REVIEWS OF GEOPHYSICS, VOL. 26, NO. 4, PAGES 659-698, NOVEMBER 1988

GEOTOMOGRAPHY WITH LOCAL EARTHQUAKE DATA

Edi Kissling

Institute of Geophysics, Eidgenössische Technische Hochschule (ETH), Zürich, Switzerland

Abstract. The inversion of local earthquake data (LED) for three-dimensional velocity structure requires the simultaneous solution of the coupled hypocenter-model problem. The Aki-Christoffersson-Husebye method (ACH) involves the inversion of large matrices, a task that is often performed by approximative solutions when the matrices become too big, as is the case for most LED, considering the coupled inverse problem. Such an approximate method (herein referred to as approximate geotomographic method) is used to perform tests with LED to obtain the best suited inversion parameters, such as velocity damping and number of iteration steps. The ACH method has been proposed for use of teleseismic data. Several adjustments to the original ACH method, which are necessary for use of LED, have been developed and are discussed. Such adjustments are the separation of the unknown hypocentral from the velocity model parameters for the inversion, the use of geometric weighting and step length weighting, the calculation of a minimum one-dimensional (1D) model as the starting three-dimensional (3D) model for the model inversion, and the display of an approximate resolution matrix (ray density tensors) before the inversion is performed. The ray density tensors allow the block cutting, e.g., the definition of the 3D velocity grid, to better correspond with the resolution capability of the specific data set. The adjustments to the method are tested by inversion of realistic LED of known variance. Synthetic LED are also used to demonstrate the effects of systematic errors, such as mislocations of seismic stations, on the resulting velocity field. Using the data sets from Long Valley, California, Yellowstone National Park, Wyoming, and Borah Peak, Idaho, the effects of improvements to the ACH method and of the data filtering process are shown. The use of the minimum 1D models for routine earthquake location improves this location procedure, as shown with the relocation of shots for the Long Valley and Yellowstone areas. The three-dimensional velocity fields obtained by the ACH method for the Long Valley and Yellowstone areas show local anomalies in the p velocity that can be correlated with tectonic and volcanic features. A pronounced anomaly of low p velocity below the Yellowstone caldera can be interpreted as a large magma chamber. However, the bulk of the paper addresses problems of the inversion method. The LED from the areas mentioned above are used to numerically and theoretically tune the inversion method for the defects that all real data contain. It is shown that one of the most important steps for any inversion of LED is the selection of the data for quality and for geometrical distribution.

1. Introduction

Since the construction of the first seismograph a primary aim of a seismic network is to obtain hypocenter locations for the recorded earthquakes. In order to do so, one has to assume a simplified velocity model of the Earth and to calculate travel times of seismic waves propagating through the model. The remaining differences between the observed and the calculated travel times of the seismic waves are composed of timing errors, of errors in the hypocenter location, of errors in the assumed velocity model or, most probably, a combination of all three sources of errors. Thus, the precision of earthquake location is closely related to the degree of our knowledge of the three-dimensional velocity structure of the Earth. The earthquake data, however, may also be used to determine the velocity field as well as the location of the hypocenters.

Already during the first half of this century, strategies to solve such problems were developed and applied to small data sets (see, for example, Byerley [1939]). It was not until the 1970's, though, that larger computers greatly increased our calculation capabilities such that the inversion of travel times from earthquakes became a powerful tool to study the detailed velocity structure of the lithosphere.

The inversion theory developed by Backus and Gilbert [1967] and improved by Crosson [1976] and others (see section 2) was applied by Aki et al. [1976, 1977] to determine the three-dimensional structure of the lithosphere in southern Scandinavia using teleseismic earthquake travel time data. Recent work shows teleseismic and surface wave data as the primary data base to illuminate the deeper part of the lithosphere and the upper mantle [e.g., Dziewonski, 1984; Woodhouse and Dziewonski, 1984; Humphreys et al., 1984; Suetsugu and Nakanishi, 1985; Nolet, 1981]. Most recently, Jordan [1986] applied the inversion technique to PKP phases of teleseismic earthquake data in order to study the nature of the core-mantle boundary.

However, because of the low temporal frequency content and the somewhat narrow angle of incidence of the different rays at one location at the Earth's surface, teleseismic travel time data have a resolving power limited to rather large structures. While the horizontal resolution of teleseismic data can be improved by a dense station network, the vertical resolution of near-surface structures will remain poor. Obviously, any method to establish a three-dimensional velocity model of near-surface structures should be able to account for known inhomogeneities. Extensive controlled source seismic experiments together with studies of the surface geology have shown that the uppermost part of the lithosphere has lateral variations of p wave velocity, which are beyond the resolution of teleseismic travel time data. With a frequency content comparable to seismic refraction signals the local earthquake travel time data suit the purpose to illuminate the lateral variations of the velocity structure in the upper lithosphere much better.

The inversion of local earthquake travel time data follows approximately the method developed by Aki et al. [1977] and Aki and Lee [1976], with variations and adjustments, which are necessary because of the inherent differences in the input data and the forward and inverse problems. Section 2 gives a brief deduction of the main formulas. In contrast to the inversion of teleseismic travel time data, the use of local earthquake data includes the problem of simultaneously locating the earthquakes while inverting for a three-dimensional (3D) velocity field. This coupling of two inverse

Paper number 88RG03754.
8755-1209/88/88RG-03754$05.00

problems is the main reason for the different inversion procedure proposed for local earthquake data.

Another difference between the original method of Aki et al. [1977] (here referred to as the ACH method for teleseismic data) and this study is the way of calculating the travel time and its derivatives, i.e., the solution to the forward problem. While the incoming wave fronts from a teleseismic event might be represented by either straight rays or rays along arcs in global models followed by a ray in a one-dimensional (1D) model structure beneath the station array, the local earthquake data (LED) require true ray tracing in a layered local one-dimensional or even two-dimensional (2D) model (see section 4). The solution of the forward problem for LED significantly increases the computational burden. An even greater increase arises from the procedure to locate all earthquakes and from the separation of the two coupled inversion problems.

The main advantages of LED over teleseismic data when inverting for crustal structures lies in the higher frequency content and the great variation of the angle of incidence compared with the relatively narrow angle of incidence (± 30°) for teleseismic data. Thus in areas of high seismic activity the upper crust may be illuminated from almost any direction by simply placing seismic stations in the desired locations. The large number of data that may be collected leads to large matrices which have to be inverted. In many cases the resolution of a LED inversion is not limited by the availability of data but rather by the size of the matrices that can be handled by the computer. Thus, the selection of data (preferably without loss of necessary information) and the search for more efficient inversion methods (with little loss in accuracy) play a key role in the inversion of LED. The approximate geotomographic inversion is proposed as a very efficient method. Results obtained with this method will be compared with results obtained with the modified ACH method (see section 2). The selection of input LED is addressed in section 6.

The main task of the method of local earthquake inversion is to illuminate the three-dimensional velocity structure of the upper lithosphere. Since we use LED as our data source, we have to solve the location problem simultaneously which, in itself, depends on the assumed velocity model. Thus, the study of the velocity field with LED will hopefully improve the precision of our earthqake locations. A short discussion of the method used to calculate a starting velocity model best suited for geotomography and for locating the earth-quakes is given in section 3.

The approximate inversion of sometimes inadequate data [Parker, 1977a, b] leads to ambiguous results. This known ambiguity has to be and sometimes can be dealt with by comparison of results from the inversion of LED with those from other geophysical studies (section 7). To diminish the chances of misinterpretation of a resulting three-dimensional velocity field, the error bounds of these results have to be established, and the inversion procedure must be rigorously tested in the first place. This includes the inversion of artificial data sets (section 5) and a thorough error analysis together with a selection of the earthquake data (section 6) used as input for the inversion.

Since seismic networks of close station spacing are very costly to emplace and maintain, LED are by no means abundant in the necessary numbers and of the desired quality for a 3D geotomo-graphic study. However, several data sets from seismically active areas were available. The Long Valley, California, and Yellowstone, Wyoming, areas have been chosen for this study because both areas have been tectonically and/or volcanically active during recent times. Volcanic structures are prime targets for 3D geotomographic studies with LED, since one might expect lateral velocity anomalies of rather large amplitude and short (of the order of several kilometers) wave-length. In both areas, large magnitudes of travel time differences have been found already by teleseismic studies [Steeples and Iyer, 1976; Evans, 1982]. However, particularly the depth resolution of the applied techniques is not sufficient for crustal structures of wave-lengths of 20 km or less [Wielandt, 1987]. In the Long Valley and Yellowstone areas, the inversion of LED illuminates a rather com-plicated crustal structure that can be related to the recent volcanic and tectonic processes (see section 8).

LED inversion is proposed and used as a method to illuminate near-surface 3D structures. Unless a seismically active deep-seated source is available, the resolution of LED for lateral changes in the velocity field below the crust-mantle boundary is inferior to that of teleseismic data. LED may, therefore, be best used to map the 3D crustal structure in order to correct the teleseismic travel times for these near-surface velocity anomalies.

The numerical method proposed in this paper is neither intrinsically different nor superior to other published inversion proce-dures [e.g., Thurber, 1981; Crosson, 1976; Aki et al., 1977; Steeples and Iyer, 1976; Dziewonski et al., 1977; Koch, 1983]. Rather, it was designed to have as much in common as possible with the established inversion methods mentioned above with the least number of adjustments. Because of these technical similarities, the methodological conclusions drawn in this paper can be applied to other inversion techniques as well. In this sense, this paper gives an overview of the problems inherent to inversion of seismic travel time data, and, further, it gives some examples on how to improve this powerful tool for studying the three-dimensional structure of the Earth.

The major goal of this paper is to make the inversion method for seismic travel time data more transparent to anyone interpreting a velocity field derived by such a method. In most cases, critical assumptions (i.e., parameterization of the model both to fit realistic physical properties of the Earth and to allow simple mathematical calculations at the same time!) cannot be tested in their effects by comparisons of resulting velocity fields derived by different inversion methods. The major problem that we face at the moment comes with the interpretation of results obtained by inversion of travel time data and is not the calculation of such results. An advantage of the particular method described in this paper is that it allows one to display intermediate results of the inversion procedure and to test and display the effects of major assumptions on the results and thus make the inversion process more transparent even to a person not involved in the calculation.

2. Method of Inversion of Local Earthquake Data for 3D Velocity Field

As in many other areas of geophysical interpretation, the inversion of LED for 3D velocity structure makes use of a complication in an otherwise simple task. With the intro-duction of sophisticated velocity models the problem of locating earthquakes becomes more complicated and costly, though the solutions obtained might not necessarily be improved. Even the three-dimensional Earth models may still oversimplify the complex, in most parts unknown, velocity structure of the Earth.

The process of locating an earthquake using a number of travel time observations is performed by calculating the hypocenter coordinates in such a way that the observations are best explained, i.e., by minimizing the misfit between calculated and observed travel times. Introducing a 3D model rather than a 1D model in the location process does not change much in principal. The hypocenter location will always be only as good as the timing errors and the estimated velocity model are combined. More velocity model parameters mean a more appropriate ray tracing and thus more realistic travel time

calculations, but more velocity model parameters possibly also mean the introduction of more sources of errors and, worse, errors that by their number might mask mislocations and thus might not be detected. Mislocations of events due to errors in the velocity model parameters are in most cases of a systematic nature and, therefore, are best visible when many earthquakes are located simultaneously and the error vector is searched for systematic areal or time de-pendency. This is called the coupled hypocenter-velocity-parameter inversion and requires the calculation of the best hypocenter and velocity parameters by simultaneously minimizing all travel time errors of a large number of earthquakes.

Numerical methods for solving such least squares problems in seismology are numerous, and many have been described and applied in recent years. The reader is referred to Aki and Richards [1980], Lee and Stewart [1981], Menke [1984], and Nolet [1987] to mention but four basic textbooks on this subject. In this paper I will discuss only some specific properties of the inverse problem of LED and elaborate in detail on the solution formalisms for the so-called ACH and approximate geotomographic methods.

We will also consider only problems with more data points (equa-tions) than unknown geophysical parameters, i.e., we will formulate the problem as overdetermined. The specific properties of the 3D velocity model (see section 2.7) are a direct consequence of this condition.

2.1. Travel Time of a Single Ray

Among the various numerical methods for calculating the travel time between a seismic source and a receiver for a given velocity field (i.e., the so-called forward problem) one of the most efficient methods is the application of ray theory. All travel times in this study are calculated by ray tracing through a one-, two-, or three-dimensional velocity model. The numerical solutions to the forward problems in this study are discussed in section 4 and are of such precision that the calculated travel times may be assumed to be free of numerical calculation errors.

The arrival times of seismic waves generated by an earthquake and observed at some seismic stations are a function of the four hypocentral parameters (origin time and spatial coordinates of hypocenter) and the velocity field sampled along each ray path:

$$t_{obs} = F(t_0, x_0, y_0, z_0, V(x,y,z)) \qquad (1)$$

where

t_{obs}	observed arrival time;
t_0	origin time;
x_0, y_0, z_0	coordinates of hypocenter;
$V(x,y,z)$	velocity field.

In general, neither the hypocentral parameters nor the velocity field are known. Thus the arrival time is the only measurable quantity in (1). However, we may always make an educated guess for the unknown parameters. Using a simple averaging velocity model, we can trace rays from the trial source location to the receivers and calculate a theoretical arrival time:

$$t_{calc} = F(t_0^*, x_0^*, y_0^*, z_0^*, V^*(x,y,z)) = F(h_j^*, m_k^*) \qquad (2)$$

where t_{calc} is the calculated theoretical arrival time, $h_j^* = t_0^*$, x_0^*, y_0^*, z_0^* the estimated hypocentral parameters, and $m_k^* = V^*(x,y,z)$ the estimated velocity parameters.The difference between the observed and the calculated arrival time is called the travel time residual. Obviously, this residual travel ime is a function of the differences between the estimated and the true hypocentral and velocity parameters.

$$t_{res} = t_{obs} - t_{calc} = F(dt_0, dx_0, dy_0, dz_0, dV(x,y,z)) \qquad (3)$$

where t_{res} is the travel time residual and d is the difference between the estimated and the 'true' unknown parameter.

To calculate the hypocentral or the velocity model parameter adjustments (d), we need to know the dependency of all these parameters on the observed travel time. For all parameters except the origin time this dependency is nonlinear. Using a Taylor series expansion, we may rewrite equation (1) as a linear approximation:

$$t_{obs} = t_{calc} + F(h_j^*, m_k^*) = \sum_{j=1}^{4} \delta F/\delta h_j \cdot \Delta h_j + \sum_{k=1}^{n} \delta F/\delta m_k \cdot \Delta m_k \qquad (4)$$

$\delta F/\delta h_j$ (at: $h_j = h_j^*$) is the partial derivative of travel time with respect to jth hypocentral parameter and $\delta F/\delta m_k$ (at: $m_k = m_k^*$) is the partial derivative of travel time with respect to kth velocity model parameter. Using equation (3), we may relate the travel time residual (t_{res})of an observation with the adjustments (d) to the hypocentral and velocity parameters:

$$t_{res} = t_{obs} - t_{calc} = \sum_{j=1}^{4} \delta F/\delta h_j \cdot \Delta h_j + \sum_{k=1}^{n} \delta F/\delta m_k \cdot \Delta m_k \qquad (5)$$

The calulation of t_{calc} and the partial derivatives is called the forward problem, while the calulation of Δh_j and Δm_k comprises the inverse problem.

2.2. The Coupled Hypocenter Velocity Model Problem

The determination of the unknown hypocentral parameters for all earthquakes and the velocity parameters from a set of arrival times is called the coupled hypocenter-velocity-parameter problem. The coupling of the two inverse problems is visible in every single ray path (Figure 1). Given a set of arrival times from one earthquake and using an estimate for the hypocenter location and a starting velocity model, we obtain a system of equations of the form of equation (5):

$$t_i = \delta F/\delta t \cdot \Delta t_0 + \delta F/\delta x \cdot \Delta x_0 + \delta F/\delta y \cdot \Delta y_0 + \delta F/\delta z \cdot \Delta z_0 + \sum_{k=1}^{n} \delta F/\delta m_k \cdot \Delta m_k \qquad (6)$$

where t_i is the ith travel time residual; i=1,...., nobs, where nobs is the total number of observations for this event.

This system of (nobs) equations for one earthquake will be written in matrix notation as:

$$\mathbf{t} = \mathbf{A}\,\mathbf{d} \qquad (7)$$

where **t** is the vector of travel time residuals, **A** the matrix with partial derivatives as elements, and **d** the vector of hypocentral and velocity model parameter adjustments.

If we assume the estimated starting velocity model to be correct and, therefore, neglect the velocity model parameter part in equation (6), we consider the problem which is routinely solved by earthquake location procedures such as HYPO71 [Lee and Lahr, 1975] , HYPOINVERSE [Klein, 1978], or HYPOCENTER [Lienert et al., 1986]. With four unknowns and a larger number of observations (NOBS > 4), this linear system of equations is usually transformed into a least squares problem that may be solved by various numerical methods (see, f.e., Lanczos [1961]). Of the two routines in this program set for inversion of LED (see Figure 9) program VELEST uses the method of Cholesky decomposition, and program RELOC the method of singular value decomposition. With the introduction of the velocity parameters as additional unknowns in equation (7), we need to solve the system of equations simul-taneously for a number of earthquakes. This formulation leads to a total number of observations (nobsto) larger than the total number of unknowns, which has been increased to four times the number of events plus the number of velocity model parameters. Adding an error vector that contains the remaining part of the travel time residual that cannot be explained by adjustments to the unknown parameters, we get :

$$\mathbf{t} = \mathbf{A}\mathbf{d} + \mathbf{e} \qquad (8)$$

now with:

$\mathbf{t}_i$ and $\mathbf{e}_i$: i = 1 ... nobsto,

$\mathbf{d} = (\Delta h_{1,j}, \Delta h_{2,j}, \ldots \Delta h_{nev,j}, \Delta m_1, \Delta m_2, \ldots \Delta m_n)$
j = 1...4

nev total number of earthquakes;

n total number of unknown velocity model parameters;

i index for observation (travel time residual);

e vector containing the timing and linearization errors.

$$\mathbf{A} = \left[\begin{array}{ccc|ccc} \frac{\delta t_i}{\delta h_{1,j}} & 0 & 0 & \frac{\delta t_i}{\delta m_1} & \cdots & \frac{\delta t_i}{\delta m_n} \\ 0 & \cdots & 0 & \frac{\delta t_i}{\delta m_1} & \cdots & \frac{\delta t_i}{\delta m_n} \\ 0 & 0 & \frac{\delta t_i}{\delta h_{nev,j}} & \frac{\delta t_i}{\delta m_1} & \cdots & \frac{\delta t_i}{\delta m_n} \end{array}\right] \qquad (9)$$

2.3. Separation of Hypocentral and Velocity Model Parameters

The matrix of partial derivatives defined in (9) is formally separated into the two matrices:

$$\mathbf{A} = \mathbf{H} + \mathbf{M} \qquad (10)$$

where **A** is the matrix with all partial derivatives as elements,

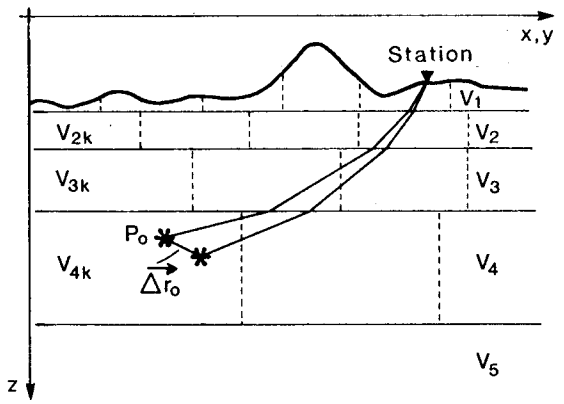

Fig. 1. The coupled hypocenter-velocity model problem illustrated for a single ray path: z, depth; V_n, velocity of nth layer; V_{2k}, velocity of kth block in second layer; P_o, hypocenter (t_o,x_o,y_o,z_o); Δr_o, hypocenter adjustement $(\Delta x_o, \Delta y_o, \Delta z_o)$.

H the matrix with partial derivatives of travel time with respect to hypocentral parameters as elements, and **M** the matrix with partial derivatives of travel time with respect to velocity model parameters as elements.

Of greater consequence than this formal separation into two matrices is the separation of the coupled inverse problem into two inverse problems:

$$\mathbf{A}\,\mathbf{d} = \mathbf{H}\,\mathbf{h} + \mathbf{M}\,\mathbf{m} \qquad (11)$$

with **d** the vector with adjustments to all unknowns as elements, **h** the vector with adjustments to hypocentral parameters as elements, and **m** the vector with adjustments to velocity model parameters as elements.

This formalism was introduced by Pavlis and Booker [1980] and allows us to solve either problem separately, i.e., the location of all hypocenters and the calculation of the velocity model parameters, without, however, neglecting the other part of the coupled problem. Neglecting the coupled inverse model problem while locating the earthquakes may only introduce systematic errors in the hypocenter locations as a result of errors in the assumed velocity model or because a station distribution does not allow the precise location of the hypocenters in some specific areas.

The number of nonzero elements in the matrix **H** is four times the number of events used in the inversion. In order to maintain a system of equations that is overdetermined, the calculation of a large number of model parameters requires the use of a large number of obser-vations. Thus, for a three-dimensional model, the separation of the coupled inverse problem consumes a considerable amount of com-puting time because of the large matrices involved. The numerical method of the parameter separation in this program set for LED follows the original method of Pavlis and Booker [1980]. The se-paration is executed by the two subroutines PARSEP and H12 of the programs CONEQ and TOMOGR (see Figure 9) performing a series of Housholder transformations on the matrix **A**.

Considering the computing time involved, the question about the necessity of parameter separation is important. In all cases where the sources are outside the model area and at greater distances (as is the case for most teleseismic studies), this problem of parameter separation can be approximatley solved by separating for the origin time only. With the use of local earthquake data in the inversion for p wave velocity

structure, four unknowns are introduced (at least theoretically) with each source, and the possible introduction of systematic errors into the resulting velocity parameters cannot be excluded. Similar to the techniques used for inversion of teleseismic data, Hearn and Clayton [1986] introduce source delays as additional unknowns. This is but an approximative solution to the problem. Thurber [1983] proposed a method of iterative simultaneous inver-sion for three-dimensional velocity structure and hypocenter para-meters (program SIMUL3D) that has been used for several appli-cations to local earthquake data [e.g., Eberhart, 1986]. The rapid increase in the size of the matrix to be inverted (four times the number of earthquakes) limits his method, however, to several hundred events. The method proposed here is the separation of the coupled problem into two separate inverse problems which will be solved separately. This separation is optional and may be avoided in special studies. However, tests with artificial data (see section 5) show that the parameter separation is necessary for most studies which use local earthquake data.

2.4. Numerical Solutions to the Separated Inverse Problem

By separation of the hypocentral parameters from the velocity model parameters in the inverse problem we obtain a system of conditional equations which contain only the velocity model para-meters as unknowns:

$$\mathbf{t_S} = \mathbf{M_S}\,\mathbf{m} + \mathbf{e_S} \qquad (12)$$

where $\mathbf{e_S}$ denotes the error vector (of the separated problem) that might be included in the transformed (by separation) travel time residual vector $\mathbf{t_S}$. $\mathbf{M_S}$ is the velocity model matrix transformed by the separation. The separation is performed by multiplying equation (8) by a matrix orthogonal to $\mathbf{H}$ (see equation (10), and Pavlis and Booker [1980]).

The original problem (see equation (3)) of calculating the velocity parameter adjustments ($\mathbf{m}$) from a set of imperfect travel time residuals ($\mathbf{t}$) is nonlinear. We solve the linearized approximation of this nonlinear problem by an iterative procedure, where each iteration step includes the task to solve a system of linear equations (12) and to apply the resulting velocity adjustments to the velocity field.

With $\mathbf{m_k}$ as the solution of the kth iteration step we obtain the final resulting velocity model:

$$\mathbf{m_{final}} = \mathbf{m_{estimated}} + \mathbf{m_1} + \mathbf{m_2} + \ldots + \mathbf{m_k} + \ldots \qquad (13)$$

The matrix of partial derivatives ($\mathbf{M_k}$) and the vector of the travel time residuals ($\mathbf{t_k}$) have to be calculated for each iteration step using the updated (by $\mathbf{m_{k-1}}$) velocity model parameters. A major difference in the inversion procedures for earthquake data proposed by Aki and Lee [1976], Crosson [1976] and Aki et al. [1977] is the calculation of the $\mathbf{m_k}$:

$$\mathbf{m_k} = (\mathbf{M_k})^{-1}\,\mathbf{t_k} \qquad (14)$$

Since there are more equations than unknowns and since our data contains reading errors, an exact solution to (14) is not available. However, we may seek the classical least squares solution. Following Levenberg [1944] the least squares solution for a system of linear equations (12) is given by the normal equations

$$(\mathbf{M}^T\,\mathbf{M})\,\mathbf{m} = \mathbf{M}^T\,\mathbf{t} \qquad (15)$$

and using the notations of the generalized inverse [Lanczos, 1961; Backus and Gilbert, 1968; Jackson, 1972; Wiggins, 1972; Aki et al., 1976], we obtain the solution by inversion of the symmetric matrix ($\mathbf{M^TM}$).

$$\mathbf{m} = (\mathbf{M}^T\,\mathbf{M})^{-1}\,\mathbf{M}^T\,\mathbf{t} = \mathbf{D}^{-g}\,\mathbf{M}^T\,\mathbf{t} \qquad (16)$$

Zandt [1978] compared three methods to solve the problem formu-lated in equation (14): the singular value decomposition of the $\mathbf{M}$ matrix in (12), the singular value decomposition of the $\mathbf{D}$ matrix in (16), and the damped least squares solution as proposed by Aki et al. [1977]. Zandt [1978, p. 60] concluded that "unless the prime interest of the user is the numerical analysis of the method the additional information supplied by the singular value decomposition is not essential," and considering the overwhelming factor of computational cost, the damped least squares solution seems adequate.

2.4.1. ACH-inversion. Following Aki et al. [1977], the damped least squares solution for (12) is given by the normal equations (17) with the diagonal $\mathbf{L}$ matrix containing the damping parameters.

$$(\mathbf{M}^T\,\mathbf{M} + \mathbf{L})\,\mathbf{m} = \mathbf{D}\,\mathbf{m} = \mathbf{M}^T\mathbf{t} \qquad (17)$$

The inverse of $\mathbf{D}$ is guaranteed to exist and, provided large enough damping parameters are chosen, we obtain the velocity adjustments ($\mathbf{m_k}$) of the kth iteration step (13) by

$$\mathbf{m_k} = (\mathbf{D_k})^{-1}\,\mathbf{M_k}^T\,\mathbf{t_k} \qquad (18)$$

Aki [1977] has shown that $\mathbf{m_k}$ obtained by (18) is an approximate solution to the generalized inverse (16), in which the eigenvectors with eigenvalues less than the damping parameters are suppressed. An alternative formulation to the generalized inverse (16) is the stochastic inverse (Franklin, 1970). It can be shown [Aki et al., 1976] that $\mathbf{m_k}$ (18) is also a special case of this stochastic inverse, where the damping parameter is the variance ratio of noise in the data to fluctuation in the model.

The numerical calculation of the ACH solution (18) is handled by the program NORMEQ (see Figure 9) using Cholesky decomposition and poses problems only if the size of the matrix $\mathbf{D}$ exceeds the limitations of the computer. The size of this matrix amounts to the square of the number of unknown velocity model parameters. At present, on a VAX/VMS 8650 computer, 2500 velocity model parameters can be handled by program NORMEQ using some 8 hours CPU time for about 80,000 travel time observations.

Following Aki et al. [1977] a large number of applications of their formalism to various types of geophysical data, such as gravity or seismic data, have led to several different numerical solutions for the inverse problem. The differences are sometimes minor in the forma-lism but might have large effects on the resulting 3D models. Neumann and Behrens [1984] give an excellent comparison of some inversion techniques that evolved from the original ACH method.

2.4.2. Approximative geotomographic inversion. In two cases where this program set has been applied, i.e., Long Valley, California, and Yellowstone, Wyoming, the total number of velocity model parameters that could be resolved

reasonably well by available LED exceeds the limit of 2500 that can be handled by program NORMEQ. Seeking an approximate solution for LED problems with very large numbers of velocity model parameters, Kissling et al. [1984] replaced the symmetric **D** matrix in (17) by a vector (**d**) that contains the diagonal elements of **D** only.The off-diagonal elements of **D** were not calculated and, therefore, are not used.

$$\mathbf{d} = \text{diagonal of } \mathbf{D} = \delta_{li} \qquad (19)$$

where δ_{li} is the element in row (l) and column (i) of **D**, with l = i = 1, ..., number of velocity model parameters. The solution vector ($\mathbf{m}_k$) of the kth iteration step may thus be approximately calculated by simple divisions rather than by full matrix inversion as in (18).

$$\mathbf{m}_k = (\mathbf{M}_k^T \mathbf{t}_k) / \mathbf{d}_k \qquad (20)$$

This approximate solution to the damped least squares formulation (17) is performed by the routine TOMOGR (see Figure 9) and is referred to as the approximate geotomographic solution to the inverse problem in contrast to the ACH solution (program NORMEQ, see Figure 9) that is obtained through matrix inversion.

Neglecting all off-diagonal elements in (**D**) for the calculation of the solution is a crude approximation that is valid only for rows (and columns) in (**D**) that show large diagonal elements relative to their nondiagonal part. Thus the approximate geotomographic solution exaggerates the solutions for poorly resolved velocity model parameters relative to the solutions for well-resolved parameters. This negative effect of the approximate geotomographic solution can be reduced by large damping parameters and by applying step length weighting (see below). Generally, the solutions obtained by the approximate geotomographic method are reached by two or more iteration steps (13) using an increasing damping value that is considerably larger even for the first iteration step than the damping parameter of the corresponding ACH solution.

Applications of the inverse theory to large data sets have been made using methods similar to the approximate geotomographic method in the past. Dziewonski et al. [1977] point out the similarity of their approach to that of Aki et al. [1977]. For computational efficiency, however, they assume [Dziewonski et al., 1977, p. 241] "... that the perturbation in the travel time is associated with that block (in a given shell) in which the ray spends the most time", an approximation that is similar in its effect to the approximate geotomographic approach, provided the geotomography takes into account the blocks hit by the rays in other layers. Hearn and Clayton [1986] use a tomographic back-projection method based on the Jacobi iteration technique to be able to process the travel time residuals on a ray-by-ray basis rather than to invert a large matrix. R.W.Clayton (personal communication, 1984) was also the first to suggest the application of this approximate geotomographic method to large sets of LED [Kissling et al., 1984].

2.4.3. Comparison of the two inversion methods: application to Borah Peak, Idaho. The ACH method can be applied to a smaller data set of local earthquake travel time data from Borah Peak; the results will be compared with the resulting velocity field from the approximate inversion technique (geotomography). To allow the direct comparison of the mathematical formulation of the solutions and the comparison of the resulting velocity fields, all other calculations (as formulated in the previous sections) were the same.

The M=7.0 October 28, 1983, Borah Peak, Idaho, earthquake on the Lost River fault was the largest earthquake since the 1959 M=7.3 Hebgen Lake, Montana, event occurring in the Intermountain Seismic Belt and attracted a great amount of scientific attention to this part of the Rocky Mountains. An overview of the large number of experiments carried out thereafter and an introduction to the tectonic setting is provided by the Redbook of the U.S. Geological Survey (open-file report 85-290, 1985, Menlo Park, California). The 1983 Borah Peak earthquake with its numerous aftershocks offers an excellent opportunity to apply the inversion method for local earthquake data to a nonvolcanic area. The comparatively small size of the data set permits the use of the ACH and the approximate geotomographic methods for inversion, thus allowing the com-parison of these two methods with real data. Beginning the day after the main event, the aftershock series was recorded on a temporary seismic network run by the University of Utah, Salt Lake City. Leu [1986] gives an excellent description of the seismic experiment and of the selection of the data. Using the approximate geotomographic method, she inverted some 4000 p wave travel time observations from 263 events recorded at 72 stations and related the resulting three-dimensional velocity field to the Bouguer gravity anomalies. The results obtained by the more recent version of the program set that performs the approximate geotomographic inversion (see Figure 9) qualitatively match Leu's results. In contrast to Leu [1986] no heavy overdamping was necessary, thanks to the step length weighting (see section 2.6.1), the use of a minimum 1D model as the starting model (see section 3), and the relocation of all events with the resulting 3D model. These features are incorporated in the newer version of the inversion program set for approximate geotomography Thus the amplitudes of the lateral variations in p wave velocity (Figure 2b) differ from the results obtained by Leu [1986].

Figure 2a shows the results from the inversion of the same data set using the ACH method (programs CONEQ and NORMEQ rather than program TOMOGR, see Figure 9). The resulting three-dimen-sional velocity fields (Figures 2a and 2b) are identical down to the level of the expected error bounds of ± 0.5%, if one considers the fact that in the ACH method for the same damping parameter, smaller amplitudes of velocity adjustments are expected as a result of the exaggeration of the diagonal element in the normal equation matrix by the approximate geotomographic method. The spatial distribution of the local anomalies in the resulting velocity field does not change significantly with additional iterations for either method.

2.5. Resolution, Covariance, and Ray Density Tensor

The use of the data variance or variance reduction as a measure of the quality of the resulting velocity model has been in discussion since the first publications appeared on inversion techniques applied to geophysical problems. As Jackson [1976] demonstrated, the most squares technique would be a more appropriate tool to calculate the expected errors in the resulting model. His method, however, is suitable only for a small number of velocity model parameters. Another method to quantitatively determine the error bounds for specific velocity model parameters would be the use of modern statistical techniques such as Bootstrap or Jacknife [Efron and Gong, 1983]. At present, these methods require too much computation time for data sets of the size of those acquired from the Long Valley or the Yellowstone region. The use of array processors and the approxi-mate geotomographic inversion may allow such calculations in the near future. In this study, we still measure the quality of the

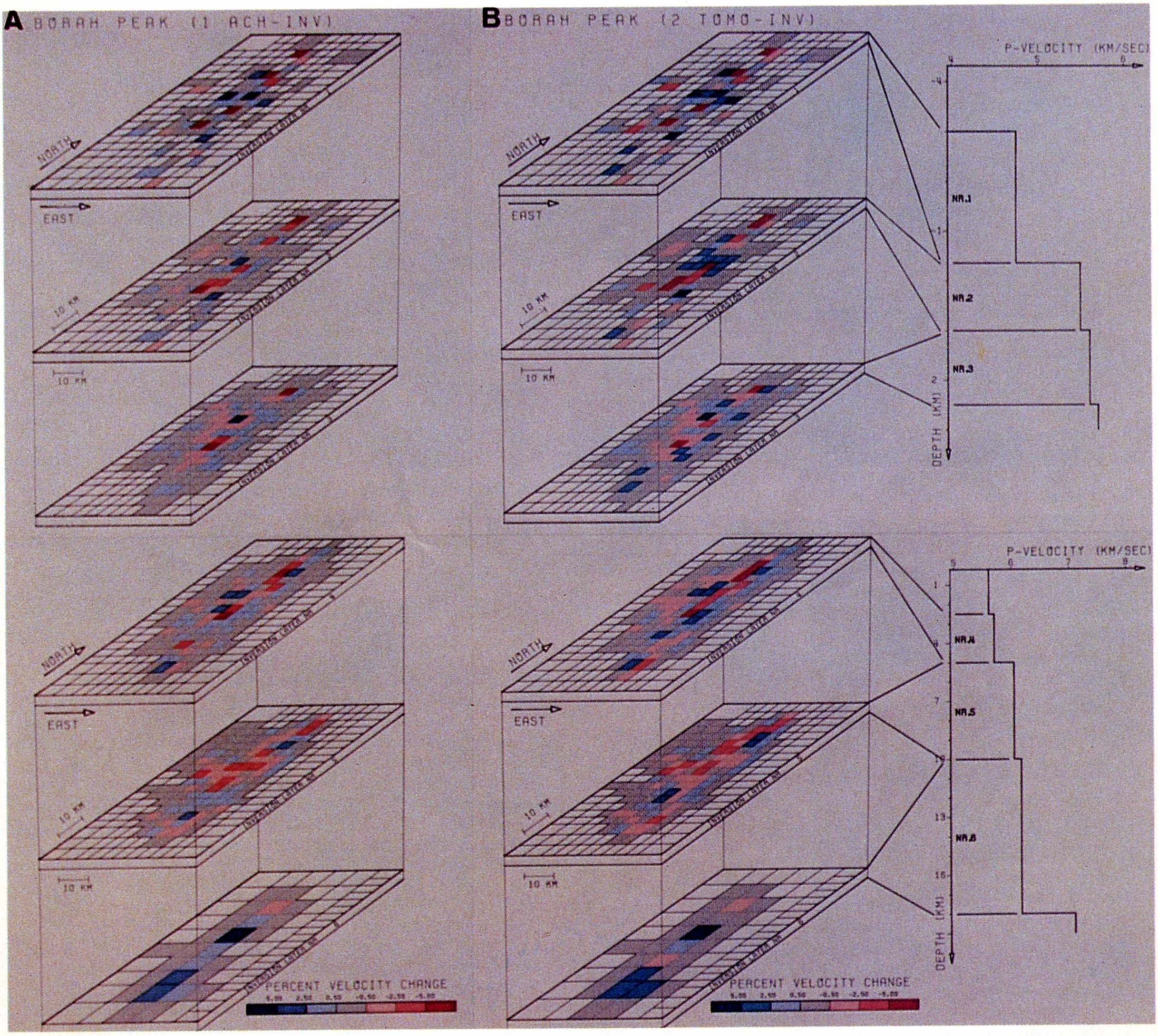

Fig. 2. Borah Peak, Idaho, area. Three-dimensional velocity field deduced from aftershock series of M7.0 October 28, 1983, earthquake. (a) Results of one iteration of the ACH method. (b) Results of two iterations of geotomographic (approximate ACH) method.

resulting velocity model by the reduction of data variance and by the resolution matrix (or the ray density tensors).

The resolution matrix **R** [Backus and Gilbert, 1968; Aki and Lee, 1976] relates the true solution to our calculated (damped least squares) estimate of it.

$$\mathbf{R} = (\mathbf{M}^T\mathbf{M} + \mathbf{L})^{-1}\,\mathbf{M}^T\mathbf{M} \qquad (21)$$

If **R** is the identity matrix, the solution is perfectly resolved. If any diagonal element of **R** is near zero, the corresponding model parameter is unresolved. Diagonal elements of **R** which are zero can and should be avoided by evicting those model parameters from the list of unknowns which are not sampled by the data. This is con-veniently done by use of a hit matrix (calculated by program RESOL, see Figure 9), which contains the total number of rays that pass through each model block, i.e., the total number of data points that refer to each model parameter.

Let us assume a teleseismic problem where **M** is calculated from the travel time spent by each ray within a single layer for the initially unperturbed model. Then the best possible resolution (Aki et al., 1976] corresponds to a diagonal element (ρ_{ii}) for the case of the generalized inverse solution (16).

$$\rho_{ii} = 1. - 1/\mathrm{nbl} \qquad (22)$$

where nbl denotes the total number of blocks sampled in the layer which contains the ith block. The off-diagonal elements (ρ_{ij}) of **R** for this case are given by Aki et al. [1976] as

$$\rho_{ij} = -1/\mathrm{nbl}$$

for the jth block in the same layer as the ith block and

$$\rho_{ij} = 0 \qquad (23)$$

for the jth block in a different layer from the ith block.

The best possible resolution for a block using LED (for the case of the generalized inverse) corresponds to a diagonal element (ρ_{ii}) = 1. With LED it is theoretically possible to resolve the absolute change in velocity (in kilometers per second) for each block, while with the inversion of teleseismic data we can only resolve fractional changes relative to the (assumed) mean velocity of each layer. In practice, however, the strong nonlinear effects of the layer geometry and mean layer velocities on the travel time requires the use of a starting 3D model that is a good approximation to the layerwise average velocity of the resulting velocity field. This starting velocity model can be obtained by a 1D inversion (see section 3: minimum 1D model) prior to the 3D inversion.

The resolution for the damped least squares solution is poorer than that for the generalized inverse, with non-vanishing elements (ρ_{ii}) for the jth block in a different layer from the ith block. However, with the damped least squares method the model standard errors are greatly reduced. Using the assumptions of linearity and linear independency and assuming that the variance is equal for all data, we may express the covariance matrix (**C**) of errors in the model (**m**) by [Ellsworth, 1977]

$$\mathbf{C} = \langle \mathbf{m}\mathbf{m}^T \rangle = \sigma_d^2\,(\mathbf{M}^T\mathbf{M} + \mathbf{L})^{-1}\,\mathbf{R} \qquad (24)$$

with σ_d the standard deviation of travel time data. The standard error of the ith model parameter is

$$\sigma_{mi} = \sqrt[2]{\gamma_{ii}} \qquad (25)$$

with γ_{ii} being the diagonal element of (**C**) corresponding to the ith model parameter. Following Ellsworth [1977] the model parameter errors for the damped least squares solution are always bound by

$$\sigma_m < \frac{\sigma_d}{2 \cdot \sqrt[2]{\mathrm{vtheta}}} \qquad (26)$$

vtheta being the damping parameter, i.e., the diagonal elements of **L** in (24).

An additional measure of the solution quality is the estimated variance improvement for the damped least squares solution.

$$|t_{k+1}|^2 = |t_k - M_k m_k|^2 = |t_k|^2 - m_k^T M_k^T t_k^T - m_k^T L m_k \qquad (27)$$

where $|t_k|^2$ is the data variance after the kth iteration step of inversion procedure and

$|t_{k+1}|^2$ is the estimated data variance after the (k+1)th iteration.

Tests with synthetic data [Ellsworth, 1977] suggest that (27) is a very good approximation which overestimates the actual improvement by less than 5%. Provided the choice of a reasonable damping parameter, our experience with estimated and true variance reduction agrees with the results of Ellsworth [1977].

2.5.1. Ray density tensor. For the maximal number of unknowns (7500 blocks) the full resolution matrix would contain over 50 million elements. Since this would be too large to store, take too long to calculate, and be too much to be controlled, a different approach to displaying the approximate resolution has been introduced by Kissling et al. [1984]. Taking the geometrical distribution of the rays through a block as the criteria for resolution, they found the ray density tensor (**RT**) to be a good qualitative measure for this.

$$\mathbf{RT} = \sum_{n=1}^{\text{all rays}} [\mathbf{V}_n] \qquad (28)$$

$$\text{with} \quad \mathbf{V}'_n = \begin{bmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

($\mathbf{V}'_n$) is given in a local coordinate system with its x axis parallel to the nth ray. The value of (a) is proportional to both the length of the ray within the block and the observation weight. ($\mathbf{V}_n$) denotes the matrix ($\mathbf{V}'_n$) in the global coordinate system. To calculate the ray density tensor, each ($\mathbf{V}'_n$) is transformed into the global coordinate system and is added up.

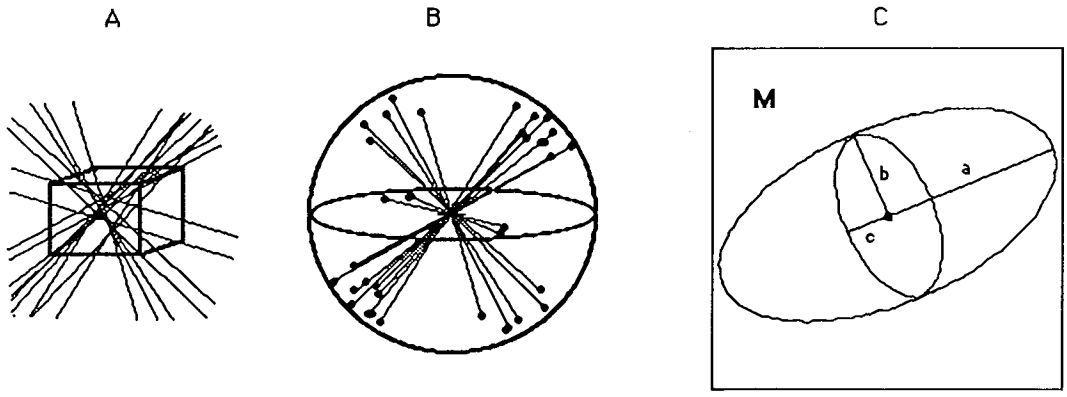

Fig. 3. Ray density tensor after Kissling et al. [1984]. (a) Rays passing through a model block. (b) Display of the directions of the same rays shifted to the center of the block. Dots denote the points of penetration of the sphere. The number of points per surface area on the sphere is called the ray density. (c) Display of the ray density on the sphere of Figure 3b. The main axes (a,b,c) of the ellipsoid are proportional to the ray density. The angle between the vertical and the axis a corresponds to the dip of this axis, i.e., the average dip of the majority of therays passing through the block. The number M corresponds to the negative logarithmus of the ratio of the largest eigenvalue of this block and the largest eigenvalue of the layer (M=-log(a/amax)).

If all the rays passing through a block (Figure 3a) are equally distributed in all directions, or in other words, if the ray density for a block is equal in all directions, a good resolution can be predicted for this block. The ray density distribution for a block can be represented conveniently by a tensor which may be displayed by two ellipses (Figure 3) with the half axis proprotional to the three eigenvalues of the tensor.

The ray density tensors show qualitatively the distribution of the resolution in an area for a given model setup (definition of blocks) and a data set before a costly inversion is performed. Therefore, the ray density tensors are determined by program RESOL (Figure 9) prior to any 3D inversion and can be looked upon as part of the data selection and the process to properly set up the 3D model.

The advantages of the ray density tensor are as follows.

1. It can be easily displayed together with the solution and thus allows the judgement of the credibility of a result.
2. It can be calculated before the actual inversion is done. This allows us to check data and model, i.e., to check the need for more data to resolve a certain area.
3. It is relatively cheap with respect to computing time.

However, as the ray density tensor contains no information about the presence of side lobes in the resolving kernels, it is only the display of an approximation to the full resolution matrix. For example, a block crossed by three orthogonal rays will yield the "optimum" ray density tensor, while the same block cannot be resolved from a neighboring block that is sampled by only one of the three rays.

2.5.2. Comparison of resolution matrix and ray density tensor. Obviously, computing time and core memory permitting, the best way to perform an inversion would be to use the ray density tensors to design the model blocks and check the data and then to calculate the full resolution matrix after the inversion. The data set from the Borah Peak, Idaho, area which was used in this study for the comparison of the two different inversion methods can also be used to compare the resolution vectors and the ray density tensors for specific blocks.

The ray density tensors for inversion layer 4 and 5 (Figure 4; see section 2.7 for definition of "inversion layer") show a good approximate resolution for the central part of the area under study (where the major local p wave velocity anomalies are located), whereas blocks in the peripheral region show poor ray density tensors. For three representative blocks (see Figure 4) with poor, fair, and good ray density tensors the corresponding rows of the resolution matrix (**R**) were calculated using program RESOL2 (see Figure 9) . The results displayed in Figures 5, 6, and 7 show the direct correlation between the ray density tensor and the resolution.

2.6. Damping Parameters

In (17) the diagonal matrix **L** contains the velocity damping parameter (vtheta) that has to be large enough to guarantee a reasonable solution of $\mathbf{D}^{-1}$ (18) and small enough not to suppress all velocity adjustments. Using the formalism of the stochastic inverse [Franklin, 1970] Aki et al. [1977] and Aki [1977] suggest a choice for the damping parameter as

$$\text{vtheta} = \frac{\text{sigdat}^2}{\text{sigmod}^2} \tag{29}$$

where sigdat^2 is the variance of the data, and sigmod^2 is the root mean squared velocity fluctuation (variance) of the true model. For applications of the ACH method to teleseismic data the damping parameter varies from 0.003 (s per cent)^2 [Oppenheimer and Herkenhoff, 1981] to 0.005 [Zandt, 1978] and to 0.02 [Aki et al., 1976] with an estimated data variance of 0.01 s^2 for all three studies.

The data variance for LED in general may be assumed to equal the variance of teleseismic data. However, for any specific data set the variance should be tested before the inversion (see section 6 and program DATVAR, Figure 9). Sigmod^2 has to be estimated mainly on the basis of information from surface geology and results from controlled source seismology. In an area of recent volcanic activity such as Long Valley, California, and Yellowstone, Wyoming, we can expect larger amplitudes of the p wave velocity anomalies but of smaller lateral extensions than, for example, in an area like Borah Peak, Idaho. In any case, the chosen value of the damping parameter (vtheta) is but one of the several assumptions necessary to perform an inversion that should be tested by doing several such inversions on the same data set while the assumed values are changed.

2.6.1. Step length weighting. The calculated velocity adjustments ($\mathbf{m}_k$ in equation (20)) that are applied to the velocity field are called the step lengths of this iteration.

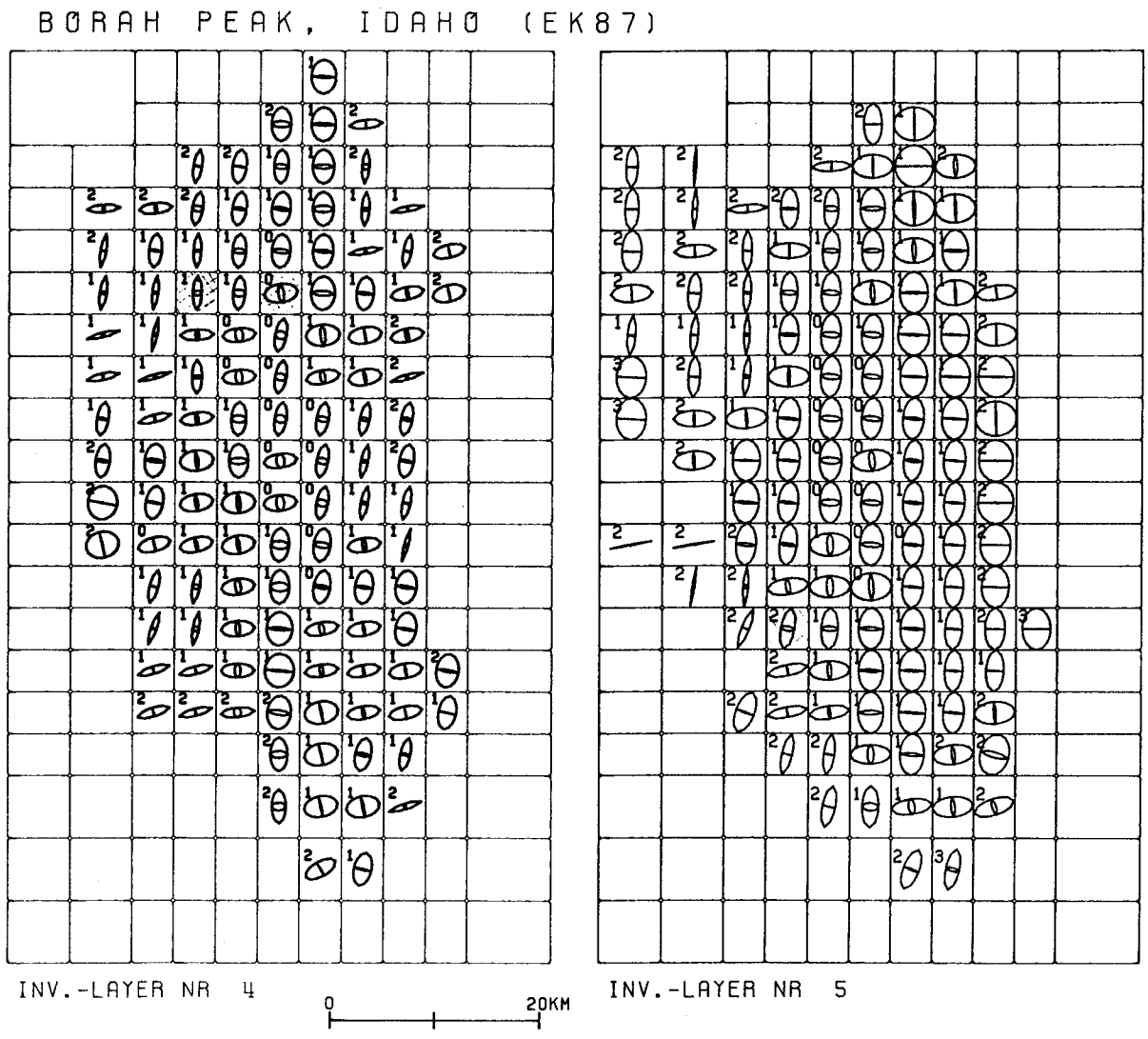

Fig. 4. Ray density tensors (see Figure 3c) of inversion layers 4 and 5 for Borah Peak, Idaho. The locations of the three representative blocks with poor (block 1012), fair (block 707), and good (block 709) ray density tensors are marked by shading.

Poorly resolved velocity model parameters usually get larger adjustments than well-resolved blocks, and these adjustments are often reversed in the following iteration step. For faster convergence, poorly resolved blocks should be allowed only a fraction of the calculated adjustments. For this purpose we introduced the step length weighting (slw). The fraction of the calculated model adjustment which is applied to the parameter at each iteration step is directly proportional to the resolution of this block. The resolution is replaced by the ray density tensor of the block and the ones of its neighboring blocks. The step length weighting is defined by

$$\frac{\text{total weighted ray length}}{\text{size of block (area)}} = \frac{RLT}{F} =$$

$$= \text{normalized ray length} = NRL \qquad (30)$$

$$lw = 1 - \frac{1}{NRL} \qquad 0<lw<1, \text{ for } NRL>1 \qquad (31)$$

$$dw = \frac{2*\frac{e_2}{e_1} + \frac{e_3}{e_2}}{3} \qquad 0<dw<1 \qquad (32)$$

where e1,e2,e3 are the ordered eigenvalues of the ray density tensor of the ith block.

$$slw_i = lw_i * dw_i \qquad \text{(for ith block)} \qquad (33)$$

Since the resolution of each block depends heavily on the resolution of its direct neighbors, but the ray density tensor does not reflect this dependency, the final step length weight for one block (slwtot) is a weighted sum over its own step length (slw_i) and the step length weight of its 26 direct neighbors (slw_k).

$$slwtot_i = 0.5*slw_i + \frac{\sum_{k=1}^{26} slw_k}{52} \qquad (34)$$

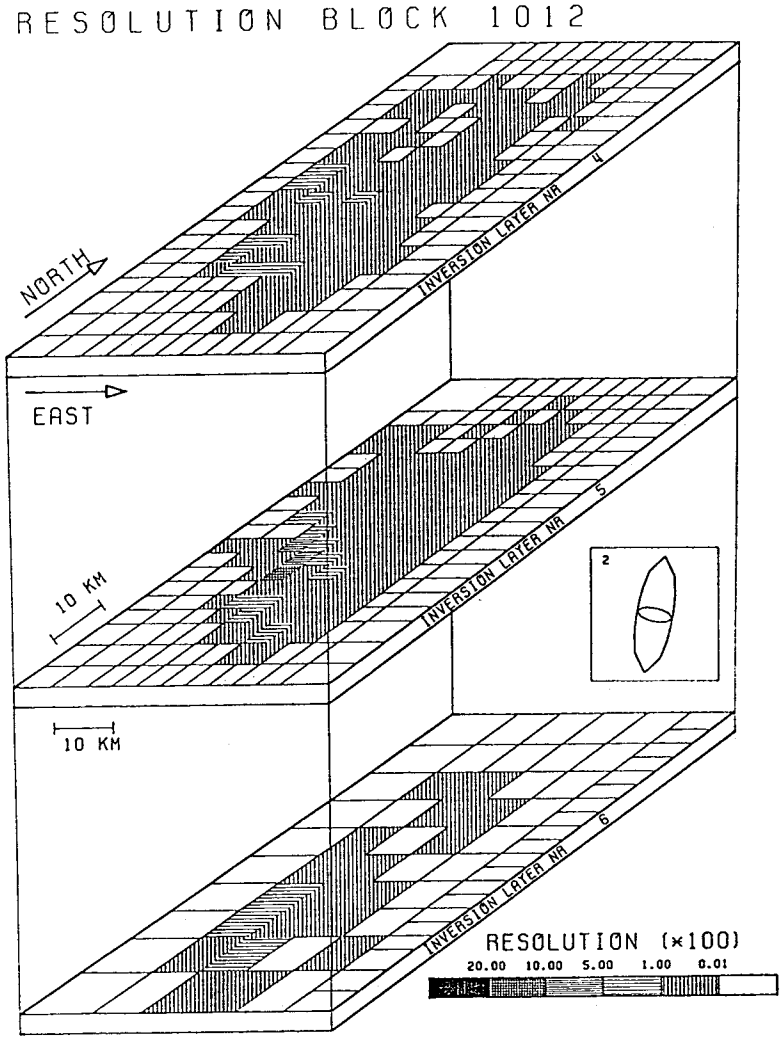

Fig. 5. Ray density tensor and row of resolution matrix for block 1012, representing poor resolution for the Borah Peak, Idaho, area. The diagonal element of the resolution matrix for this block reaches a value of 0.15 only.

Tests with artificial data (see section 5) show the advantage of the use of step length weighting for approximate geotomographic inversions. However, the costs in computing time are again considerable, since the reduction of the step length of an average resolved block may have to be compensated by an additional approximate geotomographic iteration. The effects of step length weighting are similar to those of defining a lower bound for the eigenvalues in the normal equation matrix in order to invert for a velocity model parameter. While the equation for an unknown with an eigenvalue smaller than the predefined lower bound is removed from the system of equations before the actual inversion is performed, the step length weighting is applied after the inversion. Additional tests with artificial data sets showed the step length weighting to be an unnecessary computational burden in the case of full matrix inversion.

2.7. Parameterization of 3D Model

To start an inversion of LED, an initial model for the velocity structure is required. The characteristics of the starting model, e.g., the representation of the 3D velocity field either by a continuous function or by a discretized model, are determined by the applied formalism (see, for example, Burmakov et al. [1984]). The lateral inhomogeneities in the p wave velocity field manifest themselves in travel time residuals determined with respect to the travel times corresponding to a reference velocity model. In most cases, this reference velocity model is laterally homogeneous; i.e., it is a one-dimensional velocity model. Following the initial approach by Aki and Lee [1976] and Aki et al. [1976, 1977] to parameterize the lateral inhomogeneities, we divide each layer into blocks of uniform slowness.

The block model should encompass the seismograph array and all epicenters.The minimum depth of the model has to be equal to the maximum depth of the deepest ray path in the data set. If secondary arrivals are used later, we suggest introducing several additional layers at depths below the deepest event and testing the model for rays from a source at even greater depths up to the maximum distance of the study area. This test can be performed by standard seismic refraction ray tracing programs. This procedure will ensure that no ray of the data set will ever exit the model area even after subsequent hypocenter location and velocity parameter adjustments.

The use of a priori information and the exploration of a range of initial block models are vital for the inversion of LED. Among the best suited a priori information for

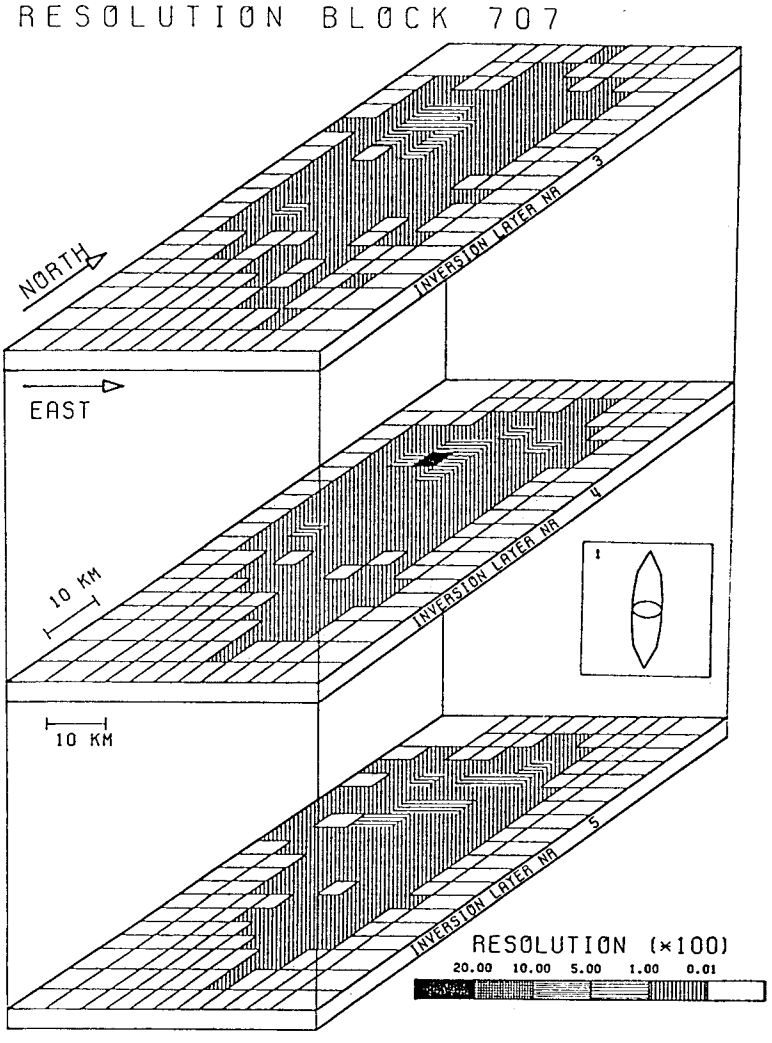

Fig. 6. Ray density tensor and row of resolution matrix for block 707, representing fair resolution for the Borah Peak, Idaho, area. The diagonal element of the resolution matrix for this block reaches a value of 0.35.

inversion of LED is refraction seismic data. One way to incorporate refraction seismic data is to add the known velocity data to an augmented matrix [Benz and Smith, 1984]. in this study, the refraction seismic data is primarly used to set up the geometry of the starting model for comparison with the tomographic results. The layers of the starting model are chosen primarily according to a priori information of the average structure and have to be as many as are necessary for a precise modeling of the travel path of a wave by 2D ray tracing. The best 1D starting models are almost exclusively those which have been tested with the help of 1D inversion programs such as VELEST and a large subset of the data which will again be used in the 3D inversion. The one-dimensional starting velocity model is used for geometrical reasons to assure horizontal layer boundaries. These layer boundaries are chosen such that a fast but still reasonably accurate tracing of the ray with refraction points at each layer boundary are included. For this reason the original layers (called seismic layers) vary in thickness and are sometimes too thin to be used as the vertical extension of the inversion blocks. Therefore, the model allows the combination of several seismic layers to one "inversion layer".

The use of blocks of uneven size (Figure 8) reduces the computation time for the inversion, since areas with fewer rays might be cut into larger blocks. Though, at the end, we will know less about the internal structure of an area with larger blocks, the obtained results are more reliable. Uneven block size might also be introduced to reflect the expected lateral variation of the velocity within a geologic unit (e.g., in Figure 8 the Sierra Nevada batholith in the area of Long Valley, California). Such assumptions, however, should always be tested by several inversions of the same data set using different sizes of blocks. Christoffersson and Husebye [1979] in their study of teleseismic data in Scandinavia used the term "option for geological modeling" for basically the same task that is performed by the option for uneven block size in this study.

The number of velocity parameters in a 3D model for inversion of local earthquake data is limited to 7500 at the moment (7500 blocks for approximative geotomography by program TOMOGR and 3000 for full inversion with the ACH method by CONEQ-NORMEQ, see Figure 9). The geological complexity of a specific area, however, might demand rather small blocks in some areas, thus leading to a total of more than 7500 blocks. The combination of several small blocks into a larger block, on the other hand, might be preferable for improvement of resolution (small total number of rays passing through some parts of the area under study) in a specific area, and the same might be done simply to reflect a large tectonic unit or to reduce the total number of blocks to less than 7500.

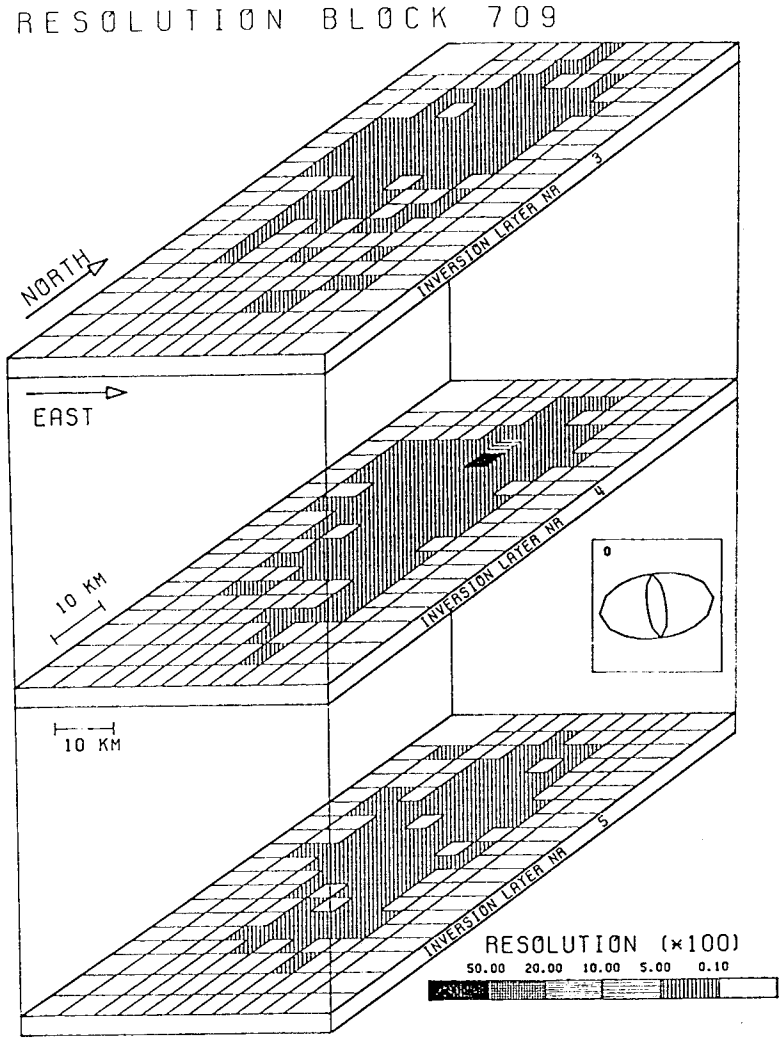

Fig. 7. Ray density tensor and row of resolution matrix for block 709, representing good resolution for the Borah Peak, Idaho, area. The diagonal element of the resolution matrix for this block reaches a value of 0.55.

This combination of small blocks is done with the interactive routine MODINP (see Figure 9). Care should be taken in laying out the block grid for the uppermost crustal layers to ensure that stations with dissimilar residual patterns or mean residual values (see section 3) are not grouped together in the same first layer block.

Within each inversion layer the blocks of uneven size may be combined before the inversion. This combination may differ from inversion layer to inversion layer. It is possible to combine poorly resolved blocks after the calculation of the ray density tensors. Since the velocity for some specific areas might be well known, it is also possible (with the interactive routine MODINP) to set the velocity for any inversion block and to hold it fixed throughout the inversion procedure.

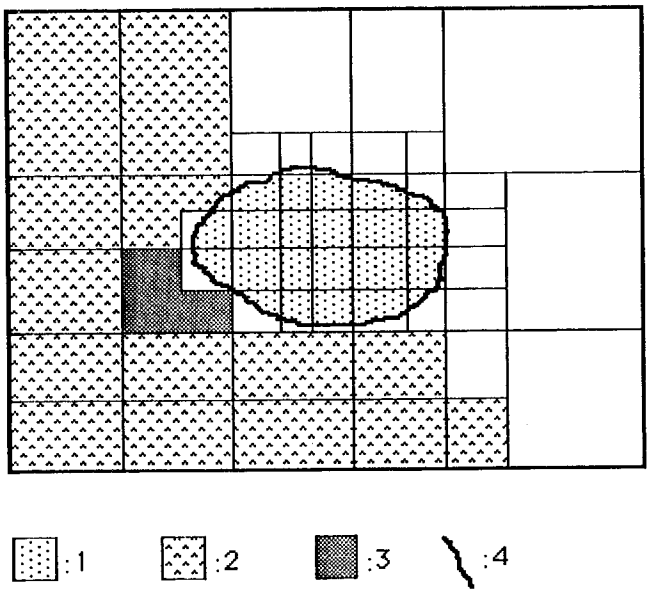

Fig. 8. Definition of inversion blocks in area with known large differences in the near-surface velocities, such as Long Valley, California: (1) Sediments of caldera with low p velocities, (2) Sierra Nevada, outcropping basement rocks, (3) Combination of three seismic blocks to inversion blocks for geometric reasons, (4) caldera bounding fault system.

2.8. Computer Programs to perform ACH and Approximate Geotomographic Inversion of LED

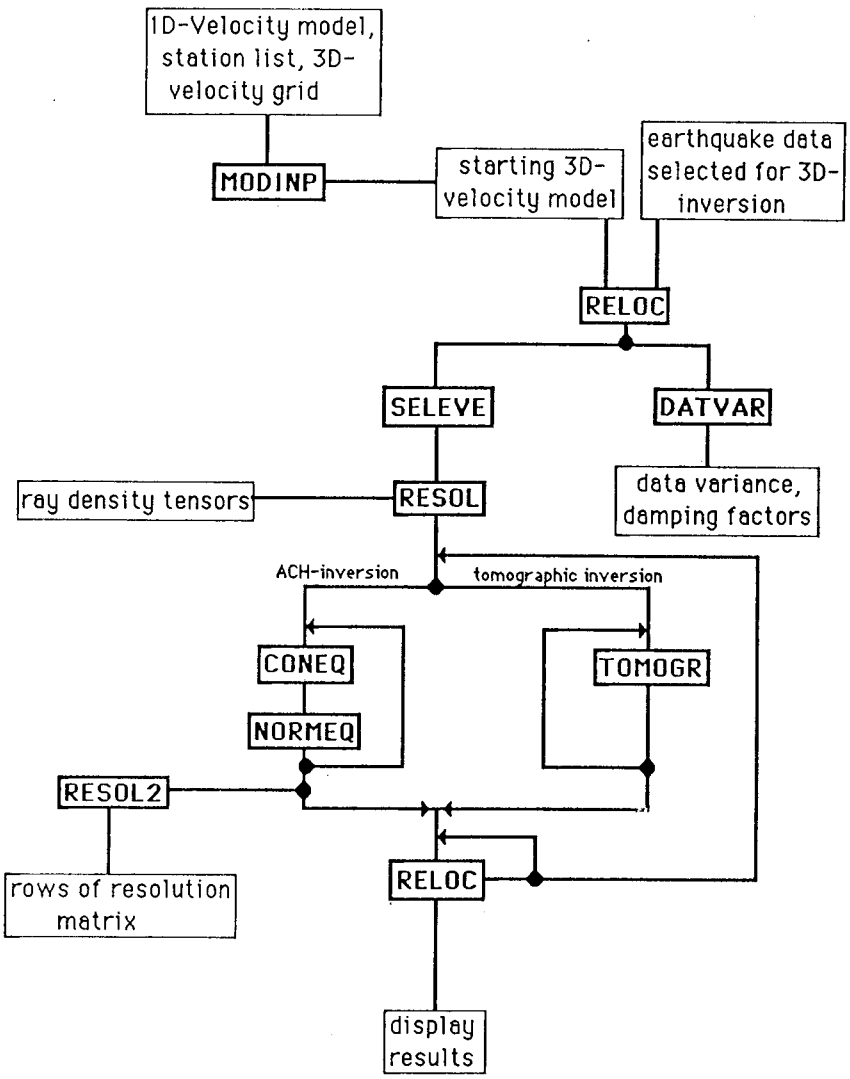

Fig. 9. Flow chart for inversion of local earthquake data to obtain a three-dimensional velocity field.

3. The Minimum 1D Model: Starting Model for 3D Inversion and Best Model for Hypocenter Location Procedure

In the 3D model the velocity field is represented by a three-dimensional grid of uneven grid spacing, superimposed on a (geometrically) one-dimensional Earth model. Except for a priori known lateral velocity anomalies, the starting 3D model is basically one-dimensional. This 1D model, which is used as the starting model for the inversion of LED to calculate the 3D velocity field, has to match the following conditions:

1. The slowness of each layer has to be the areawise weighted average of the slownesses found at that depth interval.
2. The depth of the layer boundaries and the velocities should account for the different phases observed in the data.
3. It should be possible to locate earthquakes throughout the area covered by the station array with equal precision.
4. The 1D model with its corresponding station corrections should reflect the basic features of surface geology and, where available, the overall features of the crustal structure. In Long Valley, f.e., the stations inside the caldera should all show on average delayed arrivals relative to the stations outside, since the near surface sediments of the caldera fill have a lower p wave velocity than the granite of the surrounding terrain [Hill et al., 1985].

Though representing only a crude first approximation to the real Earth, the starting 1D velocity model is of great value for the purpose of routinely locating earthquakes with a seismic network, if this model is a minimum 1D model. Such a 1D model and the corresponding station corrections are the result of a combined, simultaneous inversion of a large number of selected high-quality events for the 1D velocity model parameters and the hypocenter locations. The attribute "minimum" denotes the fact that this 1D model and the station corrections lead to a minimum average (rms) value for all earthquakes used in the inversion. Since this problem is ambiguous, usually several different models with about the same residual variance and location precision can be found. The model which coincides best with surface geology and the a priori information about the near-surface structure is then chosen as the final minimum 1D model.

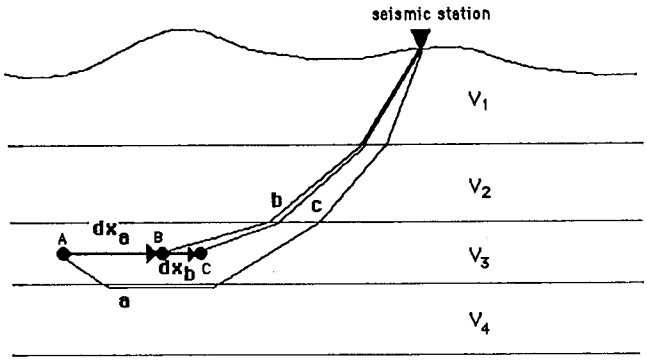

Fig. 10. The nonlinearity in the inverse problem of hypocenter location. V_k: velocity of the kth layer; A,B,C: consecutive hypocenter locations; dx_a: shift in hypocenter location for nth iteration step, from location A to B; dx_b: shift in hypocenter location for (n+1)th iteration step, from B to C; a,b,c: rays from current hypocenter location to a seismic stationwhich follow the fastest travel path. These rays are used to calculate the partial derivatives of the travel time with respect to the hypocenter parameters. Ray a is used for the nth iteration step; ray b is used for the (n+1)th iteration step.

The calculation of such a 1D model in conjunction with the checking and selection of the earthquake data is a tedious procedure and involves almost 50% of the overall computation time for the 3D inversion. However, the calculation and testing of the minimum 1D model is an integral part of the 3D inversion, since the linearization of the inverse problem is based in part on the assumption of well-located events. This includes the assumption that the hypocenter adjustments after the 3D velocity inversion (with program RELOC; see Figure 9) have normally only minor effects on the geometry of the ray paths (Figure 10). All events should, therefore, be relocated using the starting velocity model prior to the 3D inversion. Any large systematic change in the hypocenter locations after the 3D inversion for the velocity field is suspicious. Though few events might show large changes in location, on average we expect to see only small changes, similar to the average changes in the layer velocities which should remain small. The relocation vectors for the events from the Long Valley and Yellowstone areas were in average about 200 m in length after the 3D inversion. This is well within the estimated error bounds for the location procedures.

3.1. Calculation of 1D Model

The geometry of the minimum 1D model is based on a priori information, preferably on refraction seismic profiles which allow the identification and correlation of the different phases observed in LED. Studies on LED from various areas (Long Valley, Morgan Hill - Calaveras fault, Yellowstone) show the need of seismic refraction data for the construction and specifically the testing and selection of the 1D models.

The determination of a minimum 1D model starts with the selection of high-quality earthquake data (Figure 11) and the compilation of a priori information about the crustal structure. To simulate velocity gradients and in order to approximate

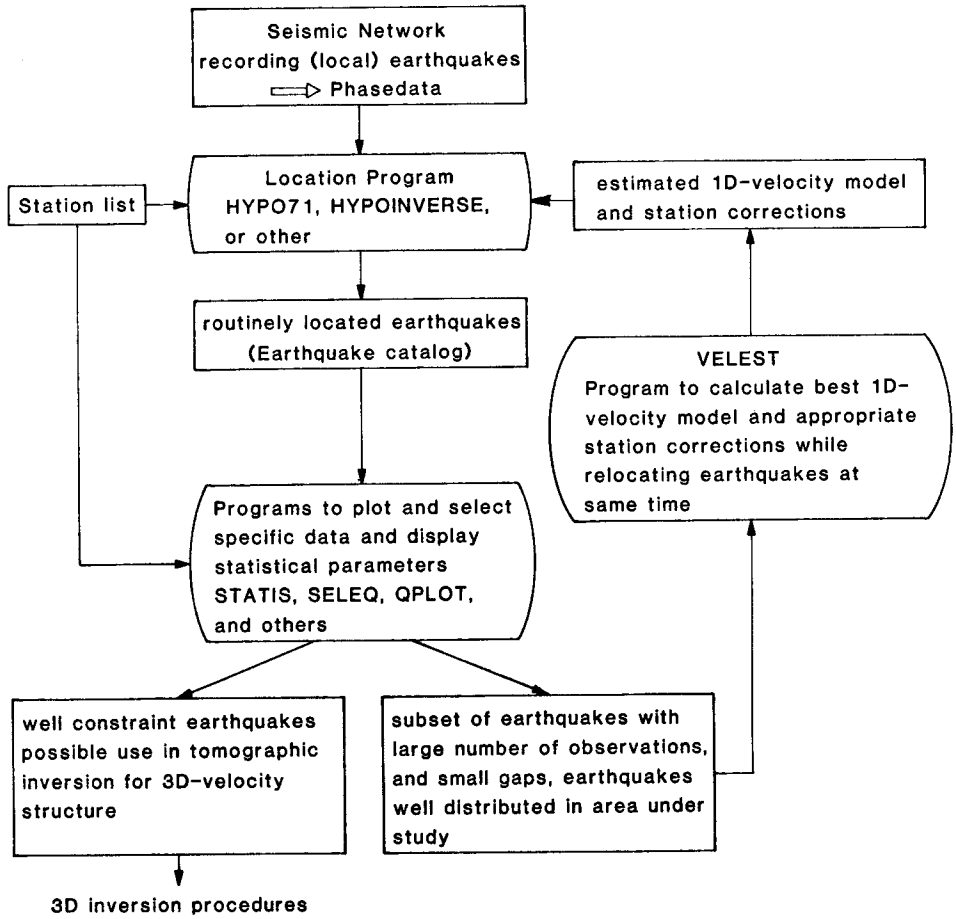

Fig. 11. Acquisition and reduction of local earthquake data sets and inversion of LED to obtain a minimum 1D model. HYPOINVERSE [Klein, 1978] and HYPO71 [Lee and Lahr, 1975] are programs for routine single-event hypocenter location, and QPLOT [Klein, 1978] is a routine to display earthquake data in various forms.

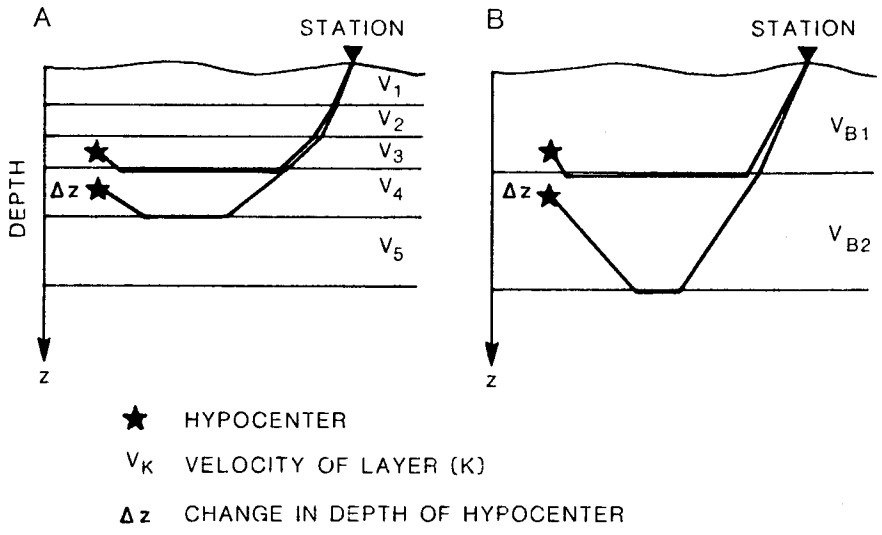

Fig. 12. One-dimensional model and hypocenter location. A change in the calculated depth of a hypocenter will cause a larger change in the calculated travel time for a 1D model with fewer layers (model B).Considering the likelihood of velocity gradients, a 1D model with several layers (model A) allows the tracing of a more realistic ray path. Using model B for hypocenter location calculations, the observed arrivals at the stations may be interpreted as belonging to many fewer phases as compared to the use of model A, which accounts better for lateral variations of the apparent velocity by introducing more phases. The dependence of the hypocenter depth on the velocity model is smaller for model A. Thus for routine earthquake location in an area of known lateral velocity variations, model A (assuming the same average velocity over the full depth range and assuming enough a priori information to establish such a multilayered model) is preferable.

arrivals from various travel time branches, the 1D velocity model should consist of a fair number of seismic layers. The use of a 1D model with too few layers for an earthquake location procedure will often result in poor depth control of the events and might produce large mislocations for some events (Figure 12).

In areas with crustal blocks with large differences in the near-surface velocity field it is advisable to separate the data into subsets with all travel paths fully within one such crustal block. In Long Valley, California, and in Yellowstone, Wyoming, the sedimentary fill of the calderas and the alteration of the rocks by hydrothermal activities on the rocks

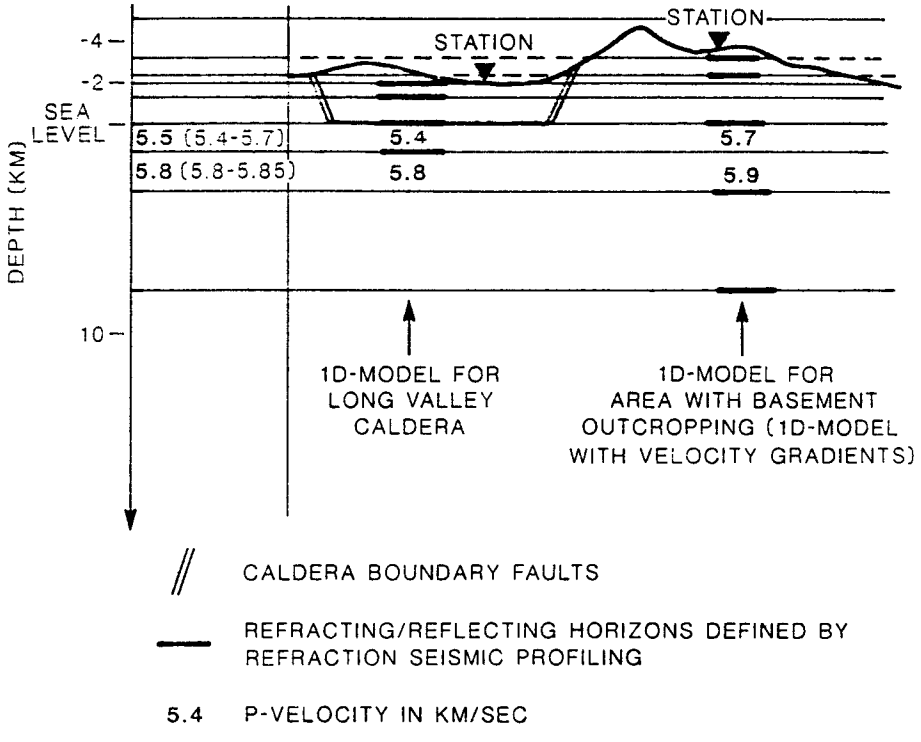

Fig. 13. Connecting two 1D models, e.g., in the area of Long Valley, California. The combined 1D model is shown in the column at left.

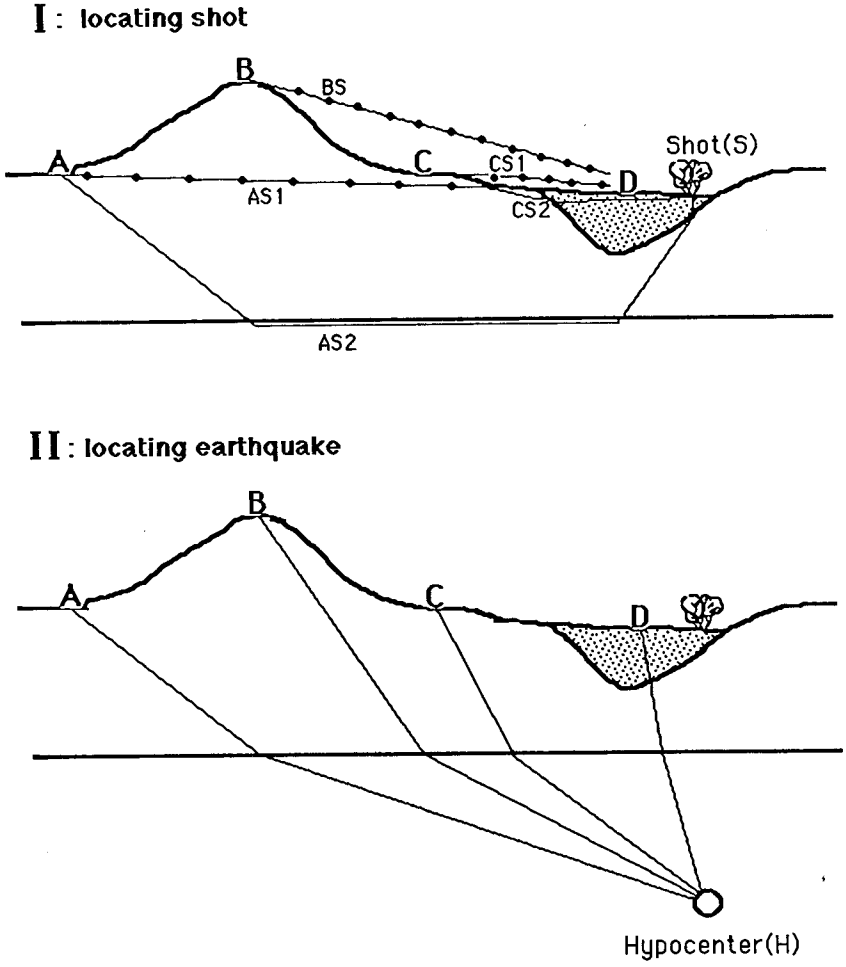

Fig. 14. Test of earthquake location procedure by relocation of shots. Display of the calculated travel paths for rays from a seismic source (I: near-surface shot; II: earthquake at depth) to four seismic stations (A,B,C,D) in an area of moderate topography and a small sedimentary basin, e.g., an alpine valley filled with Quarternary sediments (dotted area). The top model layer encompasses all topography. I: While the calculated fastest ray path BS is nonexistent in nature, the travel paths CS1 and AS1 are geometrically only partly in error.However, considering the lower velocity in the sedimentary basin, the rays CS2 and AS2 might lead to earlier arrivals but are not used in the model. Since almost all calculated travel times are more or less affected by such problems, the shot might be mislocated. II: All four rays are geometrically correct. The calculated travel time for the ray to the station D will be wrong because of the lower velocity in the sedimentary basin. This will cause a somewhat larger residual for this station with little effect on the location of the hypocenter.

beneath the calderas lead to such different velocity fields for the areas outside versus inside the calderas that this procedure had to be applied.

Once the data set and a trial 1D model have been established, repeated use of the program VELEST and appropriate changes in the 1D model will lead to one possible solution (Figure 11). The intrinsic ambiguity in such an inversion of the coupled hypocenter-velocity model problem makes the calculation and testing of several 1D models necessary. If an area had to be subdivided into different crustal blocks, these models now have to be merged (Figure 13). The combined model can then be used as the trial 1D model for inversion of the combined data set with program VELEST.

3.2. Hypocenter Location Procedure with Minimum 1D Model

The minimum 1D model is well suited for routine earthquake locations. First, after testing the location precision by means of observations from artificial sources of known locations (shots), this 1D model provides a fast way to calculate a hypocenter of known precision using standard location programs such as HYPO71 [Lee and Lahr, 1975] or HYPOINVERSE [Klein,1978]. Second, large reading errors or systematic changes in the seismic network such as differences in the time base of station groups are easily detected. Third, the station corrections for LED of known variance (azimuthal variation) can be interpreted in terms of surface geology and used to test the seismic station network.

3.2.1. Test of location procedure and 1D model. With compatible wavelengths of the seismic signals and a distance range similar to that of small local earthquakes, shots from refraction seismic experiments provide good tests for the location procedure and the minimum 1D model. Because of the poorly approximated refractions near the source at the surface, shot locations can be expected generally to be of lower accuracy than the locations of earthquakes in the same area (Figure 14).

The refraction seismic experiments carried out in Long

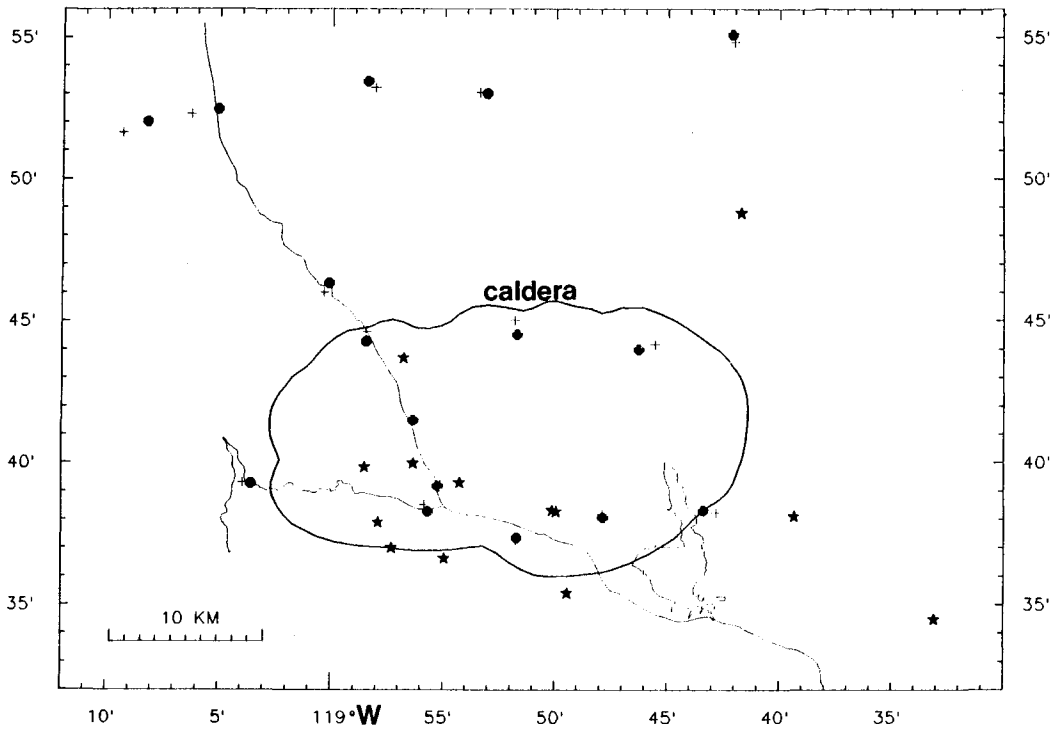

Fig. 15. Long Valley caldera, California. Comparison of the true locations (solid circles) of 16 shots from refraction seismic experiments [Hill et al., 1985] with the locations (pluses) calculated with HYPO71 [Lee and Lahr, 1975] using the minimum 1D model and appropriate station corrections. All stations (stars) that were used for at least one shot location are shown.

Valley [Hill et al., 1985] and in Yellowstone [Smith et al., 1982] provided an excellent chance to test the minimum 1D models calculated for these areas. In Long Valley (Figure 15) most shots were located within less than 1 km of their true location, and the shots situated inside the network (considering only stations with observations) were located within less than 0.5 km. The calculated focal depths for all shots were less than 2 km.

In Yellowstone National Park area (Figure 16) only five shots from the 1977/1978 refraction survey were recorded at the seismic stations used in this study, and none of these shots was located within the caldera. Despite the poor station distribution in some areas and the large differences in the near-surface velocity structure [Smith et al., 1982], the two shots outside the network were located within 5 km of their true locations, and the three shots within the network were closer than 3 km to their true locations.

3.2.2. Minimum 1D model and routine earthquake activity monitoring. A 1D model with station corrections based on a large number of observations contains a statistically significant average expected signal (travel time residual) at each station. In a microearthquake station network the difference between station corrections for neighboring stations should be explainable in terms of surface geology (directly underground of stations; see Figure 20) and/or by comparison of the timed phases. Since the noise level may differ significantly from one station location to the other, a first arrival of small amplitude may repeatedly be overlooked at the noisier station, thus causing a different station correction.

Though harder to detect, mislocations of seismic stations cause systematic changes in the station corrections. Once the seismic station network has been checked for such problems in the data acquisition system and the average observed signals at each station (travel time residual, phase correlation, even amplitude ratio) are established, the use of the minimum 1D model for earthquake location reduces the burden of the routine work. A signal that differs from the long term average at one or several stations automatically alerts the observer (or the observing system) and thus focuses attention on the more unusual events that need special treatment. For special studies such as approximate geotomographic inversions of LED, probably the most important result of the use of a minimum 1D model is the information on which events show characteristic deviations from the normal appearance of earthquakes in the area.

To recognize a deviation from the average signal, one has to know this average. The smaller the standard deviation of the average value, the smaller may be the deviation from this average that can be detected. Again, the attribute "minimum" denotes the fact that this 1D model with its station corrections was calculated by an inversion procedure that calculates these (averaging) velocity model parameters by minimizing the standard deviations of the travel time residuals at the stations. The comparison of the travel time residuals for several hundred events from the Long Valley (Figure 17) and Yellowstone (Figure 18) areas located with the same location procedure using either one or a few 1D models (and station corrections) based on refraction seismic and geologic information or using the minimum 1D models shows much smaller standard deviations for the minimum 1D models.

3.3. The 1D Model and Station Corrections as Approximation to 3D structure

A 1D model with station corrections for the hypocenter location procedure and a 1D geometry with a 3D velocity grid superimposed are approximations which are good enough for all 2D and most 3D crustal structures (Figure 19).

In cases like the Sierra Nevada (California) or like the Swiss

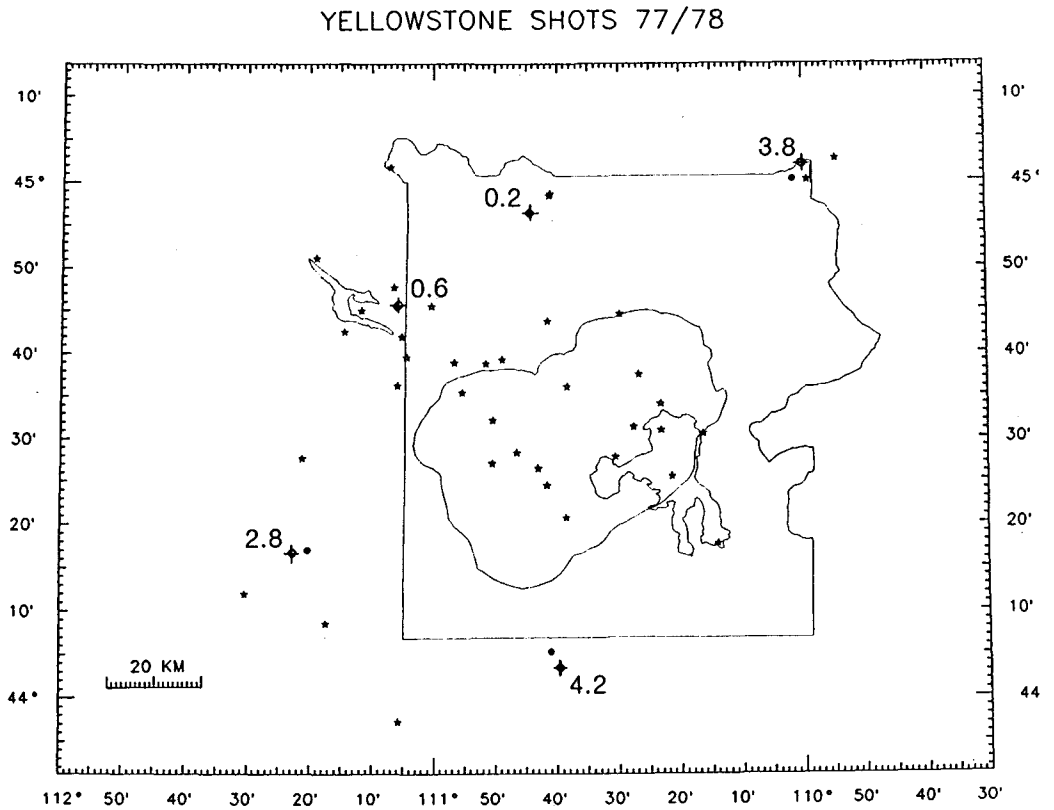

Fig. 16. Yellowstone National Park, Wyoming. Comparison of the true locations (solid circles) of five shots from refraction seismic experiments [Smith et al., 1982] with the locations (dots with pluses) calculated with HYPOINVERSE [Klein, 1978] using the minimum 1D model and appropriate station corrections. All stations (stars) that were used for the location of at least one shot are shown. The number gives the calculated hypocentral depth (kilometers). The two close stations in the northeastern corner of the park were not used for the solution of the nearby shot.

Alps (see Figure 19) the effect of the Moho topography on the travel time residuals of LED should be considerable at first sight. The reasons for the excellent performance of 1D models in both areas, though, (see U.Kradolfer, unpublished thesis, [1988], for the Swiss Alps) may include (1) in LED few rays reach the Moho, if a reasonable number of crustal layers are used (velocity gradients approximated by multilayer model) and (2) the Moho topography is only a problem for rays propagating perpendicular to the strike of the dipping layer.

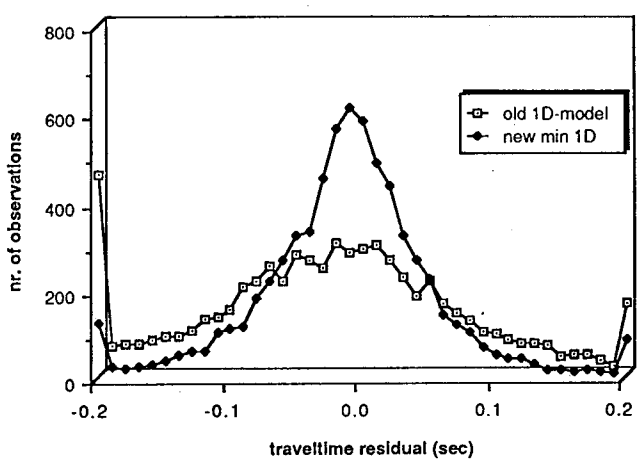

Fig. 17. Long Valley, California. Comparison of the distribution of travel time residuals ('0'-weight observations) for events located with a minimum 1D model and the same data located with the old 1D model that was also based on refraction seismic interpretations and was routinely used.

For lateral refractions the 1D geometry with a 3D velocity grid is obviously not a good approximation to the real Earth, since the procedure will tend to put the refractor (higher velocity) in the layer below rather than into the neighboring blocks. However, we do not have enough a priori information to establish and check true 3D models at the moment. Furthermore, no fast hypocenter location procedure exists that can take such a 3D velocity model into account.

Dipping layers in the starting model introduce large geometric effects (increase in nonlinearity of the travel time dependence) into the ray paths and thus the travel times. It is probably better to proceed with a 3D velocity field without discontinuities other than the (flat, horizontal) layer boundaries and to use an improved bending method [Thurber, 1986] to include all possible lateral refracted paths. This will have the same smoothing but also stabilizing effect on the inversion as the use of blocks.

In the Long Valley area, where we have refraction seismic data for a large part of the caldera and the surrounding area, the velocities calculated by 1D inversion for the minimum 1D model are syste-matically 3% to 5% lower than the velocities found by refraction seismic profiling (Hill et al., 1985; E.Kissling and D.P.Hill, unpublished work,1988] for the same depth. Apparent differences in the p wave velocities obtained from controlled sources and from earthquake

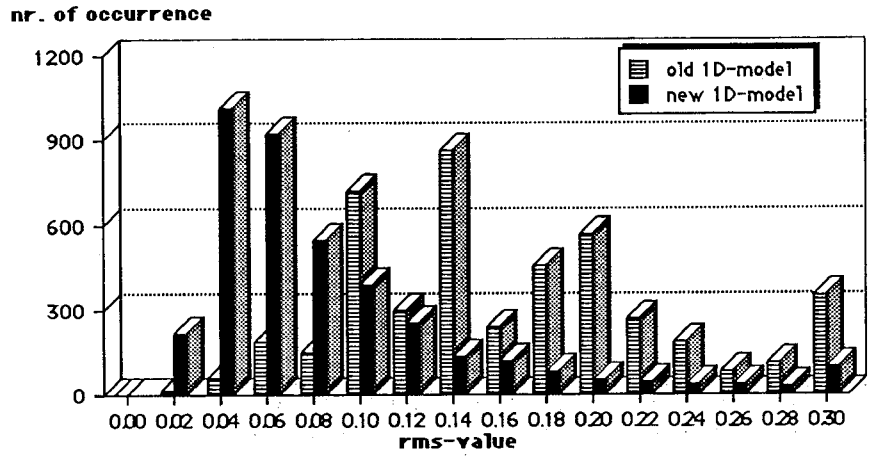

Fig. 18. Yellowstone, Wyoming. Comparison of the distribution of rms travel time residuals for some 4000 events located with a minimum 1D model and the same data located with the old 1D model that is also based on refraction seismic interpretation (value in s^2).

seismology have been reported by Gutenberg and Richter [1951] and have later been reinterpreted by Mueller and Landisman [1966] and Mueller and Peterschmitt [1981] in support of velocity reversals in the crust. In our case, however, the differences in the velocities cannot be caused by low-velocity layers, since the effect is visible for layers at all crustal depths.

The reason given by Wielandt [1987] is that diffraction around bodies of lower p wave velocity will affect the average layer velocity. This effect depends on the dimensions of the anomalous body and the direction of the rays. In general, travel times for ray paths that penetrate the body of lower p wave velocity will still be the first arrivals if the dimensions of the body perpendicular to the ray are several times the dimension of the body along the ray. With the exception of spherical bodies, the local earthquake tomography with rays from all directions and from various locations in the crust will be less affected by such diffractions than refraction seismic profiling where most rays lie within a ray tube. Since this effect will tend to hide the local (3D) anomalies of lower velocity, only the average layer velocity will be afftected, and thus in refraction seismic interpretation, the averaged layer velocities should be larger than the corresponding velocities calculated from 1D inversion of LED.

As stated earlier, the geotomographic inversions of LED for the three-dimensional velocity field may be viewed as the back projection of the station corrections calculated with the minimum 1D model. In fact, these station corrections are but the negative average values for the azimuthally and radially varying time delays at these stations relative to the starting model for the 3D inversion (Figure 20).

In teleseismic studies, the main data signal to be interpreted and inverted is the distribution of travel time delays at seismic stations with respect to the azimuth of the incoming seismic wave fronts [e.g., Steeples and Iyer, 1976]. The inversion of local earthquake data is also based on the travel time delays at seismic stations as the signal. However, with local earthquake data, the azimuthal and radial dependences are never as uniform for all stations as for the teleseismic data, even for rays passing through the same anomalous region if the hypocenters are well distributed over the area. The angle of incidence for rays from the same hypocenter varies greatly at different stations for LED. In order to display the travel time delays which have to be inverted for velocity parameters by the inversion process, we have to find rays that passed through the same region. The program DATVAR (see Figure 9), therefore, searches the data for observations at a specific station from events with hypocenters in the same small area

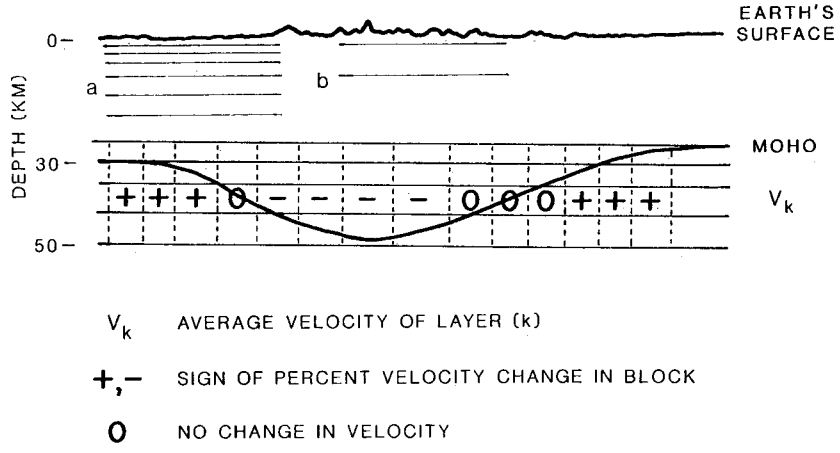

Fig. 19. Approximation of 2D structure by a 1D model. a and b represent two different one-dimensional models for the crust (see Figures 12 and 13 for details). V_k: average velocity in kth layer; -, 0, +: sign of expected percent velocity change in a 3D inversion for homogeneous ray coverage

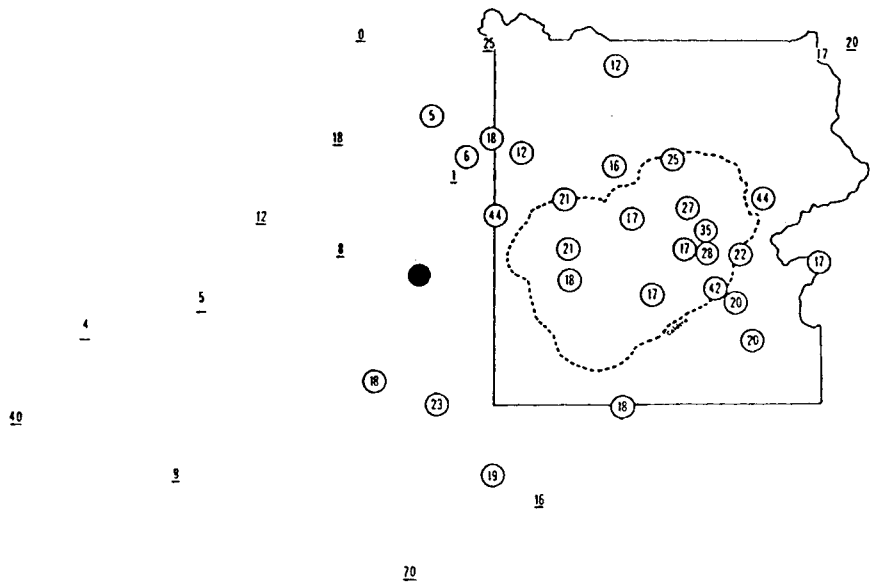

Fig. 20. Yellowstone National Park, Wyoming. Distribution of station delays (in 0.01 s) relative to reference station (solid circle). Positive values are circled; negative values are underlined.

(blocks of N*N*N (km), where N is a small integer). For each station and block, all travel time residuals from events within this block are used for the variance calculation of this specific hypocenter area (block)-station pair. While the average travel time residual for such a station-block pair is proportional to the signal (travel time delay along the ray path between source and receiver) to be inverted, the variance of these travel time residuals is a measure of the statistic input data error.

A comparison of the station delays before and after the tomographic inversion (Figure 21a) shows the effect of this back projection process. While the amplitudes of the station delays are reduced by 50% (on average), the variance of the station delays (Figure 21b) shows reductions up to 70%, demonstrating the effectiveness of the 3D model inversion.

4. Forward Problem

The earthquake location procedure and the determination of the three-dimensional velocity structure require the tracing of seismic rays in a heterogeneous isotropic medium between two given endpoints. In the 1970's several numerical techniques to solve this ray tracing problem have been developed [Jackson, 1970; Julian, 1970; Wesson, 1971; Chander, 1975; Julian and Gubbins, 1977; Pereyra et al., 1980] and applied to seismological problems. In general, the techniques can be divided into two different methods: the shooting and the bending methods.

The choice of the raytracer is crucial for all seismic inversions because of the trade-off between the precision and the required computing time. In this study, more than one third of the overall computation time is used to solve the forward problem, i.e., to calculate the travel times and the partial derivatives of the travel times with respect to the model and the hypocentral parameters.

We use a shooting method through a geometrically one-dimensional layered Earth model. The 3D model consists of horizontal layer boundaries as first-order discontinuities and with a three-dimensional velocity grid superimposed. Thus the geometry of the ray can be calculated with few iteration steps using a local averaged 1D model between source and receiver. Each layer velocity of this local 1D model is an average of the (block) velocities within a tube around the direct line between the source and the receiver. The ray tracer calculates the travel times for the direct and for the critically

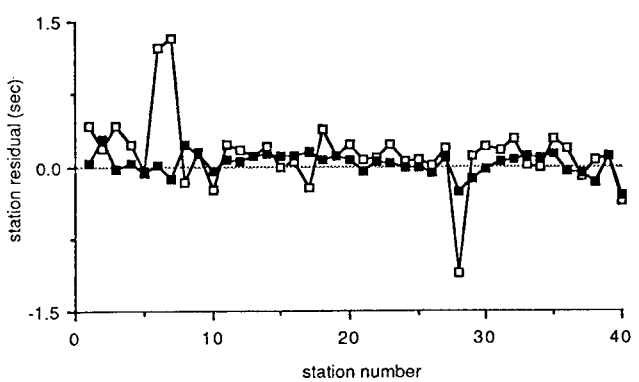

Fig. 21a. Yellowstone, Wyoming. Station residuals (seconds) before (open squares) and after (solid squares) the 3D model inversion.

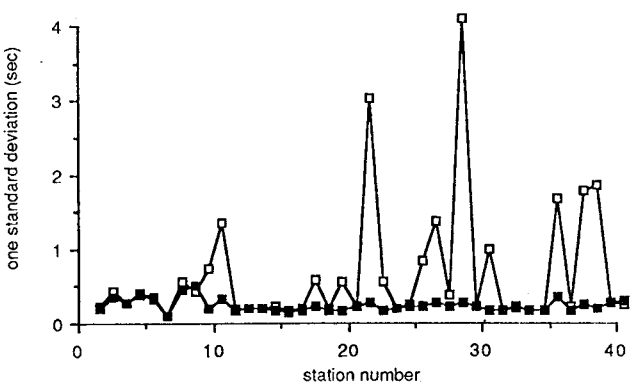

Fig. 21b. Yellowstone, Wyoming. Standard deviation of station residuals (seconds) before (open squares) and after (solid squares) the 3D model inversion.

refracted ray (Figure 22). For the critically refracted ray path the velocities of all blocks within and below the tube around the direct line between source and receiver are used for the average layer velocity.

The travel time for a seismic signal traveling along a ray path is calculated using the velocities of the 3D grid. Thus the geometry of the travel path for a given source-receiver pair is much less affected by changes in the block velocities than the travel time.

There are several reasons for using a 1D geometry with a 3D velocity grid superimposed rather than a true three-dimensional velocity structure. We are imaging the 3D velocity field as deviations from an average, simple velocity structure. This average velocity structure will account for the known and presumed discontinuities at depth. Lateral refractions along vertical block boundaries should be avoided, since the blocks denote areas of uniform slowness, rather than geometrically defined objects of constant internal velocity. The low a priori information density in most inversion studies prohibits the construction of a three-dimensional starting velocity model with well-defined lateral refractors. The geometry of the starting model is assumed to be a good average for the area studied. We do not invert for geometrical model parameters but for the 3D velocity field, locally averaged over a small volume (block). Thus the 3D geometry of the velocity field is inferred from the velocity values. Changing the geometry of the 3D velocity field during the iterative inversion process without inverting for these geometrical model parameters introduces instabilities (see, for example, Figure 10).

A newly developed bending method [Thurber, 1986] may allow future improvements on the raytracer to better control the lateral refractions. Using the 1D ray path as a starting travel path, Thurber's method approximates the true ray path by bending according to the local velocity field. The 3D velocity grid that results from an inversion of LED is thus well fit to be of use for this bending method. A flat Earth approximation [G. Mueller, 1971] is used to transform depth and velocities from the spherical 1D model into a Cartesian 1D geometry to calculate ray path, travel time, and all derivatives, that are transformed back before being used in the inversion.

5. Testing of the Inversion Procedure With Artificial Data

The calculation and use of artificial local earthquake data sets (ADS) is an intrinsic part of the inversion method discussed in this paper. Several questions concerning the method and/or the assumptions necessary to allow an inversion of a specific data set can be tested by use of ADS. In particular, the use of ADS for geotomographic inversions serves the following purposes:

1. To test the applied ray tracer and the inversion method (ACH inversion versus approximate geotomographic inversion; application of the parameter separation; use of geometric weighting; use of step length weighting; comparison between ray density tensors and resolution).
2. To select the appropriate inversion parameters (damping para-meters; number of iteration steps for hypocenter relocation and for the model inversion; minimal number of hits of a model block to be inverted for) for a given data set and type of crustal structure.
3. To test a specific data set for resolution capabilities. Where does a given data set resolve a 3D structure well and what are the best dimensions of the model blocks?
4. To address the problem of the effects of reading errors (data variance) on the hypocenter locations and on the resulting velocity field. Which local anomalies in the resulting velocity field are artefacts of the improper inversion procedure or a result of systematic errors in the input data? What are the effects of mislocated stations on a resulting 3D velocity field?
5. To compare the results of different inversion methods (approxi-mate geotomographic versus ACH method) applied to optimal and real data sets.
6. To simply test the program code of the inversion procedure for numerical or logical errors.

While the last task of finding the errors in the program code could well be performed with pure synthetic data of optimal resolution, all other tests with ADS are best addressed by a set of more realistic data with all its deficiencies. Real LED will always show an uneven travel path distribution, due to the distribution of the hypocenters or due to the uneven station coverage. As a consequence, the data set not only resolves the velocity field in some areas better than in others, but the precision of the hypocenter locations varies also. Through the coupling of the hypocenter and model inversion procedure, both deficiencies of real LED may lead to local anomalies in the resulting velocity field. To simulate a realistic situation, the ADS used in this study are copies of real data sets, as far as hypocenter and station distribution and the number of observations per event are concerned (Figure 23). The travel times for all artificial observations are calculated by ray tracing through a given three-dimensional velocity field using the 2D ray tracing system described in section 4. Random reading and timing errors can be simulated by adding random numbers to these travel times (see program ARTDAT, Figure 23). Systematic errors in the data set are introduced as systematic changes on some subset of the travel times or as specific changes in station or hypocenter coordinates (see program MODINP, Figure 23), depending on the problem to be simulated.

Random errors (even of significant size) added to travel times are of little importance for the resulting velocity field in areas of good resolution. This was proved by two approximate geotomographic inversions of artificial data, once without reading errors added and once with random errors of 0.09 s^2 variance added to the calculated travel times. Tests with different distributions of the added timing errors such as Gaussian, Poisson, or white noise (uniform) distributions showed few differences in the resulting velocity fields.

The three-dimensional velocity model used to calculate the ADS is shown in Figure 24. To allow the comparison of all three tests (see sections 5.1 through 5.3), the same ADS derived from an artificial Long Valley model was used. Local areas of anomalous high or low p wave velocity were introduced in inversion layers 3, 4, and 6. The total amount of travel time difference between this 3D model and the corresponding 1D model without the local velocity anomalies

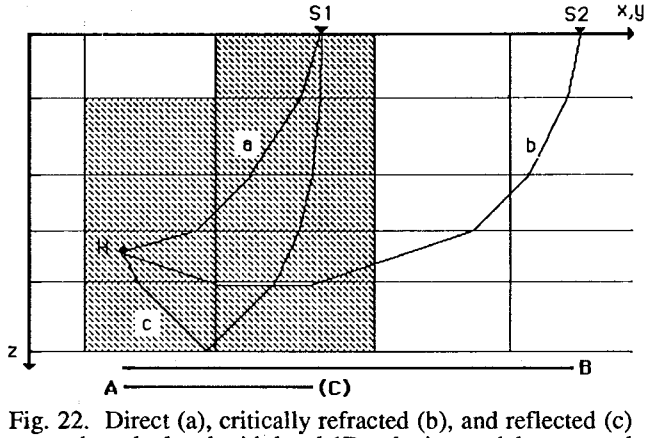

Fig. 22. Direct (a), critically refracted (b), and reflected (c) ray paths calculated with local 1D velocity models averaged over the block velocities in the ranges A (=C) and B for the direct and the refracted rays, respectively. Shaded area shows block velocities averaged for each layer of the local 1D model used for the direct ray (a). H: hypocenter; S1,S2: stations.

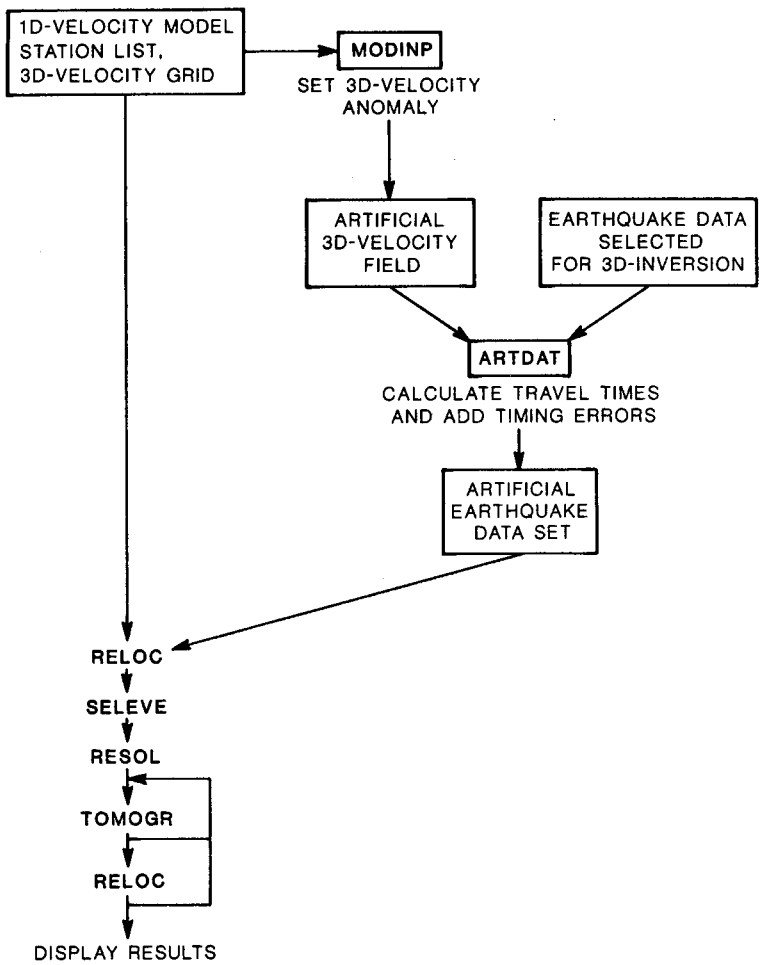

Fig. 23. Flow chart for calculation and inversion of artificial data.

(sum of delay times of each individual ray) for this ADS is 413 s. A Gaussian distributed error of 0.1 s standard deviation was added to all travel times. The ADS was derived from a subset of the Long Valley data set with similar ray coverage as the latter and is composed of 1761 events with a total of 38,476 observations. The total amount of travel time for all rays in the 1D model is 307,612 s. Thus, the total difference in travel time between the 3D model and the 1D model amounts to 0.15% of the total 1D model travel time.

5.1. Tests of Inversion Procedure

The following calculations with ADS are used to demonstrate the effects and to find the appropriate values of some parameters and options used in the inversion procedure. The main option is the possibility of separating the hypocenter parameters from the velocity parameters before the inversion for the velocity field.

In order to reduce the number of color figures the same ADS has been used for all tests with artificial data. The local anomalies in the p wave velocity field of the artificial model (Figure 24) are mainly placed to demonstrate the resolution capabilities of the Long Valley data set. For the same reason, the amount of Gaussian distributed error added to the data was comparatively small, as the data variance of the Long Valley data (Figure 35a). In most cases, however, the LED sets will not match this quality (in terms of use for inversion for the 3D velocity field) of the Long Valley data. As a consequence, for generalization of the results of this test of parameter separation, the data variance of the ADS should be larger and the resolution poorer in the areas of the local anomalies. Inversions without parameter separation on such (poorer) ADS resulted in p wave velocity fields that did not allow to recognize any geometrical features of the artificial 3D model.

The ADS from the Long Valley subset calculated for the 3D velocity model shown in Figure 24 has been inverted with the ACH method (Figure 25) and with the approximate geotomographic method keeping all inversion parameters the same but using once the option of parameter separation and step length weighting (Figure 26), whereas Figure 27 has been obtained without separation of the hypocentral and velocity model parameters and without step length weighting. The comparison of the two resulting velocity fields (Figures 26 and 27) demonstrates the need for parameter separation for inversion of LED. The purpose of the step length weighting is to increase the damping for all poorly resolved velocity parameters in order to reduce the errors in the calculated velocity adjustments. From the comparison of Figure 28 with Figure 27 it can be deduced that (1) the step length weighting is a useful tool to reduce the model variance for poorly resolved model blocks and (2) most of the errors induced by the unseparated hypocentral part of the equations are indeed induced into the poorly resolved velocity model parameters. Considering the differences between Figures 26 and 28, however, and considering the quality of the ADS used in these tests, for approximate geotomographic inversions (i.e.,

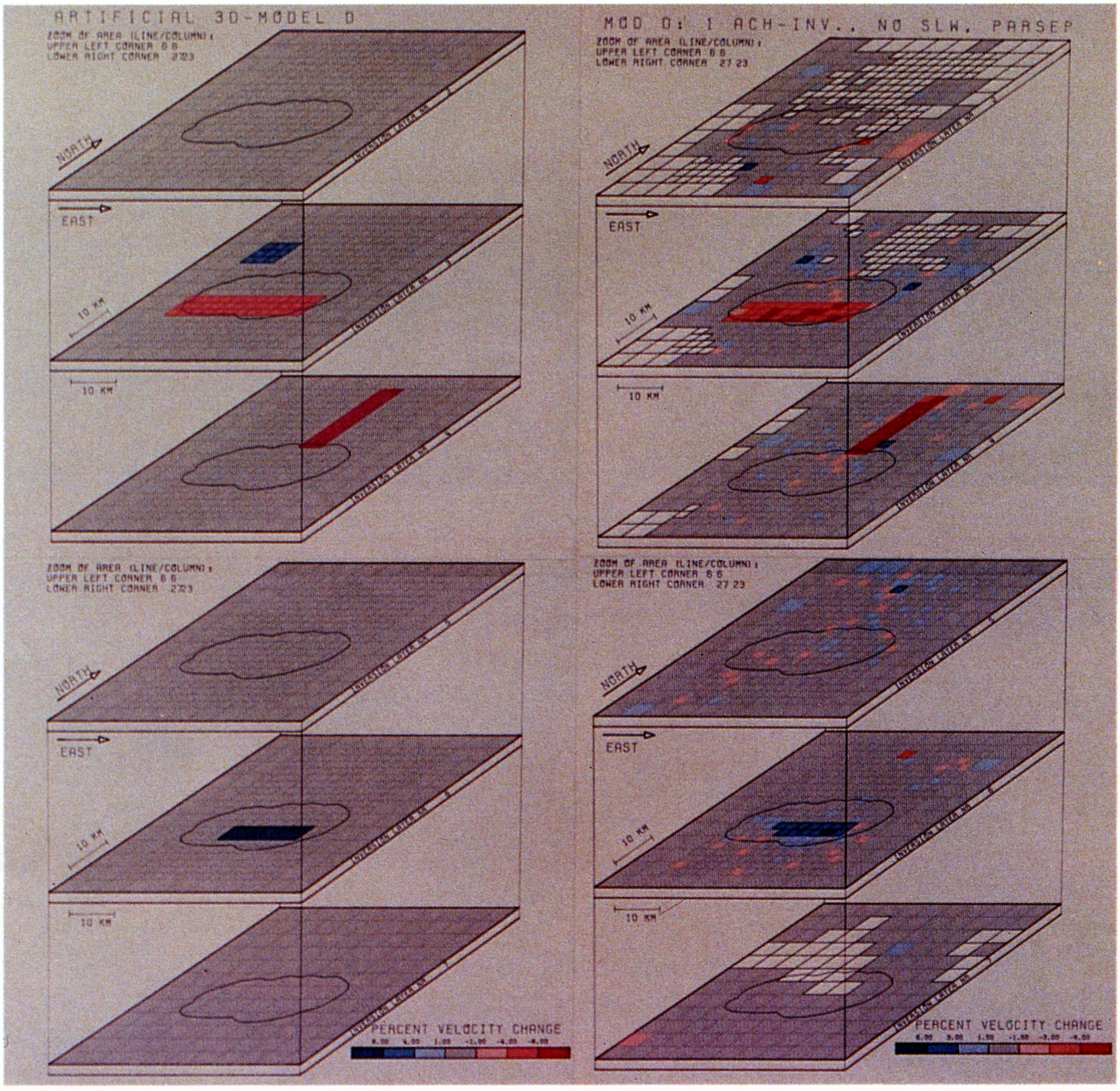

Fig. 24. Artificial 3D velocity field used to calculate the ADS (Gaussian distributed random error of 0.1 s standard deviation added) for Figures 25, 26, 27, 28, and 29.

Fig. 25. Resulting 3D velocity field obtained from inversion of ADS with ACH method (after one iteration step, without step length weighting, using a velocity damping parameter of 1.0).

approximative ACH inversion) the parameter separation must not be bypassed when using local earthquake data.

5.2. Effects of Systematic Errors in the Data

Many of the routine checks in a seismic network and most of the work done to select earthquake data for use in special applications such as inversion for velocity structure are devoted to the search for and the removal of systematic errors in the data. Systematic errors such as mislocations of seismic stations or shifts in the timing base of several stations relative to some other stations within the network might have effects on the data that are identical to the effects of 3D velocity anomalies. Systematic errors in the data, even of much smaller size than the random timing errors, must therefore be converted by the 3D inversion to velocity anomalies and will lead to serious misinterpretations. Considering the consequences, the selection (see section 6) and double-check of earthquake data thatl will be used later for an inversion to obtain a 3D velocity field, are an essential part of the inversion process.

The calculation and inversion of ADS may help in recognizing effects of systematic errors in the data. As an example, the effect of mislocated stations will be shown (Figure 29). The ADS used is the same as for the test of parameter separation and is derived from the Long Valley data set using the 3D model shown in Figure 24. For the relocation of the events (see Figure 23) and the subsequent inversion of

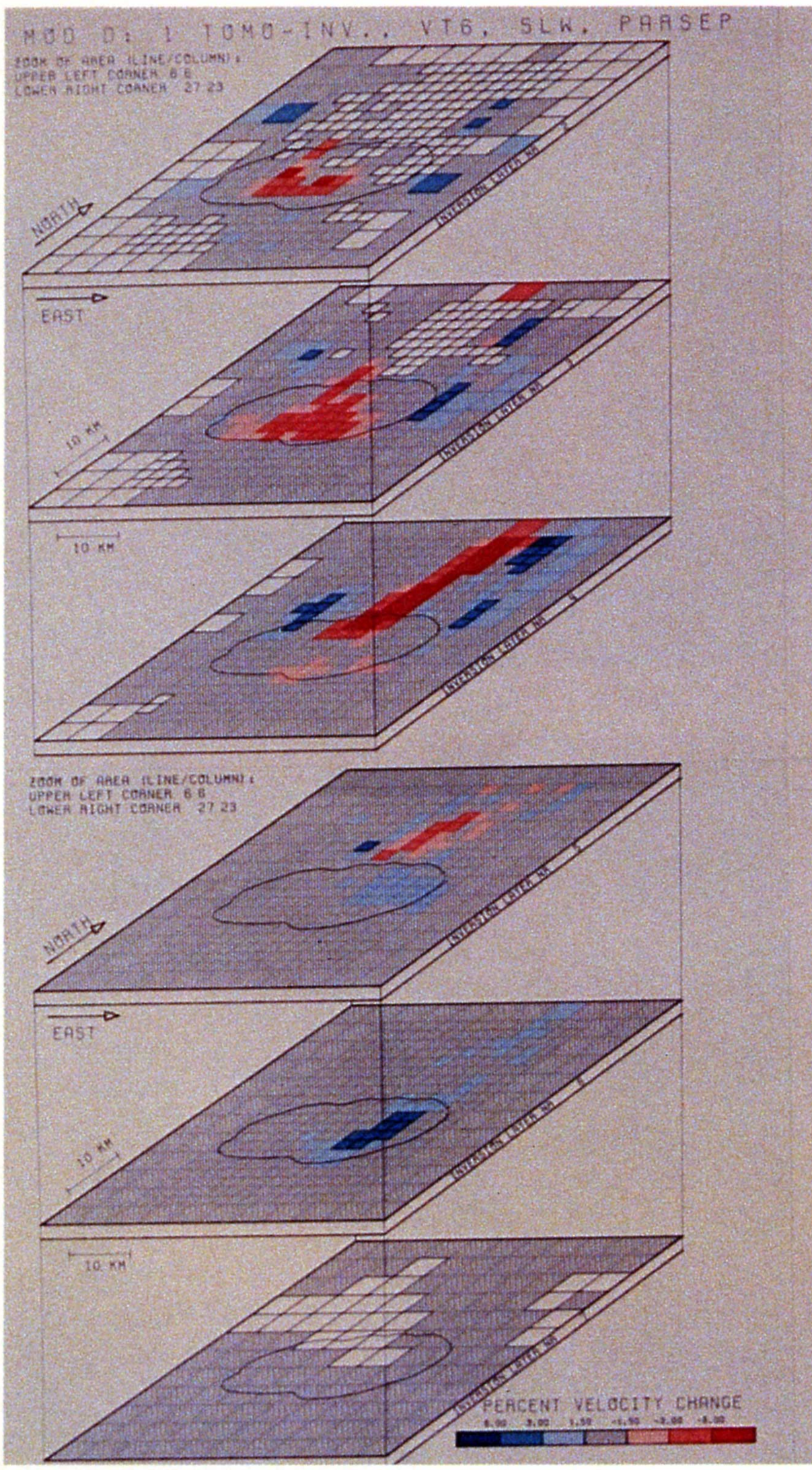

Fig. 26. Resulting 3D velocity field obtained from approximate geotomographic inversion of ADS (after first iteration step, with step length weighting and with parameter separation, using a velocity damping parameter of 6.0).

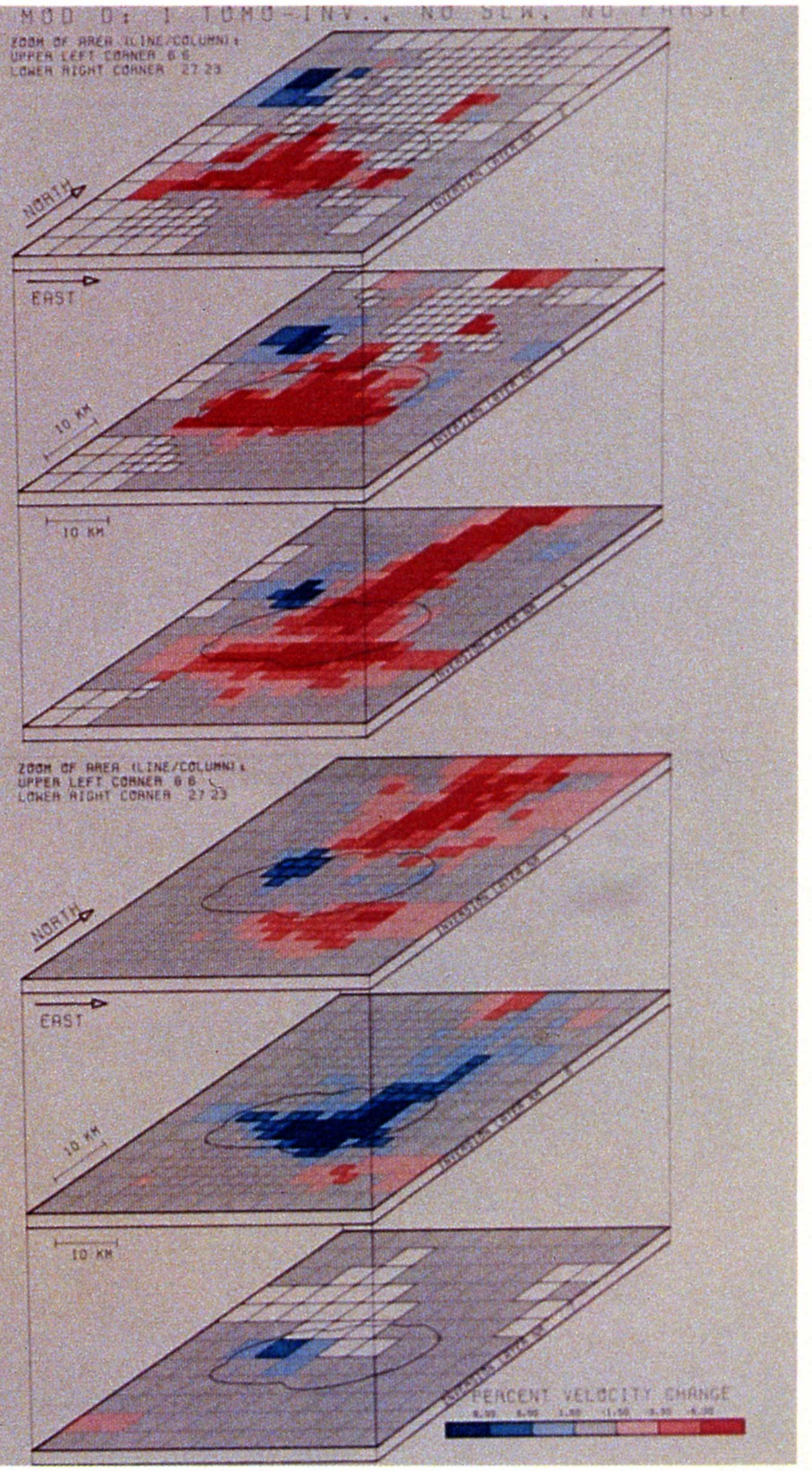

Fig. 27. Resulting 3D velocity field obtained from approximate geotomographic inversion of ADS (after first iteration step, without step length weighting and without parameter separation, using a velocity damping parameter of 6.0).

this ADS, two stations (UCHK and UMON, denoted by solid squares in Figure 29) out of a total of 225 stations have been shifted northward by 3 km each. In the 3D velocity field (Figure 29) that resulted from a ACH inversion, this systematic error, which affected less than 1% of all ray paths, causes local p wave velocity anomalies that not only mask the true 3D velocity field, but are also misleading in the structural interpretation of the resulting velocity field.

5.3. Testing a Data Set for Resolution Capabilities

In the resulting p wave velocity field for the Long Valley area (Figure 37), the region north of the caldera shows a chessboard-like pattern of local anomalies in the inversion layers 3 and 4. From the distribution of stations and hypocenters (Figures 33 and 36) in the Long Valley area, we expect to resolve the velocity structure within the caldera in more detail than in the area north and northwest of the caldera. The different block sizes (see Figure 8) reflect this expectation. However, in the resulting velocity field (Figure 37) one finds few local anomalies of small amplitude in the Sierra Nevada northwest of the caldera, whereas the area north of the caldera shows the chessboard-like pattern. The question of reliability of such results may, in part, be answered by the use of ADS. Rather than testing the applied inversion procedure, this means testing a specific data set for resolution capabilities

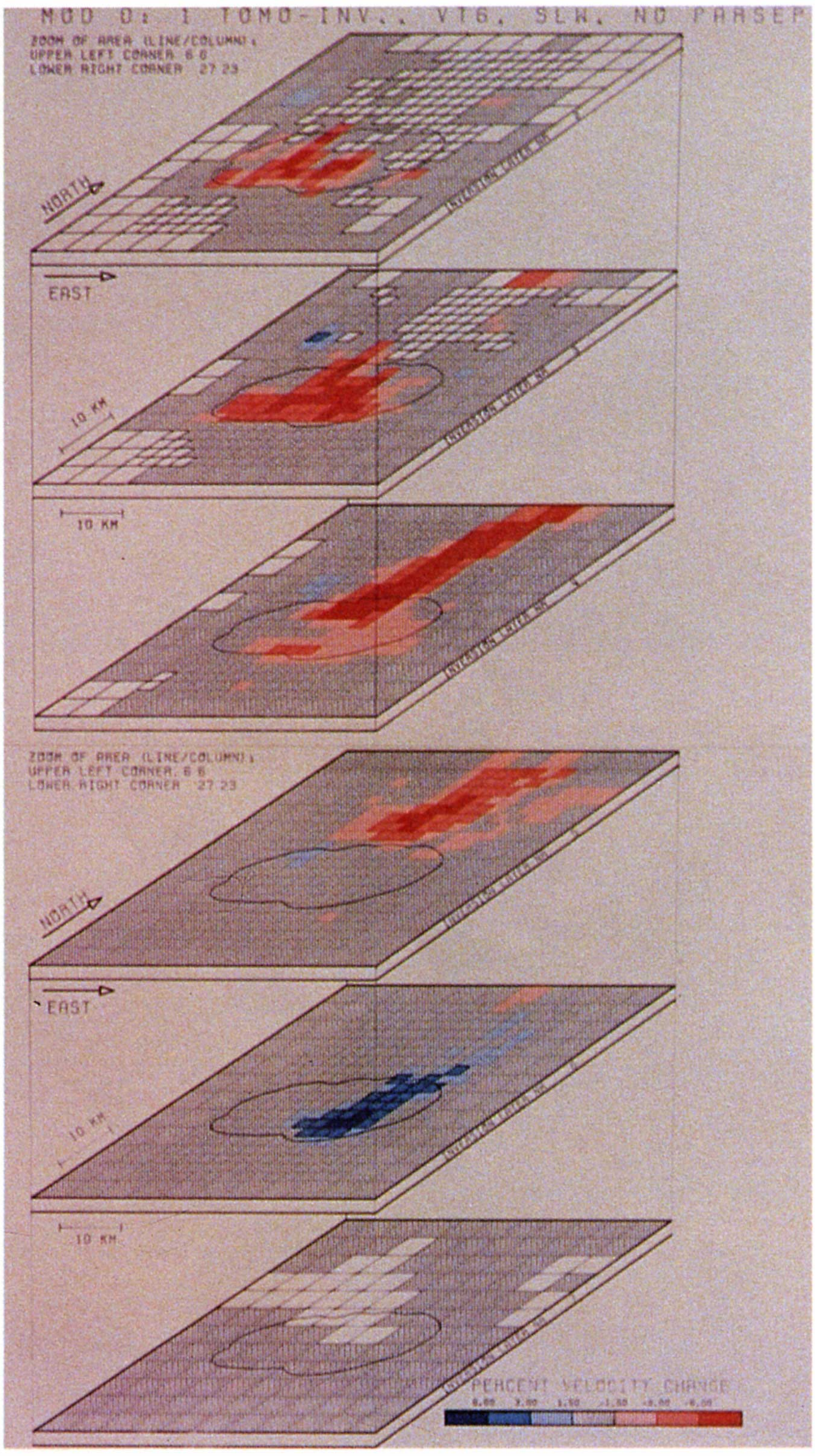

Fig. 28. Resulting 3D velocity field obtained from approximate geotomographic inversion of ADS (after first iteration step, with step length weighting but without parameter separation, using a velocity damping parameter of 6.0).

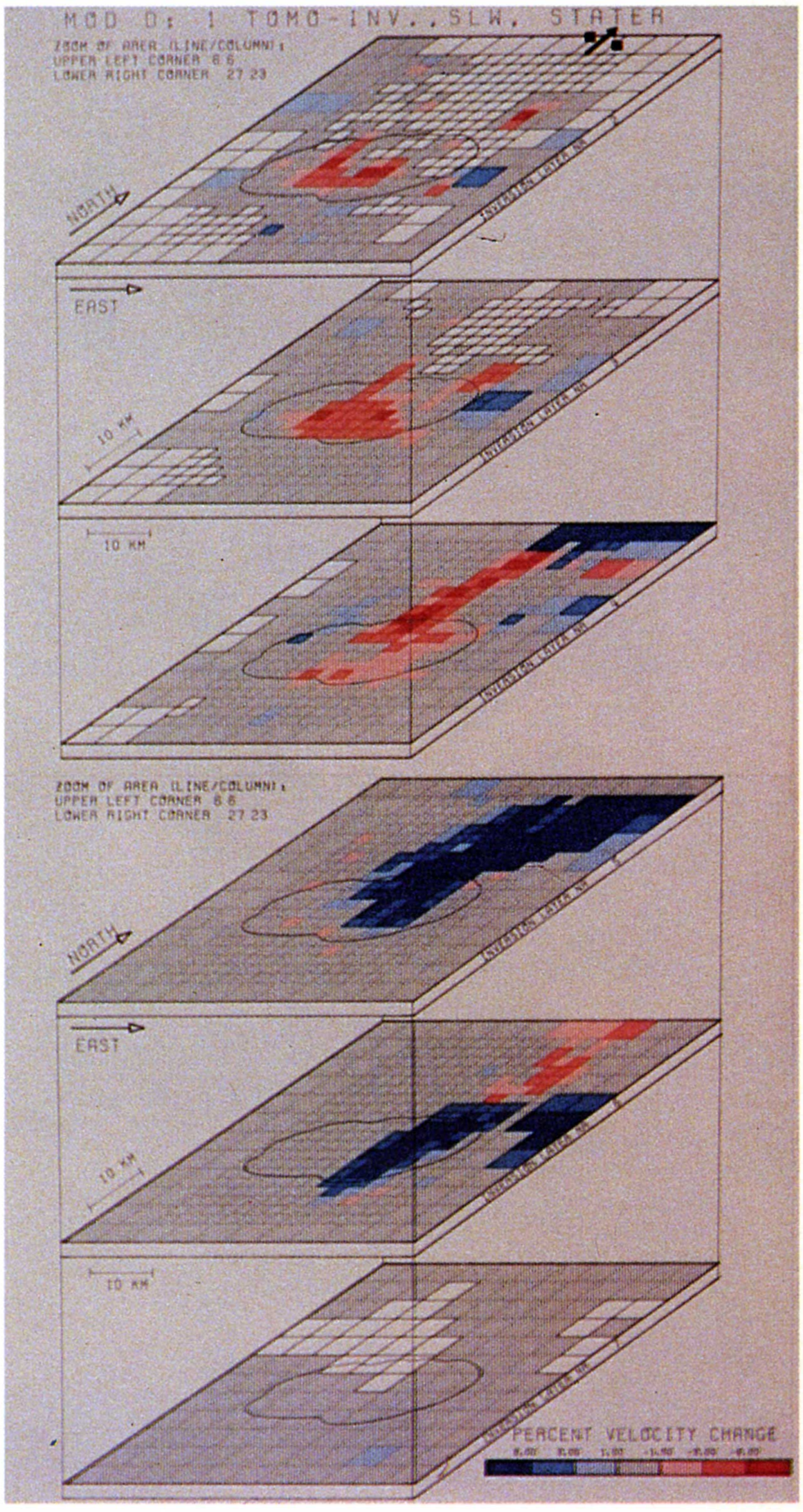

Fig. 29. Resulting 3D velocity field obtained from ACH inversion of ADS (see text) where the locations of two stations (solid squares) have been moved to wrong positions (3 km northward) deliberately. All local anomalies that differ from the resulting velocity field shown in Figure 25 are artifacts of the errors introduced into the ADS by mislocating the two stations.

in selected areas of interest. In the case of the Long Valley area several local p wave velocity anomalies were introduced in the artificial 3D model (Figure 24), and the anomalies north of the caldera were deliberately placed in areas of inhomogeneous and even poor resolution.

The results of one iteration step with the ACH inversion method (Figure 25) and with the approximate geotomographic method (Figure 26) are in general agreement, the first method being clearly superior in resolving the shape and amplitude of the local anomalies within the crustal block below the caldera. The results obtained by approximate geotomographic inversion improve with additional iterations in areas of good resolution only. Neither method, however, resolves the local anomaly of higher p wave velocity northwest of the caldera in inversion layer 3. This suggests that this data set is incapable of resolving local p wave velocity anomalies of less than 10 km diameter in that specific area, whereas within the caldera, much smaller anomalies (block sizes within the caldera equal 2.5 km by 2.5 km; see Figure 25) will be well resolved.

6. Local Earthquake Data

As are any geophysical data, LED are imperfect. The ambiguity intrinsic to the inversion procedure and the

imperfect nature of the LED used as input for such inversions make the interpretation of resulting three-dimensional velocity fields a tricky task. The impor-tance of the data check for systematic errors is obvious and has been demonstrated by the calculation of ADS to illustrate the effects of mislocated stations. Since the anomalies which are artifacts from systematic errors in the data may not be distinguishable from real velocity anomalies in the resulting 3D velocity field, in all applications of the inverse theory the criticism should first target the data-filtering process before critically addressing the results obtained or even their interpretation.

In this study, only arrival times of p waves have been used. To image attenuating bodies and to obtain further information about the local 3D structure, it would be preferable to include travel time data from s phases and amplitude data. The main reason for this (temporary) restriction to p wave data is the absence of adequate data filters for s wave and amplitude data. Such filters have to be established from theoretical assumptions followed by tuning tests on a large number of data in a standardized format. The filters applied to the p wave travel time data (see programs STATIS, VELEST, SELEQ, and SELEVE in Figures 9 and 11) are based on some 430,000 observations from the Long Valley and some 120,000 observations from the Yellowstone areas. The use of these filters has led to the detection of several systematic errors in these two data sets, such as the mislocation of a station by 3.4 km. In comparison with more than 550,000 p wave travel time observations, the total of 10,000 s wave travel times in the same two areas are insufficent to establish reliable tests independent of the p observation filtering.

6.1. Phase Correlation

The recognition and correlation of seismic phases at different stations are the key to any reliable LED set for inversion. As with the calculation of the minimum 1D model, refraction seismic profiles provide the best independent information to facilitate the correlation of phases from earthquake seismograms. Plotting the earthquake seismograms corresponding to their radial distance from the epicenter and comparing this profile with refraction seismic profiles reveals the most prominent p phases such as Pg, Pn, and PMP (the direct phase, the critically refracted phase from the crust-mantle boundary, and the reflected phase from the Moho, respectively). In this inversion study, only first p arrivals (Pg and Pn) were used, and each such travel time value was compared with the calculated travel times corresponding to both the direct and the refracted phase. With this restriction to the most reliable data, we sacrifice additional information that could be gained from strong secondary p arrivals as reflections from the Moho. In areas of lower seismic acitvity and with fewer stations than in the Long Valley area, this additional information, though less reliable in its phase correlation, is essential (U.Kradolfer, unpublished thesis, 1988).

For the Yellowstone area, R.B.Smith (unpublished written communication, 1987; for data see [Smith et al., 1982]) found that the intrinsic attenuation diminishes the amplitude of the headwaves even more drastically than that of the theoretical value, while wide angle reflections may increase in amplitude as a function of increasing distance. In both the Yellowstone and the Long Valley area, this potential problem has a higher likelihood than in a nonvolcanic area and, therefore, we must rely on the refraction seismic data for identification of phases.

The correlation of phases from different stations is not only a question of distance and the applied velocity model. The different noise levels at these stations and the existence of strongly attenuating bodies (Figure 30) pose an additional problem. In areas of volcanic activity, seismic observations which bear the most valuable infor-mation in terms of travel time signal are often attenuated to the extent that a phase correlation based on travel time data only becomes questionable. As in such cases, the uncritical use of seismic data from catalogs (that contain no information about the seismic signal) for 3D inversion may lead to serious errors in the resulting velocity field.

6.2. Selection of LED for Inversion

Considering the coupling of the hypocenter/velocity model problems, LED used in the inversion to obtain a three-dimensional velocity field should match the following criteria. The formulation of the problem as overdetermined requires at least five good p arrivals (Pg or Pn) per event to solve the location problem alone. In general, all well locatable events (gap of less than 180°) with more than seven good p observations are suited for the 3D inversion. Events with fewer observations do not add many more constraints to the model solution but unnecessarily increase the computational burden if we consider the growing size of the matrix of the coupled problem! This matrix comprises the total number of unknowns times total number of observations. While additional observations of the same earthquakes increase the overdetermination factor, every additional event also adds four more unknowns. The LED set from Long Valley used for inversion consists of about 80,000 observations from 3600 events, and these data define about 3000 velocity model parameters. Thus the original matrix of the coupled problem contains mbl (mbl = 4 * 3600 + 3000) unknowns in rows and 80,000 lines. Though this exceptional data set could be used to define more velocity model parameters (if a better distribution of the hypocenters were given) and though one could use fewer observations and events (i.e., sacrificing the high degree of overdetermination), in most inversions of LED the velocity model parameters will be far outnumbered by the hypocentral parameters. Given the importance of parameter sepa-ration (see section 5) for inversion of LED, the hypocentral parameters cannot be ignored for the model calculation. The overall computational burden and the maximum size of the required core memory are therefore

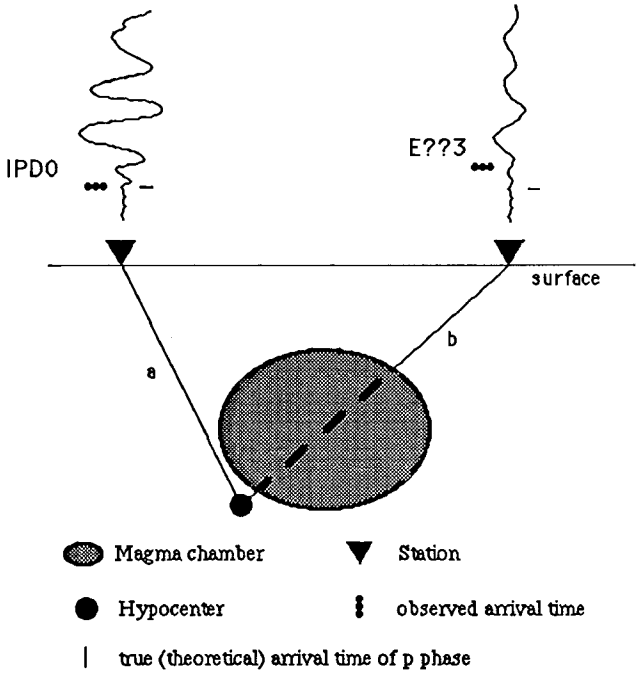

Fig. 30. Phase correlation in the vicinity of a highly attenuating body.

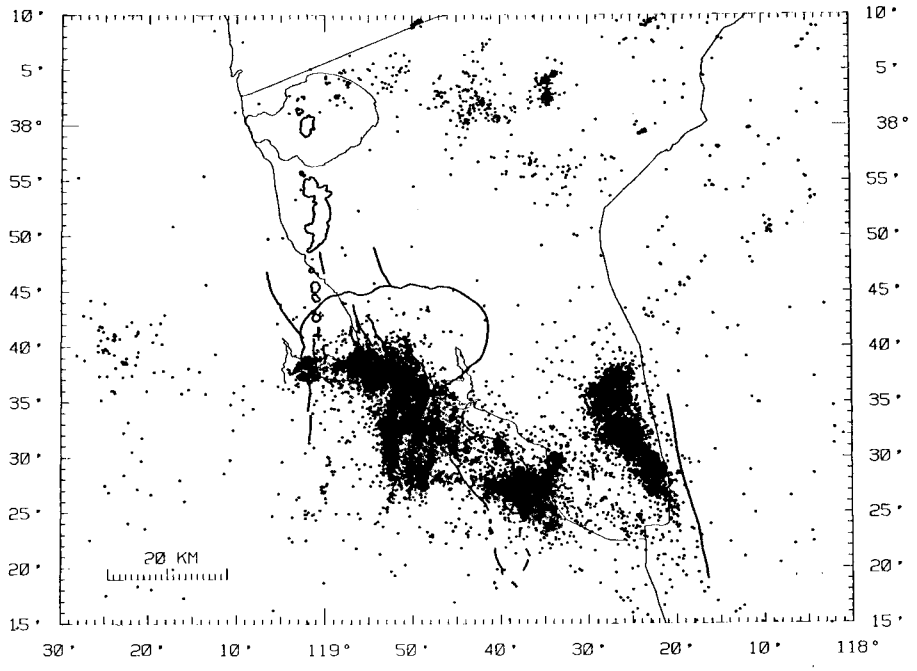

Fig. 31. Distribution of about 30,000 events located by the U.S. Geological Survey in the Long Valley, California, area between May 1980 and September 1986.

directly related to the total number of events necessary to define all model parameters.

The qualification of a travel time observation by the observer or an automatic phase picker [Allen, 1982; Baer and Kradolfer, 1987] is as important a piece of information as the travel time itself, not only for the inversion but also for the selection of the data. My experience with the compilation of LED from different seismic networks suggests the use of a standardized format with observation weights such as the international code for teleseismic data. Between May 1980 and September 1986 the seismic earthquake network of the U.S. Gelogical Survey (Menlo Park, California) recorded some 30,000 events in the magnitude range M0.5 to M6.5 in the Long Valley area [Cockerham and Corbett, 1987; R.S.Cockerham, unpublished work,1988]. This data set serves as an example of how the LED sets used in this study for inversion have been selected.

6.2.1. Step 1: relocation of events. Using the minimum 1D model technique and station corrections for the Long Valley area, all 30,000 events were relocated. The vast majority of these earthquakes occurred in one of the three swarm areas (Figure 31): the south moat of the caldera (SM) and the southerly adjacent Sierran block, the Round Valley (RV), and the Chalfant Valley (CV), from west to east.

6.2.2. Step 2: selection for quality. Repeated use of the auxiliary program STATIS (see figure 11) selects those events which match the minimal quality requirements: maximal gap (160°) and minimal number (8) of p wave travel time observations with a certain minimal reading weight ('3') and maximal travel time residual (0.5 s). Thus we obtain a data set of 9200 events (190,000 observations) in the Long Valley area (Figure 32) that can be used for inversion.

6.2.3. Step 3: selection for distribution of hypocenters. Since the seismicity pattern in Long Valley is governed by the three swarm areas, the data set obtained by step 2 is of high redundancy. To obtain the more evenly distributed final data set of the Long Valley area with some 3600 events and 80'000 observations (Figure 33), we selected (1) all events outside the three swarm areas, (2) the several hundred best (maximum number of observations, minimum gap) events within the swarm areas, and (3) all additional events within the swarm areas that have readings at (temporary) stations where we have few other travel time data.The 5600 events that were rejected in the third step as redundant data can be used to form an additional data set for inversion to test the resulting velocity field.

6.3. Geometric Weighting

Geometric weighting may be necessary to control the data redundancy to a certain degree. In areas like Long Valley caldera, California, a large number of earthquakes occur in a fairly small volume while a comparable small number of events may be more evenly distributed over a larger area and depth. With geometric weighting we try to downweight the effects of the redundant data.

Let us assume a model volume, in our case a block, being sampled by 100 subparallel rays, originating from about the same region with the same station as recording site (Figure 34). We may further assume that this block is sampled also by a few rays passing in a direction perpendicular to the 100 previous rays through the block. This is, in fact, a quite common situation in many areas with an uneven distribution of the seismic activity. In our example we would like to control the results from the 100 rays with the help of the few crossing rays (leakage problem, poor resolution of different blocks along same ray path without cross rays). However, the solution for this particular block will be extremely dominated by the many subparallel rays. If we downweight the numerous subparallel rays, we will find a more reliable solution for this specific block. This is exactly what the geometric weighting is doing.

The (optional) geometric weighting is calculated by the programs SELEVE and RESOL (Figure 9). In program SELEVE, all LED are scanned to add up the number of events that have readings at specific stations and lie within a specific model block. Then the final weight (w_{tot}) applied to each observation from an event located in the kth block is calculated in program RESOL as the product of the original observation weight (w_{obs}) and the geometric weight (w_{geo})

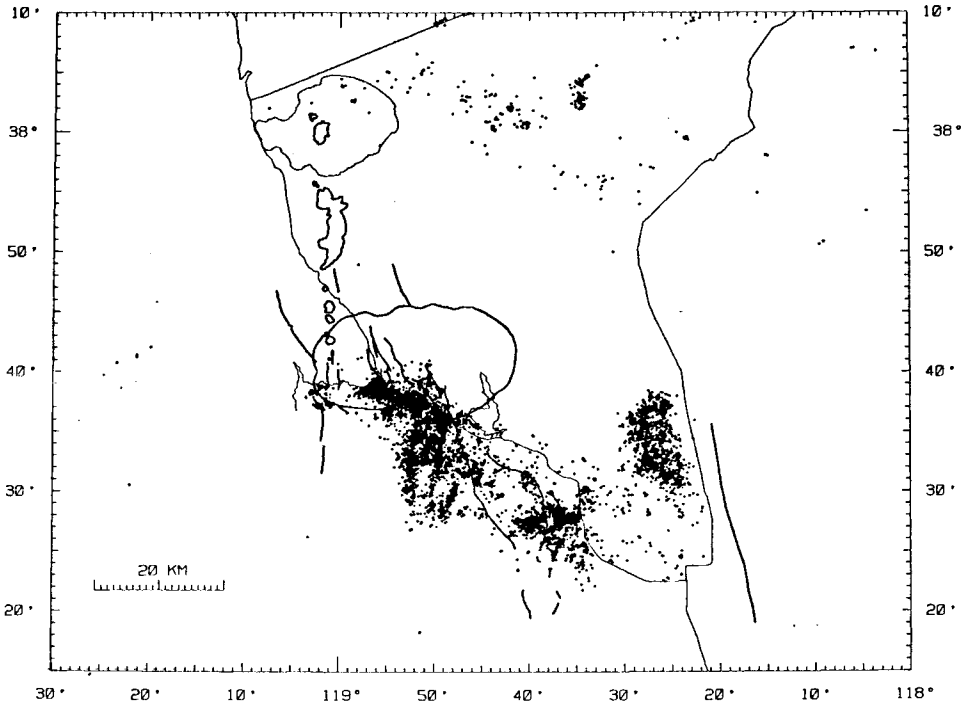

Fig. 32. Distribution of the 9200 best located events in the Long Valley, California, area between May 1980 and September 1986.

for the specific station (i) block (k) pair, given (nmin) and (N_{ik}):

$$w_{tot} = w_{obs} * w_{geo} = w_{obs} * \frac{1}{N_{ik} + nmin} \qquad (35)$$

where nmin is the scaling factor (positive integer; a value of nmin=4 has been used for the Long Valley and Yellowstone areas) and N_{ik} is the number of events from the kth block (i.e., the hypocentral area) which are recorded at the ith station

6.4. Variance of LED

Knowledge of the data variance is necessary to calculate the covariance matrix (see equation (24)) and to determine the significance level of the variance reduction (see equation (27)). The data variance of the relocated events is composed

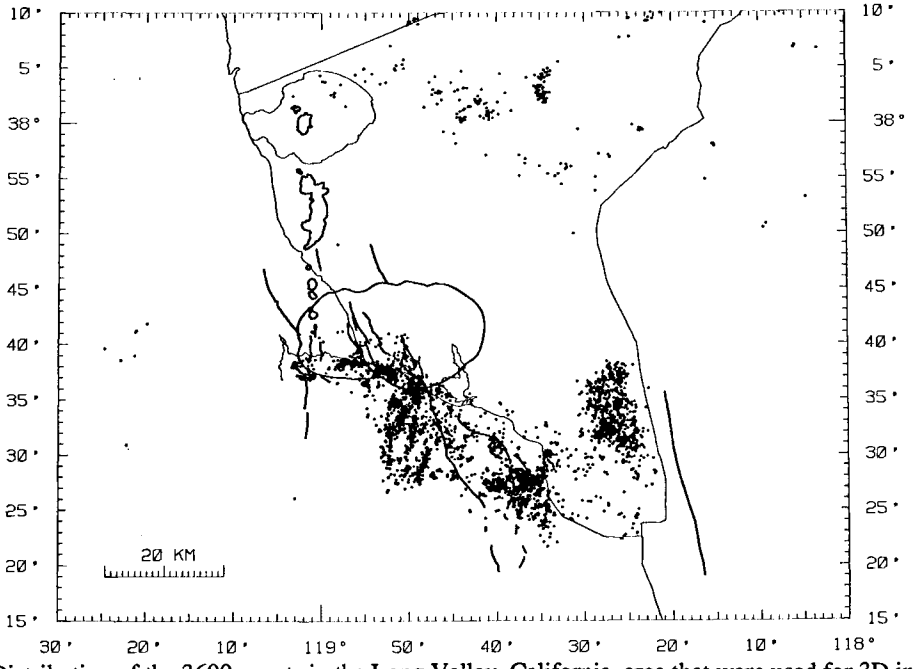

Fig. 33. Distribution of the 3600 events in the Long Valley, California, area that were used for 3D inversion.

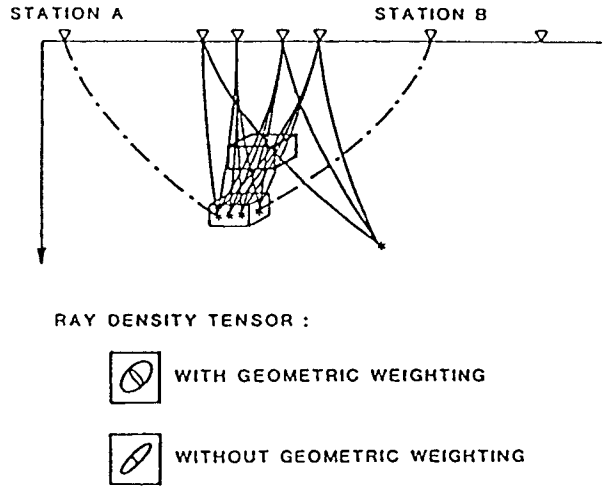

Fig. 34. Geometric weighting. Stars denote hypocenters of earthquakes used for 3D inversion. The lower block contains a large number of hypocenters that should all be used for inversion because of singular ray paths (dot-dashed lines) to stations A and B. The use of all these events, however, results in the large number of subparallel rays passing through the upper block.

of the true timing errors (for discussion of timing errors using an automatic phase picker, see Allen [1982]) and the travel time residuals due to the difference between the real velocity structure and the minimum 1D model used for the relocation of the earthquakes. These travel time residuals together with parts of the station corrections are the signal in the LED to be inverted to obtain an improved velocity model, while the true timing errors are the noise.

Velocity parameter adjustments are insignificant if they lead to a data variance smaller than the estimated true timing errors. The rms values (Figure 35a) are often used to display the overall quality of an earthquake location solution. However, for the coupled hypocenter-3D-velocity model inversion the weighting of travel time obser-vations with different reading weights (0 the best through 4 the lowest quality) may differ from the weighting in the single-event location procedure. Therefore, the input data variance for the 3D velocity inversion is best displayed as residual travel times for each class of observation weights (see Figures 35b through 35f).

Figures 35b through 35f display the data variance of the 9894 well-located events from Long Valley recorded between 1980 and 1986 (see section 6.2 for details). The estimated timing error of ±0.02 s for the 0 weight observations (Figure 35b) denotes about 50% of the displayed data variance, and this percentage remains roughly constant for the observations of lower weights considering the increase in the timing error for the lower-amplitude observations.

From the 3D inversion of LED from Long Valley a variance reduction of 50% was obtained by two iteration steps using the ACH method with a velocity damping parameter of 1.0. The remaining data variance was calculated by relocation of all events with program RELOC (see Figure 9) using the three-dimensional velocity model resulting from the 3D inversion procedure. The average station residuals which were obtained in this relocation process are smaller than the station corrections for the minimum 1D model by at least a factor of 2 (see section 3). In general, the hypocenter adjustments were less than 500 m in all directions. In contrast to pure estimates of these parameters which are commonly used, the exact calculation of the variance reduction and the reduction in the station residuals requires an enormous amount of computation time. However, such calculations may give some credibility to the resulting 3D velocity field, for which the significance is otherwise hard to judge.

The display of the data variance may also help to test and scale the reading weights applied to LED. In Figures 35b through 35f travel time residuals of the Long Valley data are displayed for each class of observations separately. With decreasing reading weight an increase in the data variance and a shift of the average is observed. There are also many more delayed arrivals (residuals of +0.2 s or more) than early arrivals (residuals of -0.2 s or less) for the classes of lower quality than for the observations with zero weight. This correlates with the expectation that readings of lower quality are more likely to be delayed than to be early, since the lower signal to noise ratio will tend to mask the low-amplitude onset of the first arriving wavelets.

7. Interpretation of Three-Dimensional p Wave Velocity Fields

The information obtained by the inversion of local earthquake data (p waves) may be summarized as relative changes in p wave velocity averaged over certain volumes (blocks). Note that in a strict sense, this information can be measured at a point; in pratice, LED give a result that is averaged within each inversion layer separately (see also discussion of resolution in section 2.5). Various reasons may account for such lateral variations in p wave velocity: (1) changes in lithology and phase transitions, (2) changes in texture (a decrease in p wave velocity due to numerous cracks in the rock, fractiles and high gas content (under high pressure)), large pore fluid volumes relating to a fault zone, such as in the Borah Peak area, (3) changes in temperature and pressure, and (4) anisotropy in p wave velocity. Any such change will neither solely affect the p wave velocity nor be detected by a specific signal in the p wave velocity field, e.g., gradients. We have to interprete a p wave velocity field which will certainly show the effects of combinations of such changes, effects we know little about at present.

To reduce the ambiguity of the petrologic and geologic interpretation of a three-dimensional p wave velocity structure, we need detailed information about the 3D structure of several other geophysical parameters, such as density, s wave velocity, and attenuation of seismic waves. However, at present we lack such information in most areas for various reasons. Rather, our knowledge of geophysical parameters other than the p wave velocity is, if detailed at all, of variable information density and in most cases of a more qualitative nature. The interpretation of a three-dimensional p wave velocity field will focus on such areas of detailed additional information and thus possibly exaggerate the importance of these places relative to the overall picture or overweight the reliability of some details in the p wave velocity field.

Whenever possible, the interpretation of p wave velocity structures must rely on qualitative additional information gained by other geophysical methods. Qualitative arguments will help to reduce the ambiguity in the interpretation. The observation of other geophysical parameters in specific locations of the area under study may allow some quantitative measure (volume percent of molten phase in a magma chamber). However, at present the qualitative information (is there clear evidence for a magma chamber?) is of greater importance.

Artificial data calculations may also give important insights for the interpretation of 3D velocity fields because the travel time residuals resulting from a local 3D velocity anomaly of the expected size might prove undetectable considering the noise of the data. In areas of recent volcanic activity, we may expect to see local anomalies of low p wave velocity, of very low s velocity, and of high attenuation in the vicinity of volcanic features on the surface. In orogenic belts like the

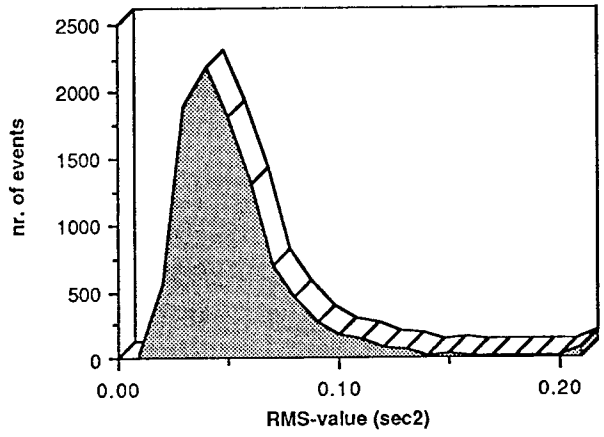

Fig. 35a. Long Valley, California: rms-distribution of 9894 best located events (rms is the root-mean-square value of travel time residuals of all observations from one event).

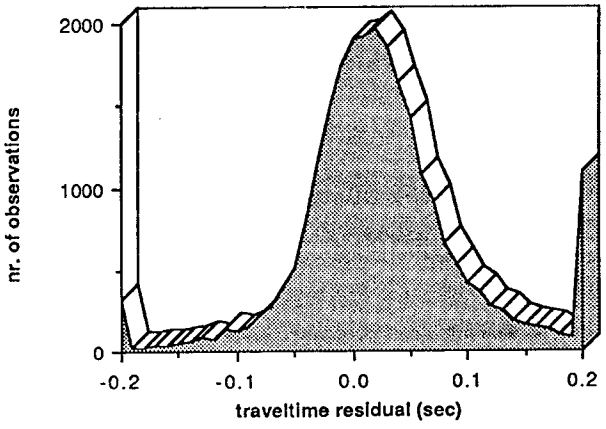

Fig. 35d. Long Valley, California. Distribution of travel time residuals (seconds) of 24,992 observations with reading weight of 2 (average: +0.015; variance: 0.001; extreme residuals have been neglected).

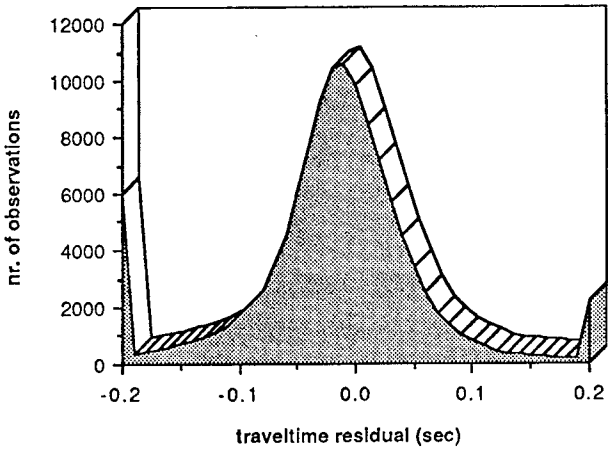

Fig. 35b. Long Valley, California. Distribution of travel time residuals (seconds) of 123,570 observations with reading weight of 0 (average: -0.002; variance: 0.0008).

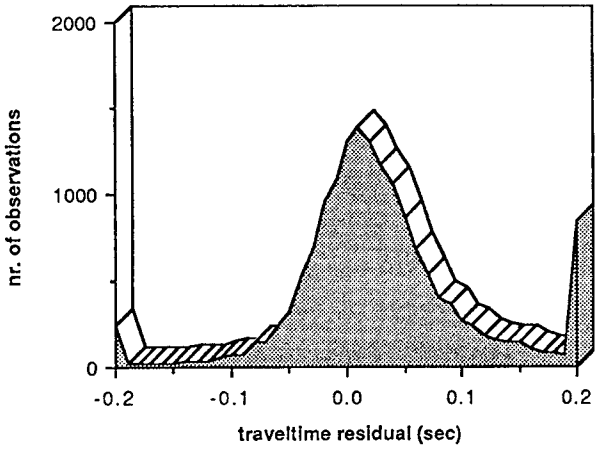

Fig. 35e. Long Valley, California. Distribution of travel time residuals (seconds) of 16,250 observations with reading weight of 3 (average: +0.022; variance: 0.002; extreme residuals have been neglected).

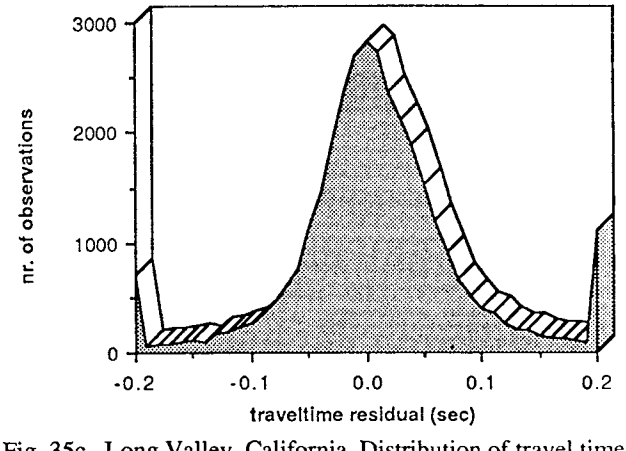

Fig. 35c. Long Valley, California. Distribution of travel time residuals (seconds) of 33,771 observations with reading weight of 1 (average: +0.001; variance: 0.0007; extreme residuals (>0.5 s) have been neglected).

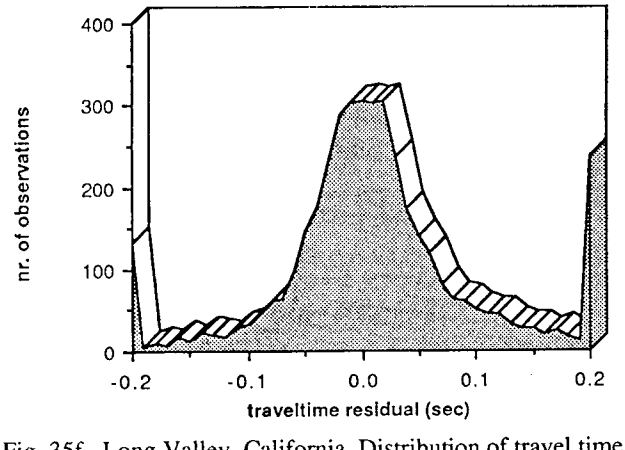

Fig. 35f. Long Valley, California. Distribution of travel time residuals (seconds) of 4,053 observations with reading weight of 4 (average: +0.08; variance: 0.002; extreme residuals have been neglected).

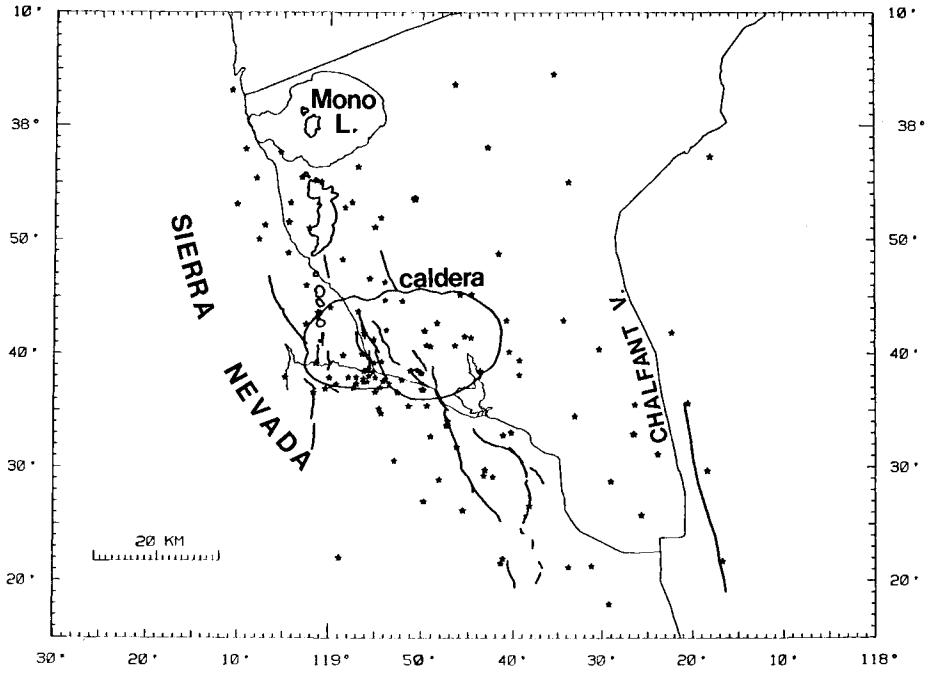

Fig. 36. Station distribution of the stations in the vicinity of Long Valley caldera which were used in the inversion of LED.

Alps, however, we obtain an excellent qualitative measure for a resulting 3D velocity field from the knowledge or expectation of an increase in the crustal thickness (Figure 19).

Of all other geophysical data, at present, Bouguer gravity data are the most easily available independent quantitative information. Ellsworth and Koyanagi [1977] and Oppenheimer and Herkenhoff [1981] rely upon arguments derived from gravity studies for their interpretations of three-dimensional velocity fields. However, as Aki [1982] points out, the velocity density relationship is affected by various factors, such as chemical composition, temperature, and pressure, and is by no means unique. This nonuniqueness specifically concerns the linear, layerwise translation of relative changes in p wave velocity into changes of bulk density. Furthermore, the smallest model unities (e.g., blocks of 2.5*2.5*2 km size), which are well suited to describe a 3D velocity field by its blockwise average slownesses, are of little use for 3D gravity models of the upper crust. The main signals to be interpreted in gravity are the gradients in the Bouguer gravity field. Given the densities of the tectonic units, the main information gained by gravity modeling is the geometry of the bodies of different density. Blocks of uniform density and size fit to describe the velocity field are very crude approximations of the variable geometry of gravitating bodies. In my opinion, the nonunique velocity-density relationship and the poor geometrical correlation between a reasonable smallest model unit of uniform density and uniform velocity put doubt on the value of simultaneous seismic-gravity inversion. Three-dimensional forward gravity modeling with bodies of arbitrary shape, which are defined by tectonic interpretations of a 3D velocity field, however, is likely to give additional contraints for the interpretation of the results from a 3D LED inversion.

8. Applications of the ACH Inversion for LED

As the various uses of LED in the previous sections implies, this method of inversion of LED has been established during its application to Long Valley, California, and Yellowstone, Wyoming. Though this paper concentrates on the inversion method, for completeness the resulting three-dimensional p wave velocity fields obtained by ACH inversion for the Long Valley and Yellowstone areas will be shown without discussion of their implications for the interpretation of the tectonic and volcanic structures.

8.1. Application of the ACH Method to Long Valley, California

The Long Valley caldera and the Inyo-Mono craters volcanic chain lie at the base of the eastern escarpment of the Sierra Nevada range in California. The escarpment is formed by large, east dipping normal faults that define the western margin of the Basin and Range province. This region is the tectonically most active area in eastern California, in terms of both persistent earthquake activity over the last century and frequent volcanism over the last 600,000 years.

In order to illuminate the middle crust beneath the Long Valley caldera from various directions, the area under study includes data from larger events in the southern Sierra Nevada and from the area south of the Tahoe basin (some 150 km northwest of the caldera). Thus the original data set encompasses all phase data from events within the area 35°20'N to 39°N and 117°W to 121°W for the time period May 1980 through September 1986. The 3800 selected events (see section 6) contain about 80,000 observations from some 200 stations. The station distribution of the Long Valley area (Figure 36) reflects the distribution of the seismicity (see Figure 32) with a con- centration of seismic stations in the southern part of the caldera. This causes great variations in the resolution capability, which was tested with artificial data calculations and has been discussed in section 5.

The many tests that were performed with the data set from the Long Valley area helped to tune the inversion procedure by setting the appropriate parameters such as the velocity damping factor. However, the resulting p wave velocity field of Long Valley, which is shown in Figure 37 is but a fraction of the results that ought to be displayed and taken into account for an interpretation and a thorough discussion. This will be

the subject of another paper (E.Kissling and W.L.Ellsworth, unpublished work, 1989).

In summary, the three-dimensional p wave velocity field of Long Valley shows several local anomalies of higher amplitude, particu-larly in the uppermost 5 km beneath the caldera and within the crustal block north and east of the caldera. On the basis of geologic data and refraction seismic profiling, some of these local anomalies can be attributed to changes in the lithology. Others can be seen in the context of the volcanic history. The most intriguing result of this 3D inversion, however, is the apparent absence of a large contem-poraneous magma chamber at midcrustal level beneath the Long Valley caldera, as was postulated by Steeples and Iyer [1976]. Small pockets of partially melted rocks at even shallow depth may exist in the vicinity of the Mono craters (between the caldera and Mono Lake) and in the central and western part of the caldera. A deep-seated anomaly of low p wave velocity exists beneath the Sierra Nevada about 30 km south of the caldera, reaching from the lower crustal level up to about 18 km below the surface. This body, which is outside the area displayed in Figure 37, will need further investi-gations with additional data because of the currently poor resolution at that depth.

8.2. Application of the ACH Method to Yellowstone, Wyoming

The Yellowstone National Park in Wyoming is one of the rare windows to the Earth's interior [Smith and Braile, 1982; Smith and Christiansen, 1980]. The volcanic-tectonic processes operating in the vicinity of this largest caldera on Earth give rise to spectacular surface features such as geysers, hot springs, mud volcanos, and young, steep canyons. A geologic history frozen into the rocks of the Snake River Plain (east-southeast of Yellowstone) marks the track of a hot spot [Smith et al., 1977]. With its current position beneath the Yellowstone caldera, this hot spot has come across the western part of the Rocky Mountains. Thus, the caldera with its young volcanic rocks is now surrounded by mostly Precambrian basement rocks. Therefore, the p velocities of the surface rocks already show large variations, a tectonic feature to keep in mind when interpreting a 3D velocity field.

The Yellowstone area is part of the Intermountain Seismic Belt [Smith and Sbar, 1974] and, as such, is a seismically active area (Figure 38). The seismicity of this region has been monitored with a varying station network since 1973 by the U.S. Geological Survey and the University of Utah, Salt Lake City. The earthquake catalog compiled by these two institutions for the time period 1973 to 1985 consists of over 10,000 events recorded at about 40 permanent and over 200 temporary stations. The station distribution of the Yellowstone area can be derived from the station delays shown in Figure 20. Though a smaller data set than that from Long Valley, the LED of the Yellowstone area provides an excellent means of illuminating the continental crust above an inferred hot spot by tomography, since the seismicity is more uniformly distributed and the travel time signals are expected to be of larger amplitude than those in Long Valley. The rather large distance to other seismically active areas outside the Intermountain Seismic Belt, however, limits the area of uniform high resolution to the direct vicinity of the Yellowstone caldera and a few more local areas in the west and in the south. Furthermore, these areas of high resolution are shifted at different depth levels. As a result the lower crust beneath the caldera is not resolved. Such features of a 3D velocity field are important for the interpretation and should therefore be included in the resulting p wave velocity field. In Figure 39 the areas that are not shaded denote blocks which are sampled by four or fewer rays and are therefore poor or not resolved at all.

The 3D velocity field of Yellowstone area (Figure 39) was obtained by ACH inversion of 3500 selected events. Again, as in the case of Figure 37, the interpretation of these results demands a thorough discussion of various other information that cannot be displayed here, which will be the subject of a different study. The exciting result of this inversion is the image of a large body of very low p wave velocity beneath the caldera. In correlation with various other geophysical data [Smith and Braile, 1982] this body of laterally low velocity can be interpreted as a magma chamber at upper crustal level (Figure 40).

9. Discussion

The method referred to as 'approximate geotomography' is a first-order approximation to the correct solution of a large system of linear equations by full inversions of the normal equation matrix (see equation (17)). With reference to the popular use of 'tomography' as a name for all inversion of observed signals from seismic waves to obtain the structure of the Earth's interior, neglecting the exact numerical solution process (see, for example, Fawcett and Clayton, [1984]; and Dziewonski and Anderson, [1984]), this method should be called approximate ACH inversion. The ACH inversion is a method of damped least squares that performs the matrix inversion by Cholesky decomposition of the normal equation matrix. In contrast ,the approximate ACH inversion (earlier called approximate geotomographic method) performs the matrix inversion of the normal equation matrix by simply using the inverse of the diagonal elements, neglecting all off-diagonal elements in the symmetric matrix. The results obtained by this approximate inversion are based on the assumption of a uniform perfect resolution, e.g., a known gross exaggeration of the true resolution. Since this assumption systematically overestimates the resolution of the poorly resolved velocity model parameters, the solution will be biased more by the poorly resolved blocks, while the solutions of the well-resolved blocks are less affected. Thus, overdamping of the system is probably a good idea. However, such overdamping may demand the use of an iterative process (see equations (18) and (20)) to obtain a reasonable estimate, and with a few iteration steps only, the computational burden reaches the same order as for the full ACH inversion by Cholesky decomposition.

Most applications that invert seismic travel time data for velocity parameters introduce some modifications of the inversion method due to data characteristics and/or for technical reasons such as computational burden. As a result, the interpretative comparison between solutions obtained with different techniques and different data sets is often misleading considering the intrinsic ambiguity in all inversions of seismic data. Different techniques often use different reliability checks for the results obtained. Since adjustments to the inversion technique for data characteristics such as the differences between teleseismic data and LED will always be necessary, a universal inversion technique for seismic data does not exist. In order to compare the inversion results for two adjoining areas and the same type of data or for two solutions of the same area but obtained from different types of data, a qualitative measure of the reliability, resolution, and ambiguity is necessary. The ambiguity can be dealt with in several ways (see, for example, Jackson [1976] and [1979]), a simple way being the calculation of more than one inversion solution. A single inversion for an area and a data set is almost useless, since all inversion parameters have to be assumed without testing their effects. Any of these inversion parameters (velocity and hypocenter location damping, block cutting, update of ray path) that might be well tuned for a given area and data set might be far off for another application. Some of these parameters have to be estimated for

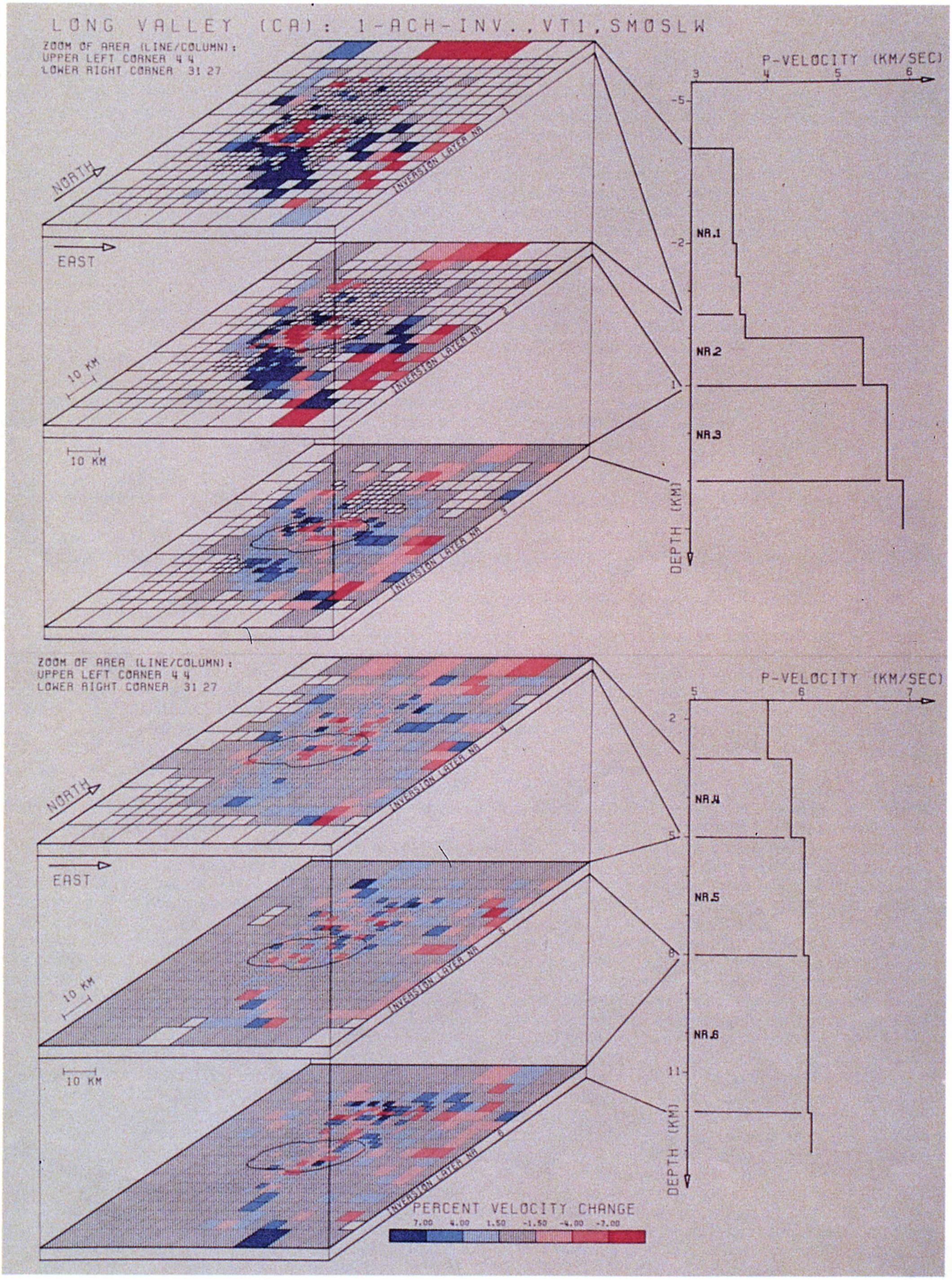

Fig. 37a. The three-dimensional p wave velocity field of the upper 15 km of the crust in the vicinity of Long Valley caldera obtained from ACH inversion (one iteration step) of about 80,000 observations of about 3800 local earthquakes using a velocity damping parameter of 1.0. The caldera is outlined by a solid black line.

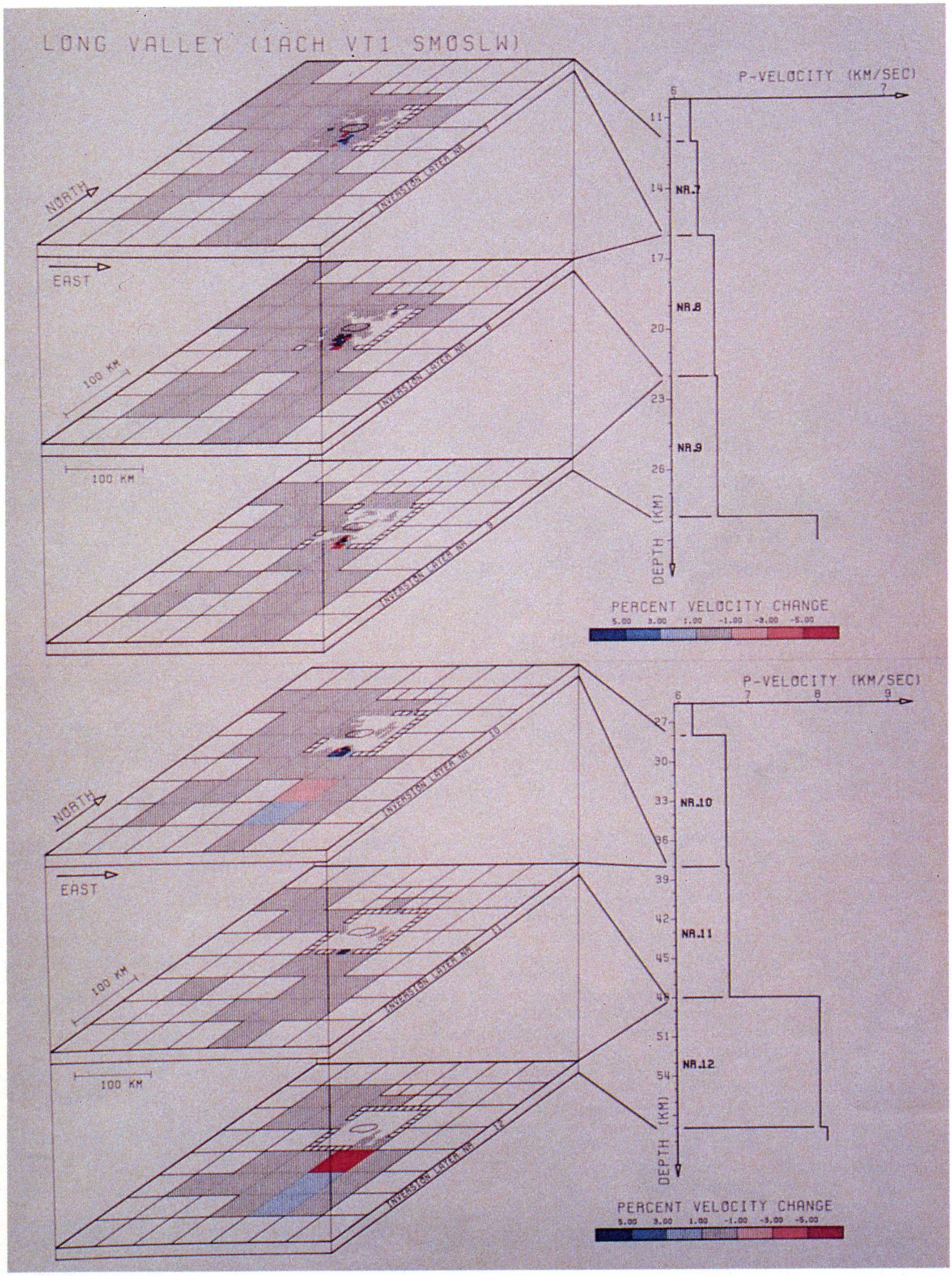

Fig. 37b. The three-dimensional p wave velocity field of the lower 35 km of the crust in the vicinity of Long Valley caldera obtained from ACH inversion (one iteration step) of about 80,000 observations of about 3800 local earthquakes using a velocity damping parameter of 1.0. The caldera is outlined by a solid black line.

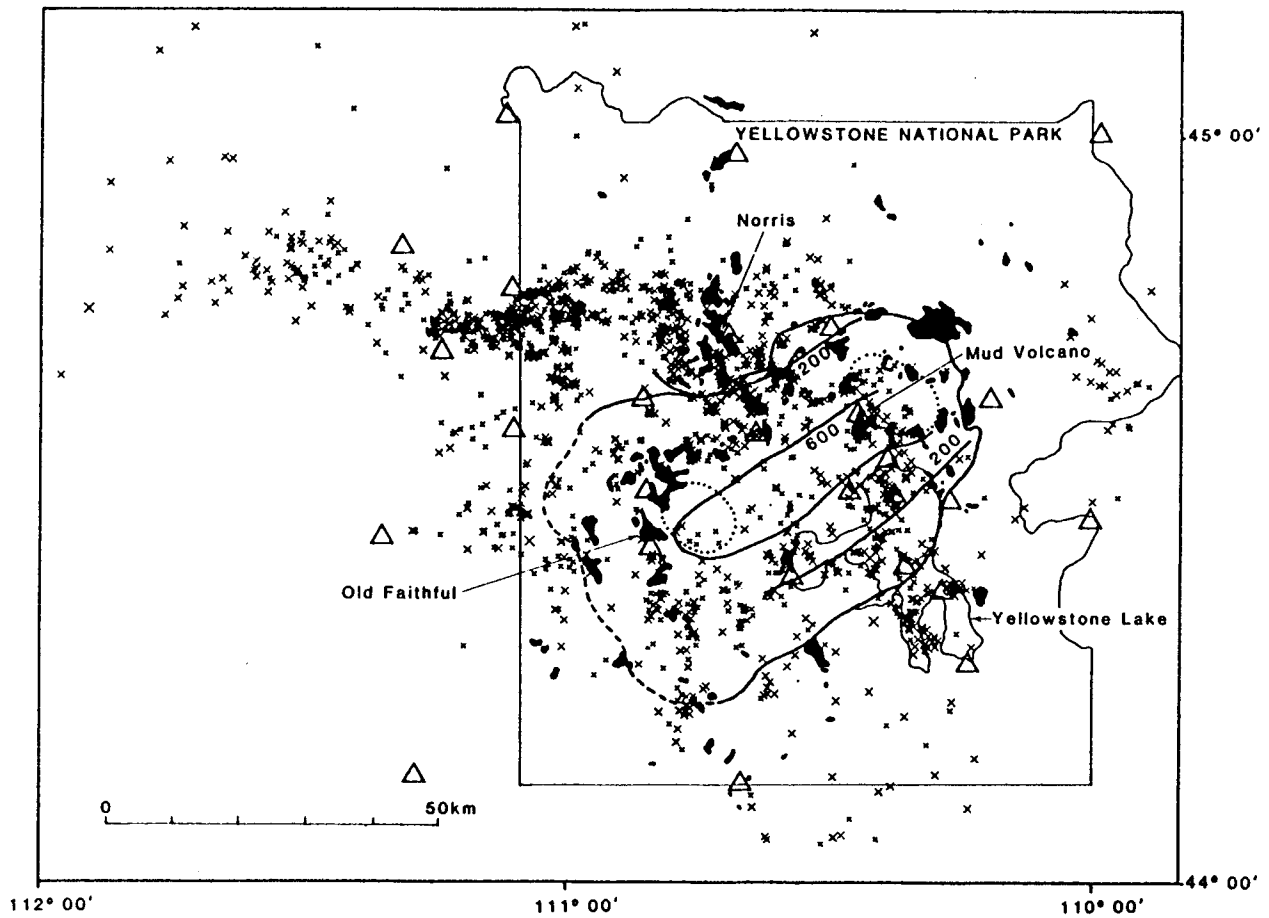

Fig. 38. Earthquake epicenters from 1973 to 1980 (indicated by crosses) in Yellowstone National Park and vicinity [Pitt and Hutchinson, 1982]. The caldera is outlined by solid and dashed line; the resurgent domes are outlined by dotted lines; the 200- and 600-mm contours of caldera uplift [from Pelton and Smith, 1979] are shown as solid lines. Triangles mark seismic stations. Hydrothermal areas are solid.

each application and may not be generalized. Therefore, test runs are necessary to compare the results obtained with different inversion parameters.

Naturally, most improvements on the inversion procedure require additional calulation time. The implementation of the resolution data into the approximate geotomographic inversion (step length weighting) and the tests with artificial data sets increase the overall calculation time for a 3D velocity inversion by at least a factor of 2. The calculation of the starting one-dimensional model with the appropriate station corrections and the relocation of all earthquakes may account for 30% of the total calculation time. As I tried to show, all these special calculations and tests are necessary to ensure a reasonable result of a 3D inversion. If the approximate geotomographic inversion as but a rough approximation to the full ACH inversion is still proposed as a reasonable method for larger data sets, this is the case, because I believe the tests mentioned and the additional calculations to be even more important than the precise numerical solution of the inverse problem. The ambiguity intrinsic to the inversion method and the restrictive assumptions necessary for the calculation of the forward and the coupled inverse problems demand tests and multiple inversion runs to gain credibility for the results.

Many tests would not be necessary and the computational burden would be greatly reduced if one could simply perform any kind of full inversion on a much smaller number of more reliable events and observations. However, what are "more reliable events"? If the true errors are distributed randomly, it should be possible to get more reliable data by building some average signal from a number of events located within the same hypocentral area. We could also build the average residual at each station for all events in the same hypocentral area. The latter is possible with the program DATVAR (see Figure 9) and results in an estimate of the data variance for each hypocentral area-station pair. This is, however, but the first step toward the calculation of synthetic events as summaries over all events that occur in given hypocentral areas. At present we lack the information necessary to perform such calculations, as there are open questions like the following:

1. What is the volume that can be assumed as the identical hypocentral area considering the random and systematic errors in our earthquake location procedure?
2. What is the travel time effect of the approximation of the 3D-wavefield by 1D or 2D ray tracing?
3. How will the different observations at a station be weighted considering the different magnitudes of the earthquakes?

To avoid the loss of information about the velocity structure and in order not to systematically mix signal and noise in the data, for the moment we have to use the larger number of data and thus perform the inversion of large matrices. We have to keep in mind, however, the assumptions about the data that were necessary first to formulate the inversion procedure and second to assure results that correlate with the true velocity structure:

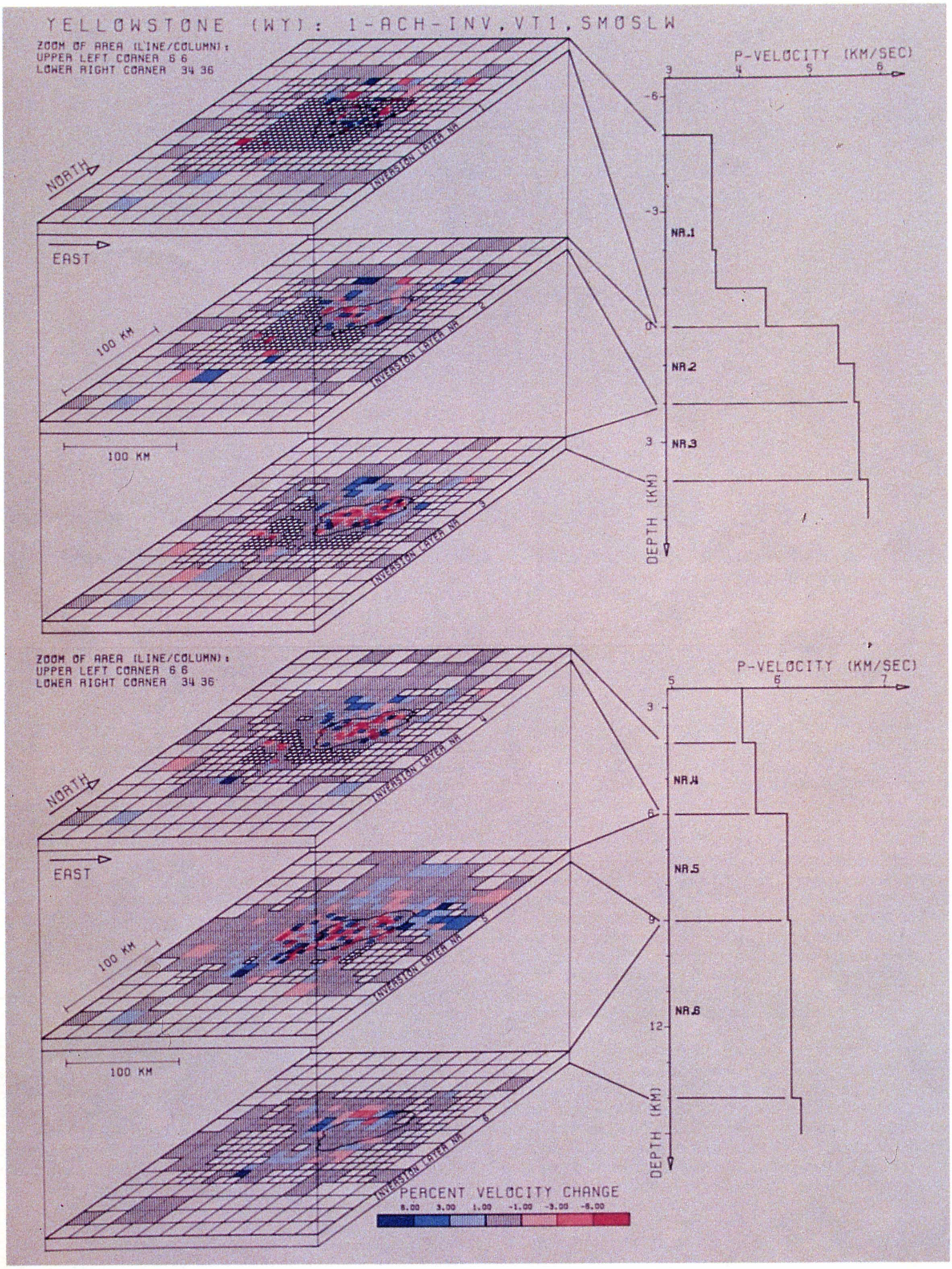

Fig. 39. The three-dimensional p wave velocity field in the vicinity of Yellowstone caldera (outlined in solid blue) obtained from ACH inversion (one iteration step) of about 40,000 observations of about 3500 local earthquakes using a velocity damping parameter of 2.0.

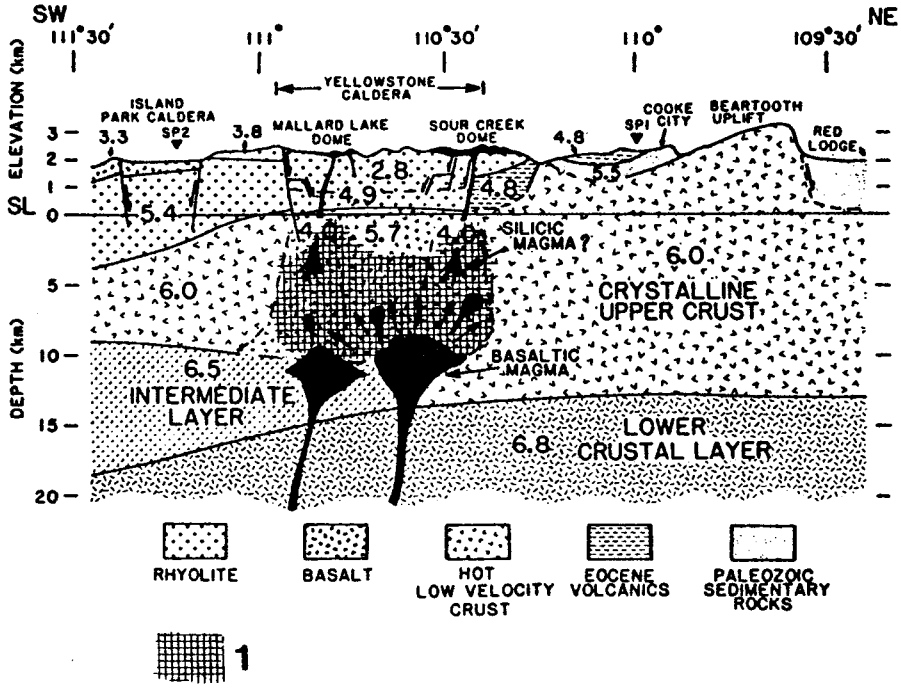

Fig. 40. Idealized NE-SW geologic seismic velocity model for the crustal structure of the Yellowstone-Island Park-Snake River Plain region [from Smith and Braile, 1982]. P wave velocities are in kilometers per second. Geologic interpretations are based on constraints of the Quaternary history and the petrologic models of R.L. Christiansen (U.S. Geological Survey, personal communication, 1978) and Lehman [1980]. Low velocities in lower crust and upper mantle are inferred from teleseismic delay studies [Iyer et al., 1981]. Model for basaltic melt rising from the upper mantle into the crust is from Hildreth [1981]. The crosshatched area (1) marks the position of the (averaged) low-velocity body resulting from this study of inversion of local earthquake data.

1. We can only image smooth variations of the slowness. Thus the minimum length of a lateral p wave velocity anomaly is several times larger than the wavelength of the main amplitude in the data. Con-sidering the model variance, a significant local p wave velocity anomaly also consists of several blocks for a discretized model, as in the ACH method. Single block anomalies might be reliable only when a consistent resulting pattern occurs for several inversions with different block cutting. The results obtained by inversion of seismic travel times are nonunique. Often this ambiguity concerns structures of less interest, while the possible solutions for the main target of the inversion show only differences that are insignificant for the interpretation. Thus several different inversions performed on the same data in search of different resulting velocity fields can help to reduce the negative effects of the ambiguity intrinsic to inversion methods.

2. As the test with mislocated stations showed, some systematic errors in the data are imaged by the inversion process into lateral changes in the velocity field that can never be separated in the inversion results from the true illuminated structures. To gain any credibility for the resulting inhomogeneous velocity field, the data should always be displayed and filtered for known and even for only possible systematic deficiencies before an inversion is performed. In the Long Valley area, where we have refraction seismic data for a large part of the caldera and the surrounding area, the velocities calculated by 1D inversion for the minimum 1D model are systematically 2-3% lower than the velocities found by refraction for the same layer. The reason given by Wielandt [1987] is that diffraction around the bodies of lower p wave velocity will affect the average layer velocity. This effect depends on the dimensions of the anomalous body and the directions of the rays. Travel times for ray paths that penetrate the body of lower p wave velocity in general will still be the first arrivals if the dimensions of the body perpendicular to the ray are several times the dimension of the body along the ray. With the exception of small spherical bodies, the local earthquake tomography with rays from all directions and from various locations in the upper crust will be less affected by such diffractions than refraction seismic profiling, where most rays lie in a ray tube. Since this effect will tend to hide the local anomalies of lower velocity, only the average layer velocity in refraction seismic interpretation should be larger than the corresponding velocities calculated from inversion of LED.

Acknowledgments. This paper is the result of several years of study in two different working groups. Funded by the Schweizerischer Nationalfonds and the National Science Foundation (82.998.0.82), I spent two exciting years with the Branch of Seismology of the U.S. Geological Survey in Menlo Park, California. I thank the many people of the USGS that were of great help. I am especially thankful for the tutorial guidance, the scientific support, and the personal friendship that I received from D.P. Hill, W.L. Ellsworth, and R.B. Cockerham. The list of dear friends that contributed in one way or the other to this study is very long. I hope they understand the sincerity of my thanks, though they shall remain anonymous. The staff of the Institut für Geophysik at the Eidgenössische Technische Hochschule in Zürich provided steady support during the second part of this study. I am thankful for this help and for the generous acceptance of a "Californized" former student. For fruitful and often long discussions on various aspects of the inversion of earthquake data I thank M. Baer, E. Wielandt, D. Mayer-Rosa, and St. Mueller. I am grateful to R.B. Smith, J. Ansorge, W.L.

Ellsworth, and A. Hill for critically reading the manuscript. Anybody familiar with the joy, the stress, and the frustrations that come with the coding of larger computer programs knows of the invaluable suggestions and the psychological support only a conspirative friend that shares the experiencies can offer. I thank U. Kradolfer for being such a friend.

References

Aki, K., Three-dimensional seismic velocity anomalies in the lithosphere, J. Geophys., 43, 235-242, 1977.

Aki, K., Three-dimensional seismic inhomogeneities in the lithosphere and asthenosphere: Evidence for decoupling in the lithosphere and flow in the asthenosphere, Rev. Geophys., 20(2), 161-170, 1982.

Aki, K., and W.H.K. Lee, 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(23), 4381-4399, 1976.

Aki, K. and P.G. Richards, Quantitative Seismology; Theory and Methods, 932 pp., San Francisco, W.H. Freeman , California,1980.

Aki, K., A. Christoffersson, and E.S. Husebye, Three-dimensional seismic structure of the lithosphere under Montana LASA, Bull.Seismol. Soc. Am., 66(2), 501-524, 1976.

Aki, K., A. Christoffersson, and E.S. Husebye, Determination of the three-dimensional seismic structure of the lithosphere, J. Geophys. Res., 82, 277-296, 1977.

Allen, R., Automatic phase pickers: Their present use and future prospects, Bull. Seismol. Soc. Am., 72(6), S225-S242, 1982.

Backus, G., and F. Gilbert, Numerical applications of a formalism for geophysical inverse problems, Geophys. J. R. Astron. Soc., 13, 247-276, 1967.

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

Baer, M., and U. Kradolfer, An automatic Phase Picker for local and teleseismic Events, Bull.Seismol. Soc. Am., 77(4), 1437-1445, 1987.

Benz, H.M., and R.B. Smith, Simultaneous inversion for lateral velocity variations and hypocenters in the Yellowstone region using earthquake and refraction data, J. Geophys. Res., 89(B2), 1208-1220, 1984.

Burmakov, J.A., A.V. Treussov,and L.P. Vinnik, Determination of three-dimensional velocity structure form observations of refracted body waves, Geophys. J. R. Astron. Soc., 79, 285-292, 1984.

Byerly, P.: Near earthquakes in central California, Bull. Seismol. Soc. Am., 29(3), 427-462, 1939.

Chander, R., On tracing seismic rays with specified end points, J. Geophys., 41, 173-177, 1975.

Christoffersson, A. and E.S. Husebye, On three-dimensional inversion of P wave time residuals: Option for geological modeling, J. Geophys. Res., 84(B11), 6168-6176, 1979.

Cockerham, R.S. and E.J. Corbett, The July 1986 Chalfant Valley, California, earthquake sequence: Preliminary results Bull. Seismol. Soc. Am., 77, 280-289, 1987.

Crosson, R. S., Crustal structure modeling of earthquake data, 1, Simultaneous least squares estimation of hypocenter and velocity parameters, J. Geophys. Res., 81(17), 3036-3046, 1976.

Dziewonski, A.M., Mapping the lower mantle: Determination of lateralheterogeneity in P velocity up to degree and order 6, J. Geophys.Res., 89(B7), 5929-5952, 1984.

Dziewonski, A.M., and D.L. Anderson, Seismic tomography of the Earth's interior, Am. Sci., 72, 483-494, 1984.

Dziewonski, A.M., B.H. Hager, and R.J. O'Connell, Large-scale heterogeneities in the lower mantle, J. Geophys. Res., 82(2), 239-255, 1977.

Eberhart, D., Three-dimensional velocity structure in northern California coast ranges from inversion of local earthquakes, Bull. Seismol. Soc. Am., 76(4), 1025-1052, 1986.

Efron, B,. and G. Gong, A leisurely look at the Bootstrap, the Jackknife, and the cross-validation, Am. Stati., 37(1), 36-48, 1983.

Ellsworth, W.L., Three-dimensional structure of the crust and mantle beneath the island of Hawaii, Ph.D. thesis, Mass. Inst. of Technol., 327 pp., Cambridge, 1977.

Ellsworth, W.L., and R.Y. Koyanagi, Three-dimensional crust and mantle structure of Kilauea volcano, Hawaii, J. Geophys. Res., 82(33), 5379-5394, 1977.

Evans, J.R., Compressional wave velocity structure of the upper 350 km under the eastern Snake River Plain near Rexburg, Idaho, J. Geophys. Res., 87(B4), 2654-2670, 1982.

Fawcett, J.A., and R.W. Clayton, Tomographic reconstruction of velocity anomalies, Bull. Seismol. Soc. Am., 74, 2201-2219, 1984.

Franklin, J.M, Well-posed stochastic extensions of ill-posed linear problems, J. Math. Anal. Appl., 31, 682-716, 1970.

Gutenberg, B., and C.F. Richter, Structure of the crust, continents and oceans, in Internal Constitution of the Earth, 2nd ed., edited by B. Gutenberg, pp. 314-339, Dover, New York, 1951.

Hearn, T.M., and R.W. Clayton, Lateral velocity variations in southern California, I, Results from the upper crust from Pg waves, Bull. Seismol. Soc. Am., 76(2), 495-509, 1986.

Hildreth,W., Gradients in silicic magma chambers, implications for lithospheric magmatism, J. Geophys. Res., 86, 10,153-10,192., 1981.

Hill,D.P., E. Kissling, J.H. Luetgert, and U. Kradolfer, Constraints on the upper crustal structure of the Long Valley-Mono craters volcanic complex, eastern California, from seismic refraction measurements, J. Geophys. Res., 90, 11,135-11,150, 1985.

Humphreys, E., R.W. Clayton, and B.H. Hager, A tomographic image of mantle structure beneath southern California, Geophys. Res. Lett., 11(7), 625-627, 1984.

Iyer, H.M., J.R. Evans, G. Zandt, R.M. Stewart, J.M. Coakley, and J.N. Roloff, A deep low-velocity body under the Yellowstone caldera, Wyoming: Delineation using teleseismic P wave residuals and tectonic interpretation: Summary, Geol. Soc. Am. Bull., 92, 792-798, 1981.

Jackson, D.D., Interpretation of inaccurate, insufficient and inconsistent data, Geophys. J. R. Astron. Soc., 28, 97-109,1972.

Jackson, D.D., Most squares inversion, J. Geophys. Res., 81(5), 1027-1030, 1976.

Jackson, D.D., The use of a priori data to resolve non-uniqueness in linear inversion, Geophys. J. R. Astron. Soc., 57, 137-157, 1979.

Jackson, P.L., Digital simulation of seismic waves, Ph.D. thesis, 84 pp., Univ.of Michigan, Ann Arbor, 1970.

Jordan,T.H., Continents at the core mantle boundary. Seminary talk presented at the Branch of Seismology, U.S. Geological Survey, Menlo Park, California, September 1986.

Julian, B.R., Ray tracing in arbitrarily heterogeneous media, Tech. Note 1970-45, 17pp., Lincoln Lab., Mass. Inst. of Technol., Cambridge, 1970.

Julian, B.R., and D. Gubbins, Three-dimensional seismic ray tracing, J.Geophys., 43, 95-113, 1977.

Kissling, E., W.L. Ellsworth, and R.S. Cockerham, Three-dimensional structure of the Long Valley Caldera, California, region by geotomography, U.S. Geol. Surv.Open File Rep. 84-939, 188-220, 1984.

Klein, R.W., Hypocenter location program HYPOINVERSE, I, Users guide to versions 1,2,3, and 4, U.S. Geol. Surv. Open File Rep. 78-694, 1978.

Koch, M., Die Bestimmung lateraler Geschwindigkeits-

inhomogenitäten aus der linearen und nichtlinearen Inversion tele- und lokalseismischer Laufzeiten - Anwendung auf die seismische Zone Vrancea, Rumänien, Ph.D. thesis, Univ. Karlsruhe, West Germany.
Lanczos, C., Linear Differential Operators, Van Nostrand, London, 1961.
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 Rep. 75-311, 1975.
Lee, W.H.K. and S.W. Stewart, Principles and applications of microearthquake networks, Academic, San Diego, California, 1981.
Lehman, J.A., Upper-crustal structure beneath Yellowstone National Park from seismic refraction and gravity observations, M.S. thesis, 92pp. Univ. of Utah, Salt Lake City, 1980.
Leu,L.L., Three-dimensional velocity structure of the 1983 M7.3, Borah Peak, Idaho, earthquake area using tomographic inversion of aftershock travel times, M.S. thesis, Univ. of Utah, Salt Lake City, 1986.
Levenberg, K., A method for the solution of certain non-linear problems in least squares, Quant. Appl. Math., 2, 164-168, 1944.
Lienert, B.R., E. Berg, and L.N. Frazer, Hypocenter: An earthquake location method using centered, scaled, and adaptively damped least squares, Bull. Seismol. Soc. Am., 76(3), 771-783, 1986.
Menke, W., The resolving power of cross-borehole t omography, Geophys.Res. Lett., 11, 105-108, 1984.
Mueller,G., Approximate treatment of elastic body waves in media with spherical symmetry, Geophys. J. R. Astron. Soc., 23, 435-449, 1971.
Mueller, S., and M. Landisman, Seismic studies of the Earth's crust in continents, I, Evidence for a low-velocity zone in the upper part of the Lithosphere, Geophys. J. R. Astron. Soc., 10, 525-538, 1966.
Mueller, S., and E. Peterschmitt, Detailed crustal structure from deep-seismic sounding and near-earthquake studies. Pure Appl. Geophys., 119, 1192-1196, 1981.
Neumann, G., and J. Behrens, Inversion of seismic data using tomographical reconstruction techniques for investigations of laterally inhomogeneous media, Geophys. J. R. Astron. Soc., 79, 305-315, 1984.
Nolet, G., Linearized inversion of (teleseismic) data, in The Solution of the Inverse Problem in Geophysical Interpretation, edited by R. Cassinis, pp. 9-37, Plenum, New York,1981.
Nolet, G., Seismic Tomography, D. Reidel, Hingham, Mass., 1987.
Oppenheimer, D.H., and K.E. Herkenhoff, Velocity-density properties of the lithosphere from three-dimensional modeling at The Geysers-Clear Lake region, California, J. Geophys. Res., 86(B7), 6057-6065, 1981.
Parker, R.L., Linear inference and underparameterized models, Rev. Geophys., 15(4), 446-456, 1977a.
Parker, R.L., Understanding inverse theory, Annu. Rev. Earth Planet. Sci, 5, 35-64, 1977b.
Pavlis, G.L., and J.R. Booker, The mixed discrete-continuous inverse problem: Application to the simultaneous determination of earthquake hypocenters and velocity structure, J. Geophys. Res., 85(B9), 4801-4810, 1980.
Pelton, J.R., and R.B. Smith, Recent crustal uplift in Yellowstone National Park, Science, 206, 1179-1182, 1979.
Pereyra, V., W.H.K. Lee, and H.B. Keller, Solving two-point seismic-ray tracing problems in a heterogenous medium, Bull. Seismol. Soc. Am., 70(1), 79-99, 1980.
Pitt, A.M., and R.A. Hutchinson, Hydrothermal changes related to Earthquake activity at Mud volcano, Yellowstone National Park, Wyoming, J. Geophys. Res., 87(B4), 2762-2766, 1982.
Smith, R.B., and L.W. Braile, Crustal structure and evolution of an explosive silicic volcanic system at Yellowstone National Park, in Symposium on Explosive Volcanism, pp. 96-109, National Academy Press, Washington, D.C., 1982.
Smith, R.B., and R.L. Christiansen, Yellowstone Park as a window on the Earth's interior, Sci. Am., 242, 104-117, 1980.
Smith, R.B., and M.L. Sbar, Contemporary tectonics and seismicity of the western United States with emphasis on the Intermountain seismic belt, Geol. Soc. Am. Bull., 85, 1205-1218, 1974.
Smith,R.B., R.T. Shuey, J.R. Pelton, and J.T. Bailey, Yellowstone hot spot: Contemporary tectonics and crustal properties from new earthquake and aeromagnetic data, J. Geophys. Res., 82, 3665-3676, 1977.
Smith, R.B., et al., The 1978 Yellowstone-eastern Snake River Plain seismic profiling experiment: Crustal structure of the Yellowstone region and experiment design, J. Geophys. Res., 87, 2583-2596, 1982.
Steeples, D.W., and H.M. Iyer, Low-velocity zone under Long Valley as determined from teleseismic events, J. Geophys. Res., 81, 849-860, 1976.
Suetsugu, D., and I. Nakanishi, Tomographic inversion and resolution for Rayleigh wave phase velocities in the Pacific Ocean, J. Phys.Earth, 33, 345-368, 1985.
Thurber, C.H., Earth structure and earthquake locations in the Coyote lake area, central California, Ph.D. thesis, Mass. Inst. of Technol, Cambridge, 1981
Thurber, C.H., Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, J. Geophys.Res., 88(B10), 8226-8236, 1983.
Thurber, C.H., A fast algorithm for three-dimensional seismic raytracing (Abstract), Eos Trans. AGU, 67(16), 304, 1986.
Wesson, R.L., Travel-time inversion for laterally inhomogeneous crustal velocity models, Bull. Seismol. Soc. Am., 61(3), 729-746, 1971.
Wielandt, E., On the validity of the straight-ray approximation for interpreting delay times, in Seismic Tomography, edited by G. Nolet, pp. 85-98, D. Reidel, Hingham, Mass., 1987.
Wiggins, R.A., The general linear inverse problem: Implication of surface waves and free oscillations for Earth structure, Rev. Geophys., 10(1), 251-285, 1972.
Woodhouse, J.H., and A.M. Dziewonski, Mapping the upper mantle:Three-dimensional modeling of Earth structure by inversion ofseismic waveforms, J. Geophys. Res., 89(B7), 5953-5986, 1984.
Zandt, G., Study of three-dimensional heterogeneity beneath seismic arrays in central California and Yellowstone, Wyoming, Ph.D. thesis, 274 pp., Mass. Inst. of Technol., Cambridge, 1978.

E. Kissling, Institut für Geophysik, ETH-Hönggerberg, CH-8093, Zürich, Switzerland.

(Received September 2, 1987;
accepted September 30, 1988.)