UNIVERSITY OF SOUTHERN CALIFORNIA

DEPARTMENT OF CIVIL ENGINEERING

A METHOD FOR ESTIMATION OF CRUSTAL VELOCITY STRUCTURE INCLUDING RELOCATION OF SEISMIC EVENTS USING FIRST P-WAVE ARRIVAL TIMES

BY

P. C. SHAH AND F. E. UDWADIA

DEPARTMENT OF CIVIL ENGINEERING
UNIVERSITY OF SOUTHERN CALIF.ORNIA

REPORT NO. CE 78-17

A REPORT ON RESEARCH CONDUCTED UNDER A GRANT SUPPORTED BY THE NATIONAL SCIENCE FOUNDATION

LOS ANGELES, CALIFORNIA

SEPTEMBER, 1978

A METHOD FOR ESTIMATION OF CRUSTAL VELOCITY STRUCTURE INCLUDING RELOCATION OF SEISMIC EVENTS USING FIRST P-WAVE ARRIVAL TIMES

by

P. C. Shah and F. E. Udwadia

ABSTRACT

This paper presents a general methodology for simultaneous estimation of the epicentral locations of local seismic events, their focal depths, their times of origin and the crustal velocity structure of the earth in their vicinity, using the first P-wave arrival times recorded at an array of sensors.

The problem is formulated within the framework of the three-dimensional ray theory. Having started with an initial "guess", a systematic method is developed for iteratively improving the estimates of the four source parameters for each event and the local velocity structure, by using the information contained in the observations. A highly effective and automatic numerical algorithm for this purpose, based on an optimal control formulation of the problem has been provided. The algorithm uses the first order gradient method to minimize a positive scalar index representing the deviations of the calculated P-wave arrival times from those observed.

The method is illustrated using two-dimensional simulations in which either all or part of the parameters are to be estimated. The nonuniqueness of the estimates is demonstrated, pointing to the fact that unless large amounts of travel time data are available, the information contained in such data may not be sufficient to pin down uniquely the source parameters and crustal velocity distribution.

INTRODUCTION

The problem of determining the local velocity structure, the location and times of occurrance of the seismic events from the observation of the times of the first p-wave arrivals at several sensors from local earthquakes has received increasing attention from the seismologists in the recent past. Wesson (1971) addressed the problem of determination of simple lateral inhomogenities in the velocity distribution and determined the parameters using an iterative least-squares approach. Lee and Engdahl (1976) treated the problem of relocation of earthquakes. They used the ray tracing approach and assumed a known heterogeneous velocity distribution in the earth medium. Aki and Lee (1976) have attempted the combined problem of determination of all four source parameters together with the three-dimensional velocity distributions; however, they obtained only approximate solutions by employing a single correction, starting with a homogeneous velocity distribution. Their method does not include a technique for iterative corrections of both the source parameters (the epicentral location, focal depth, and origin time) and the velocity distribution using the ray theory approach.

Crosson (1976) has attempted a least-squares formulation of the problem of simultaneously relocating earthquakes and determining the velocity profiles. He uses the damped least-squares approach and calculates the sensitivities of the first p-wave arrival times with respect to those parameters that model the velocity distribution as well as those that refer to the source.

In this paper we present a new iterative approach, based on the hill-climbing technique, for simultaneous determination of the epicentral locations, the focal depths and the origin times of local seismic events, and the local crustal velocity structure using the first arrival times of the p-waves from these events, recorded by an array of seismic sensors. The model for this problem employes ray-tracing in a three-dimensional region with an inhomogeneous velocity distribution.

In the approach taken here, the estimation problem is viewed as an optimal control problem (Bryson and Ho, 1969), wherein the parameters being estimated are control variables and the ray equations govern the system "response". The optimality criterion to be minimized is a positive scalar index consisting of a weighted sum of the squares of the deviations of the calculated arrival times (based on the model) from those actually observed. This approach leads to a gradient algorithm for the solution of the inverse problem. This method for the identification of dynamic systems appears to have been first employed independently by Chen et al (1974) and Chevant et al (1975), in connection with the determination of porous rock properties from oil pressure data in producing petroleum reservoirs. Later Udwadia and Shah (1976) applied this method to the problem of identification of distributed parameters in a hyperbolic equation modeling building structural motions during earthquake ground shaking. The major difference between these problems and the travel time inversion problem is that in the former set the dynamic state variables are observed as they evolve in time, whereas in the latter, the observed quantities (travel times) are fixed for a given model. Alternatively, the observed arrival times in the present problem may be viewed as functionals ofthe ray solutions, the co-ordinates of which are the evolving state variables of our "dynamic" system.

Using the optimal control formulation of the estimation problem, expressions are derived for the gradient of the scalar functional in the space of the unknown parameters; the calculated gradient is used in the hill-climbing technique of the conjugate gradient directions (Luenberger, 1973) for minimization. These gradients are obtained by considering the P-wave velocity to be a distributed parameter; they are later adapted to the situation wherein the distribution is determined by a finite number of scalar parameters. These expressions involve solutions to the model equations and the associated adjoint equations. The adjoint equations are derived considering the first variation of an expression for the minimization index; they are linear, inhomogeneous differential equations. An important advantage of this approach is its computational efficiency: the need for calculating the sensitivity of each observation with respect to each of the unknown parameters is obviated; furthermore, the effort involved in the least-square solution of large matrix equations, at each iterative step, is eliminated. However, this saving is achieved at the expense of the additional computational effort involved in searching (at each iteration) for the first local minimum of the index in a conjugate gradient direction in the parameter space.

The proposed gradient method is illustrated by application to a simulated two-dimensional problem with all of the seismic events and the sensors lying in a vertical plane. The velocity distribution is parameterized in two alternate ways: using block interpolation and horizontal layers. The two-point boundary value problem, involved in ray tracing, is numerically solved using the computer program PASVA2 by Lentini and Pereyra (1975) that employs a variable step method. Several alternative estimation problems with different "true" velocity distributions are treated; for each velocity distribution, several alternative cases are considered, with different kinds of parameters unknown in each case.

From a physical standpoint, it is clear that the observations of the travel times contain information about the velocity distribution of the medium mainly along the ray paths. Hence, in general, the velocity distributions cannot be determined in the regions which do not contain any rays. Also, associated with each seismic event are its unknown source parameters, thus further reducing the information content available to decipher the velocity distribution. Thus, the number of sensors recording the first P-wave arrivals from any seismic event must exceed the number of its source parameters that are unknowns; moreover, improved identification will be achieved by increasing the number of sensors. The simulation studies reported herein bear out these facts and show that the identification may lead to nonunique solutions. Numerical results indicate that multiple solutions which match the given set of travel time observations, with a certain accuracy, are possible.

FORMULATION OF THE PARAMETER ESTIMATION PROBLEM

For simplicity, we shall assume that the P-wave velocity distribution is independent of the y co-ordinate. Then a ray between two points in the x-z plane will lie entirely in that plane, and the velocity distribution in a bounded domain D in the x-z plane is to be estimated. However, we note that the methodology presented below is equally applicable to the general three-dimensional problem with $v = v(x,y,z)$ and the derivation of the gradient can be extended in a straight forward manner to treat that problem.

The equation of the ray between a point (x_o,z_o) where a seismic event originates and (x_s,y_s) where a sensor is located is given by,

$$\frac{d}{ds}(wx') - \frac{\partial w}{\partial x} = 0 \tag{1}$$

$$\frac{d}{ds}(wz') - \frac{\partial w}{\partial z} = 0 \tag{2}$$

with the boundary conditions,

$$x(0) = x_o,\ z(0) = z_o\ ,\ x(S) = x_s,\ z(S) = z_s \tag{3}$$

where s is the arc length measured along the trajectory measured from the origin point (x_o,z_o), S is the total length of the trajectory, and $w(x,z) = 1/v(x,z)$. For simplicity, we shall treat w as the distributed parameter to be estimated instead of v in the following.

Let there be m seismic events occuring at location (x_{o_i}, z_{o_i}) at times T_{o_i}, $i = 1,2,\ldots m$ and let there be n sensors located at $\{(x_{s_j}, z_{s_j}) : j = 1,2,\ldots,n\}$ where the times of the first arrival of the P-waves from each of the m events are recorded. Then, denoting the travel time of the signal along the ray between (x_{o_i}, z_{o_i}) and (x_{s_j}, z_{s_j})

by ΔT_{ij}, we have the observed time of the first arrival of the P-wave, to be,

$$T_{ij}^{obs} = T_{o_i} + \Delta T_{ij} + \eta_{ij} \tag{4}$$

where η_{ij} is an observation error. Furthermore, the travel time is given by,

$$\Delta T_{ij} = \int_0^{S_{ij}} w(s)\,ds \tag{5}$$

where S_{ij} is the length of the ray path from the i-th epicenter to the j-th sensor.

Assuming η_{ij} to be independent, zero-mean Gaussian random variable with variances σ_{ij}^2, the velocity estimation problem may be posed as the minimization problem:

$$\underset{\substack{\{w(x,z):\ x,z\varepsilon D\} \\ \{(x_{o_i},z_{o_i},T_{o_i}):\ i=1,\ldots,m\}}}{\text{Minimize}} \qquad J = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{1}{\sigma_{ij}^2} (T_{ij}^{obs} - T_{ij}^{calc})^2 \tag{6}$$

where T_{ij}^{calc} is the calculated value of the first arrival time of P-waves from the event i at the sensor j using the estimated values of the unknown parameters,

$$T_{ij}^{cal} = T_{o_i}^{est} + \Delta T_{ij}^{calc} \tag{7}$$

We shall assume that the domain D is sufficiently large so as to enclose all the mn ray paths at all stages of an iterative minimization procedure. For simplicity, we take D to be the rectangle ($0 \leq x \leq a$, $0 \leq z \leq L$) as shown in Figure 1.

In the following we propose an optimal control formulation of this minimization problem and derive formulae for the gradient of J with respect to the unknowns.

Optimal Control Formulation:

In optimal control formulation, the minimization problem is treated as an optimal control problem (Bryson and Ho, 1969) with the minimization criterion J as a performance index and the parameters to be estimated as control variables; the optimal control solution is to be determined subject to the system equations, which are treated as constraints on the state variables. From this standpoint, equations (1), (2) and (3) form constraints on the state variables $x_{ij}(s)$ and $z_{ij}(s)$, and the quantities w, $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial z}$ along the ray paths constitute our control variables. However, the last three quantities are not independent of each other and the relations between them must be included as constraint equations. Writing $\frac{\partial w}{\partial x} = f_1$ and $\frac{\partial w}{\partial z} = f_2$, these are:

$$\frac{dw_{ij}}{ds} = x'_{ij} f_1 + z'_{ij} f_2 \qquad (8)$$

$$\frac{\partial f_1}{\partial z} = \frac{\partial f_2}{\partial x} \quad (x,z)\varepsilon D \qquad (9)$$

where $w_{ij}(s)$ is the value of w on the $(i,j)^{th}$ ray. In order to include (8) and (9) as constraints in our formulation, $w_{ij}(s)$ and $f_1(x,z)$ may be treated as additional state variables and $f_2(x,z)$ may be treated as a control variable. Then, to determine uniquely these additional state variables for a given value of the control variable f_2, initial conditions must be supplied to equation (8) and (9). Let these be,

$$w_{ij}(0) = w_{o_i} \qquad i = 1,2,\ldots,m \qquad (10)$$

$$f_1(x,0) = h(x) \quad x\varepsilon(o,a) \qquad (11)$$

Since the values of w_{o_i} and h(x) are unknown, they must be treated as additional control variables.

Furthermore, in order to integrate the equations (1), (2) and (8) the value of the total arc length S has to be known. However, since it depends on the ray path it cannot be determined apriori. Hence, it also must be estimated and thus added on to the list of the control variables.

Thus, the equivalent optimal control problem consists of determining the contral variables $\{w_{o_i}, x_{o_i}, z_{o_i}, T_{o_i} : i=1,2,\ldots,m\}$, $\{S_{ij} : i=1,2,\ldots m; j = 1,2,\ldots n\}$, $f_2(x,z)$ and $h(x)$ such that the value of the performance index J defined in (6) is a minimum when the constraint equations (1), (2), (3), (8), (9), (10), and (11) are satisfied by the state variables. We note that instead of relation (8) we may include a more general relation

$$\frac{\partial w}{\partial x} = f_1(x,z) \qquad (x,z)\varepsilon D \tag{10}$$

with the boundary condition,

$$w(o,z) = h_1(x) \tag{11}$$

However, relation (8) is preferred over (10) because the system equations (1) and (2) depend only on the values of W and its derivatives along the ray path. Furthermore, it is evident that the observations contain information about w, f_1 and f_2 at points mainly along the ray paths. Use of relation (8) will lead to a formulation where the values of w only along the ray paths are directly estimated.

The minimization of J can be efficiently carried out using the hill-climbing technique as follows: starting with some a priori estimates of the unknowns, we make small corrections iteratively in their current estimates such that the value of J decreases locally at the greatest possible rate. Evidently, the iterative correction at any stage must be in the direction of the negative gradient of the function J in the control variable space. In the following we derive a procedure to calculate this gradient for a given set of estimates of the control variables.

Derivation of the Gradient of J with Respect to the Parameters:

Adjoining the constraint equations (1), (2), (8) and (9) to the performance index J using arbitrary multiplier functions $\lambda_{ij}(s)$, $\alpha_{ij}(s)$, $\varepsilon_{ij}(s)$ and $\Psi(x,z)$ we obtain,

$$\bar{J} = \sum_{i=1}^{m} \sum_{j=1}^{n} \left[\frac{1}{2\sigma_{ij}^2} \left(T_{ij}^{obs} - T_{o_i} - \Delta T_{ij}^{cal}\right)^2 + \int_0^{S_{ij}} \left\{\lambda_{ij}(s)\left[\frac{d}{ds}(w_{ij}x_{ij}') - f_1\right] + \alpha_{ij}(s)\left[\frac{d}{ds}(w_{ij}z_{ij}') - f_2\right] + \xi_{ij}(s)\left[\frac{dw_{ij}}{ds} - f_1x_{ij}' - f_2z_{ij}'\right]\right\} ds\right] + \int_{x=0}^{a}\int_{z=0}^{L} \Psi(x,z)\left[\frac{\partial f_1}{\partial z} - \frac{\partial f_2}{\partial x}\right] dx\, dz \qquad (12)$$

Let us consider the variation of J due to small variations in the control variables, $\delta h(x)$, $\delta f_2(x,z)$, $\{\delta T_{o_i}, \delta w_{o_i}, \delta x_{o_i}, \delta z_{o_i} : i=1,2,\ldots,m\}$, $\{\delta S_{ij}\}$. Let the corresponding variations in the state variables be denoted by δw_{ij}, δx_{ij}, δz_{ij}, and $\delta f_1(x,z)$. Then we have

$$\delta\bar{J} = \sum_i \sum_j \left\{-\frac{1}{\sigma_{ij}^2}\left(T_{ij}^{obs} - T_{o_i} - \Delta T_{ij}^{cal}\right)\left(\delta T_{o_i} + \delta\int_0^{S_{ij}} w_{ij}(s)ds\right) + \delta\int_0^{S_{ij}} \lambda_{ij}(s)\left[\frac{d}{ds}(w_{ij}x_{ij}') - f_1\right]ds + \delta\int_0^{S_{ij}} \alpha_{ij}(s)\left[\frac{d}{ds}(w_{ij}z_{ij}') - f_2\right]ds + \delta\int_0^{S_{ij}} \xi_{ij}(s)\left[\frac{dw_{ij}}{ds} - f_1x_{ij}' - f_2z_{ij}'\right]ds\right\} + \delta\int_0^a\int_0^L \Psi(x,z)\left[\frac{\partial f_1}{\partial z} - \frac{\partial f_2}{\partial x}\right]dxdz \qquad (13)$$

The first order variations (Gelfand and Fomin, 1963) of the integrals along the rays in the last expression can be evaluated using the following formula:

$$\delta\int_{x_0}^{x_1} f(x, y_1, y_1', y_1'', y_2, y_2', y_2'', \ldots, y_n'')\, dx = \sum_i \int_{x_0}^{x_1}\left[f_{y_i} - \frac{d}{dx}(f_{y_i'}) + \frac{d^2}{dx^2}(f_{y_i''})\right]\delta y_i dx$$

$$+\left[f - \sum_i f_{y_i'} y_i' + \sum_i \frac{d}{dx}(f_{y_i''})y_i' - \sum_i f_{y_i''} y_i''\right]\delta x\,\Big|_{x_0}^{x_1}$$

$$+\sum_i \left(f_{y_i'} - \frac{d}{dx} f_{y_i''}\right)\delta y_i\Big|_{x_0}^{x_1} + \sum_i f_{y_i''}\,\delta y_i'\Big|_{x_0}^{x_1} \tag{14}$$

We give the results below for a single trajectory, dropping the subscripts i and j for simplicity; we note that the lower limits of all the integrals is fixed at s=0.

$$\delta\int_0^S w(s)ds = \int_0^S \delta w ds + w(s)\delta s\Big|_0^S = \int_0^S \delta w ds + w(S)\delta S \tag{15}$$

$$\delta\int_0^S \lambda(s)\left[\frac{d}{ds}(wx') - f_1\right]ds = \delta\int_0^S \lambda(s)\left[w'x' + wx'' - f_1\right]ds$$

$$= -\int_0^S \lambda' x'\delta w ds + \int_0^S \left[\lambda'' w + \lambda' w'\right]\delta x\, ds - \int_0^S \lambda\,\delta f_1 ds$$

$$+ \lambda x'\delta w\Big|_0^S - \lambda' w\delta x\Big|_0^S + (\lambda' wx' - \lambda f_1)\delta s\Big|_0^S + \lambda w\delta x'\Big|_0^S \tag{16}$$

$$\delta\int_0^S \alpha(s)\left[\frac{d}{ds}(wz') - f_2\right]ds$$

$$= -\int_0^S \alpha' z'\delta w ds + \int_0^S \left[\alpha'' w + \alpha' w'\right]\delta z ds - \int_0^S \alpha\delta f_2 ds$$

$$+ \alpha z'\delta w\Big|_0^S - \alpha' w\delta z\Big|_0^S + (\alpha' wz' - \alpha f_2)\delta s\Big|_0^S + \alpha w\delta z'\Big|_0^S \tag{17}$$

$$\delta\int_0^S \xi(s)\left[\frac{dw}{ds} - f_1 x' - f_2 z'\right]ds$$

$$= -\int_0^S \xi'\delta w ds + \int_0^S \left[\frac{d}{ds}(\xi f_1) - \xi\frac{\partial f_1}{\partial x} x' - \xi\frac{\partial f_2}{\partial x} z'\right]\delta x ds$$

$$-\int_0^S \xi x'\delta f_1 ds + \int_0^S \left[\frac{d}{ds}(\xi f_2) - \xi\frac{\partial f_1}{\partial z} x' - \xi\frac{\partial f_2}{\partial z} z'\right]\delta z ds - \int_0^S \xi z'\delta f_2 ds$$

$$+ \xi\,\delta w\Big|_0^S - \xi f_1\delta x\Big|_0^S - \xi f_2\delta z\Big|_0^S \tag{18}$$

where $g(s)\Big|_0^S = g(S) - g(0)$. We note that since the sensor location is fixed, in the above expressions $\delta x(S) = \delta z(S) = 0$.

The first variation of the last integral in (13) can be obtained after integration by parts as,

$$\delta\int_0^a\!\!\int_0^L \Psi(x,z)\left[\frac{\partial f_1}{\partial z} - \frac{\partial f_2}{\partial x}\right] dxdz = -\int_0^a\!\!\int_0^L \left[\frac{\partial \Psi}{\partial z}\,\delta f_1 - \frac{\partial \Psi}{\partial x}\,\delta f_2\right] dxdz$$

$$+\int_0^a \Psi\delta f_1\Big|_{z=0}^{L} dx - \int_0^L \Psi\delta f_2\Big|_{x=0}^{a} dz \qquad (19)$$

Substitution of results (15) through (19) into (13) and a rearrangement yields the following expression for the first order variation in $\bar{J}$:

$$\begin{aligned}
\delta\bar{J} = \sum_i\sum_j \Bigg\{ & \int_0^{S_{ij}}\left[\lambda''_{ij}w_{ij} + \lambda'_{ij}w'_{ij} + \frac{d}{ds}(\xi_{ij}f_1) - \xi_{ij}\frac{\partial f_1}{\partial x}x'_{ij} - \xi_{ij}\frac{\partial f_2}{\partial x}z'_{ij}\right]\delta x_{ij}\,ds \\
& + \int_0^{S_{ij}}\left[\alpha''_{ij}w_{ij} + \alpha'_{ij}w'_{ij} + \frac{d}{ds}(\xi_{ij}f_2) - \xi_{ij}\frac{\partial f_1}{\partial z}x'_{ij} - \xi_{ij}\frac{\partial f_2}{\partial z}z'_{ij}\right]\delta z_{ij}\,ds \\
& - \int_0^{S_{ij}}\left[(T_{ij}^{obs} - T_{ij}^{calc}) + \lambda'_{ij}x'_{ij} + \alpha'_{ij}z'_{ij} + \xi'_{ij}\right]\delta w_{ij}\,ds \\
& - \int_0^{S_{ij}}\left[\lambda_{ij} + \xi_{ij}x'_{ij}\right]\delta f_1\,ds - \int_0^{S_{ij}}\left[\alpha_{ij} + \xi_{ij}z'_{ij}\right]\delta f_2\,ds \\
& + (\lambda_{ij}x'_{ij} + \alpha_{ij}z'_{ij} + \xi_{ij})\delta w_{ij}\Big|_0^{S_{ij}} + (\lambda'_{ij}w_{ij} + \xi_{ij}f_1)\Big|_{s=0}\delta x_{o_i} \\
& + (\alpha'_{ij}w_{ij} + \xi_{ij}f_2)\Big|_{s=0}\delta z_{o_i} + \lambda_{ij}w_{ij}\delta x'_{ij}\Big|_0^{S_{ij}} + \alpha_{ij}w_{ij}\delta z'_{ij}\Big|_0^{S_{ij}} \\
& + \left[-(T_{ij}^{obs} - T_{ij}^{calc})w_{ij} + \lambda'_{ij}w_{ij}x'_{ij} - \lambda_{ij}f_1 + \alpha'_{ij}w_{ij}z'_{ij} - \alpha_{ij}f_2\right]\Big|_{s=S_{ij}}\delta S_{ij}\Bigg\} \\
& - \int_0^a\!\!\int_0^L\left[\frac{\partial\Psi}{\partial z}\,\delta f_1 - \frac{\partial\Psi}{\partial x}\,\delta f_2\right]dxdz + \int_0^a \Psi\delta f_1\Big|_{z=0}^{L}dx - \int_0^L \Psi\delta f_2\Big|_{x=0}^{a}dz \qquad (20) \\
& - \sum_i\sum_j \frac{1}{\sigma_{ij}^2}(T_{ij}^{obs} - T_{ij}^{calc})\,\delta T_{o_i}
\end{aligned}$$

So far, the multiplier functions $\lambda_{ij}(s)$, $\alpha_{ij}(s)$ and $\xi_{ij}(s)$ are completely arbitrary. We may select them in such a manner that the expression for $\delta\bar{J}$ becomes the simplest possible. For this purpose, we let λ_{ij}, α_{ij} and ξ_{ij} satisfy,

$$\lambda''_{ij}w_{ij} + \lambda'_{ij}w'_{ij} + \frac{d}{ds}(\xi_{ij}f_1) - \xi_{ij}\frac{\partial f_1}{\partial x}x'_{ij} - \xi_{ij}\frac{\partial f_2}{\partial x}z'_{ij} = 0 \qquad (21)$$

$$\alpha''_{ij}w_{ij} + \alpha'_{ij}w'_{ij} + \frac{d}{ds}(\xi_{ij}f_2) - \xi_{ij}\frac{\partial f_1}{\partial z}x'_{ij} - \xi_{ij}\frac{\partial f_2}{\partial z}z'_{ij} = 0 \qquad (22)$$

$$\xi'_{ij} + \lambda'_{ij}x'_{ij} + \alpha'_{ij}z'_{ij} = -(T^{obs}_{ij} - T^{calc}_{ij}) \qquad s\varepsilon(0,S_{ij}) \qquad (23)$$

$$\lambda_{ij}(0) = \lambda_{ij}(S_{ij}) = \alpha_{ij}(0) = \alpha_{ij}(S_{ij}) = \xi_{ij}(S_{ij}) = 0 \qquad (24)$$

We note that relations (21) through (24) constitute a well-posed, linear inhomogeneous two-point boundary value problem in the variables λ_{ij}, α_{ij}, and ξ_{ij}. Hence, it has a unique solution.

When the multipliers are solution of the above problem, and when the functions $x_{ij}(s)$, $z_{ij}(s)$, $w_{ij}(s)$ and $f_1(x,z)$ satisfy the system equations, we obtain the following expression for $\delta\bar{J}$ after transforming the remaining integrals along rays into integrals over D by using delta and unit step (Heaviside) functions:

$$\delta J = \delta\bar{J} = -\int_0^a\int_0^L\left\{\frac{\partial\Psi}{\partial z} + \sum_i\sum_j(\lambda_{ij} + \xi_{ij}x'_{ij})\delta(x-x_{ij}(s))\delta(z-z_{ij}(s))\Big[H(0)-H(S_{ij})\Big]\right\}\delta f_1\,dxdz$$

$$+\int_0^a\int_0^L\left\{\frac{\partial\Psi}{\partial x} - \sum_i\sum_j(\alpha_{ij} + \xi_{ij}z'_{ij})\delta(x-x_{ij}(s))\delta(z-z_{ij}(s))\Big[H(0)-H(S_{ij})\Big]\right\}\delta f_2\,dxdz$$

$$-\ \xi_{ij}(0)\delta w_{o_i} + (\lambda'_{ij}w_{ij} + \xi_{ij}f_1)\Big|_{s=0}\delta x_{o_i} + (\alpha'_{ij}w_{ij} + \xi_{ij}f_2)\Big|_{s=0}\delta z_{o_i}$$

$$+ \left[-(T_{ij}^{obs} - T_{ij}^{calc})w_{ij} + \lambda_{ij}' w_{ij} x_{ij}' + \alpha_{ij}' w_{ij} z_{ij}' \right] \Bigg|_{s=S_{ij}} \delta S_{ij} - \sum_i \sum_j \frac{1}{\sigma_{ij}^2} (T_{ij}^{obs} - T_{ij}^{calc}) \delta T_{o_i}$$

$$- \int_0^L \Psi \delta f_2 \Big|_{x=0}^{a} dz - \int_0^a \Psi \delta f_1 \Big|_{z=0} dx + \int_0^a \Psi \delta f_1 \Big|_{z=L} dx \qquad (25)$$

Furthermore, we let the multiplier $\Psi(x,z)$ be determined by the initial value problem,

$$\frac{\partial \Psi}{\partial z} = - \sum_i \sum_j (\lambda_{ij} + \xi_{ij} x_{ij}') \delta(x - x_{ij}(s)) \delta(z - z_{ij}(s)) \left[H(0) - H(S_{ij}) \right],$$

$$(x,z) \varepsilon D \qquad (26)$$

$$\Psi(x,L) = 0 \qquad (27)$$

Then the first and the last terms on the right of (25) vanish.

The functional derivative $\frac{\delta J}{\delta \phi(x,z)}$ of a scalar functional J with respect to its argument function $\{\phi(x,z): x\varepsilon(x_o,x_1), z\varepsilon(z_o,z_1)\}$ is defined by (Volterra, 1959),

$$\delta J = \int_{x_0}^{x_1} \int_{z_0}^{z_1} \frac{\delta J}{\delta \phi(x,z)} \delta \phi(x,z) \, dxdz \qquad (28)$$

where δJ is the change in the value of J due to a change $\delta\phi(x,z)$ in ϕ. Then from expression (25) we obtain the following functional derivatives:

$$\frac{\delta J}{\delta f_2(x,z)} = \frac{\partial \Psi}{\partial x} - \sum_i \sum_j (\alpha_{ij} + \xi_{ij} z_{ij}') \delta(x - x_{ij}(s)) \delta(z - z_{ij}(s)) \left[H(0) - H(S_{ij}) \right]$$

$$+ \Psi(0,z)\delta(x) - \Psi(a,z)\delta(x-a), \qquad (x,z)\varepsilon D \qquad (29)$$

$$\frac{\delta J}{\delta h(x)} = \frac{\delta J}{\delta f_1(x,0)} = -\psi(x,0), \qquad x\varepsilon(0,a) \qquad (30)$$

Furthermore, we obtain from (25) the gradient of J with respect to the scalar parameters x_{o_i}, z_{o_i}, w_{o_i}, S_{ij} and T_{o_i} as,

$$\frac{\partial J}{\partial x_{o_i}} = \sum_j \left\{ \left[\lambda'_{ij} w_{ij} + \xi_{ij} f_1 \right] \Big|_{s=0} \right\} \tag{31}$$

$$\frac{\partial J}{\partial z_{o_i}} = {}_j\left\{ \left[\alpha'_{ij} w_{ij} + \xi_{ij} f_2 \right] \Big|_{s=0} \right\} \tag{32}$$

$$\frac{\partial J}{\partial w_{o_i}} = - \sum_j \xi_{ij}(0) \tag{33}$$

$$\frac{\partial J}{\partial S_{ij}} = \left[(T^{obs}_{ij} - T^{calc}_{ij}) + \lambda'_{ij} x'_{ij} + \alpha'_{ij} z'_{ij} \right] w_{ij} \Big|_{s=S_{ij}} \tag{34}$$

$$\frac{\partial J}{\partial T_{o_i}} = - \sum_j \frac{1}{\sigma^2_{ij}} (T^{obs}_{ij} - T^{calc}_{ij}) \tag{35}$$

Thus, the gradient of J with respect to all the control variables can be directly evaluated once the system equation and the adjoint equations are solved. Each of the sets of equations {(1), (2), (3), (8), and (10)} and {(21) through (24)} is equivalent to a two point boundary value problem with five first order ordinary differential equations. (In a three-dimensional problem there will be seven such first order equations in each set.) The dimensionality of these sets of equations can be reduced if we restrict x_{ij} to be a single valued function of z and use an alternative formulation of the ray equations as presented below. In addition to the two point boundary value problems, the above formulation requires integration of two ((9) and (26)) inhomogeneous, linear, first-order partial differential equations. (In a three-dimensional problem, there will be four such equations -- two compatibility conditions and two equations for the corresponding mulitplier functions.)

Alternative Formulation of the Problem:

If the velocity distribution v(x,z) is such that the ray solutions are single valued functions often expressed either as $x_{ij} = x_{ij}(z)$ or $z_{ij} = z_{ij}(x)$, then the parametric variable s may be dispensed with.

Assuming that $x_{ij} = x_{ij}(z)$ are single valued for all i and j, the equations describing the rays are,

$$\frac{d}{dz}\left(\frac{x'_{ij}}{v(1+x'^{2}_{ij})^{\frac{1}{2}}}\right) + \frac{\left(1+x'^{2}_{ij}\right)^{\frac{1}{2}}}{v^2} g_1 = 0 \qquad (36)$$

where $g_1 = \frac{\partial v}{\partial x}$, together with the boundary conditions,

$$x_{ij}\left(z_{o_i}\right) = x_{o_i} \quad , \; x_{ij}\left(z_{s_j}\right) = x_{s_j} \quad . \qquad (37)$$

In the present formulation, we will treat the velocity $v_{ij}(z)$ on a ray path and $g_2(x,z) = \frac{\partial v}{\partial z}$ as state variables, given by the additional equations,

$$\frac{dv_{ij}}{dz} = x'_{ij}g_1 + g_2 \qquad (38)$$

$$\frac{\partial g_2}{\partial x} = \frac{\partial g_1}{\partial z} \qquad (39)$$

$$v_{ij}\left(z_{s_j}\right) = v_{s_j} \qquad i=1,2, \; \ldots, \; n \qquad (40)$$

$$g_2(0,z) = h_2(z) \qquad z\varepsilon(0,L) \qquad (41)$$

Furthermore, the signal travel time from event 'i' to sensor 'j' is now given by,

$$\Delta T_{ij} = \int_{x_{o_i}}^{x_{s_j}} (1+x'^{2}_{ij})^{\frac{1}{2}}/v_{ij}dx \qquad (42)$$

Then the optimal control problem requires minimization of J in (6) with respect to the "control variables" $g_1(x,z)$, $h_2(z)$, v_{s_j}, $\{x_{o_i}, z_{o_i}, T_{o_i}\}$.

Multiplying equations (36), (38) and (39) by the functions $\lambda_{ij}(z)$, $\xi_{ij}(z)$ and $\psi(x,z)$ respectively, and adjoining them to the expression for J in (6) we get

$$\bar{J} = \sum_{i=1}^{m} \sum_{j=1}^{n} \left\{ \frac{1}{2} \frac{1}{\sigma_{ij}{}^2} (T_{ij}^{obs} - T_{ij}^{calc})^2 \right.$$

$$+ \int_{z_{o_i}}^{z_{s_j}} \lambda_{ij}(z) \left[\frac{d}{dz} \left(\frac{x'_{ij}}{v_{ij}(1+x'_{ij}{}^2)^{\frac{1}{2}}} \right) + \frac{(1+x'_{ij}{}^2)^{\frac{1}{2}}}{v_{ij}^2} g_1 \right] dz$$

$$\left. + \int_{z_{o_i}}^{z_{s_j}} \xi_{ij}(z) \left[\frac{dv_{ij}}{dz} - x'_{ij} g_1 - g_2 \right] dz \right\}$$

$$+ \iint_D \psi(x,z) \left[\frac{\partial g_1}{\partial z} - \frac{\partial g_2}{\partial x} \right] dxdz \qquad (43)$$

Taking variations of $\bar{J}$ as done previously and noting that $\bar{J} = J$ when the system equations are satisfied, we come to the following conclusions.

If $\lambda_{ij}(z)$ and $\xi_{ij}(z)$ satisfy the linear inhomogeneous two-point boundary value problem,

$$A_1 \frac{d^2\lambda_{ij}}{dz^2} - A_2 \frac{d\lambda_{ij}}{dz} + A_3\lambda_{ij} + \xi_{ij} \frac{\partial g_1}{\partial z} = 0 \tag{44}$$

$$\frac{d\xi_{ij}}{dz} + A_4\lambda_{ij} = A_5 \tag{45}$$

$$\lambda_{ij}\left(z_{o_i}\right) = \lambda_{ij}\left(z_{s_j}\right) = \xi_{ij}\left(z_{o_i}\right) = 0 \tag{46}$$

where,

$$A_1 = \frac{1}{v_{ij}} \frac{1}{(1+x'^2_{ij})^{3/2}}$$

$$A_2 = \frac{1}{v_{ij}^2} \frac{dv_{ij}}{dz} \frac{1}{(1+x'^2_{ij})^{3/2}} + \frac{3}{v_{ij}} \frac{x'_{ij}x''_{ij}}{(1+x'^2_{ij})^{5/2}}$$

$$A_3 = \frac{1}{v_{ij}^3} \frac{1}{(1+x'^2_{ij})^{\frac{1}{2}}} \left\{ v_{ij}\left(\frac{\partial g_1}{\partial x} - x'_{ij} \frac{\partial g_2}{\partial x} - \frac{x''_{ij}g_1}{(1+x'^2_{ij})} \right) - 2g_1(g_1-g_2x'_{ij}) \right\}$$

$$A_4 = \frac{1}{v_{ij}^2} \left\{ \frac{x''_{ij}}{(1+x'^2_{ij})^{3/2}} + \frac{2}{v_{ij}} \frac{1}{(1+x'^2_{ij})^{\frac{1}{2}}} (g_1-x'_{ij}g_2) \right\}$$

$$A_5 = (T^{obs}_{ij} - T^{calc}_{ij}) \frac{(1+x'^2_{ij})^{\frac{1}{2}}}{v_{ij}^2}$$

and if $\Psi(x,z)$ satisfies the inhomogeneous initial value problem,

$$\frac{\partial\psi}{\partial x} = \sum_{ij} \delta(x-x_{ij}) \left[H(z-z_{o_i}) - H(z-z_{s_j}) \right] \left[\xi_{ij} + \frac{\lambda_{ij}}{v_{ij}^2} \frac{x'_{ij}}{(1+x'^2_{ij})^{\frac{1}{2}}} \right] \tag{47}$$

$$\psi(a,z) = 0 \qquad z\varepsilon(0,L) \tag{48}$$

then the gradient is given by,

$$\frac{\delta J}{\delta g_1(x,z)} = -\frac{\partial\psi}{\partial z} + \sum_i\sum_j \left\{ \left[\frac{\lambda_{ij}}{v_{ij}^{\,2}(1+x'^{\,2}_{ij})^{\frac{1}{2}}} - \xi_{ij}x'_{ij} \right] \delta(x-x_{ij}(z)) \left[H(z-z_{o_i}) - H(z-z_{s_j}) \right] \right\}$$

$$+ \psi(x,L)\delta(z-L) - \psi(x,0)\delta(z) \tag{49}$$

$$\frac{\delta J}{\delta h_2(z)} = \frac{\delta J}{\delta g_2(0,z)} = \psi(0,z) \tag{50}$$

$$\frac{\partial J}{\partial x_{o_i}} = \sum_{j=1}^{n} \frac{1}{v_{ij}(1+x'^{\,2}_{ij})^{3/2}} \left\{ \lambda'_{ij} + x'_{ij}(1+x'^{\,2}_{ij})(T^{obs}_{ij} - T^{calc}_{ij}) \right\} \Bigg|_{z=z_{o_i}} \tag{51}$$

$$\frac{\partial J}{\partial z_{o_i}} = \sum_{j=1}^{n} \frac{1}{v_{ij}(1+x'^{\,2}_{ij})^{3/2}} \left\{ (1+x'^{\,2}_{ij})^{\frac{1}{2}}(T^{obs}_{ij} - T^{calc}_{ij}) - \lambda'_{ij}x'_{ij} \right\} \Bigg|_{z=z_{o_i}} \tag{52}$$

$$\frac{\partial J}{\partial T_{o_i}} = -\sum_{j=1}^{n} (T^{obs}_{ij} - T^{calc}_{ij}) \tag{53}$$

$$\frac{\partial J}{\partial v_{s_j}} = \sum_{i=1}^{m} \xi_{ij}(z_{s_j}) \tag{54}$$

The two point boundary value problems in the ray equations and the adjoint equations are each equivalent to three first order ordinary differential equations. This formulation is more attractive from the computational standpoint because it requires less storage. We note that the previous formulation resulted in mn additional auxillary parameters $\{S_{ij}\}$ to be estimated, whereas there are no such parameters in the present formulation. However, the assumption that $\{x_{ij}(z)\}$ are single valued functions may prove very restrictive; further, the derivatives $\{x'_{ij}\}$ may become very large and lead to numerical difficulties.

Estimation of Only Some of the Parameters

The foregoing formulae can be directly adapted to the situation where part of the parameters in the estimation problem are known; then the observations of the first arrival times may be utilized to estimate the remaining parameters. For this, the gradient direction derivation can be modified by simply setting in equation (13) the variations of the known parameters to zero and proceeding in an exactly analogous fashion. Then expressions for the gradient of J with respect to the fewer unknown parameters are identical to the corresponding expressions in the foregoing, the gradients with respect to the known parameters being ignored. This leads to a very convenient situation in the numerical study of any given problem, since a single computer program can be used to investigate the effects of considering several alternative sets of parameters as unknowns. We note that when the velocity distribution is known, the adjoint variable $\psi(x,z)$ need not be calculated; and when only $\{T_{o_i}\}$ are to be estimated, none of the adjoint variables are necessary. In fact, in the last situation, the observations are linear in the unknowns and J is quadratic; then only a single solution of each ray equation is required to solve the problem.

Parametrization to Reduce Number of Unknowns

Since the number of observations is limited, it is important to limit the number of unknown parameters in order that the resulting estimates be unique. For this purpose, the unknown functions (e.g., $g_1(x,z)$, $h_2(z)$) should be approximated using a finite number of unknowns. As detailed below, the foregoing results are easily modified to obtain the gradient of J with respect to the finite set of parameters used in the parametrization. Let the finite dimensional representation of $g_1(x,z)$ be given by,

$$g_1(x,z) = \sum_{i=1}^{I} a_i f_i(x,z) \tag{53}$$

Consider the variation δJ in J due to the variation $\delta g_1(x,z) = \sum_{i=1}^{I} \delta a_i f_i$ in $g_1(x,z)$. Then,

$$\begin{aligned} \delta J &= \iint_D \frac{\delta J}{\delta g_1(x,z)} \delta g_1(x,z) \; dxdz \\ &= \sum_{i=1}^{I} \delta a_i \iint_D \frac{\delta J}{\delta g_1(x,z)} f_i(x,z) \; dxdz \end{aligned} \tag{54}$$

Then it follows that the component of the gradient, $\partial J/\partial a_i$ is given by the integral on the right of (54). Thus, any suitable parametrization can be used in conjunction with the method presented here.

Computer Storage Requirements

The computer storage requirements in the present method as proposed, are very severe in general, since the solution of the ψ equation and the evaluation of the gradients requires the storage of all the ray solutions and the corresponding adjoint variables. This problem is further compounded by the fact that a large number of arrival time observations are needed for accurate estimation of the unknowns since each observation contains relatively little information about the unknowns. A reorganization, indicated below, of the gradient calculation can reduce these requirements to the storage of a single ray solution and the corresponding adjoint variables, at the expense of increased computational effort.

Since the adjoint equations associated with a given ray are independent of other rays, and since the expressions for the forcing term in the linear ψ-equation and the gradients involve summation of terms, each of which is dependent only on the solutions of the equations of a single ray and its associated adjoint variables, the gradients can be computed in a cumulative fashion. In this procedure, each data point is treated separately, and its contribution to the gradients is determined using the foregoing formulae specialized to the case where only that datum is available for estimation. The respective gradient components, taking into account all the available data points, are obtained by summing the contributions of each datum computed in this manner (see Figure 1). In this procedure, evidently the computations involved in the solution of the ψ equation and those associated with the parametrization in (54) have to be carried out repeatedly for each observation. Of these two, the latter involves a larger amount of labor since the integral on D has to be evaluated separately for each parameter a_i and because the solution to the ψ equation with a single ray is particularly simple.

<u>The Minimization Algorithm:</u>

The iterative first order conjugate gradient algorithm used for minimization of J is summarized in Figure 1. For convenience, all the unknown parameters are assembled in a single vector $\underset{\sim}{\pi}$, and the estimates are denoted by $\hat{\underset{\sim}{\pi}}$. The search distribution d_i in the parameter space for the i^{th} iteration are determined as follows (Polack, 1973),

$$\underset{\sim}{d}_i = - (\partial J/\partial \underset{\sim}{\pi})^i + \gamma_i \underset{\sim}{d}_{i-1}$$

where

$$\gamma_1 = 0$$

$$\gamma_i = ||(\partial J/\partial \underset{\sim}{\pi})^i||^2/||(\partial J/\partial \underset{\sim}{\pi})^{i-1}||^2 \quad , \; i > 1 \quad .$$

where the superscript i on $\partial J/\partial \underset{\sim}{\pi}$ denotes its value at the estimate of $\underset{\sim}{\pi}$

in the i^{th} iteration. The unidirectional search involves repeated evaluation of J, and thus is the step that accounts for a large fraction of the total computer time required for minimization. The search strategy that we followed involves searching with a constant step size until the minimum is included in the interval between three consecutive points; thereafter this interval is repeatedly divided until the size of the division reaches a prescribed limit. Finally, the minimum is estimated by interpolation of J values at the three consecutive points surrounding the minimum. A judicious choice of the step size and its dynamic alteration during iterations are required to limit the computational effort involved in the minimization procedure.

(1) Start by setting:

$i = 1$

current estimate $\hat{\tilde{\pi}}^i$ = initial guess $\tilde{\pi}_o$

(2) Using the current estimates:

(i) Integrate equations (39), (41) for $g_2(x,z)$

(ii) For each observation:

(a) solve ray equations (36), (37), (38)
(b) calculate travel time and model arrival time
(c) solve adjoint equations (42), (43), (44)
(d) integrate ψ-equations (47), (48)
(e) determine contribution to the gradient $\partial J/\partial\tilde{\pi}$ using (49) - (52)

(iii) Add up the contributions to J and $\partial J/\partial\tilde{\pi}$ by all observations to yield $J^i, (\partial J/\partial\tilde{\pi})^i$

(3) Determine the search direction $\tilde{d}_i$ using the conjugate gradient method.

(4) Search for the first local minimum of J in direction $\tilde{d}_i$, starting with $\tilde{\pi}^i$.

(5) Update the estimates:

$\hat{\tilde{\pi}}^{i+1}$ = the minimum found in the previous step

(6) Check the stopping criterion:

if satisfied, stop;
otherwise, set $i = i+1$ and go to step (2)

Figure 1. The Conjugate Gradient Algorithm for Minimization of J

A Comparison of the Gradient Method with the Least Square Method

First, we shall consider the computational effort required to solve a travel time inversion problem for a three-dimensional domain using the two alternative methods. Since both methods are iterative and the number of iterations required by the two methods for a satisfactory solution are generally different, the computational effort per iteration of either method will be considered.

For each iteration of the gradient method, two compatibility relations between the partial derivatives of the velocity must be integrated. These are mn ray equations and an equal number of linear, two point boundary value problems in the adjoint variables to be solved. Further, for each ray solution, the simple partial differential relations for the two space dependent adjoint variables (equation $\psi(x,z)$) must be integrated with a single forcing term, followed by the evaluation of the contribution by that ray to the gradients of J. The latter includes the effort required in accounting for the parametrization of the velocity partial derivatives. In addition, the unidirectional search in the parameter space must be carried out at each iteration. The search requires repeated simulations of all the observations with a different value of the parameter vector each time, the number of the simulations depending on the search stratgey; in our experience with the two dimensional problems, the average value of γ was approximately 6. We recall that each simulation for a three-dimensional problem consists of integration of two partial differential relations for compatibility and mn ray solutions. The average number of iterations of the gradient method required for the two-dimensional results reported here was 4.

Each iteration of the least squares method (Crosson, 1976) involves the calculation of the matrix of the sensitivity coefficients, followed by the determination of the least square solution of the resulting system of linear algebraic equations. The evaluation of the sensitivity coefficients (partial derivatives) of each of the mn observed arrival times with respect to all unknown parameters requires a large computational effort. However, the sensitivities with respect to the velocity parameters are particularly simple to calculate for the travel time inversion problem, since the variations in the ray path do not influence the travel time to the first order (Backus and Gilbert, 1969); the sensitivity with respect to each of the velocity parameters is obtained by a single quadrature along the ray path. Thus, in addition to the mn ray solutions, pmn such quadratures must be carried out, where p is the number of velocity parameters. The calculation of the sensitivities of each travel time with respect to the three epicentral coordinates requires the solution of three two-point boundary value problems, each consisting of three coupled, second order, linear differential equations for the sensitivities of the ray solution to one epicentral coordinate $\left(\frac{\partial x(\Delta)}{\partial x_o}, \frac{\partial y(\Delta)}{\partial x_o}, \frac{\partial z(\Delta)}{\partial x_o}\right)$. The sensitivities with respect to the origin times can be calculated without significant effort. Finally, if the total number of unknown parameters is N, the least square solution of mn linear algebraic equations in N unknowns must be determined at each iteration.

We note that the computational effort per iteration of the gradient method presented here does not increase significantly with N, whereas that for the least square method does; the effort for the sensitivity calculations increases as N, while that for the least square solution increases as N^3. Furthermore, although the number of times all the ray equations have to be solved during the unidirectional search in the gradient method is large, the effort involved in these repeated solutions can be reduced considerably by using for a given ray, the solution obtain in a previous step as the starting solution. Since the successive steps during the search consist of simulations using the parameter estimates which differ little from each other, this way of initialization leads to a rapid convergence of the procedure for solving the ray equations. While this makes it necessary to store the ray solutions, the storage requirements can be minimized by storing the solutions only at relatively few points along each ray.

In the iterative least square method, the coefficient matrix in the linear, algebraic system of equations is often nearly singular, loading to numerical ill-conditioning. This difficulty must be obviated by modification such as the use of 'damped least squares' method (Levenberg, 1944). On the other hand, the gradient method does not involve this difficulty and yields the smallest iterative corrections in the parameter estimates that lead to a given reduction in J. The gradient method leads to a rapid reduction in J even when the initial guess of the parameter values is grossly in error; however, the rate of convergence slows as the minimum is approached.

Illustrative Example

The results of the formulation with x as an independent variable were applied to a simulated estimation problem in two dimensions. Ten sensors located at the ground level were assumed to observe five seismic events originating at various depths in a rectangular domain 14 km deep and 15 km wide. The location of the sensors and the event origins are shown in Figure 2. The co-ordinates of the locations as well as the times of occurrance of the events are listed in Table 4 under the heading "True Parameter Values."

Details of Parameterization:

The function $g_1(x,z)$ is determined by its values at the node points a large, uniform rectangular grid covering the domain. If the co-ordinates of the corners of a rectangle of the grid are (x_i,z_j), (x_{i+1},z_j), (x_i,z_{i+1}), (x_{i+1},z_{j+1}) with $x_{i+1} > x_i$, $z_{j+1} > z_j$, the value of g_1 at any point (x,z) within the rectangle is given by the interpolation formula:

$$g_1(x,z) = g_{i,j} + \frac{(g_{i+1,j}) - g_{i,j})}{x_{i+1} - x_i}(x-x_i) + \frac{(g_{i,j+1} - g_{i,j})}{z_{j+1} - z_j}(z-z_j)$$
$$+ \frac{g_{i+1,j+1} - g_{i+1,j} - g_{i,j+1} + g_{i,j}}{(x_{i+1} - x_i)(z_{j+1} - z_j)}(x-x_i)(z-z_j) \qquad (55)$$

where $g_{i,j} = g_1(x_i,z_j)$. In the results reported here, D was divided into four divisions in the x direction and five in the z direction, yielding a total of 30 parameters representing $g_1(x,z)$ (Figure 3). The function $h_2(z)$ is parameterized by its values on the large z-grid used above; the value $h_2(z)$ is given by linear interpolation of $h_2(z_j)$ and $h_2(z_{j+1})$ for $z_j \leq z \leq z_{j+1}$. Thus, sic parameters were used to approximate $h_2(z)$.

Alternatively, a different parameterization was also used, such that $g_2(x,z)$ is uniform within a horizontal layer so that

$$g_2(x,z) = h_2(z) = c_i \qquad z_i \leq z \leq z_j \tag{56}$$

In this parametrization the value of $g_1(x,z)$ was held fixed at zero. Then all the changes in estimates of $g_2(x,z)$ are due to those in $h_2(z)$, which in turn is approximated by a piecewise constant (within each layer) function equal to $\{c_i\}$. Thus, the total number of unknowns in this parameterization replacing the functions $g_1(x,z)$ and $h_2(z)$ is equal to the number of layers in D. This parameterization implies and continuous velocity distribution that is independent of x, and piecewise linearly varying in the z direction. In the results reported here, D was divided into five horizontal layers of equal height (Figure 3).

In the sequel we shall refer to these two different descriptions of the velocity distribution as "block" and "layer" parameterization, respectively. Evidently, the layer description entails a much smaller number of scalar unknowns associated with the velocity distribution among the two parameterizations.

Numerical Solution of the Differential Equations:

The ray equations were solved using the program PASVA2 by Lentini and Pereyra which employs a variable order variable step finite difference algorithm. This program has automatic variable order and automatic mesh selection capabilities which make it very suitable to the problem of ray tracing in inhomogeneous media. The two point boundary value problem in the adjoint equation was also solved using PASVA2; since these equations are linear, the convergence is very rapid. The starting guess for the ray solution was taken as the straight line segment joining the epicenter with the sensor location for the iterative process in PASVA2. In all subsequent solutions for a given ray, the ray path of the preceeding calculation, after necessary minor modifications to account for any changes in the epicentral location, was used as the starting solution in PASVA2. For the adjoint two point boundary value problem, the uniform starting guess of zero was used in all gradient calculations.

The z-grid in the starting solution of the ray equation was always taken to be coincident with a smaller uniform grid, which divided the depth L of D into 35 equal strips, with the exception of the first division which extended from z_{o_i} to the next z level on this grid. The non-uniform z-grid output by PASVA2 in the ray solution, which always included the grid of the starting ray solution was used as the starting grid for the adjoint two point boundary value problems, the coefficients and the forcing term at all subsequent stages being determined by interpolation.

The partial differential equations for g_2 and ψ were integrated along the characteristics z=const, using a uniform grid which divided D into respectively 28 and 35 divisions in the x and z directions (Figure 5). The large grid which formed the boundaries of the blocks used for the parameterization of g_1 coincided everywhere with this small grid. The forcing term in the ψ-equation (47) was lumped suitably at the grid point on this smaller grid situated closest to the ray path for a given z. The central differences were used for approximating the derivatives. The smaller grid was also used in calculating the gradients in equations (49) and (50).

<u>Conditions of Simulation</u>:

Five different "true" unknown velocity distributions, labeled D1 through D5, were used; they are shown in Figure 4. All distributions except D2 are uniform in the x direction ($v(x,z) = V(z)$). The distribution D2 varies linearly in both x and z directions; its two sections at x=0 are shown in Figure 4.

The initial estimates of the epicentral locations and the origin times were held fixed for the simulations; they can be inferred from Table 4. The initial estimates of the unknown velocity distribution D1 through D5 are all uniform in x direction; they are also shown in Figure 4.

In all cases reported here, it is assumed that the velocities at the sensor location $\{v_{s_j}\}$ are known exactly and are held fixed at their true values during the estimation process. This assumption is justifiable on the basis that, the velocity distribution at the ground surface is available for direct observations and measurements, using auxilliary experiments. This reduces the number of unknowns in the problem and thus aids a more accurate estimation of the remaining parameters.

In order to investigate the effect on the estimation problem of having good prior estimates of some of the unknowns through some auxilliary data, simulations were also carried out where only part of the parameters were to be estimated. In these cases, the remaining parameters were assumed perfectly known, and their estimates were held fixed at the "true" values during the entire estimation process.

Results and Discussion:

The results will be presented in terms of the estimate errors for the different parameters. The errors in the x and z co-ordinates of the epicenter will be reported separately, but will be combined over all the events. For this purpose, we define:

$$\varepsilon_x = \sum_{i=1}^{5} |\hat{\underset{\sim}{x}}_i - x_i^{true}| \tag{56}$$

and

$$\varepsilon_z = \sum_{i=1}^{5} |\hat{\underset{\sim}{z}}_i - z_i^{true}| \tag{57}$$

The quantities ε_x and ε_z will be reported to indicate the improvements in the estimate (x_i, z_i) of epicenter locations. A similarly defined quantity ε_T will be employed to represent error in the estimates of the origin times of all the events. In real situations, these errors cannot be calculated since the true values are unknown; nevertheless, they will be employed for our simulated problems to obtain a measure of success of the estimation procedure.

The estimation procedure presented here yields the P-wave velocity only along the ray paths, whereas our interest is in obtaining the velocit distribution v(x,z) in D. For this purpose we will utilize the estimated distribution $g_2(x,z)$ and v(x,z) will be determined by integration of the equations:

$$\frac{\partial v}{\partial z} = g_2(x,z) \quad , \; v(x,L) = \text{given} \tag{58}$$

As mentioned previously, the velocity distribution v(x,L) at the ground surface will be assumed known since it can be directly measured. Then solution of (58) yields the distribution v(x,z) of interest. The results of the simulation runs will be evaluated using the total estimate error in v(x,z) as defined by

$$\varepsilon_v = \frac{1}{N_1} \sum_i \sum_j \left| v_{ij}^{est} - v_{ij}^{true} \right| \tag{59}$$

where the summation extends over all the points of the smaller grid, which is used in the integration of equations (39), (47), and (58) and where N_1 is the total number of terms in the summation.

It is evident that the observations do not contain any information about the velocity in those regions along the boundary of D through which no rays pass; hence the evaluation of the estimates is more properly done by excluding such regions while determining ε_v. However, this is difficult since the rays are curved. Hence, we will report an approximate measure $\bar{\varepsilon}_v$, which is determined by excluding from the summation in (59) the lowest layer and the columns of blocks adjacent to the vertical edges of D.

In Table 1 we present the results of simulation using the velocity distribution D1 and both types of parameterizations. The results for several alternative cases where different kinds of parameters are known are included in this table. In all cases except one, the final value of J is significantly smaller than the initial values, which is indicative of the success of the proposed method of matching the model behavior with the actual observations. However, in most cases, the errors ε_v, $\bar{\varepsilon}_v$, ε_x, ε_z in the parameter estimates are relatively large even for the small residual values of J. This is indicative of the ill-conditioning of the problem; for model parameter estimates significantly different from their true values, the model output is very close to the observed output.

The error ε_v does not always decrease with reduction in J, whereas $\bar{\varepsilon}_v$ invariably decreases. This confirms our expectation that in the regions of D close to the boundary, through which no rays pass, v cannot be estimated from the observations of the times of the first P-wave arrivals. The origin times are the most accurately estimated parameters. When the velocity distribution is among the unknowns to be estimated, the estimates of the epicentral locations change very little and ε_x and ε_z remain virtually unaltered from their initial values. This indicates that in this problem, the parameters in decreasing order of influence on the observations are: origin times, velocity distribution, and epicentral locations.

The estimated velocity distribution for D1, when all source parameters are also estimated, is shown in Figure 5. The numbers in the grid give the velocity in km/sec at the top left hand corner of the respective boxes.

To facilitate comparison the "true" values and the initial estimates of the velocity distribution (independent of x) are also shown. For illustration, the ray paths form a single epicenter to two sensors, for the three sets of parameter values corresponding to "true" values, initial guess and final estimates, are included in this figure. The relatively insignificant change in the epicentral location for this estimation problem is evident.

When the velocity distribution is exactly known, the x co-ordinates of the epicenters are more accurately estimated as compared to the z co-ordinates. This can be explained in terms of the geometrical properties of the problem as follows: an error in the estimate of z_{o_i} that places the estimated epicenter above the ture location of the epicenter will yield a reduction in the travel times for all the sensors, the resulting mismatch can be reduced by an estimate of T_{o_i}, that corresponds to a later occurrance of the i^{th} event compared to its true origin time. Thus, the mismatch due to an error in z_{o_i} can be compensated by a corresponding error in T_{o_i} with little increase in J. On the other hand, since each epicenter has sensors on either side in the x direction, an error in x_{o_i} will produce a reduction in the travel times of the signals to some of the sensors, while increasing the rest. Therefore, all these changes cannot be adequately compensated by an alteration in T_{o_i}, and thus any error in x_{o_i} will lead to an increment in J. Thus, when T_{o_i} are unknown, the influence on J of $\hat{x}_{o_i}$ will be greater than that of $\hat{z}_{o_i}$. This possibility of compensating errors in $\hat{z}_{o_i}$ and $\hat{T}_{o_i}$ is illustrated later in Table 5. From this discussion, it follows that when an epicenter lies on one side of all the sensors, its x and z co-ordinates both will be ill-determined if T_{o_i} is also unknown.

A comparison of the results with the two different parameterizations for distribution D1 shows that the estimate errors for all the parameters are approximately equal for corresponding cases. This is an unexpected result, because the number of parameters used to model the v distribution are greatly different (36 vs. 5) in the two cases. This probably indicates that for this velocity distribution, although there are a large number of parameters determining v distribution in the block parameterization, the gradient of J in the region of interest is significantly steep only in a very few directions in the parameter space. Then there are many directions along which no corrections are made yielding large estimate errors.

Table 2 shows the results of simulations using distribution D2 for a limited number of cases. In this case, the layer parameterization yields significantly smaller errors $\bar{\varepsilon}_v$ and ε_T than the block description for the two estimation problems considered, indicating the advantage of using a smaller number of parameters. We note that $\bar{\varepsilon}_v$ is smaller for the layer description, in spite of the fact that this description cannot exactly represent the true velocity distribution D2 (due to the linear viariation in x direction), whereas the block description can. The errors ε_x and ε_z do not reduce appreciably when all the parameters are unknown. Furthermore, the final errors $\bar{\varepsilon}_v$ and ε_T for a given parameterization are not appreciably different whether or not the epicentral locations are exactly known. This clearly shows the insensitivity of the observations to small changes in the epicentral locations.

Table 3 shows the results of simulation using D3. For this distribution, the velocity estimation problem appears to involve numerical difficulties; the ray solution algorithm often failed to converge, perhaps because the slope x' of some rays is very large in magnitude due to rapid variations in the velocity distribution. For such problems, the ray equations in the first form, with the distance s as an independent variable, would be more suitable. Hence, only the results of the estimation problems involving unknown epicenter locations and origin times are reported. As noted earlier, the final error ε_T is much smaller than ε_x and ε_z when both the locations and the origin times are estimated; furthermore, the x co-ordinates of epicenters are better determined than the z co-ordinates. When only the locations are unknown, the final value of ε_z is smaller than ε_x. This indicates that the sensitivity of observations with respect to $\{z_{o_i}\}$ is larger than that with respect to $\{x_{o_i}\}$, which is explained by the fact that for the given geometry and velocity distribution, the change in ray path length from an epicenter to a sensor (and hence the travel time) is more influenced by z_{o_i} than x_{o_i}. When only origin times are unknown, the estimation problem is linear and consequently accurate estimates, as evidenced by small ε_T are obtained.

The occurrance of mutually compensating errors in the estimates $\hat{z}_{o_i}$ and $\hat{T}_{o_i}$ is clearly illustrated in Table 4 by the detailed estimate errors for case No. 1 of Table 3. In this set of data, it is evident that the larger ΔT_{o_i} is, the greater is Δz_{o_i}. On the other hand, the error Δx_{o_i} does not show any definite trend. These detailed estimate errors were compared with those in case No. 2 of Table 3, where only sensor locations are estimated; although the final values of J are approximately equal in the two cases, the errors Δz_{o_i} are much smaller in case 2, and show no apparent trend. This underscores the influence of the unknown origin times, $\{T_{o_i}\}$. These mutually compensating errors point to the possibility of non-unique solutions when both epicenter locations and origin times are to be estimated. When velocity distribution also is unknown, the number of ways in which such errors can occur in the various estimates becomes very large and the degrees of nonuniqueness increases. As the number of parameters characterizing the v distribution increases, implying a greater flexibility in the distribution, this problem of nonuniqueness intensifies.

Tables 5 and 6 present the results of some simulations using velocity distributions D4 and D5. These results reinforce the remarks and conclusions in the foregoing discussion.

Conclusions

1. A new method for matching the model output with the observations is developed. It is applied to several two-dimensional simulated problems with different "unknown" P-wave velocity distributions.
2. When the velocity distribution, and the epicentral locations and the times of origin of the seismic events are to be estimated, the order of decreasing accuracy of the parameter estimates is: origin times, velocity distribution, epicentral locations. In the simulated problems, relatively accurate estimates of the origin times were obtained, velocity estimate errors decreased considerably whereas the errors in epicenter locations did not alter significantly.
3. The estimation problem appears to have nonunique solutions due to the occurrance of estimate errors in different parameters which can lead to little net effect on the observations. For example, when only the origin times and the epicentral locations are to be estimated, an estimate error in the origin time can nullify the effect, on the observations, of an error in the depth of the epicenter if the epicenter lies laterally to one side of all the sensors, such compensation

can also occur with the errors in the estimates of its lateral position. When the velocity distribution is also unknown, the number of possible different solutions that match the observations is much larger than in the problem with known velocity distribution; furthermore, this degree of nonuniqueness increases with the flexibility (the number of scalar parameters) in the model velocity distribution.

P.C. Shah
Research Associate
Department of Civil Engineering
University of Southern California

F.E. Udwadia
Associate Professor
Department of Civil Engineering
University of Southern California

REFERENCES

1. Aki, K., and Lee, W.H.K., 1976. Determination of Three-Dimensional Velocity Anamolies Under a Seismic Array Using First P-Arrival Times from Local Earthquakes, Part I, Journ. of Geophys. Res., 81, No. 23, 4381-4399.

2. Backus, G., and Gilbert, F., 1969. Constructing P-Velocity Models to Fit Restricted Sets of Travel Time Data, Bull. Seis. Soc. Amer., 59, 1407-1414.

3. Bryson, A.E., and Ho, Y.C., 1969. Applied Optimal Control, Ginn Waltham, Massachusetts.

4. Chen, W., Gavalas, G., Seinfeld, J., and Wasserman, M., 1974. "A New Algorithm for Automatic History Matching," Soc. of Petroleum Engg. Journ., 14, No. 6, 593-608.

5. Chevant, G., Dupuy, M., Lemmonier, P., 1975. History Matching by Use of Optimal Theory, Soc. of Petroleum Engg. Journ., 15, 74-86.

6. Crosson, R.S., 1976. Crustal Structure Modeling for Earthquake Data.1. Simultaneous Least Squares Estimation of Hypocenter and Velocity Parameters, Journal of Geophysical Res., 81, 3036-3046.

7. Engdahl, E.R., and Lee, W.H.K., 1976. Relocation of Local Earthquakes by Seismic Ray Tracing, J. Geophys. Res., 81, No. 23, 440-4406.

8. Gelfand, I.M., and Fomin, S.V., 1963. Calculus of Variations, Prentice-Hall, Englewood Cliffs, New Jersey.

9. Lentini, M., and Preyra, V., 1975. An Adaptive Finite Different Solver for Nonlinear Two Point Boundary Problems with Mild Boundary Layers, Stanford University Report STAN-CS-75-530, Computer Science Dept., Stanford, Palo Alto.

10. Levenberg, K.A., 1944. A Method for the Solution of Certain Nonlinear Problems in Least Squares, Quart. Appl. Math., 2, 164-168.

11. Luenberger, D.G., 1973. Introduction to Linear and Nonlinear Programming, Addison Wesley, 133-186.

12. Polack, E., 1973. A Historical Survey of Computational Methods in Optimal Control, SIAM Review, 15, No. 2, 553-584.

13. Udwadia, F.E., and Shah, P.C., 1976. Identification of Structures Through Records Obtained During Strong Earthquake Ground Motion, Journ. of Engineering and Industry (A.S.M.E.), 98, Series B, 1347-1362.

14. Volterra, V., 1959. Theory of Functional and of Integral and Integro-differential Equations, Dover, New York.

15. Wesson, R.L., 1971. Travel-Time Inversion for Laterally Inhomogeneous Crustal Velocity Models, Bulletin of the Seismologic Soc. of America, 61, 729-746.

TABLE 1

Simulation Results

Velocity Distribution: D1

Initial Estimate Errors: $\varepsilon_v = 0.3385$ $\bar{\varepsilon}_v = 0.2409$

$\varepsilon_T = 2.4$, $\varepsilon_x = 2.9$, $\varepsilon_z = 2.6$

	Parameters Being Estimated			J		Final Estimate Errors				
No.	v(x,z)	(x_{o_i}, z_{o_i})	T_{o_i}	Initial	Final	ε_v	$\bar{\varepsilon}_v$	ε_T	ε_x	ε_z
Block Parameterization										
1	x	x	x	9.797	0.187	0.3047	.1527	0.296	2.31	2.94
2	x	-	-	1.557	.0927	.3096	.1492	-	-	-
*3	x	-	x	7.883	5.807	.3713	.1851	2.16	-	-
4	x	x	-	3.149	0.2326	.3409	.1768	-	2.89	2.57
**5	-	x	x	6.776	.0091	-	-	0.323	0.918	2.02
Layer Parametrization										
1	x	x	x	9.797	0.108	0.370	0.193	0.330	2.86	2.55
2	x	-	-	1.557	0.0613	0.359	0.186	-	-	-
3	x	-	x	7.883	.00357	0.326	0.164	0.183	-	-
4	x	x	-	3.149	0.2944	0.411	0.220	-	2.89	2.55

* No further convergence was obtained for this initial guess; perhaps converged to a local minimum.

** In this case, since the velocity is not being estimated, the specification of its parametrization is irrelevant.

TABLE 2

Simulation Results

Velocity Distribution: D2

Initial Estimate Errors: $\varepsilon_v = 0.75$ $\bar{\varepsilon}_v = 0.60$

$\varepsilon_T = 2.4$ $\varepsilon_x = 2.9$ $\varepsilon_z = 2.6$

Parameters Being Estimated				J		Final Estimate Errors				
No.	v(x,z)	(x_{o_i}, z_{o_i})	T_{o_i}	Initial	Final	ε_v	$\bar{\varepsilon}_v$	ε_T	ε_x	ε_z
Layer Parametrization										
1	x	x	x	36.04	0.3496	0.209	0.121	0.26	2.90	2.53
2	x	-	x	27.64	0.1155	0.239	0.129	0.354	-	-
Block Parametrization										
1	x	x	x	36.04	.414	0.314	0.236	0.474	2.89	2.57
2	x	-	x	27.64	1.91	0.340	0.237	1.20	-	-

TABLE 3

Simulation Results

Velocity Distribution: D3

Initial Estimate Errors: $\varepsilon_v = 0$ $\bar{\varepsilon}_v = 0$

$\varepsilon_T = 2.4$ $\varepsilon_x = 2.9$ $\varepsilon_z = 2.6$

Parameters Being Estimated				J		Final Estimate Errors				
No.	v(x,z)	(x_{o_i}, z_{o_i})	T_{o_i}	Initial	Final	ε_v	$\bar{\varepsilon}_v$	ε_T	ε_x	ε_z
1	-	x	x	6.629	0.716×10^{-2}	-	-	0.361	0.957	2.163
2	-	x	-	0.2666	$.637 \times 10^{-2}$	-	-	-	0.498	0.257
3	-	-	x	6.10	0.643×10^{-6}	-	-	0.0004	-	-

TABLE 4

Details of Estimate Errors

Velocity Distribution: D3

Estimation Parameters: $\{(x_{o_i}, z_{o_i}), T_{o_i}\}$

Initial Value of J = 6.629

Final Value of J = 0.716×10^{-2}

Event No.	True Parameter Values			Initial Estimate Errors*			Final Estimate Errors		
	x_{o_i} km	z_{o_i} km	T_{o_i} sec	Δx_{o_i} km	Δz_{o_i} km	ΔT_{o_i} sec	Δx_{o_i} km	Δz_{o_i} km	ΔT_{o_i} sec
1	1.6	2.4	0.6	-0.4	-0.4	-0.4	-0.379	-0.473	-0.104
2	7.4	1.8	0.9	0.3	-0.3	-0.5	0.096	-0.358	-0.061
3	12.8	2.0	-0.5	-0.8	-0.8	0.6	-0.256	-0.279	-0.008
4	4.5	5.0	-0.8	-0.7	-0.5	0.6	-0.121	-0.690	-0.126
5	9.8	3.8	0.3	-0.7	-0.6	-0.3	0.105	-0.363	-0.062

* $\Delta x_{o_i} = \hat{x}_{o_i} - x_{o_i}^{true}$, etc.

TABLE 5

Simulation Results

Velocity Distribution: D4

Initial Estimate Errors: $\varepsilon_v = 2.463$ $\bar{\varepsilon}_v = 2.34$

$\varepsilon_T = 2.4$ $\varepsilon_x = 2.9$ $\varepsilon_z = 2.6$

Layer Parametrization

No.	Parameters Being Estimated			J		Final Estimate Errors				
	$v(x,z)$	(x_{o_i}, z_{o_i})	T_{o_i}	Initial	Final	ε_v	$\bar{\varepsilon}_v$	ε_T	ε_x	ε_z
1	x	x	x	166.5	0.415×10^{-1}	0.206	0.230	0.821	2.877	2.323
2	x	-	x	147.0	$.614 \times 10^{-2}$	0.220	0.260	0.761	-	-

TABLE 6

Simulation Results

Velocity Distribution: D5

Initial Estimate Errors: $\varepsilon_v = 0.50$ $\bar{\varepsilon}_v = 0.40$

$\varepsilon_T = 2.4$ $\varepsilon_x = 2.9$ $\varepsilon_z = 2.6$

Layer Parametrization

No.	Parameters Being Estimated			J		Final Estimate Errors				
	$v(x,z)$	(x_{o_i}, z_{o_i})	T_{o_i}	Initial	Final	ε_v	$\bar{\varepsilon}_v$	ε_T	ε_x	ε_z
1	x	x	x	11.13	0.1723	0.215	0.152	0.371	2.882	2.634

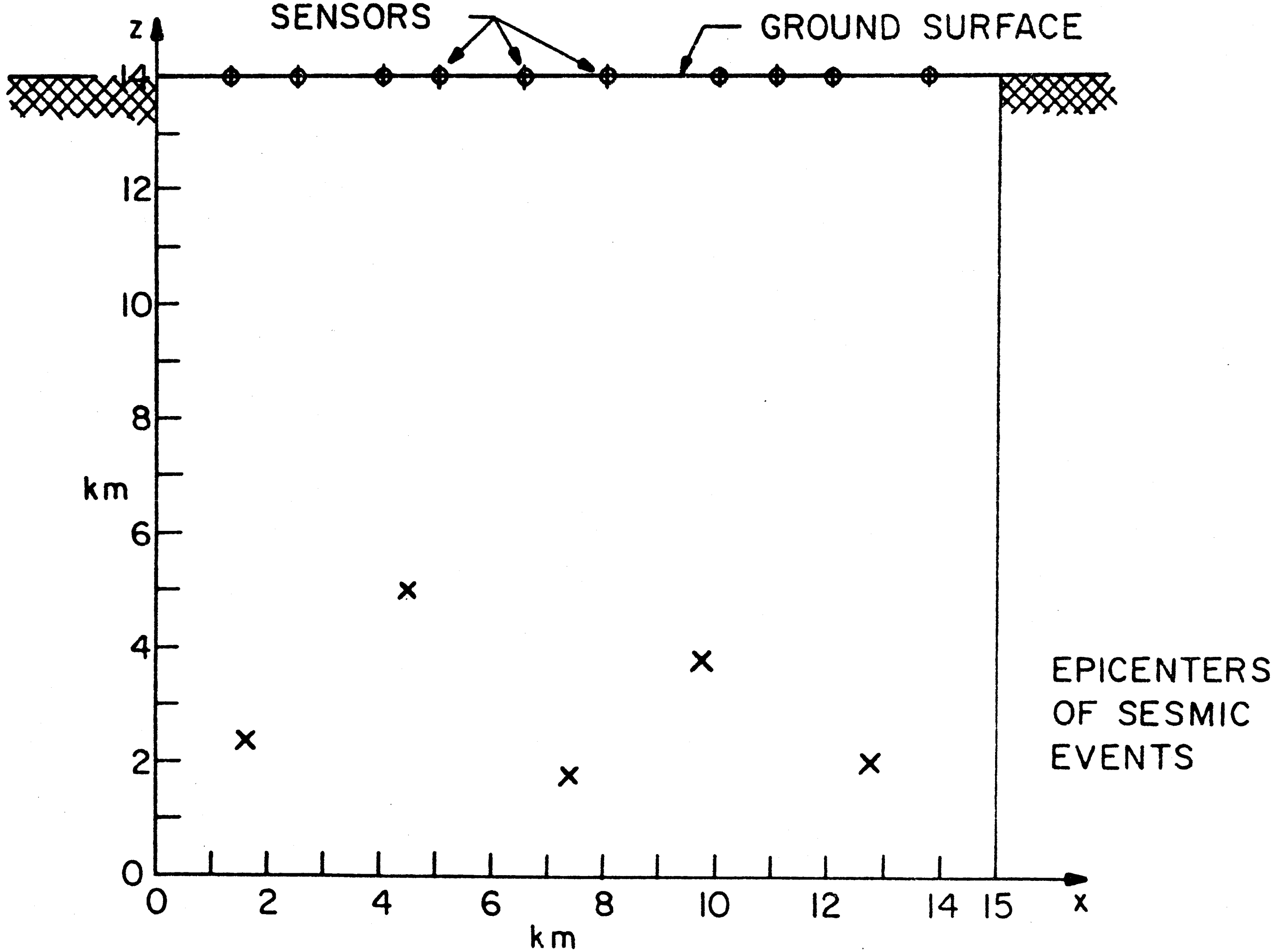

Figure 2

Conditions of the simulated two-dimensional estimation problem

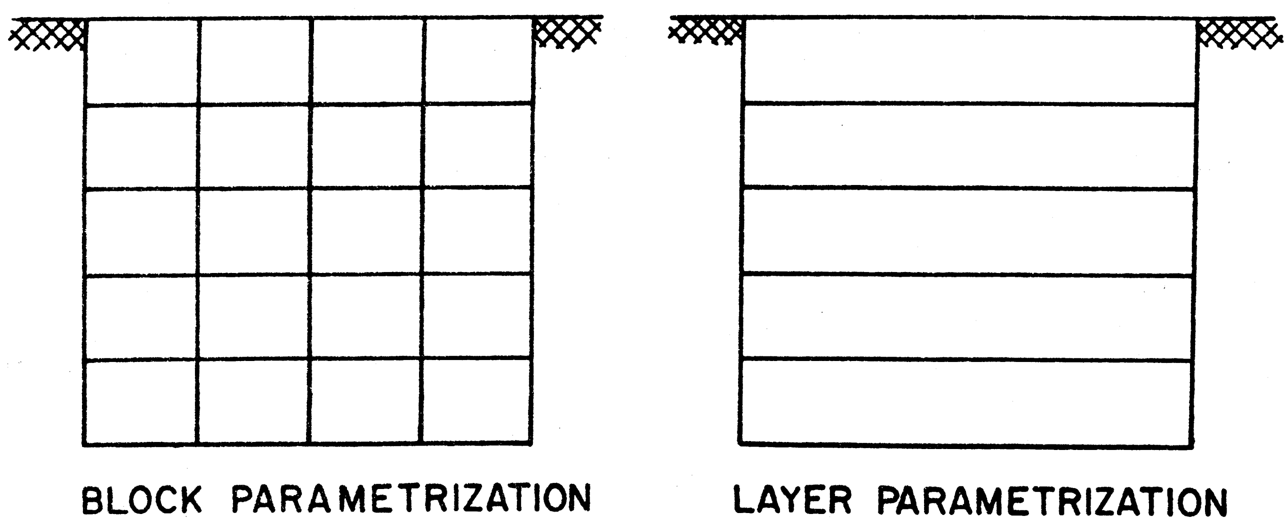

Figure 3

Two different parametrizations of the spatial domain under study

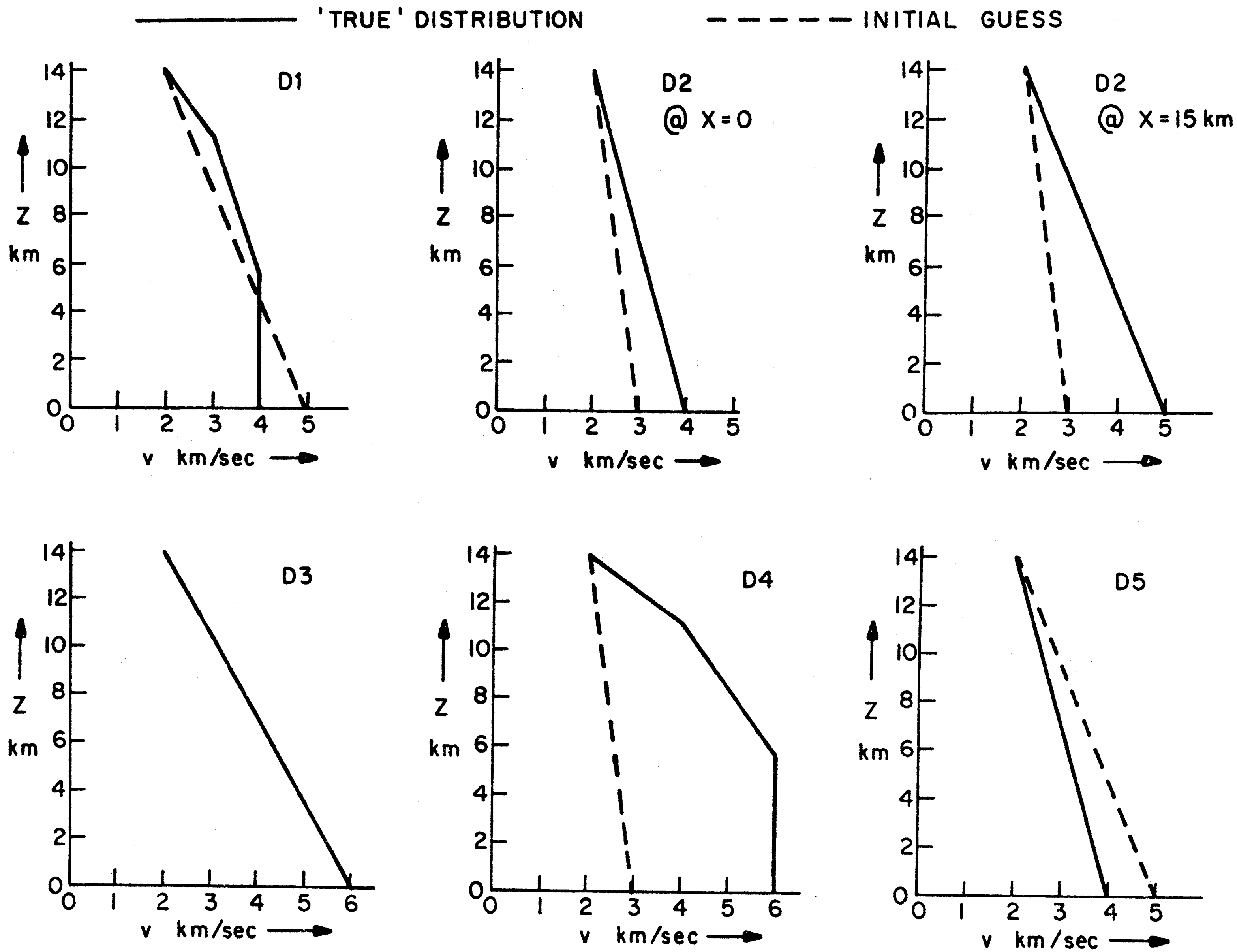

Figure 4

Velocity distributions and their initial guesses used in simulated estimation problems

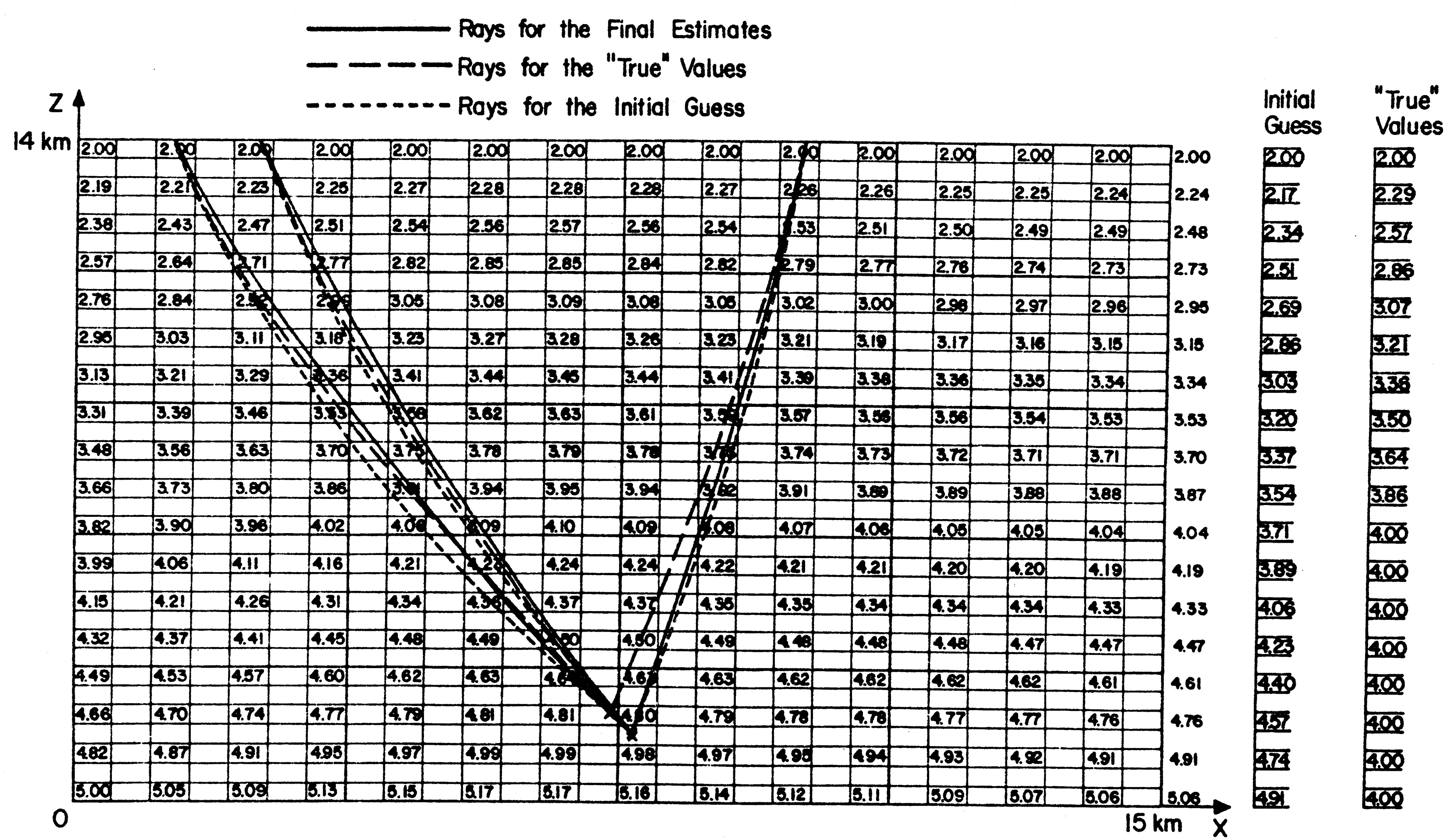

Figure 5

Typical results of identification procedure showing a visual picture of three rays