F(\mathbf{x}_0)$, and $\mathbf{H}$ is the Hessian matrix of $F(\mathbf{x}_0)$.

In Geiger's method, only first order terms in $\delta\mathbf{x}$ are used in an iterative search for the minimum of (2.1). Lee and Stewart (1981) point out that Geiger's method is an example of the Gauss-Newton method of optimization, which is known to be fallible. Thurber (1985) gives examples of some situations where Geiger's method fails. He shows that Newton's method, which retains the Hessian in (2.2), can give more stable results. For example, shallow earthquakes recorded only be nearby stations sometimes became "airquakes", that is, they locate above the Earth's surface, when using an earthquake location program based on Geiger's method. In synthetic tests, Thurber showed that use of Newton's method eliminated "airquake" location artifacts. It is widely known that linearized approaches to solving nonlinear problems can be unstable and produce erroneous results. However, virtually all approaches to earthquake location have used Geiger's method, usually with satisfactory success (Flinn, 1965; Bolt, 1960, 1970; Lee and Lahr, 1972, Buland, 1976; Klein, 1978). Other approachs are outlined by Lomnitz (1977), Tarantola and Valette (1982), and Thurber (1985).

In the context of local earthquake location, Geiger's method generally works well when locating earthquakes that occur inside seismic recording networks. Lee and Stewart (1981) review cases where the method performs poorly, for instance, when an earthquake is located well outside a recording network. Geiger's method will be used here, not because it is the best method in general, but because its form permits the use of linear projection operators which form the foundation for development of progressive inversion.

By keeping only first order terms in $\delta \mathbf{x}$, equation (2.2) can be rewritten in a different form which is more conducive to discussions of linear projection operators

$$\mathbf{A}\,\delta\mathbf{x} = \mathbf{r} \tag{2.3}$$

where $\mathbf{A}$ (related to (2.2) through $\mathbf{g} = -2\mathbf{A}^T\mathbf{r}$) is the matrix of partial derivatives of calculated travel time with respect to the Cartesian coordinates of the hypocenter and has the form

$$\mathbf{A} = \begin{bmatrix} \frac{\partial T_1}{\partial x} & \frac{\partial T_1}{\partial y} & \frac{\partial T_1}{\partial z} & 1 \\ \frac{\partial T_2}{\partial x} & \frac{\partial T_2}{\partial y} & \frac{\partial T_2}{\partial z} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial T_m}{\partial x} & \frac{\partial T_m}{\partial y} & \frac{\partial T_m}{\partial z} & 1 \end{bmatrix} \tag{2.4}$$

the partial derivatives being evaluated at the hypocenter for m data.

There are a variety of ways to solve (2.3). Our implementation, BERQLY written by L. R. Johnson, utilizes the Levenburg-Marquardt inverse which is usually defined as

$$\mathbf{A}^{*} = (\mathbf{A}^T\mathbf{A} + \lambda^2\mathbf{I})^{-1}\mathbf{A}^T \tag{2.5}$$

where λ is an adjustable constant referred to as the damping parameter. The matrix $\mathbf{A}$ is calculated using finite differences. For a particular iteration, an appropriate value for λ is found using the method of forced descent. This implementation has proven to be robust. Thurber (1985) noted that a damped version of Geiger's method was successful in avoiding the aforementioned "airquake" problem if the damping were chosen properly. BERQLY only produces "airquakes" when there are serious blunders in the input arrival-time data, input velocity model, or starting location.

The Levenburg-Marquardt inverse can be related to the method of singular value decomposition as a means of forming a generalized inverse. Singular value decomposition (SVD) of **A** can be written (Lawson and Hanson, 1974)

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{V}^{\mathrm{T}} \tag{2.6}$$

where **U** is an $m \times m$ orthogonal matrix, **V** is an $n \times n$ orthogonal matrix, and Λ is an $m \times n$ diagonal matrix containing the singular values of **A**. The generalized inverse of **A** in terms of its SVD is (Lanczos, 1961)

$$\mathbf{A}^{+} = \mathbf{V}\Lambda^{+}\mathbf{U}^{\mathrm{T}} \tag{2.7}$$

where $^{+}$ denotes matrix inverse and Λ^{+} contains $k \leq n$ nonzero reciprocal singular values. An alternative form for $\mathbf{A}^{m}$ that has a simple relation to the SVD of **A** is (Lawson and Hanson, 1974)

$$\mathbf{A}^{m} = \mathbf{V}\Lambda^{m}\mathbf{U}^{\mathrm{T}} \tag{2.8}$$

where the diagonal elements of Λ^{m} have the form

$$\frac{\Lambda_i}{\Lambda_i^2 + \lambda^2}$$

In spite of differences in their forms, $\mathbf{A}^{+}$ and $\mathbf{A}^{m}$ share a fundamental property; Pavlis (1982) showed that they use identical orthogonal projection operators **V** and $\mathbf{U}^{\mathrm{T}}$. This result makes it possible to use the separation technique developed by Pavlis and Booker (1980) to solve the problem of progressive multiple-event location and slowness inversion, the topic of the next section.

2.3. Multiple-Event Location and Slowness Inversion

The goal of progressive inversion is to determine hypocenters, station corrections, and velocity structure using seismic arrival-time data from multiple events. Suppose we have a set of arrival-time measurements for m_e earthquakes recorded by n stations. Some or all of these stations provide S-wave arrival times in addition to first P arrivals. The total number of stations in the context of station corrections is n_s, the number of stations in the network plus the number of those stations that record S-waves (i.e., record on three components). The total number of discrete slowness variables is N_s which corresponds to the combined totals from the P-wave and S-wave models. There are a total of M arrivals in the data set. The general problem has the form (Pavlis, 1982)

$$\mathbf{r} = \mathbf{A}\delta\mathbf{h} + \mathbf{S}\delta\mathbf{s} + <\mathbf{G} \mid \delta\mathbf{u}> + \mathbf{e} \tag{2.9}$$

where

$\mathbf{r} \in R^{M \times 1}$ = residual vector.

$\mathbf{A} \in R^{M \times 4m_e}$ = matrix of partial derivatives for all events (see (2.4) for single event form).

$\delta\mathbf{h} \in R^{4m_e \times 1}$ = vector of perturbations to the hypocenters.

$\mathbf{S} \in R^{M \times n_s}$ = matrix of partial derivatives for stations corrections.

$\delta\mathbf{s} \in R^{n_s \times 1}$ = vector of station correction perturbations.

$\mathbf{G} \in R^{M \times N_s}$ = matrix of discretized Frechet derivatives (kernels).

$\delta\mathbf{u} \in R^{N_s \times 1}$ = vector of perturbations to slowness model(s).

$\mathbf{e} \in R^{M \times 1}$ = column vector of observational standard errors

Subsequent forms of (2.9) are implicitly weighted by $\mathbf{e}^{-1}$. The notation $<\mathbf{G} \mid \delta\mathbf{u}>$ indicates the inner product (2.11). The term $<\mathbf{G} \mid \delta\mathbf{u}>$ represents the component of $\mathbf{r}$ due to slowness model errors. While $\mathbf{G}$ and $\delta\mathbf{u}$ are a discrete matrix and vector, respectively, during actual computations, the notation $<\mathbf{G} \mid \delta\mathbf{u}>$ emphasizes the fact that $\mathbf{G}$ and $\delta\mathbf{u}$ are conceptually functions.

The term $\mathbf{G}\delta\mathbf{u}$ is the discrete representation of a linearization based on Fermat's principle (Pavlis, 1982; Backus and Gilbert, 1969). Let an incremental travel time, $\delta\tau$, represent the component of $\mathbf{r}$, due to slowness model errors. $\delta\tau$ is related to slowness perturbations along the ray path between hypocenter, $\mathbf{h}$, and station, Δ_i, by

$$\delta\tau_i = \int_{h \to \Delta_i} \delta u \; ds = \int_{h \to \Delta_i} \delta u\,(x\,,y\,,z\,)ds\,(x\,,y\,,z\,) \tag{2.10}$$

where ds is incremental path length. In a general approach δu would be a function of three dimensions. In our implementation we assume the velocity varies with depth, z, only. Then (2.10) can be written

$$\delta\tau_i = \int_0^L G_i(z\,)\delta u\,(z\,)dz \tag{2.11}$$

where $G_i(z\,)$ depends on the source depth z_h and takes the form

$$G_i(z\,) = \begin{cases} \dfrac{u}{(u^2 - p_i{}^2)^{1/2}} & 0 < z \le z_h \\ \dfrac{2u}{(u^2 - p_i{}^2)^{1/2}} & z_h < z \le \text{depth of ray bottom } (u = p_i\,) \\ 0 & z > \text{depth of ray bottom} \end{cases} \tag{2.12}$$

where L is chosen in (2.11) as some depth below the bottom of the deepest rays.

Adding S-wave data increases the number of station corrections and the dimension of δu. The column dimensions of **S** and **G** increase accordingly. The addition of S-wave data has not changed the basic problem form from that derived for P-wave data by Pavlis (1982).

The most general problem form for progressive inversion is shown is (2.9). Situations may arise where the use of station corrections is not required and the terms involving **S** and δs can be omitted. If the term <**G** | δu> is omitted, (2.9) takes the problem form for PMEL, the progressive multiple event location method of Pavlis and Booker (1983b).

A more compact form for (2.9) is

$$\mathbf{r} = \mathbf{B}\mathbf{y} \tag{2.13}$$

where

$$\mathbf{B} = [\mathbf{A} \mid \mathbf{S} \mid <\mathbf{G} \mid]$$

$$\mathbf{y} = [\delta\mathbf{h} \mid \delta\mathbf{s} \mid \delta\mathbf{u}>]^{\mathrm{T}}$$

One approach to solving (2.9) is to invert for all the unknowns (δh, δs, and δu>)

simultaneously. For example, Crosson (1976) formulated a procedure that solved for hypocenter perturbations and velocity perturbations simultaneously; that is, all parameter perturbations are determined at once using a full matrix solution. He solved a matrix equation of the form of (2.13) (excluding station correction terms), in an iterative least squares approach. There are two reasons to avoid a simultaneous solution to (2.13). On the practical side, the matrix B can become quite large even for a moderate number of events thereby limiting the amount of data that can be used and the number of parameters that can be solved for. More importantly, Pavlis and Booker (1980) showed that the partitioned form of (2.9) could be exploited to partition the problem into independent parts using an "annulling transformation".

The solution to (2.9) by progressive inversion involves five steps

1. event location by the single event method described in Section (2.2), using the current estimate of the station corrections and slowness models,
2. calculation of an annulled version of the residuals unbiased by the hypocenter perturbations,
3. utilization of this annulled data to estimate a perturbation to the station corrections,
4. calculation of a second annulled version of the residuals unbiased by the station correction perturbations,
5. utilization of this second annulled set of data to estimate perturbations for the slowness model(s).

The first step was outlined in Section (2.2). Subsequent steps require the use of special properties of partitioned matrices and orthogonal projections, the "annulling transformation". The use of annulling transformations on partitioned matrix problems was developed by Pavlis and Booker (1980), Rodi *et al.* (1980), Spencer and Gubbins (1980), and Jordon and Sverdrup (1981). Spencer (1985) gives an overview of the use of partitioned matrices, and various projection operators that produce annulling transformations as applied to geophysical inverse problems. What follows is a brief outline of the basis for the annulling transformation and how it is used in a progressive process leading to a solution of (2.9).

2.4. The Basis for Progressive Inversion

2.4.1. Orthogonal Projection Operators and Overdetermined Least Squares.

We will limit our discussion to two orthogonal projection operators. A more complete discussion of orthogonal projection operators is given by Spencer (1985) and Pavlis (1982).

Consider the problem defined by (2.3) where $m > n$. Let the range space of the matrix $\mathbf{A}$ be represented by $R(\mathbf{A}) = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{y} = \mathbf{A}\delta\mathbf{x} \text{ for some } \delta\mathbf{x} \in \mathbf{R}^n\}$. Let the null space of the matrix $\mathbf{A}^T$ be represented by $N = (\mathbf{A}^T) = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{A}^T\mathbf{y} = 0\}$. Any vector in $R(\mathbf{A})$ is perpendicular to any other vector in $N(\mathbf{A}^T)$: they are orthogonal compliments and together they generate all of $\mathbf{R}^m$. Any vector $\mathbf{y} \in \mathbf{R}^m$ can always be decomposed into two orthogonal vectors

$$\mathbf{y} = \mathbf{y}_R + \mathbf{y}_N \tag{2.14}$$

where $\mathbf{y}_R \in \mathrm{R}(\mathbf{A})$ and $\mathbf{y}_N \in \mathrm{N}(\mathbf{A}^T)$ (Pavlis, 1982).

A generalized inverse solution (Penrose, 1955) of (2.3) will be a minimum-norm (Euclidean length) solution and will give a minimum sum of squares of residuals. One method of solution is to use a generalized inverse defined in terms of the SVD of A (2.7). Another method of solution could employ the QRK decomposition computed using Householder transformations (Lawson and Hanson, 1974)

$$\mathbf{A} = \mathbf{QRK}^T \tag{2.15}$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

here $\mathbf{R}_{11}$ is a $k \times k$ nonsingular triangular matrix, and $k \leq n$, is the rank of A. Q is an $m \times m$ orthogonal matrix and K is an $n \times n$ orthogonal matrix. The generalized inverse in terms of (2.15) which could be used to solve (2.3) is

$$A^{+} = KR^{-1}Q^{T} \tag{2.16}$$

Denote a generalized inverse solution vector to (2.3) as $\delta\hat{x}$. Let a solution be obtained using (2.7) or (2.16). In either case, from (2.14) the estimate of the solution will have two properties: firstly, the solution vector will have no projection on $(N(A^T))$, and secondly, the residual vector $e = r - A\delta\hat{x}$ will have no projection on $(R(A))$. These results mean that n pieces of information are extracted from the residuals in the least squares solution, and a linear combination of n pieces of data are mapped into $R(A)$. There are $m-n$ pieces of information left over in the residuals that are completely independent of the solution, and a linear combination of $m-n$ pieces of data are mapped into $N(A^T)$.[2.1]

How are the $m-n$ locally independent parts of the residual separated from the input residual vector? If (2.7) is used, multiply (2.3) by U^T where U^T is partitioned as

$$U^T = \begin{bmatrix} U_R^T \\ \hline U_N^T \end{bmatrix} \tag{2.17}$$

with $U_R^T \in R^{n\times m}$ and $U_N^T \in R^{(m-n)\times m}$. U_R^T is the orthogonal projector that projects data into $R(A)$. U_N^T is the orthogonal projector that projects data into $N(A^T)$. Operating on (2.3) with (2.17) gives

$$\begin{bmatrix} r_R \\ \hline r_N \end{bmatrix} = \begin{bmatrix} A_R \\ 0 \end{bmatrix} \begin{bmatrix} \delta x \end{bmatrix} \tag{2.18}$$

where

$$r_R \in R^{n}$$

$$A_R \in R^{n\times n}$$

and

$$r_N \in R^{m-n}$$

r_N are the independent data we seek. It is clear from (2.18) that U_N^T annihilates A. U_N^T maps

$m-n$ components of the original data into $\mathrm{N}(\mathbf{A}^{\mathrm{T}})$ ensuring their independence from the n components of the data used in $\mathrm{R}(\mathbf{A})$ to constrain the solution. These two results, obtained by Pavlis and Booker (1980) and Pavlis (1982), form the basis for their separation procedure.

The same results can be had using the $\mathbf{QRK}^{\mathrm{T}}$ decomposition represented by (2.15). In this case, $\mathbf{Q}^{\mathrm{T}}$ takes the place of $\mathbf{U}^{\mathrm{T}}$ in (2.17), and the same partion forms for $\mathbf{Q}^{\mathrm{T}}$ are obtained. The two projectors, $\mathbf{U}^{\mathrm{T}}$ and $\mathbf{Q}^{\mathrm{T}}$ are not identical, but they both satisfy the projection properties required to achieve parameter separation and data "annulment". Either projector, $\mathbf{U}^{\mathrm{T}}$, or $\mathbf{Q}^{\mathrm{T}}$, could be used in each step of the development of progressive inversion. We will use these two projectors in the same manner given in the development by Pavlis (1982). They will be the only projection operators used in the development of a solution to (2.9). The utility of orthogonal projectors will now be related to the problem of progressive inversion.

2.4.2. Applying Projection Operators to the Partitioned Matrices of the Multiple Event-Slowness Inversion Problem

To see how the annulling transformation works, we return to (2.9). The following derivation closely follows that of Pavlis (1982). First, (2.9) is expanded to show the contributions of individual events

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_{m_e} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_{m_e} \end{bmatrix} \begin{bmatrix} \delta\mathbf{h}_1 \\ \delta\mathbf{h}_2 \\ \vdots \\ \delta\mathbf{h}_{m_e} \end{bmatrix} + \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \\ \vdots \\ \mathbf{S}_{m_e} \end{bmatrix} \begin{bmatrix} \delta\mathbf{s} \end{bmatrix} + \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \vdots \\ \mathbf{G}_{m_e} \end{bmatrix} \begin{bmatrix} \delta\mathbf{u} \end{bmatrix} \tag{2.19}$$

where the subscripts refer to the contribution of each event of the total $i = 1, m_e$. Pavlis and Booker (1980) showed that each of partitions, $\mathbf{A}_i$, has an inverse SVD decomposition of the form (2.7). By using the individual event matrices, $\mathbf{U}_i^{\mathrm{T}}$, an orthogonal premultiplication matrix of the form

$$\begin{bmatrix} (\mathbf{U}_R^T)_1 & 0 & \cdots & 0 \\ 0 & (\mathbf{U}_R^T)_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (\mathbf{U}_R^T)_{m_e} \\ \hline (\mathbf{U}_N^T)_1 & 0 & \cdots & 0 \\ 0 & (\mathbf{U}_N^T)_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (\mathbf{U}_N^T)_{m_e} \end{bmatrix} \tag{2.20}$$

is used to premultiply (2.19) which yields the result

$$\begin{bmatrix} (\mathbf{r}_R)_1 \\ (\mathbf{r}_R)_2 \\ \vdots \\ (\mathbf{r}_R)_{m_e} \\ \hline (\mathbf{r}_N)_1 \\ (\mathbf{r}_N)_2 \\ \vdots \\ (\mathbf{r}_N)_{m_e} \end{bmatrix} = \begin{bmatrix} (\mathbf{A}_R)_1 & 0 & \cdots & 0 \\ 0 & (\mathbf{A}_R)_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (\mathbf{A}_R)_{m_e} \\ \hline 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \delta\mathbf{h}_1 \\ \delta\mathbf{h}_2 \\ \vdots \\ \delta\mathbf{h}_{m_e} \end{bmatrix} + \begin{bmatrix} (\mathbf{S}_R)_1 \\ (\mathbf{S}_R)_2 \\ \vdots \\ (\mathbf{S}_R)_{m_e} \\ \hline (\mathbf{S}_N)_1 \\ (\mathbf{S}_N)_2 \\ \vdots \\ (\mathbf{S}_N)_{m_e} \end{bmatrix} \begin{bmatrix} \delta\mathbf{s} \end{bmatrix} + \begin{bmatrix} (\mathbf{Y}_h)_1 \\ (\mathbf{Y}_h)_2 \\ \vdots \\ (\mathbf{Y}_h)_{m_e} \\ \hline (\mathbf{G}_N)_1 \\ (\mathbf{G}_N)_2 \\ \vdots \\ (\mathbf{G}_N)_{m_e} \end{bmatrix} \begin{bmatrix} \delta\mathbf{u} \end{bmatrix} \tag{2.21}$$

where

$$(\mathbf{r}_R)_i = (\mathbf{U}_R^T)_i \, \mathbf{r}_i \in R^4 \tag{2.22}$$

$$(\mathbf{A}_R)_i = (\mathbf{U}_R^T)_i \, \mathbf{A}_i \in R^{4\times 4} \tag{2.23}$$

$$(\mathbf{S}_R)_i = (\mathbf{U}_R^T)_i \, \mathbf{S}_i \in R^{4\times n_s} \tag{2.24}$$

$$(\mathbf{Y}_h)_i = (\mathbf{U}_R^T)_i \, \mathbf{G}_i \in R^{4\times N_g} \tag{2.25}$$

$$j = 1, 2, \ldots, m_e$$

and

$$(\mathbf{r}_N)_i = (\mathbf{U}_N^T)_i \, \mathbf{r}_i \in R^{m_i - 4} \tag{2.26}$$

$$(\mathbf{S}_N)_i = (\mathbf{U}_N^T)_i \, \mathbf{S}_i \in R^{m_i - 4\times n_s} \tag{2.27}$$

$$(\mathbf{G}_N)_i = (\mathbf{U}_N^T)_i \, \mathbf{G}_i \in R^{m_i - 4\times N_g} \tag{2.28}$$

$$j = 1, 2, \ldots, m_e$$

Multiplication of (2.19) by (2.20) has produced a lower partition of (2.21) where the data, $\mathbf{r}_N$, that are independent of the hypocenter perturbations. Condensing the lower partition of (2.20) into compact form gives the new reduced problem form

$$\mathbf{r}_N = \mathbf{S}_N \delta\mathbf{s} + <\mathbf{G}_N \mid \delta\mathbf{u}> \tag{2.29}$$

where

$$\mathbf{r}_N \in R^{M-4m_e} \tag{2.30}$$

$$\mathbf{S}_N \in R^{M-4m_e \times n_s} \tag{2.31}$$

$$\mathbf{G}_N \in R^{M-4m_e \times N_f} \tag{2.32}$$

The next step involves solving for the station correction perturbations. We defer discussion of a specific method to solve for the station correction perturbations. Here, we concentrate on the next step in the process, which is to reduce (2.29) to the slowness inversion problem. As part of the station correction solution process, an orthogonal projector is again used, this time to reduce (2.29) to the slowness inversion problem. Following Pavlis (1982), this is done by applying the orthogonal projector $\mathbf{Q}^T$ of the $\mathbf{QRK}^T$ decomposition (2.16) of $\mathbf{S}_N$ to (2.29), which gives

$$\begin{bmatrix} \mathbf{r}_s \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_s \\ 0 \end{bmatrix} \begin{bmatrix} \delta\mathbf{s} \end{bmatrix} + \begin{bmatrix} \mathbf{Y}_s \\ \mathbf{N} \end{bmatrix} \begin{bmatrix} \delta\mathbf{u} \end{bmatrix} \tag{2.33}$$

where

$$\mathbf{r}_s = \mathbf{Q}_R^T \mathbf{r}_N \in R^{n_s} \tag{2.34}$$

$$\mathbf{R}_s = \mathbf{Q}_R^T \mathbf{S}_N \in R^{n_s \times n_s} \tag{2.35}$$

$$\mathbf{Y}_s = \mathbf{Q}_R^T \mathbf{G}_N \in R^{4m_e + n_s \times N_f} \tag{2.36}$$

and

$$\mathbf{a} = \mathbf{Q}_N^T \mathbf{r}_N \in R^{M-4m_e - n_s} \tag{2.37}$$

$$\mathbf{N} = \mathbf{Q}_N^T \mathbf{G}_N \in R^{M-4m_e - n_s \times N_f} \tag{2.38}$$

Examination of the lower partition of (2.31) reveals that the data (2.37) are independent of the earthquake hypocenter perturbations and the station correction perturbations as are the kernels (2.38). Now we have all the components needed to form the slowness inversion problem. The slowness inversion problem can be written

$$\mathbf{a} = <\mathbf{N} \mid \delta\mathbf{u}> \tag{2.39}$$

At this juncture we can see how adding S-arrival times has altered the slowness inversion problem. One way to ascertain the affect of adding S-arrival times is to look at the form of the kernels **G** in (2.9) for a single event before annulment and after annulment.

Let $\mathbf{G}^P$ denote P-kernels and $\mathbf{G}^S$ denote S-kernels. Each of these kernels is discretized in depth z. The total number of discrete depths for $\mathbf{G}^P$ is N_P and for $\mathbf{G}^S$ is N_S. The slowness perturbations for the P-model are $\delta\mathbf{u}_P$ and those for the S-model are $\delta\mathbf{u}_S$. Then $N_s = N_P + N_S$. For the illustration of a single event here, let there be m_P P-arrival times and m_S S-arrival times with the m_P residuals ordered above the m_S residuals. Taking a single event from (2.21) we have

$$\begin{bmatrix} r_{1_P} \\ r_{2_P} \\ \vdots \\ r_{m_P} \\ \hline r_{1_S} \\ r_{2_S} \\ \vdots \\ r_{m_S} \end{bmatrix} = \cdots + \left[\begin{array}{cccc|cccc} G_1^P(z_0) & G_1^P(z_1) & \cdots & G_1^P(z_{N_P}) & 0 & 0 & \cdots & 0 \\ G_2^P(z_0) & G_2^P(z_1) & \cdots & G_2^P(z_{N_P}) & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ G_m^P(z_0) & G_m^P(z_1) & \cdots & G_m^P(z_{N_P}) & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & G_1^S(z_0) & G_1^S(z_1) & \cdots & G_1^S(z_{N_S}) \\ 0 & 0 & \cdots & 0 & G_2^S(z_0) & G_2^S(z_1) & \cdots & G_2^S(z_{N_S}) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & G_m^S(z_0) & G_m^S(z_1) & \cdots & G_m^S(z_{N_S}) \end{array}\right] \begin{bmatrix} \delta\mathbf{u}_P \\ \delta\mathbf{u}_S \end{bmatrix} \tag{2.40}$$

The hypocenter and station correction terms are omitted to achieve a compact form. Now do the annulling transformation by premultiplying (2.40) by (2.17) and then by $\mathbf{Q}^T$ of (2.33) which gives

$$\begin{bmatrix} (r_R)_1 \\ (r_R)_2 \\ (r_R)_3 \\ (r_R)_4 \\ \hline r_1 \\ r_2 \\ \vdots \\ r_{m_A} \end{bmatrix} = \cdots + \left[\begin{matrix} (Y^P)_1(z_0) & (Y^P)_1(z_1) & \cdots & (Y^P)_1(z_{N_P}) \\ (Y^P)_2(z_0) & (Y^P)_2(z_1) & \cdots & (Y^P)_2(z_{N_P}) \\ \cdot & \cdot & \cdot & \cdot \\ (Y^P)_4(z_0) & (Y^P)_4(z_1) & \cdots & (Y^P)_4(z_{N_P}) \\ \text{—} & \text{—} & \text{—} & \text{—} \\ (N^P)_1(z_0) & (N^P)_1(z_1) & \cdots & (N^P)_1(z_{N_P}) \\ (N^P)_2(z_0) & (N^P)_2(z_1) & \cdots & (N^P)_2(z_{N_P}) \\ \vdots & \vdots & \vdots & \vdots \\ (N^P)_{m_A}(z_0) & (N^P)_{m_A}(z_1) & \cdots & (N^P)_{m_A}(z_{N_P}) \end{matrix}\right.$$

(columns continued below)

$$\left.\begin{matrix} (Y^S)_1(z_0) & (Y^S)_1(z_1) & \cdots & (Y^S)_1(z_{N_S}) \\ (Y^S)_2(z_0) & (Y^S)_2(z_1) & \cdots & (Y^S)_2(z_{N_S}) \\ \cdot & \cdot & \cdot & \cdot \\ (Y^S)_4(z_0) & (Y^S)_4(z_1) & \cdots & (Y^S)_4(z_{N_S}) \\ \text{—} & \text{—} & \text{—} & \text{—} \\ (N^S)_1(z_0) & (N^S)_1(z_1) & \cdots & (N^S)_1(z_{N_S}) \\ (N^S)_2(z_0) & (N^S)_2(z_1) & \cdots & (N^S)_2(z_{N_S}) \\ \vdots & \vdots & \vdots & \vdots \\ (N^S)_{m_A}(z_0) & (N^S)_{m_A}(z_1) & \cdots & (N^S)_{m_A}(z_{N_S}) \end{matrix}\right] \begin{bmatrix} \delta \mathbf{u}_P \\ \delta \mathbf{u}_S \end{bmatrix} \tag{2.41}$$

where :

$$(r_R)_i \in R^4 = \text{data that constrain the hypocenter.} \tag{2.42}$$

$$(Y^P)(z_k) \in R^{4 \times N_P} = \text{incremental ray paths used by P-component of } (r_R)_i\,. \tag{2.43}$$

$$(Y^S)(z_k) \in R^{4 \times N_S} = \text{incremental ray paths used by S-component of } (r_R)_i\,. \tag{2.44}$$

$$i = 1,2,3,4c$$

and

$$r_j \in R^{m_A} = \text{annulled data.} \tag{2.45}$$

$$(N^P)(z_k) \in R^{m_A \times N_P} = \text{annulled P-component of data kernels.} \tag{2.46}$$

$$(N^S)(z_k) \in R^{m_A \times N_S} = \text{annulled S-component of data kernels.} \tag{2.47}$$

$$j = 1,2,\ldots,m_A \qquad k = 1,2,\ldots,(N_P \text{ or } N_S)$$

$$m_A = m_P + m_S$$

The annulled data r_j (2.45) are a linear combination of originally distinct P and S-residuals. Similarly, the P and S-kernels that were originally block diagonal (Golub and Plemmons, 1981) in (2.40) are coupled in (2.41). Just as the block diagonal form of **A** in (2.9) (see (2.19)) made it possible to treat each event separately, the elimination of block diagonal form in (2.41) necessitates *simultaneous* inversion for P and S-slowness perturbations.

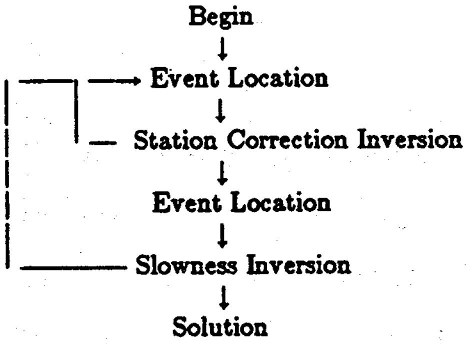
The annulled data in (2.45) can no longer be thought of as representing contributions from one wave type. Rather, they assume the character of a mixed body wave phase such as ScP. The use of mixed wave kernels is new but does not pose significant problems.

As with the station corrections, adding S-data has not profoundly changed the problem of estimating slowness perturbations. Some modifications are required in a P-only procedure to solve jointly for P and S-slowness perturbations. These modifications will be discussed as they arise.

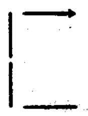
We have completed our discussion of how (2.9) is partitioned into three (locally) distinct problems.[2.2] The next step is to specify how we solve each of the three partitioned problems. We have already discussed how we solve the first problem, event location, in Section (2.2). Now, we will describe our particular approaches to solving for station correction perturbations and slowness perturbations.

2.5. Method of Solution

For the sake of clarity and simplicity several practical matrix operations were omitted from the discussion in Section (2.4.2). This was done to avoid obscuring the fundamental operations that are the essence of the separation process of progressive inversion. Since (2.9) is based on a linearization of a nonlinear problem its solution requires an iterative approach. The sequence of iterations used is shown schematically as



where



indicates a loop and ↓ indicates loop convergence.

This diagram shows that iterative solution to (2.9) is being done on three levels: (1) event by event location; (2) station correction estimation; and (3) slowness inversion. Iterative event by event location is an implicit loop and its solution was described in Section (2.2). Now, we will outline how we solve for the station correction perturbations and slowness model perturbations within each loop. We start with the station corrections.

2.5.1. Station Corrections

Much of our approach towards solving for station corrections perturbations is essentially the same as PMEL, Progressive Multiple Event Location, developed by Pavlis and Booker (1983b). They in turn, used ideas presented in Jordon and Sverdrup (1981), in the course of their development of PMEL. The addition of S-arrival times to the problem does not require significant changes in PMEL if only hypocenters and station corrections are sought. If an inversion for slowness models is included, some modifications to PMEL are required. We develop a solution for station correction perturbations taking into consideration that a slowness inversion will follow. As a result, the design philosophy differs somewhat from PMEL of Pavlis and Booker (1983b). Here, we present the essential elements of a solution algorithm based on parts of PMEL, highlighting significant differences as they arise.

Referring back to (2.21), the contributions of each event to the data and coefficient matrix for the station correction problem are found in (2.26) and (2.27), respectively. The routine SVDRS in Lawson and Hanson (1974) is used to form the products indicated in (2.26) and (2.27).

The final coefficient matrix (2.31) would attain the size $(M-4m_e) \times n_s$ if the contributions from each event (2.27) were accumulated. Pavlis and Booker (1983b) showed that the matrix $\mathbf{R}_s$ in (2.35) could be formed from the individual event partions (2.27) instead of $\mathbf{S}_N$ of (2.31) using sequential accumulation (algorithm SEQHT of Lawson and Hanson (1974)). The key point is that SEQHT uses a QR decomposition (see (2.15)) to reduce $\mathbf{S}_N$ to the upper triangular form $\mathbf{R}_s$. Instead of requiring $(M-4m_e) \times n_s$ rows of storage to solve for station correction perturbations, just $2n_s$ rows are required. This is done by applying a QR decomposition to each event in the manner outlined in Lawson and Hanson (1974). They show that the end result is equivalent to doing a QR decomposition on $\mathbf{S}_N$ of (2.31). This is why the matrix $\mathbf{Q}^T$ comes to be used in (2.33) to produce the annulled data and kernels for a subsequent slowness inversion.

From (2.33) the basic station correction perturbation problem becomes

$$\mathbf{r}_s = \mathbf{R}_s \delta\mathbf{s} \tag{2.48}$$

Pavlis and Booker (1983b) solve (2.48) using a pseudoinverse of $\mathbf{R}_s$, denoted $\mathbf{R}_s^+$ (see (2.7)), obtained by singular value decomposition.$^{2.3}$ They found that, in general, $\mathbf{R}_s^+$ has a pseudorank k of n_s-1. The pseudorank $k = n_s-1$ reflects the ambiguity produced by the tradeoff between earthquake origin times and the dc component of the station corrections. This ambiguity is inherent when only P-arrival times are used since P-arrival times do not put strong constraints on origin time. If only S-P times were used this source of ambiguity would be removed because origin time is eliminated from the problem. Using S-P times does not ensure that $\mathbf{R}_s$ will have a pseudorank of n_s. As pointed out in Pavlis and Booker (1983b), if one tries to jointly locate a set of earthquakes that occurred in a small cluster, one will generally find rank $(\mathbf{R}_s) \leq n_s-4$ or ($k \leq n_s-3$ if S-P data are used). This reflects the ambiguity of the absolute location of the cluster as well as the origin time. Often S-arrival times are only available for a subset of the stations with P-arrival times. In this case origin time must be included in the problem and the aforementioned origin time ambiguity persists. Consequently, as a practical matter, we expect that rank $(\mathbf{R}_s) < n_s$ in general, whether or not S-arrival times are included with P-arrival times, or S-P times are used.

There is a case when the dc station correction component can be constrained. Arrival-time data from explosions with known origin times can produce rank $\mathbf{R}_s = n_s$ thereby constraining the dc component of the station corrections. However, we prefer that the final station corrections have zero mean, independent of the constraints provided by the data. We will present a justification for this point of view shortly. First, we discuss the modifications to PMEL due to the addition of S-arrival times.

Since dc station corrections cannot be constrained in general, we impose the condition that the sum of the station corrections equal zero. If S-arrival times are used this condition is applied separately to each wave type's station corrections. During tests with synthetic data (to be discussed in Chapter 3), we found a tendency for errors in assumed V_p/V_s models to be compensated by the station corrections unless the dc component was separately calculated and removed from P-station corrections and S-station corrections. It is easy to see how this happens by considering a simple example.

Let us assume that the true P-station corrections and true S-station corrections have zero mean. Further, take the case when the assumed V_p/V_s model is in error and V_p/V_s is too large. Let the dc component be determined using combined P and S-station corrections and then remove this dc component from all P and S-station corrections. Since V_p/V_s is too large in the assumed model, S-P times will be too large. By adding a positive dc component to all the P-station corrections and a negative dc component to all the S-station corrections, S-P times are produced that satisfy the data. In addition, the sum of the dc P-station correction component and the dc S-station correction component is zero. The inversion process has produced an erroneous result. But if the dc components are calculated independently for each wave type's station corrections and then removed from the P and S-station corrections separately, the erroneous V_p/V_s ratio will not be incorporated into the station corrections.

Removing the dc station correction component separately from P and S-station corrections is essential to ensure the integrity of the annulled data (2.48) used to calculate perturbations to the slowness models. Since solving for station corrections is just one step in an overall procedure to locate earthquakes and determine slowness models, it is important that implementation of a solution for station corrections be consistent with the overall process. It may be argued that incorporating whole model V_p/V_s errors into the station corrections is acceptable because station corrections are meant to correct for deviations of the assumed slowness model from the true model. This point of view is unacceptable here because, in the context of our use, station corrections represent lateral velocity heterogeneity local to the station only. They assume the role of static corrections in reflection seismology (Dobrin, 1976). Consequently, we can place a bound on the largest possible station correction by specifying the maximum variation of near surface slowness and the maximum depth that constitutes the "near surface" zone.

In the next section we develop a slowness inversion that includes perturbations to the surface slowness. Therefore, requiring the respective P and S-station corrections to have zero mean is reasonable, since we determine the best least squares estimate of the near surface and surface slownesses. Station corrections should reflect the deviations from these mean surface slownesses.

In light of these requirements, we specify an upper bound on the absolute value that any single station correction can attain. We found this bound was difficult to satisfy using a psuedoinverse where the rank was decreased until the perturbations satisfied the bounds. However, a Levenburg-Marquardt solution (see (2.8)) was effective in satisfying our requirements. In synthetic tests (described in Chapter 3), we found that a damped solution retained the same pattern of perturbations that would be obtained using a pseudoinverse of pseudorank $n_s - 1$. By increasing the damping the magnitude of the perturbations was decreased, while preserving the "correct" pattern of perturbations. The drawback is that the generalized inverse with damping does not satisfy the symmetry conditions of Penrose (1955). Consequently, the overall form of the generalized inverse for PMEL will have a form slightly more complicated than found in

Pavlis and Booker (1983b). The utility of the Levenburg-Marquardt solution outweights this minor drawback.

For each iteration, a solution to (2.48) is obtained in the following way. First, a singular value decomposition of $\mathbf{R}_s$ is computed using routine SVDRS of Lawson and Hanson (1974). Next, the psuedorank k is determined from the SVD of $\mathbf{R}_s$ using a tolerance that requires that the smallest singular value be no smaller than $10^{\frac{-n}{2}}$ the largest singular value, where n is the total number of significant figures available to represent real numbers. A modification of the routine SVA from Lawson and Hanson (1974) is used to determine 21 values of damping λ uniformly spaced in $\log_{10} \lambda$ increments between $10\lambda_1$ and $\lambda_k/10$. Candidate solutions for are determined using the 21 damping values. Candidate perturbations are added to the previous values of the station corrections and the means are removed from each wave type's station corrections. The minimum value of λ is used that satisfies the bound constraints on station corrections.

As in PMEL of Pavlis and Booker (1983b), solution for station correction perturbations is part of an iterative process of solving (2.9). Once station correction perturbations are calculated, all events are relocated with the new station corrections. This process continues until: (1) subsequent station correction perturbations become small, i.e., when $\| \delta \mathbf{s} \| \leq \epsilon$ where ϵ is a small number, (2) the residual norm $\| \mathbf{r}_s \|$ does not decrease significantly in successive iterations, or (3) the residual norm $\| \mathbf{r}_s \|$ becomes smaller than the a priori standard error of the input data, in which case the solution to (2.9) is complete. Pavlis and Booker (1983b) give a theoretical justification for criteria (1) and suggest the use of an F test (Hoel, 1971, pp. 285-295) to detect case (2).

Pavlis and Booker (1983b) discuss the analysis of error through the resolution matrix, covariance matrix, and information distribution matrix. The basic results of their resolution analysis has been discussed in the context of station correction dc component ambiguity. The covariance matrix for the station corrections, $\mathbf{C}_s$, can be calculated from the SVD of $\mathbf{R}_s$ by the relation

$$\mathbf{C}_s = \mathbf{R}_s^{+}(\mathbf{R}_s^{+})^{T} \tag{2.49}$$

where $\mathbf{R}_s^{+}$ is the generalized inverse of $\mathbf{R}_s$. The complete covariance matrix for hypocenters and station corrections can now be written (Pavlis and Booker, 1983b)

$$\mathbf{C} = \begin{bmatrix} \mathbf{H}_A \mathbf{H}_A^T + \mathbf{HSC}_s \mathbf{S}^T \mathbf{H}_A^T & -\mathbf{HSC}_s \\ -\mathbf{C}_s \mathbf{S}^T \mathbf{H}^T & \mathbf{C}_s \end{bmatrix} \tag{2.50}$$

where $\mathbf{H}_A$ is the generalized inverse for all the events. They also show that when the number of arrival time data is much larger than the number of parameters, $4m_e + n_s$, the covariance matrix for the hypocenter and station corrections C has the form

$$\mathbf{C} \approx \begin{bmatrix} \mathbf{H}_A \mathbf{H}_A^T & 0 \\ 0 & \mathbf{C}_s \end{bmatrix} \tag{2.51}$$

A more complete discussion of resolution and errors for hypocenters and station corrections is deferred until Section (2.6.2).

Having outlined a means to solve for hypocenters and station corrections in a progressive iterative approach, we are ready to discuss the final step in each iteration of a solution to (2.9), solving for slowness perturbations.

2.5.2. Slowness Models

All information constraining slowness perturbations is contained in (2.39). We must decide how to utilize it. We choose to use a Backus and Gilbert approach; that is, we over-parameterize the slowness models. This ensures two things: firstly, the resulting slowness models will be free of artifacts of course parameterization, and, secondly, we can determine the type of averaging functions required to obtain acceptable standard errors in different portions of the model using well developed concepts of resolution and error analysis.

The kernels N in (2.39) are singular at the bottoming points of rays. The singularity precludes the application of a least squares approach because the error norms are not square integrable (Backus and Gilbert, 1969). As a practical matter, the kernels N are always finite due to finite discretization. Thus, a numerical least squares solution to (2.39) is possible. However, least squares solutions to (2.39) tend to be highly oscillatory. Rewritting (2.39) we have

$$a_i = \int_L^0 \mathrm{N}_i(z)\delta u(z)dz \qquad i=1,2,\ldots,M_n \tag{2.52}$$

The singularities in N can be removed by integrating (2.52) by parts, a process Backus (1970b) calls "quelling by integration". Integration of (2.52) by parts gives

$$a_i = \overline{\mathrm{N}}(0)_i\,\delta u(0) - \int_L^0 \overline{\mathrm{N}}(z)_i\,\delta' u(z)\,dz \tag{2.53}$$

where

$$\overline{\mathrm{N}}(z) = \int_L^z \mathrm{N}(\rho)d\rho \tag{2.54}$$

and we define

$$\delta' u(z) = \frac{d}{dz}(\delta u(z)) \tag{2.55}$$

The kernels in (2.54) are the total path lengths from z to the ray bottom. They are calculated using linear velocity (not slowness) gradients (see (A.18) of appendix A).

The term $\overline{\mathrm{N}}(0)_i\,\delta u(0)$ in (2.53) is always included in the problem for the following reasons. Oldenburg (1984) demonstrated that when $u(0)$ is incorrectly specified, profound errors are introduced into estimates of the model. Further, as stated in the previous section, we want to determine $u(0)$ in the best least squares sense so that station corrections will reflect local deviations from least square surface slowness and consequently will have zero means. Even if good estimates of surface slowness are available, the first term in (2.53) should be retained to allow adjustments to $u(0)$ dictated by the overall surface slowness as sampled by all the recording stations. Good estimates of $u(0)$ help ensure proper convergence to the true solution as shown in Oldenburg (1984). Pavlis (1982) and Pavlis and Booker (1980) remove the term $\overline{\mathrm{N}}(0)_i\,\delta u(0)$ from (2.53) by imposing a constraint that causes as estimate of the perturbations to the rest of the model to be independent of $u(0)$. As demonstrated in Oldenburg (1984), this approach is not valid unless the assigned surface slowness $u(0)$ is the true value.

Following Pavlis and Booker (1983a), Pavlis (1982), Kennett (1976), and Backus and Gilbert (1969), we choose to solve for the "flattest perturbation". Thus we minimize

$$\frac{1}{2}\int(\delta' u(z))^2\,dz + \frac{1}{2}\,\delta u^2(0) \tag{2.56}$$

subject to the constraints (2.53).

Pavlis and Booker (1983a), Pavlis (1982), and Kennett (1976) used Johnson and Gilbert's (1972) variational approach to solve the problem. The resulting set of normal equations are numerically ill-conditioned. In order to obtain a stable result, orthogonal transformations (Gilbert, 1971; Parker, 1977a) or damping (Pavlis, 1982) must be used.

An alternative approach is available that satisfies our requirements: a direct least squares solution that minimizes the quantity (from (2.53))

$$\| \mathbf{a} - (\overline{\mathbf{N}}(0)\delta u(0) + \int_0^L \overline{\mathbf{N}}(z)\delta' u(z)\,dz) \|^2 \tag{2.57}$$

using a generalized inverse (Wiggins, 1972). Remember that $\overline{\mathbf{N}}$ in (2.41) is discretized finely enough (in accordance with a Backus-Gilbert approach) that the problem defined by (2.57) is underdetermined; $\overline{\mathbf{N}}$ will always be rank deficient. If $\overline{\mathbf{N}}$ is not rank deficient we have not parameterized the model finely enough, violating the stated requirements that the slowness model be free of artifacts of parameterization. Thus, if the problem is formulated correctly, (2.57) will be rank deficient.

A pseudoinverse solution to a rank deficient problem has the property that the solution vector will be the (unique) minimum length solution (Lawson and Hanson, 1974). Consequently, a pseudoinverse solution to (2.57) satisfies the minimization conditions in (2.56) and the sum of the squares of the residuals defined by (2.57) will also be minimized. Since arrival time reading errors from local networks are best described by a Gaussian probability density (Buland, 1976), minimization of the norm in (2.57) is appropriate.

Lawson and Hanson (1974) show that the solution vector norm (2.56) is a nondecreasing function of pseudorank while the squared squared residual norm (2.57) is a decreasing function of pseudorank. Specifying a pseudorank involves a trade-off between minimizing (2.56) and (2.57). A solution that neither overfits nor underfits the data is obtained by proper specification of the pseudorank.

One practical advantage of this approach is that formation of the normal equations is not required. Lawson and Hanson (1974) show that forming normal equations squares the condition number of the equivalent problem solved using a direct solution approach. A direct generalized inverse least squares solution is better conditioned and requires half the significant figures required to obtain an equivalent solution of the normal equations.

A second consequence of a direct solution is that the kernels $\overline{\mathrm{N}}$ in (2.57) can be sequentially accumulated event by event using algorithm SEQHT of Lawson and Hanson (1974), thus eliminating the need for large matrix storage of the size $M_n \times N_g$. By using sequential accumulation of the kernels $\overline{\mathrm{N}}$, storage requirements are reduced to $(N_g + n_e) \times N_g$.

Using synthetic data we solved (2.57) using singular value decomposition pseudoinverses and found the resulting perturbations contained spurious oscillations and were very sensitive to pseudorank. We found that a Levenburg-Marquardt solution to (2.57) was much more satisfactory; oscillatory effects were greatly diminished and the pseudorank sensitivity was virtually eliminated. Since a Levenburg-Marquardt inverse (2.8) does not satisfy all the Penrose (1955) conditions for a generalized inverse it will now be shown that a Levenburg-Marquardt solution satisfies the minimization conditions on (2.56) and (2.57).

It has already be shown that a pseudoinverse with $k < N_g$ satisfies the minimization conditions on (2.56) and (2.57). For a particular pseudorank k let a true pseudoinverse solution vector norm be $\| \delta \mathbf{u}_k \|$. Marquardt (1970) showed that for a nonzero value of λ the damped solution vector norm $\| \delta \mathbf{u}_k \|_\lambda \leq \| \delta \mathbf{u}_k \|$. The norm of a damped solution vector is always less than the corresponding solution vector norm of a true pseudoinverse solution guaranteeing the satisfication of the minimization condition (2.56). From Lawson and Hanson (1974, pp. 193) (modified for the notation used here) we have:

Theorem 2.1

For a fixed nonnegative value of λ, say, $\bar{\lambda}$, let $\delta\bar{\mathbf{u}}$ be the solution for the problem (2.57) using the damped inverse of the form (2.8) for $\bar{\mathbf{N}}$. Then the residual norm obtained is the minimum value of (2.57) for all vectors $\delta\mathbf{u}$ satisfying $\| \delta\mathbf{u} \| \leq \| \delta\bar{\mathbf{u}} \|$.

Theorem 2.1 guarantees the minimization of (2.57).

Damped least squares is a robust approach to solving nonlinear least squares problems. This is due to its search properties that enable a reasonable descent direction to be found using damping even when a pseudoinverse fails to find a proper descent path. A damped solution allows us to bound the magnitude of perturbation solutions. This bound represents a bias or a priori information about the maximum size the perturbations should attain. Since the problem is nonlinear, placing bounds on perturbation size is sensible to provide reasonable step sizes.

In each iteration a solution is found as follows. First, a singular value decomposition of $\bar{\mathbf{N}}$ is calculated. Then, 21 damping values are determined using the method outlined in Section (2.5.1). A solution is chosen that produces reasonable solution and residual norms. Slowness perturbations are obtained by integrating the slowness gradient perturbation solution. Inversion for P-slownesses and S-slownesses is done separately, that is, various damping values are used to find acceptable P-slowness perturbations, then various damping values are used to find acceptable S-slowness perturbations. We do this because S-slownesses are larger than P slownesses so that different solution perturbation bounds are used for each wave type. Consequently, different values of damping will generally be needed to obtain solutions for each wave type. The value of damping used is that which produces optimal solution and residual norms: the largest projected decrease of residual norm that has an acceptable solution norm. If the new slowness models exhibit fluctuations on scales smaller than resolving lengths, the new models are smoothed. The new slowness models are used to start a new iteration.

Convergence is detected when either: (1) the root mean square (RMS) residual norm is $\leq$ the a priori standard error of the observed data, or (2) the RMS residual norm does not decrease significantly for several iterations. Criteria (1) requires reliable estimates of the standard errors of the observed data. Criteria (2) may indicate convergence to a local minimum. If (1) is not satisfied then the model is insufficiently parameterized to fit the data or the a priori standard errors of the data are incorrect. Since we use a slowness parameterization that varies only with depth, failure to satisfy (1) may indicate that significant lateral slowness variations exist. Reliable estimates of data quality are required to ascertain whether this is the case. Even then, only by using several different starting models can the problem of converging to local minima be investigated. As noted by Pavlis (1982), the problem of determining when convergence has occurred is a difficult problem that has thus far defied rigorous solution.

Once it has been determined that convergence has occurred the results must be appraised. This is the topic of the next section.

2.6. Resolution-Error Analysis

The analysis of resolution and error that will be presented in the next two sections is based on linearized solutions to nonlinear problems. The pitfalls of this approach have been investigated by Pavlis and Booker (1983a), by studying the problem using synthetic data. They found that linearized estimates of resolution can be unreliable when nonlinear effects are significant. However, they also found that solutions linearly close to the truth are obtained whenever the intrinsic resolution of the data is sufficient to resolve the structure of the true model. It is difficult to know which assessment applies to resolution-error analysis of inversions with real data. In Chapter 3 we address this problem by doing progressive inversions using various synthetic data sets. For the present, let it be kept in mind that the following discussions of resolution-error analysis are only strictly valid for truly linear problems.

2.6.1. Slowness Models

Backus and Gilbert (1967, 1968, and 1970) showed that estimates of seismic velocity are intrinsically nonunique. Constraints on model values are only determined from resolution-error trade-off analysis. The objective of resolution-error analysis is to determine averaging functions which produce model values with acceptable uncertainties and at the same time show what model features are resolvable.

We use methods outlined in Wiggins (1972) to assess resolution and error for slowness and slowness gradient. Wiggins' (1972) development is based on a discrete (matrix) representation of the data kernels. As mentioned in Section (2.4.2) and Section (2.5.2), P and S-models are coupled after data annulment because the kernels N consist of P and S-data kernels. We want to determine the degree of interplay between P and S-models. The resolving kernels defined in Wiggins (1972) facilitate direct determination of coupling between models. Methods of Backus-Gilbert model assessment as presented in Pavlis and Booker (1983a), Parker (1977a), and Johnson and Gilbert (1972) cannot be used because multiple models and parameter types indigenous to the kernels used here make specification of a physical misfit norm impossible.

Slowness resolution and error is of obvious interest. Slowness gradient resolution and error is also investigated because it is the quantity actually solved for in the course of estimating a slowness model. Resolving kernels for slowness gradient can show how different parts of a model trade-off, thereby alerting us to features that are artifacts of the inversion process. Assessment of slowness gradient resolution and error is important because slowness gradients have a more profound effect on the amplitudes of propagating waves than does slowness magnitude. Later, we will use estimated slowness models to calculate Green functions for source moment tensor inversion. Consequently, we want to assess resolvability and uncertainty of the input parameters that the Green functions are most sensitive to. Secondly, the slopes of earthquake travel time curves reflect slowness magnitudes while the changes in slope reflect slowness gradient magnitudes. In a synthetic test of Pavlis and Booker (1983a), a trade-off between velocity magnitudes and origin times allowed the same earthquake, placed a different depths in two different velocity models, to produce similar travel-time curves. The two velocity models had similar overall velocity gradients, only velocity magnitudes differed substantially. We suspect that slowness gradient is better resolved by earthquake arrival-time data than slowness.

When estimating slowness resolution and error the data kernels N are used[2.4]. The data kernels $\overline{\mathbf{N}}$ are required to determine slowness gradient resolution and error. Singular value decompositions of N and $\overline{\mathbf{N}}$, are used in subsequent resolution and error computations. Let the SVD of N be

$$\mathbf{N} = \mathbf{U}\Lambda\mathbf{V}^{\mathrm{T}} \tag{2.58}$$

where

$$\Lambda = diag(s_1, s_2, \ldots, s_{N_g})$$

s_i are the singular values and the dimensions of the orthogonal matrices, U and $\mathbf{V}^{\mathrm{T}}$ are $M_u \times M_u$, and $N_g \times N_g$, respectively. Let the SVD of $\overline{\mathbf{N}}$ have the same form as (2.58) except the orthogonal matrices, U, Λ, and V will be denoted as $\overline{\mathbf{U}}$, $\overline{\Lambda}$, and $\overline{\mathbf{V}}$ and the singular values as $\overline{s}_i$. In the following discussion of slowness resolution and error, the results for slowness gradient can be obtained by substitution of overbar equivalent quantities.

From (2.58) the resolution matrix $\mathbf{R}_k$ for slowness is

$$\mathbf{R}_k = \mathbf{V}_k \mathbf{V}_k^{\mathrm{T}} \tag{2.59}$$

where k is an expansion order. The expansion order of (2.59) refers to the number of eigenvectors retained in the calculation of $\mathbf{R}_k$. These eigenvectors correspond to the k largest singular values of N. The effect of statistical errors in the data on the variance of slowness perturbations is calculated from

$$\sigma_k^2(\delta u(z_i)) = \sum_{j=1}^{k} V_{ij}^2 / s_j^2 \tag{2.60}$$

V and s_i are as defined in (2.58). A trade-off between width of resolving kernels in (2.59) and variance in (2.60) exists as a function of k. Plots of standard error as a function of depth versus expansion order are used to determine the range of expansion orders that produce acceptable errors. Next, resolving kernels are plotted as a function of depth for each of the acceptable expansion orders The optimal expansion order is chosen as the one that produces the "best" combination of resolution and error.

The j th column of the resolution matrix represents the least-squares solution for maximizing the j th parameter (Wiggins, 1972). Each column of R represents a particular target depth. Each column of R is computed using the k that produces an acceptable standard error for that target depth. The expansion order is not, in general, the same for all columns of R. Inspection of the columns of R reveals the extent of parameter averaging and what parts of the model are being averaged. Columns of R are displayed in two ways. Both ways of displaying the resolving kernel use plots which make it possible to display all target depths and model depths simultaneously.

In one manner of presentation, each column of R is normalized to have unit area. An example of this kind of plot is shown in Figure (4.8) of Chapter 4. The target depths with the largest peaks represent the portions of the models that are best constrained by the data. These depths correspond to regions with the largest number of turning waves.

A second approach is to plot each column of R normalized to its maximum value. A plot of R of this type is shown if Figure (4.9) of Chapter 4. This plot shows how a velocity estimates at particular target depths trade-off with other portions of the velocity models. In this way the amount of smearing and coupling can be seen throughout the models. Sometimes the averaging kernels of R have significant side lobes. Side lobes provide useful information about inherent trade-offs in a problem. Consequently we have not used alternative approachs to attempt to minimize them. In fact, the interpretation of side lobes is essential in determining the coupling within and between P and S models. This is due to the fact that the annulling transformation produces mixed wave kernels for each annulled data point. Both sets of parameters must be included in the resolution analysis to determine how much each model trades-off with the other. A more complete discussion of the interpretation of the resolving kernels is deferred to Chapters 3 and 4.

The methods just described are demonstrated using synthetic data in Chapter 3 and using real data in Chapter 4. Practical aspects and results will be described there. Having outlined how to assess results of slowness inversions, our attention turns to the assessment of errors in the hypocenters and station corrections.

2.6.2. Assessment of Hypocenter and Station Correction Uncertainties

Pavlis and Booker (1980) and Pavlis (1982) has presented the only development, thus far, of a systematic means to estimate the combined effects of errors in arrival-time data and uncertainties in slowness model on estimates of earthquake locations in a progressive or joint inversion approach. Tarantola and Valette (1982) and Pavlis (1986) provide alternative approachs in the context of single event location. In most cases, seismologists have chosen to ignore the effect of slowness or velocity uncertainties on hypocenter determination or have at best, tried to account for these uncertainties in an ad hoc approach. Pavlis and Booker (1980) and Pavlis (1982) developed a means of making predictions about the magnitude of error in hypocenter and station correction estimates caused by velocity model errors using the Backus' (1970a,b,1971) method of generalized prediction. What follows is a modified development of their approach. Discussion of the method of Tarantola and Valette (1982) is contained in Appendix B.

Traditionally, seismologists have considered errors in observed arrival time data as the only source of errors in hypocentral estimates. Pavlis and Booker (1980) and Pavlis (1982) showed that there are two additional sources of error that must be considered: statistical errors in the assumed velocity model and model errors caused by unresolved velocity structure. These two additional errors follow naturally from Section (2.6.1) where they were discussed in the context of slowness model appraisal. Appraisal of errors in the discrete parameters, hypocenters and station corrections, requires a different interpretation than that used for continuous slowness models.

Earthquake hypocenters are constrained by observed arrival times that correspond to propagating waves traveling between the focus and seismic recorders. We would like to make a prediction of the error in estimates of travel time along these wave paths. Here, the earthquake location problem has been approximated as a linear process. In Section (2.4.1) it was demonstrated that n components ($\mathbf{Y}_h$ of (2.25)) of the m wave paths representing the observed data, are used to constrain a hypocenter. Consequently, we would like to have have an estimate of errors, due to uncertainties in the assumed velocity model, in the travel times of these n wave paths. The following discussion focuses on the problem of hypocenter appraisal; Station correction appraisal is done in the same manner by using the station correction kernels $\mathbf{Y}_s$ (2.36) in place of $\mathbf{Y}_h$.

At the outset it is important to note what errors are neglected in this analysis. An obvious source of error is the inherent nonlinearity of the problem which is neglected here. Nonlinearity means that wave path constraints cannot be simply mapped into n components as in (2.25). The matrix **A** (2.3) will be in error due to two sources: (1) nonlinearity and (2) errors in the partial derivatives that depend on the imperfectly known velocity model. Effects of model prediction errors on solutions obtained using **A** can be bounded (Lawson and Hanson, 1974, p. 41-52) but are beyond the scope of the work here. Pavlis (1986) has addressed the problem of nonlinearity by investigating the contribution of Hessian (see (2.2)). A full nonlinear appraisal is done in Appendix B where model prediction errors in travel times for all m wave paths will be considered.

The annulling transformations in (2.21) and (2.33) allow a simple outline of one way to look at discrete parameter error estimation. Let the hypocenter and station correction parameters be grouped together as $\mathbf{x}$ to allow a simpler and more compact presentation. After data annulment (2.21 and 2.33), we have

$$\mathbf{r}_R = \mathbf{A}_R \mathbf{x} + \langle \mathbf{Y} \mid \delta \mathbf{u} \rangle$$

Rearranging terms gives

$$\mathbf{r}_R - \langle \mathbf{Y} \mid \delta \mathbf{u} \rangle = \mathbf{A}_R \mathbf{x} \tag{2.61}$$

Although $\mathbf{x}$ is solved for using $\mathbf{r}_R$ (see 2.3), it is clear that we can estimate $\mathbf{x}$ from $\langle \mathbf{Y} \mid \delta \mathbf{u} \rangle$

also. This peculiarity provided motivation to develop the annulling transformation to allow independent estimation of slowness models (Pavlis and Booker, 1980). Unfortunately, $\mathbf{x}$ depends on $\mathbf{r}_R$ and $<\mathbf{Y} \mid \delta\mathbf{u}>$ reflecting the fundamental ambiguity of locating earthquakes. This form emphasizes that there are two sources of error in $\mathbf{x}$: observational errors in $\mathbf{r}_R$ and theoretical calculated travel time errors due to errors in the slowness models from the term $<\mathbf{Y} \mid \delta\mathbf{u}>$.

Let observational travel time reading errors be described by the covariance matrix $\mathbf{C}_t$ and theoretical travel time errors be described by the covariance matrix $\mathbf{C}_T$. Error contributions from $\mathbf{C}_t$ are straightforward to estimate (see 2.50 or 2.51). We use the development by Pavlis and Booker (1980) and Pavlis (1982) with several modifications to determine how to assess error contributions from $\mathbf{C}_T$.

We start by rewritting (2.13) is a slightly expanded form

$$\mathbf{r} = \mathbf{B}\mathbf{x} + <\mathbf{G} \mid \delta\mathbf{u}> \tag{2.62}$$

where $\mathbf{B}$ is the partial derivative matrix for all discrete parameters, $\mathbf{r}$ and $\mathbf{G}$ are as defined in (2.9). Let $\mathbf{H}_B \in R^{N \times M}$ be the generalized inverse used to estimate a solution to (2.62) for $\hat{\mathbf{x}}$. [2.5] A solution of (2.62) using $\mathbf{H}_B$ gives

$$\mathbf{R}_B\mathbf{x} = \hat{\mathbf{x}} + <\mathbf{X} \mid \delta\mathbf{u}> \tag{2.63}$$

where

$$\mathbf{X} = \mathbf{H}_B\mathbf{G} \in R^{N \times N_s} \tag{2.64}$$

$$\mathbf{R}_B = \mathbf{H}_B\mathbf{B} \in R^{N \times N} \tag{2.65}$$

$$N = 4m_e + n_s$$

$\mathbf{R}_B$ is the resolving matrix for the discrete parameters $\mathbf{x}$. The function $<\mathbf{X} \mid$ contains information about the ray paths used to constrain the discrete parameters. If $\hat{\mathbf{x}}$ consists of well constrained hypocenter parameters $\mathbf{R}_B \approx \mathbf{I}$. Pavlis (1982) showed that when station corrections are included in $\hat{\mathbf{x}}$, $\mathbf{R}_B \not\approx \mathbf{I}$ due to the tradeoff between station correction means and hypocenters. Since Pavlis (1982) considered only P-data, adding S-data mitigates this problem but does not eliminate it. This subject is addressed in Chapter 3 using synthetic data.

From (2.63) it is clear that estimates of $\delta\mathbf{x}$ are dependent on the slowness model through the term $<\mathbf{X} \mid \delta\mathbf{u}>$. Pavlis (1982) showed that (2.63) can be reduced to the form

$$\delta\mathbf{x} = \mathbf{X}\delta\mathbf{u} \tag{2.66}$$

using the substitution $\delta\mathbf{x} = \mathbf{R}_B\,\mathbf{x}-\hat{\mathbf{x}}$.

The N numbers δx_i can be estimated from a linear combination of the annulled data using Backus' generalized prediction procedure (Backus, 1970a,b, 1971). To get an estimate of $\delta\mathbf{x}$, we use (Pavlis, 1982)

$$\delta\hat{x}_i = \sum_{j=1}^{M_a} w_j\, a_j \tag{2.67}$$

where w_j are weights to be determined and a_j are the annulled data from (2.52). The kernels $<\mathbf{X} \mid$ in (2.66) are quelled by integration to ensure a finite error norm

$$\overline{X}_i(z) = \int_L^z X_i(\rho)d\rho \tag{2.68}$$

The model prediction error, $<\epsilon \mid$, is the difference between the kernels used to constrain the discrete parameters and predictions of these kernels using the annulled data,

$$<\epsilon_i \mid = \overline{\mathbf{X}}_i - \sum_{j=1}^{M_a} w_j\,\overline{\mathbf{N}}_j \tag{2.69}$$

where $<\epsilon_i \mid$ has a finite norm.

To determine the contribution of prediction error it is necessary to have a bound B on the derivative of deviations of the estimated model from the true true model such that

$$\| \frac{d}{dz}(\mathbf{u}_{true} - \hat{\mathbf{u}}) \| = \| \frac{d}{dz}(\delta\mathbf{u}_{true}) \| \leq B \tag{2.70}$$

where

$\mathbf{u}_{true}$ = true but unknown slowness function.

$\hat{\mathbf{u}}$ = final estimate of $\mathbf{u}_{true}$.

Pavlis (1982) uses B to bound the error, Δx_i, in x_i due to intrinsic slowness model nonuniqueness. Using Schwartz's inequality he specifies B such that

$$\Delta x_i \leq \parallel \epsilon_i \parallel B \tag{2.71}$$

where $\parallel \epsilon \parallel = \sqrt{<\epsilon \mid \epsilon>}$.

We cannot estimate B from the observed data since information about B is contained in a subspace orthogonal to the model subspace constrained by the data (Backus, 1970a). Thus B can represent a subjective bias about the range of fluctuations from the estimated model that we consider to be plausible.

A difficulty with Pavlis' (1982) approach is that B is also used as a damping parameter in calculations of prediction error and statistical error. The use of damping in resolution-error analysis causes overall errors to be underestimated. Nonzero values of damping overestimate resolution and underestimate slowness model statistical errors. We modify Pavlis' (1982) development to eliminate the use of B as a damping parameter by using an expansion order approach.

Pavlis (1982) showed that statistical errors in the annulled data used to estimate the slowness model produce errors in the discrete parameters of the form

$$\mathbf{C}_S = \mathbf{W}\mathbf{W}^{\mathrm{T}} \tag{2.72}$$

where $\mathbf{W}$ is the same as in (2.67) and (2.69). Pavlis (1982) used B as a damping parameter in the calculation of the weights $\mathbf{W}$ of the form

$$\mathbf{W}^{\mathrm{T}} = (\breve{\mathbf{N}} + \frac{1}{B^2}\mathbf{I})^{-1}\,\breve{\mathbf{X}} \tag{2.73}$$

where

$$\breve{\mathbf{N}} = \overline{\mathbf{N}\mathbf{N}^{\mathrm{T}}} \in R^{M_s \times M_s} \tag{2.74}$$

$$\breve{\mathbf{X}} = \overline{\mathbf{N}\mathbf{X}^{\mathrm{T}}} \in R^{M_s \times N}. \tag{2.75}$$

Here, the use of B as a damping parameter will be eliminated and an orthogonal expansion using a singular value decomposition of $\overline{\mathbf{N}}$ (2.58) will be substituted for $\hat{\mathbf{N}}$ in (2.73). Now (2.73) takes the form

$$\begin{aligned}\mathbf{W}_k^T &= (\overline{\mathbf{N}\mathbf{N}^T})^{-1}\,\overline{\mathbf{N}\mathbf{X}^T} \\ &= (\overline{\mathbf{U}\Lambda}_k^2\overline{\mathbf{U}^T})^{-1}\,\overline{\mathbf{U}\Lambda}_k\,\overline{\mathbf{V}}^T\overline{\mathbf{X}}^T \\ &= \overline{\mathbf{U}\Lambda}_k^{-1}\overline{\mathbf{V}^T\mathbf{X}^T}\end{aligned} \tag{2.76}$$

Now, using (2.76) ,$\mathbf{C}_S$ in (2.72) can be written

$$\begin{aligned}\mathbf{C}_S &= \mathbf{W}_k\,\mathbf{W}_k^T \\ &= \overline{\mathbf{X}\mathbf{V}\Lambda}_k^{-2}\overline{\mathbf{V}^T\mathbf{X}^T} \\ &= \overline{\mathbf{X}}\mathbf{C}_k^{s\prime}\,\overline{\mathbf{X}}^T\end{aligned} \tag{2.77}$$

where $\mathbf{C}_k^{s\prime}$ is the Covariance matrix for model slowness gradients. From (2.64), (2.77) can be rewritten as

$$\mathbf{C}_S = \mathbf{H}_B\,\overline{\mathbf{G}}\mathbf{C}_k^{s\prime}\,\overline{\mathbf{G}}^T\mathbf{H}_B^T \tag{2.78}$$

and we see that the term $\overline{\mathbf{G}}\mathbf{C}_k^{s\prime}\,\overline{\mathbf{G}}^T$ corresponds to an estimate of calculated travel time errors due to statistical errors in the slowness gradient models.

Using $\mathbf{W}_k$ (2.76) obtaining from the singular value decomposition of $\overline{\mathbf{N}}$, the model prediction error in (2.69) can be written

$$\begin{aligned}<\epsilon_k \mid &= \overline{\mathbf{X}} - \overline{\mathbf{X}\mathbf{V}}_k\,\overline{\mathbf{V}}_k^T \\ &= \overline{\mathbf{X}} - \overline{\mathbf{X}\mathbf{R}}_k \\ &= \overline{\mathbf{X}}(\mathbf{I} - \overline{\mathbf{R}}_k)\end{aligned} \tag{2.79}$$

where $\overline{\mathbf{R}}$ is the normalized resolving function for the slowness gradient.

The last form of (2.79) helps to demonstrate what $<\epsilon_k \mid$ is. The term $(\mathbf{I} - \overline{\mathbf{R}}_k)$ acts as a projector on $\overline{\mathbf{X}}$ and produces the component of $\overline{\mathbf{X}}$, $\overline{\mathbf{X}}^{\perp}$, that is orthogonal to the range of $\overline{\mathbf{N}}$

(2.54). As k decreases $(\mathbf{I} - \overline{\mathbf{R}}_k) \rightarrow \mathbf{I}$ and $\overline{\mathbf{X}}^{\perp}$ increases. Since $\overline{\mathbf{X}}^{\perp}$ maps from the subspace of the model unconstrained by the data, an estimate of the magnitude of unresolved deviations of the estimated model from the true model, such as B, must be specified to find the contribution of $\overline{\mathbf{X}}^{\perp}$ to the total parameter error.

In terms of projection properties it is now clear why damping should not be used to compute $<\epsilon\,|$ and $\mathbf{C}_S$. Nonzero value of damping overestimate the dimension of the range space of $\overline{\mathbf{N}}$ and consequently underestimate $<\epsilon\,|$. Nonzero damping also causes $\mathbf{C}_S$ to be underestimated by eliminating large error contributions of small singular values.

The total error in the discrete parameters due to errors in the slowness gradient model, $\mathbf{C}_T$, is the sum of the model prediction errors (2.79) and statistical errors (2.77) which gives

$$\mathbf{C}_T = <\epsilon_k\,|\,\epsilon_k> B^2 + \mathbf{C}_S \tag{2.80}$$

$$\mathbf{C}_T = \sigma_\beta^2 + \mathbf{C}_S$$

where σ_β^2 is the model prediction error component. The model prediction error will always be nonzero because the resolving function is never a unit delta function and B is always nonzero. It is clear that the contributions of the first and second terms in (2.80) trade-off with expansion order k and B. At this point the the bound B must be invoked to determine the prediction error contribution to $\mathbf{C}_T$. We now specify how B is interpreted and used.

It would be relatively straightforward to make a guess at the maximum possible variation of slowness. Typical values of slowness range from 1.0 to 0.125 sec/km (1.0 to 8.0 km/sec) from the Earth's surface to the bottom of the crust. We could safely assume that slowness deviations from the estimated model are no larger than 0.1-0.2 sec/km. However, B is a bound on the spatial derivative of slowness deviations (2.70). Consequently, it is difficult to confidently specify small values for B because the derivative in (2.70) could be very large. For example, take the common situation of two rocks with different slownesses in welded contact. Slowness is virtually discontinuous at the contact; the slowness changes rapidly over dimensions of the order of centimeters. Since we can confidently specify the magnitude of the largest plausible slowness variations, $\Delta\delta \mathbf{u}_{true}$, we could investigate the error contribution in (2.80) as a function of length scales Δs using (from 2.70)

$$\| \frac{d}{dz}(\delta \mathbf{u}_{true}) \| \approx \frac{\Delta\delta \mathbf{u}_{true}}{\Delta s} \leq B \tag{2.81}$$

This is essentially the approach taken by Pavlis (1982). One aspect that differs here is that the optimal expansion order must also be determined. This suggests that the following approach be used to determine the optimal combination of k and B.

A range of k are used to determine $\mathbf{C}_S$. Then, only those values of k that produce values of $\mathbf{C}_S$ smaller than $\mathbf{C}_t$, the conventional statistical error, are considered further. This condition forces us to consider nonzero values of B. Next, the remaining values of k are used to determine the expansion order that allows the maximum value of B such that $\mathbf{C}_T \leq \mathbf{C}_S$. By specifying a value for $\Delta\delta \mathbf{u}_{true}$, the resulting maximum B is used to determine the minimum Δs. If the minimum Δs is found to be very small, say on the order of millimeters or less, we can confidently conclude that model prediction errors have a negligible effect. If the resulting minimum Δs is larger, then an appropriate smaller value of Δs must be specified. The resulting larger value of B could be used to calculate $\mathbf{C}_T$, and added to $\mathbf{C}_S$ to give the total prediction error component of hypocenter error. However, it has been our experience that model error contributions are overestimated using this approach. This is not surprising since Parker (1977b) has noted that bounds, such as B, tend to produce overly pessimistic error estimates. We can demonstrate that this is indeed the case here.

To determine the errors due to unresolved slowness structure, travel times are calculated for the two models shown in Figure (2.1). The "true" model has large fluctuations of slowness and the "test" model represents a smooth gradient average to the true model. Travel times are investigated because, inspection of the form of (2.64) shows that our predictions of model induced errors are fundamentally related to the problem of predicting errors in travel times calculated from the estimated model. Pavlis (1982) developed an expression for B that relates B to slowness deviations and their length scales

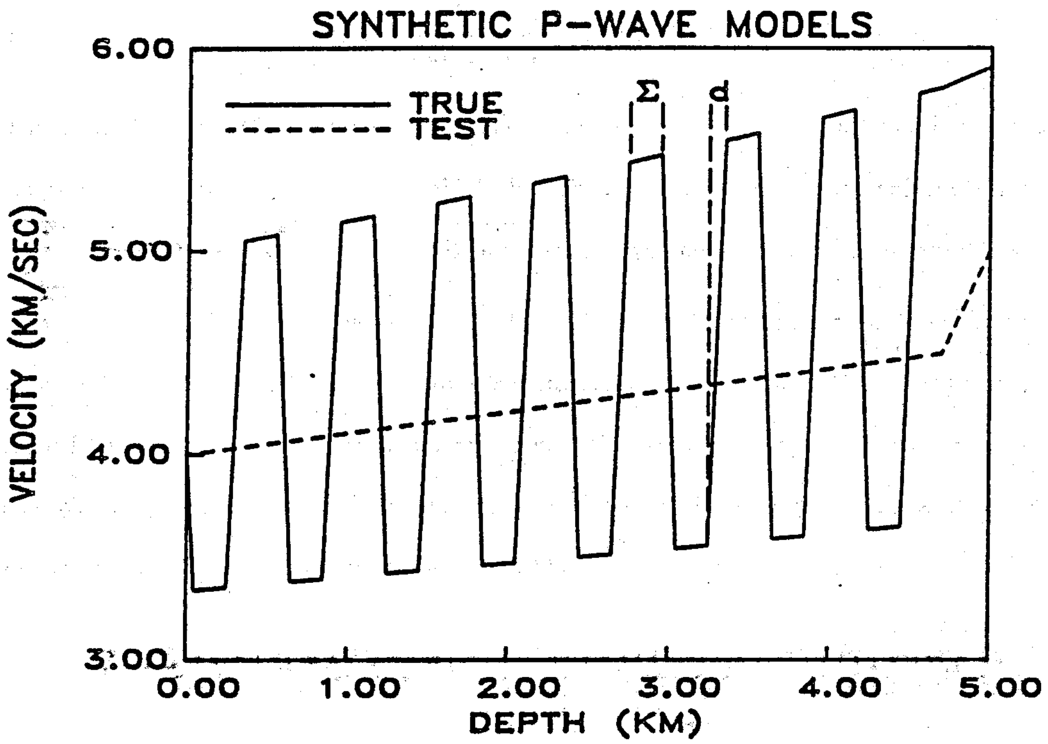


Figure 2.1 Contrast of a true synthetic model (solid line) and a smooth model approximation to it (dashed line). The large velocity fluctuations of the true model are produced by (see (2.82)) $\Delta u = 0.05$ sec/km, $\Sigma = 0.2$ km, $d = 0.05$ km, and $L = 5.0$ km.

$$B \approx \frac{\Delta u \sqrt{L}}{\Sigma \sqrt{d}} \tag{2.82}$$

where L is the depth dimension of the slowness model, Δu is as defined above, Σ is the depth duration of segments of unresolved slowness variations with amplitude Δu, and d is distance over which slowness changes from true model values to the amplitude variation of Δu. This definition of B is used to estimate how close bound estimates of travel time errors are to true errors.

The true travel time error is calculated as the difference between travel times calculated for the true model and test model. These travel time errors are compared to estimates of the travel time error obtained using B and an estimate of $<\epsilon|$. We use $\| \overline{\mathbf{G}}_{true} - \overline{\mathbf{G}}_{test} \|$ as an estimate of $<\epsilon|$ since $\| \overline{\mathbf{G}}_{true} - \overline{\mathbf{G}}_{test} \|$ is the norm of path length errors between the true and test models. Results are shown in Figure (2.2) for a source at 2.2 km depth with Δu =0.05, Σ=0.2 km, d =0.05 km, and L =5.0 km corresponding to B =2.5 (Figure (2.1)). The bound estimates of travel time error have been divided by 100 to plot on the same scale. It is clear from Figure (2.2) that for nonvertically propagating rays, the bound estimate of error is several orders of magnitude too large. Distances of 10 to 12 km correspond to rays that leave the source almost horizontally. The bound estimate of travel time error becomes increasingly overly pessimistic as ray paths begin to have turning points. While the magnitude of bound errors is much to large, the trend of errors with increasing distance (takeoff angle) is approximately correct. It makes intuitive sense that the bound estimate of travel time error will overestimate the true error by the greatest amount for turning rays.

In light of these results, it is clear that absolute hypocenter error due to model errors cannot be reliably determined using B. The fact that B is intrinsically unknown anyway has already precluded estimating absolute errors. However, the relative importance of model errors can be assessed between individual hypocenters and station corrections. Our previous proposed method of determining optimal values of k and B for each discrete parameter is undermined by the gross overestimation of the contribution of prediction error, σ_P^2. Erroneously large values of k are required to produce the minimum sum of $\mathbf{C}_S$ and σ_P^2. Because the results of this kind of

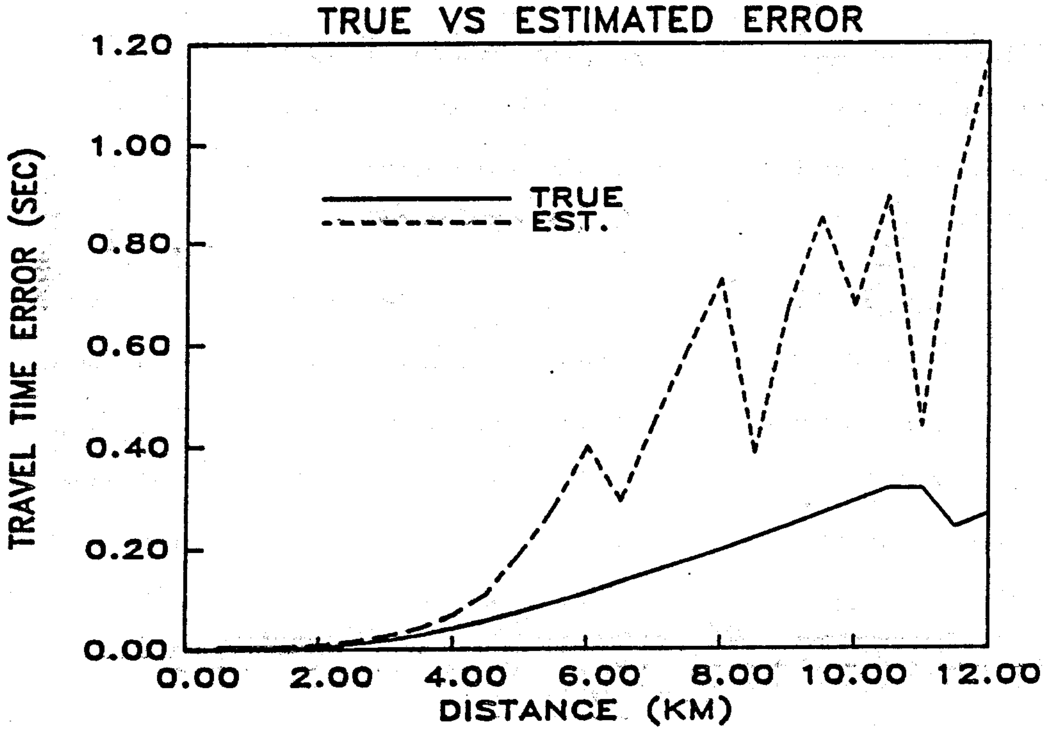


Figure 2.2 Comparison of the absolute values of true travel time errors (solid line) and bound estimates of the same travel time errors (dashed line). The bound error estimates have been divided by 100 to plot at this scale.

approach defy meaningful interpretation, a different tack was taken.

It is desirable to be able to interpret hypocenter position with respect to features in the slowness and slowness gradient models. This suggests that value of expansion order used for discrete parameter error assessment be set to the expansion order used for the slowness gradient model that produced the best combination of resolution and error. In this way, hypocenter position can be assessed in the context of resolvable features in the slowness gradient models. This is the approach we finally chose. The value most representative value of k used in slowness gradient resolution and error is used to calculate $\mathbf{C}_S$. Then, for each hypocenter parameter and each station correction, minimum values of Δz are calculated four ways.[2.6] Four values of Δz are found such that: (1) $\sigma_\beta^2 = \sigma_t^2$, (2) $\sigma_\beta^2 = \sigma_S^2$, (3) $\sigma_\beta^2 = \sigma_t^2 + \sigma_S^2$, and (4) $\sigma_\beta^2 = constant$. By comparing the sizes of the resulting values of Δz, for cases (1), (2), and (3) the relative importance of model errors can be assessed between station corrections and between spatial parameters for individual hypocenters. Bound test (4) is the most useful for comparing the sensitivity of parameters from different hypocenters since a constant error magnitude is used for all events. Parameters with large test (4) values of Δz relative to other hypocenters are more likely to be influenced by model errors.

The total discrete parameter covariance $\mathbf{C}_z$ is the sum of (2.80) and the contribution of statistical errors in data used to constrain the discrete parameters which gives

$$\mathbf{C}_z = \mathbf{C}_t + \mathbf{C}_S + O(\sigma_\beta^2) \tag{2.83}$$

If the $\mathbf{C}_t$ is dominant in (2.83) then the conventional approach of ignoring model-induced errors is appropriate. If $\mathbf{C}_S$ is large, then model induce errors are significant and we must be careful to include an estimate of model-induced errors in assessing discrete parameter errors. The important result is that (2.83) makes it possible to make an assessment of how different sources of error effect the final estimate of the discrete parameters.

This concludes our discussion of parameter error assessment. We would like to determine how well the approach works in practice. Tests of the method are done in Chapter 3 using synthetic data. In Appendix B the effects of nonlinearity on parameter error estimates are assessed by comparing results of the linearized approach outlined here against results obtained using a full nonlinear approach provided by Tarantola and Valette (1982)

2.7. Summary and Conclusions

A method of progressive hypocenter-velocity inversion has been developed to incorporate S-wave arrival time data and estimate S-wave velocities in addition to P-wave velocities. The ability to utilize S-wave data improves constraints on hypocenters and provides more complete information on seismic structure than P-wave data alone. The growing amount of three-component seismic data being collected by the seismological community ensures that immediate benefits will be realized by the new capability to incorporate S-wave data in a progressive hypocenter-velocity inversion.

The effect of slowness model errors on hypocenters and station corrections can be calculated using a modification of Pavlis and Booker's (1980) and Pavlis' (1982) approach. Their approach has been modified to eliminate damping from the error appraisal process. Contributions of model prediction errors to hypocenter and station correction errors are assessed with respect to a priori bounds on unresolved slowness model errors. The relative importance of statistical errors and model prediction errors on discrete parameter error is determined with respect to features in estimated slowness gradient models.

The one-dimensional velocity model assumption used here is not an inherent limitation of the development. The process developed here for estimating hypocenters, velocity models, and estimating their errors can be generalized to solve for three-dimensional velocity models if resources are available to solve the three dimensional travel time and wave-path problems. The three-dimensional velocity-hypocenter inversion method with data annulment developed by Thurber (1983) for P-wave data could be extended to include S-wave data using the ideas outlined here.

Notes

[2.1] These results are only completely true for linear problems. Multiple event - slowness inversion is inherently a non-linear problem. Consequently, while the results here have proven useful (Pavlis and Booker, 1983a; Pavlis, 1982), this limitation should be kept in mind.

[2.2] The problems are all coupled by the inherent nonlinearity ignored in (2.9). Further, Pavlis (1982) and Pavlis and Booker (1983b) showed that while the station correction perturbations are (locally) independent of the hypocenters, the hypocenters are inherently coupled to estimates of the station corrections. Consequently, a solution to (2.9) must be iterative to reduce the effect of coupling between hypocenters, station corrections, and slowness models.

[2.3] From Lawson and Hanson (1974, pp. 77-78), "the pseudorank k of a matrix A is the rank of the rank deficient matrix $\tilde{A}$ that replaces A as the result of a specific computational algorithm. Note that the pseudorank is not a unique property of the matrix A but also depends on other factors, such as the details of the computational algorithm, the value of tolerance parameters used in the computation, and effects of machine round-off errors." For example, determining k requires specifying which small singular values should be set to zero, an approach requiring some degree of subjective judgement. Lawson and Hanson (1974, pp. 180-206) outline approaches to make the decisions as objective as possible.

[2.4] The discrete data kernels N are finite at turning points only because N was approximated by incremental path lengths for a velocities model parameterized by linear velocity gradient layers (see appendix A, equation (A.18)). Consequently, subsequent slowness (velocity) resolution and error is dependent on model parameterization and the results are not expected to be completely reliable in an absolute sense. Rather, this approach helps to indicate which portions of models are most sensitive to errors in the data and shows patterns of trade-offs within and between models due to inadequate constraints provide by the data. These results are not used to support arguments that a particular feature exists and has acceptable error; they help to highlight features that should be viewed with skepticism due to inherent trade-offs or instabilities. This approach has been used frequently (Crosson, 1976; Aki and Lee, 1976; Spencer and Gubbins, 1980; Thurber, 1983) and could be called "quelling by parameterization", not a particularly rigorous approach, but one that has proved useful.

The data kernels for slowness gradient are finite (2.54) which makes them amenable for resolution and error analysis. However, progressive inversion is a nonlinear process; resolution and error estimates are only reliable if the final model estimate is linearly close to the true model.

[2.5] $\mathbf{H}_B$ is never explicitly formed. Solution for each hypocenter and for the station corrections, is done separately using annulling transformations. Pavlis (1982) showed that the final result can be viewed as a single operation by $\mathbf{H}_B$.

[2.6] The substitution of Δz is made for B using (2.81). Since an estimate of $\Delta\delta\mathbf{u}_{true}$ is used in the analysis, (2.81) allows us to solve for the minimum length scale, Δz that produces a particular error magnitude. If we find that Δz is indeed very small for a particular discrete parameter, we can have some confidence that that parameter is not significantly effected by unknown slowness gradient model errors.

References

Aki, K., and Lee, W. H. K., 1976. Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes, 1, A homogeneous initial model, *J. Geophys. Res.*, **81**, 4381-4399.

Backus, G. and F. Gilbert, 1968. The resolving power of gross earth data, *Geophys. J. R. astr. Soc.*, **16**, 169-205.

Backus, G. and F. Gilbert, 1969. Constructing P-velocity models to fit restricted sets of travel-time data, *Bull. Seism. Soc. Am.*, **59**, 1407-1414.

Backus, G. and F. Gilbert, 1970. uniqueness in the inversion of inaccurate gross earth data, *Phil. Trans. Roy. Soc. London, Ser. A*, **266**, 123-192.

Backus, G., 1970a. Inference from inadequate and inaccurate data, I, *Proc. Nat. Acad. of Sciences*, **65**, 1-7.

Backus, G., 1970b. Inference from inadequate and inaccurate data, II, *Proc. Nat. Acad. of Sciences*, **65**, 281-287.

Backus, G., 1971, Inference from inadequate and inaccurate data, III, *Proc. Nat. Acad. of Sciences*, **67**, 282-289.

Bolt, B. A., 1960. The revision of earthquake epicenter, focal depths, and origin times using a high-speed computer, *Geophys. J. R. astr. Soc.*, **3**, 433-440.

Bolt, B. A., 1970. Earthquake location for small networks using the generalized inverse matrix, *Bull. Seism. Soc. Am.*, **60**, 1823-1828.

Buland, R., 1976. The mechanics of locating earthquakes, *Bull. Seism. Soc. Am.*, **66**, 173-187.

Crosson, R. S., 1976. Crustal structure modeling of earthquake data, 1, Simultaneous least square estimation of hypocenter and velocity parameters, *J. Geophys. Res.*, **81**, 3036-3046.

Dobrin, M. B., 1976. *Introduction to Geophysical Prospecting*, 3rd ed., McGraw-Hill, New York.

Domenico, S. N., 1984. Rock lithology and porosity determination from shear and compressional wave velocity, *Geophysics*, **49**, 1188-1195.

Flinn, E. A., 1965. Confidence regions and error determinations for seismic event location, *Rev. Geophys.*, **3**, 157-185.

Geiger, L., 1910. Herdbestimmung bei erdbeden aus den ankunfzeitzen, *K. Gessell. Wiss. Goett.*, **4**, 331-349.

Golub, G. H. and R. J. Plemmons, 1981. Large-scale geodetic least-squares adjustment by dissection and orthogonal decompositions, in *Large Scale Matrix Problems*, ed. H. Schneider, Elsevier, New York.

Hoel, P. G., 1971. *Introduction to Mathematical Statistics*, Wiley, New York.

Jordan, T. H. and K. A. Sverdrup, 1981. Teleseismic location techniques and their application to earthquake clusters in the South-Central Pacific, *Bull. Seism. Soc. Am.*, **71**, 1105-1130.

Klein, F. W., 1978. Hypocenter location program HYPOINVERSE, Part 1: user's guide to versions 1, 2, 3, and 4, *U.S. Geol. Surv. Open-File Rept.*, 78-694.

Lanczos, C., 1961. Linear differential operators, Van Nostrand, New York.

Lawson, C. L. and R. J. Hanson, 1974. Solving least squares problems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Lee, W. H. K., and J. C. Lahr, HYPO71, A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes, *U.S. Geol. Surv. Open-File Rept.*

Lee, W. H. K., and S. W. Stewart, 1981. Principles and applications of microearthquake networks, in *Advances in Geophysics*, **23**, Academic Press, New York.

Lomnitz, C., 1977. A fast epicenter location program, *Bull. Seism. Soc. Am.*, **67**, 425-431.

Moos, D. and M. D. Zoback, 1983. *In situ* studies of seismic velocity in fractured crystalline rocks, *J. Geophys. Res.*, **88**, 2345-2358.

Oldenburg, D. W., 1984. An introduction of linear inverse theory, *IEEE Trans. Geoscience and Remote Sensing*, GE-22, 665-674.

Parker, R. L., 1977a. Understanding inverse theory, *Ann. Rev. Earth Planet. Sci.*, **5**, 35-64.

Parker, R. L., 1977b. Linear inference and underparameterized models, *Rev. Geophys. Space Phys.*, **15**, 446-456.

Pavlis, G. L. and J. R. Booker, 1980. The mixed discrete-continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res.*, **81**, 4801-4810.

Pavlis, G. L., 1982. Progressive inversion, *Ph.D. Dissertation*, University of Washington, Seattle, Washington.

Pavlis, G. L. and J. R. Booker, 1983a. A study of the importance of nonlinearity in the inversion of earthquake arrival time data for velocity structure, *J. Geophys. Res.*, **88**, 5047-5055.

Pavlis, G. L. and J. R. Booker, 1983b. Progressive multiple event location (PMEL), *Bull. Seism. Soc. Am.*, **73**, 1753-1777.

Pavlis, G. L., 1986. Appraising earthquake hypocenter location errors: a complete, practical approach for single event locations. *Eos*, **67**, 306.

Penrose, R., 1955. A generalized inverse for matrices, *Proc. Camb. phil. Soc.*, **52**, 406-413.

Rodi, W. L., T. H. Jordan, J. F. Masso, and J. M. Savino, 1980. Determination of the three-dimensional structure of eastern Washington from the joint inversion of gravity and earthquake data, *Rep. SSS-R-80-4516*, Systems Science and Software, La Jolla.

Spencer, C. and D. Gubbins, 1980. Travel-time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media, *Geophys. J. R. astr. Soc.*, **63**, 95-116.

Spencer, C., 1985. The use of partitioned matrices in geophysical inversion problems, *Geophys. J. R. astr. Soc.*, **80**, 619-629.

Stump, B. W. and L. R. Johnson, 1977. The determination of source properties by the linear inversion of seismograms, *Bull. Seism. Soc. Am.*, **67**, 1489-1502.

Tarantola, A. and B. Valette, 1982. Inverse problems = quest for information, *J. Geophys.*, **50**, 159-170.

Tatham, R. H. and P. L. Stoffa, 1976. V_p/V_s– A potential hydrocarbon indicator *Geophysics*, **41**, 837-849.

Tatham, R. H., 1982. V_p/V_s and lithology, *Geophysics*, **47**, 336-344.

Thurber, C. H., 1983. Earthquake locations and three-dimensional crustal structure in the Coyote Lake Area, Central California, *J. Geophys. Res.*, **88**, 8226-8236.

Thurber, C. H., 1985. Nonlinear earthquake location: theory and examples, *Bull. Seism. Soc. Am.*, **75**, 779-790.

Wiggins, R. A., 1972. The general linear inverse problem: implications of surface waves and free oscillations for earth structure, *Rev. Geophys.*, **10**, 251-285.

Chapter 3

Synthetic Tests of Progressive Inversion

3.1. Introduction

Progressive inversion was extended to incorporated S-wave arrival time data in Chapter 2. Since the approach required linearization of a nonlinear problem we need to determine how well the method works in practice. Pavlis and Booker (1983) investigated the problem of estimating P-wave velocity structure using P-wave arrival time data from earthquakes. Using synthetic data they focused on determining the reliability of linearized resolution and error analysis of estimated P-wave velocity structure. Effects of nonlinearity on hypocentral estimates and station corrections were not presented. Here, synthetic data is used to access effects of nonlinearity on estimates of P and S-wave velocity structure, hypocenters, and station corrections. The scope is broader in order to determine how strongly nonlinearities in the problem effect each set of parameters in an absolute and relative sense.

In some situations, reliable earthquake locations are a primary objective, whereas estimating velocity structure is not as critical. A common example is the problem of determining fault position and mode of faulting using earthquake locations and focal mechanisms. Of course accurate velocity models are necessary to accurately locate earthquakes, but we would like to know what effect inaccurate velocity models have on estimates of earthquake locations. A method was developed in Section (2.6.2) of Chapter 2 to estimate the errors in hypocenters due to imperfect velocity models. Since the approach required a linearized approximation and a rough bound approach, tests with synthetic data are needed to investigate well how these approximations predict true hypocenter error.

Progressive inversion as developed in Chapter 2 is very similar to the methods of Pavlis and Booker (1980) and Pavlis (1982), except S-wave data and velocity structure are added. They conducted a thorough study of the effects of nonlinearity on inversions for P-velocity structure. They showed that linearized inversions can produce incorrect models for which subsequent resolution-error analysis overestimate resolution and underestimates errors. The same problem can arise when S-wave arrival times are added. The results of Pavlis and Booker (1983a) will be used as a starting point. Consequently, the present series of synthetic tests will concentrate more on determining the effects of including S-waves in the problem than focusing solely on studying the importance of nonlinearity on inversions for velocity structure, hypocenters, and station corrections.

Although velocity structure is an end product of progressive inversion, it is slowness gradient structure that is directly solved for. Resolution and error are only rigorously defined for slowness gradient kernels. Resolution and error can be estimated for velocity structure but the results are highly dependent on model parameterization. Consequently, we use slowness gradient models to investigate the reliability of linearized resolution-error analysis. The problem of estimating velocity structure is considered first. Slowness gradient, station correction, and hypocenter assessment follow in subsequent sections.

3.2. Synthetic Data

A synthetic data set was constructed using earthquake locations and station distributions from The Geysers, California. The locations of 39 earthquakes were estimated using P and S-wave arrival times recorded by a temporary network of 9 three-component stations and the USGS permanent seismographic network. Earthquake locations were estimated using an assumed set of velocity models and no attempt was made to estimate station corrections or invert for velocity structure. These locations were used to calculate synthetic travel times by employing prescribed station corrections and velocity models. The velocity models and station corrections used to generate the synthetic data are listed in Tables (3.1) and (3.2), respectively. True earthquake epicenters and station locations are shown in Figure (3.1). Figure (3.2) shows true earthquake elevations in cross section. Earthquake depths are shallow reflecting the

Synthetic Model		
Velocity (km/sec)		
P-wave	S-wave	Depth (km)
3.08	1.65	0.0
3.11	1.87	0.4
3.50	2.20	0.6
3.80	2.28	0.7
4.80	2.43	0.9
4.94	2.90	1.5
5.30	3.20	3.0
6.00	3.60	6.0

Table 3.1 Velocity models used to generate synthetic travel times. Linear velocity gradients are used between each velocity-depth pair. The depth of 0.0 km corresponds to an elevation of -0.379 km in Figure (3.2). Velocity values at 0.0 km depth are used to define the constant velocity elevation correction layer described in Appendix A.

Synthetic Model		
	Station Corrections (sec)	
Station	P-wave	S-wave
GAX	-0.10	-
GBO	0.15	-
GCM	0.10	-
GCR	-0.15	-
GDX	-0.05	-
GGP	0.10	0.15
GMM	-0.10	-
GSM	-0.10	-
TRA	0.05	0.10
TSP	0.05	0.10
TOT	-0.05	-0.10
TPU	0.10	0.15
TPS	0.00	-0.10
TPT	-0.10	-0.20
TPL	-0.10	-0.20
THR	0.10	0.15
TMZ	0.10	0.15
TPR	-0.10	-0.20

Table 3.2 Station corrections used to generate synthetic data. Station corrections were chosen large enough to reflect variations of near surface velocities over a range of 1-4 km/sec and thicknesses of 0.1-0.5 km that have been observed at The Geysers.

shallow geothermal reservoir at The Geysers. The regular pattern of epicenters in Figure (3.1) is due to a formatting error when generating the input file to the program that calculated the synthetic travel times. Decimal latitude and longitude coordinates for the epicenters were rounded to the nearest hundredth of a degree, producing the distribution in Figure (3.1). The net result is a decrease in distinct epicenters from 39 to 17. Since the depths are distinct (Figure (3.2)), data from overlapping epicenters are not completely redundant. Still, this geometry produces data that contain less information than more realistic geometries with more distinct epicenters, because fewer station-receiver distances are represented. Velocities from the surface down to 1.5 km depth were taken from a velocity inversion of VSP data described in Chapter 4. P-wave velocities below 1.5 km depth were modified from Eberhart-Phillips and Oppenheimer (1984). S-wave velocities below 1.5 km depth were obtained by assuming a slowly decreasing V_p/V_s ratio with depth to simulate anticipated V_p/V_s decreases in the steam reservoir at The Geysers. The reason The Geysers data were used to formulate synthetic data was to better assess results of progressive inversions on real earthquake data from The Geysers. Also, this data set is typical of many microearthquake studies and aftershock studies in terms of recording geometry and event locations.

All data sets consist of error free (errors <0.0001 second) travel times calculated using a method described in appendix A. We consider two types of data sets. One data set consists of P-wave arrival times only. This data set, dubbed P-only, is obtained by deleting all S-wave arrival times from the synthetic data set. A second data set, dubbed joint P and S, consists of the same P-wave arrival times plus the original S-wave arrival times. Since only 11 of 18 recording stations at The Geysers operated with three-component geophones, there are only about half as many S-wave arrival times. The second data set reflects the common situation whereby S-wave arrival times are only available from a subset of stations; those that record three components of ground motion.

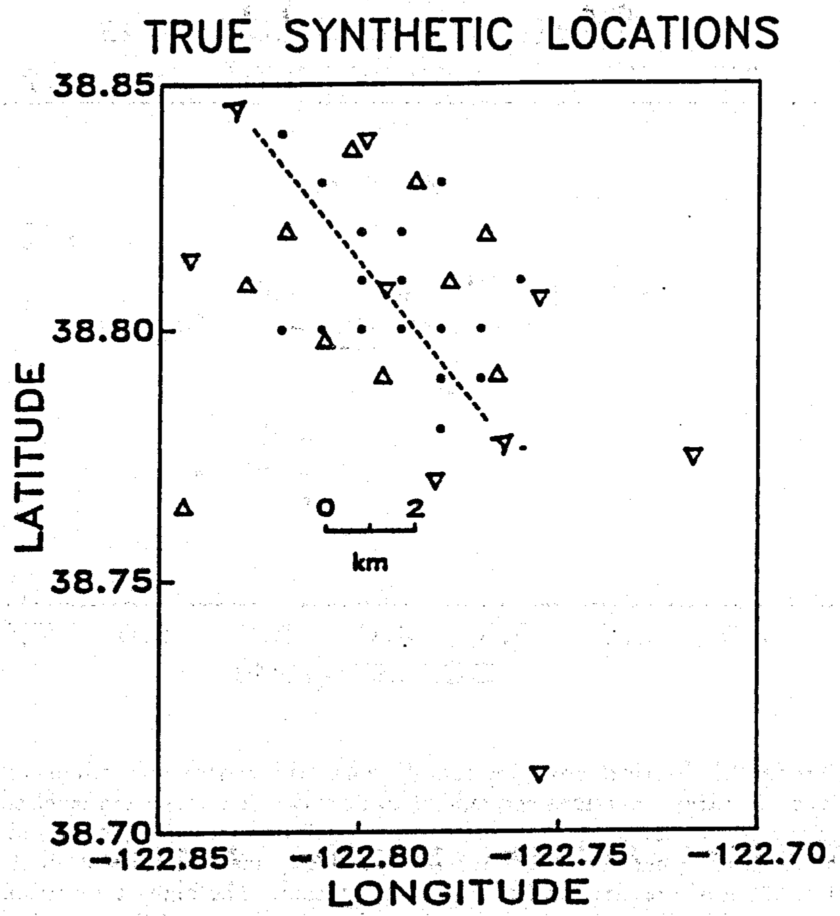


Figure 3.1 True epicenter positions of synthetic earthquakes are shown as small circles. Stations that provide P and S-wave arrival times are denoted by Δ and station that provide P-wave arrival times only are shown as ▽. The dashed line from A to A′ is the surface projection for the vertical cross section shown in figure 3.2.

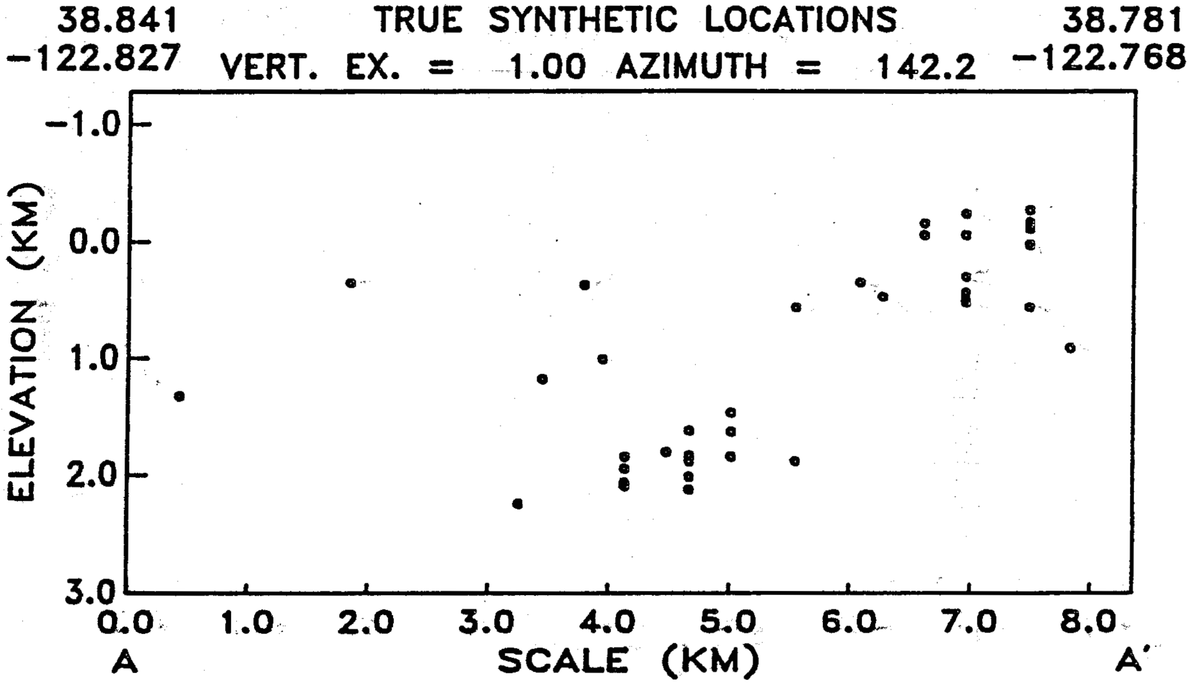


Figure 3.2: Vertical cross section showing true synthetic earthquake elevations. Positive elevations correspond to depths below mean sea level and zero elevation corresponds to mean sea level. The cross section ends, A and A' are the same as in figure 3.1. The top of the velocity models starts at an elevation of -0.379 km; the elevation of the lowest station. The highest elevation, -1.28 km, corresponds to the elevation of the highest station. Differences in station elevations were accounted for using methods described in Appendix A.

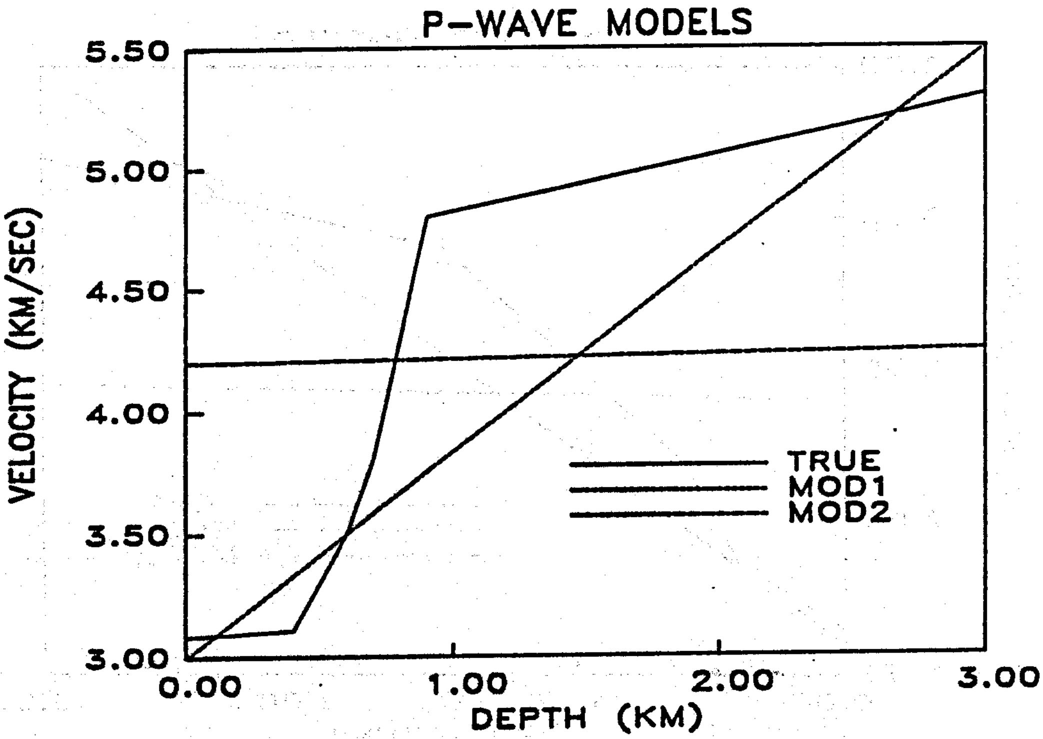


Figure 3.3 P-wave starting models shown with the true model. The "close" starting model, model 1, is shown as MOD1 and the "far" starting model, model 2, is shown as MOD2.

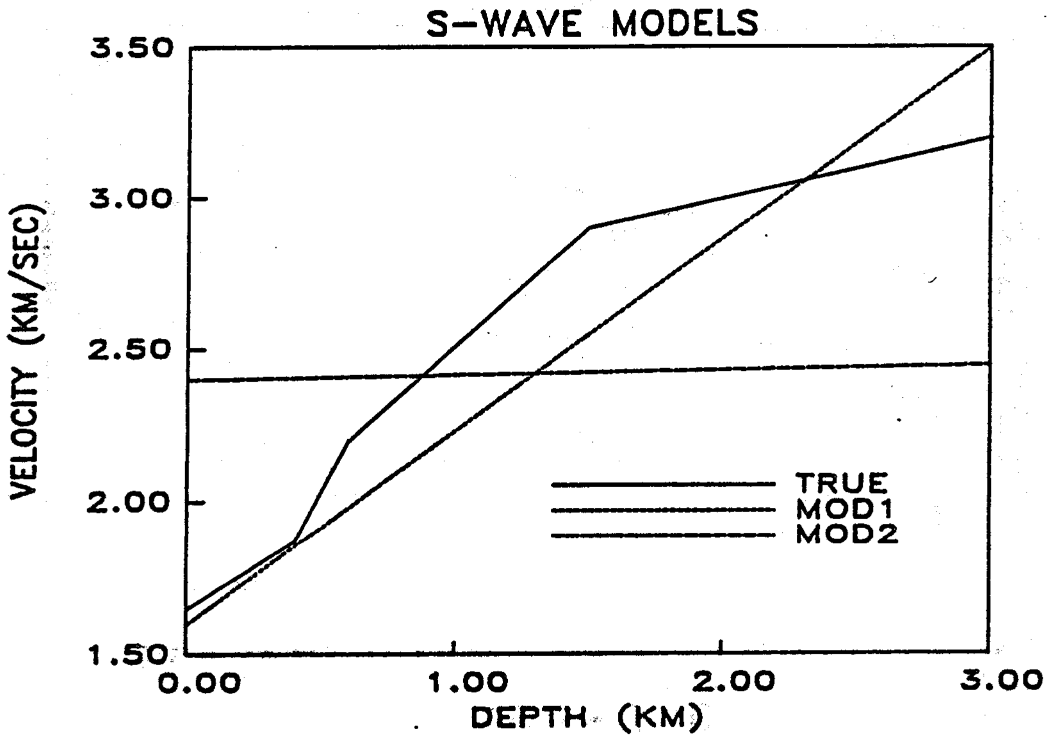


Figure 3.4 S-wave starting models shown with the true model. The "close" starting model, model 1, is shown as MOD1 and the "far" starting model, model 2, is shown as MOD2.

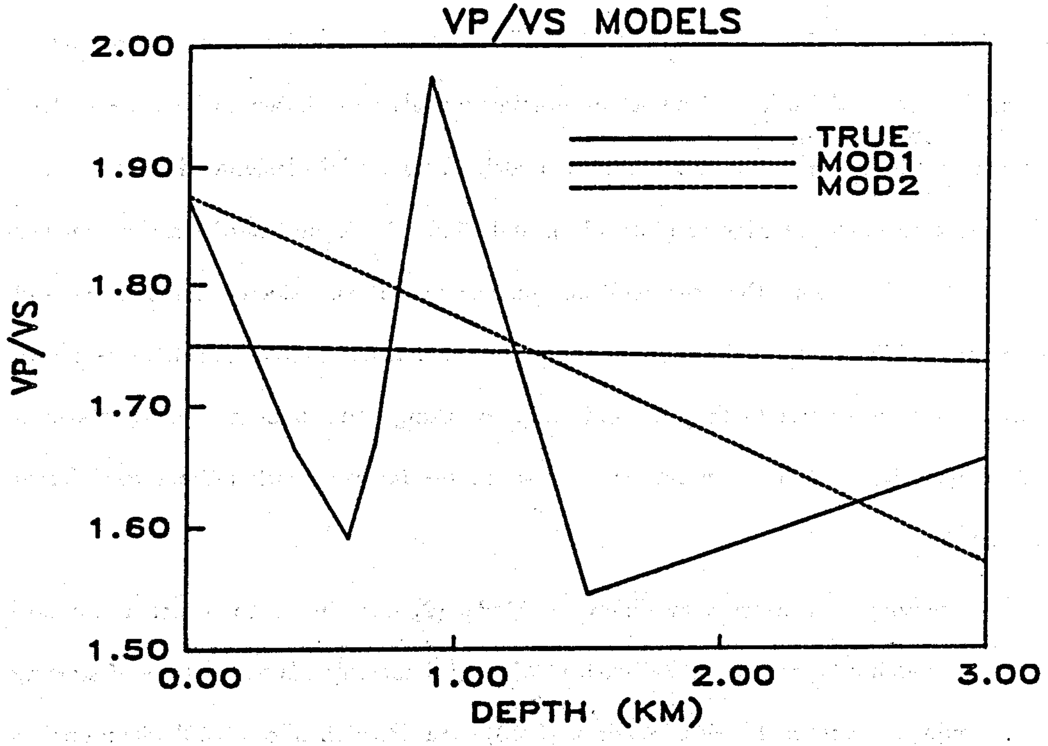


Figure 3.5 V_P/V_S starting models shown with the true model. The "close" starting model, model 1, is shown as MOD1 and the "far" starting model, model 2, is shown as MOD2.

3.2.1. Starting Models

Results of a nonlinear least-squares inversion can depend strongly on the starting model. To study how strongly inversion results depend on initial models, two sets of starting models were used. Both sets of models consisted of a single linear velocity gradient from the surface to maximum model depth of 3.0 km. One set of starting models was chosen to be close to true surface velocities and to be "close" to the true model. This model, designated model (1), is shown with the true model in Figures (3.3), (3.4), and (3.5). A second set of starting models was chosen to be "far" from the true models; surface velocities, velocity magnitudes, and Vp/Vs were much different than the true models. This model, designated model (2), is shown with the true model in Figures (3.3), (3.4), and (3.5). By using quite different staring models it was hoped to quantify how stable progressive inversion results were with respect to different starting models.

The two starting models are very different. Model (2) was chosen to reflect a standard type of starting model: a homogeneous half-space. Model (1) corresponds to the type of starting model that could be obtained using Wadati (1933) and Riznichenko (1958) diagrams as described by Nicholson and Simpson (1985). The advantages of constructing good starting models, such as would be determined using Nicholson and Simpson's (1985) approach, becomes clearly apparent in subsequent sections.

3.3. Velocity structure

Estimating seismic velocity structure using earthquake sources is inherently an inferior experimental design. Seismic velocity determination is best done using seismic reflection, VSP, and refraction experiments with controlled sources. The reasons are obvious; neither travel times or source locations are known for earthquakes, but are well determined in controlled experiments.

Some situations dictate that earthquake data be used to estimate seismic velocity structure. Many times earthquake data are the only data available. In other cases earthquake data sometimes contain information that is not contained in reflection and refraction data. Reflection and refraction data represent virtual two-dimensional samples of seismic structure. The use of overlapping profiles produces improved spatial sampling, but the extent of three-dimensional sampling is limited. Earthquake data recorded by typical networks may sample over a larger volume of a region because earthquake locations are distributed over a wide range of depths and epicentral locations. Consequently, overall three-dimensional seismic velocity information may be contained in in earthquake data that is absent from reflection and refraction data. (Thurber, 1983; Eberhart-Phillips, 1986).

There are two objectives when inverting for seismic velocity structure using earthquake data. Obviously, precise determination of seismic velocity structure is one objective. It is already clear that earthquake data are not ideal in this regard. It will be demonstrated that, in general, absolute velocity magnitudes are not well constrained by earthquake data. Thus, a primary objective is to reduce the component of earthquake location errors due to errors in assumed seismic velocity structure. This topic will be disussed in Section (3.6). In the following sections we investigate how well seismic velocity structure is constrained by earthquake data.

3.3.1. Factor Analysis

Progressive inversion involves iterative solutions on three levels; hypocenters, station corrections, and velocity structure. Here, the primary interest is on how velocity estimates are effected, by being required to solve for hypocenters and station corrections as a prelude to solving for velocity structure. Not only are more unknowns added to the problem, but some information becomes unavailable for the velocity inversion due to data annulment. The goal of this factor analysis is to determine the effects on velocity inversions, of removing data through annulment and adding hypocenters and station corrections as variables.

Assessment of seismic velocity inversions using synthetic earthquake data is done in a series of steps (Table 3.3). Each step is designed to isolate the effect of removing data through annulment and/or adding more unknowns to the problem. In case (A), all parameters are fixed at their true values; inversion is done for velocity models only. Case (B) differs from case (A) only in that data is removed from the subsequent velocity inversion through hypocenter annulment. Next, in case (C), earthquake locations are estimated in addition to velocity models. This case was investigated by Pavlis and Booker (1983a) and Pavlis (1982) in the case of P-wave data only. Case (D) investigates the effect of removing more data from the subsequent velocity inversion by means of station correction annulment. Case (E) is similar to case (D) except that hypocenters must be solved for also. Finally, in case (F), station corrections and hypocenter are solved for along with velocity structure. Specific cases will be discussed in subsequent sections.

All cases, (A-F), were only done for the combined P and S inversions with starting model (1). A subset of cases were investigated with starting model (1) for P-only data and starting model (2) data sets. This was done to reduce the overall computational burden and the fact that case (F) was of primary interest.

In the synthetic tests that require event location, the Inglada (1928) method was used to obtain a starting epicenter and origin time. Initial event elevation was set to the same value, 1.12 km, for all events. In the synthetic tests which required station corrections to be estimated, all initial station corrections were set to zero.

3.3.1.1. Case (A)

This test provides insight into inherent strengths and weaknesses of the synthetic data's source-receiver geometries when inverting for velocity structure. Iterative velocity inversions were done with both starting models for P- and S-wave velocity structure. In this case, the data set is like a combination of seismic refraction, reflection, and VSP data. Event location and annulment of data are not required. Results of this inversion indicate fundamental strengths and weaknesses of the data set. For instance, if convergence to true model values cannot be obtained in some depth intervals, we can anticipate that hypocentral depths and velocity structure will be poorly constrained in those depth intervals.

Case	Annulment	Hypocenters	Station Corrections
A	none	fixed at true values	fixed at true values
B	hypocenters	fixed at true values	fixed at true values
C	hypocenters	solved for	fixed at true values
D	hypocenters station corrections	fixed at true values	fixed at true values
E	hypocenters station corrections	solved for	fixed at true values
F	hypocenters station corrections	solved for	solved for

Table 3.3 List of progressive inversion test cases. The annulment column indicates what data components were removed from the potential pool of data for velocity inversion using orthogonal transformations.

For this special case, P-wave and S-wave inversions are completely independent so only one inversion need be done to obtain complete P and S results for each starting model set. From the results in Figures (3.6) and (3.7), it is clear that a slightly smoothed version of the true P and S models can be recovered using starting model (1). Results with starting model (2) are not as satisfactory (Figures (3.6) and (3.7)). Shallow (0-0.5 km) model (2) near-surface P-wave velocities are too high and the position and magnitude of the large velocity gradient between 0.4 and 0.7 km are incorrect. The is due to the large error in the initial surface velocity of model (2), (1.115 km/sec), and the lack of hypocenters shallower than 0.15 km model depth. Since we are inverting for the flattest model, incorrect surface velocities produces an incorrect boundary condition for the inversion which produces errors in other parts of the model (see Oldenburg (1984) for some simple examples). This does not bode well for cases (C-F) with the starting model (2), when hypocenters and station corrections are added as free variables to the problem.

3.3.1.2. Cases (B-E)

Sparse data are available to constrain near surface velocities and velocities near the bottom of the models. Near surface velocities are poorly constrained because nearly all wave paths traverse the near surface at a small range of nearly vertical incidence angles. Model bottom velocities are constrained by a small number of rays that bottom there. Consequently, we would expect that data annulment would adversely effect estimates of velocities in these portions of the model, relative to case (A).

Cases (B) and (D) reveal effects of data annulment on velocity estimates. P and S-wave velocity estimates for case (B) (Figures 3.8 and 3.9) reproduce the true models well except for the near surface portions of the models. The deepest portion of the P-wave model is not faithfully reproduced for case (B). Since these portions of the models were recovered in case (A), lack of convergence to the true model in the deepest portion of the estimated model, indicates that constraining data were removed during hypocenter annulment. Model misfit increase throughout the velocity models in case (D) indicating that more data, fundamental to

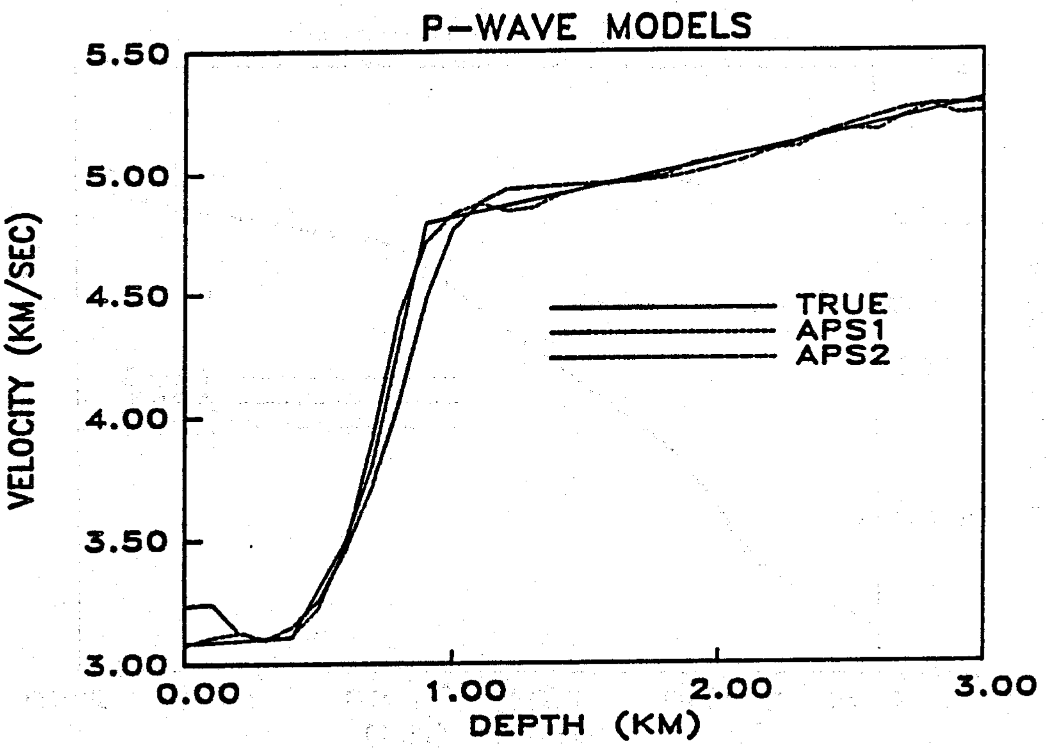


Figure 3.6 Results of P-wave velocity inversions for case A. Final estimates of velocities for model one are identified by APS1 and those for model two correspond to APS2. The final RMS residual for model one was 0.001 sec after 7 iterations. Final RMS residual for model two was 0.014 sec after 18 iterations.

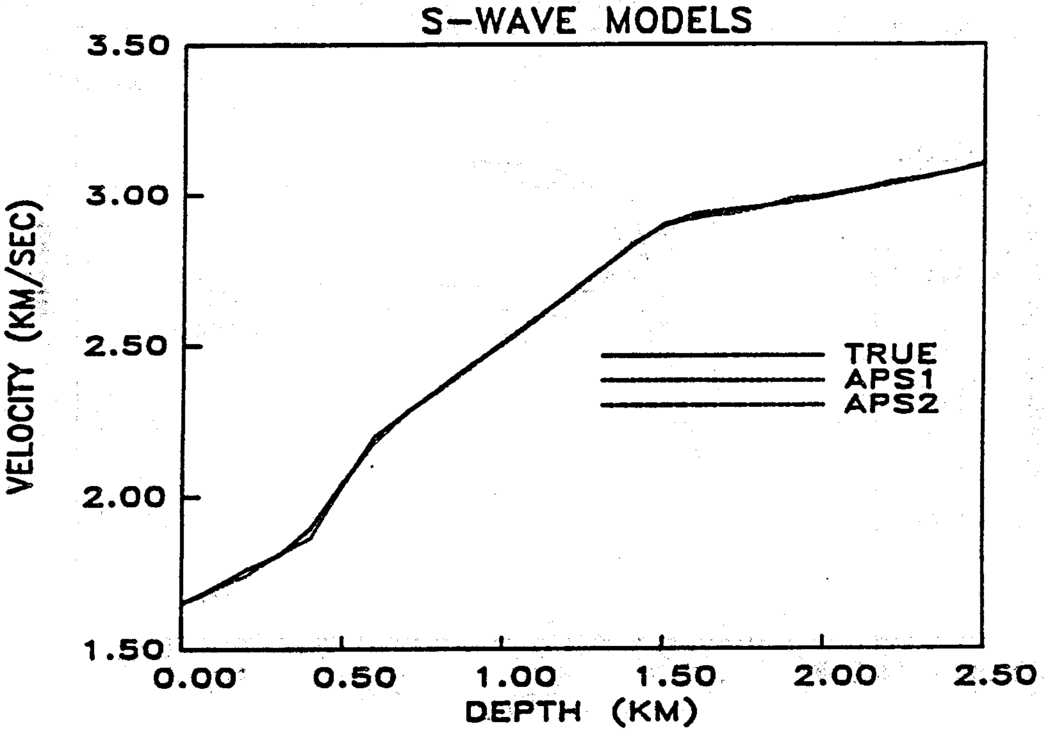


Figure 3.7 Results of S-wave velocity inversions for case A. Final estimates of velocities for model one are identified by APS1 and those for model two correspond to APS2. The final RMS residual for model one was 0.007 sec after 7 iterations. Final RMS residual for model two was 0.008 sec after 18 iterations.

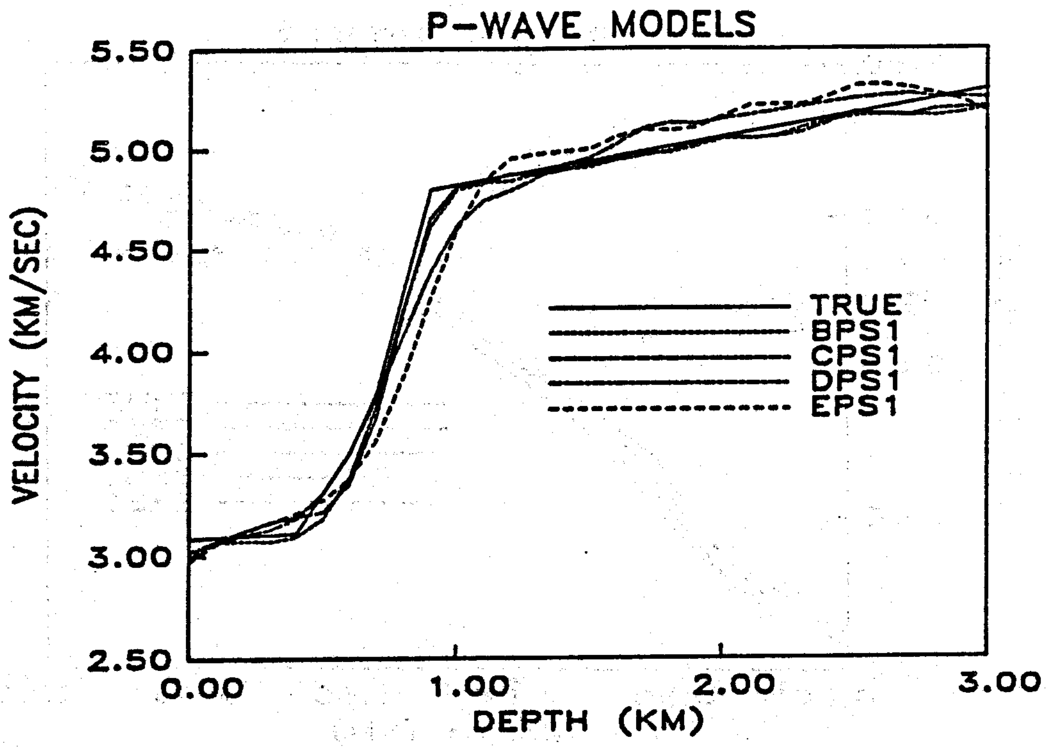


Figure 3.8 P-wave results for joint P and S velocity inversions with starting model one, cases B-E. Case B (0.018 sec RMS, 9 iterations) is designated by BPS1, case C (0.002 sec RMS, 5 iterations) by CPS1, case D (0.024 sec RMS, 3 iterations) by DPS1, and case E (0.010 sec RMS, 4 iterations) by EPS1.

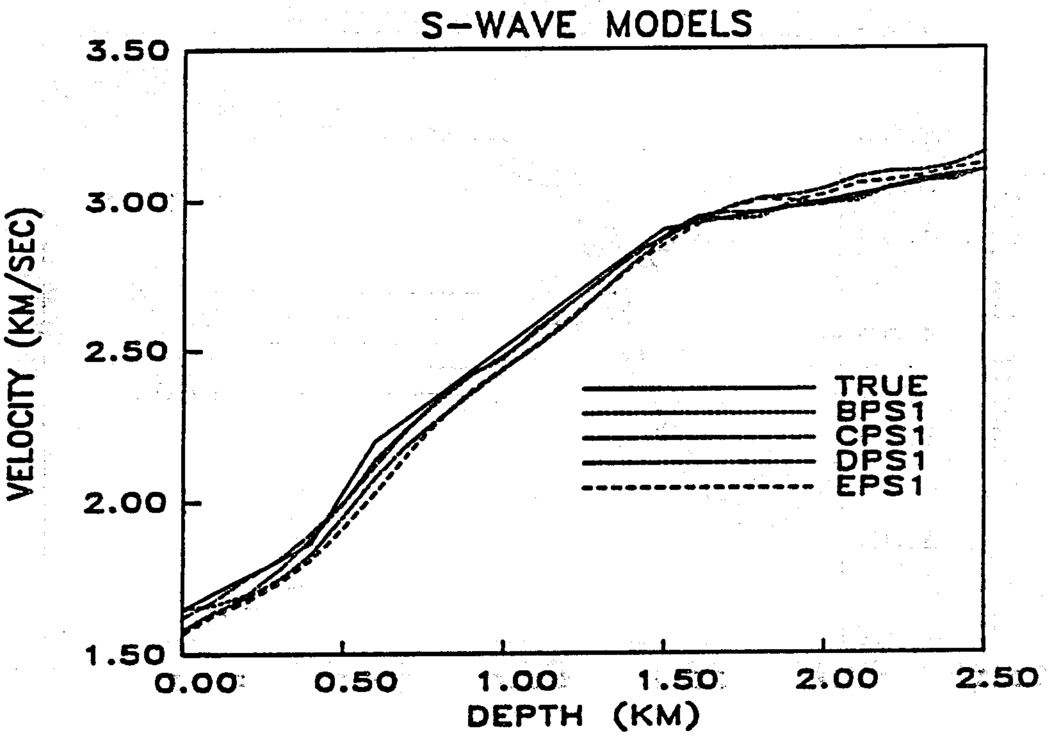


Figure 3.9 S-wave results for joint P and S velocity inversions with starting model one, cases B-E. Case B (0.017 sec RMS, 9 iterations) is designated by BPS1, case C (0.004 sec RMS, 5 iterations) by CPS1, case D (0.056 sec RMS, 3 iterations) by DPS1, and case E (0.020 sec RMS, 4 iterations) by EPS1.

constraining whole-model velocity structure, were removed during station correction annulment. Results for case (B) bear out our expectations concerning patterns of near-surface model misfit. However, case (D) results indicate that data with potential to constrain all portions of the velocity models is removed to constrain station corrections.

Inversions for cases (C) and (E) produce results similar to cases (B) and (D) (Figures (3.8) and (3.9)). These results are not surprising; the hypocentral estimates at convergence in cases (C) and (E), were nearly identical to the true locations. This indicates that the determining velocity structure along with hypocenter locations is a robust process when the starting model is not too far from the true model.

Case (C) corresponds to the problem considered by Pavlis and Booker (1983a). Figure (3.10) shows cases (C) results for P-wave structure from the joint P and S velocity-hypocenter inversion and the P-only inversion. Despite doing 11 velocity inversion iterations for the P-only case, versus just 5 iterations for the joint P and S case, the P-only velocity model estimate does not reproduce the true P model as well as the joint P and S inversion estimate. The high velocity gradient in Figure (3.10) has been systematically moved deeper in the P-only case. This result can be explained by the origin time-depth-dc velocity tradeoff discussed in Pavlis and Booker (1983a).

Inspection of Figure (3.34) reveals that estimated hypocentral depths for P-only case (C) are systematically too deep. Estimated origin times for P-only case (C) were systematically early. Thus, the depth offset of P-only velocity model and estimated hypocenters has been compensated by a dc shift in origin times. The origin time-depth-dc velocity tradeoff was greatly diminished using joint P and S data in case (C). Hypocentral depths errors are much smaller (Figure (3.32)), and the P velocity model is closer to the true model (Figure (3.10)).

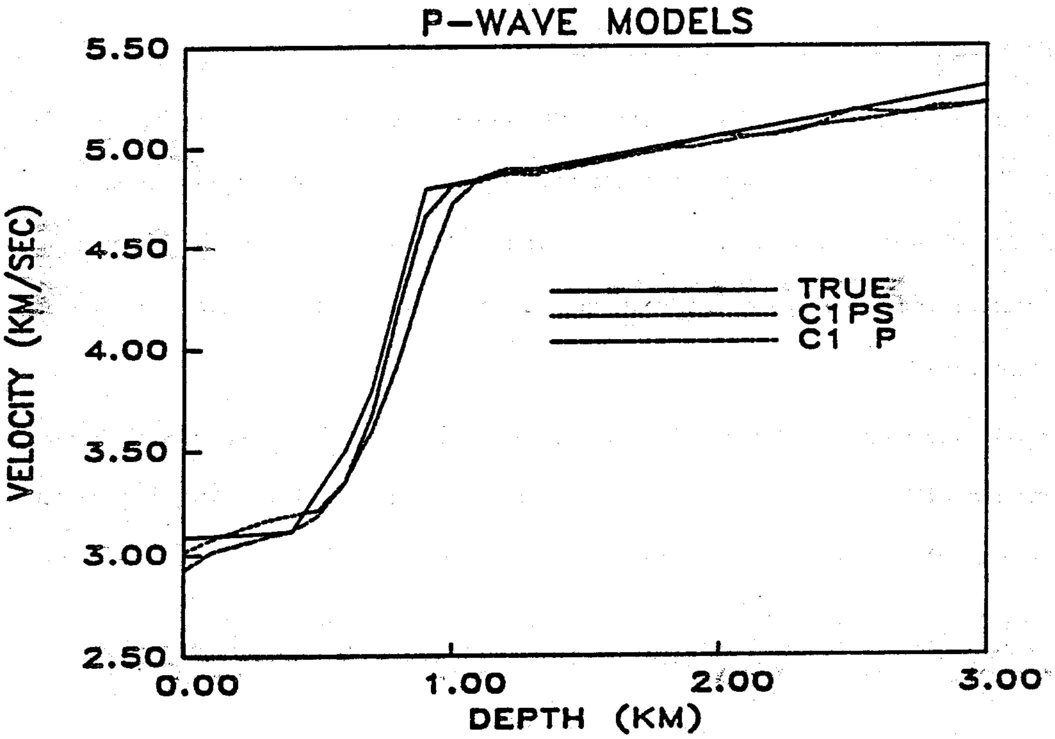


Figure 3.10 P-wave case C velocity inversion results using starting model one. The P-wave model estimate from the joint P and S inversion is denoted as C1PS (0.002 sec RMS, 5 iterations), and the P-wave model estimated using P arrival data only is denoted as C1 P (0.005 sec RMS, 11 iterations).

3.3.1.3. Case (F)

Since case (F) is the end product of primary practical interest, inversions were done for both P-only and joint P and S data for starting models (1) and (2). Since case (F) involves solving for the most variables, errors due to nonlinearity are more likely to manifest themselves in this test. Inversion results are likely to be sensitive to starting models. These expectations are confirmed.

Convergence problems were pronounced in case (F) when using starting model (2). Satisfactory convergence could not be obtained for either the P-only or joint P and S inversion with model 2 (Figures (3.11) and (3.12)). Lack of convergence was not readily apparent from the RMS data misfit of 0.028 sec for the P-only starting model (2) inversion. An RMS misfit of 0.028 sec would clearly be within the uncertainty of real data that had arrival time errors on the order 0.03-0.04 sec. What indicates a lack of proper convergence is the tendency for some shallow earthquakes to locate at or near the free surface (Figure (3.22)). If this where real data, one would have to detect this problem to avoid erroneous results. This would require knowing that very shallow earthquake locations are unreasonable, not a trivial assumption if shallow seismicity is of interest.

When S-wave arrival times are incorporated into the far model inversion, the lack of convergence is more clearly apparent in the RMS data misfit. P-wave RMS misfit is 0.048 sec and S-wave misfit is 0.118 sec in this case. The tendency for some shallow earthquakes to locate near the free surface is still apparent (Figure (3.24)), but the large RMS S-wave data misfit helps detect lack of proper convergence. Including S-wave information allowed detection of convergence to a local rather than global minimum through inspection of RMS data misfit magnitude. Convergence to a local minima was not clearly apparent when only P-wave arrival times were used.

Large RMS misfit magnitudes, relative to uncertainties in the data, can be used reevaluate the starting model. If too little information were utilized to construct the starting model, a more comprehensive effort could be made to improve it and the new starting model used to

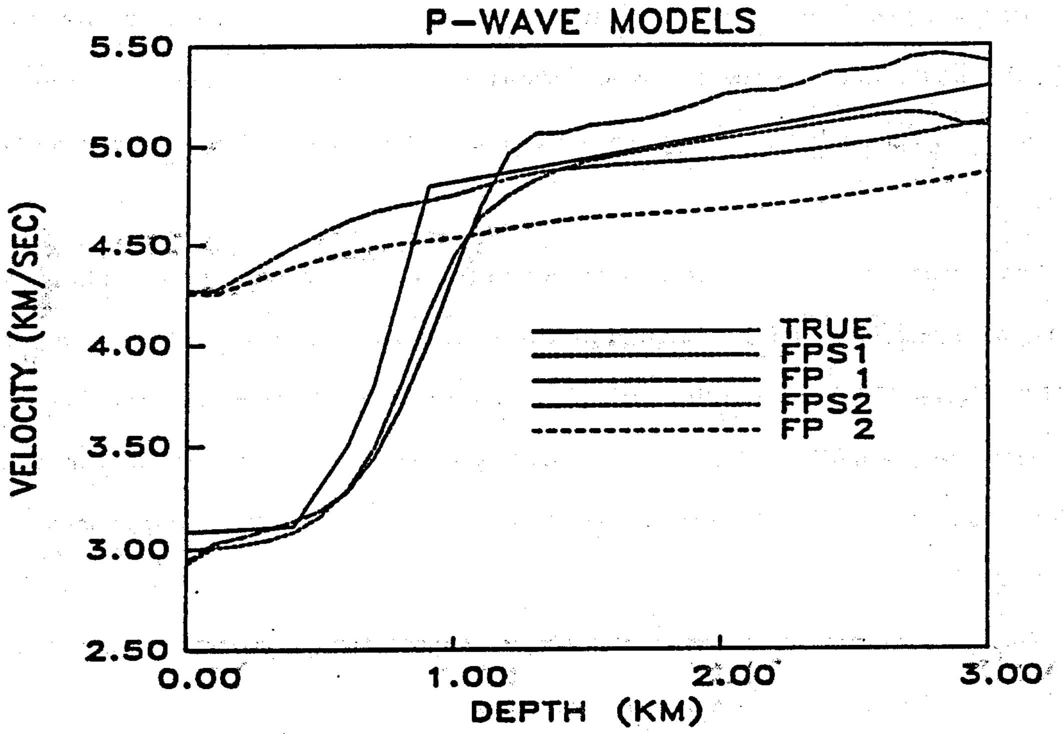


Figure 3.11 P-wave case F velocity inversion results for both starting models and P only and joint P and S inversions. Starting model one results for the joint P and S inversion are denoted as FPS1 (0.010 RMS, 3 iterations) and as FP 1 (0.003 RMS, 5 iterations) for the P only inversion. Starting model two results for the joint P and inversion are shown as FPS2 (0.048 RMS, 3 iterations) and as FP 2 (0.028 RMS, 2 iterations) for the P only inversion.

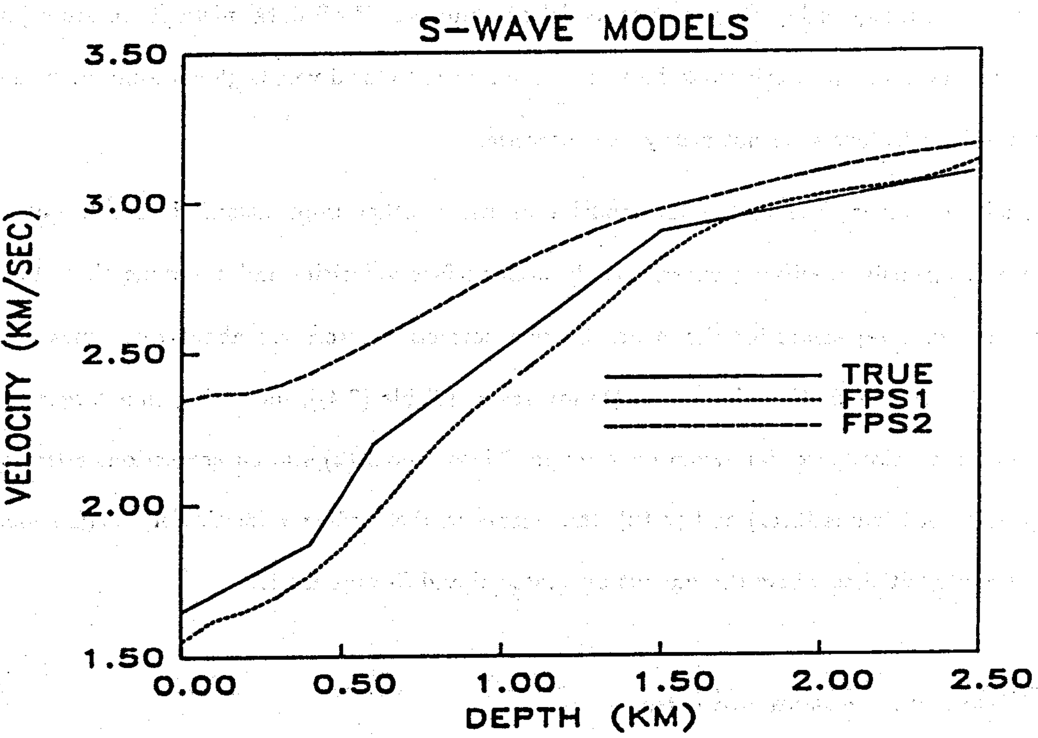


Figure 3.12 S-wave case F velocity inversion results for both starting models. Starting model one results are denoted as FPS1 (0.025 RMS, 3 iterations). Starting model two results are shown as FPS2 (0.118 RMS, 3 iterations).

reinvert. Lack of convergence to acceptable RMS data misfit levels could also indicate that the one-dimension model assumption is invalid. However, careful investigation of the raw data (inspection of travel time curves, etc.) should reveal whether or not model assumptions are appropriate. Consequently, given proper model assumptions, RMS data misfit levels from joint P and S inversions help discriminate between convergence to local versus global minima whereas P-only RMS misfit levels are not nearly as diagnostic.

Lack of convergence in the far model case has another implication. It points out the danger of arbitrarily specifying unrealistically high surface velocities and assuming that station corrections can compensate for the error. Station correction errors are about two times larger for model (2) inversions than for model (1) inversions (Table (3.4)), indicating that larger components of the velocity model errors have mapped into model (2) station corrections estimates. It is clear from Figures (3.11) and (3.12), that errors in the surface velocities of starting model (2) have contributed to errors throughout estimated P and S-wave models.

3.3.2. Velocity resolution and error

Estimates of velocity structure should be equivalent to the true model smoothed by the resolving functions for the solution within predicted error bounds.[3.1] Resolving functions and standard errors can be calculated from equations (2.59) and (2.60) of Chapter 2. Figure (3.13) compares starting model (1) estimated P-wave velocities with the true P-wave velocity model smoothed by resolving kernels for the case (F) P-only inversion. The true model and smoothed true model fall outside the estimated model bounds virtually throughout the entire model depth extent. Yet the RMS data residual for this inversion is very small (0.003 sec). In contrast, slowness gradient for this example (Figure (3.17)) is very close to the true model for most of the model depth extent. This example demonstrates that dc velocities can be incorrectly predicted using earthquake data whereas slowness gradients errors are much smaller.

The problem is most pronounced when only P-wave data are used. If is clear from Figure (3.11) that estimated P-wave velocities are much closer to the true model when combined P and

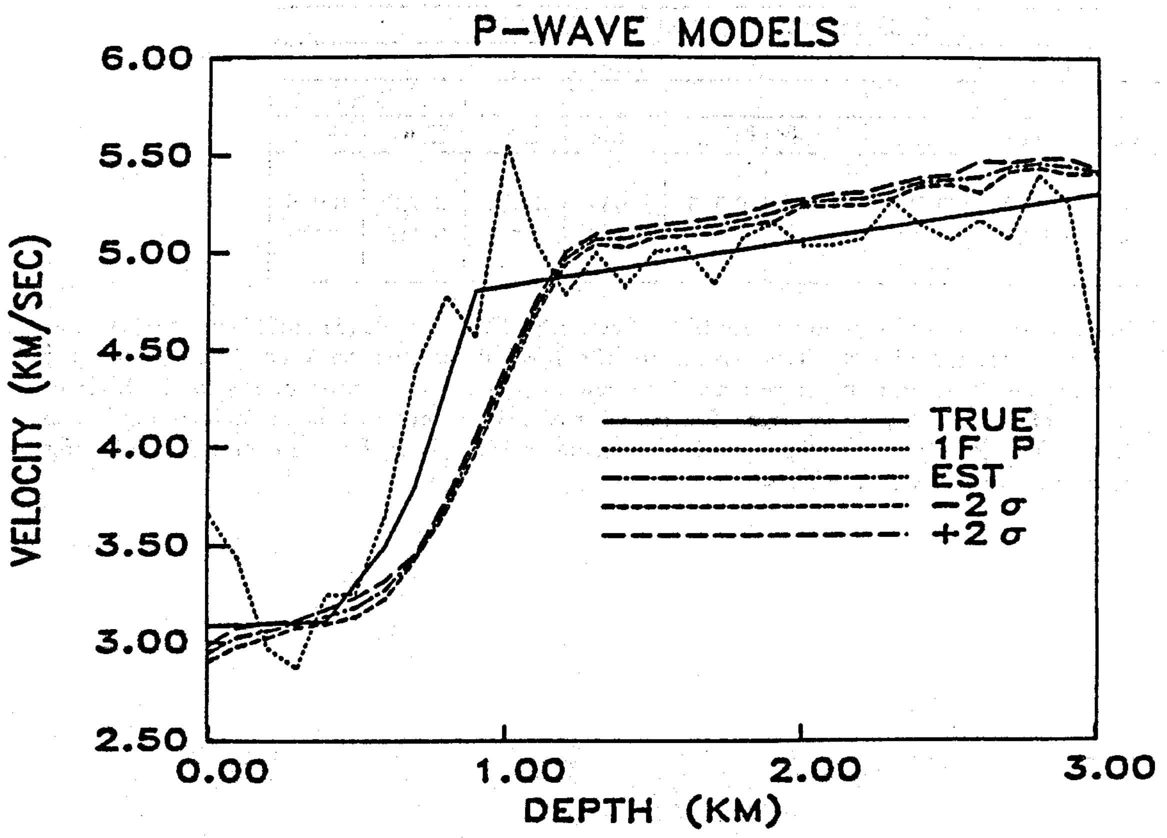


Figure 3.13 The true P-wave velocity model (solid line) filtered by the resolving kernels of the estimated model for the starting model (1) case (F) inversion is denoted by 1F P. The estimated model is denoted by EST where -2 σ and + 2σ are the minimum and maximum 2 standard error limit estimates.

RMS Station Correction Error (sec)								
	P only		Joint P and S					
	P		P		S		Total	
Model	True	Est.	True	Est.	True	Est.	True	Est.
1	0.023	0.024	0.019	0.035	0.059	0.038	0.039	0.036
2	0.059	——	0.055	——	0.084	——	0.067	——

Table 3.4 All tests correspond to case (F) of Section (3.3.2). Models (1) and (2) are the same as described in Section (3.3.1). True errors are the RMS differences between estimated station corrections and true station corrections. The column labeled Est., corresponds to the RMS of estimates of station correction error. Estimated errors were only calculated for starting model (1) inversions. P-only refers to progressive inversions using only P data and joint P and S refer to simultaneous P and S inversions.

S data were used in case (F). Since only a single synthetic model could be considered here, it is not known whether this result will be true in general. However, we suspect that adding S-wave data to progressive inversion will reduce dc velocity errors in general.

The quantity V_p/V_s provides an important constraint on material properties of rocks. Figure (3.14) show estimated V_p/V_s models for starting model (1) inversion cases (C) and (F). A smoothed version of the shape of the true V_p/V_s model is recovered except for the V_p/V_s minimum at 1.5 km depth. The V_p/V_s maximum is displaced with depth in a pattern similar to that found for the P and S velocity model estimates. The absolute V_p/V_s level is not well determined but overall changes of V_p/V_s with depth are obtained. The inability to closely reproduce the true V_p/V_s model indicates that the dc velocity ambiguity is amplified in V_p/V_s. Consequently, interpretation of V_p/V_s estimates should concentrate more on relative variations of V_p/V_s with depth than absolute V_p/V_s magnitude.

This concludes the velocity structure section. Next, slowness gradient will be used to investigated how well linearized resolution and error predictions work. Also, estimates of slowness gradient are compared as a function of the number of unknowns inverted for and the types of data used in inversions.

3.4. Slowness Gradient Structure

3.4.1. Effects of Adding S-waves, Station Corrections, and Hypocenters

There are two objectives here. The first is to determine what effect adding S-wave data to progressive inversion has on estimates of P and S-wave slowness gradient. The second is to determine the effects of adding hypocenter and then station corrections as unknowns (cases (C) and (F) of Section 3.3.1) on estimates of slowness gradient.

P-wave slowness gradient is more accurately recovered in joint P and S inversions than P-only inversions. Figure (3.15) shows the P-only inversion results for case (C) and Figure (3.16) shows P-wave slowness gradient estimates for the case (C) joint P and S inversion. P-wave

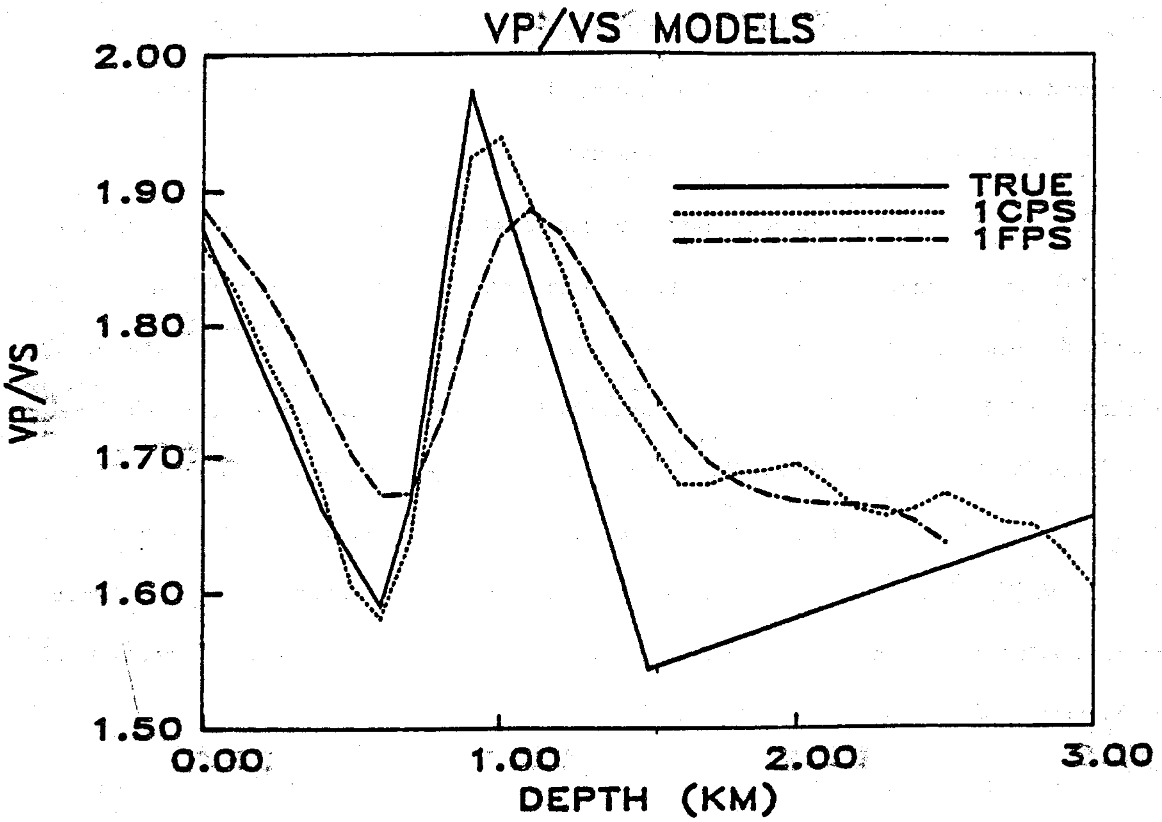


Figure 3.14 The true V_p/V_s model (solid line) is compared to estimated V_p/V_s for case (C) starting model (1) denoted as 1CPS, and case (F) starting model (1) denoted as 1FPS.

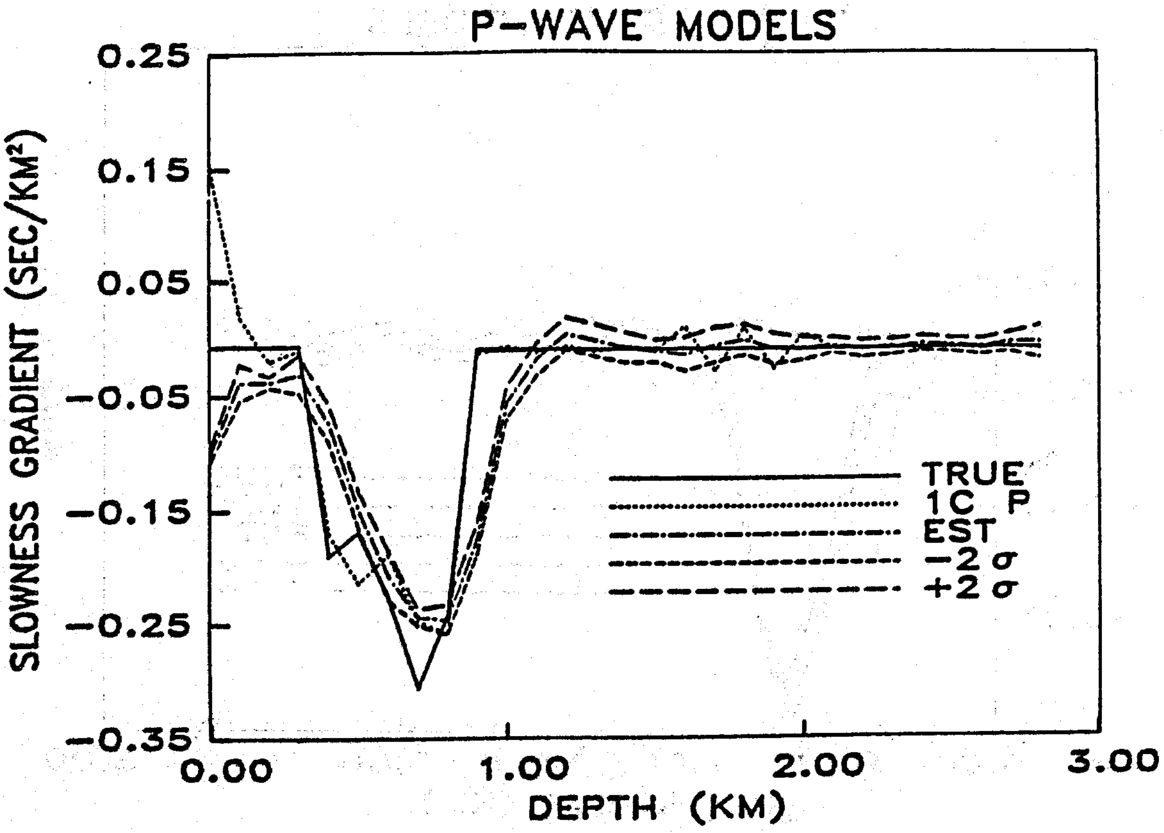


Figure 3.15 The true P-wave slowness gradient model (solid line) filtered by the resolving kernels for the case (C) starting model (1) P-only inversion is denoted as 1C P. The estimated model is denoted by EST where -2σ and $+2\sigma$ are the minimum and maximum 2 standard error limit estimates.

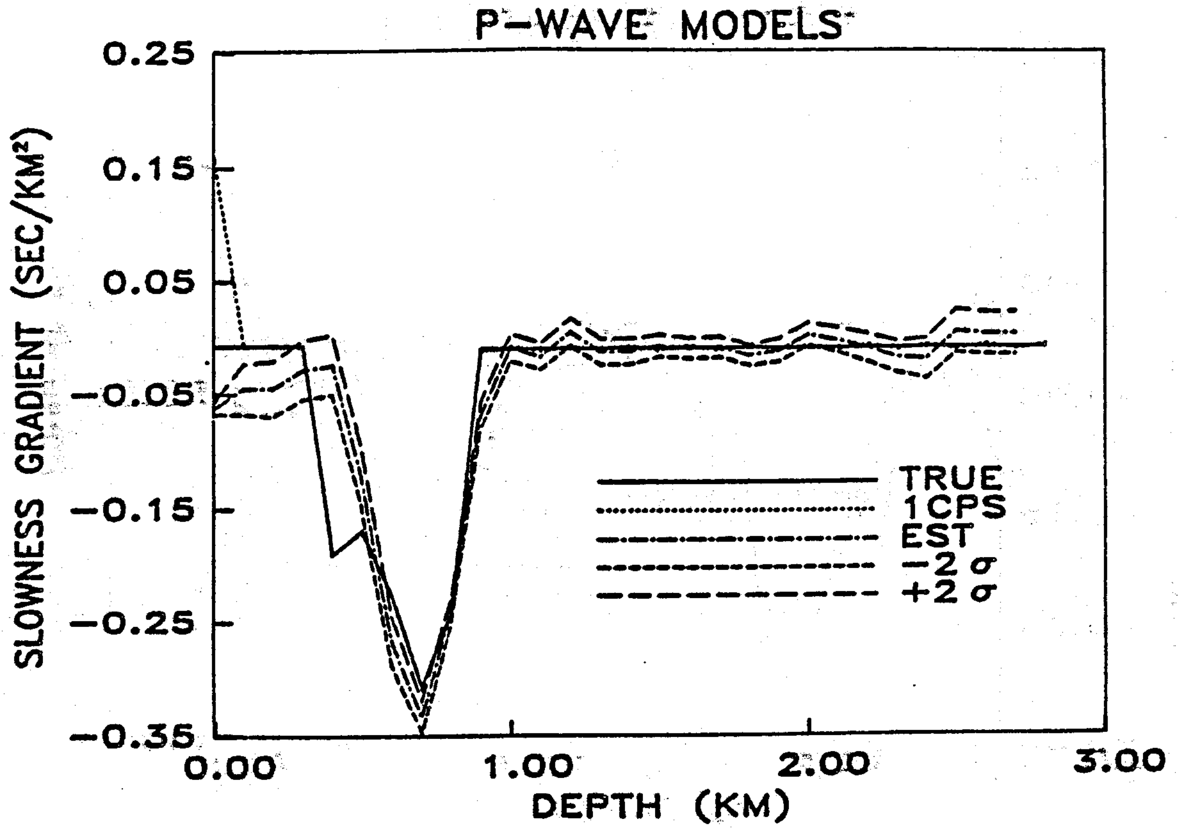


Figure 3.18 The true P-wave slowness gradient model (solid line) filtered by the resolving kernels for the case (C) starting model (1) joint P and S inversion is denoted as 1CPS. The estimated model is denoted by EST where $-2\,\sigma$ and $+2\,\sigma$ are the minimum and maximum 2 standard error limit estimates.

slowness gradient is very accurately estimated throughout most of the model depth extent in the joint inversion whereas the large negative slowness gradient is underestimated and shifted in depth in the P-only inversion. This is in spite of the fact that more than twice as many iterations were done (11 versus 5) for the P-only inversion.

The same pattern can be seen in the case (F) results. The P-only inversion for slowness gradient (Figure (3.17)) underestimates the magnitude of the large negative slowness gradient and its depth position is incorrect. These errors are much smaller for the joint inversion P-wave slowness gradient estimate shown in Figure (3.18). The P-only inversion misfit is larger even though the final estimate RMS residual is one third that of the joint inversion, and more iterations were done (5 versus 3).

The addition of S-wave data substantially improved estimates of P-wave slowness gradient structure over that estimated using P-wave data alone. Since case (C) requires hypocenter estimation and case (F) requires hypocenter and station correction estimation, it is clear that the S-wave data have help reduce errors in slowness structure estimates that arise from hypocenter and station correction errors.

Case (C) results can be compared to the results of Pavlis and Booker (1983a). They noted difficulties with tradeoffs between hypocentral depth estimates and velocity model features. The dc shift of P-wave slowness gradient with depth for the case (C) P-only inversion correlates with systematic hypocentral depth mislocations to deeper depths (Figure (3.34)). The dc slowness gradient depth shift is expected to positively correlate with hypocentral shifts. There are numerous hypocenters in the depth interval corresponding to the large negative slowness gradient (Figure 3.2). Estimates of hypocentral depth are shifted in the case (C) P-only inversion to produce travel time curves with a shape a close as possible to the true travel time curves. In contrast, the hypocenter depth errors are much smaller for the case (C) joint inversion (Figure (3.32)), and the correct slowness gradient model is almost completely recovered.

The fact that the case (F) joint inversion (Figure (3.18)) did not reproduce the true slowness gradient model as well as the case (C) joint inversion did, indicates that estimating station

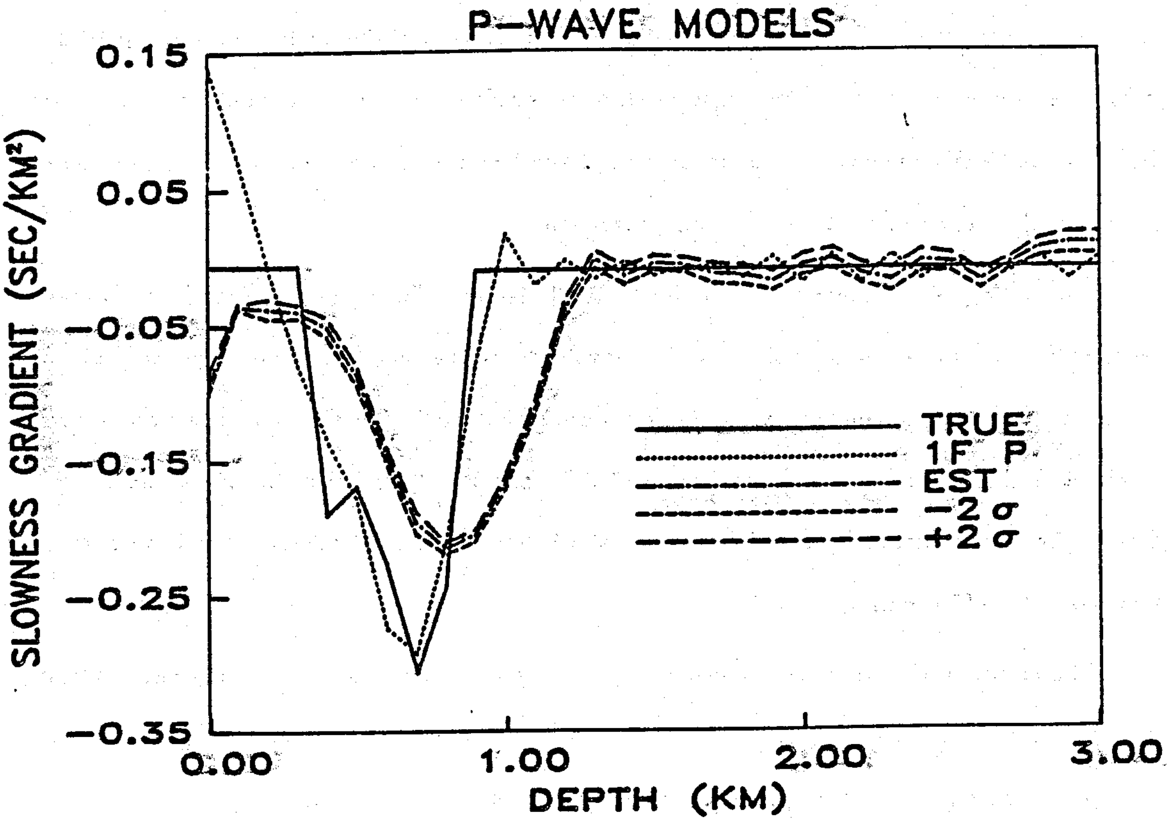


Figure 3:17 The true P-wave slowness gradient model (solid line) filtered by the resolving kernels for the case (F) starting model (1) P-only inversion is denoted as 1F P. The estimated model is denoted by EST where −2 σ and +2 σ are the minimum and maximum 2 standard error limit estimates.

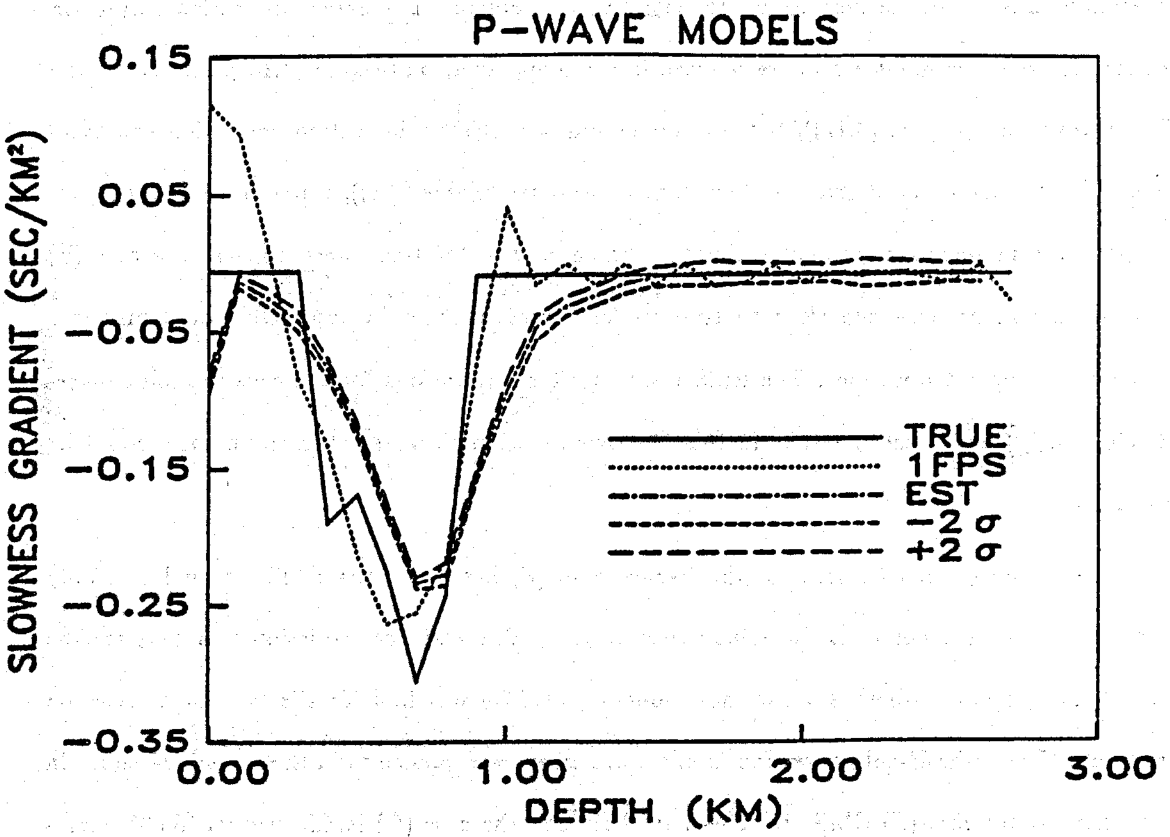


Figure 3.18 The true P-wave slowness gradient model (solid line) filtered by the resolving kernels for the case (F) starting model (1) joint P and S inversion is denoted as 1FPS. The estimated model is denoted by EST where -2σ and $+2\sigma$ are the minimum and maximum 2 standard error limit estimates.

corrections has added another set of tradeoffs to the problem. The extent of station correction induced slowness gradient errors are reduced in the joint inversion (Figure (3.18)), relative to the P-only inversion (Figure (3.17)), but are not eliminated. Errors in station correction estimates (Table (3.4)) have induced errors in hypocenter estimates (Table (3.5)), which have in turn, produced errors in slowness gradient estimates. Since only 3 iterations were done in the case (F) joint inversion, convergence close to true model values, such as in case (C), would probably occur in subsequent iterations. The RMS residual after 3 iterations for the case (F) joint inversion was 0.01 sec, indicating that further improvements in slowness gradient models could be obtained.

The same cannot be said for the P-only case (F) inversion; the RMS residual was only 0.003 sec, much smaller than the joint inversion case. The variance reduction was 28% for the last P-only iteration whereas a variance reduction of 59% was had for the last joint inversion iteration. Thus, the P-only inversion is much closer to convergence than the joint inversion. In fact, the P-only case (F) RMS residual is smaller than the case (C) RMS residual (0.003 versus 0.005 sec). Yet, the estimates of P-wave slowness gradient in case (C) are much more accurate (Figure (3.15)). This indicates that the case (F) P-only inversion converged to a local minimum.

It appears that P-wave data alone do not provide enough information to completely recover slowness gradient structure and station corrections using earthquake data. However, except for dc-depth shift, the true slowness gradient magnitude is well estimated in Figure (3.17)). This contrasts sharply with analogous tests of estimated velocity structure (Figure (3.13)), where estimated velocity magnitudes are incorrect throughout virtually the entire model depth extent.

RMS Hypocenter Error								
	Origin time (sec)		Latitude (km)		Longitude (km)		Depth (km)	
Case	True	Est.	True	Est.	True	Est.	True	Est.
1C P	0.034	0.050	0.025	0.189	0.014	0.094	0.081	0.314
1CPS	0.014	0.023	0.008	0.072	0.005	0.046	0.031	0.130
1F P	0.043	0.054	0.071	0.240	0.069	0.115	0.226	0.342
1FPC	0.084	—	0.125	—	0.214	—	0.326	—
1FPS	0.053	0.051	0.074	0.168	0.115	0.097	0.086	0.282
2F P	0.135	—	0.288	—	0.336	—	0.514	—
2FPS	0.148	—	0.390	—	0.415	—	0.544	—

Table 3.5 True and estimated RMS hypocenter errors as a function of inversion case. The column labeled Est., contains the RMS estimated errors and the column labeled True, contains the RMS misfit between true and estimated hypocenter parameters. The first number in the column labeled case, corresponds to the starting model used, the second character to the case list in Table (3.3), and the last two characters indicate where a joint P and S inversion was done (PS) or if only P data were used (P). Case (1FPC) is the P-only inversion result with a RMS data misfit comparable to case (1FPS).

3.4.2. Resolution and Error

The data kernels for slowness gradient are finite and square integrable facilitating resolution-error analysis using equations (2.59) and (2.60). Resolution and error were determined using the RMS data misfit from the inversions to scale error estimates. Expansion order in equations (2.59) and (2.60) were chosen as the maximum rank, below which, most model errors flattened out. This approach uses true data misfit as an estimate of statistical errors in the data. Since the data are error free, model uncertainties calculated in this manner should be close to true misfit errors.

In each of Figures (3.15-3.20), the true slowness gradient is shown along with the result of filtering the true slowness gradient by the resolving functions calculated from the corresponding estimated model. The estimated model should fit the filtered true model to within its error bounds. If it does not, this indicates that nonlinear effects are important; the estimated model is not linearly close to the truth.

Figure (3.16) shows the only case were linearized resolution and error estimates make completely accurate predictions. Slowness gradient is recovered everywhere except near sharp discontinuities and close to the free surface where predicted resolution is poor. Predictions of resolution and error are less accurate near slowness gradient discontinuities in the other cases. Resolution errors are reduced in the case (C) joint inversion (Figure (3.18)) compared to the P-only inversion (Figure (3.17)). The case (C) S-wave slowness gradient estimates (Figure (3.19)) do not recover structure as well as the P-wave structure in the neighborhood of slowness gradient discontinuities. Since there are only 294 S-wave arrival times versus 476 P-wave arrival times, and S-wave arrivals are confined to a smaller distance range, poorer S structure resolution is to be expected.

The problem of incorrect resolution and error predictions is more pronounced for the case (F) S-wave inversion (Figure (3.20)), indicating that station corrections have had a more profound effect on S-wave structure than P-wave structure. The larger S-wave station correction errors (Table (3.4)) bear this out. However, the amount of smearing out of slowness gradient

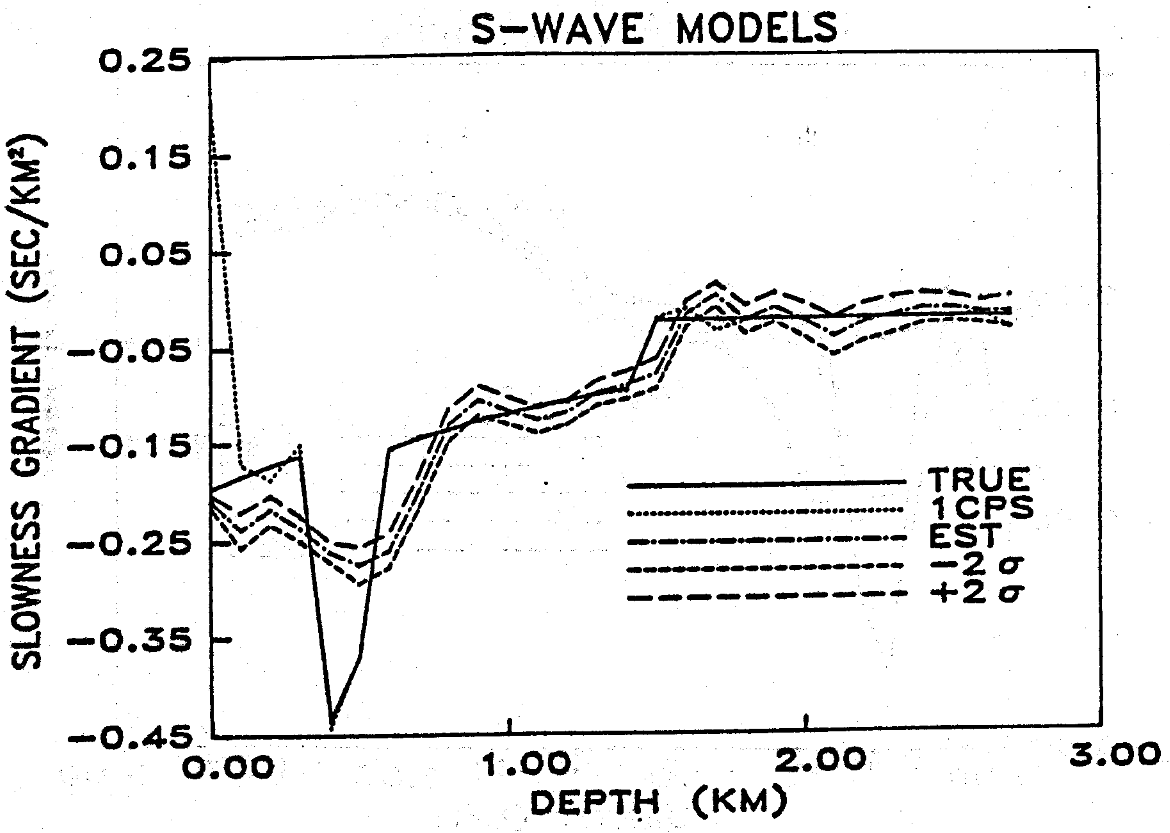


Figure 3.19 The true S-wave slowness gradient model (solid line) filtered by the resolving kernels for the case (C) starting model (1) joint P and S inversion is denoted as 1CPS. The estimated model is denoted by EST where -2 σ and +2 σ are the minimum and maximum 2 standard error limit estimates.

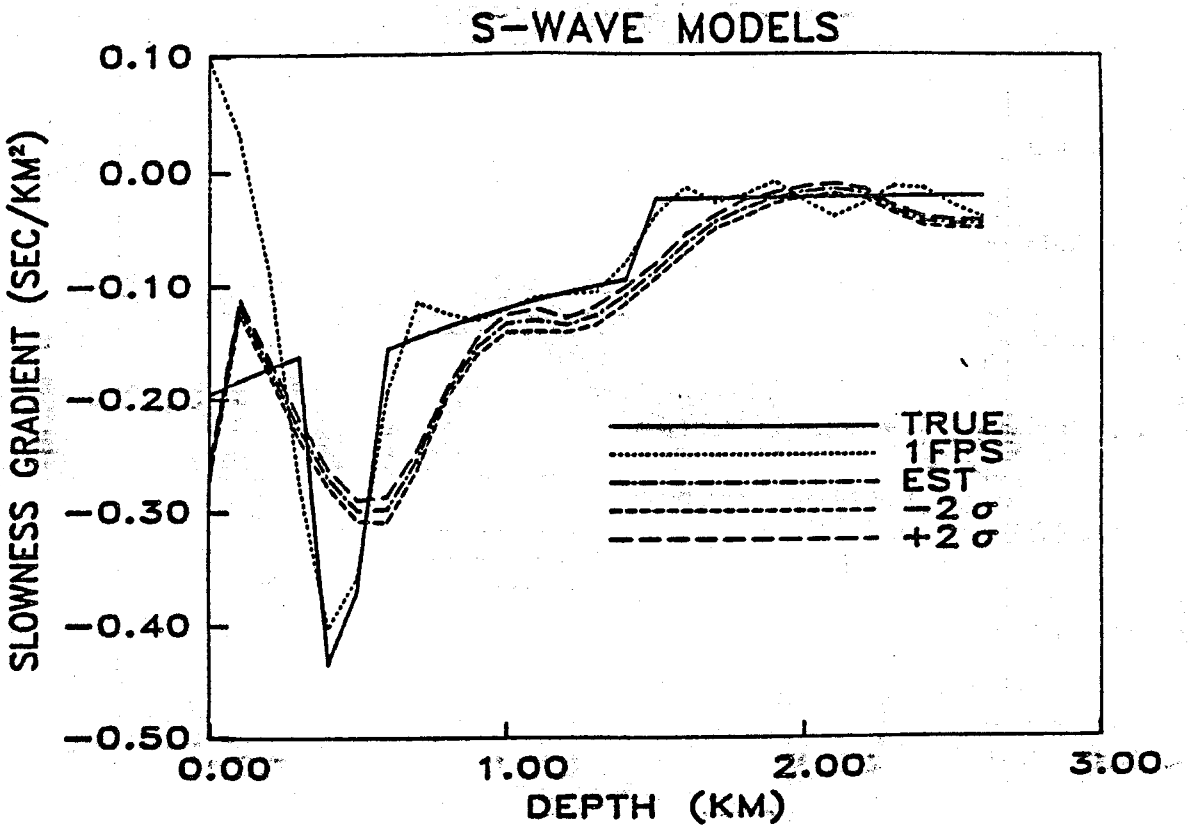


Figure 3.20 The true S-wave slowness gradient model (solid line) filtered by the resolving kernels for the case (F) starting model (1) joint P and S inversion is denoted as 1FPS. The estimated model is denoted by EST where -2σ and $+2\sigma$ are the minimum and maximum 2 standard error limit estimates.

structure for P and S models for case (F) is comparable; the S model simply has more slowness gradient discontinuities in its true model.

Linearized resolution and error analysis tends to produce overly optimistic estimates of resolution, especially near model discontinuities. Pavlis and Booker (1983) demonstrated this in a much more complete investigation of this problem than was considered here. The important point is that the inclusion of S-wave data has reduced errors made in linearized resolution and error analysis. These errors still exist when S-wave data are used. Probably the safest means to determine what features are truly resolvable in model estimates is to do synthetic tests similar to those done here.

3.5. Station Corrections

Station corrections are required to account for lateral velocity variations that are not included in one-dimensional velocity model parameterization. Many times the most pronounced lateral velocity variations are confined to the near surface. It would be preferable to determine actual receiver-site velocity variations than to use station corrections to compensate for them. Often this in not feasible and station corrections must be employed. As was found in Section (3.3), the price paid is reduced resolution of overall velocity structure.

As with velocity inversions, assessment of station correction estimates with synthetic data involves factor analysis, although the scope is reduced. The error prediction approach developed in Section (2.6.2) of Chapter 2 is also tested. Finally, station correction resolution is briefly addressed.

3.5.1. Factor Analysis

Only two cases need be considered: (1) inversion for station corrections with true velocity models, and (2) inversion for station corrections as a component of progressive inversion for velocity structure. Case (1) corresponds to progressive multiple event location (PMEL) of Pavlis and Booker (1983b). Case (2) corresponds to case (F) of Section (3.3.1.3). As in Section (3.3), the cases of joint P and S inversion and P-only inversion are also considered.

The initial values for all station corrections were set to zero for all tests. This represents the minimal starting model assumption and is a good test of the robustness of station correction estimation. Considering the substantial true synthetic station correction magnitudes (Table (3.2)), the use of zero initial station corrections is a strong test. The true synthetic station corrections for both P and S have zero mean as do the initial values.

3.5.1.1. Case (1): PMEL Results

Convergence to true hypocenters and station corrections was obtained in all cases. Fewer station correction inversion iterations were required to attain convergence for joint P and S PMEL inversions than with P-only PMEL inversions. The fastest convergence was obtained using a synthetic data set which consisted of matched P and S arrival times at all recording stations used for each event location. Overall, PMEL proved to be robust; hypocenters and station corrections were accurately recovered despite using starting locations and initial station corrections that were not particularly close to the true values.

3.5.1.2. Case (2): Progressive Inversion

Results of progressive inversions involving station correction estimation are summarized in Table (3.4). The model numbers correspond to the same starting velocity models described in Section (3.2.1). Errors in station corrections depended weakly on whether joint P and S data or only P-data were used. The most significant factor controlling station correction errors was starting velocity model. Station correction estimates obtained using starting model (2) had much larger errors than those obtained using starting model (1) (Table (3.4)). This result serves to emphasize the importance of using the best possible starting model when doing progressive inversions.

3.5.2. Station Correction Error Estimates

Estimates of station correction errors obtained using the method outlined in Section (2.6.2) of Chapter 2 are shown for starting model (1) in Table (3.4). The observed data misfit of the synthetic data was used to scale the error estimates. The error estimates in Table (3.4) are within a factor of 2 of the true errors. The primary contribution to the estimated errors comes from the velocity model error contribution, $\mathbf{C}_S$, of equation (2.83). The conventional statistical error term, $\mathbf{C}_t$, of equation (2.83) is 2 to 3 orders of magnitude smaller than $\mathbf{C}_S$ in all cases. The estimated contribution of $\mathbf{C}_t$ should be small since the data are error free. The entire station correction error is due to velocity model induced errors since true station corrections are recovered in PMEL tests. Consequently, if only conventional statistical errors had been considered, station correction errors would be grossly underestimated.

The same expansion orders were used to compute $\mathbf{C}_S$ as were used in slowness gradient resolution and error appraisal of Section (3.4.2). What is a bit surprising, is that the estimates of station correction errors are so close to the true errors. Station correction error estimates may be fairly accurate because station corrections are linear variables in progressive inversion. It is not known whether these synthetic error results would be as accurate in general.

3.5.3. Station Correction Resolution

In all the cases considered in Table (3.4) the station correction matrix $\mathbf{R}_s$ of equation (2.48) had $n_s - 1$ nonzero singular values and a corresponding null vector with elements equal to one over the square root of the number of station corrections as predicted by Pavlis and Booker (1983b). Even the matched P and S arrival time data set did not produce a full rank station correction inversion. A full rank station correction inversion might be attained in only S-P times were used thereby eliminating origin time from progressive inversion. However, mean P and S station corrections would still need to be constrained to zero to ensure the integrity of subsequent velocity inversions.

3.6. Hypocenters

The accurate determination of hypocenter position is critical to afford maximum possible utility for geological interpretations. These synthetic tests of hypocenter estimation are designed to determine the sensitivity of hypocentral estimates errors in velocity models and station corrections. The accuracy of hypocentral error estimates determined using the approach of Section (2.6.2) are also investigated.

3.6.1. Factor Analysis

The cases considered here are listed in Table (3.5) and correspond to cases (C) and (F) of Section (3.3.2). Starting model (2) results are only considered for case (F) inversions.

The most significant factor controlling hypocentral errors is starting velocity model. Starting model (2) mislocation errors are much larger than those for starting model (1) (Table (3.5)). For case (F) starting model (2), this can be clearly seen in Figures (3.21) and (3.22) of the P-only inversion, and Figures (3.23) and (3.24) of the joint inversion. Hypocentral errors are actually larger for the joint inversion than for the P-only inversion (Table (3.5)). This is due to the large S-wave station correction errors in the starting model (2) joint inversion (Table (3.4)). In contrast, the case (F) starting model (1) joint inversion (Figures (3.27) and (3.28)) and P-only inversion (Figures (3.29) and (3.30)) have much smaller mislocation errors (Table (3.5)).

An interesting feature of the mislocations in Figures (3.21-3.24) is their systematic nature. Large epicentral errors are not expected for earthquakes located using well distributed recording stations. One-dimensional velocity model errors alone do not produce significant epicentral errors for the recording geometry used here. The systematic epicentral mislocations in Figures (3.21) and (3.23) are due to combined station correction and velocity model errors. This illustrates yet another deleterious effect of using an incorrect starting model with artificially high near surface velocities and assuming that velocity model errors will be absorbed into the station corrections. If these errors are undetected using real data (and they might well go undetected if only P data are used (see Section (3.3.1.3))) the results can be misinterpretation of earthquake

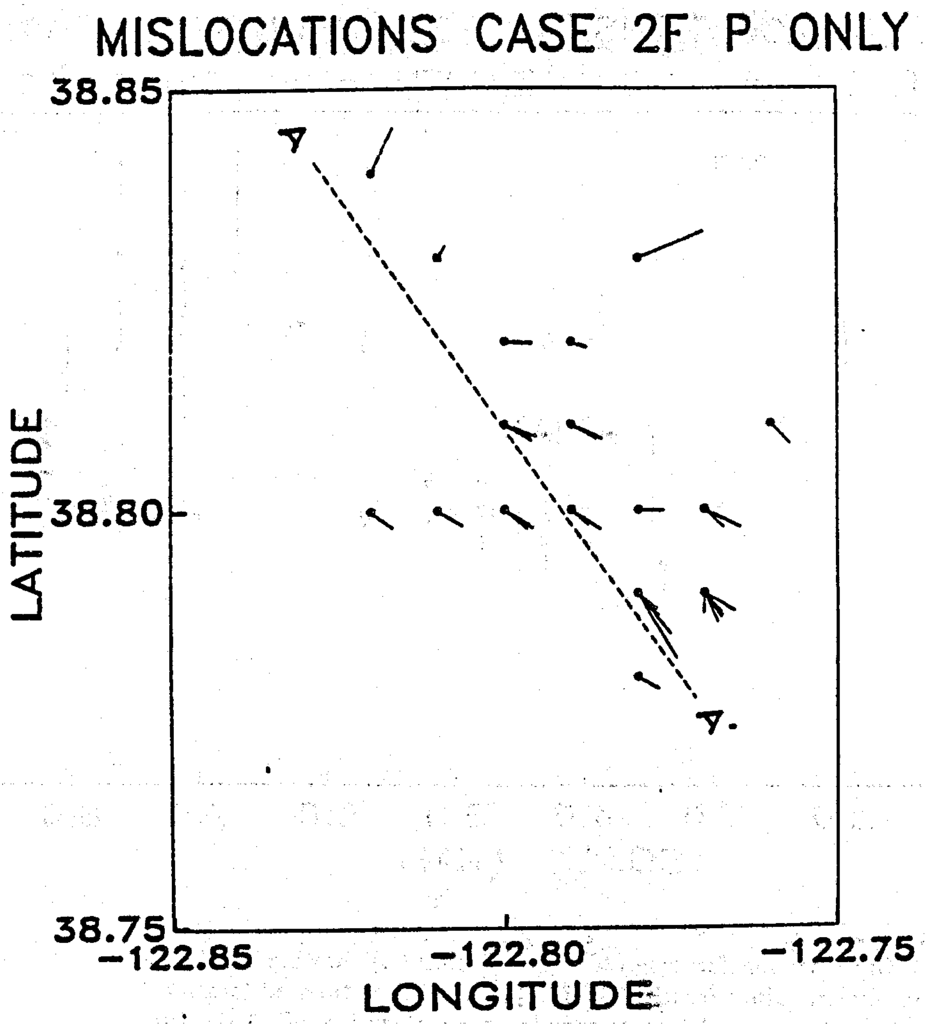


Figure 3.21 Epicenter mislocations for case (2F P) of Table (3.5). Small circles show true epicentral locations and lines point from the circles to estimated epicenter locations. The circles with multiple lines correspond to multiple events at the same true epicenter. The dashed line from A to A′ is the surface projection of the accompanying cross section shown in Figure (3.22).

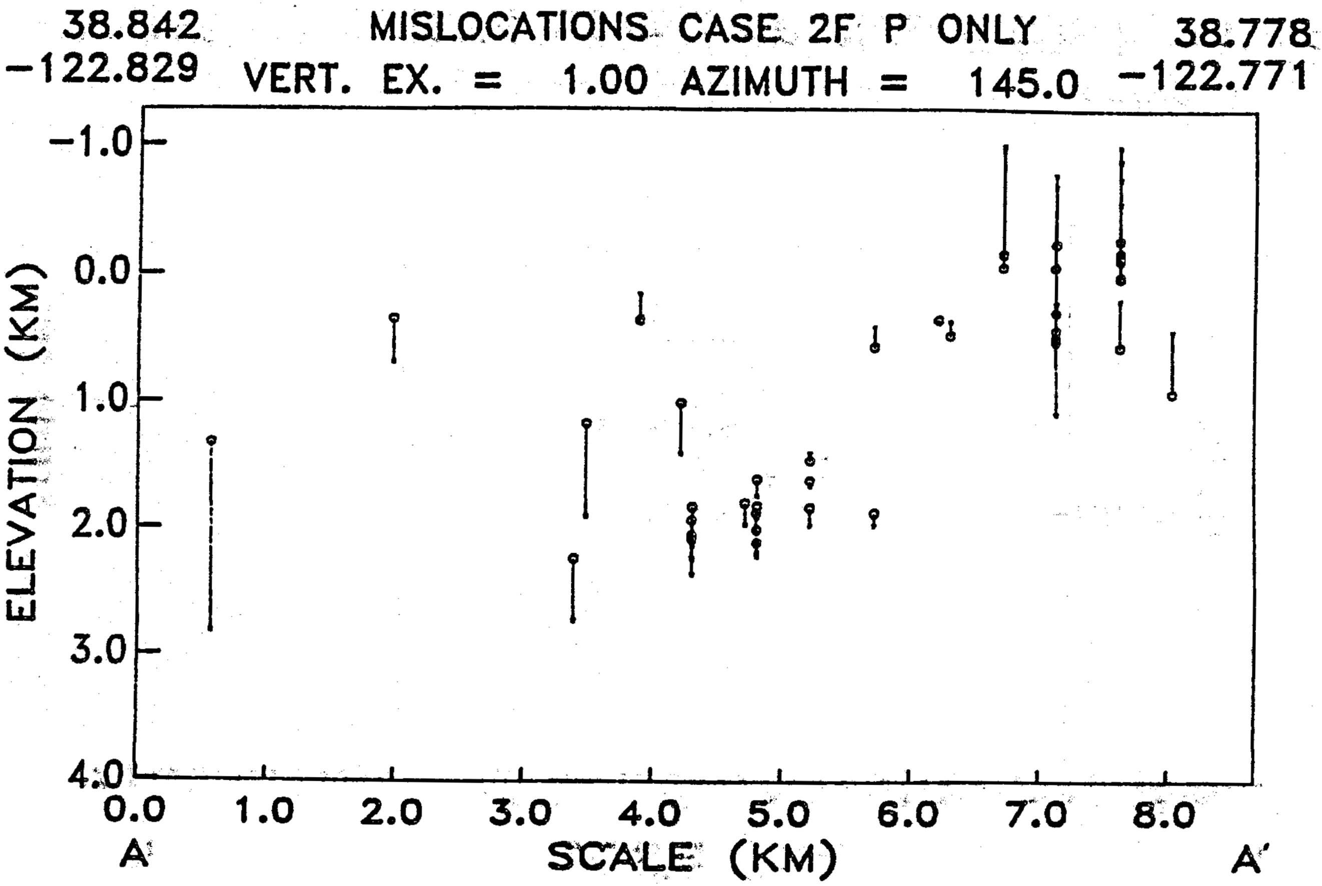


Figure 3.22 Cross section for case (2F P) of Table (3.5) showing true locations as small circles. Lines from the circles point to the estimated locations. The top of the velocity models corresponds to an elevation of -0.379 km. Note that several events have located above the top of the velocity model in the constant velocity elevation correction layer.

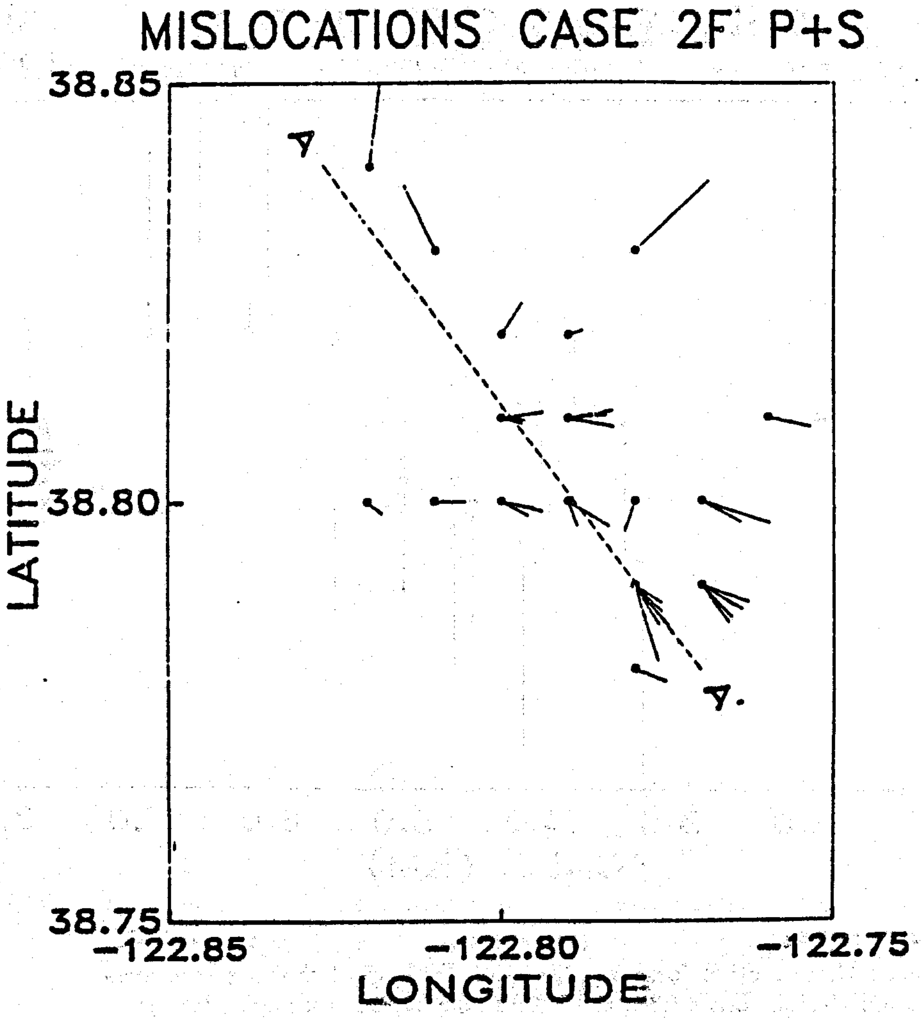


Figure 3.23 Epicenter mislocations for case (2FPS) of Table (3.5). Small circles show true epicentral locations and lines point from the circles to estimated epicenter locations. The circles with multiple lines correspond to multiple events at the same true epicenter. The dashed line from A to A′ is the surface projection of the accompanying cross section shown in Figure (3.24).

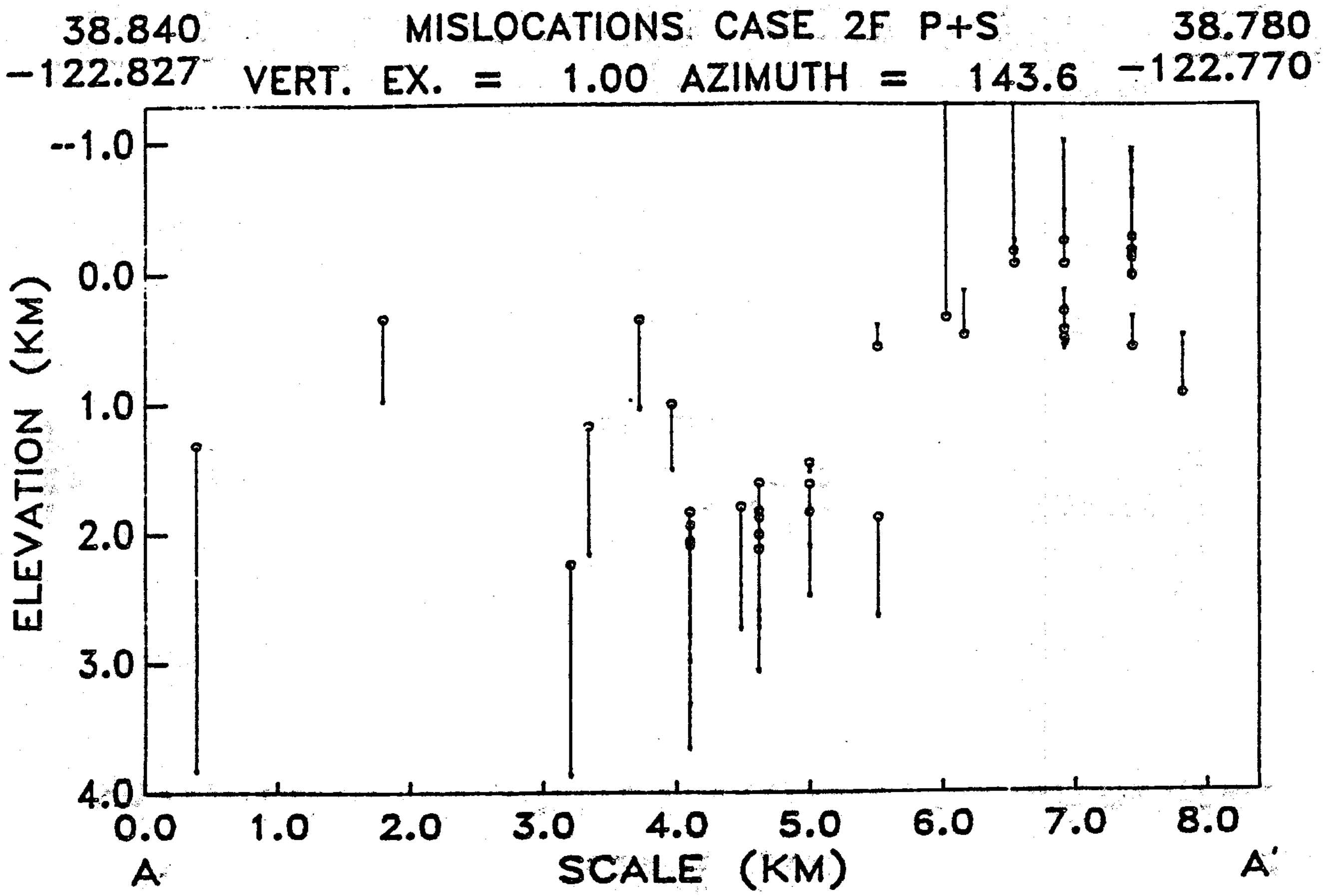


Figure 3.24 Cross section for case (2FPS) of Table (3.5) showing true locations as small circles. Lines from the circles point to the estimated locations. The top of the velocity models corresponds to an elevation of -0.379 km. Note that many of the shallowest events have located above the top of the velocity model in the constant velocity elevation correction layer.

locations with respect to geologic features.

The second strongest factor affecting epicentral estimates is whether station corrections are included as unknowns. This is not the case for hypocentral depth where the joint P and S inversion with station correction has an RMS misfit comparable to the P-only inversion without station corrections (cases (1FPS) and (1C P) of Table (3.5), respectively). RMS depth misfits are a factor of 2.6 greater for P-only inversions (cases (1C P) and (1F P) of Table (3.5), respectively) relative to joint P and S inversions of the same case (cases (1CPS) and (1FPS) of Table (3.5), respectively). This is in spite of the fact that the case (1F P) inversion has an RMS data misfit of only 0.003 sec, whereas the joint inversion (1FPS) has an RMS data misfit of 0.01 sec. When a case (F) P-only inversion with comparable misfit, (case (1FPC), obtained by using results of an iteration preceding convergence), is compared to the joint P and S inversion (1FPS), the error disparity increases substantially as can be seen in Figures (3.25) and (3.26) for case (1FPC) versus Figures (3.27) and (3.28) for case (1FPS) (see also Table (3.5)). Similar results are obtained for case (C) P-only versus joint P and S inversions (Table (3.5)).

3.6.2. Estimates of True Misfit

Estimates of hypocenter misfits errors are shown in Table (3.5) for starting model 1 cases (C) and (F). Both station correction induced errors, equation (2.50) (which also include their own component of velocity model induced errors), and velocity model induced errors, $\mathbf{C}_S$, of equation (2.83), are added to the conventional statistical error estimates, $\mathbf{C}_t$ of equation (2.83), to produce the error estimates in Table (3.5). The same expansion orders were used to compute $\mathbf{C}_S$ that were used in Sections (3.3.4) and (3.3.5).

For all events, the velocity model error component was dominant, followed by station correction error contributions approximately a factor of 3-4 smaller, and conventional statistical error approximately an order of magnitude smaller. Estimates of total hypocenter error tend to be larger than true misfit errors indicating that velocity model contributions to total hypocentral errors are overestimated. Total hypocenter misfit errors are especially overestimated in case

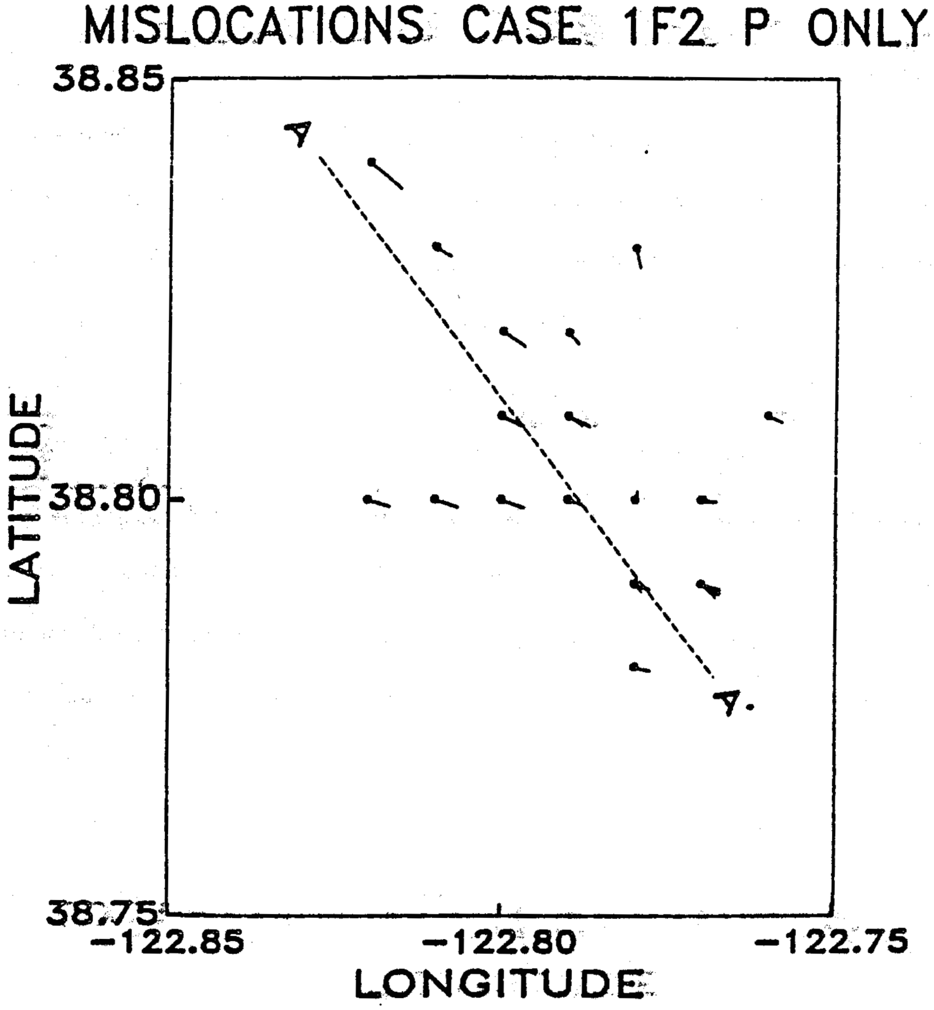


Figure 3.25 Epicenter mislocations for case (1FPC) of Table (3.5). Small circles show true epicentral locations and lines point from the circles to estimated epicenter locations. The circles with multiple lines correspond to multiple events at the same true epicenter. The dashed line from A to A′ is the surface projection of the accompanying cross section shown in Figure (3.26).

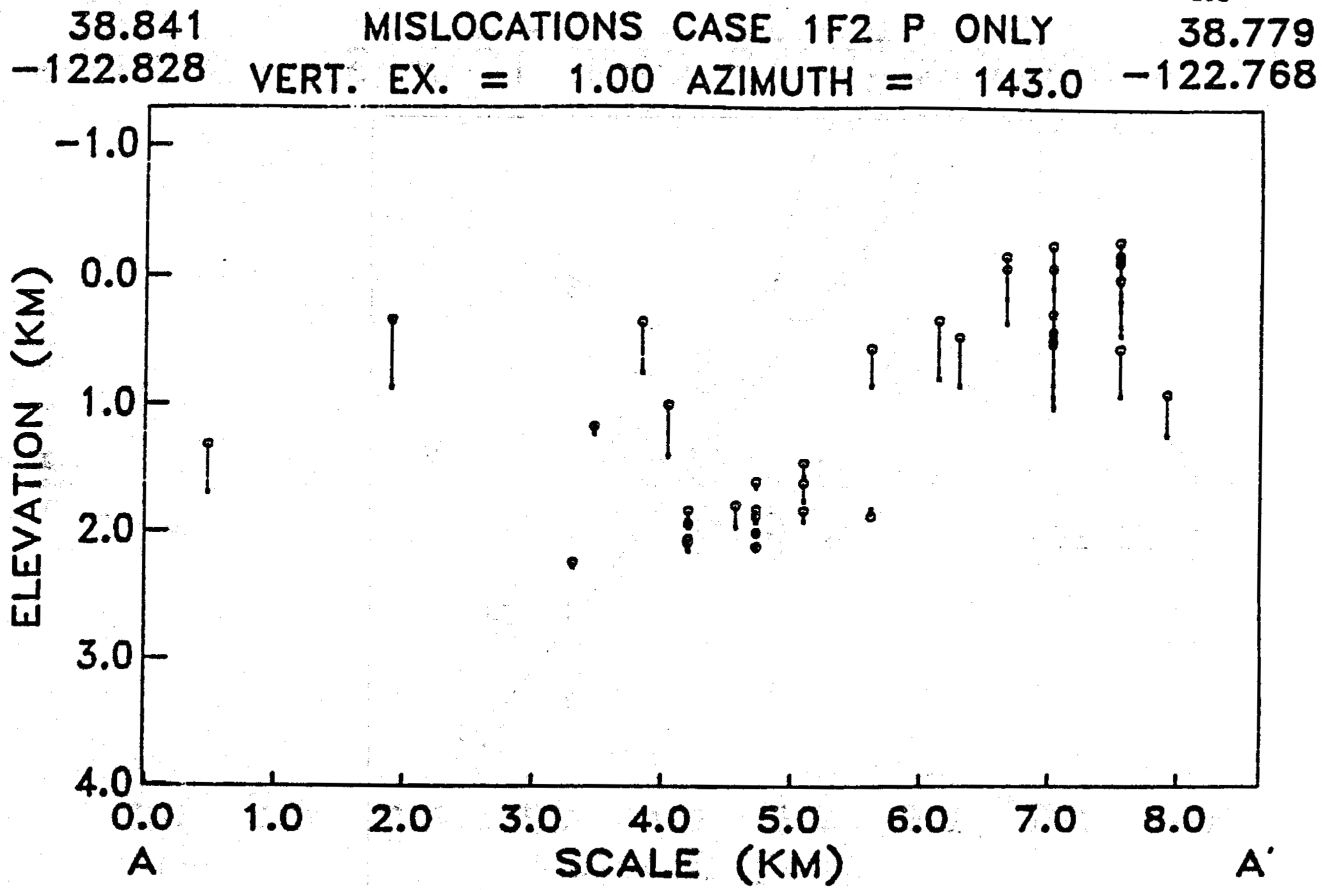


Figure 3.26 Cross section for case (1FPC) of Table (3.5) showing true locations as small circles. Lines from the circles point to the estimated locations. The top of the velocity models corresponds to an elevation of -0.379 km. Note the systematic mislocation of many of the shallowest events.

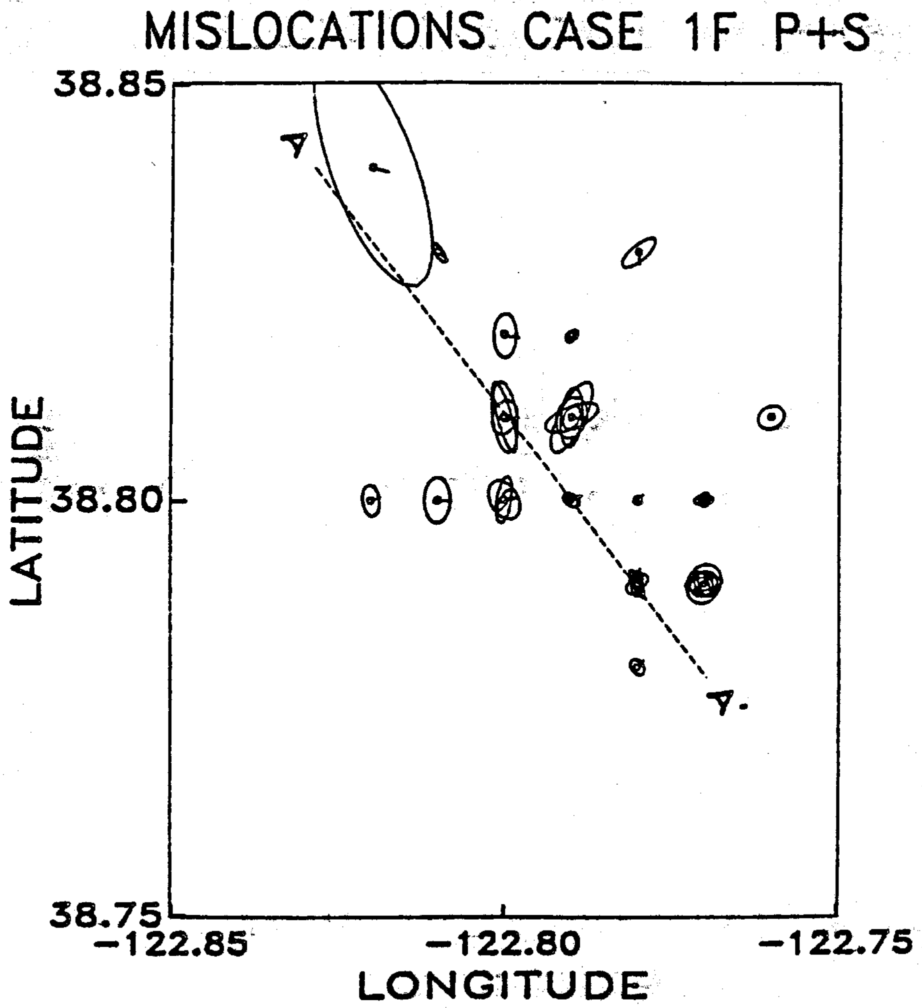


Figure 3.27 Epicenter mislocations for case (1FPS) of Table (3.5). Small circles show true epicentral locations and lines point from the circles to estimated epicenter locations. The circles with multiple lines correspond to multiple events at the same true epicenter. Two standard error confidence ellipses calculated using all sources of error are also shown. The dashed line from A to A′ is the surface projection of the accompanying cross section shown in Figure (3.28).

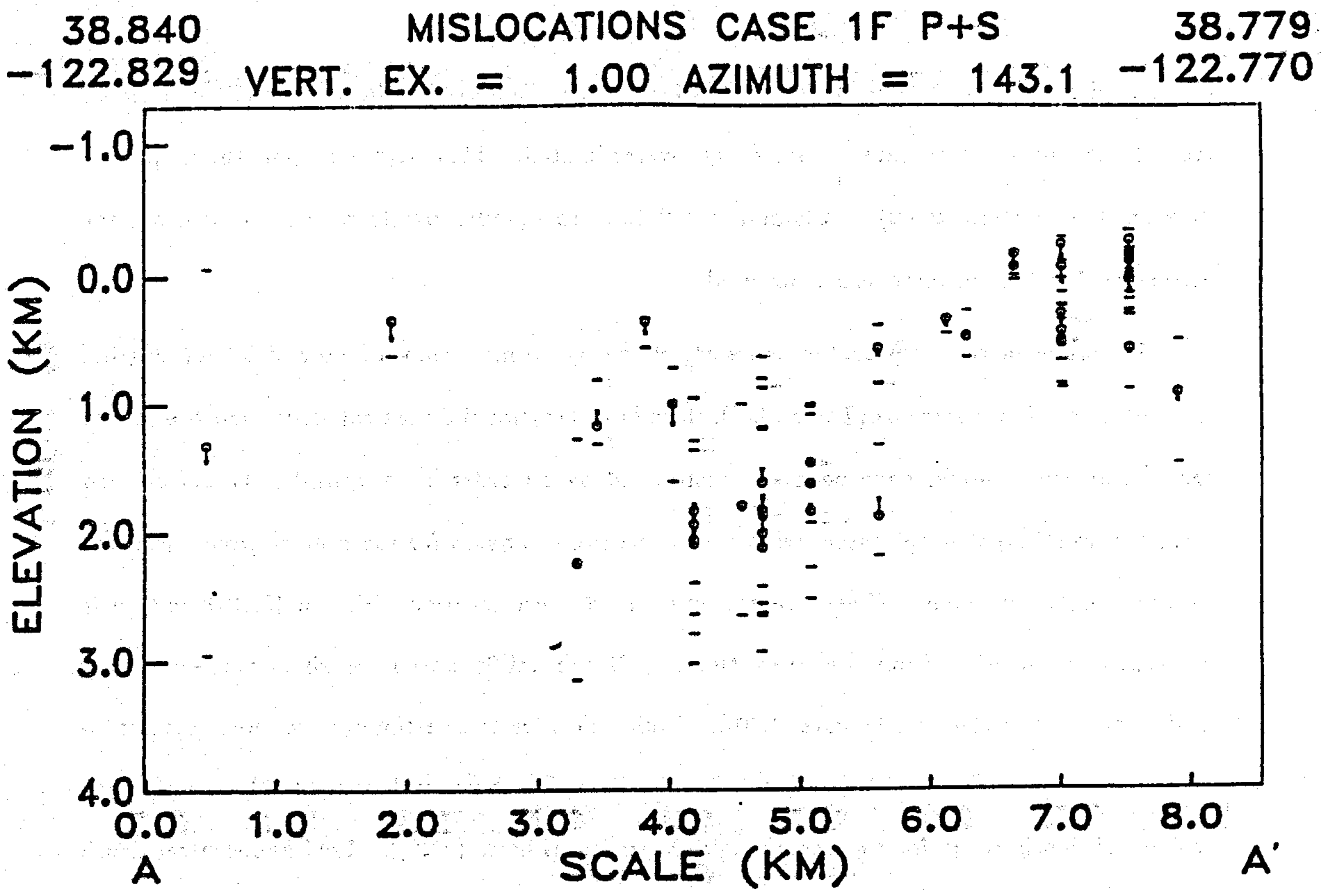


Figure 3.28 Cross section for case (1FPS) of Table (3.5) showing true locations as small circles. Lines from the circles point to the estimated locations. The top of the velocity models corresponds to an elevation of -0.379 km. Two standard error limits estimated using all sources of error are plotted as small bars about the estimated location. Note that all true locations lie between the error limits.

(C) inversions as can be seen in Figures (3.31-3.34) and Table (3.5). In all case (C) inversions, true depth misfit errors are considerably overestimated. This suggests that the approach developed in Section (2.6.2) of Chapter 2 will tend to produce overly pessimistic depth error estimates if station corrections are not used.

Estimates of true hypocenter errors are wrong by no more that a factor of 3.4 for starting model (1) case (F) inversions (Table (3.5)). If only conventional statistical errors had been used, true misfit errors would have been underestimated by an order of magnitude. It can be seen from Figure (3.29) that epicenter errors tend to be overestimated for the central group of events of the P-only inversion. These events have the deepest locations (Figure (3.30)) and their corresponding depth misfits are overestimated. Depth misfit errors for shallower events are underestimated in this case (Figure (3.30)). While epicenter error estimates are more representative of true epicenter errors for the joint P and S inversion shown in Figure (3.27), the overestimation of depth errors for the deepest events persists (Figure (3.28)). True hypocentral depth errors are not a function of depth in either Figure (3.28) or Figure (3.30).

This discrepancy points out a fundamental weakness of the error appraisal approach developed in Section (2.6.2) of Chapter 2 as applied to hypocenter mislocations. The inconsistency is the result of two causes. Firstly, ray path lengths to deeper events are longer than those for shallow events (most ray paths are upgoing). Consequently, the estimated error of calculated travel times will be correspondly higher.[3.2] This tends to inflate hypocenter error estimates as source depth increases. Secondly, the ray paths for shallow events have much larger path percentages in shallow less well resolved (and estimated) portions of the velocity models than do ray paths from deeper events. Shallow event hypocentral mislocation errors are underestimated (especially for the P-only inversions) because linearized error estimates are incorrect in the shallow portions of the velocity models that are not linearly close to the true models.

The scale length comparisons of sensitivity of hypocenter estimates to unresolved slowness gradient model features overestimated the sensitivity of hypocenter depth to potential unknown errors in the model for all cases.

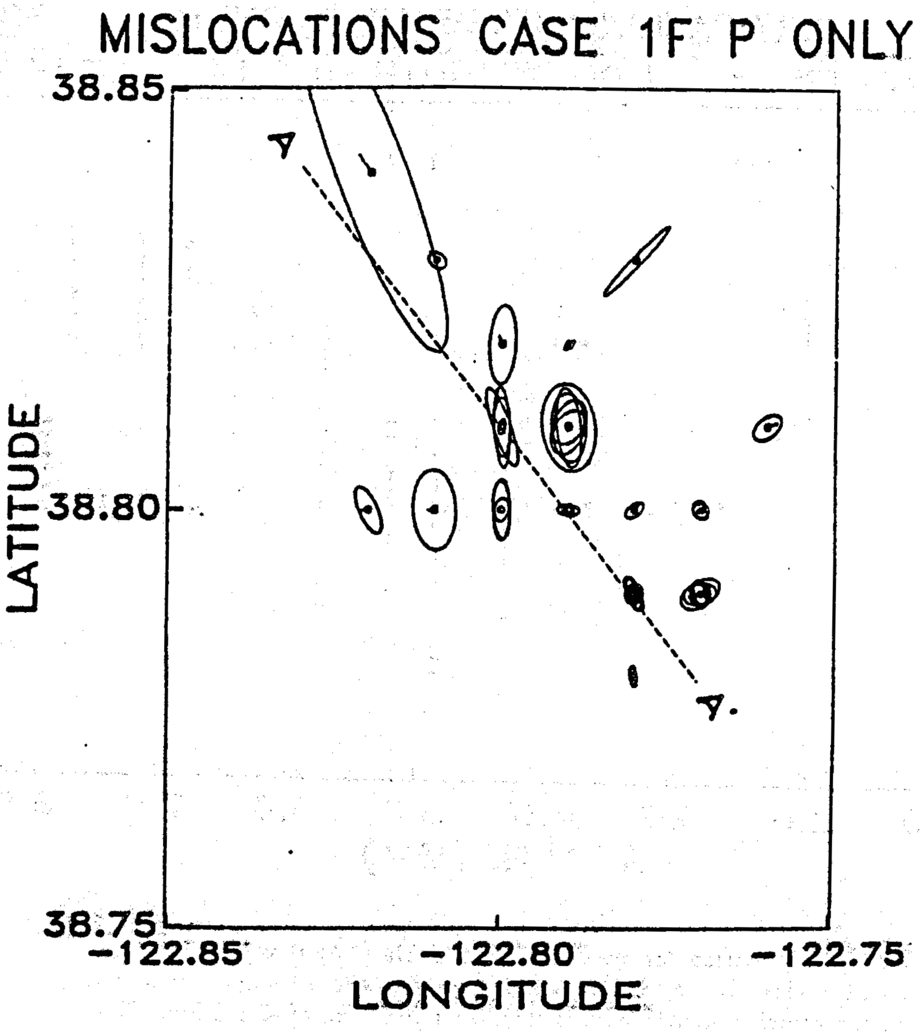


Figure 3.29 Epicenter mislocations for case (1F P) of Table (3.5). Small circles show true epicentral locations and lines point from the circles to estimated epicenter locations. The circles with multiple lines correspond to multiple events at the same true epicenter. Two standard error confidence ellipses calculated using all sources of error are also shown. The dashed line from A to A′ is the surface projection of the accompanying cross section shown in Figure (3.30).

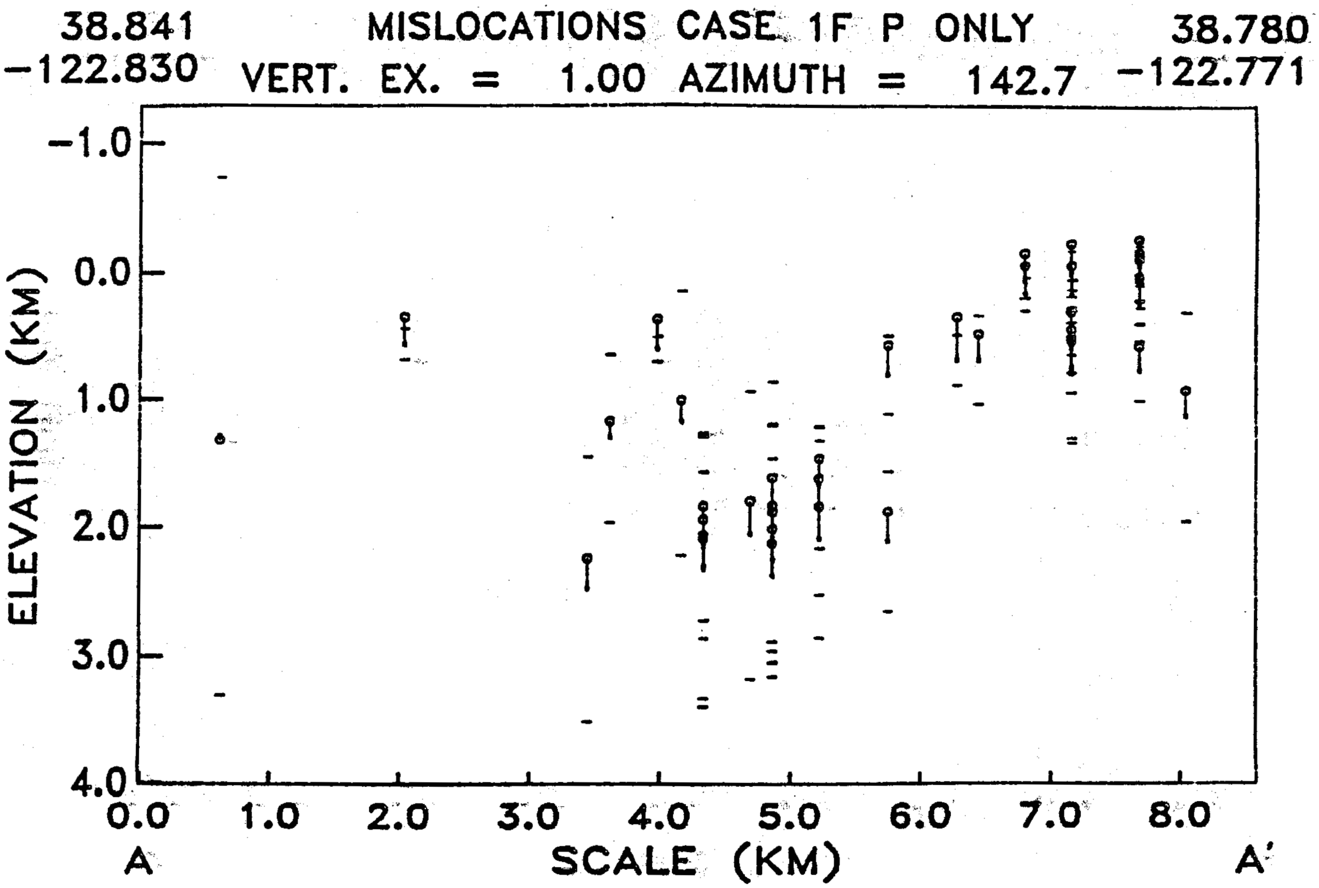


Figure 3.30 Cross section for case (1F P) of Table (3.5) showing true locations as small circles. Lines from the circles point to the estimated locations. The top of the velocity models corresponds to an elevation of -0.379 km. Two standard error limits estimated using all sources of error are plotted as small bars about the estimated location. Note that some true locations lie outside the error limits for the shallow events.

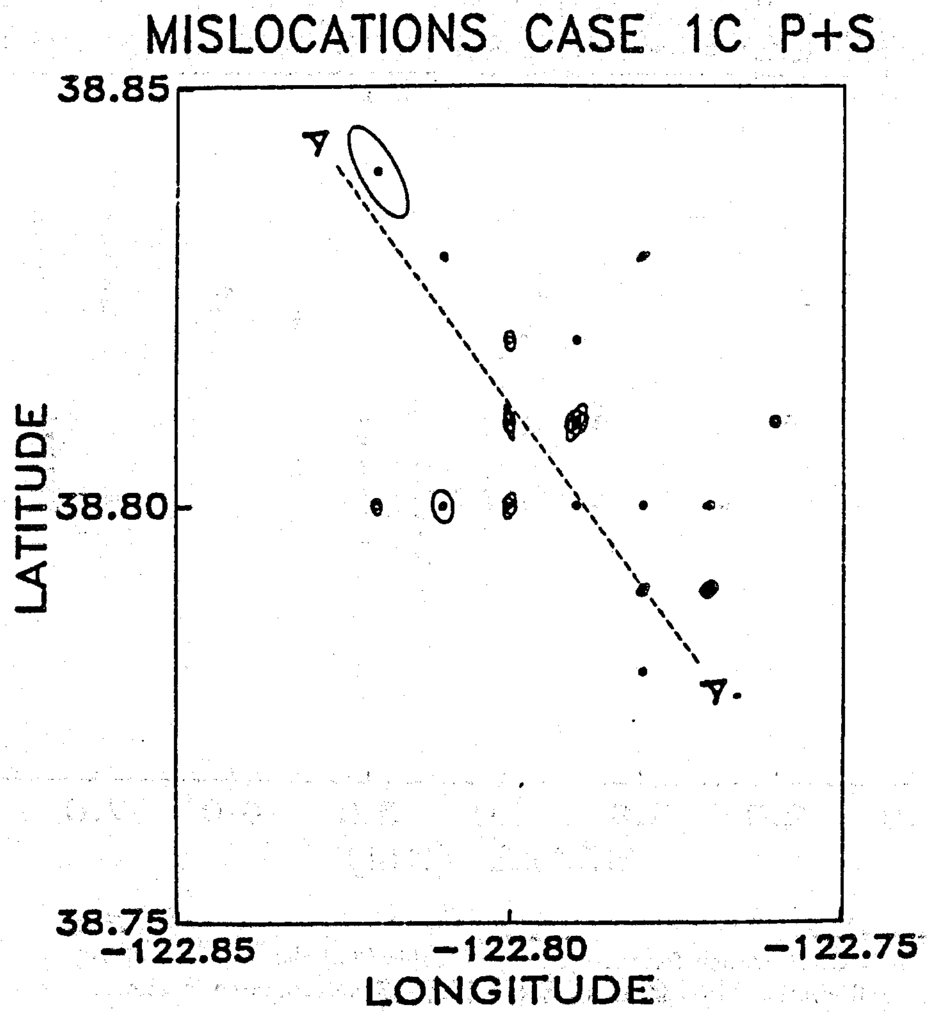


Figure 3.31 Epicenter mislocations for case (1CPS) of Table (3.5). Small circles show true epicentral locations and lines point from the circles to estimated epicenter locations. The circles with multiple lines correspond to multiple events at the same true epicenter. Two standard error confidence ellipses calculated using all sources of error are also shown. The true mislocations are much smaller the the circles. The dashed line from A to A' is the surface projection of the accompanying cross section shown in Figure (3.32).

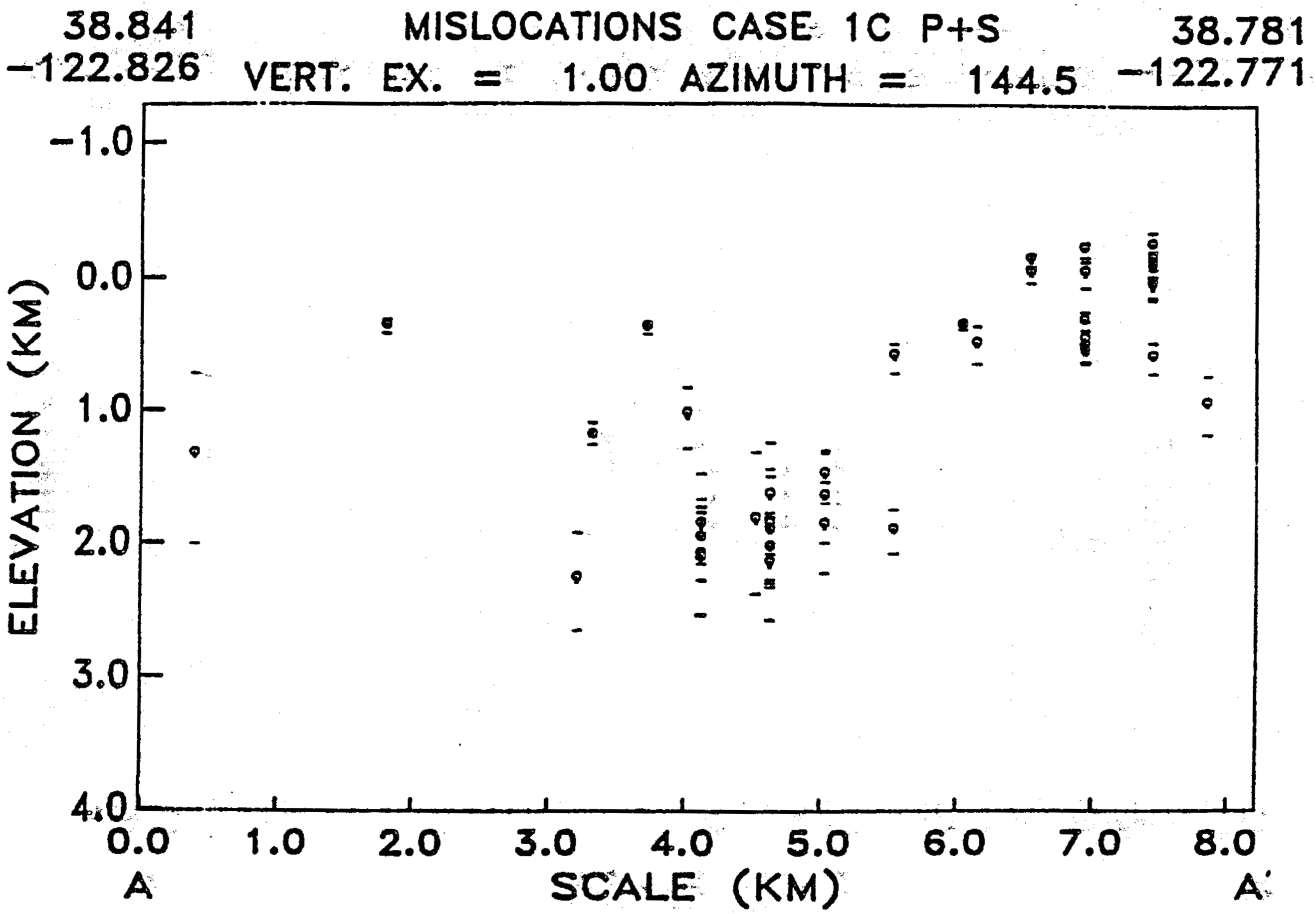


Figure 3.32 Cross section for case (1CPS) of Table (3.5) showing true locations as small circles. Lines from the circles point to the estimated locations. The top of the velocity models corresponds to an elevation of -0.379 km. Two standard error limits estimated using all sources of error are plotted as small bars about the estimated location. Note that the error limits tend to be overly pessimistic.

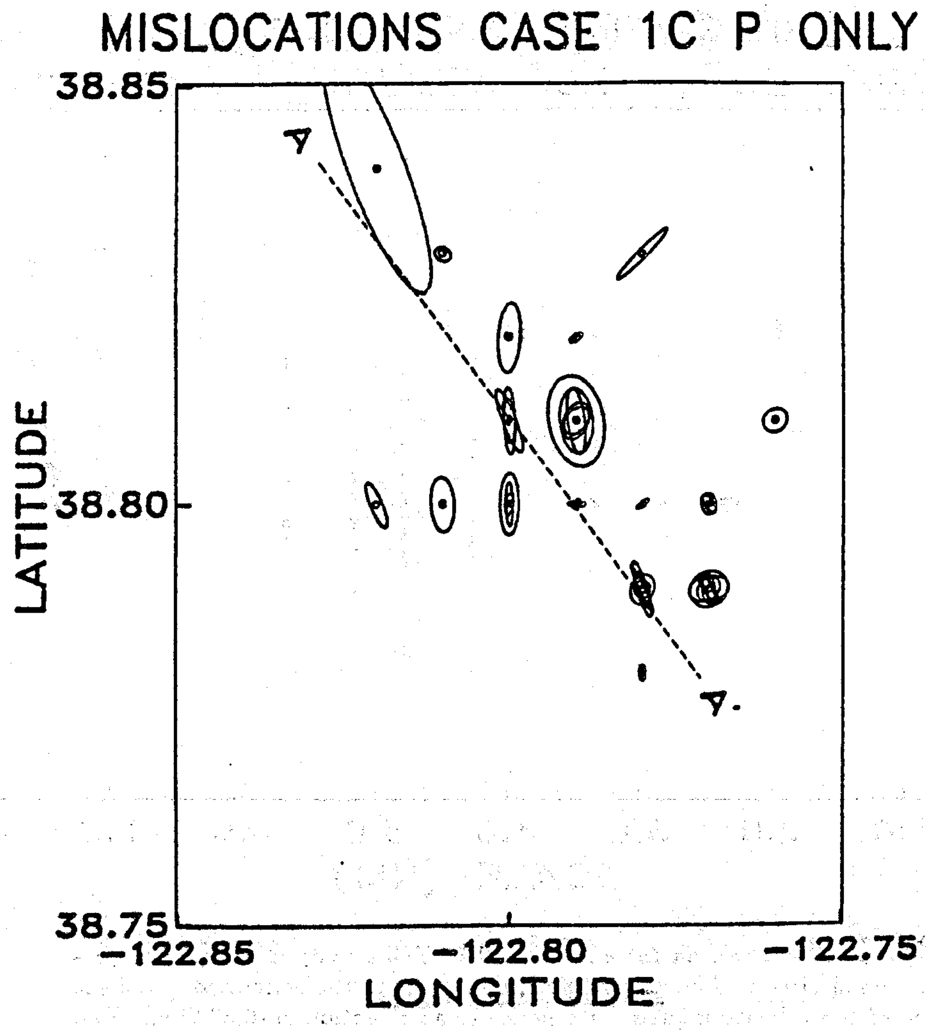


Figure 3.33 Epicenter mislocations for case (1C P) of Table (3.5). Small circles show true epicentral locations and lines point from the circles to estimated epicenter locations. The circles with multiple lines correspond to multiple events at the same true epicenter. Two standard error confidence ellipses calculated using all sources of error are also shown. The true mislocations are much smaller the the circles. The dashed line from A to A′ is the surface projection of the accompanying cross section shown in Figure (3.34).

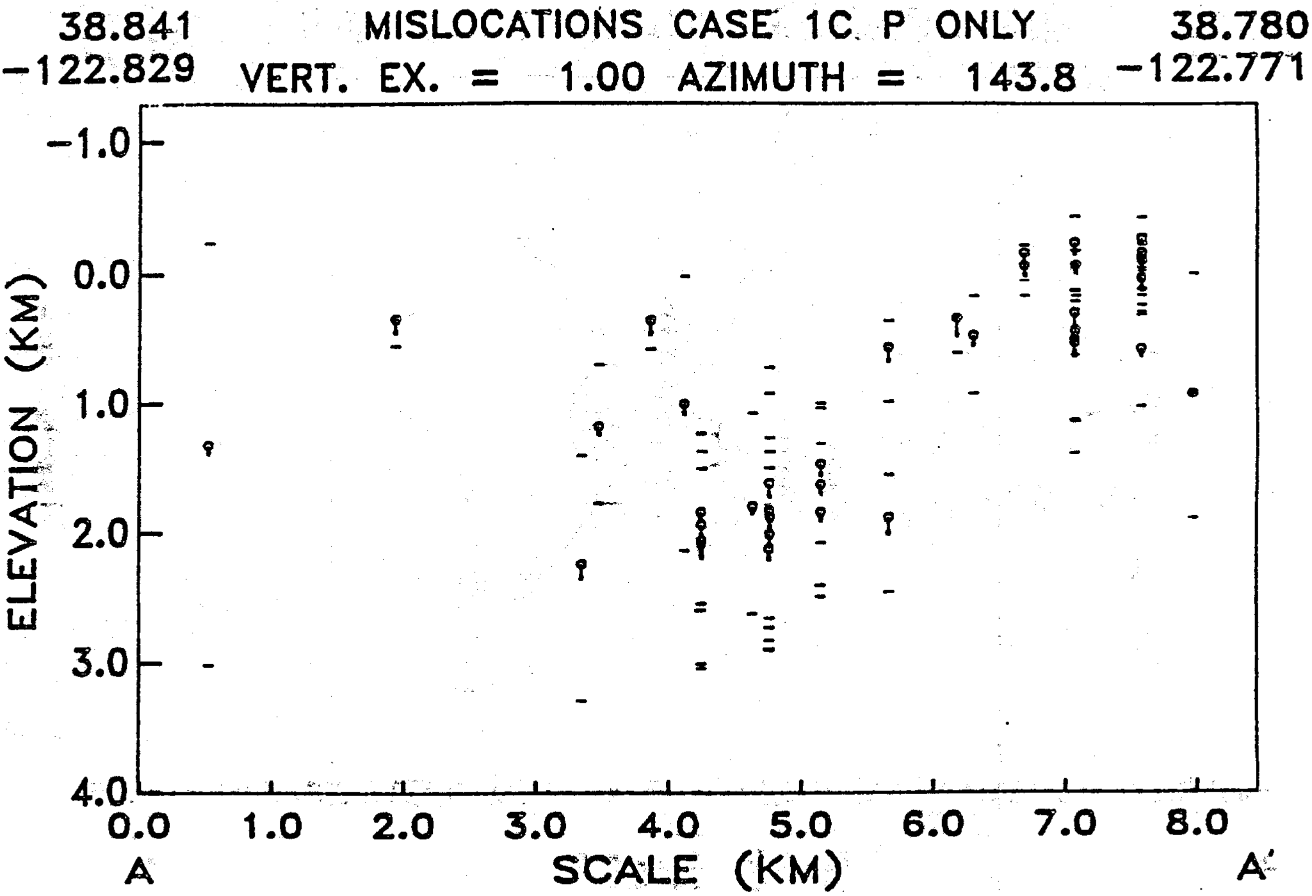


Figure 3.34 Cross section for case (1C P) of Table (3.5) showing true locations as small circles. Lines from the circles point to the estimated locations. The top of the velocity models corresponds to an elevation of -0.379 km. Two standard error limits estimated using all sources of error are plotted as small bars about the estimated location. Note that mislocation errors are systematic and larger than in Figure (3.32).

Even with these problems, the linearized error estimation approach gave overall estimates of true hypocenter mislocation that were much closer to the truth than the conventional statistical error estimation approach. Thus, the approach of Section (2.6.2) may be useful when applied to real data to get a rough idea of absolute error magnitudes, keeping the aforementioned caveats in mind.

3.7. Summary and Conclusions

These synthetic tests have clearly demonstrated the advantages of including S-wave arrival time data in progressive inversions for hypocenters, station corrections, and velocity structure. Joint use of P and S wave arrival time data has the following advantages over the use of P-wave data alone:

(1) P-wave velocity and slowness gradient structure are more accurately estimated;

(2) Hypocenter mislocation errors are substantially reduced, especially hypocentral depth;

(3) Convergence of progressive inversions to local minima is more detectable using RMS data misfits of P and S wave data;

(4) Velocity model and hypocenter estimates are much more accurately determined when station corrections are used;

(5) Complete elastic properties are estimated providing greater constraints for geologic interpretation of velocity structure.

Slowness gradient is better resolved using earthquake data than velocity. Interpretation of progressive inversion for velocity structure should concentrate more on spatial variations of velocity by looking at slowness gradient resolution and error. It was clearly demonstrated that velocity magnitude is not nearly as well determined.

The dangers of using artificially high surface velocities to locate earthquakes and invert for velocity structure are clearly demonstrated. High surface velocities are not harmlessly incorporated into station corrections but in fact produce substantial errors in estimated velocity structure and earthquake locations.

The error prediction approach developed in Section (2.6.2) of Chapter 2 for station corrections and hypocenters provide somewhat pessimistic estimates of true errors. However, these error estimates are much more realistic than conventional error estimates that ignore model induced errors.

Adding S-waves to progressive inversion does not completely eliminated hypocenter-velocity tradeoffs, but they are substantially reduced. Linearized resolution and error analysis will still yield unreliable results near strong discontinuities of slowness gradient. The errors in linearized resolution and error estimates are reduced when P and S data are used together. The only way to ensure that model features are resolvable is to do synthetic tests using the estimated model, as done here. While this may appear to involve too much effort, the information obtained is of fundamental importance for reliably interpreting inversion results with real data.

Notes

[3.1] Slowness is actually the quantity estimated in Chapter 2. Resolving functions are calculated for slowness and used to smooth the true slowness model. Velocity models are then taken by inverting slowness at each discretized depth and assuming linear velocity gradients between depth points.

[3.2] From equation (2.78) we see that longer path lengths for a particular ray parameter mean that $\overline{\mathbf{G}}$ will have more nonzero elements. Therefore, the term, $\overline{\mathbf{G}}\mathbf{C}_k^{s'}\,\overline{\mathbf{G}}^{T}$, the estimate of calculated travel time error, will correspondingly increase.

References

Eberhart-Phillips, D. and D. H. Oppenheimer, 1984. Induced seismicity in The Geysers geothermal area, California, *J. Geophys. Res.*, **89**, 1191-1207.

Eberhart-Phillips, D., 1986. Three-dimensional velocity structure in Northern California Coast Ranges from inversion of local earthquake arrival times, *Bull. Seism. Soc. Am.*, **76**, 1025-1052.

Inglada, V. O., 1928. Calculo de las coordenadas del foco y del instante inicial de un sismo por medio del las horas de las ondas S registradas en las estaciones proximas. *Revista del la Real Acadmia de Ciencias exactas, fisicas y naturales*, de Madrid, **24**, 175-201.

Nicholson, C. and D. W. Simpson, 1985. Changes in V_p/V_s with depth: implications for appropriate velocity models, improved earthquake locations, and material properties of the upper crust, *Bull. Seism. Soc. Am.*, **75**, 1105-1123.

Oldenburg, D. W., 1984. An introduction to linear inverse theory, *IEEE Trans. Geoscience and Remote Sensing*, **GE-22**, 665-674.

Pavlis, G. L. and J. R. Booker, 1980. The mixed discrete-continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res.*, **81**, 4801-4810.

Pavlis, G. L., 1982. Progressive inversion, *Ph.D. Dissertation*, University of Washington, Seattle, Washington.

Pavlis, G. L. and J. R. Booker, 1983a. A study of the importance of nonlinearity in the inversion of earthquake arrival time data for velocity structure, *J. Geophys. Res.*, **88**, 5047-5055.

Pavlis, G. L. and J. R. Booker, 1983b. Progressive multiple event location (PMEL), *Bull. Seism. Soc. Am.*, **73**, 1753-1777.

Riznichenko, Y. V., 1958. Methods for large-scale determination of focus coordinates of nearby earthquakes and velocities of seismic waves in the focal region, *Tr. Inst. Fiz. Zemli. Akad. Nauk SSR*, **4**.

Thurber, C. H., 1983. Earthquake locations and three-dimensional crustal structure in the Coyote Lake Area, Central California, *J. Geophys. Res.*, **88**, 8226-8236.

Wadati, K., 1933. On the travel time of earthquake waves, II, *Geophys. Mag.*, **7**, 101-111.

Chapter 4

Progressive Inversion for Hypocenters and P-wave and S-wave Velocity Structure: Application to The Geysers

4.1. Introduction

The Geysers geothermal field is the site of intense microseismicity. The rate of seismicity at The Geysers is 45 times the regional rate (Ludwin et al., 1982). The Geysers is also the world's largest generator of electricity using geothermal energy. It has been suggested that seismicity in The Geysers is induced by some aspect of steam production but the specific mechanism has not been determined (Oppenheimer, 1986; Eberhart-Phillips and Oppenheimer, 1984; Bufe et al., 1981). Of all the inducing mechanisms proposed to explain the seismicity, perturbation of the regional stress field by volumetric contraction of the reservoir due to a net mass (water) withdrawal (Majer and McEvilly, 1979) and conversion of aseismic slip due to an increase in frictional stength (Allis, 1982) remain the most plausible (Oppenheimer, 1986).

A long standing question is whether The Geysers geothermal field has a distinctive seismic signature. Eberhart-Phillips and Oppenheimer (1984) and Eberhart-Phillips (1986) have done the only formal inversions for velocity structure in The Geysers area. Eberhart-Phillips and Oppenheimer (1984) used arrival times from earthquakes and explosions to derive a P-wave velocity model for a region of which the primary producing portion of The Geysers geothermal reservoir comprised a small part (e.g. <1%) of the total area. Eberhart-Phillips (1986) inverted for three-dimensional P-wave velocity structure on three scales. The smallest scale inversion was at only twice the scale of Figure (4.1), but comprised only three depth samples, allowing limited interpretation of velocity depth variations in the production zone. Eberhart-Phillips (1986) and Majer and McEvilly (1979) found that P-wave velocities in The Geysers production area between the Mercuryville fault to the southwest and the Collayomi fault to the northeast

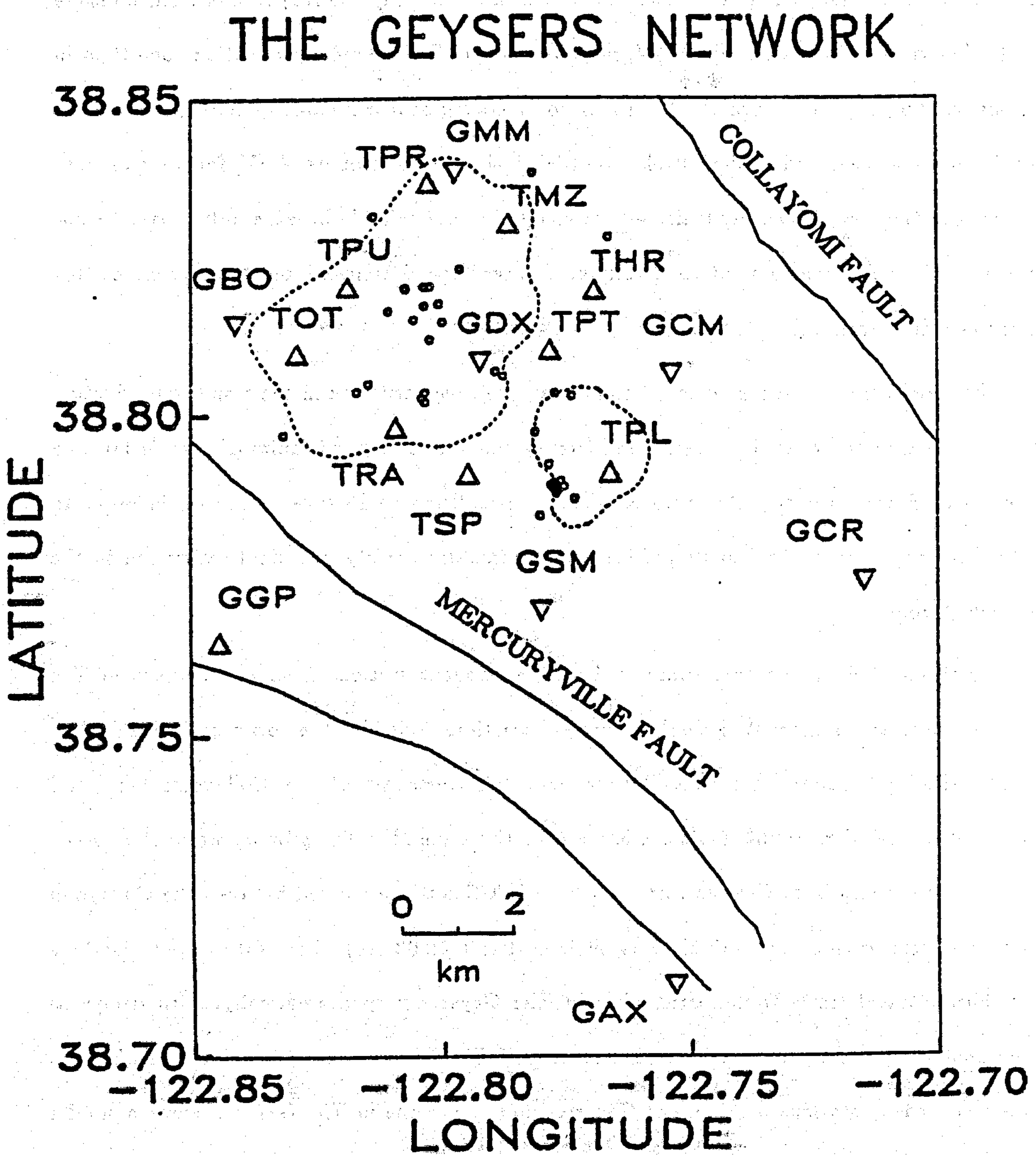


Figure 4.1 Temporary station names begin with a T and USGS station names begin with a G. Three-component stations are plotted as Δ and single vertical component stations are plotted as ▽. The dashed lines are 34.5 bar contours (Lipman et al., 1978) enclosing areas of pressure decline for the year 1977 (not available for other years). These pressure decline contours roughly define the region referred to as the primary production zone throughout Chapter 4. Faults are shown as solid lines. Small circles are earthquake locations estimated from progressive inversion. Note that, with the exception of stations GGP and GAX, all stations lie between the Collayomi and Mercuryville faults, inside or in close proximity to the primary production zone.

were faster than regional P-wave velocities. Majer and McEvilly's (1979) results from multiple event Wadati plots indicated that V_p/V_s may be lower in The Geysers production zone than in the surrounding region. Gupta et al., (1982) obtained regional estimates of half-space P-wave and S-wave velocities and used multiple event Wadati plots to estimate V_p/V_s for a region containing The Geysers. However, their S-wave results are not reliable because only vertical component seismograms where used to determine S-wave arrival times, a point that will be discussed in more detail later.

The goal here is to focus on the P and S-wave velocity structure in the primary production zone of The Geysers located between the Mercuryville and Collayomi faults (Figure (4.1)). An estimate of S-wave velocity structure is needed, in addition to P-wave velocities, in order to make inferences about physical properties such as fracture density and fluid saturation in the production zone.

Analysis of P-wave travel times to USGS permanent stations from earthquakes in The Geysers indicate significant lateral velocity variations outside the primary steam field. Eberhart-Phillips (1986) finds that P-wave velocities northeast of the Collayomi fault and southwest of the Mercuryville fault are lower than those found in the primary production zone. P-wave travel times from Geysers earthquakes to USGS stations located between the Collayomi fault and Mercuryville fault exhibit very little scatter ($\leq$0.03 sec). For stations outside these boundaries, travel times from earthquakes in The Geysers vary dramatically, with delays as large as 0.4 seconds to stations located northeast of The Geysers near Clear Lake. Refraction data recorded by stations outside The Geysers from explosions in The Geysers show a similar travel time pattern (Eberhart-Phillips, 1986). Majer and McEvilly's (1979) analysis of refraction data from explosions outside The Geysers are consistent with these results.

Our goals were to determine P-wave and S-wave velocity structure in the primary production zone of The Geysers and to locate microearthquakes there. The results of prospective approachs to estimate velocity structure at The Geysers were required to meet the following set of requirements: (1) velocity models must be estimated in a form suitable for calculation of synthetic seismograms; (2) S-wave velocity structure must be estimated along with P-wave velocity structure; (3) estimation of microearthquake locations must be carried out as part of the process; (4) seismic sources and recording stations must be located primarily in the producing steam field; and (5) recording stations must be distributed in a way that provides adequate azimuthal and range coverage to constrain earthquake locations and provide velocity information. The methods used in Chapter 5 to calculate synthetic seismograms require laterally homogeneous velocity models. Consequently, seismic velocity models vary with depth only. The assumption of lateral velocity homogeneity is justified because, as discussed above, P-wave travel time data from stations within the steam field, exhibit very little scatter over a wide range of azimuths. Requirement (4) follows from the desire to determine seismic properties of the steam field, but is also forced upon us by the obvious lateral velocity variations outside the field.

No inversion approach existed that satisfied the first three requirements. The results in Majer and McEvilly (1979), Gupta, et al. (1982), Eberhart-Phillips and Oppenheimer (1984), and Eberhart-Phillips (1986) do not satisfy the last two objectives. Eberhart-Phillips and Oppenheimer's (1984) parameterization of P-wave velocities required the use of constant velocity layers. Layer thicknesses were determined to obtain acceptable resolution. Consequently, velocity discontinuities at layer boundaries are somewhat artificial making their results unsuitable for calculating synthetic seismograms. Consequently, in Chapter 2, a progressive hypocenter-velocity inversion method was developed to satisfy the first three requirements by extending an approach developed by Pavlis (1982) exclusively for P-waves, to incorporate S-wave arrival time data to estimate hypocenters, station corrections, and P-wave and S-wave velocity models.

4.2. Data

Although the USGS operates a permanent network of eight seismic stations in close proximity to The Geysers, only one station records three components of ground motion. To provide reliable S-wave arrival time data, a nine station temporary network of three-component digital event recorders was deployed in The Geysers from July 21, 1982 to August 15, 1982. Data were recorded at 200 samples/sec. Requirement (4) placed constraints on station locations; all temporary stations were deployed in the primary production zone. Temporary station locations are shown in Figure (4.1) along with permanent USGS stations. It should be noted that throughout this chapter the convention is adopted that negative elevations correspond to elevations above mean sea level. Station elevations are listed in Table (4.1). Since station elevations variations exceeded 1.0 km, a special travel time calculation procedure, described in Appendix A, was developed to calculate travel times for stations at varying elevations. Actual ray paths were calculated instead of using vertical path approximations. Proper accounting for station elevation differences is important to obtain reliable hypocentral depth estimates, given that station elevation differences are of the same order as source depths. The temporary network had an aperture of approximately six kilometers. This limited the resolution of S-velocity inversions at deeper depths, due to restricted wave depth penetration at small offsets. Larger offsets provided by USGS permanent stations helped provide somewhat better resolution of deeper P-wave velocity structure. However, more distant stations are generally outside the primary production volume (Figure (4.1)), so portions of wave paths to these stations were outside the zone of primary steam production.

Arrival time data were obtained from the USGS stations in one of two ways. Due to a change of dubbing policy during the recording period by the USGS, waveform data were available only for earthquakes with coda magnitudes $M_c \geq 1.5$. P-wave arrival times for smaller earthquakes were provided by the USGS P-picker. Impulsive P-wave arrival could be accurately read to 0.01-0.02 sec on inkjet playbacks of USGS waveforms. Uncertainties of impulsive P-picker readings are 0.01-0.04 sec (Oppenheimer, personal communication, 1983). Since P-picker

Station Corrections							
		P-wave			S-wave		
Station	Elevation	Estimate	Error 1	Error 2	Estimate	Error 1	Error 2
name	(km)	(sec)	(sec)	(sec)	(sec)	(sec)	(sec)
GAX	-0.379	-0.045	0.041	0.019	—	—	—
GBO	-0.879	0.173	0.028	0.011	—	—	—
GCM	-1.286	-0.046	0.033	0.014	—	—	—
GCR	-0.731	-0.142	0.042	0.019	—	—	—
GDX	-0.931	-0.005	0.015	0.007	—	—	—
GGP	-1.054	0.031	0.031	0.015	0.084	0.0456	0.023
GMM	-0.963	0.063	0.022	0.010	—	—	—
GSM	-1.017	-0.110	0.043	0.013	—	—	—
TRA	-0.509	0.051	0.031	0.011	0.000	0.045	0.018
TSP	-0.549	-0.001	0.035	0.012	-0.085	0.042	0.017
TOT	-0.549	0.106	0.020	0.017	0.169	0.036	0.027
TPU	-0.948	0.078	0.027	0.011	0.111	0.035	0.020
TPS	-0.997	-0.054	0.087	0.067	0.117	0.109	0.087
TPT	-1.020	-0.061	0.032	0.011	-0.186	0.069	0.015
TPL	-0.908	-0.096	0.037	0.013	-0.243	0.045	0.017
THR	-1.032	-0.038	0.039	0.013	-0.087	0.027	0.019
TMZ	-0.988	0.046	0.038	0.013	0.050	0.050	0.018
TPR	-0.927	0.048	0.039	0.013	0.082	0.057	0.018

Table 4.1 Station names that start with a G are USGS permanent station and those that start with a T are temporary stations. The convection is adopted that negative elevations correspond to elevations above sea level and positive elevations to distance below sea level. Standard error estimate 1 includes an estimate of velocity model induced errors by adding those error estimates to standard error estimate 2, the conventional statistical errors calculated using equation (2.50).

readings for emergent arrivals are highly unreliable, only the two highest quality P-picks were used and were assigned uncertainties of 0.03 and 0.04 sec, respectively.

Noise levels at the temporary stations were very low allowing a picking precision of 0.005-0.01 sec for P-waves and 0.01-0.03 sec for S-waves. Accuracy of arrival time readings from the temporary digital stations were primarily limited by clock corrections. Clock drift uncertainties produce a lower bound on temporary station arrival time uncertainties of 0.01-0.03 sec. A conservative estimate of the combined clock drift uncertainties and picking uncertainties for the most impulsive arrivals at temporary stations is 0.04 sec. Consequently, the best USGS P-arrival times were assigned uncertainties of 0.02 sec and the best temporary station P-arrival times assigned 0.04 sec uncertainties.

S-wave arrival times were read exclusively from horizontal component seismograms. S-wave arrival times estimated from vertical component seismograms at The Geysers are not reliable. Figure (4.2) shows the large errors that can result if first-S-arrival times are picked from vertical component seismograms. After analyzing over 500 three-component seismograms from The Geysers area, it was clear that S-wave arrival times estimated from vertical component seismograms are not reliable. Similar results have been observed in the Mississippi embayment near New Madrid (Andrews et al., 1985).

Of all events recorded, 39 earthquakes were recorded that had at least 10 P and S-arrival time readings and located in the recording network (Figure (4.1) and Table (4.2)). These earthquakes were used as input for progressive inversion. Assigned pick uncertainties were used to weight arrival time readings. The best quality picks had uncertainties of 0.02 sec. Therefore, 0.02 sec is used as the data standard error for χ^2 tests of goodness of fit for progressive inversion.

A total of 469 P-wave arrival times and 294 S-wave arrival times were used. Earthquake coda magnitudes ranged from 1.0 to 3.0 for the data set. The relatively small number of earthquakes recorded over the three week period can be attributed to the low gains used at temporary stations, and telemetry problems with the USGS network during the recording period.

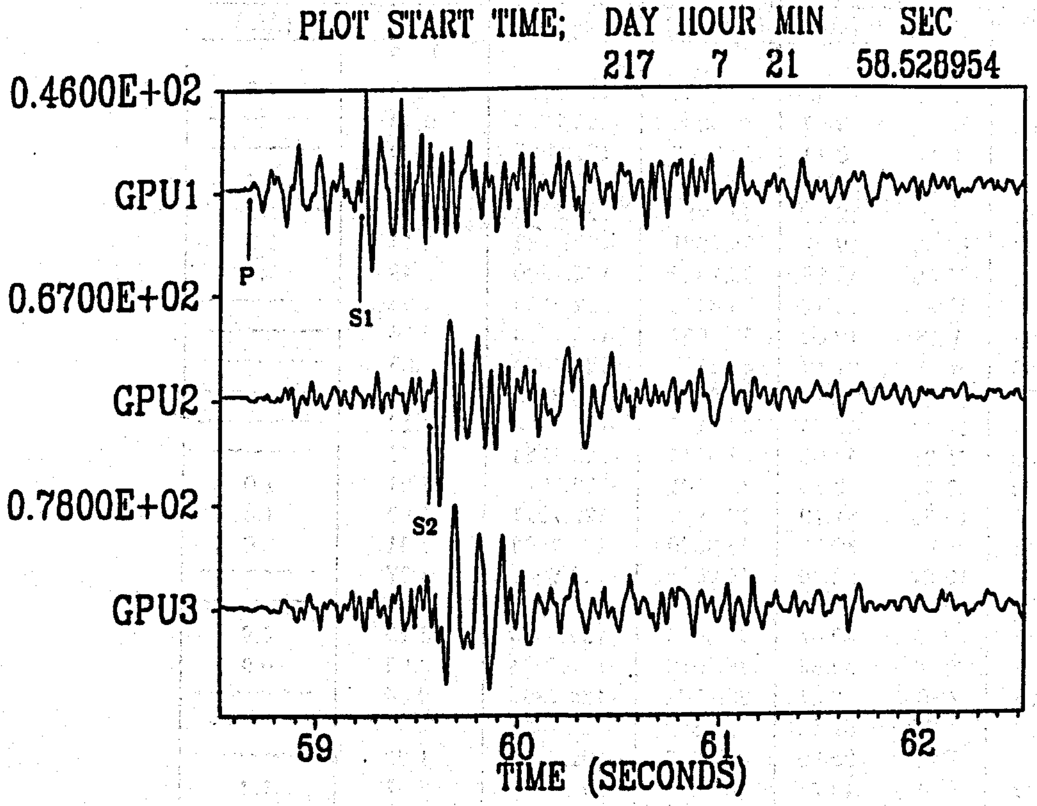


Figure 4.2 A three-component seismograms recorded at station TPU (Figure (4.1)) from an earthquake in The Geysers that occurred on 8-5-82 (Table (4.2)). The vertical component is labeled GPU1, the east-west oriented horizontal component is labeled GPU2, and the north-south oriented horizontal component is labeled GPU3. The true S-P time (S2-P), obtained using the horizontal components to read S-wave arrival time, is 0.90 sec. If only the vertical component were available, the estimate of S-P time (S1-P) would be 0.57 sec. This would results in a total S-P time error of 0.33 sec which represents 37% of the true S-P time.

Earthquake Locations

Date	Origin	Time	Latitude	Longitude	Elevation	Magnitude
(MDY)	(HM)	(sec)	(degrees)	(degrees)	(km)	(M_c)
7-21-82	9 11	48.49	38.8533	-122.8368	0.80	1.6
7-22-82	22 2	52.96	38.8030	-122.7770	0.62	—
7-23-82	1 31	0.13	38.8202	-122.8035	3.40	—
7-23-82	19 30	31.49	38.8034	-122.8171	3.05	1.2
7-26-82	5 6	12.59	38.7969	-122.8316	1.22	1.0
7-26-82	12 59	19.57	38.7921	-122.7782	1.14	1.3
7-26-82	22 50	45.58	38.7842	-122.7800	2.49	1.1
7-28-82	23 33	16.42	38.8315	-122.8138	1.06	—
7-29-82	1 16	12.00	38.8057	-122.7876	0.76	—
7-29-82	6 41	21.16	38.8116	-122.8024	3.23	—
7-30-82	11 40	15.73	38.8163	-122.8108	3.35	—
7-31-82	11 36	37.85	38.8199	-122.8073	3.16	—
7-31-82	21 21	56.06	38.8231	-122.7964	1.47	—
7-31-82	22 41	52.58	38.8201	-122.8074	3.21	1.0
8- 1-82	16 38	51.19	38.7970	-122.7811	1.06	1.6
8- 2-82	8 14	39.69	38.8280	-122.7663	3.31	2.0
8- 3-82	16 27	52.46	38.8032	-122.8034	3.37	—
8- 3-82	19 8	44.45	38.8203	-122.8024	3.35	—
8- 3-82	22 6	30.61	38.8024	-122.8040	3.19	0.9
8- 3-82	22 6	39.63	38.8018	-122.8034	3.06	0.9
8- 4-82	7 6	40.01	38.8064	-122.7892	0.80	—
8- 5-82	7 21	57.28	38.8382	-122.7815	2.45	—
8- 5-82	7 27	28.80	38.8026	-122.7737	1.29	—
8- 6-82	9 4	14.77	38.8046	-122.8148	2.57	1.3
8- 6-82	9 40	38.09	38.8148	-122.8056	2.99	2.2
8- 6-82	18 51	40.34	38.7887	-122.7752	0.88	1.5
8- 6-82	18 55	4.65	38.7896	-122.7759	0.92	—
8- 6-82	18 57	28.42	38.7889	-122.7779	0.72	—
8- 6-82	18 59	51.00	38.7885	-122.7774	0.93	2.1
8- 6-82	19 36	27.90	38.7881	-122.7763	1.06	1.4
8- 6-82	19 37	18.39	38.7876	-122.7769	1.05	1.2
8- 6-82	19 37	20.55	38.7883	-122.7766	1.00	—
8- 6-82	19 37	51.76	38.7888	-122.7775	1.26	1.8
8- 6-82	19 48	20.49	38.7893	-122.7772	1.08	—
8- 6-82	23 20	55.35	38.7889	-122.7767	0.94	—
8- 7-82	1 49	13.54	38.7868	-122.7732	0.99	1.7
8- 8-82	9 8	54.15	38.8174	-122.8005	3.18	2.3
8- 8-82	11 22	31.55	38.8171	-122.8035	2.85	1.0
8-11-82	2 21	33.60	38.8144	-122.7998	3.50	1.8

Table 4.2 Earthquake locations estimated from progressive inversion. Positive elevation denotes distance below mean sea level. To convert to depth in the velocity and slowness gradient models add 0.379 km. Latitude and longitude are in degrees north and west, respectively. Coda wave magnitudes, M_c, are USGS estimates for these events.

Relatively low gains were required to prevent clipping of waveforms for $M_c \approx 1.5$–2.5 earthquakes.

4.3. Application of Progressive Inversion

The method of progressive inversion developed in Chapter 2 and tested in Chapter 3 with synthetic data, is used to estimate earthquake locations, station corrections, and P and S-wave velocity structure. Assuming that arrival time data do not contain large undetected errors, the most important aspect affecting the success of progressive inversion is the use of good starting models.

4.3.1. Starting Models

The importance of using realistic starting models for velocity-hypocenter inversions was clearly demonstrated in Chapter 3. Good estimates of near surface velocity structure are particularly important to achieve accurate results. Fortunately, VSP data from a neighboring portion of The Geysers geothermal field were available to estimate shallow (0-1.5 km depth) P and S-wave velocity structure.

Velocities for P and S-wave from the free surface to a depth of 1.5 km were obtained from a one-dimensional trial and error inversion of the multi-offset P and S-wave VSP data of Majer et al., (1987). Models for P and S velocities in the VSP inversion were developed by interactively perturbing velocity models until they fit the VSP first-arrival travel time data. P-wave arrival times had uncertainties of 0.001 sec and S-wave arrival times had uncertainties of 0.004 sec. The small pick uncertainties and small (30.5m) receiver spacing in the well provided good constraints on P and S velocities. The models were constructed by including the smallest number of linear velocity gradient layers required to fit the data. The resulting models are the simplest piecewise-continuous velocity models that reproduce the observed travel times within their standard error.

Majer et al, (1987) observed a 11% velocity variation for shallow (0.3-0.06 km depth) wave paths between *SH* and *SV* waves generated by rotating the vibrator orientation to two orthogonal polarizations for each survey level in the well. Shear wave splitting was not observed at the temporary recording stations used in The Geysers. The absence of shear wave splitting could imply that shear wave velocity anisotropy is not significant in the primary production zone. However, the surface temporary station locations preclude reliable detection of shear wave splitting, even if anisotropy were present (Crampin, 1985). Consequently, the VSP S-wave travel times used to estimate S-wave velocities were formed from the average of *SH* and *SV* travel times for the shallow wave paths. Only one polarization of S-waves were available for deeper wave paths and those travel times were used directly.

The resulting estimates velocities are shown in Figures (4.3) and (4.4). The highly variable velocities in the top 100 meters were excluded from these models. P-wave velocities in the near surface layer ranged from less than 1.0 km/sec to more than 4.0 km/sec. Ridley and Vantine (1978) observed surface P-wave velocities of 0.7-2.6 km/sec in landslide and hydrothermally altered terrain at power plants 1 and 2 (Figure (4.5)). Denlinger and Kovach (1981) obtain average P-wave velocities of 2.38-2.59 km/sec for the top 200 meters in the area near station GCR. Given the wide range of near surface velocities, the surface values shown in Figures (4.3) and (4.4) were chosen as appropriate averages for the region. Local variations from these values can be accounted for with station corrections. We wanted to be sure to use a representative average of near surface velocities to avoid problems associated with the use of unrealistic near-surface velocities in progressive inversion (see Chapter 3).

The VSP data provided important information on near surface velocities. P-wave velocities below 1.5 km depth were taken from Eberhart-Phillips and Oppenheimer (1984) and corresponding S-wave velocities calculated assuming a constant V_p/V_s of 1.65. The starting P and S-wave velocities thus obtained are shown in Figures (4.3) and (4.4), respectively. The VSP velocity models were obtained for an area between the Mercuryville and Collayomi faults but not in the primary production zone defined by the pressure decline contours shown in Figure

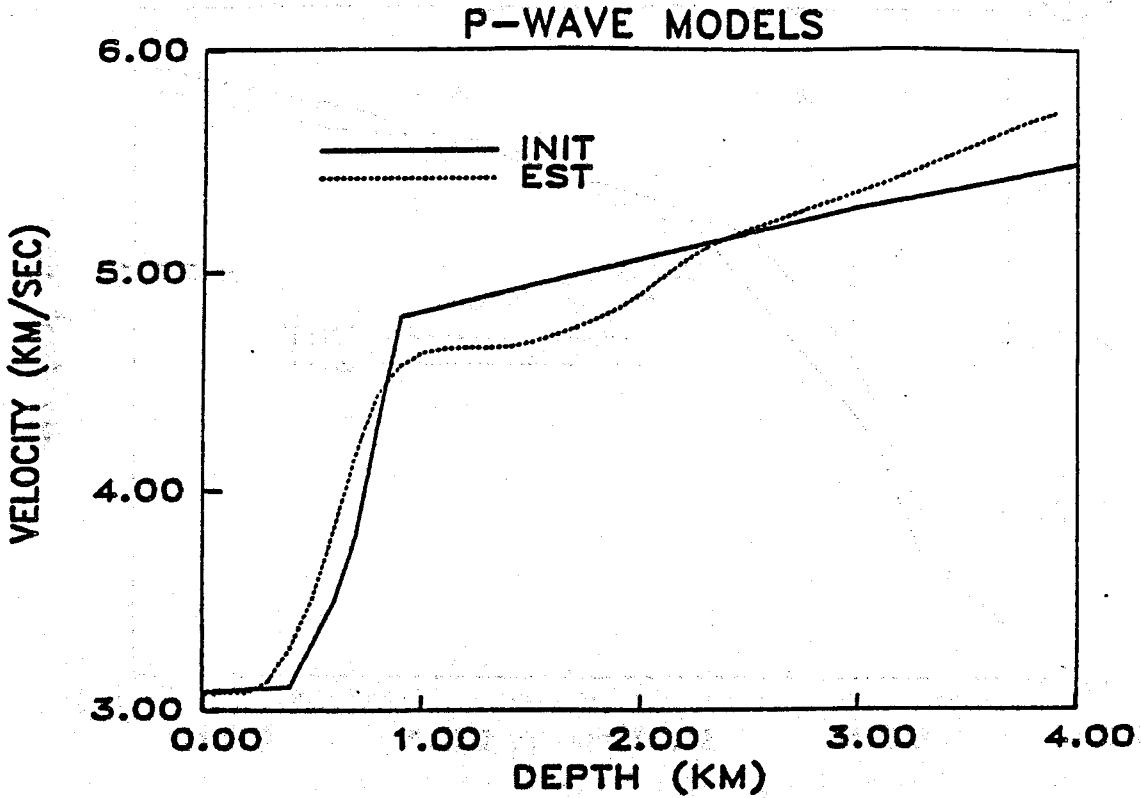


Figure 4.3 The P-wave starting velocity is denoted by INIT and the model estimated from progressive inversion is denoted by EST. The portion of the starting model between 0.0 and 1.5 km depth was derived from VSP data. To convert from model depth to elevation, subtract 0.379 km from the depth values. A depth node spacing of 0.1 km was used for the progressive inversion.

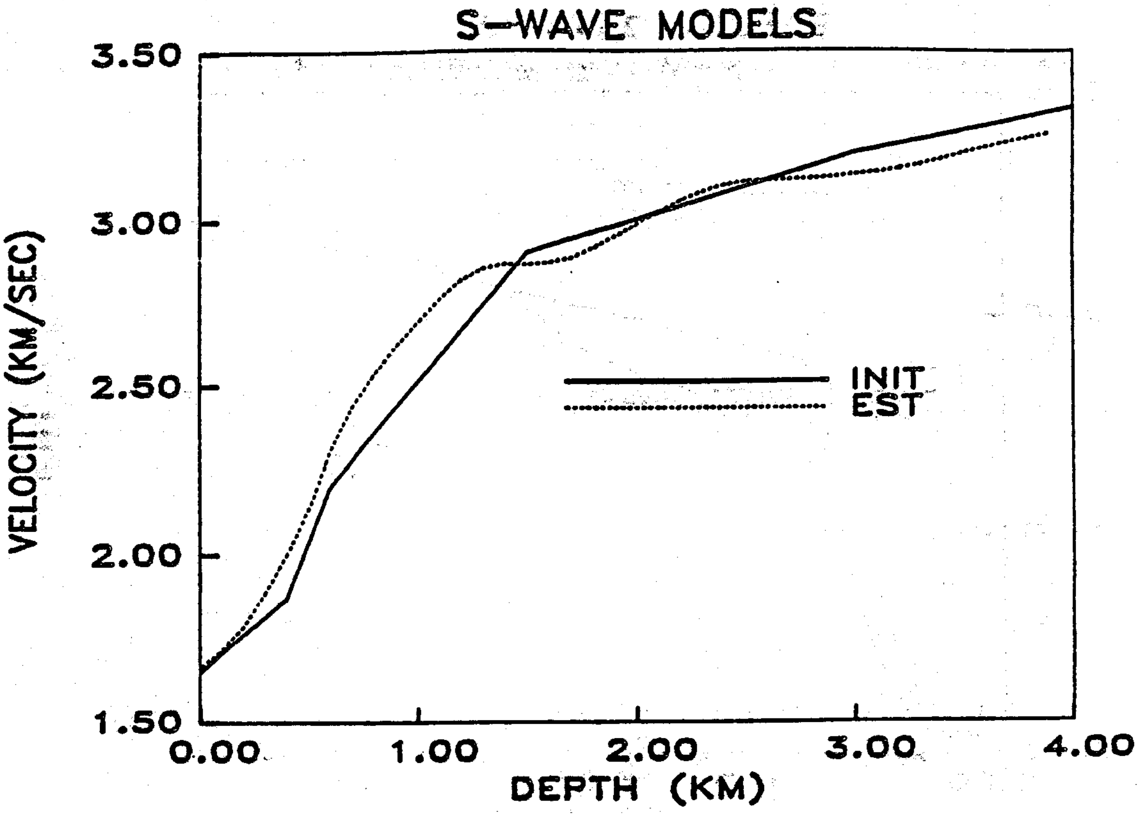


Figure 4.4 The S-wave starting velocity is denoted by INIT and the model estimated from progressive inversion is denoted by EST. The portion of the starting model between 0.0 and 1.5 km depth was derived from VSP data. To convert from model depth to elevation, subtract 0.379 km from the depth values. A depth node spacing of 0.1 km was used for the progressive inversion.

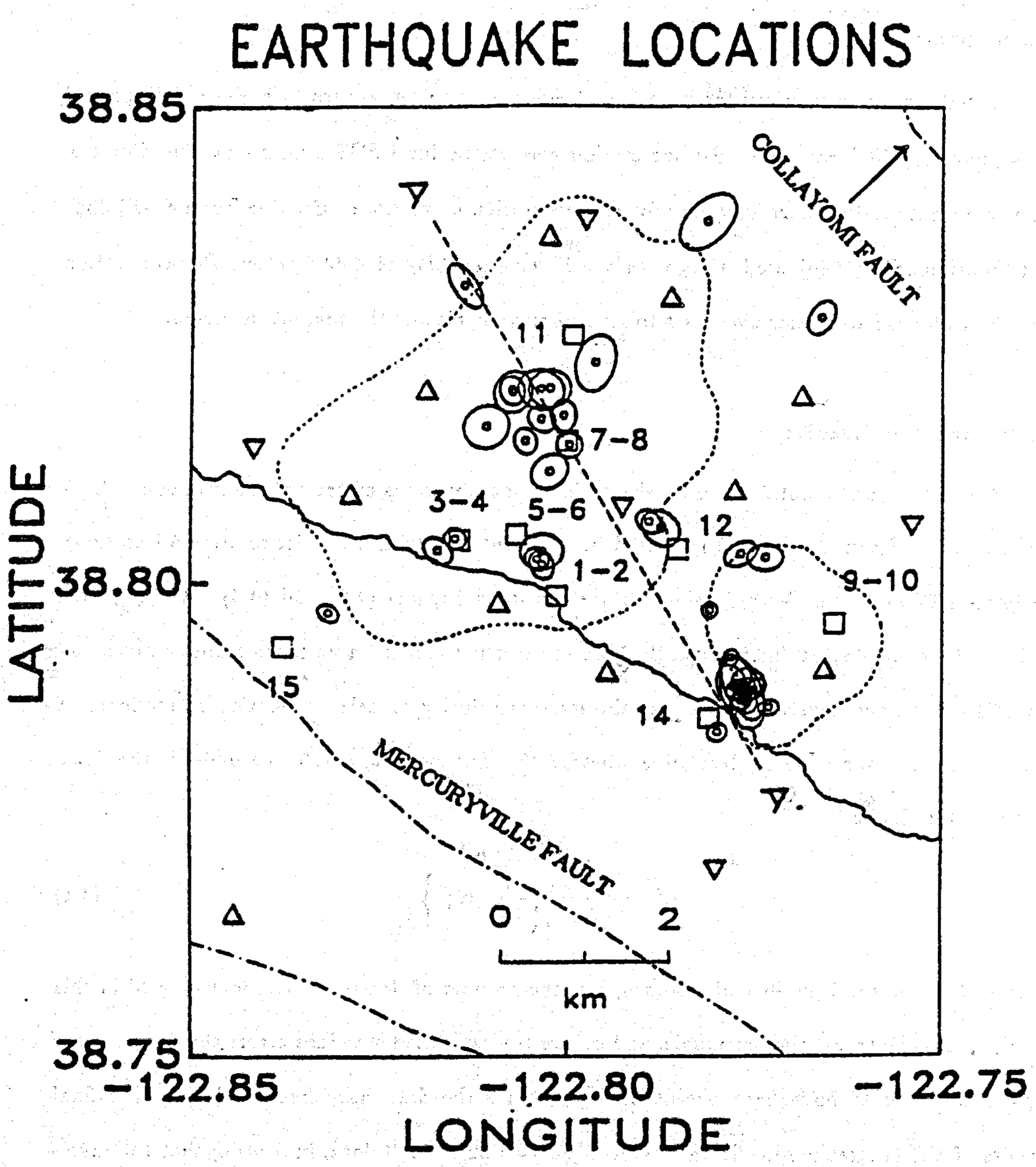


Figure 4.5 Earthquake locations are shown as small circles inside one-standard-error ellipses. The numbered squares are power plants that were operating at the time of the experiment. Three-component stations are shown as Δ and vertical component stations as ∇. The solid trace is Big Sulfur Creek. Fine dashed lines are the 1977 pressure decline contours (Lipman et al., 1978) described in Figure (4.1). Dot-dash lines are faults. The medium-dashed line labeled A and A′ is the surface projection of the cross section in Figure (4.10).

(4.1). Consequently, the starting models were expected to be close to the true structure in the primary production zone.

Initial station corrections for both P and S-waves were set to zero. Eberhart-Phillips and Oppenheimer (1984) computed P-wave station corrections for USGS stations at The Geysers. They were not used as starting values for P-wave station corrections. Because Eberhart-Phillips and Oppenheimer (1984) used a high surface P-wave velocity of 4.43 km/sec, P-wave station corrections for USGS stations were set to zero to prevent biasing the present inversion.

4.3.2. Inversion Results

Convergence was obtained in 5 velocity inversion iterations of progressive inversion. An F test at the 95% confidence level indicated insignificant reduction of variance after 5 iterations. Estimated P-wave and S-wave velocities are shown in Figures (4.3) and (4.4). Although the estimated models do not differ markedly from the starting models, a variance reduction of a factor of 12.8 was obtained with respect to the starting velocity models and station corrections. A reduced χ^2 test was done to determine whether the inversion fit, overfit, or underfit the data. Define reduced χ^2 as

$$\chi_\nu^2 \equiv \frac{1}{M} \sum_{i=1}^{M} \left\{ \frac{1}{\sigma_i^2} [r_i]^2 \right\} \tag{4.1}$$

where M is the total number of readings, ν is the number of degrees of freedom (ν= M in this case), r_i are the travel time residuals, and σ_i are the estimated standard errors the data. When $\chi_\nu^2 << 1.0$ the data have been overfit or if $\chi_\nu^2 >> 1.0$ the data have been underfit. The final results of the progressive inversion produced $\chi_\nu^2 = 1.02$ for 663 data, indicating that the inversion neither overfit or underfit the data. This demonstrates that a one-dimensional model with station corrections satisfactorily fits the arrival time data within the estimated standard error of 0.02 sec.

Eberhart-Phillips (1986) finds that a one-dimensional P-wave velocity model with station corrections produces earthquake locations in The Geysers steam field comparable to those found using a three-dimensional P-wave model for The Geysers-Clear Lake region. She concludes that a laterally homogeneous P-wave model with station corrections can be used to obtain accurate earthquake locations in The Geysers. The results here collaborate her findings for P-waves and also show that a one-dimensional S-wave model with station corrections is adequate for The Geysers primary production zone as well.

The fact that the earthquake locations, station corrections, and velocity models fit the data does not alone ensure that the estimated earthquake locations, station corrections, and velocity models are accurate. To determine what features of the velocity structures are meaningful, we turn to a discussion of resolution and error. Since velocities models very similar to those estimate by progressive inversion, namely the starting models, were used in synthetic tests of progressive inversion, results of those synthetic tests help delineate what features are truly resolvable. Results of the synthetic tests in Chapter 3 are also used to help assess uncertainties in earthquake locations. We begin by discussing velocity and slowness gradient estimates.

4.3.3. Appraisal of Velocity and Slowness Gradient Results

In Chapter 3 it was demonstrated that slowness gradient (the quantity directly solved for in progressive inversion) is better constrained by earthquake data than velocities. Linearized resolution and error analysis was found to be accurate for appraising slowness gradient structure with the exceptions that overall slowness gradient depth profiles may be shifted in depth and sharp discontinuities are smoothed and smeared. The amount of depth offset of slowness gradient structure was found to be quite comparable for both P and S-wave models, allowing meaningful comparisons of relative changes between models as a function of depth.

The estimated P-wave slowness gradient model with 95% error estimates is shown in Figure (4.6) and the estimated S-wave slowness gradient model is shown in Figure (4.7). To help determine what features are significant, resolving kernels (calculated from equation (2.59) using the same expansion orders used to estimate the errors in Figures (4.6) and (4.7)) are plotted in Figures (4.8) and (4.9). The resolving kernels in Figure (4.8) show what portions of the slowness

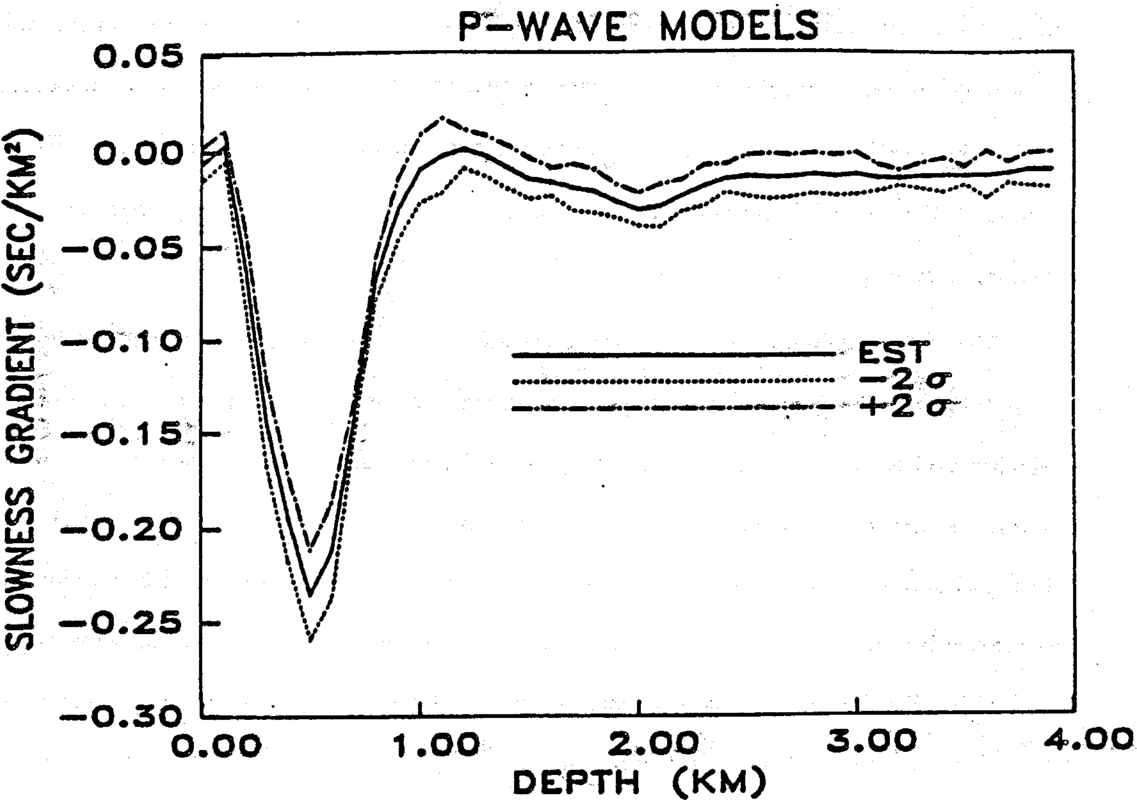


Figure 4.6 The estimated P-wave slowness gradient model, denoted as EST, from progressive inversion is shown with positive and negative 95% confidence limits, denoted as +2 σ and -2 σ, respectively. To convert from model depth to elevation, subtract 0.379 km.

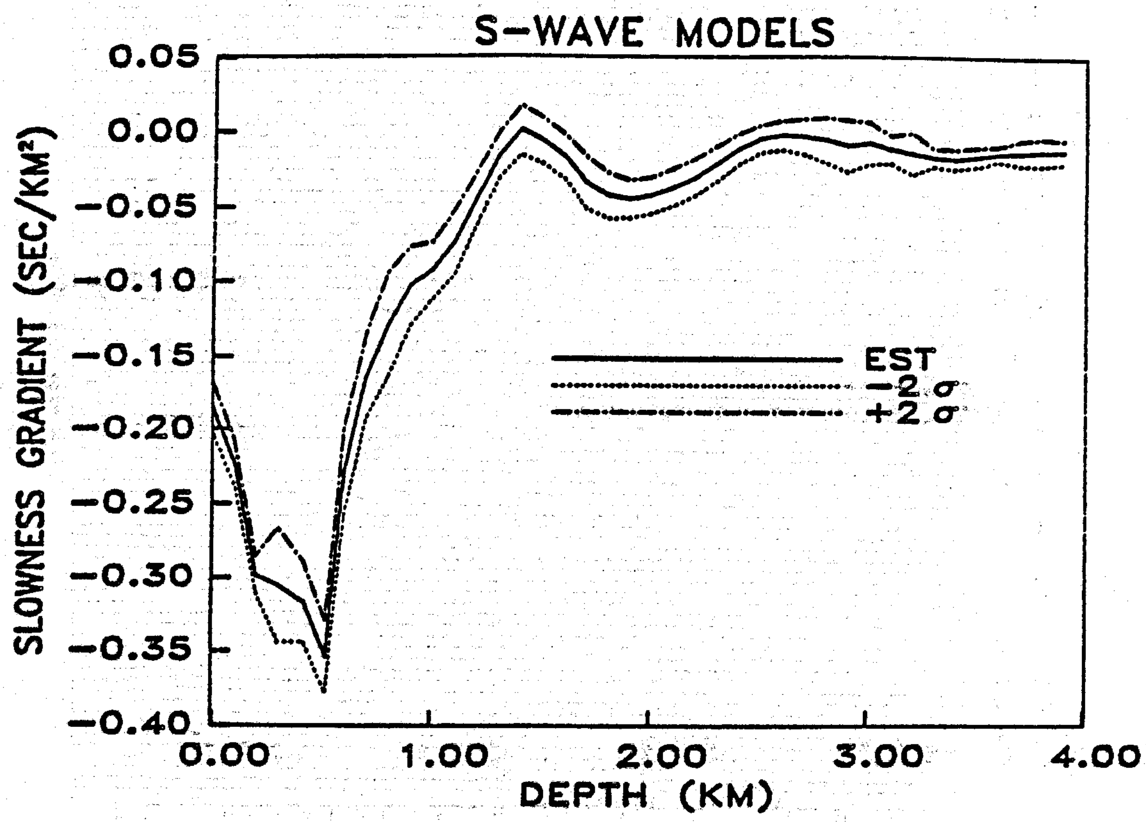


Figure 4.7 The estimated S-wave slowness gradient model, denoted as EST, from progressive inversion is shown with positive and negative 95% confidence limits, denoted as $+2\sigma$ and -2σ, respectively. To convert from model depth to elevation, subtract 0.379 km.

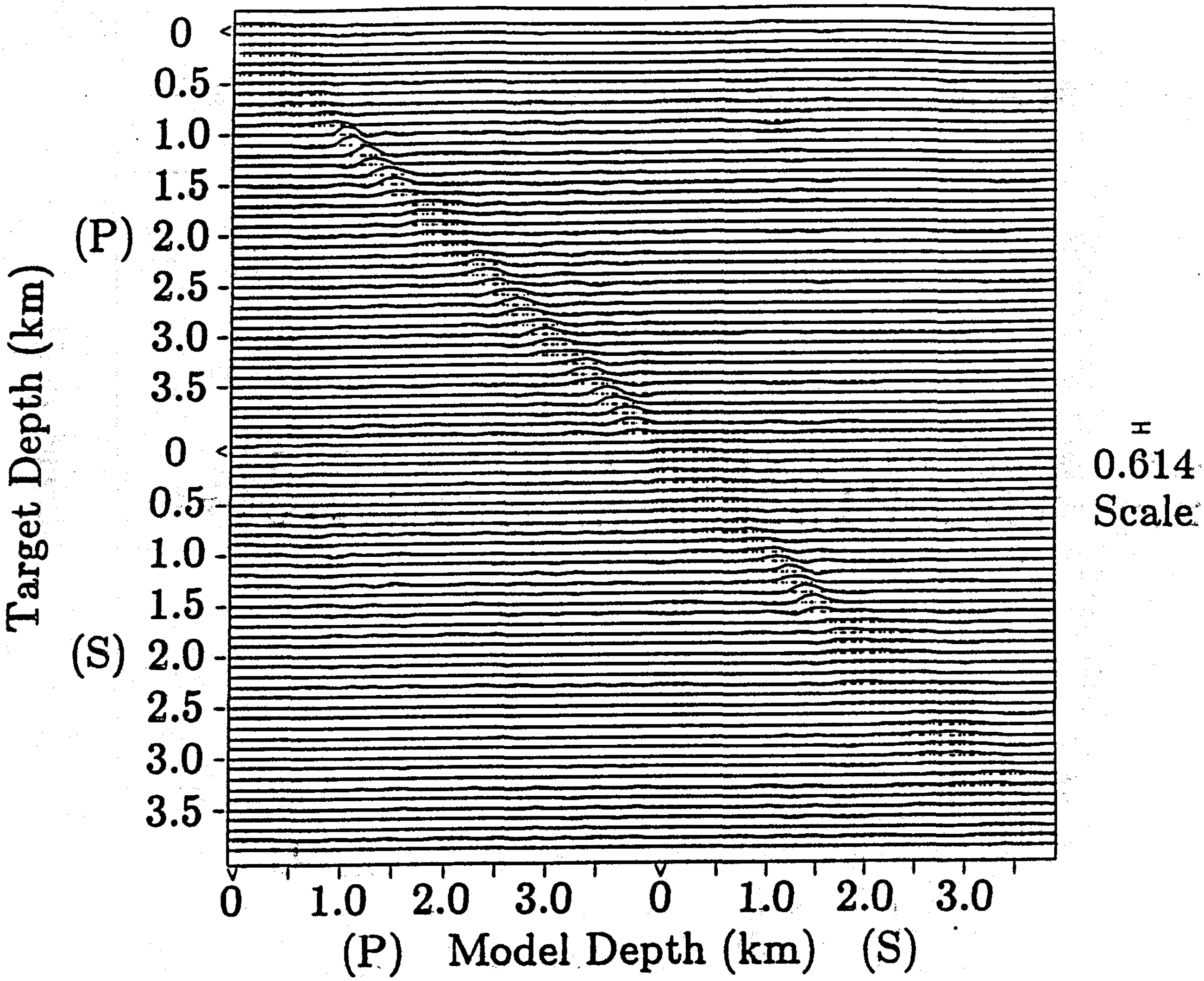


Figure 4.8 Resolving kernels for estimated P and S-wave slowness gradient models. All kernels are normalized to unit area. Each resolving kernel contains both models because P and S models are coupled in progressive inversion. The labels (P) and (S) plotted adjacent to the depth scales, denote which model type corresponds to the model and target depths. There are two resolving kernels at 0.0 km depth for each model; the first kernel corresponds to surface slowness, and the second to surface slowness gradient. The largests amplitude peaks correspond to depths that where data provide strong constraints. Model depths can be converted to elevations by subtracting 0.379 km.

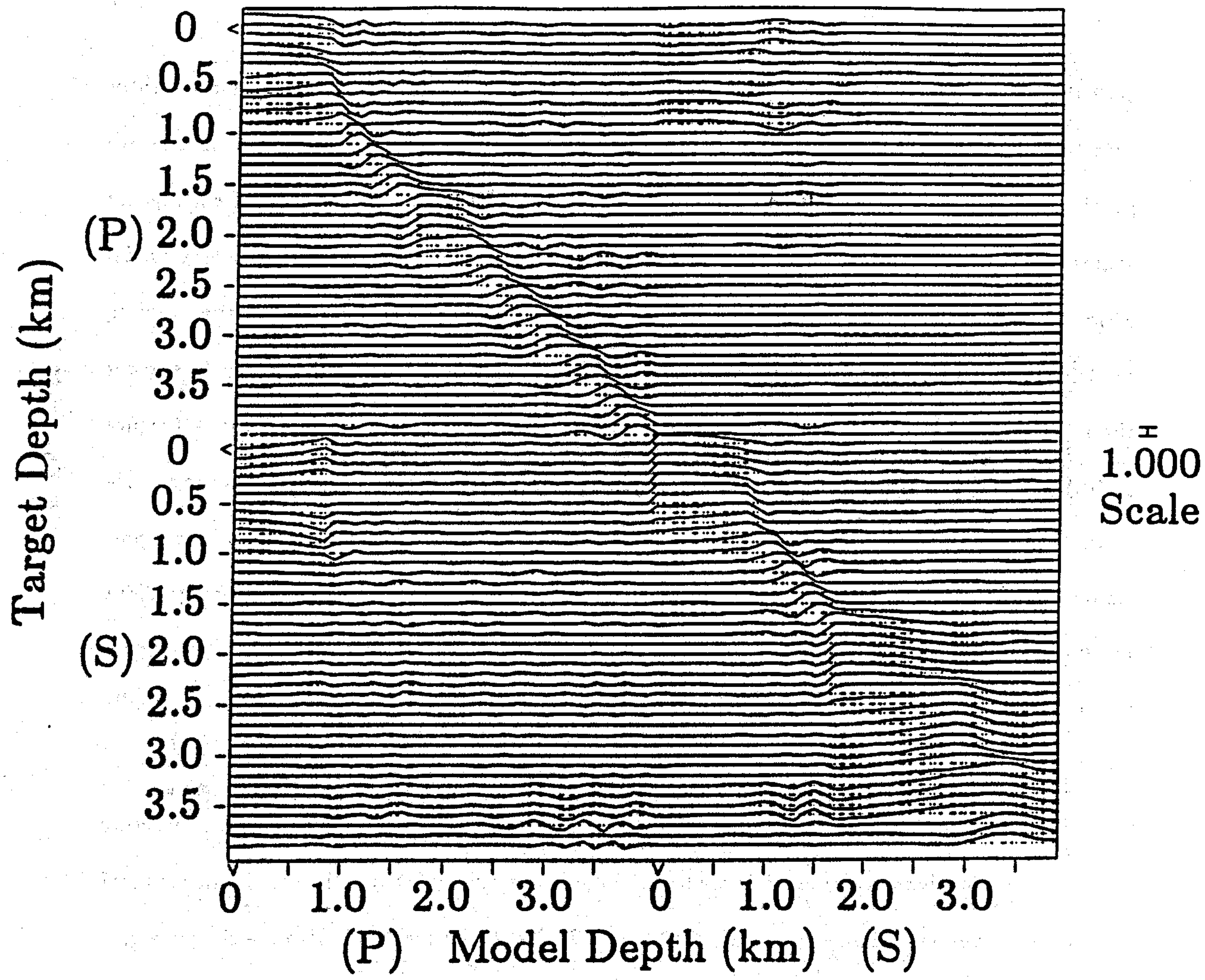


Figure 4.9 Resolving kernels for P and S-wave slowness gradient are plotted with each kernel's maximum value at full scale. This plot helps show averaging widths for small amplitude kernels in Figure (4.8). It also helps show model tradeoffs that are difficult to see in Figure (4.8). Model positions and kernels are the same as in Figure (4.8). Model depth can be converted to elevation by subtracting 0.379 km.

gradient models are best constrained by the data. P and S-slowness gradient are most tightly constrained between 1.0 and 1.5 km depth. Slowness gradient structure for both models are poorly constrained between 0.0 and 0.7 km depth. This is the result of the lack of earthquake sources in this depth interval (see Figure (4.10)). Details in the final slowness gradient model estimates and velocity models in this depth range primarily reflect features of the starting models. P and S-slowness gradient resolution between depths of 1.6 and 2.0 km depth is lower than between 0.8 and 1.5 km due to the pacity of earthquakes in this depth range (Figure (4.10)). Below 2.0 km depth, P-slowness gradient resolution is good due to bottoming wave paths recorded at more distant USGS vertical component stations. The price paid is that a portion of these wave paths lie outside the primary steam production zone. S-slowness gradient resolution below 2.0 km depth is lower than for P-waves because almost all S-wave paths constraining this part of the model correspond to upgoing wavepaths; turning S-waves are virtually absent in this depth range.

Figure (4.9) shows the averaging widths of the resolving kernels more clearly and shows how different portions of the models tradeoff. Slowness gradient features in the depth range of 1.0-1.5 km have very narrow resolving kernels and are very well resolved. P-wave slowness gradient features have broader averaging widths in the depth range of 1.6-2.3 km, indicating that the small increase in negative slowness gradient at 2.0 km depth is only marginally significant. Below 2.3 km depth, P-wave slowness gradient resolution is fairly compact and has small errors. S-wave slowness gradient resolving widths are almost 1.0 km wide between 1.6 and 2.5 km depth indicating that the increase of negative slowness gradient between 1.5 and 2.5 km is marginally significant. S-wave slowness gradients between 2.5 and 3.0 km depth have more compact resolving kernels indicating that the decrease of slowness gradient between 2.5 and 3.0 km is significant. Below 3.0 km depth S-wave slowness gradient resolving widths are broad and model values there reflect the starting model.

Tradeoffs between P and S models can be seen in Figure (4.9). The strongest tradeoffs between models are between both shallow, poorly resolved portions. While these tradeoffs do not

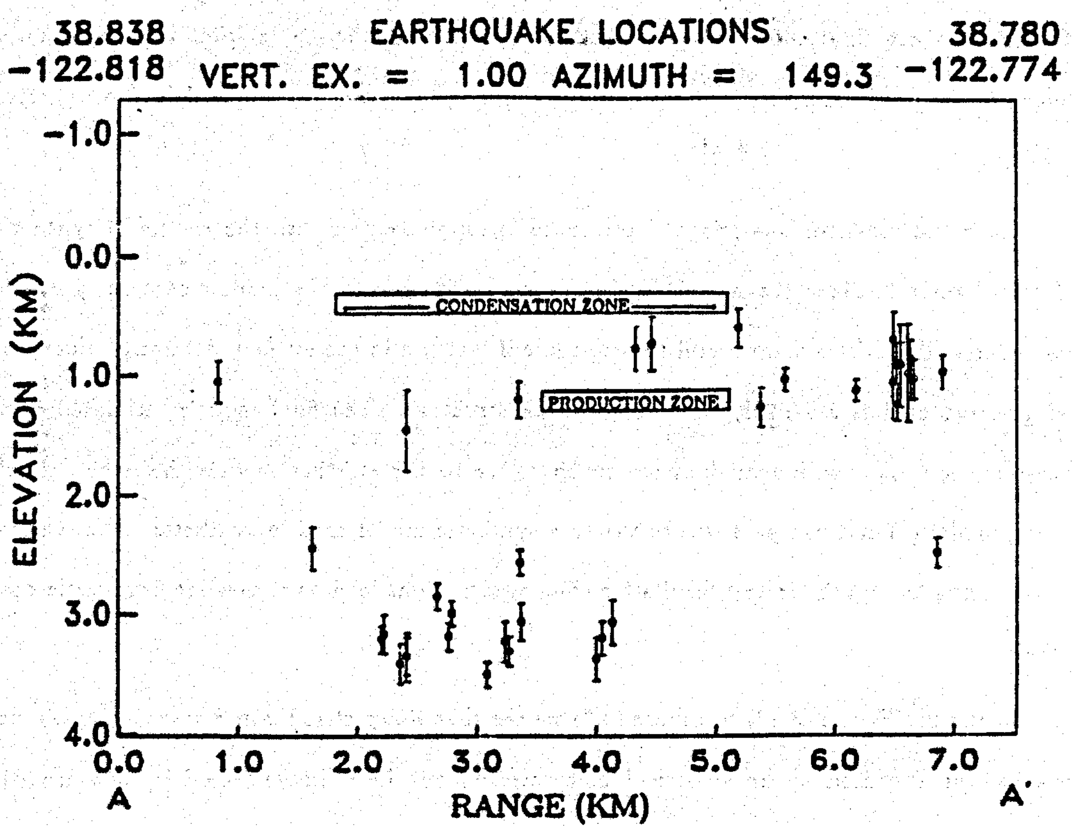


Figure 4.10 Cross section showing earthquake elevations with respect to mean sea level. The cross section ends A and A′ correspond to the surface projection shown in Figure (4.5). Positive elevations correspond to distances below sea level. The ability to plot with respect to true elevation is provided by calculating wave paths to stations at different elevations. This facilitates direct comparison of earthquake positions with features inferred from the well logs and cross sections of Stockton et al., (1984) and Heblein (1986). Inferred positions of the shallow condensation zone and shallow production zone are denoted by labels. Labels sizes do not to imply vertical or horizontal extent (see text).

significantly effect final model estimates (these portions of the model are not significantly changed in the inversion), they demonstrate the coupling between P and S models in progressive inversion.

The resolution-error analysis just presented is supplemented with the results of synthetic tests of Chapter 3. Since the models estimated here are close to the models used in synthetic tests, results of those tests are useful to determine if features in the estimated velocity and slowness gradient models are likely to represent true structure. The final models estimated from progressive inversion with real data are much closer to the starting models than the "close" starting model (1) of Chapter 3 was to the true synthetic model used in synthetic tests. Consequently, using the synthetic test results for comparison might lead to somewhat pessimistic conclusions.

Comparing Figure (3.20) to Figure (4.7) we see that sharp changes in S-wave slowness gradient can be expected to be smoothed and smeared out in estimated models, but are still representative of the true model. Comparison of Figure (3.12) with Figure (4.4) indicates that the S-wave velocity variations between 0.8 and 3.0 km depth represent somewhat smeared out properties of true velocity structure. Below 3.0 km depth, S-wave velocities are poorly constrained. Comparison of Figure (3.18) and Figure (4.6) indicate that the large negative P-wave slowness gradient between 0.4 and 0.7 km depth is probably a smoothed and smeared feature of the true P-wave slowness gradient structure. The estimated P-wave velocity model is probably representative of smoothed P-wave velocity structure between 0.8 and 3.5 km depth as indicated by comparison of Figure (4.3) and Figure (3.11), although a small (0.1-0.3 km) dc shift of velocity with depth is a good possibility. A similar dc velocity profile can be expected for estimated S-wave velocity structure (Figure (3.12)).

Figure (4.11) show estimated V_p/V_s structure for The Geysers. The most striking feature is the rapid decrease of V_p/V_s between 0.8 and 1.2 km depth. The key question is: Is this feature real?. The answer is that it is the most tightly constrained feature in the entire V_p/V_s model and is very well resolved. Inspection of Figures (4.6) and Figure (4.7) reveal that the sharp

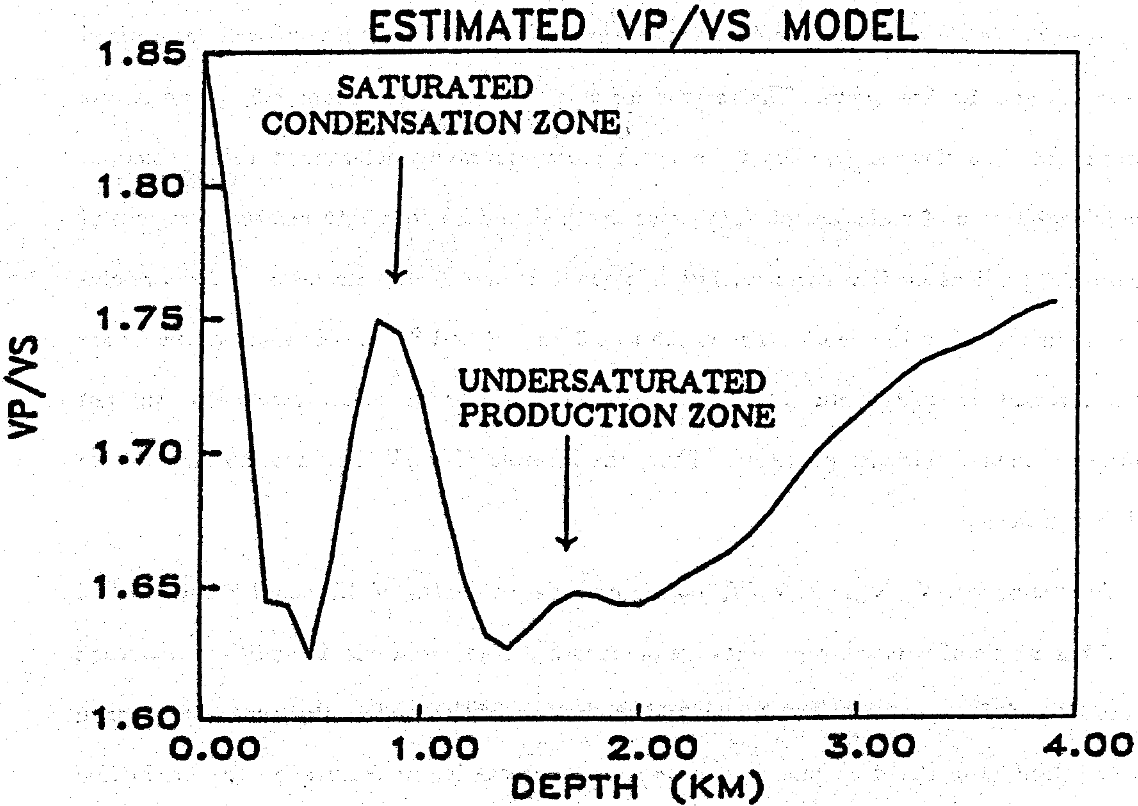


Figure 4.11 V_p/V_s structure estimated from progressive inversion. Model depths can be converted to elevation by subtracting 0.379 km. The shallow zones of high and low fluid saturation discussed in the text are indicated on the plot.

decrease in V_p/V_s is caused by combining a very small negative P-wave slowness gradient with a large negative S-wave slowness gradient over the same depth interval. Both slowness gradients have excellent resolution in this depth interval (Figure (4.8)). While absolute magnitude of V_p/V_s at any point is not tightly constrained, changes of V_p/V_s with depth are well constrained between 0.8 and 1.5 km depth. The reasons for this are clear from Chapter 3, where it was demonstrated that slowness gradients are much more accurately determined than velocities. Thus, interpretation of variations of V_p/V_s with depth should be done with respect to regions of the slowness gradient models where resolution is good. P and S-wave slowness gradient resolution also fairly good in the depth range of 2.5 to 3.0 km. P and S-wave slowness gradients are almost constant in this depth interval, so smearing due to imperfect resolution will not significantly change slowness gradients. Thus, the increase of V_p/V_s between 2.5 and 3.0 km depth is significant.

To summarize, V_p, V_s, and V_p/V_s variations between depths of 0.7 and 1.5 km, and 2.5 and 3.0 km are significant and representative of smoothed true variations. In addition, resolution of variations of V_p is good between 3.0 km and the bottom of the model. However, variations in P-wave velocities in the deep model may be less representative of features in the production zone because portions of constraining wave paths lie outside the immediate production zone. Features between 0.0 and 0.7 km depth are not well resolved and are basically representative of the starting models. Consequently, features in this depth interval such as the shallow steam anomaly near power plants 1 and 2 (see Figure (4.5)) are not discussed.

4.3.4. Assessment of Earthquake Locations

Earthquake locations with estimated error ellipses are shown in Figure (4.5) along with power plants that were producing during the recording period. Earthquake elevations with estimated standard errors are shown in Figure (4.10). A complete list of hypocenters and their dates in provided in Table (4.2). Estimates of velocity model induced location errors (see Section (2.6.2)) were comparable to conventional statistical standard errors calculated using equation (2.50). Since estimates of velocity model induced hypocenter errors proved to be overly pessimistic for joint P and S synthetic inversion in Chapter 3, they were not included in Figures (4.5) and (4.10). Results of full nonlinear synthetic tests of earthquake locations estimates in Appendix B, indicate that nonlinear errors are not significant for earthquakes located inside the recording network.

Station correction errors are listed in Table (4.1). Both conventional estimates of station correction errors obtained using equation equation (2.50) and total error, which includes an estimate of velocity model induced errors, calculated using equation (2.82) are included in Table (4.1). Synthetic tests in Chapter 3 indicated that including estimates of velocity model induced errors produced the most realistic station correction error estimates. Consequently, the larger estimates of station correction error in Table (4.1) were used as the station correction errors included in hypocenter error estimates using equation (2.50). Tests of hypocenter locations with synthetic data in Chapter 3 indicated that estimates of hypocenters, particularly hypocentral depth, are robust when using P and S-wave arrival time data in progressive inversion. By including an estimate of the velocity model induced errors, via the station corrections, estimates of hypocenter error in Figures (4.5) and (4.10) represent conservative estimates.

4.4. Interpretation of Estimated Station Corrections

Estimated station corrections generally correlated well with surficial geology. However, some station corrections are too large to be adequately explained by near surface velocity variations alone. Locations of station with large negative station corrections (GCR, GSM, TSP, TPT, AND TPL of Table (4.1)), correlate with the high velocity P-wave anomaly found by Eberhart-Phillips (1986) in the southeastern portion of the steam field (see Figure (4.1)). The same anomaly might have high S-wave velocities also since stations TSP, TPT, and TPL have large negative S-wave station corrections. Stations with the largest positive station corrections (GBO, GMM, TOT, TPU, AND TPR) are located in the northwestern portion of the steam field. A component of the northwest-southeast variation of station corrections can be explained by the fact that the northwestern stations are sited on alluvial landslides, graywackes, and cherts, whereas, the southeastern stations tend to be sited on basalts, greenstones, and serpentinites that have somewhat higher velocities than the sedimentary units. The overall pattern of station corrections is well explained by near surface velocity heterogeneities combined with a high velocity anomaly in the southeast portion of the production zone as seen in the results of Eberhart-Phillips (1986).

4.5. Overview of Geology at The Geysers

The primary reservoir rock at The Geysers in comprised of Franciscan graywacke and metagraywacke (Stockton et al., 1984, McLaughlin, 1981). The reservoir rock is overlain by a complex assemblage of melanges, greenstones, serpentinites, graywackes, and metagraywackes (McLaughlin, 1981). The main reservoir graywacke unit may be underlain by extensive intrusives (Heblein, 1986; Stockton et al., 1984). What lies below the intrusives is unknown. There is seismic evidence (Majer et al., 1987; Denlinger and Kovach, 1981) for a velocity discontinuity at elevations of 3.0-4.0 km that may represent a change in lithology or a tectonic boundary.

Extensive folding and faulting are apparent in the primary production zone (Stockton et al., 1984). Several northwest-trending fault zones are evident, the most prominent being the Big Sulfur Creek fault zone which consists of near vertical and steeply dipping faults (Stockton et al., 1984). The Big Sulfur Creek fault zone may merge with the northwest-trending Squaw Creek fault zone in the northwestern part of the field. Extensive thrust faults in the caprock are not thought to be important in influencing hydrothermal circulation (McLaughlin, 1981).

4.6. Interpretation of V_p/V_s Variations with Depth

The variations of V_p/V_s with depth in Figure (4.11) are well explained by variations in degree of fluid saturation. Toksöz et al. (1976) and Gregory (1976) showed that V_p/V_s is proportional to degree of fluid saturation. The peak of V_p/V_s at $\sim$ 1.0 km depth in Figure (4.11) corresponds to the saturated condensation zone observed to lie above the primary steam production zone in many portions of The Geysers (Mogen et al., 1985; Heblein, 1985). Due to the high degree of fluid saturation in the condensation zone, relative to other parts of the reservoir, a maximum of V_p/V_s is observed. The minimum of V_p/V_s at depths between 1.0 to 2.5 km in Figure (4.11) (elevations of 0.6 to 2.1 km) corresponds to the depth range of primary steam production (Stockton et al., 1984). Due to undersaturated conditions, V_p/V_s is lowest in the production zone, reflecting depletion of pore fluids and dominance of vapor static conditions.

Observed V_p/V_s variations are not explained by variations in fracture density. Moos and Zoback (1983) found that high fractures densities produce high V_p/V_s and decreased V_p and V_s. The highest fracture densities at The Geysers are expected to be found in the primary production zone at depths of 1.0-2.5 km in Figure (4.11). Yet, V_p/V_s is actually a minimum in this depth interval. The caprock assemblage is certainly no more fractured than the reservoir rock. In fact, Heblein (1985) proposes that fractures in the caprock are sealed by sericitic alteration. Thus, if only fracture density where considered, we would expect to find a V_p/V_s variation opposite of that actually observed. The fact that observed variations of V_p/V_s with depth are the opposite of that predicted by fracture density alone, strengthens the interpretation that observed variations of V_p/V_s between 0.5 and 2.5 km depth are caused by variations of fluid saturation.

It is interesting to note that Majer et al., (1987) observed a sharp drop of V_p/V_s at an elevation of 0.6 km where producing steam entries in the VSP well started. Although the VSP was done several kilometers outside the primary production zone, the elevation of their observed V_p/V_s decline coincides with the elevation of sharp V_p/V_s decline estimated here for the primary production zone.

4.7. Discussion of Models for The Geysers Geothermal Reservoir

Condensation zones immediately above the primary production zone are found in many parts of The Geysers (Stockton, et al., 1984, Mogen, et al., 1985, Heblein, 1985). These condensation zones are the result of vapor upflow from the main Geysers reservoir (Mogen, et al., 1985). The V_p/V_s maximum at an elevation of 0.5 km is the seismic signature of the condensation zone. Immediately below the condensation zone, is the zone of primary steam production. The vast majority of steam entries at The Geysers (Stockton, et al., 1984) lie in the region of lowest V_p/V_s. The minimum of V_p/V_s at depth of primary production is due to maximum depletion of pore fluid there, due to long term production of the porous low permeability rock there (Pruess and Narasimhan, 1982). Pore fluid depletion decreases with depth (increasing V_p/V_s). Seismicity in the most depleted portion of the production zone (depths of 1-2 km in Figure (4.10)) could be caused by volume changes as fluid is extracted from the reservoir rock (Majer and McEvilly, 1979). Using seismic focal mechanisms and geodetic information, Oppenheimer (1986) has concluded that the shallow earthquakes are induced primarily by volume changes in the reservoir. This conclusion, is consistent with the findings here, which indicate that the shallowest seismicity is located in the depth of maximum volume change (pore fluid depletion). As the production zone (and thus pore fluid depletion) extends to greater depths over time, the shallow seismicity should extend downward, filling the seismic gap evident in the elevation range of 1.5 to 2.5 km in Figure (4.10) and in the more extensive data of Eberhart-Phillips and Oppenheimer (1984).

Oppenheimer (1986) suggests that induced seismicity could be actively extending the vertical fracture system thereby enlarging the source region of steam withdrawal. If the shallow seismicity is primarily a result of volume contraction due to production of pore fluid, this has the following implication. Production of steam induces earthquakes, which then expand the potential volume of steam production by expanding fracture networks (steam production is primarily controlled by fractures). Thus, steam production and induced earthquakes would act as a feedback pair, helping to sustain each other as long as production could be sustained.

An alternative interpretation, is that the increase of V_p/V_s and decrease of seismicity in the elevation interval between 1.5 and 2.5 km could be caused by a major change in rock type or properties. Heblein (1986), postulates the existence of an extensive felsite intrusive lying immediately below the main reservoir graywacke rock. The shallow seismicity is generally confined to the depth range of the reservoir graywacke unit. Even though permeabilities are rather low in the reservoir graywacke, the largest volume of steam production could be confined to the graywacke, with much smaller amounts being removed from the felsite intrusive, due to its much lower matrix permeability and less extensive fracturing. Then, most of the volume change associated with production of pore fluid would be confined to the graywacke unit. Lower seismicity would then be expected in the depth range of the felsite intrusive due to much lower volume changes there. The deeper seismicity (at 3 km elevation) could be related to tectonic processes. For instance, The Geysers might be in a pull-apart basin between to strike-slip fault systems (the Maacama and Collayomi faults), but this interpretation is not well supported by available data (Oppenheimer, 1986). Steam entries that are observed in the intrusive (Stockton, et al., 1984) could correspond to extensive fracture zones.

Earthquake locations appear to be confined to two distinct depth intervals. Shallow seismicity at elevations of 0.60 to 1.5 km are associated with production of steam in the main reservoir graywacke. The deeper seismicity at elevations of 3.0 to 3.5 km may be associated with the depletion of fluid from a second reservoir associated with a change of rock type or tectonic boundary. Seismicity is low in the elevation interval of 1.5 to 3.0 km between these two reservoirs. This pattern can be seen in the extensive earthquake data of Eberhart-Phillips and Oppenheimer (1984), but is partially obscured by scatter of their hypocentral depths, particularly mislocation of shallow seismicity. Their earthquake locations between the free surface and an elevation of 0.5 km are suspect, due to errors in assumed shallow P-wave velocity structure, lack of accounting for large variations of station elevations, and the fact that only P-wave data were used. As demonstrated in Chapter 3 with synthetic data, the combination of these factors produces hypocentral errors of up to 1.0 km with error free data for shallow earthquakes at The

Geysers. Thus, the earthquakes located between 0.0 and 1.0-1.5 km depth by Eberhart-Phillips and Oppenheimer (1984) are probably mislocated and instead lie over a fairly narrow elevation interval of 0.5-1.5 km. The deep cloud of hypocenters of Eberhart-Phillips and Oppenheimer (1984) are also probably more diffuse than the true seismicity in that depth interval due to limited earthquake depth resolution of P-wave data.

The following model is proposed for The Geysers which satisfies all currently available seismic, gravity, geodetic, and geologic data. The Geysers is characterized by two distinct reservoirs separated by a more fluid saturated unit of very low permeability, possibly an extensive intrusive, hydrothermally sealed unit, or unfractured unit of low permeability. Hydrothermal fluid circulates from the deep reservoir though vertical fractures into the shallow reservoir. Low V_p/V_s indicates that steam is also being produced from the pore fluid of shallow reservoir rock as proposed by Pruess and Narasimhan (1982). The depletion of fluid in both reservoir horizons and resulting volume change have induced seismicity due to subsidence, contraction, and possibly conversion of aseismic slip to stick-slip due to deposition of silica on producing fractures (Allis, 1982). Shallow seismicity is primarily confined to the elevation interval of the shallow productions zone due to mass withdrawal amplification of contraction and subsidence effects there. Since lithostatic stresses are proportionally small compared to horizontal contraction and regional stresses for these shallow elevations, resulting earthquakes have reverse and strike-slip mechanisms as proposed by Oppenheimer (1986). The mass withdrawal and resulting subsidence in the deep reservoir induces earthquakes immediately above and inside it. Normal faulting is observed here because subsidence is localized and lithostatic stresses are much larger at these depths (Oppenheimer, 1986). The deep earthquakes appear to cluster about a possible tectonic or lithological boundary suggested by seismic reflection data.

The mechanism proposed by Allis (1982) could also explain the correlation of earthquake locations with zones of fluid depletion. In his model, dehydration of the reservoir is associated with movement of fluid to fractures. Silica is deposited on producing fracture surfaces and clay and fault gauge are hardened due to dehydration. The resulting increase is the coefficient of friction would convert aseismic creep to stick-slip movement.

The deep earthquakes are confined to a much smaller area of the field than the shallow seismicity (Eberhart-Phillips and Oppenheimer, 1984 and Figure (4.10)). This is due to the fact that deep reservoir depletion is dependent on extensive vertical fracture systems to allow migration of fluid up to the shallow production zone. The location of the deep earthquakes suggests that hydrothermal fluid migration is associated with a fault zone between the Big Sulfur Creek and Squaw Creek fault systems. It appears that the extent of these vertical fractures is limited. Thus, a smaller volume of the deep reservoir is being depleted due to restrictions on flow and a smaller volume of deep seismicity is observed.

This model of The Geysers is obviously much simplified because of the limited data available, especially about lateral variations of material properties of the steam field. The elevation and extent of the condensation zone are likely to vary laterally. This model resembles that proposed by White et al., (1971) in that it involves migration of fluid from a deep reservoir. Production in the shallow reservoir is probably the result of hydrothermal circulation of fluid from the deep reservoir combined with the production of steam from the shallow reservoir rock as described by Pruess and Narasimhan (1982) resulting in pore fluid depletion in the shallow field and low V_p/V_s there.

The proposed model is quite speculative. Determination of the rock properties immediately below the zone of primary production will reveal how realistic this model is. The most speculative aspect of the model is the postulation of a second deep reservoir that is located immediately below a seismic discontinuity; the existence of which is far from conclusively proven at this point. There are indications of a post-critical reflection arriving several seconds after the first S-wave arrival at epicentral distance of more than 5.0 km from earthquakes recorded at the edges of the temporary network. Waveform modeling of these phases may provide some constraints on seismic properties of the structure between 3.0 to 4.0 km elevation in the steam field.

4.8. Summary and Conclusions

The top of the steam reservoir is clearly defined by a large decreases of V_p/V_s at the condensation zone-production zone contact. The depth interval of maximum steam production coincides with minimum observed V_p/V_s, and V_p/V_s increases below the shallow primary production zone suggesting that reservoir rock becomes more fluid saturated. These results suggest that VSP and crosshole combined P and S-wave surveys could be used to map hydrothermal convections cells by delineating steam-condensation zone boundaries.

If the model proposed above proves to be valid, it could have profound consequences for determining the potential to expand production to greater depths. If the graywacke unit has considerable unproduced portions at greater depths, the prospects for expanding production are good, especially if the production-earthquake feedback mechanisms is actually occuring. If, however, the graywacke unit is substantially depleted and increasing production requires producing steam from an intrusive unit or impermeable graywacke unit, the prospect are not as good. Producing steam entries are much rarer below the main production zone than in the main producing graywacke and metagraywacke (Stockton, et al., 1984) and seem to require large intersecting fracture systems. Obtaining production from the underlying units would seem to entail the difficult task of delineating zones of extensive fractures. While this might become feasable to some extent using multi-offset P and polarized S VSP surveys (Majer, et al., 1987), it would be difficult.

The delineation of condensation zones is critical in terms of deciding where to drill production wells. Heblein (1985) emphasized the importance of determining the horizontal and vertical boundaries of hydrothermal cells which define the maximum volume dimensions of production in specific portions of the reservoir. These boundaries may be characterized by condensation zones that have prominent V_p/V_s signatures. This suggests that high resolution P and S VSP and crosshole surveys may prove very useful for finding the horizontal and vertical boundaries of hydrothermal cells.

Finally, this study highlights the practical benefits of doing progressive inversion using both P and S-wave data. Careful accounting for station elevation differences and low near surface velocities resulted in well constrained earthquake locations. Improved earthquake locations and estimated V_p/V_s structure allowed more definitive correlation of seismicity with specific features in The Geysers geothermal field than would have been possible using P-wave data alone.

References

Allis, R. G., 1982. Mechanism of induced seismicity at The Geysers geothermal reservoir, California, *Geophys. Res. Lett.*, **9**, 629-632.

Andrews, M. C., Mooney, W. D., and Meyer, R. P., 1985. The relocation of microearthquakes in the northern Mississippi embayment, *J. Geophys. Res.*, **90**, 10,223-10,236.

Crampin, S., 1985. Evaluation of anisotropy by shear wave splitting, *Geophysics*, **50**, 142-152.

Denlinger, R. P., and Kovach, R. L., 1981. Seismic-reflection investigations at Castle Rock Springs in The Geysers geothermal area, *U. S. Geol. Surv. Prof. Pap. 1141*, 117-128.

Eberhart-Phillips, D., and Oppenheimer, D. H., 1984. Induced seismicity in The Geysers geothermal area, California, *J. Geophys. Res.*, **89**, 1191-1207.

Eberhart-Phillips, D., 1986. Three-dimensional velocity structure in Northern California Coast Ranges from inversion of local earthquakes, *Bull. Seism. Soc. Am.*, **76**, 1025-1052.

Gregory, A. R., 1976. Fluid saturation effects on dynamic elastic properties of sedimentary rocks, *Geophysics*, **41**, 895-921.

Heblein, J. J., 1985. Discussions on a type of reservoir cell boundary in The Geysers steam field, *Proceedings, Tenth Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, California*, **SGP-TR-84**, 89-96.

Lipman, S. C., Strobel, C. J., and Gulati, M. S., 1978. Reservoir performance of The Geysers field, *in*, *Proceedings of the Larderello workshop on geothermal resource assessment and reservoir engineering, Sept. 12-16, 1977, Pisa, Italy, Geothermics*, **7**, 209-219.

Ludwin, R. S., Cagnetti, V., and Bufe, C. G., 1982. Comparison of seismicity in The Geysers geothermal area with the surrounding region, *Bull. Seism. Soc. Am.*, **72**, 863-871.

Majer, E. L., and McEvilly, T. V., 1979. Seismological investigations at The Geysers geothermal field, *Geophysics*, **44**, 246-269.

Majer, E. L., McEvilly, T. V., Eastwood, F., and Myer, L., 1987. Fracture detection using P- and S-wave VSP's at The Geysers geothermal field, *Geophysics*, (in press).

McGarr, A., 1976. Seismic moments and volume changes, *J. Geophys. Res.*, **81**, 1487-1494.

McLaughlin, R. P., 1981. Tectonic setting of pre-Tertiary rocks and its relation to geothermal resources in the Geysers-Clear Lake area, *U. S. Geol. Surv. Prof. Pap. 1141*, 3-24.

Mogen, P., Pittinger, L, and Magers, M., 1985. 1982 thermal shallow reservoir testing, *Proceedings, Tenth Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, California*, **SGP-TR-84**, 133-139.

Moos, D., and Zoback, M. D., 1983. In situ studies of velocity in fractured crystalline rocks, *J. Geophys. Res.*, **88**, 2345-2358.

Oldenburg, D. W., 1984. An introduction to linear inverse theory, *IEEE, Trans. Geoscience and*

Remote Sensing, GE-22, 665-674.

Oppenheimer, D. H., 1986. Extensional tectonics at The Geysers geothermal area, California, *J. Geophys. Res.*, **91**, 11463-11476.

Pavlis, G. L., 1982. Progressive inversion, Ph.D. Dissertation, University of Washington, Seattle, Washington.

Pruess, K., and Narasimhan, T. N., 1982. On fluid reserves and the production of superheated steam from fractured, vapor-dominated geothermal reservoirs, *J. Geophys. Res.*, **87**, 9329-9339.

Ridley, A. P., and Vantine, J. V., 1978. Geologic exploration methods for siting power plants and other structures in landslide and hydrothermally altered terrain at The Geysers, California, *Geothermal Resources Council, Transactions*, **2**, 563-565.

Stockton, A. D., Thomas, R. P., Chapman, R. H., and Dykstra, H., 1984. A reservoir assessment of The Geysers geothermal field, *J. Pet. Tech.*, **36**, 2137-2159.

Toksöz, M. N., Cheng, C. H., and Timur, A., 1976. Velocities of seismic waves in porous rocks, *Geophysics*, **41**, 621-645.

White, D. E., Muffler, L. J. P., and Truesdell, A. H., 1971. Vapor-dominated hydrothermal systems compared with hot water systems, *Economic Geology*, **66**, 75-97.

Zoback, M. D., and Hickman, S., 1982. In situ study of the physical mechanisms controlling induced seismicity at Monticello Reservoir, South Carolina, *J. Geophys. Res.*, **87**, 6959-6974.

Chapter 5

Inversion for the First Degree Moment Tensor of Microearthquakes at The Geysers Geothermal Field

5.1. Introduction

Estimates of seismic source properties are one of the most important pieces of information extracted from recordings of microearthquakes. Source studies using microearthquakes can constrain the mechanism and orientation of faulting in an area and provide estimates of principal stress orientations. These estimates form the basis for interpretations about ongoing deformation associated with the earthquakes and possible relations between seismicity and tectonic stresses. At The Geysers the relationship between seismicity, tectonic and locally induced stresses is unclear. Based on alignment of the extensional principal stress direction estimated from the seismicity as a whole with that obtained from regional geodetic data (Prescott and Yu, 1986), Oppenheimer (1986) concludes that regional tectonic stresses are much larger than the stresses induced locally through geothermal activities. His ability to drawn firm conclusions is hampered by frequent ambiguities in fault plane solutions used to infer principal stress directions.

Oppenheimer (1986) has demonstrated that understanding of the relationship between seismicity and steam production at The Geysers requires determining the source properties of microearthquakes there. Bufe et al., (1981) inferred that the wide variations of fault plane solutions found at The Geysers using P-wave first motion data were a function of time. Oppenheimer (1986) estimated the stress field orientation at The Geysers from 210 fault plane solutions. He concludes that focal mechanisms of earthquakes at The Geysers geothermal field are a function of focal depth. He suggests that shallow earthquake focal mechanisms are dominantly strike-slip and reverse whereas deeper focal mechanisms predominantly exhibit normal faulting.

His estimation of stress field orientation and variation of focal mechanisms with depth are hampered by the common problem of nonunique fault plane solutions. For many shallow events there is an ambiguity between pure strike-slip and pure dip-slip mechanisms. These ambiguities are not due to a paucity of P-wave first motion data. Numerous P-wave first motions are available for earthquakes at The Geysers owing to the extensive seismic recording network operated in The Geysers area by the USGS. Often however, P-wave first motions are absent from the central portions of the focal sphere for shallow events because none of the stations are close enough to the epicenter. Observations from distant stations that sample the central portion of the focal sphere are absent due to attenuation of microearthquake signals. Oppenheimer (1986) noted that for some events, fault plane solutions were completely ambiguous; strike-slip, reverse-slip, and normal-slip solutions could fit the same first motion data. An alternative approach to estimate source mechanisms and principal stress orientations is to invert for the seismic moment tensors of microearthquakes at The Geysers. The method of Stump and Johnson (1977) is used to estimate first order seismic moment tensors for several microearthquakes at The Geysers geothermal field.

5.2. Source Characterization

The moment tensor formulation is used to represent the seismic source in space and time. Assuming a seismic source can be represented as a set of equivalent body forces, the source can be written as a series of moments. For small sources or large wavelengths, only the first term of the series is retained (point source approximation), and the displacement at any point and time can be written as

$$U_k(\underline{x}',t') = G_{ki,j}(\underline{x}',t';\underline{0},0) \otimes M_{ij}(\underline{0},t') \tag{5.1}$$

where U_k is the displacement in the k direction, G_{ki} is the Green function, M_{ij} is the moment tensor, $_{,j}$ indicates derivative with respect to x_j, and $\otimes$ represents temporal convolution. A more complete derivation of (5.1) is given in Stump and Johnson (1977).

In the frequency domain, equation (5.1) reduces to

$$U_k(\underline{x}',f) = G_{ki,j}(\underline{x}',f;\underline{0},0)\cdot M_{ij}(\underline{0},f) \tag{5.2}$$

If the propagation paths effects ($G_{ki,j}$) are known, one can determine the source (M_{ij}) from a set of observational data (U_k) by solving this set of linear equations.

In the implementation used here, Fourier transforms of the data and Green functions are calculated, and the moment rate tensor ($\dot{M}_{ij}$) is solved for in the frequency domain. An inverse Fourier transform is used to obtain $\dot{M}_{ij}$ in the time domain. Then $\dot{M}_{ij}$ is detrended to eliminate spurious dc offsets. The resulting estimates of $\dot{M}_{ij}$ are integrated to yield M_{ij}. Detrending of $\dot{M}_{ij}$ is physically justified because the moment rate tensor elements cannot have a permanent dc offset. If perfect data were available, detrending would not be required but all seismic data are intrinsically bandlimited, and the instruments used here (4.5 Hz velocity transducers) have limited low frequency responses. Consequently, detrending of $\dot{M}_{ij}$ is used.

Since the complex frequency dependence is obtained for each moment tensor element, moment tensor elements are not required to have a common time function. This allows inversion for complex sources that could have several physical source components with different time histories. An alternative approach is to solve for the moment tensor element time functions using the multichannel vector decomposition (MVD) method developed by Oldenburg (1982), as presented by Sipkin (1986). Allowing all moment tensor elements to have their own time functions eliminates errors in moment tensor estimates, in cases with source multiplicity, that are inherent in time domain approachs that assume a common time function for all moment tensor elements (see Sipkin (1986) for some examples).

The moment tensor characterization of seismic sources provides a means for estimating source properties of microearthquakes. Stump and Johnson's (1977) approach is completely general; no restrictive assumptions are required about physical source types or the time dependence of moment tensor elements. All physical source types can be included; isotropic (volume) sources, compensated linear vector dipole (CLVD) sources, and double-couple sources. One practical advantage of the moment tensor approach over first motion methods is that it does not require nearly as many recording stations to constrain seismic source properties. Only 6 components of ground motion (two three-component stations) are required in theory, although in practice about 15 components of ground motion are recommended to ensure reliable results. Another advantage is that azimuth and takeoff angle coverage need not be as comprehensive as for P-wave first motion approaches. Consequently, it is possible to estimate source properties of earthquakes that are not completely surrounded by recording stations, something that is not possible when using only P-wave first motion data.

This method of moment tensor estimation has not been applied to microearthquakes before. Stump and Johnson (1984) have used the method to characterize nuclear explosion sources using near-field data. Moment tensor inversions with restrictions on physical source type (pure deviatoric) and moment tensor time dependence have been applied to microearthquake data by Saikia and Herrmann (1986). However, their approach requires assuming that all moment tensor elements have the same time function and that the time function is known. The result of this type of inversion is simply a static estimate of the moment tensor elements. Since the source time function is intrisically unknown, any errors in the assumed time function will produce errors in the static moment tensor estimate. Further, if all moment tensor elements do not actually have the same time function, another component of error will be added to the static moment tensor estimate. The approach used here allows each moment tensor to have an independent time function. This requires more data than the time domain approach of Saikia and Herrmann (1986), Langston (1981), and Langston and Helmberger (1977), but yields more complete information about source properties.

Estimated moment tensors can be decomposed into isotropic and deviatoric components. The relations are

$$M_{TR}\,(\text{isotropic}) = \frac{1}{3} M_{ij}\,\delta_{ij} \tag{5.3}$$

$$D_{ij}\,(\text{deviatoric}) = M_{ij} - M_{TR}\,\delta_{ij}\,. \tag{5.4}$$

If prior knowledge is available about the source, then appropriate constraints can be placed on equation (5.2). Constraints on equation (5.2) were not used when inverting for microearthquake moment tensors. We wanted to investigate if unconstrained moment tensor inversions would produce moment tensor estimates consistent with the common assumption for earthquakes of a single double-couple source. The frequency domain approach was used to avoid errors due to possible source multiplicity. The point source assumption is valid for the microearthquakes used here, since source dimensions are small compared to the wavelengths represented in the observed data.

In general, D_{ij} is comprised of three double couples. The eigenvalues and corresponding eigenvectors of D_{ij} describe the magnitude and orientation, respectively, of the principal stress axes acting at the source. These principal stress axes represent the quantity that is uniquely determined (within a range of uncertainties due to errors in U_k and $G_{ki,j}$) by moment tensor inversion. Decomposition of D_{ij} into physical source components is fundamentally nonunique (unless D_{ij} happens to consist of a single double couple) (Geller, 1976). Julian (1986) uses linear-programming methods to investigate the range of possible physical source mechanisms that a particular moment tensor solution is consistent with. The common approach of decomposing D_{ij} into double couple and compensated linear vector dipole (CLVD) components is not particularly meaningful due to its intrinsic nonuniqueness unless it is believed that both components are truly contained in the seismic source. A simple shear dislocation earthquake source can have nonzero isotropic and CLVD components if the rupturing fault plane has nonzero curvature (Backus and Mulcahy, 1976). The decomposition of D_{ij} into CLVD and double couple components does give a measure of the departure of the estimated source from a planar faulting single double-couple earthquake model. A simple measure of the departure of D_{ij} from a single double couple is the ratio of the smallest and largest eigenvalues of D_{ij}.

5.3. Data Analysis

A 9 station network was deployed in The Geysers as described in Chapter 4 (see Figure(4.1)). Data were recorded at 200.32 samples/sec using three-component 4.5 Hz velocity-transducer geophones. The data were anti-alias lowpass-filtered using a 5-pole Butterworth filter at 50 Hz and high-pass filtered at 0.2 Hz with two 1-pole Butterworth filters before sampling. The rapid decrease of displacement magnification of the velocity transducers below 4.5 Hz, combined with the 12 bit resolution of the recording system, limited the frequency band of good signal to noise to the range of 1.0 to 50 Hz.

Electrical problems with some of the recorders resulted in increased noise at low (<1 Hz) frequencies, so moment tensor estimates below 1 Hz are considered to be primarily noise. This does not significantly effect the results of the moment tensor inversions because source corner frequencies for the microearthquakes used here are in the range of 6 to 10 Hz. It does however, necessitate detrending of the moment rate tensor in the time domain as discussed earlier.

The Green functions were calculated using a spectral wavenumber-frequency approach similar to the reflectivity method of Fuchs and Müller (1971). An importance difference is that the entire model between the free surface and the model bottom is the reflectivity zone; all reverberations, including free surface reflections, are included. The resulting Green functions represent the complete medium response. Since the wavenumber response is computed as a function of frequency, the frequency domain Green functions used in equation (5.2) are had directly.

To reduce problems associated with station site variations of corner frequency, the data were lowpass filtered using a 2-pole Butterworth filter at 10 Hz. This also mitigated the problems of P-wave and S-wave corner frequency differences observed at some stations. It also decreased the burden of Green function computations by reducing the maximum frequency required. The price paid is that moment tensor time functions will represent lowpass-filtered versions of true source time function(s).

The P and S-wave velocity models estimated for The Geysers in Chapter 4 are used to specify the velocity structure for Green function calculations. Anelastic attenuation was included by specifying a Q model for The Geysers consistent with the results of Majer and McEvilly (1979). Values of 50-100 were used for Q_P, and values of 40-80 were used for Q_S. The low Q values were used near the free surface and the higher values used in the production zone. These low Q values were used because attenuation at The Geysers is not realistically represented by Q values larger than 100.

Proper phase matching of observed S-P times with Green function S-P times is important to ensure the success of the moment tensor inversions. Earthquake locations estimated in Chapter 4 were used to define initial hypocenter-receiver azimuths and distances. Since recording station were located at different elevations, the predicted S-P times for initial hypocenter-receiver distance did not always match observed S-P times. Hypocenter-receiver distances were modified so as to produce correct Green function S-P times. Hypocenter-receiver azimuths are preserved but takeoff angles are slightly different. For a few stations take-off-angles were altered by as much as 30°, but for most stations take-off-angles were not changed by more than 5° - 10°.

Since a frequency domain inversion is used, it would be difficult to use windows about certain phases in the inversion. Small time windows about the first P and S-wave pulses would not provide the frequency bandwidth or resolution that help to produce reliable estimates of the moment tensor. Truncation effects due to windowing are accentuated for short time windows. Consequently, complete seismograms were used for all components in inverting for the moment tensor. Ten seconds of data were used in the inversions. For stations that had shorter records zeros were added to give total lengths of 10 seconds.

5.4. Inversion Results

Moment tensor inversions were done for three earthquakes at The Geysers. Two of the events were shallow, approximately 2 km below station elevations, and the third event was deeper, approximately 4 km below station elevations. The shallow events correspond to the depth interval where the strike-slip, normal-slip, reverse-slip ambiguity is most pronounced. Oppenheimer (1986) found that most events in this depth range had strike-slip solutions with a smaller number of events having reverse-slip mechanisms. The deeper event corresponds to the depth interval where predominantly normal faulting mechanisms are found (Oppenheimer, 1986).

Results of moment tensor inversion are displayed in the following manner. The orientations of the eigenvectors of D_{ij} are plotted on stereographic lower hemisphere projections along with available P-wave first motion data. The P-wave first motion data come from the temporary network and USGS stations. In order to obtain as many first motions as possible, USGS stations outside the primary production zone at The Geysers were used. The P-wave velocity model estimated in Chapter 4, is not adequate to accurately determine azimuth and takeoff angles for stations outside The Geysers for two reasons. Firstly, P-wave velocities are only estimated to a depth of 4.0 km and more distant station arrivals correspond to rays bottoming below this depth, in a part of the model that is only a guess at the true velocity structure. Secondly, Eberhart-Phillips (1986) has found significant lateral variations of P-wave velocity structure outside The Geysers so azimuthal estimates may be in error due to out-of-vertical-plane propagation paths. The estimated position of first motions on the focal sphere of the distant USGS readings may have substantial uncertainties.

While record lengths of 10 seconds were used in the moment tensor inversions, moment tensor results are plotted for times less than one second. This was done because source durations are short, approximately 0.1 sec and the fact that the Green functions do not contain coda waves durations as long as seen in the observed data. Also, some components of the observed data had small noise glitches approximately two seconds after the primary S-wave arrival and these glitches contaminate the moment tensor time functions after several seconds. Windows of less that one second were used to detrend the moment rate tensor. Since source durations were short for these microearthquakes, this approach is reasonable.

5.4.1. Event 2131638

The location of this event is shown in Figure (5.1) along with the 7 stations used in the moment tensor inversion. Three-component seismograms were available at all stations yielding 21 components of ground motion for the inversion. This event had a USGS M_c of 1.6 and the inversion yielded a scalar moment estimate of $\approx 2.0\times 10^{19}$ dyne-cm. The first 0.5 seconds of the principal stress axes orientations, estimated from the moment tensor inversion, are displayed in Figure (5.2). It can be seen that the orientations of the principal stress axes show only small variations with time except for the intermediate axis at early times.

The fault plane solution shown in Figure (5.2) is well constrained by the distribution of P-wave first motions. Note that the P and T principal stress directions predicted by the P-wave first motions agree with the moment tensor estimates (Table (5.1)). The magnitude of the intermediate stress axis was about 0.25 that of the maximum stress for much of the time interval shown in Figure (5.2). Since the intermediate stress is nonzero, the intermediate axis is not required to coincide with the intersection of the nodal planes.

If the north-most and west-most dilatations were unavailable in Figure (5.2), a normal faulting mechanism would be compatible with the remaining P-wave first motions. The moment tensor solution precludes this mechanism as being significant for this event.

5.4.2. Event 2181937

The location of this event is shown in Figure (5.3) along with the 6 stations used in the moment tensor inversion. A total of 17 components, consisting of 6 vertical and 11 horizontal components, were used in the inversion. This event had a USGS M_c of 1.8 and the inversion yielded a moment estimate of $\approx 8.0\times 10^{19}$ dyne-cm. The first 0.5 seconds of the principal stress

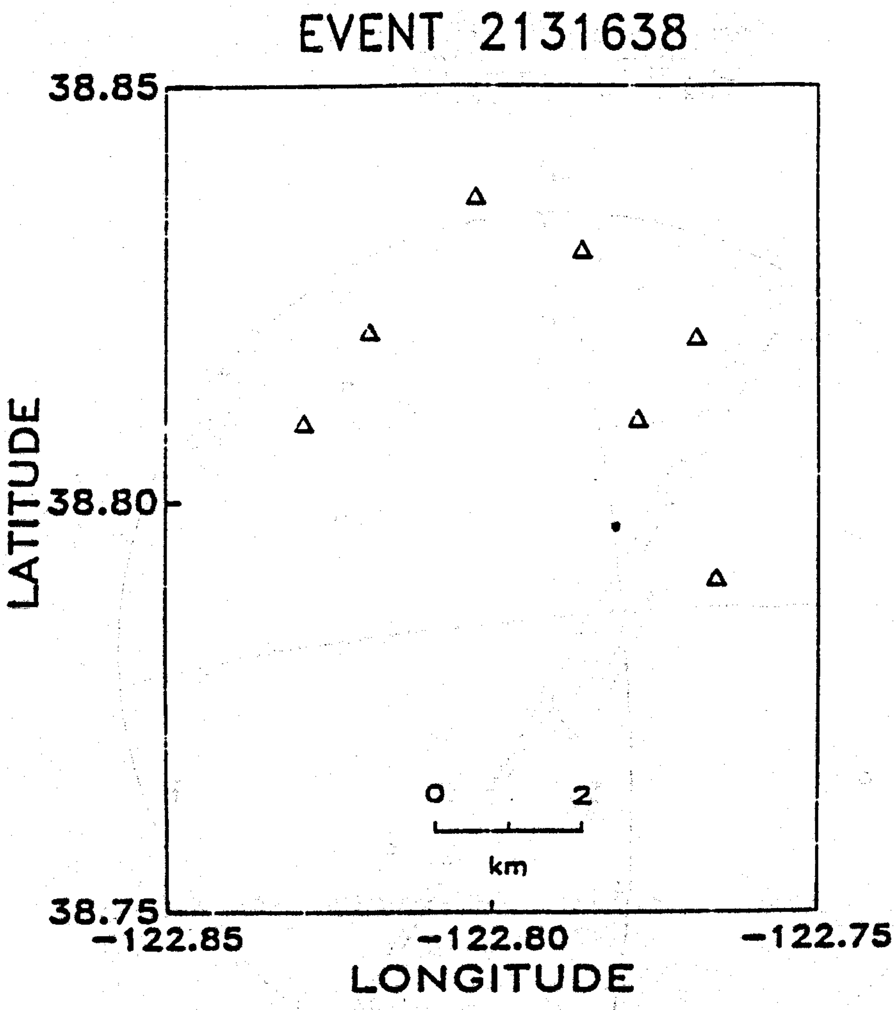


Figure 5.1 Map showing relation of earthquake 2131638 epicenter (small circle) to stations (Δ) used in moment tensor inversion. The hypocentral depth is 2.34 km below an elevation of 1.28 km above sea level, representing a ≈ 2 km depth below the recording station elevations. Note the epicenter-station recording geometry has an azimuthal gap of ≈ 180°. The range of epicenter-station distances provides good takeoff-angle coverage.

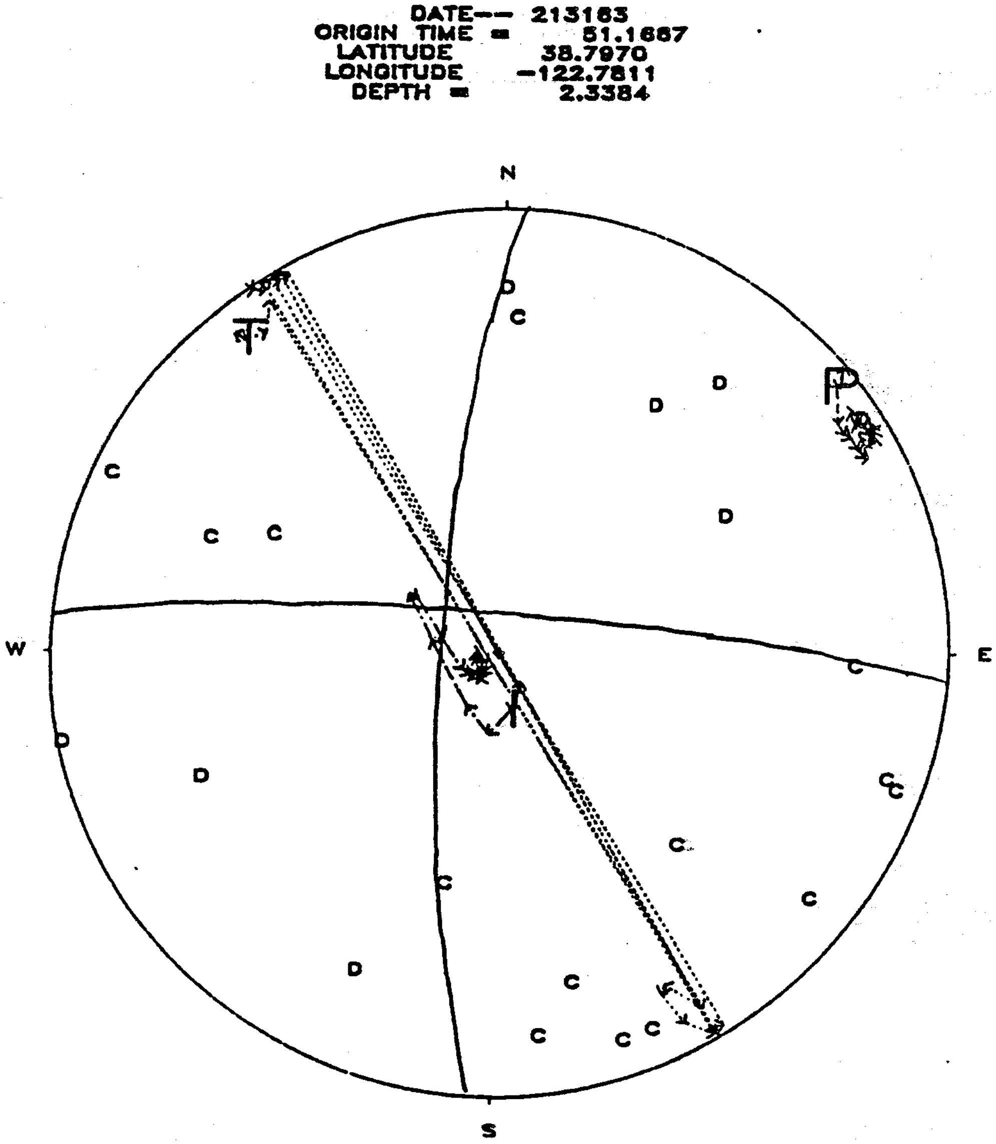


Figure 5.2 Lower-hemisphere equal-area-stereographic plot of P-wave first motions, time varying principal stress axes estimated from the moment tensor inversion, and double-couple fault plane solution for event 2131638. Compressional and dilatational P-wave first motions are plotted as (C) and (D), respectively. The tension axis starting time point is denoted by the large (T) and the fine-dashed line is its time history, where the arrow heads point toward the next point in time. Broken lines spanning the plot represent excursions required to plot stress axes points solely on the lower hemisphere. The compression axis starting time is denote by a large (P) and its time history by a medium-dashed line. The intermediate stress axis start time is denoted by a large (I), and its time history by a dot-dash line. The solid lines are nodal planes drawn to satisfy the first motion data. The inconsistent compression in the upper right quadrant corresponds to a distant station and its position is subject to errors as described in the text.

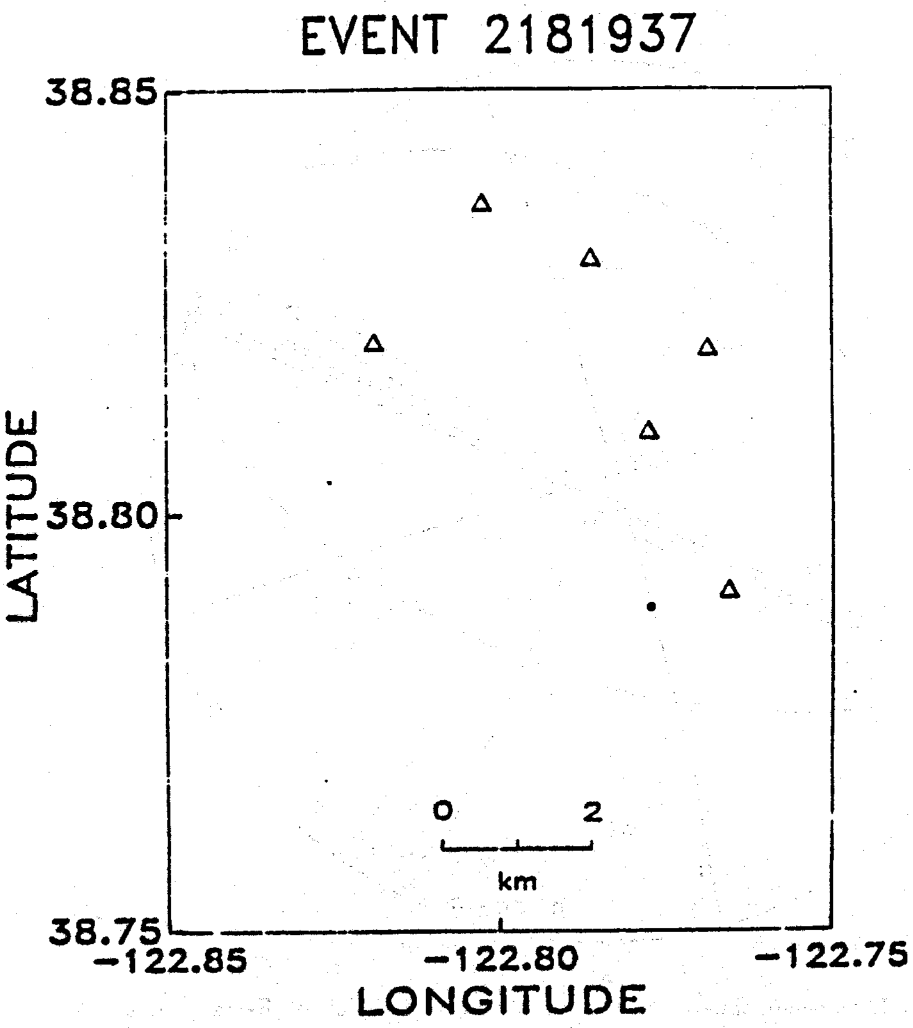


Figure 5.3 Map showing relation of earthquake 2181937 epicenter (small circle) to stations (Δ) used in moment tensor inversion. The hypocentral depth is 2.42 km below an elevation of 1.28 km above sea level, representing a ≈ 2 km depth below the recording station elevations. Note the epicenter-station recording geometry has an azimuthal gap of ≈ 250°. The range of epicenter-station distances provides good takeoff-angle coverage.

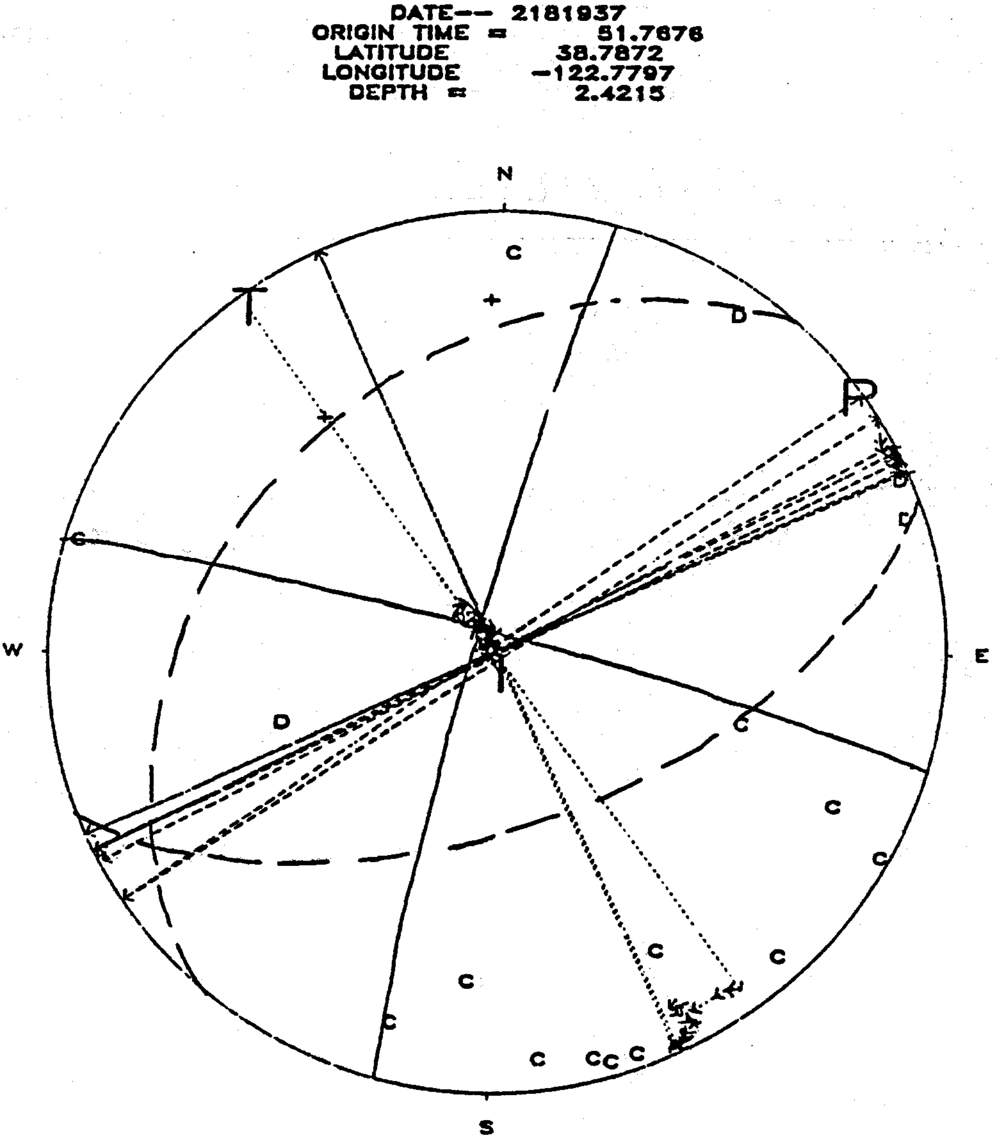


Figure 5.4 Lower-hemisphere equal-area-stereographic plot of P-wave first motions, time varying principal stress axes estimated from the moment tensor inversion, and double-couple fault plane solution for event 2181937. Compressional first motions are plotted as (C) for impulsive arrivals and (+) for emergent arrivals. Dilatational P-wave first motions are shown as (D). Convention for principal stress axes is the same as in Figure (5.2). Broken lines spanning the plot represent excursions required to plot stress axes points solely on the lower hemisphere. The solid lines are nodal planes drawn to satisfy the first motion data and the moment tensor estimates of principal stress orientations. The long-dashed line nodal planes represent a solution compatible with the first motions. Note that a wide range of fault plane solutions, ranging from almost pure normal faulting (dashed nodal lines) to pure strike-slip faulting (solid nodal lines) are compatible with the P-wave first motions.

axes orientations, estimated from the moment tensor inversion, are displayed in Figure (5.4). It can be seen that the orientations of the principal stress axes show only small variations with time. This is in spite of the fact that there is a large azimuthal gap in station coverage.

The fault plane solutions shown in Figure (5.4) demonstrate that a wide range of focal mechanisms are consistent with the P-wave first motion data. Solutions ranging from nearly pure normal-slip to pure strike-slip are consistent with the first motion data (Table(5.1)). The moment tensor inversion solution has a intermediate axis that is small indicating that a single double couple is dominant. Consequently, the intersection of the nodal planes should coincide closely with intermediate axis position for any fault plane solution consistent with the first motion data. Thus, the moment tensor solution constrains the solution to be dominantly strike-slip.

This is the type of event that made Oppenheimer's (1986) reduction of his fault plane solution data difficult. The P-wave first motions are consistent with both strike-slip mechanisms (postulated shallow event mechanism) and normal-slip mechanisms (postulated deep event mechanism). He would have been forced to discard this event since it contains no constraints with respect to the depth-dependent-focal-mechanism hypothesis. Here, the moment tensor estimate constrains the focal mechanism to be dominantly strike-slip, consistent with Oppenheimer's (1986) predictions for shallow events.

5.4.3. Event 2200908

The location of this event is shown in Figure (5.5) along with the 7 stations used in the moment tensor inversion. A total of 17 components, consisting of 7 vertical and 10 horizontal components, were used in the inversion. This event had a USGS M_c of 2.3 and the inversion yielded a moment estimate of $\approx 3.0\times10^{20}$ dyne-cm. The first 0.8 seconds of the principal stress axes orientations, estimated from the moment tensor inversion, are displayed in Figure (5.6). It can be seen that the orientations of the principal stress axes show only small variations with time. This is spite of the fact that there is a very limited range of takeoff angles represented in

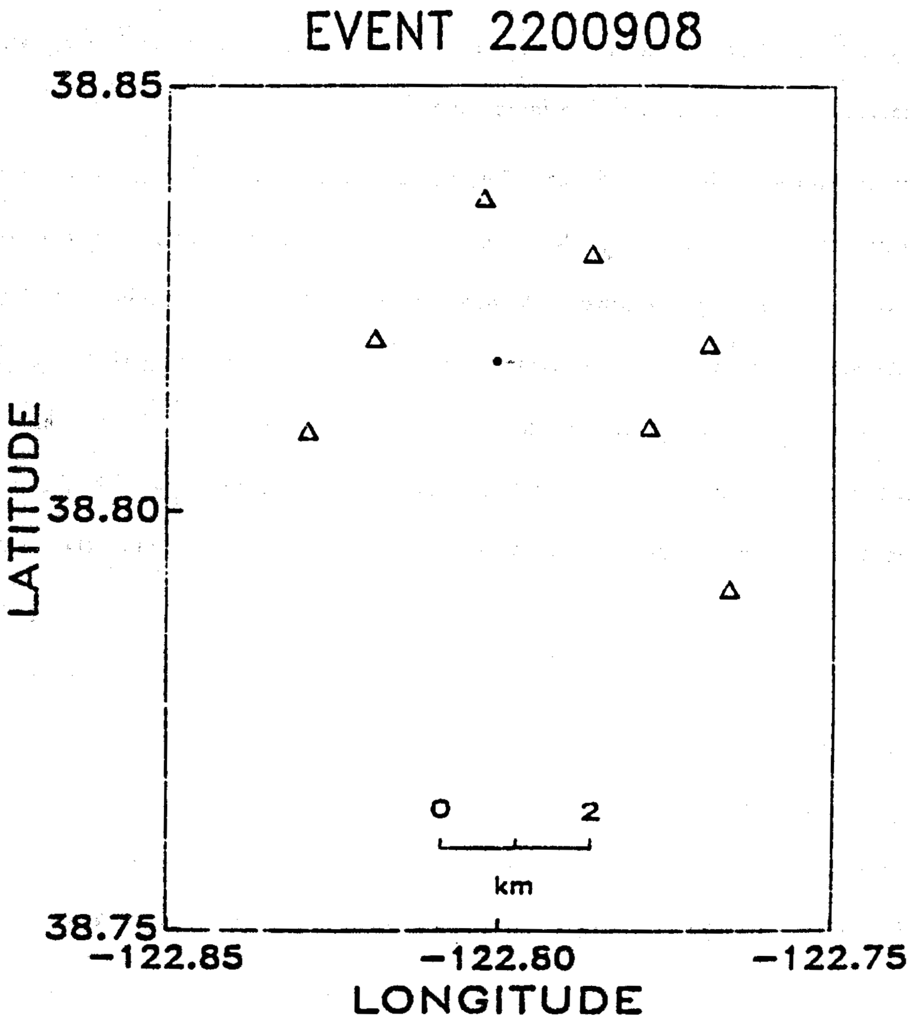


Figure 5.5 Map showing relation of earthquake 2200908 epicenter (small circle) to stations (Δ) used in moment tensor inversion. The hypocentral depth is 4.46 km below an elevation of 1.28 km above sea level, representing a ≈ 4.1 km depth below the recording station elevations. Note that the azimuthal coverage is good (≈ 90° azimuthal gap), but a smaller range of takeoff angles are represented due to the deeper event depth and the small range of epicenter-station distances.

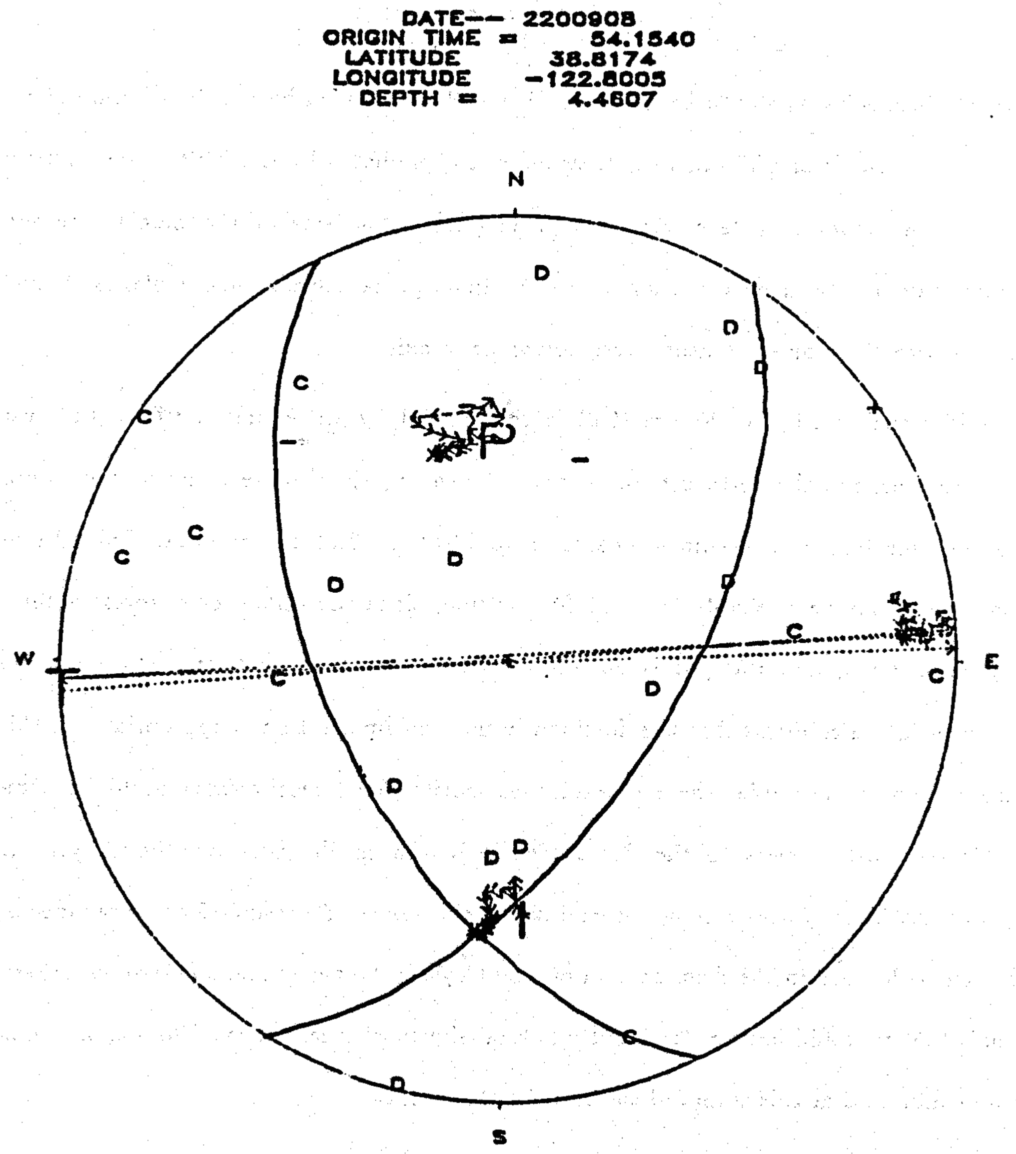


Figure 5.8 Lower-hemisphere equal-area-stereographic plot of P-wave first motions, time varying principal stress axes estimated from the moment tensor inversion, and double-couple fault plane solution for event 2181937. Compressional first motions are plotted as (C) for impulsive arrivals and (+) for emergent arrivals. Dilatational P-wave first motions are shown as (D) for impulsive arrivals and (-) for emergent arrivals. Convention for principal stress axes is the same as in Figure (5.2). Fine-dashed lines spanning the plot represent excursions required to plot stress axes points solely on the lower hemisphere. The solid lines are nodal planes drawn to satisfy the first motion data and the moment tensor estimates of principal stress orientations. The inconsistent compression in the upper dilatational quadrant corresponds to a distant station and its position is subject to errors as described in the text.

the data.

The fault plane solution shown in Figure (5.6) is well constrained by the distribution of P-wave first motions. The P and T principal stress directions predicted by the fault plane solution agree well with the moment tensor estimates (Table(5.1)). The size of the moment tensor estimated intermediate stress axis was small, so the fault plane solution nodal planes should coincide closely with the moment tensor intermediate stress axis.

The fault plane solution in Figure (5.6) is constrained by three first motions, the two west-most dilatations, and the northeast compression shown as (+). If these 3 first motions were unavailable, an almost pure strike-slip mechanisms would fit the first motion data. These three first motions would not be available for a USGS solution, since the dilatations represent temporary station readings, and the (+) is an ambiguous reading from a distant station. The moment tensor solution confirms that the dominantly normal-slip solution is appropriate for this event. This is another example where P-wave first motion focal mechanisms could be completely ambiguous with respect to the depth-dependent-focal-mechanism hypothesis, but the moment tensor inversion places strong constraints on the range of possible focal mechanisms. Since this event is located in the deep portion of The Geysers, Oppenheimer's (1986) hypothesis would predict that it would have a dominantly normal-slip focal mechanism. The moment tensor inversion confirms that this is indeed the case for this event.

5.4.4. Physical Source Component Decomposition

For event 2131638, the isotropic component (equation (5.3)) of the moment tensor was much smaller ($< 5\%$) than the deviatoric component (equation (5.4)) for the first 0.1 sec, which corresponds to the rise time of the dominant moment tensor element. After 0.1 sec, the isotropic component was as large as 25% of the deviatoric component. The size of the intermediate principal stress varied between 10% and 25% of the maximum principal stress. For event 2181937, the size of the isotropic component was about 25% to 30% of the size of the deviatoric component. The size of the intermediate principal stress varied between 10% to 15% of the

Principal Stress Orientations								
	First Motion				Moment Tensor			
	P		T		P		T	
Event	strike	dip	strike	dip	strike	dip	strike	dip
2131638	47°	2°	133°	14°	50°,62°	0°,10°	145°,150°	-10°,20°
2181937	59°	0°	149°	5°	53°,64°	-2°,3°	142°,154°	-4°,11°
	18°	72°	132°	11°				
2200908	355°	52°	92°	5°	338°,358°	40°,50°	80°,88°	-1°,14°

Table 5.1 Comparison of maximum compressional (P) and tensile (T) axes orientations estimated from first-motion fault plane solutions and moment tensor inversions. The range of time variations of orientations of the stress axes are listed for the moment tensor estimates. Event 2181937 has an ambiguous fault plane solution so two possible first motion solutions are listed.

maximum principal stress. For event 2200908, the size of the isotropic component was about 40% to 55% of the size of the deviatoric component. The size of the intermediate principal stress varied between 0% and 20% of the maximum principal stress.

The estimated principal stress orientations for these three Geysers earthquakes proved to be quite stable and provided results consistent with observed P-wave first motion distributions. The decompositions of the estimated moment tensors into isotropic and deviatoric components were not as satisfactory. None of the estimated moment tensors corresponded to a single double-couple source. The widely observed quadrapole radiation pattern observed for numerous earthquakes at The Geysers precludes a substantial isotropic component in the moment tensors of earthquakes there.

These results could be due to the rather unfavorable recording geometries used in all three moment tensor inversions. Event 2200908 had the moment tensor inversion with the largest condition number due to the small range of takeoff angles represented in the data. It is also the event with the largest isotropic component. The nonzero intermediate stress axis values are small enough that they could be explained by shear faulting on curved or bumpy fault surfaces (Backus and Mulcahy, 1976). Further investigations of moment tensor inversions with more advantageous station geometries should help clarify whether the non single-double-couple components of the moment tensor inversion estimates obtained are representative of true source complexity or are merely inversion artifacts.

5.4.5. Comparison of Observed and Predicted Seismograms

Excellent agreement between relative P and S-wave amplitudes were obtained for all inversions. In large part, this probably reflects the fact that good estimates of seismic velocity structure (obtained by progressive inversion in Chapter 4) were used. Efforts to invert for moment tensors during a feasibility study using data from a 4 station three-component network at The Geysers (conducted prior to collecting the data used here and in Chapter 4) failed to produce the correct amplitude pattern of P and S-wave phases on any components of ground

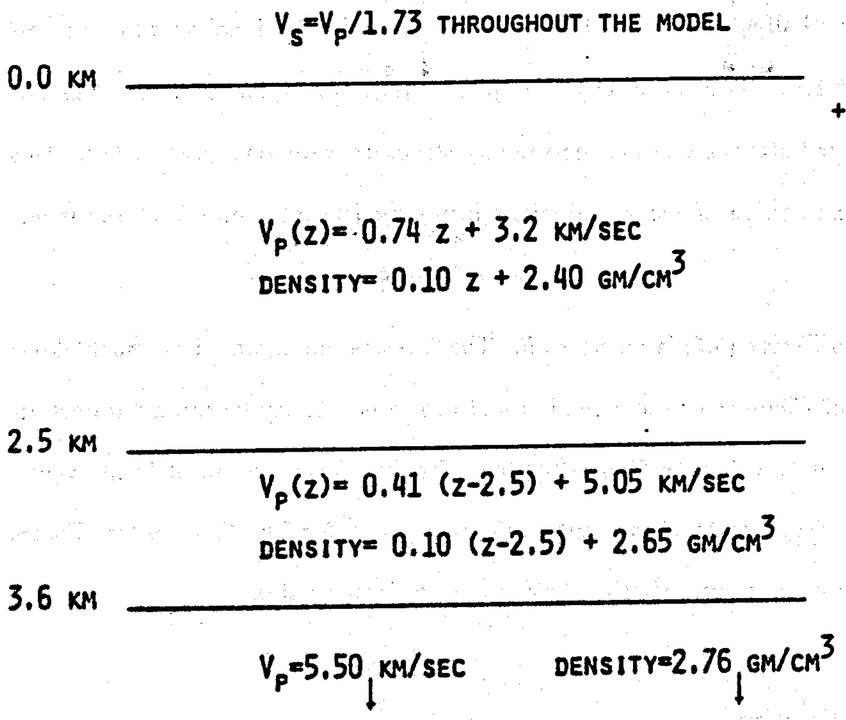


Figure 5.7 Velocity models used for a feasibility study of moment tensor inversions at The Geysers. Results of this feasibility study were used to design the experiment at The Geysers that ultimately provided the data for Chapter 4 and the moment tensor inversions done in this chapter.

motion. A course three-layer P-wave velocity model (Figure (5.7)), consisting of two linear gradient velocity layers over a half-space, was used and S-wave velocities were estimated by assuming a constant V_p/V_s of 1.73. It is clear from Chapter 4 that this velocity model was not correct. The results obtained after estimating the velocity structure were very satisfactory. This points out the critical importance of using realistic velocity models for moment tensor inversions.

The model shown in Figure (5.7) is incorrect for The Geysers but it would be much closer to the estimated models of Chapter 4 than a small set of constant velocity layers, a parameterization commonly employed to calculate Green functions for waveform studies of local events. The assumption made in Figure (5.7) of constant V_p/V_s was wrong for The Geysers (see Figure (4.11), but also reflects an assumption sometimes made in waveform studies.

5.5. Summary and Conclusions

Moment tensor inversions were successfully done for three earthquakes from The Geysers geothermal field. Estimated principal stress orientations were comparable to those estimated using P-wave first motions as constraints. In the case of one event, P-wave first motions could not constrain the focal mechanism to the degree that almost pure normal-slip and pure strike-slip focal mechanisms were consistent with the first-motion data. The estimated moment tensor principal stresses constrained the focal mechanism to be almost purely strike-slip. The moment tensor estimates of principal stress orientations were obtained using far fewer stations than required for first-motion focal mechanisms solutions. Seven stations were used for two inversions and 6 stations for the other. In contrast, an ambiguous first-motion focal mechanism was obtained for event 2181937, even though first motions were used from 17 stations.

The three focal mechanisms obtained here support Oppenheimer's (1986) hypothesis that focal mechanisms are a function of depth at The Geysers. Specifically, strike-slip focal mechanisms were obtained for the two shallow events and a predominantly normal-slip focal mechanism was obtained for the deep event, as his model predicts. The orientation of the minimum compressive stress axes for the shallow events (events 2131638 and 2181937 in Table (5.1)) is rotated 40° to 50° clockwise, with respect to his estimate of 105° as the azimuth of least compressive stress for The Geysers. Since only two events are available for comparison the differences may not be significant. The potential of moment tensor inversions to provide well constrained principal stress orientations for individual events may make it possible to map local variations of stress by using earthquakes distributed throughout The Geysers geothermal field.

The results obtained here were critically dependent on good estimates of velocity structure so as to minimize errors in calculated Green functions. The velocity model determined in Chapter 4, when used in the moment tensor inversions, produced the correct ratio of P and S-wave amplitudes on all components of ground motion, which strengthens the arguments in Chapter 4 that the estimated models are good one-dimensional representations of the velocity structure at The Geysers.

The good results obtained here are a direct consequence of using P and S-wave data together to estimate velocity structure, hypocenter locations, and moment tensors. Although moment tensor inversions were not done to compare the effects of using just P-wave data to those using both P and S-wave data, Stump and Johnson (1977) found that inversions that just used P-wave maximum amplitudes were not as well conditioned as inversions using complete seismograms. It was clear from the synthetic fits to the data, that the S-wave phases significantly constrained the moment tensor estimates.

A final note is that The Geysers represents a difficult area to do moment tensor inversions. Topographics variations are large, reflected in the 0.6 km variation of station elevations, and near surface velocity variations are profound. These factors were ignored in the moment tensor inversions, and good results were obtained despite rather unfavorable station recording geometries for all three events considered. This shows that, with the proper attention to determining realistic velocity structure and event locations, moment tensor estimates for microearthquake sources can be very robust.

References

Backus, G. and Mulcahy, M., 1976. Moment tensor and other phenomenological descriptions of seismic source -II. Discontinuous displacements, *Geophys. J. R. astr. Soc.*, **47**, 301-329.

Bufe, C. G., Marks, S. M., Lester, F. W., Ludwin, R. S., and Stickney, M. C., 1981. Seismicity of The Geysers-Clear Lake region, *U. S. Geol. Surv. Prof. Pap. 1141*, 129-137.

Fuchs, K., and Müller, G., 1971. Computation of synthetic seismograms with the reflectivity method and comparison of observations, *Geophys. J. R. astr. Soc.*, **23**, 417-433.

Geller, R. J., 1976. Body force equivalents for stress-drop seismic sources, *Bull. Seism. Soc. Am.*, **66**, 1801-1804.

Julian, B. R., 1986. Analysing seismic-source mechanisms by linear-programming methods, *Geophys. J. R. astr. Soc.*, **84**, 431-443.

Langston, C. A., 1981. Source inversion of seismic waveforms: the Konya India, earthquake of September 13, 1967, *Bull. Seism. Soc. Am.*, **71**, 1-24.

Langston, C. A., and Helmberger, D. V., 1975. A procedure for modeling shallow dislocation sources, *Geophys. J. R. astr. Soc.*, **42**, 117-130.

Majer, E. L., and McEvilly, T. V., 1979. Seismological investigations at The Geysers geothermal field, *Geophysics*, **44**, 246-269.

Oldenburg, D. W., 1982. Multichannel appraisal deconvolution, *Geophys. J. R. astr. Soc.*, **69**, 405-414.

Oppenheimer, D. H., 1986. Extensional tectonics at The Geysers geothermal area, California, *J. Geophys. Res.*, **91**, 11463-11476.

Prescott, W. H., and Yu, S. B., 1986. Geodetic measurement of horizontal deformation in the northern San Francisco Bay region, California, *J. Geophys. Res.*, **91**, 7475-7484.

Saikia, C. K., and Herrmann, R. B., 1986. Moment-tensor solutions for three 1982 Arkansas swarm earthquakes by waveform modeling, *Bull. Seism. Soc. Am.*, **76**, 709-723.

Sipkin, S. A., 1986. Interpretation of non-double-couple earthquake mechanisms derived from moment tensor inversion, *J. Geophys. Res.*, **91**, 531-547.

Stump, B. W. and L. R. Johnson, 1977. The determination of source properties by the linear inversion of seismograms, *Bull. Seism. Soc. Am.*, **67**, 1489-1502.

Stump, B. W., and Johnson, L. R., 1984. Near-field source characterization of contained nuclear explosions in tuff, *Bull. Seism. Soc. Am.*, **74**, 1-26.

Chapter 6

Summary and Conclusions

The method of progressive inversion has been modified to include S-wave data. By extending the method of progressive inversion to include S-wave information, improved estimates of earthquake locations and seismic velocity structure are obtained. The synthetic tests in Chapter 3 showed that the joint use of P and S-wave arrival time data has the following advantages over the use of P-wave data alone: (1) P-wave velocity and slowness gradient structure are more accurately estimated; (2) Hypocenter mislocation errors are substantially reduced, especially hypocentral depth; (3) Convergence of progressive inversions to local minima is more detectable using RMS data misfits of P and S-wave data; (4) Velocity model and hypocenter estimates are much more accurately determined when station corrections are used; (5) Complete elastic properties are estimated providing greater constraints for geologic interpretation of velocity structure.

The success of progressive inversion is dependent on the use of reasonable starting velocity models and synthetic tests showed that assuming unreasonably high surface velocity produces substantial errors in estimated velocity structure and earthquake locations. Adding S-wave data to progressive inversions does not completely eliminate hypocenter-velocity tradeoffs, but they are substantially reduced.

Progressive inversion was used to determine microearthquake locations and P and S-wave velocity structure at The Geysers geothermal field. The top of the steam reservoir is clearly defined by a large decrease of V_p/V_s at the condensation zone- production zone contact. The depth interval of maximum steam production coincides with minimum observed V_p/V_s, and V_p/V_s increases below the shallow primary production zone suggesting that reservoir rock becomes more fluid saturated. The correlation of V_p/V_s variations with the upper condensation zone suggests that determining V_p/V_s variations using P and S-wave VSP and crosshole surveys may delineate hydrothermal convection cell boundaries.

Most earthquakes appear to occur in two depth intervals. The majority of events occur in the depth interval (0.5-2.5 km elevation) of primary steam production. A gap in seismicity between the primary production zone and an elevation of ≈ 3.0 km is evident here and in the results of Eberhart-Phillips and Oppenheimer (1984). The second largest concentration of seismicity is found in an areally restricted region at elevations >3.0 km. Both regions of seismicity appear to be associated with zones of fluid depletion. Seismicity could be induced by volume contraction due to fluid withdrawal or conversion of aseismic slip to stick-slip due to deposition of silica at fractures or a combination of these mechanisms. The apparent confinement of earthquakes to regions of fluid withdrawal favor the increase of friction mechanisms of Allis (1982) over the volume contraction mechanism of Majer and McEvilly (1979). The relatively short time period of the present data set precludes definite conclusions concerning long term patterns of seismicity. A much larger data set needs to be investigated to determine if longer term seismicity is confined to zones of steam production. The data set used by Eberhart-Phillips and Oppenheimer (1984) is suited for this purpose. However, proper accounting for differences in station elevations and realistic near-surface velocities must be used to obtain reliable estimates of earthquake locations.

One-dimensional P and S-wave velocity models with station corrections satisfactorily fit arrival time data to within a standard error of 0.02 sec. This supports Eberhart-Phillips' (1986) conclusion that a one-dimensional P-wave model with station corrections is adequate to locate earthquakes in the primary production zone at The Geysers and also indicates that a one-dimensional S-wave velocity model with station corrections is adequate as well.

Moment tensor estimates were obtained for three earthquakes from The Geysers geothermal field. Estimated principal stress directions were comparable to those estimated using P-wave first motions as constraints. The moment tensor estimates were consistent with the P-wave first motion data and constrained the focal mechanism of an event where the P-wave first motion data could not. The moment tensor estimates of principal stress orientations were obtained using far fewer station than required for first-motion focal mechanism solutions.

The three focal mechanisms obtained here support Oppenheimer's (1986) hypothesis that focal mechanisms are a function of depth at The Geysers. Focal mechanisms for the shallow earthquakes were almost pure strike-slip and the focal mechanism for the deep earthquake was predominantly normal-slip, as predicted by his model. The potential of moment tensor inversions to provide well constrained principal stress orientations for individual events may make it possible to map local variations of stress by using earthquakes distributed in different portions of The Geysers geothermal field.

The satisfactory results of moment tensor inversions using microearthquake data from The Geysers was critically dependent on the good estimates of velocity structure and earthquake locations provided by the progressive inversion. This ensured that Green function errors would be minimized. The synthetic seismograms produced using the estimated moment tensors and Green functions matched the observed data well, particularly the ratio of P and S-wave amplitudes on all components of ground motion. This strengthens the argument that the estimated velocity models are good one-dimensional representations of the true velocity structure at The Geysers.

The Geysers represents a difficult area to do a progressive inversion and moment tensor inversions. Topographic variations are large and near surface velocity variations are profound. By properly accounting for differences in station elevations and using realistic near surface velocities combined with station corrections, progressive inversion results clearly delineated significant features at The Geysers. These factors were ignored in the moment tensor inversions and good results were obtained despite rather unfavorable station recording geometries for all three earthquake considered. This shows that moment tensor estimates for microearthquake sources can be very robust.

Progressive inversion as developed here and the moment tensor inversion method of Stump and Johnson (1977) provide a complete approach to determine earthquake locations, P and S-wave velocity structure, and earthquake source mechanisms. The entire process is internally consistent; earthquake locations and velocity structure are estimated using progressive inversion without making any assumptions about source properties. Results of progressive inversion are then used to calculate Green functions for moment tensor inversions. Because progressive inversion was extended to include S-wave data, complete three-component data can be used in moment tensor inversions. Since the joint use of P and S-wave data minimizes earthquake location errors and velocity structure errors, the best estimate of Green functions are obtained and moment tensor inversion errors are minimized.

The importance of using S-wave information cannot be over- emphasized. Not only does including S-wave data reduce errors in progressive inversion, but it facilitates reliable moment tensor inversions when using data from a relatively small number of three-component recording stations.

The one-dimensional velocity model assumptions used throughout this dissertation is not an inherent limitation of the development. The method of Thurber (1983) for three-dimensional velocity-hypocenter inversion can be simply extended to include S-wave data. In contrast to the one-dimensional case, Green function calculations for three-dimensional velocity structures must be approximate or very time consuming if finite difference approaches are used. Thus, while moment tensor inversions can be done in cases of three-dimensional velocity variations, the most profound problems are accurate determination of three-dimensional velocity variations and accurate calculation of Green functions in three-dimensional inhomogeneous media. The satisfactory moment tensor inversion results obtained at The Geysers using one-dimensional assumptions indicates that three-dimensional velocity modeling and Green function calculations will probably not be required in many regions that the approach of progressive inversion and moment tensor inversion, outlined here, may potentially be applied.

Appendix A

Travel Time Calculations

Hypocenter location and velocity inversion require determining travel times for first P and S-wave arrivals as a function of distance between sources and receivers. Accurate calculated travel times are required for reliable hypocenter location and velocity inversion. Errors in calculated travel times must be small in relation to errors in observed arrival times. Travel time calculations must be computationally efficient since progressive hypocenter-velocity inversion involves many travel time evaluations. Buland and Chapman (1983) developed a method of travel time determination based on interpolation of $\tau - p$ curves that combines accuracy with computational efficiency. Their method eliminates the computational burden involved with shooting methods of ray tracing and has reduced storage requirements relative to table lookup approaches. In addition, $\tau - p$ interpolation provides the flexibility to satisfy special requirements concerning elevation corrections. Their approach forms the basis for the τ–p interpolation method outlined here.

The basic problem is: given source and receiver distances we need to calculate travel times for first P and S-wave phases and their associated wavepaths. We start by briefly outlining the relations between travel time, distance, ray parameter, tau, and the theta function (Buland and Chapman, 1983). Distances considered here are small, less than 100 km, so the Earth's sphericity is ignored. Let x define horizontal distance and z define depth (positive downward). Velocity, $v(z)$, varies as a function of depth only. Define the ray parameter or horizontal ray slowness

$$p = \frac{\sin(i(z))}{v(z)} = \sin(i(z))u(z) \tag{A.1}$$

where $u(z)$ is medium slowness, $i(z)$ is the angle the direction of ray propagation makes with the vertical, and vertical ray slowness is given by

$$q(p,z) = (u^2(z) - p^2)^{1/2} \tag{A.2}$$

Travel time is given by

$$T(p) = \oint \frac{u^2(z)dz}{q(p,z)} \tag{A.3}$$

where $\oint$ indicates the integral over the complete ray path. All ray segment contributions are taken as positive. The range integral is given by

$$X(p) = \oint \frac{pdz}{q(p,z)} \tag{A.4}$$

The delay-time function, $\tau(p)$, is related to $T(X)$ and is given by

$$\tau(p) = T(p) - pX(p) \tag{A.5}$$

and

$$p = \frac{dT}{dX}$$

$$X = -\frac{d\tau}{dp} \tag{A.6}$$

Define the theta function as

$$\theta(p,x) = \tau(p) + px \tag{A.7}$$

or substituting (A.5)

$$\theta(p,x) = T(p) + p(x - X(p)) \tag{A.8}$$

where x is a distance of interest.

Stationary points of the theta function correspond to geometrical wave group arrivals:

$$\frac{\partial\theta(p,x)}{\partial p} = \frac{d\tau(p)}{dp} + x = x - X(p) \tag{A.9}$$

so

$$\left.\frac{\partial\theta(p,x)}{\partial p}\right|_{p_0} = 0 \tag{A.10}$$

when

$$x = X(p_0) \tag{A.11}$$

From (A.8), the value of the theta function at the stationary point is just the travel time

$$T(p_0) = \theta(p_0, X(p_0)) \tag{A.12}$$

Buland and Chapman (1983) showed that given $\tau(p)$ and a distance x at which a travel time of a suite of phases is desired, it is sufficient to examine the stationary points of the corresponding theta function. Neither $X(p)$ nor $T(p)$ are required. Caustics produce multiple stationary points allowing all geometrical arrivals to be determined. The ray parameter corresponding to each arrival is determined as a byproduct of the travel time calculation through (A.10). The estimated ray parameter is needed to calculate the ray-path kernels required for velocity inversion and to determine take-off angles at the source.

In the absence of low velocity zones (LVZ), only downgoing transmitted and refracted-wave tau branches and upgoing direct-wave tau branches are required to determine first-arrival travel times for any distance. When a velocity model contains a LVZ, a transmitted wave shadow occurs over a range of distances. If only transmitted and refracted path tau branches are used, no geometrical arrivals will be found in the shadow, and observed data in the LVZ shadow zone must be discarded. A solution to this problem is to include reflected-wave tau branches corresponding to reflections from velocity model discontinuities in and below the LVZ. Geometrical reflected phases from these tau branches provide travel times for stations in the shadow zone.

It is difficult to determine the magnitude of velocity model discontinuities required to produce observable reflected phases without calculating synthetic seismograms appropriate for the problem. Since this was not feasible, the following approach was used. Reflected tau branches were constructed for all velocity or velocity gradient discontinuities in or immediately below the bottom of each LVZ. This guaranteed that reflected geometrical phases fill the transmitted and refracted path shadow zone. It is likely that phases from some of the reflected tau branches used would not be observable, but the approach was used for the following reasons. Firstly, we were concerned that if data were discarded from LVZ shadows, the removal of the data misfit associated with the deleted data may result in erroneous convergence of the velocity inversion. This might occur if the removed data had contributed significantly to the total data misfit. Secondly, the fact that waves are observed in apparent shadow zones implies that these arrivals correspond to reflected, diffracted, or scattered waves. We simply choose to model them as reflected waves.

A possible side effect of this approach is that estimated LVZs from a velocity inversion could have smaller depth extent and larger velocity contrast than true LVZs. This would occur when travel times for data constraining the presence of a LVZ were calculated from an erroneous reflected branch originating just below the LVZ lid, when the observed arrivals actually corresponded to a much deeper reflected phase.

It may appear that we are implying that we can solve for the velocity structure within a LVZ. This is not the case. Gerver and Markushevitch (1966) showed that only the maximum thickness of the LVZ can be constrained from travel time data. Our approach will tend to estimate that a LVZ has a thickness less than the maximum bound provided by the travel time data. With these caveats it was decided that using reflected tau branches was preferable to the elimination of perfectly good data. If a LVZ were found in a velocity inversion, a more rigorous criteria could be developed to decide what (if any) reflected branches should be included.

Given the aforementioned choices of tau branches, we must find an appropriate ray parameter sampling to construct their discrete representation. Buland and Chapman (1983) suggest a quadratic ray parameter spacing

$$p_j = p_{\mathrm{end}} - j^2 \cdot \delta p \qquad j = 1,2,\ldots,k. \tag{A.13}$$

where p_{end} and p_k are critical points. We found that (A.13) works well to produce $\tau(p_j)$ points approximately equally spaced in range for downgoing branches. A cubic ray parameter spacing

$$p_j = p_{\mathrm{end}} - j^3 \cdot \delta p \qquad j = 1,2,\ldots,k. \tag{A.14}$$

was found to produce more evenly spaced $\tau(p_j)$ in range for upgoing branches from the shallow source depths (0-5 km) considered here.

In the neighborhood of caustics δp was fixed at a small value to ensure monotonic variation of range within discrete tau intervals (between $\tau(p_j)$ and $\tau(p_{j+1})$). For very shallow sources δp for upgoing branches was made a function of source depth to ensure proper distance spacing. Ray parameter spacing for a velocity model need be determined only once for downgoing branches between critical points. However, when a source does not reside at a depth corresponding to a critical point, optimal ray parameter spacing is determined anew for each new source depth between the source critical ray parameter and the next critical point below it. The same approach is used to determine optimal ray parameter spacing for the upgoing direct wave branch at a new source depth. Thus, only incremental layer $\tau(p)$ and $X(p)$ values (to be discussed below) for a range of ray parameters corresponding to all downgoing paths of interest are stored in a table. For upgoing tau branches, the optimal ray parameter spacing, $\tau(p)$, and $X(p)$ are computed specifically for each new source depth.

The velocity models were discretized using plane layers consisting of linear velocity gradients. Travel time, range, and tau contributions for the ith layer with ray parameter p_j are given by

$$\Delta T_{ij} = \frac{\Delta z_i}{\Delta v_i} \ln \left[\frac{u_i + (u_i^2 - p_j^2)^{1/2}}{u_{i+1} + (u_{i+1}^2 - p_j^2)^{1/2}} \right] \tag{A.15}$$

$$\Delta X_{ij} = \frac{\Delta z_i}{\Delta v_i} \left[\frac{(u_i^2 - p_j^2)^{1/2}}{u_i p_j} - \frac{(u_{i+1}^2 - p_j^2)^{1/2}}{u_{i+1} p_j} \right] \tag{A.16}$$

$$\Delta \tau_{ij} = \Delta T_{ij} - p_j \Delta X_{ij} \tag{A.17}$$

where the layer is bounded by the depths z_i and z_{i+1} with corresponding slownesses $u_i = u(z_i) \geq u(z_{i+1}) = u_{i+1}$, $\Delta z_i = | z_i - z_{i+1} |$, and $\Delta v_i = u_i^{-1} - u_{i+1}^{-1}$. Tau branches are calculated by summing appropriate incremental layer contributions. The incremental ray path distances, $G(z_i)$ (2.12), needed for the slowness inversion are approximated using arcs of circles produced by linear velocity gradients given by

$$G(z_i) \approx \frac{\Delta v_i}{\Delta z_i \, p_j} \left[\sin^{-1}(p_j \, v_{i+1}) - \sin^{-1}(p_j \, v_i) \right] \tag{A.18}$$

The velocity models used here are piecewise continuous; velocity is continuous, velocity gradient may be discontinuous at layer boundaries. Consequently, head-wave tau branches are not included in this parameterization. Initial discretization was chosen to satisfy requirements for slowness inversions; normally 0.1 km layer spacing was used. For interpolation of tau branches, only velocity-depth points associated with velocity gradient discontinuities were retained to define the velocity models critical ray parameters.

Now, a means to interpolate the theta function must be specified. Actually, the theta function never has to be formed; only $\tau(p)$ need be interpolated to find stationary points of the theta function (Buland and Chapman, 1983). They introduce piecewise τ splines to interpolate $\tau(p_j)$ of the form

$$\tau(p) = a_i + b_i(p_{pend} - p) + c_i(p_{pend} - p)^2 + d_i(p_{pend} - p)^{3/2}, \quad p_i \le p \le p_{i+1} \tag{A.19}$$

This interpolant is preferable to piecewise cubic splines because it incorporates the square root singularity in the derivative of range, $X'(p)$, at the largest ray parameter, p_{pend}, of each tau segment. The range (A.6) is given by

$$X(p) = b_i + 2c_i(p_{pend} - p) + 3/2 d_i (p_{pend} - p)^{1/2}, \quad p_i \le p \le p_{i+1} \tag{A.20}$$

Theta will have an extremum when

$$2c_i(p_{pend} - p) + 3/2 d_i (p_{pend} - p)^{1/2} + (b_i - x) = 0 \tag{A.21}$$

which is quadratic in $(p_{pend} - p)^{1/2}$.

Either H splines (Hermite splines fit both the function and its first derivative exactly at each discrete point) or B splines (smooth splines fit the function exactly at each discrete point, the first derivative exactly at the first and last discrete point, and have continuous first and second derivatives everywhere) could be used. Buland and Chapman (1983) choose B splines because they perform better in the neighborhood of broad caustics such as point B of PKP. Here, we are primarily interested in travel times for first arrivals, so evaluation of $\tau(p_i)$ near caustics does not present a significant problem. The potential problem of H splines producing numerical caustics over several ray parameter intervals was mitigated by using high precision for ray parameter and small ray parameter intervals near the caustic. This ensured that range was monotonic in each ray parameter interval, guaranteeing a single contribution for each interval, except in an interval containing an actual caustic.

H spline interpolation of (A.19) is much faster than for the B spline analog; H splines are local and solving for the coefficients is simple and fast. Only the spline coefficients for range intervals corresponding to distances of interest need be calculated for the H spline approach. B spline interpolation requires solving a system of equations for all the coefficients, although Buland and Chapman (1983) present an approach using basis functions that reduce the computation burden somewhat. Twice as much information must be stored when using H splines, $\tau(p_j)$ and $X(p_j)$, than when using B splines where just $\tau(p_j)$, $X(p_{pend})$, and $X(p_k)$ are needed. However, only a small number of tau branches need be retained to determine first arrival travel times so the addition storage requirements are not excessive. Computational speed was deemed more important than keeping storage requirements to an absolute minimum. Computational speed is the limiting factor for doing hypocenter-velocity inversions with large amounts of data given the modest storage requires of the progressive inversion algorithm of Chapter 2.

We performed extensive tests to determine optimal δp spacing for tau branch interpolation and to investigate the magnitude of errors in estimates of travel time and p. We were unable to uncover any problems associated with using H splines in the neighborhood of caustics. This could be because we used a much finer ray parameter spacing than Buland and Chapman (1983) or the fact that the geometry used here (small distance and plane layers) is less likely to produce problems at caustics than for the earth flattened models they used.

Due to the stationarity of the theta function, estimated travel times are more accurate than tau or distance. Consequently, estimated ray parameters used to determine ray paths produce estimates of total range that have relatively larger errors than errors in corresponding travel time estimates. The absolute range estimate errors were small, ranging from .1 to 10 meters. Thus, when we used the ray parameter estimates to trace rays to check the travel time estimates we found that we were required to use a shooting approach to vary the interpolated ray parameter estimate to produce the correct range. As the range estimate converged to the correct distance the ray traced travel time tended to converge to the interpolation estimate travel time. Consequently, we used the ray traced travel time, estimated using the interpolation estimate of ray parameter, as an upper bound on the interpolation estimating travel time error. It is ironic that the interpolation estimates of travel time are more accurate than corresponding ray traced estimates, but follows naturally from the stationarity of the theta function. We found that using a δp spacing that produced range intervals of 2-3 km (except near caustics where much smaller δp s were used), produced travel time errors no larger than 1 millisecond. Integration estimates of range using interpolation ray parameter estimates had errors no larger than 10 meters.

The H spline approach afforded us an addition degree of flexibility that was essential to accurately and efficiently determine travel times to recording stations at widely differing elevations. This was done in the following way. The elevation at the top of the velocity models was defined by the elevation of the lowest station used. When a station was at a higher elevation, an additional layer was added between the top of the velocity model and that station. The layer thickness was defined by the elevation difference between the station and the elevation of the top of the model. The layer was assumed to have a constant velocity defined by the velocity at the top of the velocity model. Incremental $\tau(p_j)$ and $X(p_j)$ contributions of this extra layer were added to all tau curves. Then, only the spline coefficients of range intervals containing the station distance are calculated and these intervals used to determine travel times to the station. A B spline interpolation would require solving for spline coefficients for all range intervals of all pertinent tau branchs. It is clear that the local nature of H splines is crucial to the efficient implementation of this method. Layer thickness is adjusted to appropriate values for subsequent stations by adding or removing appropriate incremental $\tau(p_j)$ and $X(p_j)$ contributions.

This approach allows accurate determination of travel times in regions with large variations of station elevations by determining the geometrical paths to stations at differing elevations instead of using vertical path approximations. If something is known about velocities under each station, then determined velocities could be used to define the velocities in the elevation correction layer beneath each station.

Further computational savings are realized if certain tau branches can be excluded from consideration. If all recording station are located at distances less than the critical distance for upgoing rays, only the upgoing tau branch and spline coefficients need be calculated to determine first arrivals. If all recording stations are beyond the critical distance for upgoing rays, only downgoing tau branch contributions must be considered.

To provide a rough estimate of the time required to calculate travel times with this approach the follow example is provided. Velocity models consisted of 9 layers for P-waves and 8 layers for S-waves. For a source at fixed depth, it took 50 CPU seconds to compute 2500 travels times using a Digital VMS/VAX 780 computer.

References

Buland, R., and Chapman, C. H., 1983. The computation of seismic travel times, *Bull. Seism. Soc. Am.*, **73**, 1271-1302.

Gerver, M. L., and Markushevitch, V., 1966. Determination of a seismic wave velocity from the travel time curve. *Geophys. J. R. astr. Soc.*, **11**, 165-173.

Appendix B

Nonlinear Earthquake Error Appraisal

B.1. Introduction

Earthquake location is in general a nonlinear problem. One goal of any earthquake location scheme is to determine the maximum likelihood location; the spatial coordinates with the greatest probability of corresponding to the actual position of the hypocenter. A second goal is to determine the significance of the maximum likelihood location relative to other prospective spatial locations. The second goal arises because determining an earthquake's location is a statistical problem. All information used to constrain an earthquake's location contain errors. We know something about the nature of errors in seismic arrival times, station locations, velocity structure, and travel time calculations. By specifying the statistical properties of input data errors, their effect on estimates of an earthquake's location can be calculated.

The most common approach used to locate earthquakes is Geiger's method (see Chapter 2). While Geiger's method can successfully determine an earthquake's maximum likelihood location, estimates of effects of input data errors on location solution statistics are not as reliable. Geiger's method uses only one term in a Taylor series expansion to estimate hypocentral position. Consequently, effects of nonlinearities in the problem are completely ignored. Thurber (1985) improves on Geiger's method by including the next higher order term in the Taylor series, the Hessian (Newton's method) when solving for hypocenter position and hypocenter uncertainty. Inclusion of the Hessian makes it possible to approximately incorporate the effects of intrinsic nonlinearity. Tarantola and Valette (1982) present an alternative to Taylor series expansion approaches to earthquake location.

Tarantola and Valette (1982) used methods of probability calculus to derive a general nonlinear earthquake location method devoid of linearizations. The result of their approach (called

PDFLOC here) is a probability density function, p.d.f., for the spatial position of an earthquake hypocenter. PDFLOC satisfies both goals stated at the outset; maximum likelihood position is determined and the shape of the p.d.f. provides all information about the importance of the maximum likelihood position relative to other possible candidate hypocenter positions. Results provided by PDFLOC are ideal for doing statistical hypothesis testing of absolute or relative positions of different events. For instance, an apparent decrease in focal depth with time of seismicity may be associated with rising magma in a volcanic region. Using p.d.f.s for individual events determined by PDFLOC the statistical significance of focal depth changes can be evaluated.

In order to use PDFLOC, the nature of errors in the input information to the problem must be determined. We begin with a discussion of errors in input information. Then the development of PDFLOC is presented. The linearized method presented in chapter 2 is currently the only other method that systematically accounts for velocity model errors in hypocenter error estimation.[1] Since hypocenter location is inherently a nonlinear problem, estimates of hypocenter errors using linearized approaches are only approximate. We would like to determine how reliable linearized hypocenter error estimates are. In particular, we want to determine the reliability of hypocenter error estimates obtain using the method of progressive inversion in Chapter 2. Using synthetic data, PDFLOC is used to investigate the significance of nonlinearities with respect to standard error levels in the input information and earthquake-station recording geometries.

B.2. Sources and Classification of Errors for Earthquake Location

Errors in event location are attributed to two general sources: data errors; and model errors.[2] Data errors consist of arrival time errors, clock errors, and station location errors. Data errors are confined to the observed information that constrain the problem. Model errors consist of errors in calculated travel times due to velocity model errors and numerical limitations of digital computations. Model errors produces errors in the forward problem, calculating theoretical travel times. The distinction between the two classes of errors is useful when considering sources of errors in the earthquake location problem. In Chapter 2 a similar classification of errors was used. There, data errors were referred to as statistical errors following the convention of Pavlis (1982). Since all classes of errors have statistical properties the term, data error, is used here in place of the term statistical error. Data errors can be estimated for arrival times, clock errors, and station location errors.

Proper appraisal of hypocentral estimates requires considering the effects of model errors in calculated travel times. Errors in calculated travel times arise primarily from ignorance of the true velocity structure. All methods of earthquake location require travel times calculated from earth models. This includes the distance residual approach of Lomnitz (1977). Consequently, we need a means to estimate the component of errors in calculated travel times due to imprecise knowledge of velocity structure. This in turns requires an estimate of the errors in assumed velocity structure.

In the approach developed in Section (2.6.2), it was seen from equation (2.61), that model errors contribute to the total errors in the residuals used to estimate earthquake locations in a linearized process. A Gaussian error distribution was assumed for all errors in Chapter 2 and all error estimates there have a Gaussian distribution.

Tarantola and Valette (1982). derived expressions for estimating hypocenter p.d.f.'s by explicitly assuming a Gaussian distribution of data errors and model errors. Since the same Gaussian assumptions are made in Chapter 2, results of nonlinear tests can be directly compared with the linearized predictions of Chapter 2.

B.3. Hypocenter Location from Probability Calculus, (PDFLOC)

The following derivation is from Tarantola and Valette (1982). Let there be n parameters $\mathbf{p}$. The a posterior p.d.f. for the parameters, $\nu_{\mathrm{p}}(\mathbf{p})$, can be calculated from the expression

$$\nu_{\mathrm{p}}(\mathbf{p}) = \rho_{\mathrm{p}}(\mathbf{p})\cdot\int \frac{\rho_{\mathrm{d}}(\mathbf{d})\cdot\theta(\mathbf{d}\mid\mathbf{p})}{\mu_{\mathrm{d}}(\mathbf{d})}\,d\,\mathbf{d} \tag{B.1}$$

where

$\rho_p(\mathbf{p})$ = a priori p.d.f. for the parameters.

$\rho_d(\mathbf{d})$ = a priori p.d.f. for the data.

$\theta(\mathbf{d} \mid \mathbf{p})$ = conditional p.d.f. for d given p

$\mu_d(\mathbf{d})$ = null information d.f. for the parameters.

All information about the parameters is contained in $\nu_p(\mathbf{p})$. If $\nu_p(\mathbf{p})$ is unimodal the maximum of $\nu_p(\mathbf{p})$ corresponds to the maximum likelihood solution for the parameters. In most applications, we are interested in information about particular parameters. All information about a single parameter, p_1 is contained in its marginal d.f. defined as

$$\nu_1(p_1) = \int \nu_p(p)\cdot dp_2 \cdot dp_3 \cdot \cdots dp_n . \tag{B.2}$$

Hypocenter position is specified by X, Y, Z corresponding to longitude, latitude, and depth, respectively, and T corresponds to origin time. Let:

$$\mathbf{t} = \mathbf{f}(X,Y,Z,T) \tag{B.3}$$

be the inexact theoretical relationship between arrival times and the spatio-temporal coordinates of the focus, which depends on assumptions about wave propagation theory and the velocity model. Let $\mathbf{C_T}$ be the covariance matrix for model errors. Then the theoretical relationship between data and parameters is

$$\theta(\mathbf{t} \mid X,Y,Z,T) = \exp\{-1/2[\mathbf{t}-\mathbf{f}(X,Y,Z,T)]^{\mathrm{T}}\cdot \mathbf{C_T}^{-1}\cdot[\mathbf{t}-\mathbf{f}(X,Y,Z,T)]\} \tag{B.4}$$

Let the observed arrival time data have mean $\mathbf{t}_0$ and covariance matrix $\mathbf{C}_t$. Then the a priori d.f. for the data is

$$\rho(\mathbf{t}) = \exp\{-1/2(\mathbf{t}-\mathbf{t}_0)^{\mathrm{T}}\cdot \mathbf{C}_t^{-1}\cdot(\mathbf{t}-\mathbf{t}_0)\} \tag{B.5}$$

Since all data and parameters are specified in Cartesian coordinates the null information function is constant and need not be considered further.

The a posteriori density function for the parameters obtained by analytical intergration is

$$\nu(X,Y,Z,T) = \rho(X,Y,Z,T)\cdot\exp\{-1/2[\mathbf{t}_0-\mathbf{f}(X,Y,Z,T)]^T\cdot(\mathbf{C}_t+\mathbf{C}_T)^{-1}\cdot[\mathbf{t}_0-\mathbf{f}(X,Y,Z,T)]\} \quad \text{(B.6)}$$

The spatial location of an earthquake can be obtained by determining the marginal density function

$$\nu(X,Y,Z) = \int_{-\infty}^{+\infty} \nu(X,Y,Z,T)\,dT \quad \text{(B.7)}$$

by integration of origin time T. In general, we do not have a priori information about origin time, the a priori d.f. for origin time can be assumed constant. The a priori d.f. for the spatial position of the event remains the same:

$$\rho(X,Y,Z,T) = \rho(T)\cdot\rho(X,Y,Z) = \rho(X,Y,Z) \quad \text{(B.8)}$$

The computed arrival time at a station i, $f_i(X,Y,Z,T)$ can be written:

$$f_i(X,Y,Z,T) = h_i(X,Y,Z) + T \quad \text{(B.9)}$$

where h_i is the travel time between the point (X,Y,Z) and the station i.

The a posteriori d.f. for spatial location is

$$\nu(X,Y,Z) = K\cdot\rho(X,Y,Z)\cdot\exp\{-1/2[\tilde{\mathbf{t}}^0-\tilde{\mathbf{h}}(X,Y,Z)]^{\mathrm{T}}\cdot\mathbf{P}\cdot[\tilde{\mathbf{t}}^0-\tilde{\mathbf{h}}(X,Y,Z)]\} \quad \text{(B.10)}$$

where

$$\mathbf{P} = (\mathbf{C}_t + \mathbf{C}_T)^{-1} \quad \text{(B.11)}$$

is a "weight matrix"

$$p_i = \sum_j P_{ij} \quad \text{(B.12)}$$

are "weights", and

$$K = \sum_i p_i = \sum_{ij} P_{ij}. \quad \text{(B.13)}$$

The quantity, $\tilde{t}_j^{\,0}$, is the observed arrival time minus the weighted mean of observed arrival times,

$$\tilde{t}_i{}^0 = t_i{}^0 - \frac{\sum_j p_j \, t_j{}^0}{\sum_j p_j} \tag{B.14}$$

and $\tilde{h}_i$ is the computed travel time minus the weighted mean of computed travel times

$$\tilde{h}_i = h_i - \frac{\sum_j p_j \, h_j}{\sum_j p_j} \tag{B.15}$$

It should be noted that C_T may depend on (X,Y,Z) and therefore P_{ij}, p_i, and K will also.

The marginal density function for epicenter is

$$\nu(X,Y) = \int_0^{z\,\max} \nu(X,Y,Z)\,dZ \tag{B.16}$$

and the marginal density function for depth is

$$\nu(Z) = \int_{-\infty}^{+\infty} dX \int_{-\infty}^{+\infty} dY \, \nu(X,Y,Z) \tag{B.17}$$

From the marginal density functions the maximum likelihood point(s) are determined as well as the form of the likelihood function.

B.4. Synthetic Tests with PDFLOC

Equation (B.6) shows that the total error is just the sum of the data errors and model errors and that each error type makes contributions to the overall error in the same manner. Rather than investigate how varying proportions of data error and model error contributions to total hypocenter uncertainty, it is sufficient to investigate how total input error magnitudes influences hypocenter marginal density functions. In fact, the distinction between data error and model error is essentially meaningless in this context because only the total error is used to estimate marginal density functions. This fact points out why reducing picking uncertainties of arrival time phases may not reduce uncertainties in hypocenters. If calculated travel time errors due to incorrect velocity models are larger than picking errors, the reduction in picking error uncertainties does not further reduce hypocenter uncertainties. The method of progressive inversion developed in Chapter 2 allows maximum utilization of data picking accuracy by improving velocity models as earthquakes are located. In this way, the minimum total error is obtained, and hypocenter uncertainties are minimized.

Two cases were investigated. Case (1) was chosen to reflect the situation where an earthquake is located on the edge of a recording network and is shown in Figure (B.1). Case (2) was chosen to correspond to an earthquake located inside a recording network and is shown in Figure (B.1). Event depths were 2.0 km in both cases. All station elevations correspond to zero model depth to simplify calculations. Since relative variations of uncertainty estimates are the goal, accounting for station elevation differences is not critical.

Two standard error sizes were investigated. A standard error of 0.02 sec was used in one set of tests because it corresponds to the standard error reference for the data in Chapter 4. The second standard size considered was 0.05 sec. This represent a conservative estimate of the total error due to data errors (0.02 sec) and model errors.

The synthetic P and S-wave velocity models of Chapter 3 were used to calculate synthetic travel times and when estimating $\nu(X,Y,Z)$. The synthetic data were error free. Investigating effects of using incorrect velocity models and data are outside the scope of inquiry. We only want to investigate the effects of nonlinearity on hypocenter uncertainty estimates. Consequently, the maximum likelihood points estimates for all cases correspond to the true hypocenter position.

Nonlinear effects manifest themselves as asymmetries in the marginal density functions for epicenter and depth and in the nonlinear scaling of different input errors sizes on confidence bounds. When asymmetries are not particularly large, the linearized error estimates in Chapter 4 can be safely assumed to be representative of true uncertainties. Hypocentral depth is used to quantify the results because it is most effected by nonlinearities.

The a posteriori d.f. in equation (B.10) is calculated on a three-dimensional grid and then numerically integrated using equations (B.16) and (B.17) to obtain marginal density functions for epicenter and depth, respectively. Once the marginal density function has been computed

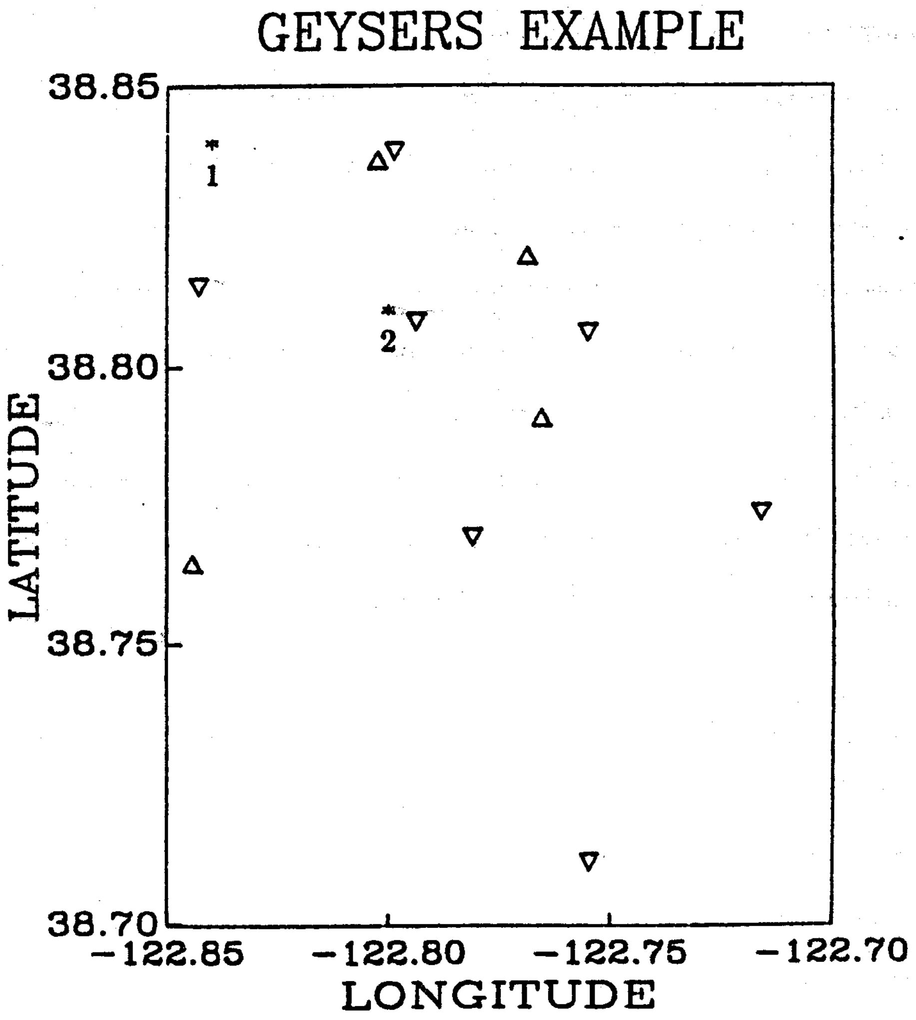


Figure B.1 Map showing the relation of the two event locations, denoted as (1) and (2), considered in the synthetic tests and stations used to compute $\nu(X,Y,Z)$. P and S-wave arrival times were used for stations plotted as Δ and P-wave arrival times were used for stations plotted as ∇. The stations correspond to those used to locate the earthquake in Chapter 4 that occurred on 7-21-82.

for depth, it is normalized to unit area. Confidence limits are determined by integration from infinity to the limits that produce the desired probabilities. A two-sided confidence interval is determined by integrating inward from the extreme values of depth to the depths that produce the desired confidence value. For example, if the two-sided 95% confidence bounds on depth are needed, the minimum depth limit is obtained by integrating from the minimum depth down to the depth limit that produces 2.5% of the area, and the maximum depth limit is obtained by integrating from maximum depth upward to the depth limit that produces 2.5% of the area.

Results of the synthetic tests are summarized in Table (B.1). The most prominent results is the asymmetry of depth uncertainties for computations that used a standard error estimate of 0.05 sec. The effect is most pronounced in case (1) and $\nu(Z)$ is shown for this case in Figure (B.2) and $\nu(X,Y)$ is shown in Figure (B.3). As can be seen from Figures (B.2) and (B.3), nonlinear effects on epicenter uncertainties are much smaller than for depth, indicating that linearized error ellipses are adequate for epicentral epicenter uncertainties. Figure (B.2) shows that hypocenter depth uncertainties are underestimated for events on the edge and outside the recording network in Chapter 4 if model error contributions are 2.5 times larger than arrival time standard errors. If model errors are not significant with respect to arrival time uncertainties, it can be seen from Figure (B.4) and Table (B.1) that linearized error estimates are adequate, although the tendency for a larger uncertainty with respect to shallow locations, will be underestimated. Nonlinear effects cannot be seen in the epicenter marginal density function when a standard error of 0.02 sec is used (Figure (B.5),) indicating negligible nonlinear effects.

Nonlinear effects for case (2) are much smaller than in case (1) (Table (B.1)). Even when the 0.05 sec standard error is used, nonlinear effects are almost negligible. Figure (B.6) shows $\nu(Z)$ and Figure (B.7) shows $\nu(X,Y)$ for 0.05 sec standard error. It is interesting to note that in case (2), depth uncertainty asymmetries are opposite those of case (1). This reflects the fundamental differences in recording geometries for the two cases. The nonlinearity of error scaling in case (2) is small indicating that linearized error estimates are adequate for earthquakes located inside the recording network in Chapter 4 for all plausible error levels.

95% Confidence Deviations				
	Standard error (sec)			
	0.02		0.05	
Case	Shallow	Deep	Shallow	Deep
1	0.53	0.45	1.58	1.08
2	0.28	0.31	0.68	0.81

Table B.1 Differences between the maximum likelihood depth and two-sided 95% confidence level depths in kilometers as a function of standard errors for the two locations cases considered. The columns labeled shallow are the differences with respect to shallower (minimum Z) 95% confidence locations depths and the columns labeled deep are the differences with respect to deeper (maximum Z) 95% confidence depths.

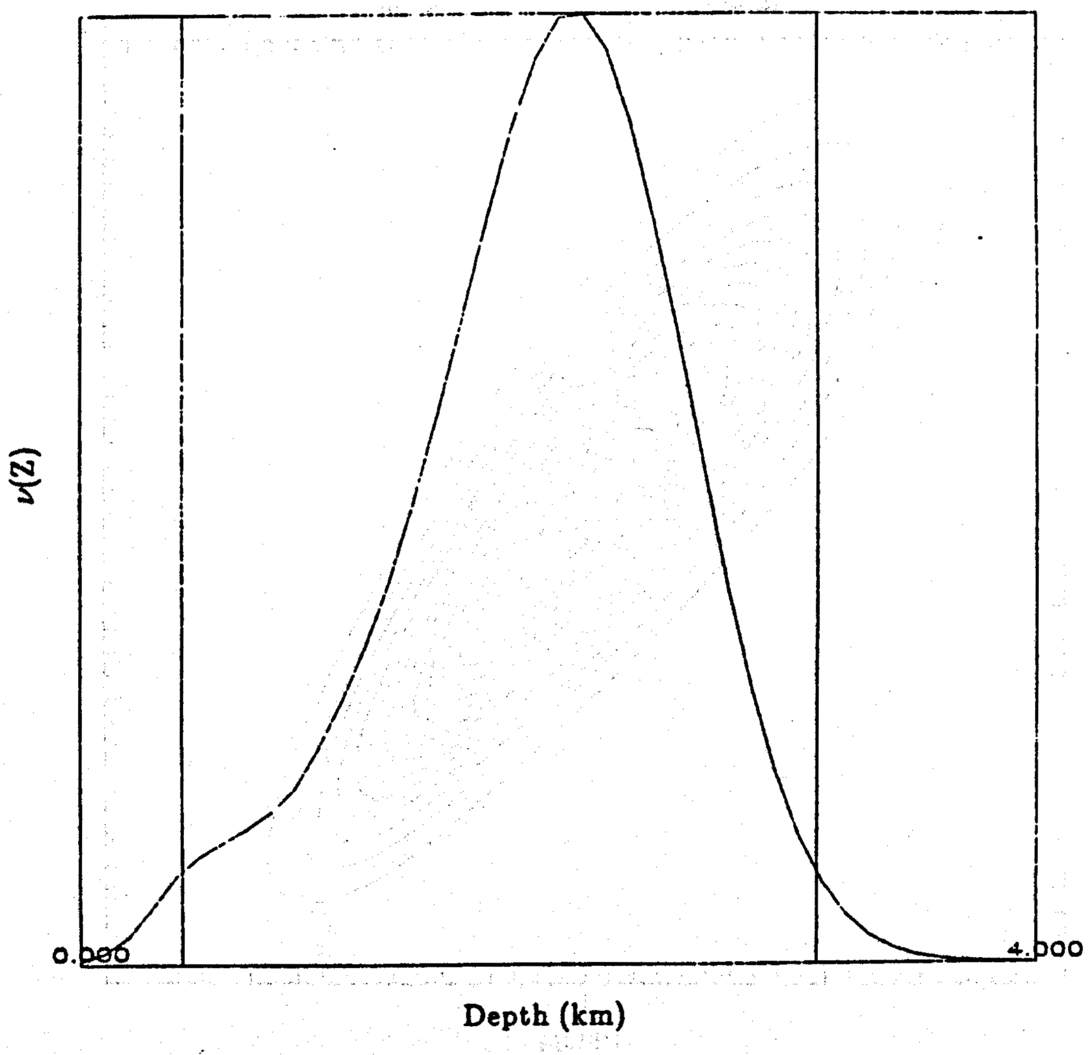


Figure B.2 Marginal density function for depth for case (1) computed using a 0.05 sec standard error. The vertical lines denote the positions of the 95% confidence limits.

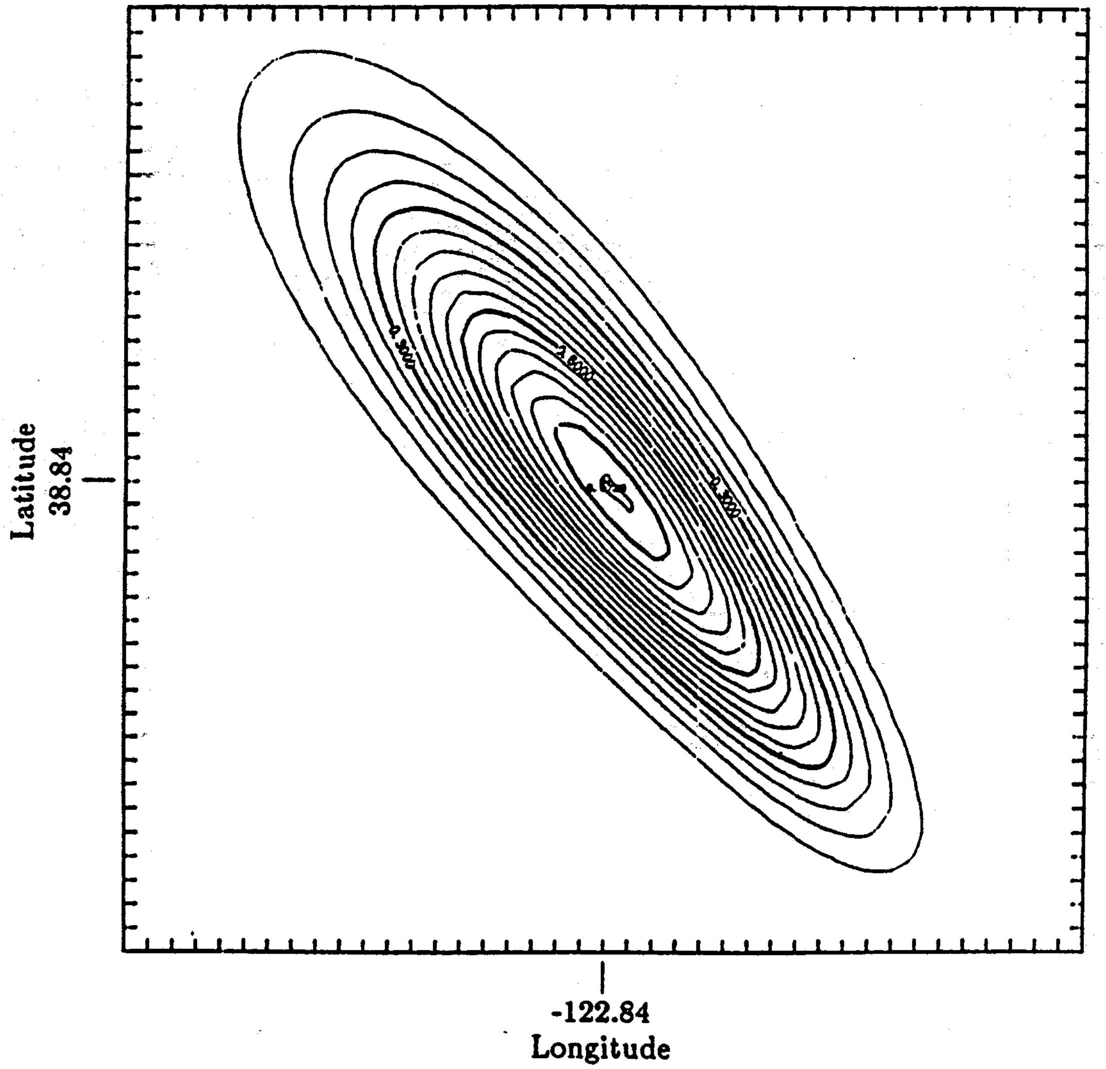


Figure B.3 Marginal density function for epicenter for case (1) computed using a 0.05 sec standard error. The total dimension of each axis is 1.5 km. Notice the slight asymmetry such that the end closer to the recording network is narrower than the far end.

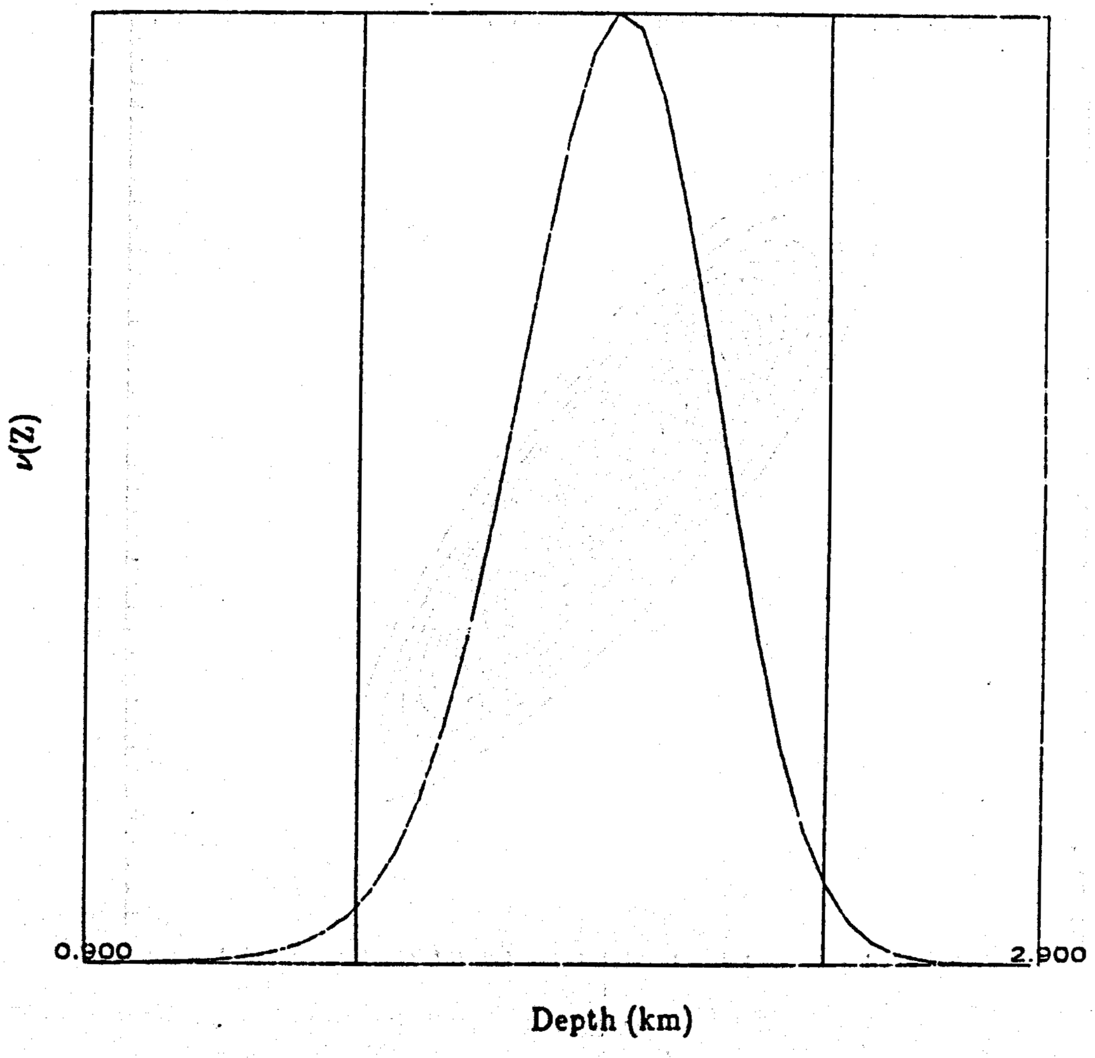


Figure B.4 Marginal density function for depth for case (1) computed using a 0.02 sec standard error. The vertical lines denote the positions of the 95% confidence limits.

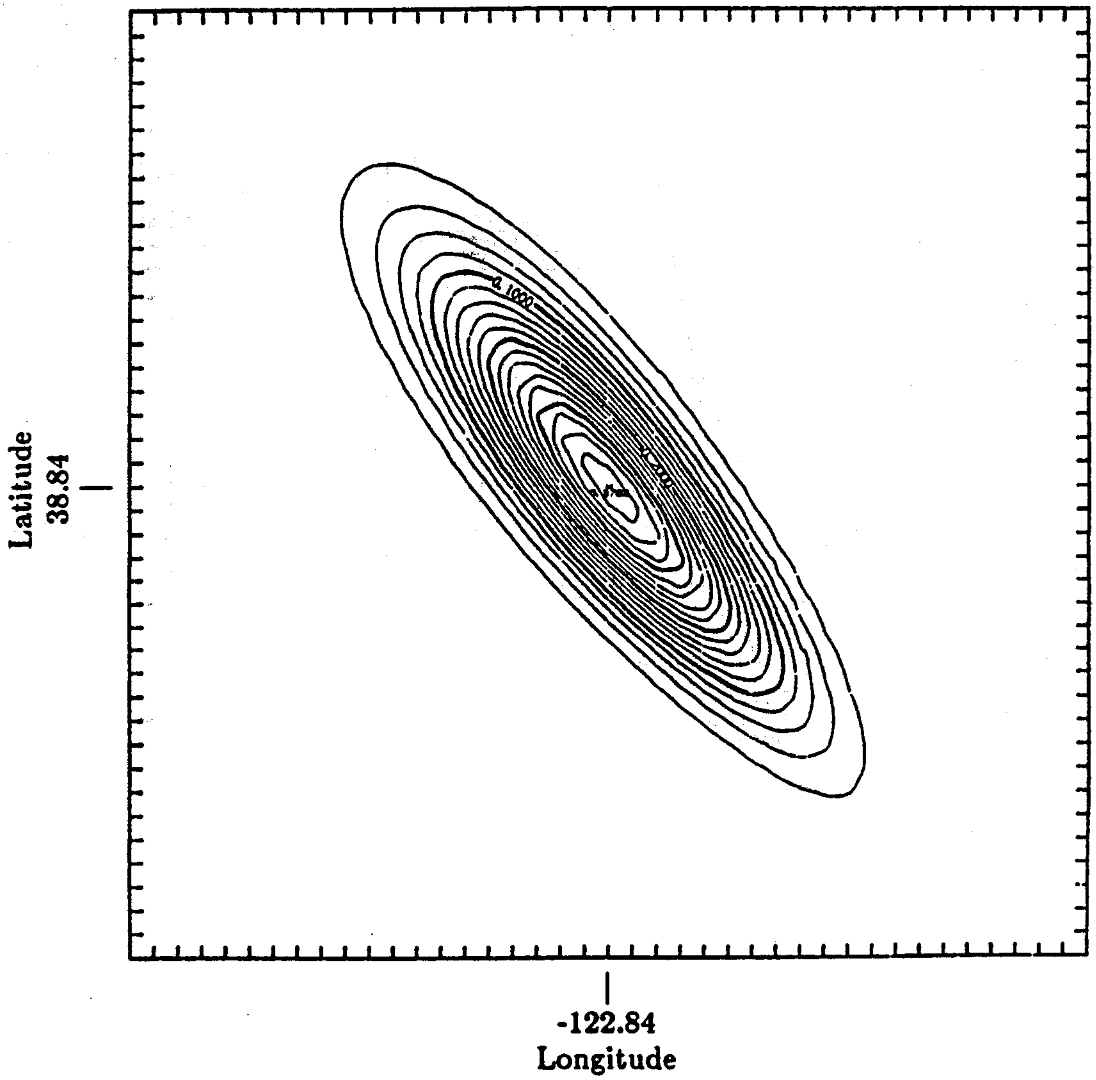


Figure B.5 Marginal density function for epicenter for case (1) computed using a 0.02 sec standard error. The total dimension of each axis is 0.8 km.

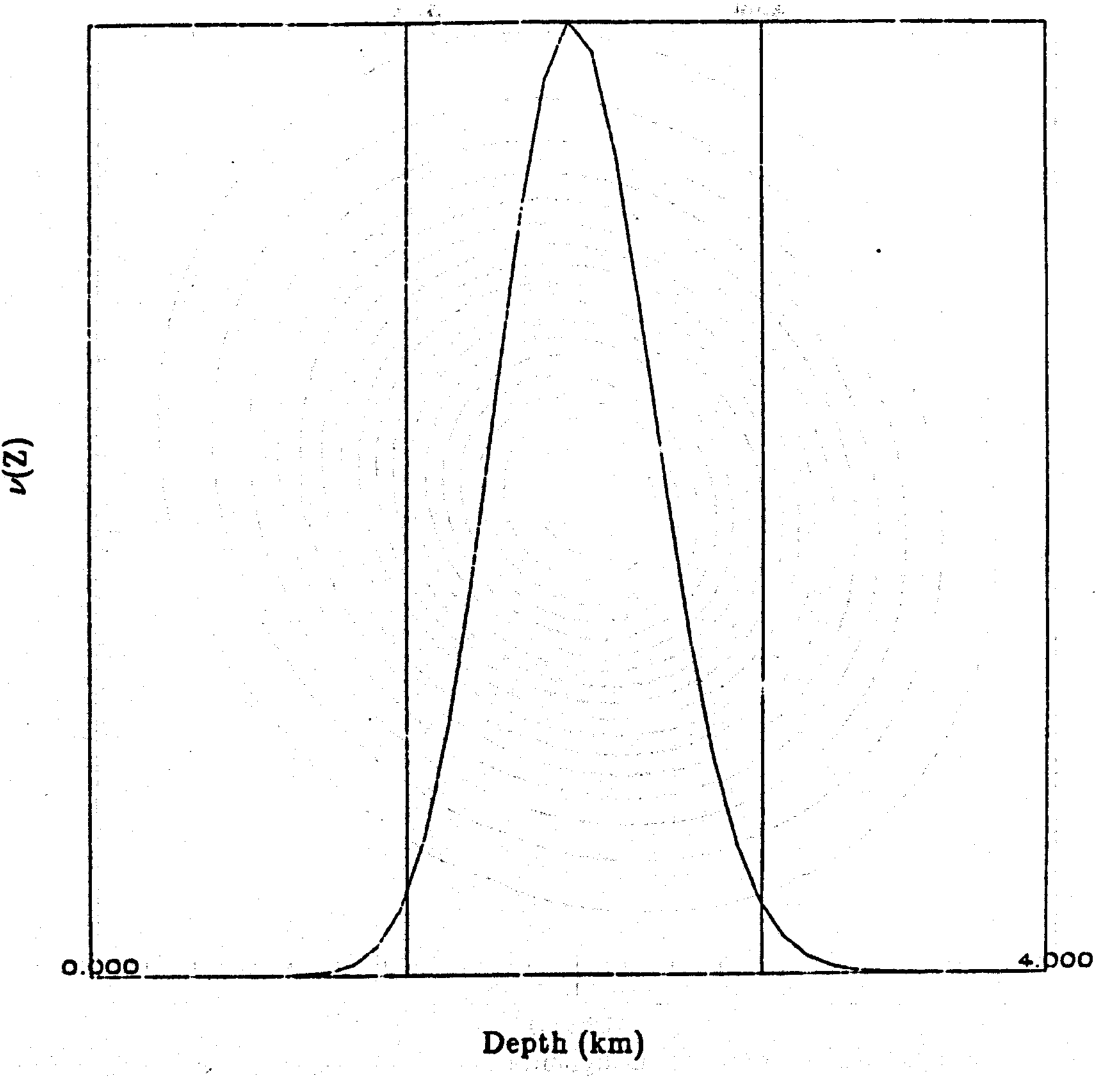


Figure B.6 Marginal density function for depth for case (2) computed using a 0.05 sec standard error. The vertical lines denote the positions of the 95% confidence limits.

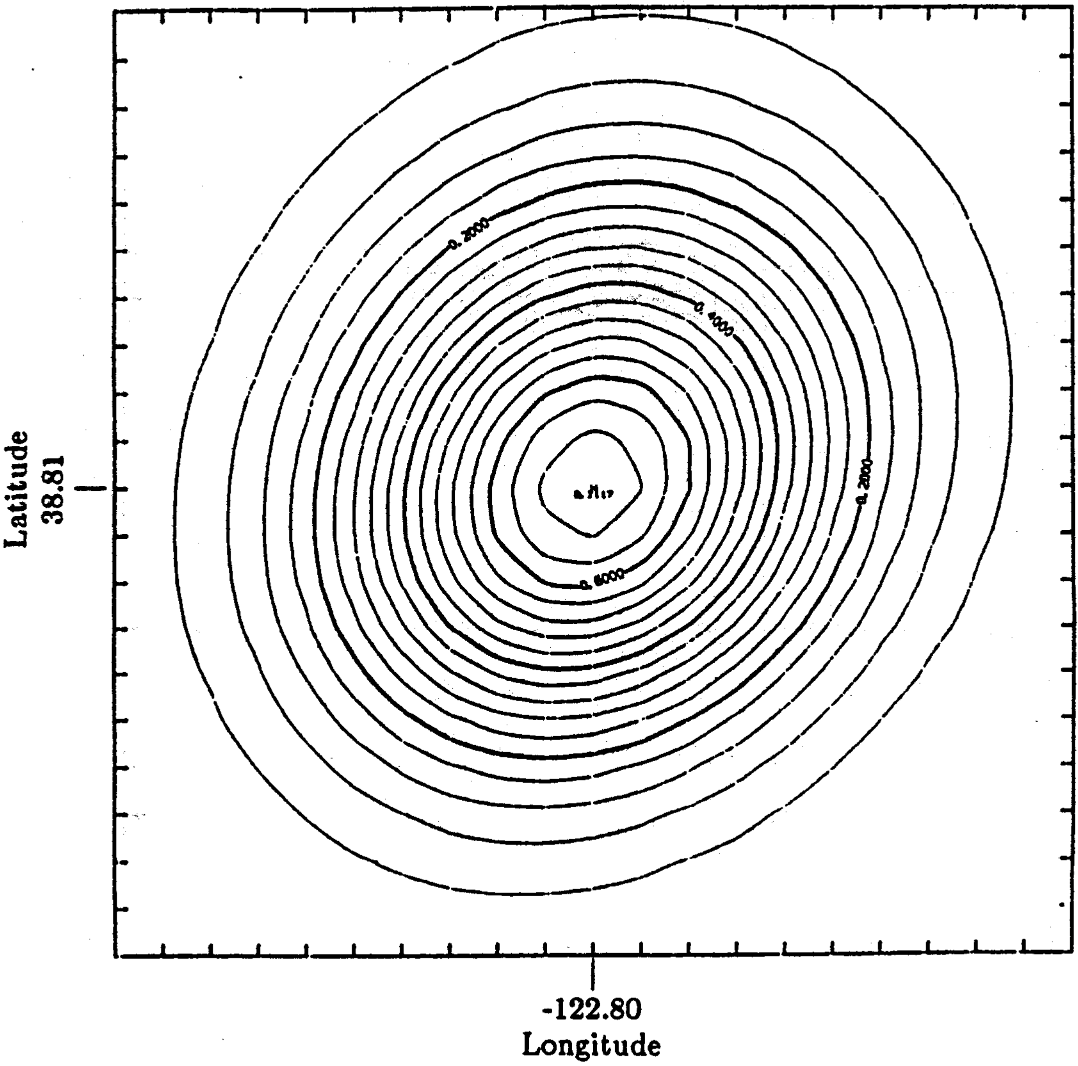


Figure B.7 Marginal density function for epicenter for case (2) computed using a 0.05 sec standard error. The total dimension of each axis is 0.8 km. The contours are nearly circular, reflecting the position of this event in the center of the recording network.

Notes

[1] Pavlis (1986) has outlined a method of hypocenter error analysis that incorporates the Hessian to estimate bounds on nonlinear errors.

[2] Model error is used here in place of the more general classification "theoretical error" in Tarantola and Valette (1982). This serves to emphasize that the primary source of error in calculated travel times is due to errors in the specified velocity model, not errors in numerical solution of the eikonal equations. Because we use a one dimensional velocity structure, the forward problem of calculating travel times is solved with negligible numerical integration errors (see Appendix A). The forward problem of calculating travel times in two and three-dimensional velocity models is more likely to contain significant errors due to numerical methods employed to solve the eikonal equations. In this case the term "theoretical error" is appropriate because the forward problem of solving the eikonal equations is difficult; the theoretical problem of solving for the travel time between two points cannot always be solved with negligible error.

References

Lomnitz, C., 1977. A fast epicenter location program, *Bull. Seism. Soc. Am.*, **67**, 425-431.

Pavlis, G. L., 1982. Progressive inversion, *Ph.D. Dissertation*, University of Washington, Seattle, Washington.

Pavlis, G. L., 1986. Appraising earthquake hypocenter location errors: a complete, practical approach for single event locations. *Eos*, **67**, 306.

Tarantola, A. and B. Valette, 1982. Inverse problems = quest for information, *J. Geophys.*, **50**, 159-170.

Thurber, C. H., 1985. Nonlinear earthquake location: theory and examples, *Bull. Seism. Soc. Am.*, **75**, 779-790.