INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE CENTRATOM TRIESTE

H4.SMR/782-7

Second Workshop on
Three-Dimensional Modelling of Seismic Waves
Generation, Propagation and their Inversion

7 - 18 November 1994

Adaptive Grid Tomography

A. Michelini

Università di Trieste
Istituto di Geodesia e Geofisica
Trieste, Italy

Geophys. J. Int. (1993) **115**, 337–343

Velocity model inversion using parametric curves

Alberto Michelini

CNR-Istituto Talassografico Sperimentale, Viale Romolo Gessi 2, 34110 Trieste, Italy
Dipartimento di Ingegneria Navale, del Mare e per l'Ambiente, Università di Trieste, Via A. Valerio 10, 34127 Trieste, Italy

Accepted 1993 March 12. Received 1993 March 9; in original form 1993 February 1

SUMMARY

The use of parametric curves to define velocity models in seismic traveltime tomography is proposed. This approach provides increased flexibility in shaping the velocity models because it allows inversion for both velocity and grid point adjustments. The method is applied to a data set generated from a 1-D profile having sharp discontinuities. The results are compared with those obtained from other methods currently in use for traveltime tomography. It is found that large gradients and discontinuities can be retrieved with improved accuracy and that instabilities caused by non-optimal prior selection of the inversion grid can be avoided. The results appear to be robust when the sizes of both data and model space are reduced which makes the method appealing for the solution of large-scale tomographic problems.

Key words: crustal models, model parameterization, seismic inversion.

1 INTRODUCTION

The goal of seismic traveltime tomography is to recover the *true* Earth's velocity structure. However, the Earth displays a variety of 3-D inhomogeneities that are difficult to model accurately even when the data set has enough resolving power to resolve them. This difficulty derives mainly from the inability of the model parameterization and/or discretization to shape properly the *true* velocity model (Michelena & Harris 1991). In fact, both parameterization and discretization are generally held fixed throughout the inversion so that the velocity adjustments alone (or any other coefficient depending on the adopted type of basis function expansion) may not be sufficient for accurate shaping of the true velocity field. For example, a first-order discontinuity which finds the velocity grid mesh ill-positioned for its imaging will produce artificial fluctuations or other effects (depending on the adopted basis functions) onto the resolved model. This is a well-known problem which is commonly alleviated by the introduction of dense grid meshes and the selection of more accurate interpolation schemes (e.g. Sambridge 1990). Nonetheless, these 'cures' may not be optimal because: (1) the larger size of the parameter space can make the inversion computationally intractable; (2) the inversion needs to be closely monitored and regularized to avoid the insurgence of artefacts; (3) whilst providing increased flexibility in shaping the velocity models, denser grid meshes still do not insure optimal positioning of the inversion grid points; and (4) prior selection of a basis function for parameterization constrains the resulting models to the degree of continuity offered by the selected basis functions.

Ideally, it would be desirable to vary the grid spacing and the parameterization locally, the local adjustments being based on the resolving power of the data set and on the presence of large gradients. Recently, Michelena & Harris (1991) have proposed the adoption of 'natural pixels' whose calculation derives from the geometry of the ray paths, to obviate the *a priori* selection of the inversion grid. The quality of the images obtained with their method is comparable to that obtained with standard constant-velocity pixel parameterizations, but their technique has the advantage of decreasing the number of the model parameters by two orders of magnitude.

In this paper, an approach based on the use of parametric curves to define the velocity model is pursued. This approach, which is commonly adopted in computer graphics design, has at least two features relevant to tomography. First, it offers increased flexibility in defining the model grid spacing with any type of basis function. This can be especially useful when 3-D grids are used in that the user is not obliged to duplicate particular grid geometries on parallel sections of the model. Secondly, this approach allows inversion for the position of the grid points so that it is possible to use a smaller number of grid points and benefit from the reduced number of parameters in the model space.

To this regard, Zelt & Smith (1992) have investigated the inversion for both velocities and layer boundaries in their analysis of crustal refraction/wide-angle reflection data sets. Although they do not use parametric curves, their model parameterization offers increased flexibility in shaping the blocks of the velocity model. Relying heavily on the use of secondary phases in the determination of the layer boundary depths, their method appears well suited for their type of data, whereas the method proposed here is primarily designed for first-arrival data which are commonly used in earthquake seismology and cross-hole set ups.

Use of parametric curves has been recently introduced in seismology by Moser, Nolet & Snieder (1992) for their two-point bending ray tracer. In this paper, the feasibility of this approach for 1-D model velocity inversions is explored, and the results are compared with a standard technique which solves solely for velocity adjustments using fixed grid geometries (e.g. Thurber 1983).

METHOD

Parameterization

A laterally homogeneous, vertically varying velocity profile can be represented in parametric form as a 2-D polynomial curve

$$\mathbf{Q}(u) = \sum_{i=1}^{N} \mathbf{P}_i \rho_i(u) = \sum_{i=1}^{N} [z_i \rho_i(u),\ v_i \rho_i(u)]. \tag{1}$$

where $\mathbf{P}_i = (z_i, v_i)$ are the control vertices of the curve, that is, z_i is the vertical position of the grid point and v_i is its associated velocity. $\rho_i(u)$ are the basis functions adopted for the interpolation and u is the parameter which varies between a minimum, u_1, and a maximum, u_N. The sequence of knots u_i $(i = 1, \ldots, N)$, needs to be non-decreasing and, for convenience, equally spaced (i.e. $u_{i+1} = u_i + 1$). Linear B splines are defined as

$$\rho_i(u) = \begin{cases} 0, & u_1 < u < u_{i-1} \\ u - u_{i-1}, & u_{i-1} < u < u_i \\ u_{i+1} - u, & u_i < u < u_{i+1} \\ 0, & u_{i+1} < u < u_N \end{cases} \tag{2}$$

are used as basis functions. To avoid unphysical fluctuations of the velocity model, the sequence of vertex depth values, z_i $(i = 1, \ldots, N)$ is imposed to be non-decreasing.

Partial derivatives

To solve the inverse problem, we must determine the partial derivatives of the velocity profile with respect to the position of control vertices, $\mathbf{P}_i$.

Equation (1) establishes that the velocity profile $\mathbf{Q}(u)$ is a running average of the control vertices $\mathbf{P}_i$, weighted by the basis functions, $\rho_i(u)$. The two components of $\mathbf{Q}(u)$ are

$$Z = Z(u) = \sum_{i=1}^{N} z_i \rho_i(u)$$
$$= 0 + \cdots + (u_{i+1} - u) z_i + (u - u_i) z_{i+1} + \cdots + 0 \tag{3}$$

and

$$V = V(u) = \sum_{i=1}^{N} v_i \rho_i(u)$$
$$= 0 + \cdots + (u_{i+1} - u) v_i + (u - u_i) v_{i+1} + \cdots + 0. \tag{4}$$

We solve for u eq. (3)

$$u = \frac{Z - u_{i+1} z_i + u_i z_{i+1}}{z_{i+1} - z_i} \tag{5}$$

and substitute into (4) to obtain

$$V = \frac{v_{i+1} - v_i}{z_{i+1} - z_i} (Z - u_{i+1} z_i + u_i z_{i+1}) + u_{i+1} v_i - u_i v_{i+1}. \tag{6}$$

Eq. (6) can be differentiated with respect to z_i and v_i so that

$$\frac{\partial V(u)}{\partial z_i} = -\frac{v_{i+1} - v_i}{z_{i+1} - z_i} \rho_i(u) \tag{7}$$

and

$$\frac{\partial V(u)}{\partial v_i} = \rho_i(u). \tag{8}$$

Equations (7) and (8) provide the partial derivatives needed for the determination of the control vertex adjustments in the non-linear iterative inversion.

Inversion

A solution to the non-linear inverse problem can be obtained by local linearization and adopting an iterative inversion scheme. The residual δt (i.e. observed–calculated times) of each phase arrival used in the inversion can be expressed to the first order as follows

$$\delta t = \sum_{i=1}^{N} \frac{\partial t}{\partial z_i} \delta z_i + \sum_{i=1}^{N} \frac{\partial t}{\partial v_i} \delta v_i \tag{9}$$

where t is the traveltime determined with the current model, $\frac{\partial t}{\partial z_i}$ and $\frac{\partial t}{\partial v_i}$ are the partial derivatives of the traveltime with respect to the control vertices, and δz_i, δv_i are the control vertex adjustments to the current velocity model that we seek. The partial derivatives of the traveltime with respect to the control vertices are:

$$\frac{\partial t}{\partial z_i} = -\int_{\text{source}}^{\text{receiver}} \frac{1}{V^2} \frac{\partial V}{\partial z_i} ds = \int_s^r \frac{1}{V^2} \left(\frac{v_{i+1} - v_i}{z_{i+1} - z_i} \right) \rho_i(u)\, ds \tag{10}$$

and

$$\frac{\partial t}{\partial v_i} = -\int_{\text{source}}^{\text{receiver}} \frac{1}{V^2} \frac{\partial V}{\partial v_i} ds = \int_s^r \frac{1}{V^2} \rho_i(u)\, ds \tag{11}$$

where ds is an infinitesimal element along the ray path.

Given M residual times, we can write the problem in matrix form as follows

$$\delta \mathbf{t} = [\mathbf{A} \mid \mathbf{B}] \begin{bmatrix} \delta \mathbf{z} \\ \delta \mathbf{v} \end{bmatrix} \tag{12}$$

where $\delta \mathbf{t}$ is a vector of dimensions M, $\mathbf{A}$ and $\mathbf{B}$ are the matrices of partial derivatives of the traveltimes with respect to the control-vertex positions having individual dimensions $M \times N$, and $\delta \mathbf{z}$, $\delta \mathbf{v}$ are N-dimensional vectors.

Equation (12) represents a classic inverse problem which can be solved using least-squares techniques. However, we note that this inverse problem is of mixed type because we solve simultaneously for two sets of parameters. Multiparameter inversion is often a difficult task because inappropriate relative scaling of the different parameter types may retard and bias convergence in non-linear and iterative schemes (Kennett, Sambridge & Williamson 1988). For this reason, I have adopted the 'subspace method' proposed by Kennett *et al.* (1988). This technique is based on a local minimization on a local subspace spanned by a limited number of vectors in model space. The basis vectors are opportunely chosen along the directions determined by the variations of the misfit functional and by the additional vectors representing the rate of change of the gradient partitions. I refer to the exhaustive description of the methodology given in Kennett *et al.*'s paper for details.

APPLICATION

I have appraised the technique using a synthetic data set generated from the model shown in Fig. 1(b). This model was calculated using the parametric curves formalism of eq. (1) in which the linear basis functions have been replaced by cubic B splines. It displays two abrupt velocity variations located at 2.5 and 3.75 km, with a strong linear gradient interconnecting them (0.8 s^{-1}). A third smoother discontinuity was introduced at 7.5 km which, however, attains a velocity gradient of approximately 1.0 s^{-1} at its flexural point (see Fig. 1b). Similar velocity profiles are found in places where a soft alluvium is deposited above a hard-rock basement—difficult geologic structures to be imaged with existing traveltime-inversion algorithms.

The source–receiver geometry consists of nine receivers and twenty sources (Fig. 1a). The sources were placed along a vertical array at 0.5 km interval and between 0.5 and 10 km depth.

All the ray path calculations were made using the conjugate gradient bending ray tracer introduced by Moser *et al.* (1992). Uniformly distributed random noise ranging between ±0.01 s was added to the traveltimes which were used as input data in the inversion.

Inversions

We show the results obtained using linear B-splines interpolation for both *parametric* (i.e. $[Z(u), V(u)]$) and non-parametric (standard) *velocity-depth function* representations (i.e. $V(Z)$). In order to appraise the technique, the inversion has been repeated using four different grid spacings—3.0, 2.0, 1.0 and 0.5 km. It should be understood, however, that these spacings were free to change when the parametric formalism was used. The results obtained with the entire data set are displayed in Figs 2 to 4 and summarized in Table 1. For conciseness, the results obtained with the 2 km spacing are shown only for the inversions that adopted the depleted data sets.

The initial model used for all inversions consists of a constant gradient of 0.25 s^{-1} (Fig. 1b). Traveltimes calculated with this initial model produced a root-mean square value (RMS) of the residual times misfit of 0.035 s. It was also found that, in order to avoid poor convergence of

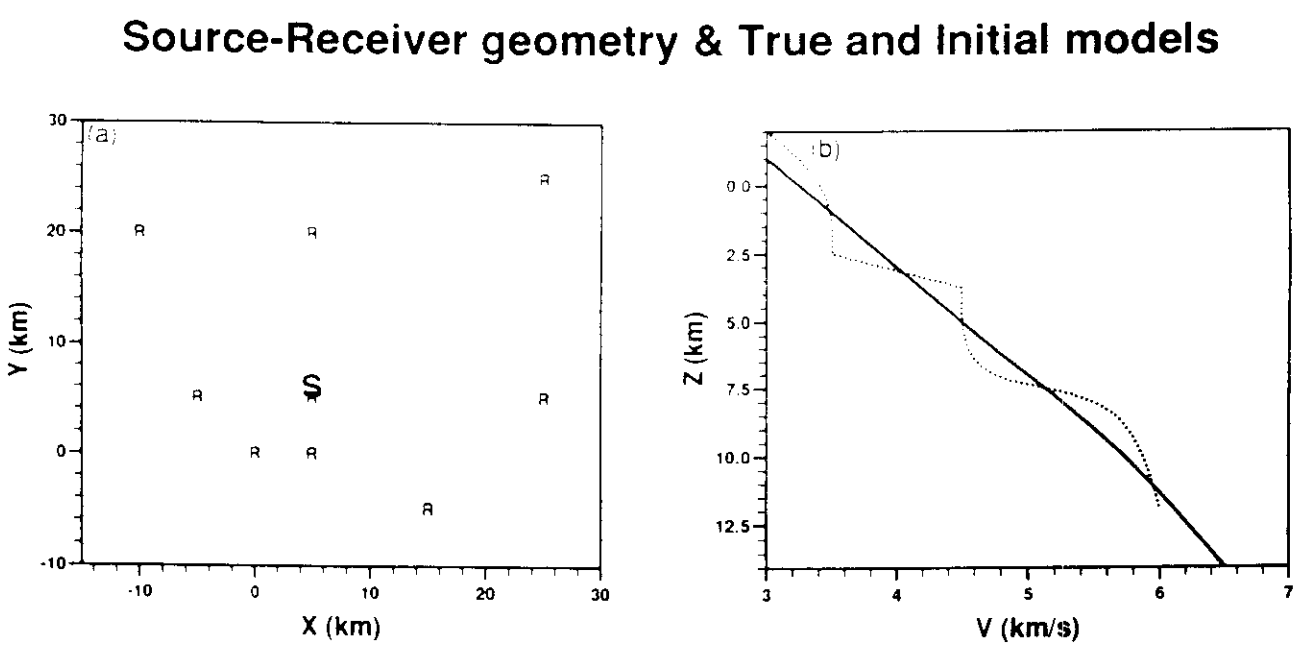


Figure 1. (a) Source–receiver geometry adopted for the tests. **S** is the position of the vertical array of sources which were set equally spaced at 0.5 km intervals between 0.5 and 10 km depth. **R** receiver locations at the surface ($Z = 0$). (b) The dotted line corresponds to the *true* velocity profile that was used to calculate the traveltimes used as input data set for the inversion; the solid line defines the initial model used for all runs of the inversion.

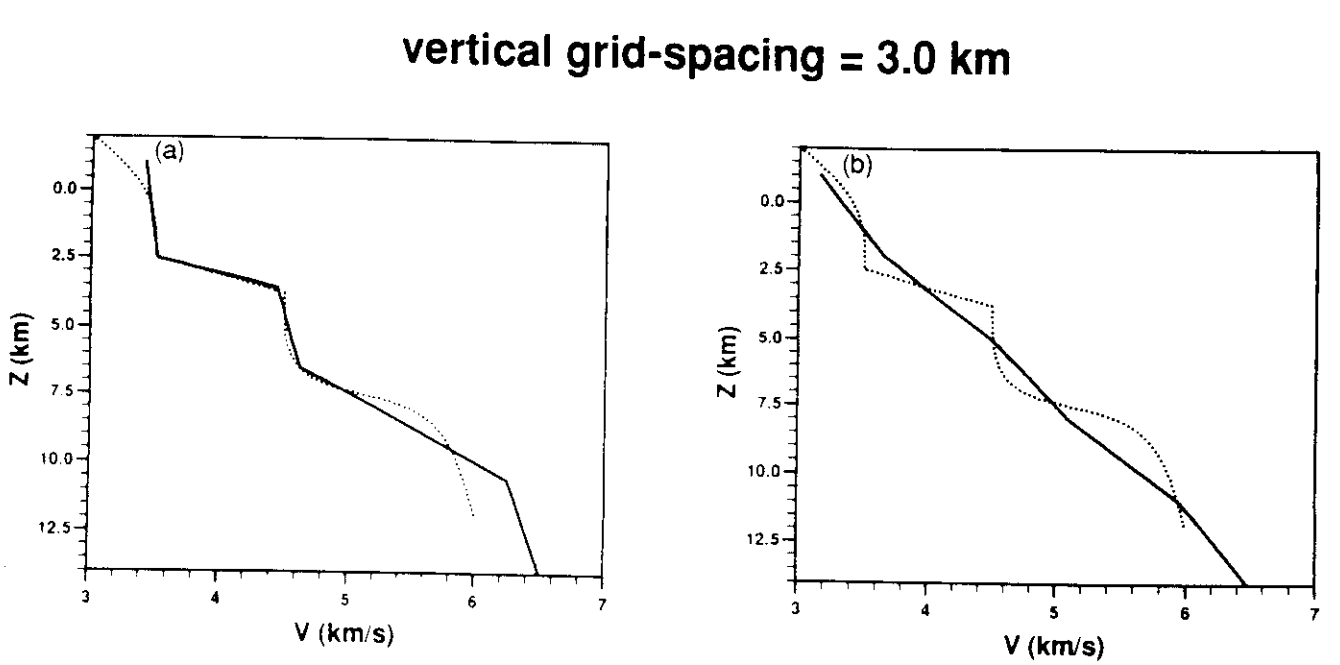


Figure 2. Inversion results when the grid spacing was set to 3 km. (a) Inversion results obtained with the *parametric* curves formalism; (b) inversion results obtained with standard *velocity-depth function* techniques having fixed grid-point positions. The dotted line is the *true* velocity profile of Fig. 1(b).

the iterative inversion, some amount of row weighting was needed. In practice, both residual time and distance weighting was applied. Rows whose traveltime residuals fell in the range $0.1 \leq \delta t \leq 0.15$ s were linearly down-weighted from one to zero and, similarly, for epicentral distances in the range between 15 and 30 km.

Figure 2(b) shows the results of the velocity-depth function inversion when the grid spacing was set to 3 km. This coarse vertical spacing of the grid points inhibits the inversion and the RMS does not decrease beyond 0.029 s. Adding values of damping ranging between one tenth of the largest eigenvalue and the largest eigenvalue to the matrix of partial derivatives produced nearly identical results.

The results obtained with the parametric formalism are shown in Fig. 2(a). This technique produces a model that closely matches the true one down to depths of about 7 km. At greater depths the resolving power of the data set decreases and the inversion procedure is not capable of

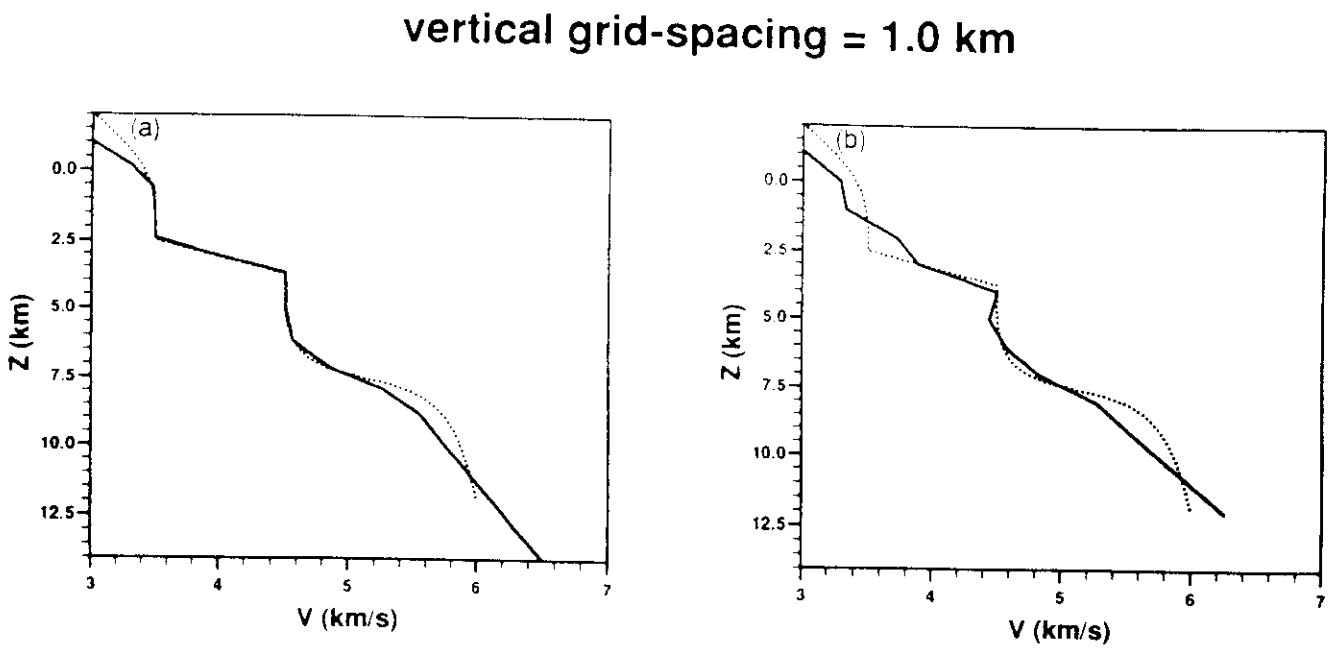


Figure 3. Grid spacing of 1 km. See the caption to Fig. 2 for details.

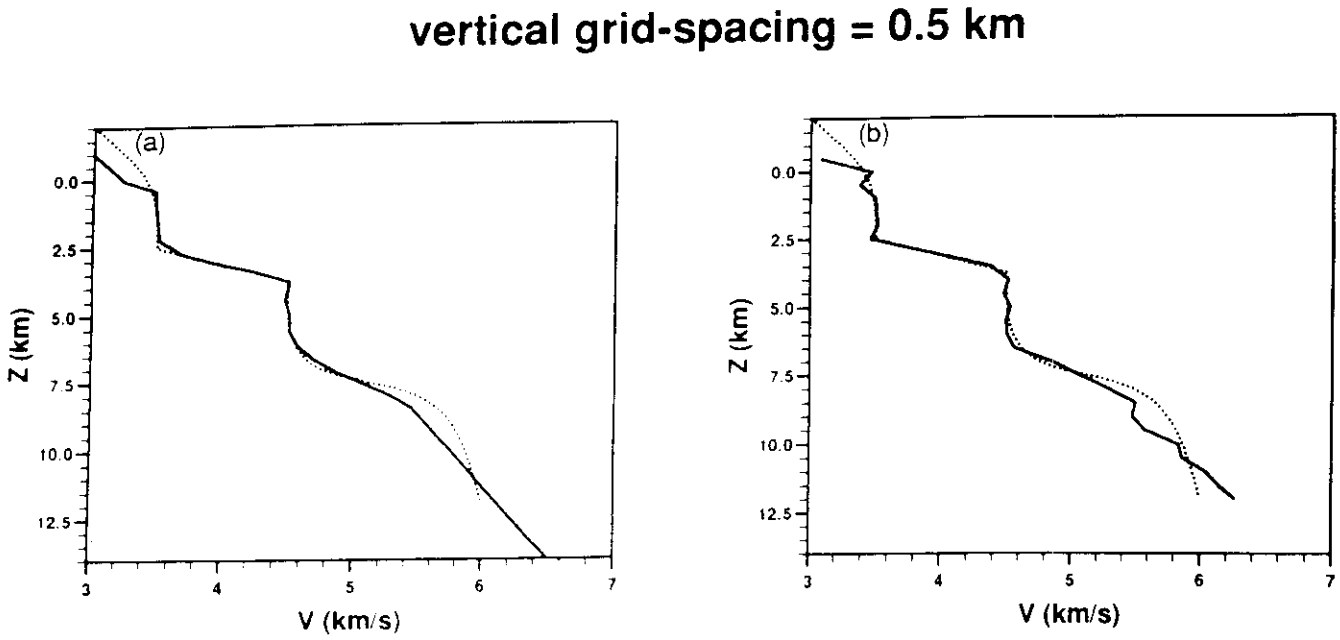


Figure 4. Grid spacing of 0.5 km. See the caption to Fig. 2 for details.

accurately recovering the true model leaving the velocity values there nearly unchanged. It was found that velocity (v) damping values approximately two orders of magnitude smaller than those applied for grid (z) adjustments were needed for proper convergence of the algorithm (i.e. 0.001 versus 0.1 times the largest eigenvalue for velocity and grid adjustments, respectively).

Figure 3 shows the results obtained when the grid spacing was set to 1 km. The fit to the true model obtained with the velocity-depth function method improves but the inversion still has difficulty in retrieving the sharp velocity corner at 2.5 km depth. Conversely, the parametric formalism produces a very good fit to the true model. In Fig. 4, the grid spacing has been decreased to 0.5 km and the results obtained with the two methods are nearly comparable. In fact, the velocity-depth function technique, while producing some spurious oscillations, is now capable of recovering the principal features of the true model with reasonable accuracy. The spurious oscillations could have been mitigated by applying some smoothing operator to the matrix of partial derivatives. However, the purpose of these tests was to compare the performance of the two inversion techniques under nearly identical *a priori* conditioning.

Table 1. Final RMS residual times obtained with the parametric (PAR) and the standard velocity-depth function (V-D) representations. The initial RMS was in all cases 0.035 s.

RMS DATA MISFIT (s)

Grid Spacing (km)	PAR	V-D
3.0	0.011	0.029
2.0	0.008	0.027
1.0	0.008	0.021
0.5	0.008	0.007

The results obtained using the parametric technique in these last two tests provide an explanation for the occurrence of oscillations in the velocity-depth function results. Before doing this, it should be clarified what is meant for resolving power of the data set. Given a source–receiver geometry, the resolving power is closely related to the degree of model determinacy. Williamson (1990) points out that, while increasing the number of model parameters, the model determinacy decreases but the inversion progressively exploits more of the information contained in the data set. Depending on the rank of the matrix of partial derivatives at the increase of the model parameters, the inversion is commonly defined as fully over-determined, mixed-(under- over-) determined and fully under-determined (see also Menke 1984). Owing to the non-linearity of seismic inversions, noise in the data and inaccuracy of the ray tracers, the mixed- and under-determined inversions have to be closely monitored and properly damped by various means according to preconceptions about the final model. In practice and with standard methods, an increase in size of the parameter space results in a decrease of the scale length of the parameterization (e.g. smaller constant velocity pixels) and the associated ray coverage per parameter (e.g. pixel) decreases. It follows that the resolving power of the data set onto the parameters is similarly decreased.

In the test examples above, it appears that the main cause for the instabilities near the velocity discontinuities is the mispositioning of grid nodes with respect to those needed to recover the true structure. In these circumstances, the input data have information about the true location of the discontinuity but the positioning of the grid mesh inhibits its recovery. At this stage, standard inversion methods may, for

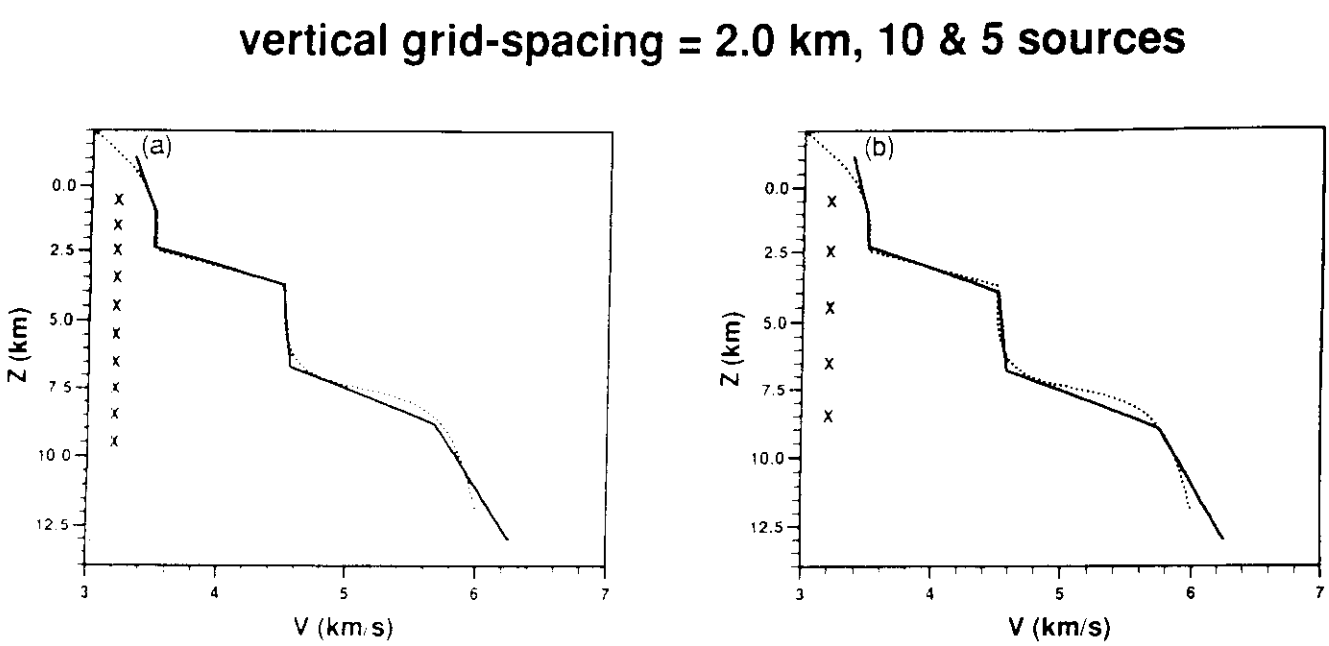


Figure 5. Inversion results obtained with parametric curves using depleted data sets and 2.0 km initial vertical grid spacing. (a) 10 equally spaced sources (spacing interval of 1 km); (b) 5 equally spaced sources (spacing interval of 2 km). In both cases, the uppermost source was set at 0.5 km depth. The vertical position of the sources is marked by crosses.

example, overshoot the velocity of one grid point and, at the subsequent iteration, complete the minimization by perturbing the velocities of the neighbouring grid points while attempting to compensate also for the previous overshooting. This effect is magnified when the number of grid points is large because the data set has diminished resolving power and the resulting models are more prone to instabilities (e.g. Fig. 4b). These oscillations can be mitigated, however, by adding a smoothing operator. Conversely, this problem does not occur when the vertical gridding is coarser, but in these cases the inversion is unable to proceed (e.g. Fig. 2b). The small velocity fluctuations observable in the parametric formalism results of Fig. 4(a) are also caused mainly by the diminished resolving power of the data set.

Another way to decrease the resolving power of the data set is by reducing the size of the data space. In Fig. 5, I show the results of the parametric inversion when only 10 and 5 sources of the original data set (i.e. 1.0 and 2.0 km spacing along the vertical array of sources, respectively) were used and the initial grid spacing was set to 2 km. The technique resolves well the true model in both data depleted inversions. When the same data sets were used as inputs for the standard velocity-depth inversion with 0.5 km spacing (i.e. the only one capable of reproducing the main features of the true model), the reduced resolving power of the data sets combined with the large number of parameters to be determined, caused the inversion algorithm to stop at its first iteration. Similar results were obtained with values of damping ranging between 0.25 and 10 times the largest eigenvalue of the singular value decomposition of the matrix of partial derivatives.

DISCUSSION

The main advantage of the parametric method presented in this study lies in the greatly reduced number of grid points needed to model accurately an unknown velocity structure. In fact, parametric curves allow the velocity model to 'shape itself' and find, where the resolving power of the data set is sufficient, the optimal position of its control vertices. The reduced number of grid points results in a significant saving in computation time. While doubling the number of parameters (in the 1-D cases), the parametric method appears to make optimal use of the increased size of the model space by adjusting the position of the control vertices to enable accurate retrieval of any velocity profile, avoiding the instabilities caused by fixed discretization grids. It should be noted that use of the 'subspace method' avoids inversion of large-scale matrices altogether, the entire problem being projected into a small-dimensional subspace (Sambridge 1990). Therefore, doubling the size of the model space is not really a burden.

Damping is a factor that can severely affect any non-linear, iterative inversion. In the parametric inversion, it was found that when the initial gridding is coarse (i.e. 3 and 2 km in our test) it is preferable to apply larger damping to the node (z) adjustments than to the velocity (v) adjustments. When the initial grid spacing was reduced, nearly identical damping of node and velocity adjustments was needed for proper convergence.

Selection of the initial grid spacing appears less troublesome than with standard methods because the true velocity discontinuities are recovered even with rather coarse spacing (e.g. Fig. 2). Successive inversions with finer

grid spacing should confirm the initial results while adding more detail. However, occurrence of some small fluctuations near the discontinuities remains an inevitable burden when the grid spacing is chosen too fine (e.g. Fig. 4a) although very dense inversion meshes are not required with the parametric technique. In order to maintain model determinacy, it appears more advantageous to select intermediate values of initial grid meshing (e.g. see Fig. 5).

The inversion results with the depleted data sets indicate that the method is robust (see Fig. 5) which can lead again to significant saving in computation time.

In the application of inversion algorithms to observed data, the only quantity that can be monitored to provide an estimate of the goodness of the fit to the true model is the RMS of the residual times. Models that generate RMS residuals within the estimated errors of the data set are generally chosen as reasonable answers to the inverse problem. Table 1 shows that all the inversions using the parametric method result in models having RMS values that fall within or close to the noise value of 0.01 s added to the input data. In contrast, only the 0.5 km spacing inversion with fixed nodes was capable of converging to RMS values within the noise level.

The method presented here adopts linear B-splines basis functions but, in principle, higher-order basis could be used. For example, parametric curves which adopt cubic B splines are widely used in computer graphics (e.g. Bartels, Beatty & Barsky 1987) because they provide increased flexibility in shaping objects. Preliminary inversions using this class of curves and an approximate solution for the partial derivatives of the traveltime with respect to the control vertex positions, yield encouraging results.

Use of parametric curves also appears quite suitable for 2-D and 3-D geometries. In fact, existing algorithms are generally limited by the use of simple and repetitive grid meshes while the extension of the present methodology to the 2-D and 3-D cases is capable of meeting the degrees of freedom needed for modelling existing geological structures (Michelini, in preparation).

Finally, the results of this study show that existing methods may suffer heavily from vertical undersampling owing to the often coarse *a priori* imposed discretization. This burden can be obviated with the method presented above.

ACKNOWLEDGMENTS

The author is indebted to F. Bocchio for discussions on the method and to W. Foxall, I. Marson and an anonymous reviewer for carefully reading the manuscript and providing suggestions.

REFERENCES

Bartels, R. H., Beatty, J. C. & Barsky, B. A., 1987. *An introduction to Splines for use in Computer Graphics and Geometric Modeling*, Morgan Kaufmann Publishers, Inc, Los Altos, California.

Kennett, B. L. N., Sambridge, M. S. & Williamson, P. R., 1988. Subspace methods for large inverse problems with multiple parameter classes, *Geophys. J. Int.*, **94**, 237–247.

Menke, W., 1984. *Geophysical data analysis: discrete inverse theory*, Academic Press, London.

Michelena, R. J. & Harris, J. M., 1991. Tomographic traveltime inversion using natural pixels, *Geophysics*, **56**, 635–644.

Moser, T. J., Nolet, G. & Snieder, R., 1992. Ray bending revisited, *Bull. seism. Soc. Am.*, **82**, 259–288.

Sambridge, M. S., 1990. Non-linear arrival time inversion: constraining velocity anomalies by seeking smooth models in 3-D, *Geophys. J. Int.*, **102**, 653–677.

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

Williamson, P. R., 1990. Tomographic inversion in reflection seismology, *Geophys. J. Int.*, **100**, 255–274.

Zelt, C. A. & Smith, R. B., 1992. Seismic traveltime inversion for 2-D crustal velocity structure, *Geophys. J. Int.*, **108**, 16–34.

An adaptive-grid formalism for traveltime tomography

Alberto Michelini

Istituto di Geodesia e Geofisica
Università di Trieste,
Via dell'Universitá, 7, 34127 Trieste, Italy

Revised for the *Geophys. J. Int.*
24 October, 1994

Key words: seismic inversion, parametric representation, model parameterization, ray-tracing
Short title: adaptive-grid tomography

SUMMARY

The use of an adaptive-grid formalism for seismic, non-linear, traveltime tomography is proposed. The method is based on a parametric representation of the velocity model and involves the simultaneous inversion for both velocity and position of the grid points of the model discretization mesh. Therefore, the method seeks the optimal grid configuration to define the model. Cubic B-splines basis functions have been used for model representation as they are particularly versatile in the reproduction of geologically complex structures. The traveltimes are calculated using a new initial value ray-tracer that calculates the ray trajectories directly in the parametric domain.

The method is tested against synthetic data generated for various cross-hole geometries. It is found that, in parts of the model having good ray-coverage, the method accurately retrieves velocity anomalies of arbitrary shape using a generally small number of grid points of the inversion discretization mesh. With the exception of initial meshes which are too coarse to describe accurately the complexity of the true structure, the method retrieves nearly identical final models regardless of the predefined node spacing and node configuration. Therefore, the method can avoid very fine discretization and, matrix sparseness, one of the main sources of indeterminacy, is similarly avoided. When compared to standard methods entailing a similar total number of inversion parameters, our results show the pitfalls that may derive from *a priori* assigning a fixed-grid mesh. Overall, the parametric representation leads to some saving in the total number of inversion parameters when the structure consists of sparse distributions of irregular features. Other styles of model may be better recovered by regular grids for a given number of inversion parameters.

Introduction

The goal of seismic traveltime tomography is retrieval of the *true* Earth's velocity structure. In general, the images of the Earth's interior that are resolved by traveltime tomography depend both upon the quality of the data set and upon the inversion technique employed. The inversion technique itself can be subdivided into two main parts: first, the *a priori* selection of the parameterization and discretization used to represent the velocity model; and, second, the data functional misfit minimization criterion that is applied.

Parameterization and discretization define the "filter" through which the model is viewed and, because of our scarce knowledge of the target velocity field, it is critical to minimize the effect of this filter in order to obtain accurate reconstructions of the true model (e.g., Michelena and Harris, 1991; Spakman, 1993).

In principle, this can be achieved by decreasing the scale length of the model discretization and by selecting powerful interpolation schemes (e.g., Sambridge, 1990). Although this procedure yields more accurate fits of the retrieved velocity models to the true ones and fuller exploitation of the information contained in the data sets, it decreases the model determinacy; i.e., it increases the size of the parameter space and may make the inverse problem strongly under-determined. Therefore, the solutions need to be properly regularized to avoid instabilities and artifacts (Williamson, 1990). Alternatively (see Thurber, 1993, for a comprehensive review), the velocity structure can be treated explicitly as a continuous function of the spatial coordinates (Chou and Booker,1979; Tarantola and Nercessian, 1984), but this approach, while allowing faithful reconstructions of the inhomogeneous structure, does not differ significantly from those based on fine discretization, requiring the same kind of regularization.

Various factors contribute to the rise of instabilities and artifacts through non-linear and iterative tomographic inversions. In addition to the noise contained in the data sets, poor positioning of the discretization grid with respect to the positions of velocity anomalies and choice of inappropriate basis functions for interpolation may introduce artifacts. For example, velocity model oscillations may arise near a velocity discontinuity due to inappropriate distribution of the grid points (nodes) there—the inversion algorithm attempts minimization by creating artificial oscillations, or discontinuities may be not resolved if the discretization interval is too coarse (Michelini,

1993), leading to biased solutions (Snieder *et al.*, 1991; Godmundsson and Clayton, 1991; Spakman, 1993). These problems could be avoided if some *a priori* information on the position and order of the velocity anomalies was available. This information is, in fact, contained in the data set, and the aim of this paper is to show how it can be used to optimize node positions through an adaptive grid technique. For this purpose we (1) use a parametric formalism to represent the velocity model, (2) present a ray tracer for the determination of traveltimes through such a model and (3) introduce an adaptive gridding procedure that makes optimization of the node positions an active part of the traveltime misfit minimization.

Curvilinear coordinates (closely related to the parametric representation; Courant, 1936) are commonly employed in computational fluid dynamics because they allow arbitrary irregular fields to be modelled using an economical distribution of grid points. Adaptive gridding schemes have been developed that use the physics of a problem to direct the grid points to their optimal positions (Thompson *et al.*, 1985). Using curvilinear coordinates, the field is mapped from the "physical" to a "computational" (parametric) domain and the grid is adapted according to the behaviour of the field itself while satisfying the boundary conditions. The problem of representing bodies of arbitrary shape also occurs in computer graphics, where it is standard practice to use parametric surfaces to render objects while finding the most economical distribution of support points (e.g., Bartels *et al.*, 1987).

In seismology, Fornberg (1988) introduced the use of curvilinear coordinates for pseudospectral calculation of the elastic wavefield. This allows accurate representation of velocity interfaces that avoids the non-physical diffractions resulting from the discretization (see also Nielsen *et al.*, 1992). More recently, Carcione and Wang (1993) have proposed a Chebyshev collocation method in generalized coordinates, and Moser *et al.*, (1992) have developed a bending ray-tracer which makes use of parametrically-defined curves and cubic β-splines basis functions for interpolation. Their approach appears very flexible for modelling arbitrarily curved ray-paths using a small number of control vertices along the ray-path. Pereyra (1992) has developed a 3-D ray tracer that uses a parametric representation to define velocity interfaces. He also suggests that the smoothly varying material properties within blocks can be represented parametrically in a way similar to that described in this paper.

An adaptive-grid approach is attractive for tomography where it is desirable to

use the minimum number of nodes and determine their optimal positions, through inversion. An approach of this kind has been recently pursued by Michelini (1993, hereafter cited as Paper I), who has shown that the 1-D velocity inversion problem can be formulated using parametric curves; traveltime minimization being achieved by adjusting both the velocities and vertical positions of the nodes. Paper I showed that this technique makes it possible to use relatively coarsely-spaced discretization grids to produce accurate velocity model reconstructions. In Paper I, the parametric analysis was considerably simplified by limiting it to laterally homogeneous, vertically-varying models (1-D) and using linear B-splines basis functions to interpolate between nodes. In this paper, I develop a general parametric (or, equivalently, adaptive-grid) formulation which can be used when the reconstruction problem is posed in one, two or three dimensions and employs arbitrary local basis functions.

Method

The development of the parametric formalism is discussed for two-dimensional (2-D) tomographic problems, but the same formalism can be followed for 1-D and 3-D cases.

Parameterization

A 2-D velocity model $v(x,z)$ can be expressed parametrically in the form

$$x = x(X_{ij},\xi,\eta), \quad z = z(Z_{ij},\xi,\eta), \quad v = v(V_{ij},\xi,\eta), \tag{1}$$

where

$$\begin{aligned} x(X_{ij},\xi,\eta) &= \sum_{i=0}^{M}\sum_{j=0}^{N} X_{ij}\rho_i(\xi)\rho_j(\eta) \\ z(Z_{ij},\xi,\eta) &= \sum_{i=0}^{M}\sum_{j=0}^{N} Z_{ij}\rho_i(\xi)\rho_j(\eta) \\ v(V_{ij},\xi,\eta) &= \sum_{i=0}^{M}\sum_{j=0}^{N} V_{ij}\rho_i(\xi)\rho_j(\eta) \end{aligned} \tag{2}$$

are functions of the parameters ξ and η. $\rho_i(\xi)$ and $\rho_j(\eta)$ $(i = 0,1,\cdots,M;\ j = 0,1,\cdots,N)$ are arbitrary local basis functions used to determine a surface from $(M+1)\times(N+1)$ control vertices $\mathbf{P}_{ij} \equiv (X_{ij}, Z_{ij}, V_{ij})$.

Thus, X_{ij} and Z_{ij} are the horizontal and vertical position, respectively, of the ij^{th} control vertex (i.e., grid point) of the model discretization mesh and V_{ij} is the associated velocity value. Therefore, in setting up the model, one would typically

define X_{ij}, Z_{ij} and V_{ij} on a regular grid in $\xi\eta$-space. We have used cubic B-splines basis functions as defined in Appendix A.

In general, Equation (1) establishes that to a point (ξ, η) that ranges over a given region R in the $\xi\eta$-plane (see Appendix A) corresponds a point (x, z, v) that ranges over a configuration in the xzv-space (Courant, 1936). Therefore, the parametric representation of a surface may be regarded as the "mapping" of the region R of the $\xi\eta$-plane on the corresponding surface where the word mapping is understood to mean a point-to-point correspondence. In addition and for our purposes, we assume that the mapping from the parametric $\xi\eta$-plane to the physical coordinates xz-plane (i.e., by neglecting v in Equation (1)) is always *one-to-one*. This prevents the velocity surface to be multivalued at points in the xz-plane (i.e., unphysical topological folding of the velocity surface). To prevent this situation when setting up the velocity model or when inverting for the positions of the control vertices, it is necessary to check that curvilinear lines (physical space) of the same family do not cross, and lines of different families do not cross more than once. Each family of curvilinear lines is generated by assigning constant values to one parametric coordinate and varying the other (e.g., Thompson *et al.*, 1985). In the following, we will often refer to the parametric $\xi\eta$-plane and the Cartesian xz-plane as the "computational" and "physical" domains, respectively, in agreement with the terminology in use in computational fluid dynamics (Thompson *et al.*, 1985).

In Figure 1, we provide an example that illustrates the *one-to-one* correspondence described above for a grid consisting of 5×5 control vertices, P_{ij} $(i, j = 0, 1, \cdots, 4)$. In this figure, we have also plotted the two families of curvilinear lines generated by assigning constant values to one parametric coordinate and allowing the other to vary (i.e., $\xi = 0, 1, \cdots, 4$ while varying η, and viceversa). Curvilinear lines may be regarded as a curvilinear $\xi\eta$-system of coordinates in the xz-region. Note also that in the parametric representation with cubic B-splines basis and with the exception of regularly spaced grid geometries (see examples below) the curvilinear lines approximate the corresponding grid points but do not pass necessarily through them.

A second example of the parametric representation is shown in Figure 2. In Figure 2a and 2c, we show the perspective and plan views of the model whereas in Figure 2b, we show the two families of curvilinear coordinate lines obtained using the grid points (solid circles) listed in Table 1. In addition to the greatly reduced

number of control vertices needed to represent accurately the velocity surface, it should be noted that the curved sharp discontinuities near $z = -0.6$ and $z = -0.3$ are obtained by simply making coincide three sets of control vertices (see Bartels *et al.*, 1987 and Table 1). This feature follows from the adoption of cubic B-splines basis functions and it has the potential advantage when compared to standard methods for model parameterization, to reproduce any discontinuity while using the same set of basis functions so that, at least in principle, one can reproduce any model shape and be independent from the initial choice of parameterization. The same feature has been used by Moser *et al.* (1992) to make a ray-path to go through sharp velocity discontinuities or to reflect away at a velocity interface.

Partial derivatives

A prerequisite for the solution of inverse problems which are based on local linearization of the data functional misfit is the calculation of the partial derivatives of the velocity surface with respect to the control vertices, $\mathbf{P}_{ij}$; i.e., with respect to both the position of the grid point (X_{ij}, Z_{ij}) and the associated velocity value V_{ij}. In Paper I, for the 1-D case employing linear basis functions, this was accomplished by eliminating the single parameter ξ (u in Paper I), and expressing the velocity profile at any depth as a function of the control vertices (Z_i, V_i) (see Equation 6, Paper I). However, the same procedure cannot be followed when basis functions of degree higher than one (linear) are used for the 1-D inversion, or when basis functions even of degree one are used in 2-D and 3-D cases. This procedure would generate equations in ξ and η of order greater than one, which would greatly complicate the analytical solution.

An alternative approach in which we do not need to solve explicitly for ξ and η in Equation (1) takes advantage of the differentiation rules for inverse functions (see Courant vol. 2, 1936, p. 142). We can write

$$\xi = \xi(X_{ij}, Z_{ij}, x, z) \tag{3}$$

and

$$\eta = \eta(X_{ij}, Z_{ij}, x, z) \tag{4}$$

so that, from Equation (1), v at any location (x, z) can be expressed as

$$v = v\left(V_{ij},\ \xi(X_{ij}, Z_{ij}, x, z),\ \eta(X_{ij}, Z_{ij}, x, z)\right). \tag{5}$$

In Equations (3) to (5) it is implied that for some x and z, ξ and η are functions of all the contributing ij terms (see Appendix A). The number of the contributing terms will vary according to the type of local basis functions employed (e.g., 4 and 16 for linear and cubic B-splines, respectively).

In addition to the derivatives of v with respect to V_{ij} which can be easily determined from Equation (2):

$$\frac{\partial v}{\partial V_{ij}} = \rho_i(\xi)\rho_j(\eta) \qquad \left(= \frac{\partial x}{\partial X_{ij}} = \frac{\partial z}{\partial Z_{ij}} \right) , \tag{6}$$

we seek the partial derivatives of v with respect to the vertex positions X_{ij} and Z_{ij}, which, using Equation (5) and the chain rule for differentiation, can be written as

$$\frac{\partial v}{\partial X_{ij}} = \frac{\partial v}{\partial \xi}\frac{\partial \xi}{\partial X_{ij}} + \frac{\partial v}{\partial \eta}\frac{\partial \eta}{\partial X_{ij}} \tag{7}$$

and

$$\frac{\partial v}{\partial Z_{ij}} = \frac{\partial v}{\partial \xi}\frac{\partial \xi}{\partial Z_{ij}} + \frac{\partial v}{\partial \eta}\frac{\partial \eta}{\partial Z_{ij}}. \tag{8}$$

Differentiation with respect to the parameters ξ and η can be easily done directly from Equation (2):

$$\begin{aligned} \frac{\partial v}{\partial \xi} &= \sum_{i=0}^{M}\sum_{j=0}^{N} V_{ij}\rho_i{}'(\xi)\rho_j(\eta) \\ \frac{\partial v}{\partial \eta} &= \sum_{i=0}^{M}\sum_{j=0}^{N} V_{ij}\rho_i(\xi)\rho_j{}'(\eta) \end{aligned} \tag{9}$$

where $\rho_i{}'(\xi)$ and $\rho_j{}'(\eta)$ indicate $\frac{d\rho_i(\xi)}{d\xi}$ and $\frac{d\rho_j(\eta)}{d\eta}$, respectively. Similar relations hold for $\frac{\partial x}{\partial \xi}$, $\frac{\partial x}{\partial \eta}$, $\frac{\partial z}{\partial \xi}$ and $\frac{\partial z}{\partial \eta}$.

The derivatives of the parameters, ξ and η, with respect to X_{ij} and Z_{ij} can be obtained by substitution of Equations (3) and (4) into (2). The following identities are obtained:

$$x = x\,(X_{ij},\ \xi(X_{ij}, Z_{ij}, x, z),\ \eta(X_{ij}, Z_{ij}, x, z)) \tag{10}$$

and

$$z = z\,(Z_{ij},\ \xi(X_{ij}, Z_{ij}, x, z),\ \eta(X_{ij}, Z_{ij}, x, z)) \tag{11}$$

where the subscripts ij again refer to all the contributing terms. Note also that the left hand side of the identities (10) and (11) cannot depend upon X_{ij} and Z_{ij}, respectively (i.e., $\frac{\partial x}{\partial X_{ij}} = 0$ and $\frac{\partial z}{\partial Z_{ij}} = 0$).

If the identities above are differentiated with respect to X_{ij} and Z_{ij}, we can form the following systems of equations:

$$\begin{cases} \dfrac{\partial x}{\partial X_{ij}} + \dfrac{\partial x}{\partial \xi}\dfrac{\partial \xi}{\partial X_{ij}} + \dfrac{\partial x}{\partial \eta}\dfrac{\partial \eta}{\partial X_{ij}} = 0 \\ \dfrac{\partial z}{\partial \xi}\dfrac{\partial \xi}{\partial X_{ij}} + \dfrac{\partial z}{\partial \eta}\dfrac{\partial \eta}{\partial X_{ij}} = 0 \end{cases} \tag{12}$$

and

$$\begin{cases} \dfrac{\partial x}{\partial \xi}\dfrac{\partial \xi}{\partial Z_{ij}} + \dfrac{\partial x}{\partial \eta}\dfrac{\partial \eta}{\partial Z_{ij}} = 0 \\ \dfrac{\partial z}{\partial Z_{ij}} + \dfrac{\partial z}{\partial \xi}\dfrac{\partial \xi}{\partial Z_{ij}} + \dfrac{\partial z}{\partial \eta}\dfrac{\partial \eta}{\partial Z_{ij}} = 0 \end{cases} \tag{13}$$

which can be solved for $\dfrac{\partial \xi}{\partial X_{ij}}$, $\dfrac{\partial \eta}{\partial X_{ij}}$ and $\dfrac{\partial \xi}{\partial Z_{ij}}$, $\dfrac{\partial \eta}{\partial Z_{ij}}$, respectively. When these derivatives are substituted into (7) and (8), we obtain the needed derivatives,

$$\frac{\partial v}{\partial X_{ij}} = -\frac{\partial x}{\partial X_{ij}}\left(\frac{\partial z}{\partial \eta}\frac{\partial v}{\partial \xi} - \frac{\partial z}{\partial \xi}\frac{\partial v}{\partial \eta}\right) J^{-1} \tag{14}$$

and

$$\frac{\partial v}{\partial Z_{ij}} = -\frac{\partial z}{\partial Z_{ij}}\left(\frac{\partial x}{\partial \xi}\frac{\partial v}{\partial \eta} - \frac{\partial x}{\partial \eta}\frac{\partial v}{\partial \xi}\right) J^{-1} \tag{15}$$

where J is the Jacobian of the coordinate transformation defined as

$$J = \frac{\partial x}{\partial \xi}\frac{\partial z}{\partial \eta} - \frac{\partial x}{\partial \eta}\frac{\partial z}{\partial \xi} \quad . \tag{16}$$

Inversion

In travel time tomography, it is customary to solve the non-linear inverse problem by local linearization and adoption of an iterative inversion scheme. Within the parametric formalism outlined above, the residual δt (i.e., observed - calculated traveltime) of each phase arrival used in the inversion can be expressed to first order as:

$$\delta t \simeq \sum_{i=0}^{M}\sum_{j=0}^{N}\frac{\partial t}{\partial X_{ij}}\delta X_{ij} + \sum_{i=0}^{M}\sum_{j=0}^{N}\frac{\partial t}{\partial Z_{ij}}\delta Z_{ij} + \sum_{i=0}^{M}\sum_{j=0}^{N}\frac{\partial t}{\partial V_{ij}}\delta V_{ij} \tag{17}$$

where t is the traveltime determined with the current model and δX_{ij}, δZ_{ij}, δV_{ij} are the control vertex adjustments to the current velocity model that we seek. The partial derivatives of the travel time with respect to the control vertices are:

$$\frac{\partial t}{\partial X_{ij}} = -\int_{source}^{receiver}\frac{1}{v^2}\frac{\partial v}{\partial X_{ij}}ds \ , \tag{18}$$

$$\frac{\partial t}{\partial Z_{ij}} = -\int_{source}^{receiver}\frac{1}{v^2}\frac{\partial v}{\partial Z_{ij}}ds \tag{19}$$

and

$$\frac{\partial t}{\partial V_{ij}} = -\int_{source}^{receiver}\frac{1}{v^2}\frac{\partial v}{\partial V_{ij}}ds \tag{20}$$

where ds is an infinitesimal distance along the ray-path.

If we rearrange the ij terms as column vectors and suppose that we are given L residual times, we can write Equation (17) in matrix form as:

$$\delta \mathbf{t} \simeq [\mathbf{A} \mid \mathbf{B} \mid \mathbf{C}] \begin{bmatrix} \delta \mathbf{x} \\ \delta \mathbf{z} \\ \delta \mathbf{v} \end{bmatrix} \tag{21}$$

where $\delta \mathbf{t}$ is a vector of dimensions L; $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are the matrices of partial derivatives of the traveltimes with respect to the control-vertex positions having individual dimensions $L \times ((M+1) \times (N+1))$, and $\delta \mathbf{x}$, $\delta \mathbf{z}$, $\delta \mathbf{v}$ are $((M+1) \times (N+1))$-dimensional vectors.

Equation (21) represents a classic inverse problem of mixed type because we solve simultaneously for three sets of parameters (the two spatial adjustments $\delta \mathbf{x}$ and $\delta \mathbf{z}$ and the velocity adjustments $\delta \mathbf{v}$). Multiparameter inversion is often a difficult task because inappropriate relative scaling of the different parameter types may retard and bias convergence in non-linear and iterative schemes (Kennett *et al.*, 1988). As in Paper I, we follow here the "subspace method" proposed by Kennett *et al.* (1988).

According to this method, the full matrix problem in (21) is projected onto a specified set of orthonormal vectors (the subspace) which, in our case, includes the three ascent directions (i.e., each vector corresponds to a parameter type) of the misfit functional and their nine associated rates of change along and among themselves (Kennett *et al.*, 1988). Therefore, in our problem the subspace consists of a total of twelve orthonormal vectors.

Damping can severely affect the results of a non-linear, iterative inversion. In the adaptive-grid formalism a solution is sought simultaneously for three groups of parameters (2 spatial, 1 velocity), and the corresponding damping values must be set to achieve proper regularization of the solution. In our inversion procedure, the three damping values are determined as a fraction of the largest eigenvalue of the singular value decomposition of the appropriate partition of the matrix of traveltime partial derivatives. Specifically, at the n^{th} iteration of the tomographic reconstruction and for one group of parameters, the fraction of the largest eigenvalue to be applied to the derivatives matrix is

$$\epsilon_n = \epsilon_0 b^{-n} \tag{22}$$

where ϵ_0 is the initial fractional value of damping and b is a convergence factor. When b is greater than one, the damping is a progressively smaller fraction of the largest

eigenvalue. In the tomographic tests below, we have applied initial fractional velocity and position damping values of the order 0.01 and 0.1, respectively. Thus damping velocity adjustments fractionally less than position adjustments. Similarly, the convergence factor b was set to 1.5 and 1.2 for velocity and position adjustments, respectively. This (relative) underdamping of the velocity with respect to the position adjustments (in terms of largest eigenvalues) reflects the need to retrieve the initial size of the velocity anomaly first, the boundaries of which are later refined by grid position adjustments. This procedure prevents the rise of large, uncontrolled, grid position adjustments in the early stages of the inversion which would be deleterious for successful final reconstructions. It also prevents the formation of the unphysical topological foldings of the velocity surface mentioned above.

Ray-tracing

Efficient and accurate ray tracing is critical for any traveltime tomography algorithm. In our scheme, Equation (2) establishes the mapping from the parametric to the physical space. However, there is no simple analytical solution for the inverse mapping for which one has to rely on non-linear, computationally intensive search methods. Therefore, because any ray-tracing method that relies heavily on this inverse transformation would be prohibitive computationally, we have developed a method that carries out the ray-tracing in the computational domain and requires minimal physical-parametric mapping.

The ray-tracer solves the ray equation numerically (e.g., Lee and Stewart, 1981)

$$\frac{d}{ds}\left[u(\frac{d\mathbf{r}}{ds})\right] = \nabla u \tag{23}$$

where $\mathbf{r}$ is a position vector, ds an infinitesimal increment along the ray-path, $u = \frac{1}{v}$ is the slowness and ∇ is the gradient operator. A numerical solution to this equation can be found by casting it as an initial value problem and by rewriting Equation (23) in terms of simultaneous first order differential equations (e.g., Cerveny, 1987).

We define

$$\omega_1 = \xi, \quad \omega_2 = \frac{dx}{ds}, \quad \omega_3 = \eta, \quad \omega_4 = \frac{dz}{ds}$$

and arbitrarily set $s = 0$ at the source. The ray trajectory in the computational

$\xi\eta$-plain is a solution of the following ray-tracing system:

$$\begin{aligned}
\frac{d\omega_1}{ds} &= \left(\frac{\partial z}{\partial \eta}\omega_2 - \frac{\partial x}{\partial \eta}\omega_4\right) J^{-1} \\
\frac{d\omega_2}{ds} &= v\left[\frac{\partial u}{\partial x} - (\frac{\partial u}{\partial x}\omega_2 + \frac{\partial u}{\partial z}\omega_4)\omega_2\right] \\
\frac{d\omega_3}{ds} &= \left(-\frac{\partial z}{\partial \xi}\omega_2 + \frac{\partial x}{\partial \xi}\omega_4\right) J^{-1} \\
\frac{d\omega_4}{ds} &= v\left[\frac{\partial u}{\partial z} - (\frac{\partial u}{\partial x}\omega_2 + \frac{\partial u}{\partial z}\omega_4)\omega_4\right]
\end{aligned} \tag{24}$$

where

$$\begin{aligned}
\frac{\partial u}{\partial x} &= -\frac{1}{v^2}\left(\frac{\partial v}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial v}{\partial \eta}\frac{\partial \eta}{\partial x}\right) \\
\frac{\partial u}{\partial z} &= -\frac{1}{v^2}\left(\frac{\partial v}{\partial \xi}\frac{\partial \xi}{\partial z} + \frac{\partial v}{\partial \eta}\frac{\partial \eta}{\partial z}\right)
\end{aligned} \tag{25}$$

and

$$\begin{aligned}
\frac{\partial \xi}{\partial x} &= \frac{\partial z}{\partial \eta} J^{-1} \\
\frac{\partial \xi}{\partial z} &= -\frac{\partial x}{\partial \eta} J^{-1} \\
\frac{\partial \eta}{\partial x} &= -\frac{\partial z}{\partial \xi} J^{-1} \\
\frac{\partial \eta}{\partial z} &= \frac{\partial x}{\partial \xi} J^{-1} .
\end{aligned} \tag{26}$$

In this system, we provide the position of the source $(\omega_{10}, \omega_{30})$ in the parametric domain and the ray take-off angle $(\omega_{20}, \omega_{40})$ at the source in the physical domain as initial conditions. Note that the components of the initial ray take-off angle are not independent and they must satisfy the eikonal equation that in our notation takes the form

$$\omega_2^2 + \omega_4^2 = 1. \tag{27}$$

The eikonal equation is satisfied along the whole ray as long as it is satisfied at the initial point of the ray (Cerveny, 1987).

The inverse transformation needed to map initially the position of the source into the parametric space is performed using the *downhill simplex method* (Nelder and Mead, 1965) for non-linear search given by Press *et al.* (1986). However, this computationally expensive calculation is performed only once for each ray. In our implementation, the take-off angle is adjusted using root bisection until the trajectory "hits" the receiver. A fourth order Runge-Kutta scheme has been used for the numerical integration. The infinitesimal increment along the ray-path is calculated in the computational space then mapped into the physical space. Therefore, the distance between the ray-trajectory and the receiver point, necessary to trace the ray properly (i.e., hit the receiver), is calculated in the physical domain. Alternatively, it is also

possible to avoid the computational-physical mapping of each infinitesimal increment by specifying also the position of the receiver in the $\xi\eta$-plane and make the ray hit the receiver in the computational space.

Examples of rays calculated with this ray-tracer are shown in Figure 3. The velocity model was set to a constant value of 2.0 km/s and computational-physical space mapping was performed using various geometries for the grid points (X_{ij}, Z_{ij}) while keeping V_{ij} constant. Thus, the velocity field is maintained constant regardless of the grid point distribution. The curvilinear lines associated with the various grid point geometries are shown in Figure 3. Regardless of the grid geometry, the ray-tracer produces the correct straight ray-paths through the model and the traveltimes for the three curvilinear grids differ by less than 0.001% of the total traveltime.

Ray-trajectory stepping with Equations (24) involves 27 operations, that is, three times those required by standard methods on a regular grid (e.g., Lee and Stewart, 1981). However, the computational cost is dominated by the analytical calculation of the partial derivatives of Equation (9) (i.e., $\frac{\partial v}{\partial \xi}$, $\frac{\partial v}{\partial \eta}$, $\frac{\partial x}{\partial \xi}$, $\frac{\partial x}{\partial \eta}$, $\frac{\partial z}{\partial \xi}$ and $\frac{\partial z}{\partial \eta}$). Each partial derivative calculation involves $2 \times (M+1) \times (N+1)$ operations which, however, can be reduced to $32 = 2 \times 4 \times 4$ because cubic B-splines are defined only on 4 discretization intervals. Overall, the ray-tracing proposed here performs a total of 219 calculations per step—three times those required by standard methods. Therefore, the required computational-physical mapping makes the traveltime calculation CPU intensive but, nonetheless, tractable for inversion problems whose sizes do not differ significantly from those shown in the tests below. In our computer implementation, for models consisting of 10×10 control vertices, it is found that the ray-tracer takes on the order of 1 s of CPU time per ray (source-receiver) on a Sun SPARC10 but this computational time can be consistently reduced if the computational-physical mapping, as explained above, is avoided; if the optimization is improved, or other methods (see below) are used to define the velocity model.

Application

The aim of this section is to assess the performance of the parametric technique when compared to standard fixed-node inversion methods. Note, however, that the same computer algorithm was used for both techniques. By fixing the positions of the grid points, the technique becomes *de facto* a traditional fixed-grid one. To fix the mesh,

we overdamped the spatial adjustments by using $\epsilon_0 = 100,\ b = 1.0$.

Our first objective is to test whether the method is capable of retrieving velocity anomalies of arbitrary shape with a relatively small number of model discretization points. This first series of tests is carried out using a data set having a large information content because we wish to verify the efficacy of the parametric method through the non-linear inversion. In the second series of tests our objective is to evaluate the performance of the method when the information contained in the data set is greatly reduced. Our concern in this case is that the increased number of degrees of freedom inherent in the parametric method, combined with poorer ray-coverage may adversely affect the data misfit minimization through the iterative inversion. In the third test, we address the imaging of more complex models.

To illustrate the performance of the method for the first two series of tests, we have selected the model, MODEL1, shown in Figure 4a which features an oblique S-shaped, 50% maximum positive velocity anomaly. A total of 8×7 grid points were used to generate this model. Note that in the figures displaying the grid geometry, the outermost control vertices which lie outside the target area have not been plotted. The third series of tests used the traveltimes generated with the MODEL2 shown in Figure 12a, which features two oblique S-shaped, 20% maximum velocity anomalies (positive and negative, respectively). A grid consisting of 8×12 knots with two overlapping rows of grid points at $z = -0.8$ was used to represent this velocity model. Uniformly distributed random noise ranging between ± 0.5 ms was added to all the traveltimes which ranged approximately between 0.3 and 0.5 s for MODEL1 and between 0.4 and 0.9 s for MODEL2. This distribution of random noise implies an expected root mean square (RMS) value of the residual times of $0.5/\sqrt{3}$ ms. Inversions were stopped when the RMS was less than this value or when the RMS decrease was less than 2.5% of the current value.

In the parametric inversion algorithm, it is necessary to check that the mapping between the parametric and the physical space is always *one-to-one*. However, in the inversion tests described below, this was always the case owing to the relatively large amount of spatial adjustment damping imposed and to the use of cubic B-splines parameterization, which has an overall stabilizing effect (because cubic B-splines basis functions are supported over four discretization intervals; see Appendix A). Alternatively, it appears possible to add some term to the objective function in

order to penalize the irregularity of the grid mesh and prevent unphysical topological mappings.

MODEL1

This model was obtained by obliquely shifting the grid points along the x-direction (Figure 4b). The background has a constant velocity of 2 km/s whereas the anomaly features a maximum value of 3.0 km/s.

For the first series of tests the acquisition geometry consists of two groups of 9 sources and 9 receivers arranged on the opposite sides of the square target area (Figure 4c). This acquisition geometry results in the calculation of 162 ray-paths and it will be referred in figures and text as the "medical" set-up—by analogy with the acquisition geometries in use in medical tomography. Overall, this set-up simulates situations in which the information content of the data set allows accurate model reconstructions. A somewhat similar data acquisition setup has been planned for the cross-volcano experiment that will be conducted on the flanks of Mount Etna (ETNAseis). The second series of tests uses the depleted source-receiver geometry shown in Figure 4d—a cross-hole experiment. This second acquisition set-up has a poorer ray coverage and only a smeared image of the true model is expected.

The initial velocity distribution for all the inversions was set to a constant 2.2 km/s and initially the control vertices were equally spaced along the horizontal and vertical coordinates. The RMS misfits obtained with this initial model were 26.2 and 14.9 ms for the medical and cross-hole acquisition setup, respectively. In Figures 5, 8, 9 and 10 we show the initial and final position of the grid points together with the curvilinear grid generated using the parametric relations of Equation (2). Each family of curves was generated by varying one coordinate (e.g., $\xi = 0, 1, 2, \cdots, M$) while holding the other constant.

The inversion was carried out using four meshes—control vertices spaced 0.2, 0.15, 0.10 and 0.05 km along x and z for a total of 6×6, 7×7, 9×9 and 17×17 grid points, respectively (Figure 5). In addition, inversions were performed using unevenly spaced and skewed, 6×6, initial distributions of the grid points (Figure 9) and with configurations that entailed (for the adaptive- and the fixed-grid methods) a similar total number of inversion parameters (Figure 7).

"Medical" Inversions

The results obtained using the parametric formalism are compared with those from the fixed-grid method in Figure 6. Even when the inversion grid is coarse (i.e., 0.2 and 0.15 km) the adaptive grid technique is capable of fully exploiting the information contained in the data set. Very similar velocity models are obtained regardless of the initial grid spacing (Figure 6e-h). Conversely, the "rigidity" of fixed grids inhibits proper convergence to the true model (Figure 6a,b) unless the number of discretization points is increased (Figure 6c,d). In this example, accurate reconstructions are obtained with fixed nodes only when grid spacings of 0.1 km or less are employed (Figure 6c,d). There are, however, only minor differences between the results obtained using the adaptive- and fixed-grid techniques when the spacing is reduced to 0.05 km (Figure 6d,h). This behavior is not unexpected and it can be explained as follows. When the inversion grid (fixed or adaptable) is sufficiently dense, there is a greater likelihood that the grid points are properly positioned, and node position adjustments will contribute less significantly toward data misfit minimization, the velocity perturbations playing the predominant role. Consequently, as the data misfit reaches values close to the noise level, the velocity model refinement achievable by grid position adjustments becomes negligible. Conversely, with a wider-spaced grid, the velocity adjustments alone are unable to reduce the misfit to values close to the noise level and the grid point perturbations become effective (see Table 2).

Table 2 shows that the only case in which the parametric formalism performs (slightly) worse than the fixed grid approach is when the spacing is reduced to 0.05 km. The slight image deterioration evident in Figure 6h compared with Figure 6d results from the competeing roles that velocity and position adjustments play in misfit minimization. This creates the hole in the node distribution at the anomaly, which inhibits full recovery of the true model (Figure 5h). Note, however, that a larger value of grid position damping would have probably avoided this problem. Table 2 lists also the total number of iterations required for convergence. This number is generally larger for coarser meshes because some of the grid points have to undergo larger total displacements from their original positions.

In the tests above, we have compared the performance of the adaptive- and fixed-grid inversion methods using identical-size grid meshes. This approach, however,

neglects that the effective number of inversion parameters is three times larger in the adaptive-grid method than in the fixed one. Therefore, in Figure 7, we show the results obtained when a similar total number of inversion parameters is used, that is, we compare inversions with dense, fixed, meshes and relatively coarse, adaptive ones. In order to display the effective differences between the true and the calculated models, we plot the absolute value of the percentual difference between the two models (i.e., $\| (true - calculated)/true \| \times 100$). In detail, we have selected four test cases. In the first two (Figure 7a and 7e, and Figure 7b and 7f) the number of inversion parameters is smaller than that used to generate the true model, that is, 168 ($= 8 \times 7 \times 3$) parameters were used to define the true model whereas 100 ($= 10 \times 10$) and 144 ($= 12 \times 12$) were used in the inversions with the fixed-grid mesh and, 108 ($= 6 \times 6 \times 3$) and 147 ($= 7 \times 7 \times 3$) for the adaptive-grid inversion. The third test involves a total number of parameters similar to that of the true model ($169 = 13 \times 13$ and $168 = 8 \times 7 \times 3$ for the fixed- and adaptive-grid inversions, respectively; Figure 7c and 7g) and the fourth was completed using a total number of parameters much larger than that used to define the true model ($289 = 17 \times 17$ and $300 = 10 \times 10 \times 3$ for the fixed- and adaptive-grid inversions, respectively, Figure 7d and 7h). Although the differences between the true and the calculated models using the two methods are never larger than 10%, the fit of the adaptive-grid method is generally superior because the method can effectively find the best node configuration to reproduce the true model. In contrast, the fixed grid method seems to create some spurious horizontal and vertical bands having larger misfit. Presumably, the origin of these bands is to be found in the non-optimal positioning of the grid-points and in the increased indeterminacy that results from reducing the scalelength of the parameterization when using fixed meshes.

In general, the parametric approach permits use of relatively coarse grids for accurate reconstructions while minimizing the insurgence of instabilities. However, the use of a dense grid mesh can accentuate the insurgence of instabilities owing to the increased indeterminacy resulting from the large number of parameters sought simultaneously. Therefore, it appears that the robustness of the results should always be tested by repeating the inversion using various grid spacings—a procedure that should be also followed with fixed grid meshes. One advantage of the adaptive grid approach is that, with the exception of the very dense meshes described above, almost identical final models are retrieved regardless of the grid spacing whereas with fixed-

grid methods the final model always depends to some extent upon the predefined node spacing.

In all the inversions using the adaptive grid method described above, we have solved simultaneously for velocity and position adjustments using a constant velocity initial model that is significantly removed from the true model in terms of RMS residual misfit. Figure 8 shows the results obtained when the models obtained from the fixed-grid inversions (Figures 6a-c) were subsequently refined by control vertex adjustments using the parametric inversion. These images are nearly identical to those displayed in Figure 6e-g (see also the RMS values listed in Table 2), but the distribution of grid points and the associated curvilinear grid (Figure 8a-c) now differ from those obtained when both spatial and velocity adjustemnts are simultaneously inverted for throughout (Figure 5e-g). Therefore, there appears to be no unique distribution of a given set of control vertices that is capable of reproducing the true model accurately. This fact can be visualized by comparing the final models obtained in Figures 6e-g and 8d-f and the grids of Figures 5e-g and 8a-c. In Figure 9, the inversion has been repeated using two additional 6×6 initial configurations of the control vertices (Figure 9a,b). Similar final models (Figure 9e,f) are obtained regardless of the final grid node configuration (Figure 9c,d). The RMS residual misfit is 0.78 and 0.80 ms for the models shown in Figures 9e and 9f, respectively. Therefore, by redistributing the grid points and readjusting the associated velocities, the adaptive gridding technique formalism can converge to different final configurations of the control vertices while still reproducing the model faithfully. I speculate that in non-linear inversions this feature can be advantageous because it enlarges the range of successful paths toward minimization of the data misfit functional.

The number of iterations (Table 2) is comparable for parametric inversions starting from fixed and relaxed initial models. Thus, refinement by adaptive gridding, although effective is, nonetheless, as expensive computationally as using adaptive gridding from the start, but this does not preclude the method from being used to refine results obtained using fixed-grid methods. For example, in local earthquake tomography, it rarely occurs that the final data misfit at convergence of the inversion falls within the observational errors of the data set. This is probably due to inadequate meshing of the target volume which the adaptive-grid technique obviates.

Cross-hole Inversions

The second data acquisition geometry is designed to investigate the performance of the parametric approach when the resolving power of the data set is such that only a smeared image of the true model can be expected. As we remarked above, the combination of a large number of degrees of freedom and a partial lack of resolving power might retard and/or bias the solution. Figure 10 shows the initial fixed grids and the curvilinear grid resulting from the parametric inversion. The results are compared in Figure 11. As expected, neither approach is capable of reproducing the true model accurately. However, the parametric inversion with a relatively small number of grid points provides results which are comparable to those obtained with the denser fixed meshes (compare Figures 11d and 11e-h). Therefore, it appears that the advantages of the adaptive-grid method are maintained without introducing significant bias. The pattern of RMS data misfits in Table 3 is also similar to that of Table 2. However, the slightly smaller RMS values in Table 3 and the poorer image reconstruction attained with this set-up illustrate the pitfalls inherent in under-determined inversion problems. This is a classic case of inversion non-uniqueness where, given the poor constraints on the model provided by the data set, minimization stops when it reaches values within the noise level, but there is no guarantee that the resolved model is the true one.

MODEL2

MODEL2 and the data acquisition geometry used to image it is shown in Figures 12a and 12c. This acquisition set-up is similar to that used by Berryman (1989) and by Ammon and Vidale (1993). Again, the S-like character of the anomalies was generated by shifting the corresponding grid points along the x-direction (Figure 12b). A constant 2.0 km/s initial model was used for all the inversions. Traveltimes calculated with this model produced an RMS misfit of 22.1 ms. In order to avoid poor convergence of the algorithm, some amount of row weighting was needed. Rows of (21) whose source-receiver distances fell in the range between 1.2 and 2.0 km where down-weighted from one to zero. Therefore, the rows of the longest rays (from top to bottom) were scaled down to approximately one fourth of their original value. In Table 4, we summarize the final RMS misfit values obtained at convergence for the various cases that were analysed. Again, the number in parentheses indicates the total number of iterations for each inversion test. In general, the number of iterations varies according to the

amount of damping that is added to regularize the solution and to the number of grid points in the mesh. In this test, about 20 iterations were needed for the parametric method. However, only three to eight iterations, depending on the grid mesh, were needed to decrease the residual misfit to about 10% of the initial one. This number of iterations is typical for non-linear inversions of this kind.

Figures 13 and 14 compare representative results obtained with the adaptive- and the fixed-grid methods, respectively. Adaptive inversions having relatively coarse meshes (0.2 and 0.15 km; Figure 13) are capable of reproducing adequately the anomalies of the true model, although the sharpness that distinguishes the medical inversion experiment cannot now be attained owing mainly to the greater model complexity. As expected, regardless of the method, it is more difficult to image the low velocity anomaly because of the inherently sparse ray-coverage there (see Figure 12c), but the parametric method does substantially better (compare Figures 13a and 13c with Figure 14a). This is due to the way the parametric method uses and expands the parameter space. With standard methods, an increase in the number of parameters is always associated with a decrease in the discretization interval. This increases the problem indeterminacy because fewer rays sample each node. This is accentuated by the already sparse ray density within the low-velocity zone. In contrast, the model indeterminacy resulting from triplication (in 2-D) of the size of the parameter space in the adaptive-grid method is offset by the fact that this method can accurately reproduce the velocity field using a coarse mesh. Therefore, fine discretization is unnecessary and matrix sparseness, one of the main sources of indeterminacy, is avoided altogether.

Discussion

The approach taken in this study differs significantly from those previously proposed for traveltime tomography. Here, the inverse problem involves determination of the optimal position of the grid points and their associated velocities to represent the velocity model. In contrast, standard methods rely on an initial and fixed choice of the discretization mesh.

In general, there is a variety of ways to express a specific surface parametrically. Each can be broadly classified as being based on "interpolation", in which the surface is required to pass through the control vertices, or on "approximation" in which the

surface passes "near" the control vertices (Bartels *et al.*, 1987). For example, linear B-splines basis functions ensure interpolation between the control vertices of the surface whereas cubic B-splines—the basis used in this study—approximate the surface to the control vertices. Note, however, that the use of cubic B-splines does not preclude the generation of sharp variations of the velocity surface in correspondence with the control vertices (e.g., Figure 2). This is an extension to parametric surface representation of the node multiplicity property used by Moser *et al.* (1992) to model first order discontinuities for ray-tracing, and is one of the principal motivations for selecting such basis.

A second remarkable feature of the parametric representation with cubic B-splines basis is that nearly identical models can be reproduced using different configurations of the grid points (see the 6×6 meshes of Figures 5 to 9). In fact, differences in grid point configurations are balanced by variations in the associated velocities and, as the grid point configuration varies, complementary variations in the associated velocities produce similar final velocity models. This effective trade-off between nodal positions and velocity values would indicate that in our tests, even with coarse meshes, the velocity model is *de facto* over-parameterized so that different configurations of the control vertices can effectively reproduce similar images. In this context, an inversion test carried out with a 5×5 variable grid (not shown, final RMS=2.37 ms) was still capable of reproducing the main features of the true structure whereas the 4×4 adaptive-grid inversion failed being too under-parameterized. The point is that the complexity of a structure should be quantified by the number of parameters needed to define it (to within some degree of "closeness"). Therefore, the models presented in this study are not really that "complex" if the grid points are well-positioned. Conversely, these models are "complex" and require more parameters if fixed, rectangular, grid geometries are used. The results shown in Figure 7 well expose that over-parameterization using fixed-grids may not be the best strategy to follow in inversion problems of this kind. However, the models used here are well-suited to demonstrate the efficacy of the method in that they comprise isolated, irregular, velocity anomalies which can be reproduced using a relatively small number of well-positioned grid points. In contrast, there may be other styles of model which can be well defined by a relatively dense fixed grid but not by a coarse adaptive one for the same total number of parameters. For example, a model consisting of Nyquist sinusoids defined on a regularly spaced

dense grid cannot be represented by an adaptive one using a comparable total number of inversion parameters (e.g., $10 \times 10 = 100$ for the fixed- and $6 \times 6 \times 3 = 108$ for the adaptive-grid, respectively).

It is well known that, apart from the methodology, tomographic reconstructions depend both upon the quality of the data and their resolving power. In practice and with standard methods, an increase in size of the model space results in a decrease of the scale length of the model discretization. Therefore, the associated ray coverage per model parameter decreases so that the resolving power of the data onto the model parameters is similarly decreased (e.g., Williamson, 1990; Michelini, 1993). The parametric method triples the number of model parameters but avoids the indeterminacy caused by the decrease in scale length of the model discretization. This is probably one of the main features of this approach because it allows the use of coarse discretization meshes obviating the inversion of sparse, often near singular, matrices. However, the mentioned trade-off between nodal positions and velocity values indicates that some near-zero singular values of the parameter derivatives matrix still occur but do not seem to affect significantly the model in xzv-space if their contribution is properly damped as explained below.

Because the parametric formalism involves simultaneous inversion for both velocity and position adjustments, to assign the relative amount of damping for position and velocity adjustments in the inversion of the partitioned matrix (see Equation 21) is probably the most intricate task for successful model reconstructions. We have used Kennett *at al*'s (1988) subspace method and applied damping based on the value of the largest eigenvalue of each class of parameters but other matrix inversion methods and damping criterions could have been adopted. Our selection of the damping values reflects the observation that during the first iterations the dominant role toward misfit minimization is played by the velocity adjustments whereas grid point adjustments become preponderant only in the final iterations that is, when the size of the velocity anomalies is close to that of the true model and position adjustments alone can substantially improve the model fit by modifying the shape of the anomalies. A second motivation for overdamping the position adjustments follows from preventing the formation of unphysical topological foldings of the velocity surface.

The convenient features of the parametric method, however, trade-off with the total computational time needed to complete the non-linear inversion. In our appli-

cations the CPU time for the iterative inversion is tractable (15 to 60 minutes on a SPARC10 depending on the total number of iterations for convergence, the grid size and, on the integration step along the ray-path) but, for tomographic problems involving several thousends of traveltime calculations, the ray tracing scheme proposed here may become a burden to the methodology. The most costly part of the ray-tracer is the analytical calculation of the spatial derivatives (i.e., Equation (9)) which depends on the basis functions used for model representation. Therefore, some significant computational saving can be achieved by a more thorough optimization of the derivatives calculation or, by using simpler basis functions (e.g., linear B-splines would reduce the computational cost to one fourth). In addition, some CPU saving is expected by (1) redefining the model in terms of square of slowness (Psencik, written communication) or, (2) by bending the ray trajectory adopting the approach proposed by Um and Thurber (1987).

In the application of the method, we suggest starting from relatively coarse initial meshes and testing the robustness of the results by using successively denser meshes. In light that some styles of model can be ill-suited to be defined by relatively coarse meshes, it is also advisable to test the inversion using the fixed-grid method with a dense mesh. In any event, selection of the initial configuration of the grid mesh appears less troublesome than with standard methods because the true shape of the velocity anomalies can be recovered regardless of the initial distribution of the grid points. The method is also rather flexible because the nodal positions can be constrained at will. For example, the grid could be held fixed during the initial iterations and the results refined by adaptive gridding during the final iterations.

In all our tests, the noise added to the data is a small fraction of the total traveltime and larger values of noise would have certainly deteriorated (smeared) the reconstructions. Various amounts of uniformly distributed random noise in the range 2.5 to 10 ms have been to the MODEL1, *medical inversions*, data set. The results obtained with the fixed- and the adaptive-grid methods are shown in Figure 15 where we have inverted the data using a nearly equivalent number of effective model parameters (i.e., 169 and 168 for the fixed- and adaptive-grid methods, respectively; see also Figure 7). All inversions were stopped when the data misfit decreased to values less than the expected RMS misfit. As predicted, the sharpness of the reconstructions depends on the noise level but, because of the larger scalelenght in model discretization

and the optimal positioning of the nodes, the parametric formalism appears less prone to create the instabilities observable in the fixed-grid inversion results.

In Paper I and here, the number of control vertices used for the interpolation is maintained fixed throughout the iterations of the inversion—the parametric technique seeks their optimal position. A different strategy would consist of using rectangular grid methods and devise schemes that "sense" the formation of velocity gradients through the inversion to increase or deplete locally the number of grid points accordingly. This is a promising approach to the tomographic problem but it relies on the selection of a robust criterion for isolating those parts of the model where the density of the mesh should be varied. The adaptive grid formalism presented here inverts for the position of the grid points and the criterion above is not needed.

In order to solve the non-linear inverse problem of traveltime tomography by local linearization of the misfit functional and through iterations, we have presented the analytical relations required for the determination of the traveltime partial derivatives with respect to the model control vertex positions. This local linearization approach is required by the computationally intensive forward calculation of the traveltimes. In inversion problems where the forward calculation is less demanding, local linearization and calculation of the partial derivatives are not needed but the parametric representation can still be used to describe the model and fully non-linear methods employed to determine the optimal mesh configuration.

Conclusions

The underlying idea of this study is that a generalized curvilinear coordinate basis can be more effective in defining complex velocity structures than a rectangular one. To this purpose it is introduced: (1) the use of velocity models defined parametrically; (2) a ray-tracer through such models and (3) an adaptive-grid formalism for non-linear traveltime tomography.

In order to define the velocity model parametrically, we have chosen cubic B-splines basis functions because they are particularly versatile in the reproduction of complex geologic structures with high fidelity while using a minimal number of discretization grid points.

Because the bulk of the computational effort required by non-linear tomography is in the ray-tracing itself and existing ray-tracing algorithms are designed for rectangu-

lar or spherical geometries whereas the parametric representation involves generalized curvilinear coordinates, we have developed an exact, initial value, ray-tracer which relies on numerical integration along the ray-path in the parametric (computational) domain. Computationally, this ray-tracer performs three times as many operations than its correspondent in rectangular coordinates which makes the parametric formalism of practical use for tomographic problems of approximately the same size as those presented here. Although in its current, unoptimized, version the ray-tracer is moderately slow, faster schemes are possible and are under investigation.

The following conclusions are drawn from application of the adaptive-grid technique to various synthetic data sets:

i.) The method always produced good fits to the true model while using relatively coarse grid meshes. When compared to the standard fixed-grid method and for models defined using the same total number parameters, the method is shown to be generally more accurate. However, I remark that the models used here are well suited to demonstrate the methodology but there may be other kinds of model featuring continuous patterns of short wavelength anomalies which cannot be represented using coarse adjustable grids for the same total number of inversion parameters.

ii.) Adaptive gridding partly obviates the model indeterminacy that afflicts standard methods using dense grids. Because the adaptive grid method seeks the optimal position of the grid points, relatively coarse meshes can still be used avoiding the inversion of large, sparse, matrices. Nevertheless, the derivatives matrix can still display some near-singular values whose damaging effects can be removed by damping.

iii.) In our applications, different configurations of the same set of control vertices appear to reproduce accurately the true model (to within a certain degree of "closeness"). This effective trade-off between node positions and velocity values follows from the versatility of cubic B-splines for model definition. In addition, it would indicate that the true models adopted here are not particularly complex to reconstruct if the inversion grid points are well-positioned. I conjecture that this added freedom can be beneficial to non-linear inversions because it increases the probability of reproducing the true model regardless of the starting velocity model.

iv.) Inversion grid meshes having different densities of the grid points can reproduce the true model with equal fidelity. This can be an advantage with respect to standard fixed-grid methods when testing the robustness of the results.

In this paper, the adaptive-grid method has been applied to a 2-D problem but it appears particularly suited also for 3-D tomographic inversion problems, such as simultaneous determination of earthquake locations and velocity structure. Overall, the parametric representation can find application in other kinds of geophysical problem.

Acknowledgments

The author is indebted to G. Caristi and G. Cassiani for discussions on the theory of inverse functions. The thorough review of Bill Foxall and the suggestions and criticism provided by Ivan Psencik, Jose Carcione and two anonymous reviewers greatly improved the original manuscript. The author feels indebted to one of the anonymous reviewers for his (her) particularly constructive criticism. All the figures were generated using the GMT-graphics package of Wessel and Smith (1991). The work has been supported by CNR contract 9400193CT05 and through MURST 40% and 60% funds.

References

Ammon, C. J. and J. E. Vidale (1993). Tomography without rays, *Bull. Seism. Soc. Am.*, **83**, 509-528.

Bartels, R.H., J.C. Beatty and B.A. Barsky (1987). *An introduction to Splines for use in Computer Graphics and Geometry Modeling*, Morgan Kaufmann Publishers, Inc, Los Altos, California, 476 pp.

Berryman, J. G. (1989). Fermat's principle and non-linear traveltime tomography, *Phys. Rev. Lett.*, **62**, 2953-2956.

Carcione, J. M. and J-P Wang (1993). A Chebyshev collocation method for the elastodynamic equation in generalized coordinates, *Comp. Fluid Dynam. Jour.*, **2**, 269-290.

Cerveny, V. (1987). Ray tracing algorithms in three-dimensional laterally varying layered structures, in *Seismic Tomography*, G. Nolet, (Editor), D. Reidel Publishing Company, 99-133.

Chou, C. W. and J. R. Booker (1979). A Backus-Gilbert approach to the inversion of traveltime data for three-dimensional velocity structure, *Geophys. J. R. Astr. Soc.*, **59**, 325-344.

Courant, R. (1936). *Differential and integral calculus (Vol. 2)*, Wiley-Interscience , 682 pp.

Gudmunsson, O. and Clayton, R. W. (1991). A 2-D synthetic study of global traveltime tomography, *Geophys. Journ. Int.*, **106**, 53-66.

Fornberg, B. (1988). The pseudospectral method: Accurate representation of interfaces in elastic wave calculations, *Geophysics*, **53**, 625-637.

Kennett, B.L.N., M.S. Sambridge and P.R. Williamson (1988). Subspace methods for large inverse problems with multiple parameter classes, *Geophys. Journ. Int.*, **94**, 237-247.

Lee, W. H. K. and S. W. Stewart (1981). *Principles and applications of microearthquake networks*, Academic Press, Inc., 293 pp.

Michelena, R.J. and J.M. Harris (1991). Tomographic traveltime inversion using natural pixels, *Geophysics*, **56**, 635-644.

Michelini, A. (1991). Fault zone structure determined through the analysis of earthquake arrival times, *Ph.D. Thesis*, University of California, Berkeley, 191 pp.

Michelini, A. (1993). Velocity model inversion using parametric curves, *Geophys. Journ. Int.*, **115**, 337-343.

Moser, T.J., G. Nolet and R. Snieder (1992). Ray bending revisited, *Bull. Seismol. Soc. Am.*, **82**, 259-288.

Nelder, J.A., and R. Mead (1965). A simplex method for function minimization, *Computer J.*, **7**, 308.

Nielsen, P., Berg, P., If, F. and Skovgaard, O. (1992). Using the pseudospectral method on curved grids for seismic forward modelling, European Association of Exploration Geophysicists, *54th Meeting and Technical Exhibition, Expanded Abstracts*, , 150-151.

Pereyra, V. (1992). Two-point ray tracing in general 3-D media, *Geophys. Prospect.*, **40**, 267-287.

Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Vetterling (1986). *Numerical recipes, the art of scientific computing*, Cambridge University Press, 818 pp.

Sambridge, M.S. (1990). Non-linear arrival time inversion: constraining velocity anomalies by seeking smooth models in 3-D, *Geophys. Journ. Int.*, **102**, 653-677.

Snieder, R., Beckers, J. and F. Neele (1991). The effect of small scale structure on normal ,mode frequencies and global inversions, *Journ. Geophys. Res.*, **96**, 501-515.

Spakman, W (1993). Iterative strategies for non-linear travel time tomography using global earthquake data, in *Seismic Tomography: Theory and practice*, H. M. Iyer and K. Hirahara, (Editor), Chapman & Hall, 190-226.

Tarantola, A. and A. Nercessian (1984). Three-dimensional inversion without blocks, *Geophys. J. R. Astr. Soc.*, **76**, 299-306.

Thompson, J. F., Warsi, Z. U. A., and C. W. Mastin (1985). *Numerical grid generation, foundations and applications*, North Holland, 483 pp.

Thurber, C.H. (1993). Local earthquake tomography: velocities and Vp/Vs—theory, in *Seismic Tomography: Theory and practice*, H. M. Iyer and K. Hirahara, (Editor), Chapman & Hall, 563-583.

Um, J. and C. H. Thurber (1987). Rapid solution of ray tracing problems in heterogeneous media, *Bull. Seism. Soc. Am.*, **77**, 972-986.

Wessel, P., and W. H. F. Smith (1991). Free software helps map and display data, *EOS Trans. AGU*, **72**, 441.

Williamson, P.R. (1990). Tomographic inversion in reflection seismology, *Geophys. Journ. Int.*, **100**, 255-274.

Appendix A: Cubic B-splines

We follow the formulation given by Lancaster and Salkauskas (1986). Cubic B-splines are a set of basis functions ρ_i taken from the vector space S_M of piecewise cubic functions which has dimensions $M+3$. To determine uniquely a cubic B-spline function through a set of $M+1$ nodes, it is constraint the second spatial derivatives at the boundary nodes ($i = 0, M$) to be equal to zero. The smoothness in cubic B-splines (continuity up to the second spatial derivative) demands the support of at least four consecutive intervals for internal knots and three or two intervals at the boundaries. The normalized cubic B-splines functions satisfying these boundary conditions are defined as follows:

$$\rho_0(\xi) = \begin{cases} \frac{1}{6}(\xi - \xi_0)^3 - (\xi - \xi_0) + 1 & \xi_0 \le \xi \le \xi_1; \\ -\frac{1}{6}(\xi - \xi_1)^3 + \frac{1}{2}(\xi - \xi_1)^2 + & \\ -\frac{1}{2}(\xi - \xi_1) + \frac{1}{6}, & \xi_1 \le \xi \le \xi_2; \\ 0, & \xi_2 \le \xi \le \xi_M; \end{cases} \tag{A1}$$

$$\rho_1(\xi) = \begin{cases} -\frac{1}{3}(\xi - \xi_0)^3 + (\xi - \xi_0), & \xi_0 \le \xi \le \xi_1; \\ \frac{1}{2}(\xi - \xi_1)^3 + (\xi - \xi_1)^2 + \frac{2}{3}, & \xi_1 \le \xi \le \xi_2; \\ -\frac{1}{6}(\xi - \xi_2)^3 + \frac{1}{2}(\xi - \xi_2)^2 + & \\ -\frac{1}{2}(\xi - \xi_2) + \frac{1}{6}, & \xi_2 \le \xi \le \xi_3; \\ 0, & \xi_3 \le \xi \le \xi_M; \end{cases} \tag{A2}$$

$$\rho_i(\xi) = \begin{cases} 0, & \xi_0 \le \xi \le \xi_{i-2}; \\ \frac{1}{6}(\xi - \xi_{i-2})^3, & \xi_{i-2} \le \xi \le \xi_{i-1}; \\ -\frac{1}{2}(\xi - \xi_{i-1})^3 + \frac{1}{2}(\xi - \xi_{i-1})^2 & \\ +\frac{1}{2}(\xi - \xi_{i-1}) + \frac{1}{6}, & \xi_{i-1} \le \xi \le \xi_i; \\ \frac{1}{2}(\xi - \xi_i)^3 - (\xi - \xi_i)^2 + \frac{2}{3}, & \xi_i \le \xi \le \xi_{i+1}; \\ -\frac{1}{6}(\xi - \xi_{i+1})^3 + \frac{1}{2}(\xi - \xi_{i+1})^2 + & \\ -\frac{1}{2}(\xi - \xi_{i+1}) + \frac{1}{6}, & \xi_{i+1} \le \xi \le \xi_{i+2}; \\ 0, & \xi_{i+2} \le \xi \le \xi_M; \end{cases} \tag{A3}$$

$$\rho_{M-1}(\xi) = \begin{cases} 0, & \xi_0 \le \xi \le \xi_{M-3}; \\ \frac{1}{6}(\xi - \xi_{M-3})^3, & \xi_{M-3} \le \xi \le \xi_{M-2}; \\ -\frac{1}{2}(\xi - \xi_{M-2})^3 + \frac{1}{2}(\xi - \xi_{M-2})^2 & \\ +\frac{1}{2}(\xi - \xi_{M-2}) + \frac{1}{6}, & \xi_{M-2} \le \xi \le \xi_{M-1}; \\ \frac{1}{3}(\xi - \xi_{M-1})^3 - (\xi - \xi_{M-1})^2 + & \\ +\frac{2}{3}, & \xi_{M-1} \le \xi \le \xi_M; \end{cases} \tag{A4}$$

$$\rho_M(\xi) = \begin{cases} 0, & \xi_0 \le \xi \le \xi_{M-2}; \\ \frac{1}{6}(\xi - \xi_{M-2})^3, & \xi_{M-2} \le \xi \le \xi_{M-1}; \\ -\frac{1}{6}(\xi - \xi_{M-1})^3 + \frac{1}{2}(\xi - \xi_{M-1})^2 + & \\ \frac{1}{2}(\xi - \xi_{M-1}) + \frac{1}{6}, & \xi_{M-1} \le \xi \le \xi_M; \end{cases} \tag{A5}$$

where ξ is the independent variable which varies between some minimum, ξ_0, and a maximum ξ_M. The sequence of knots, ξ_i $(i = 0, \cdots, M)$, must be non-decreasing and, for convenience, are equally spaced (i.e., $\xi_{i+1} = \xi_i + 1$). Analogous relations apply along the η parametric coordinate.

Figure Captions

Figure 1. Example of *one-to-one* mapping from region R in the $\xi\eta$-plane ($0 \leq \xi \leq 4, 0 \leq \eta \leq 4$) onto the xz-plane. The solid circles indicate the grid point positions (X_{ij}, Z_{ij}) used in the functional relationships of Equation (2). The thin lines represent the two families of curvilinear grid lines obtained from the chosen geometry of grid points. Each line is generated by fixing one parametric coordinate and varying the other (e.g., ξ is fixed to 0 whereas η is increased mononically from 0 to M; see text).

Figure 2. Example of velocity model generated using the parametric representation. *a)* Perspective view; *b)* grid points used to generate the model (solid circles) and the curvilinear lines (thin solid lines). Each family of curvilinear lines is drawn using the mapping relations of Equation (2); *c)* plan view.

Figure 3. Example of ray-tracing through a velocity model defined parametrically. The velocity is constant ($v = 2.0$ km/s) throughout the model in all cases. Ray paths are calculated for: *a)* an evenly spaced mesh of grid points; *b)* two coinciding grid points at $x = 0.25$, $z = -0.1$ and $x = 0.25$, $z = -0.25$ and *c)* the two coinciding grid points of the previous case plus two additional coinciding points at $x = 0.55$, $z = -0.1$ and $x = 0.7$, $z = -0.1$. Stars indicate the coinciding grid points.

Figure 4. Velocity model and data acquisition geometry for MODEL1: *a)* model used to generate the traveltimes; *b)* curvilinear grid lines (solid circles indicate the position of the grid points); *c)* ray coverage for the "medical" data acquisition; *d)* ray coverage for the "cross-hole" data acquisition. Sources and receivers are indicated by solid stars and triangles, respectively.

Figure 5. "Medical" inversion results: curvilinear grid lines and grid point positions. *a-d)* fixed-grid and initial configuration of the grid points for the adaptive-grid inversion; *e-h)* final grid point positions for the adaptive-grid inversion. Note that the boundary grid points are not plotted because they lie outside the shown physical space.

Figure 6. "Medical" inversion results: velocity models resulting from the inversion using the fixed- (*a-d*) and adaptive-grid methods (*e-h*).

Figure 7. "Medical" inversion results obtained with the adaptive- (*e-h*) and fixed-grid (*a-d*) methods for a similar total number of inversion parameters. It is plotted the absolute value percentual difference between the true and the calculated models. The whitened areas show percentual differences larger than 5%. The number in parenthesis above the fixed-grid panels indicates the node spacing in km.

Figure 8. "Medical" inversion results obtained with the best fitting, fixed-grid initial models of Figure 6a-c: *a-c)* curvilinear grid lines and grid point positions; *d-f)*

correspondent velocity models.

Figure 9. "Medical" inversion results obtained using two additional 6×6 initial configurations of the grid points *(a,b)*. Panels *(c,d)* show the final configurations of the grid points whereas the final velocity models are shown in *(e,f)*.

Figure 10. "Cross-hole" inversion results: same format as Figure 5.

Figure 11. "Cross-hole" inversion results: same format as Figure 6.

Figure 12. Velocity model and data acquisition geometry for MODEL2: *a)* model used to generate the traveltimes; *b)* curvilinear grid lines (solid circles represent the grid point positions); *c)* ray coverage for the data acquisition. Sources and receivers are indicated by solid stars and triangles, respectively.

Figure 13. MODEL2 inversion results using the adaptive-grid formalism. *a,c)* Final velocity models obtained from the indicated mesh size and spacing; *b,d)* final configurations of the grid points and of the correspondent curvilinear lines.

Figure 14. MODEL2 inversion results using the fixed-grid formalism. *a,c)* Final velocity models with the indicated mesh size and spacing; *b,d)* equally spaced mesh and associated grid.

Figure 15. "Medical" inversion results obtained by adding different amounts of uniformly distributed random noise to the synthetic traveltimes. The amount of noise is indicated above each panel. A nearly identical number of parameters are used in inverting the data (i.e., $169 = 13 \times 13$ and $168 = 8 \times 7 \times 3$ for the fixed- and adaptive-grid methods, respectively). Final velocity models obtained with the fixed- (*a-c*) and the adaptive-grid (*d-f*) method.

Table Captions

Table 1. Position of the control vertices used to generate the model shown in Figure 2 using the cubic B-splines basis functions of Equation 2.

Table 2. MODEL1: final RMS residual times obtained with with the parametric (PAR), the standard fixed (FIXED) node representations and by relaxing the position of the grid-nodes when the final fixed-grid results are used as starting velocity models (PARFIN). The initial RMS is 26.2 *ms* for the PAR and FIXED inversions. For the PARFIN inversions, the initial RMS is that of correspondent FIXED value. The integer provided in parenthesis is the total number of iterations needed for convergence.

Table 3. MODEL1: final RMS residual times obtained using the cross-hole set-up. The initial RMS data misfit is 14.9 *ms*. The integer provided in parenthesis is the total number of iterations needed for convergence.

Table 4. MODEL2: final RMS residual times obtained using the data acquisition geometry of Figure 11. The initial RMS data misfit is 22.1 *ms*. The integer provided in parenthesis is the total number of iterations needed for convergence.

CONTROL VERTICES				
i	j	X_{ij}	Z_{ij}	V_{ij}
0	0	-0.5	-1.5	2.0
1	0	0.3	-1.5	2.0
2	0	0.5	-1.5	2.0
3	0	1.5	-1.5	2.0
0	1,2,3	-0.5	-0.5	2.0
1	1,2,3	0.3	-0.8	2.0
2	1,2,3	0.5	-0.5	2.0
3	1,2,3	1.5	-0.8	2.0
0	4,5,6	-0.5	-0.3	3.0
1	4,5,6	0.3	-0.4	3.0
2	4,5,6	0.5	-0.2	3.0
3	4,5,6	1.5	-0.3	3.0
0	7	-0.5	-0.2	1.0
1	7	0.3	-0.3	1.0
2	7	0.5	-0.1	1.0
3	7	1.5	-0.2	1.0
0	8	-0.5	0.5	2.0
1	8	0.3	0.5	2.0
2	8	0.5	0.5	2.0
3	8	1.5	0.5	2.0

Table 1:

MODEL1: MEDICAL RMS DATA MISFIT (ms)				
Mesh	Spacing (km)	PAR	FIXED	PARFIN
6x6	0.2	0.74 (19)	3.62 (10)	0.61 (23)
7x7	0.15	0.48 (20)	2.08 (14)	0.65 (15)
8x7	0.114 (x) 0.15 (z)	0.40 (21)		
9x9	0.1	0.48 (14)	0.98 (16)	0.45 (13)
10x10	0.089	0.38 (24)	0.69 (16)	
12x12	0.073		0.56 (14)	
13x13	0.067		0.44 (17)	
17x17	0.05	0.41 (17)	0.31 (17)	

Table 2:

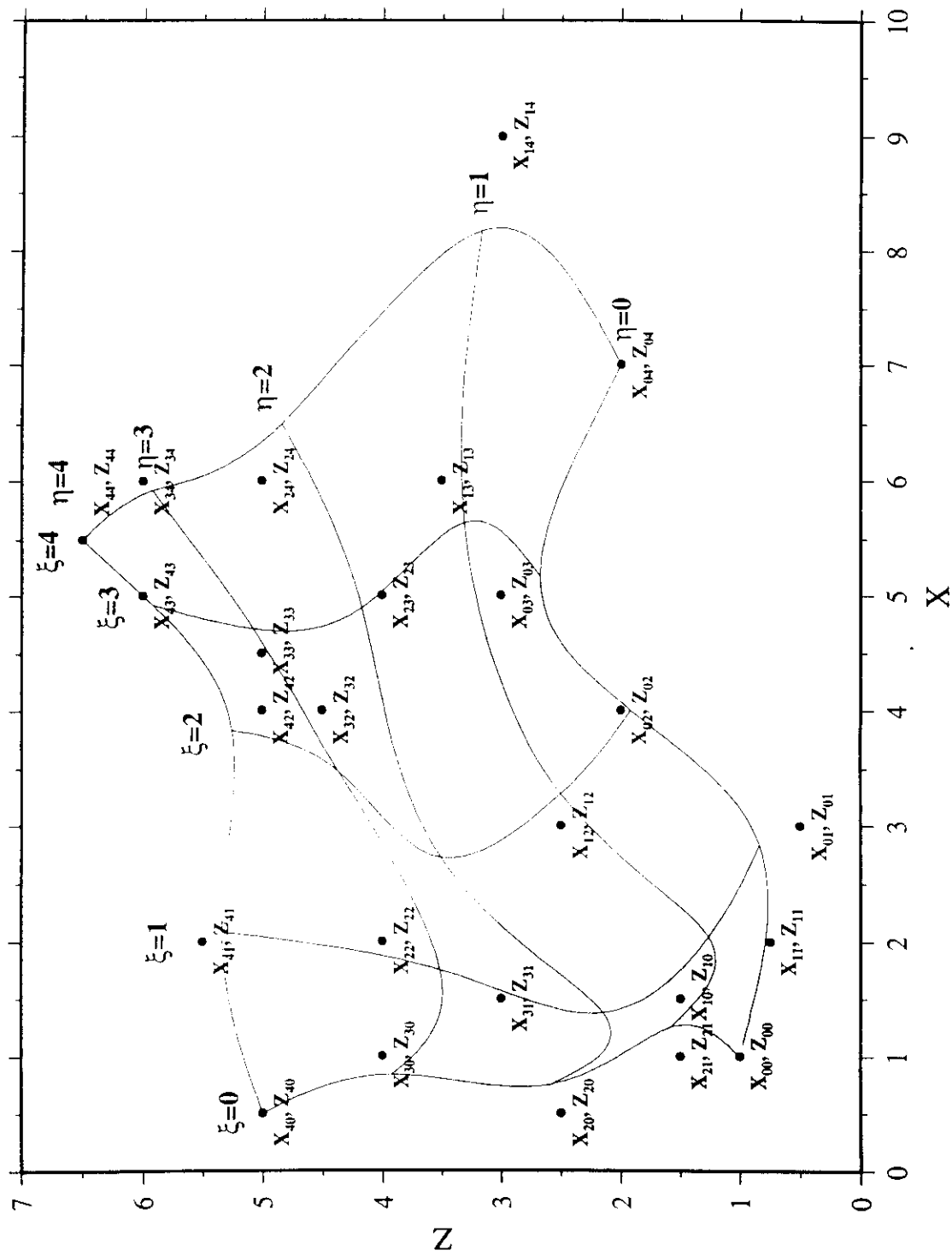
MODEL1: CROSS-HOLE RMS DATA MISFIT (ms)			
Mesh	Spacing (km)	PAR	FIXED
6x6	0.2	0.87 (13)	2.29 (6)
7x7	0.15	0.46 (18)	2.06 (6)
9x9	0.1	0.43 (15)	0.96 (10)
17x17	0.05	0.27 (13)	0.30 (12)

Table 3:

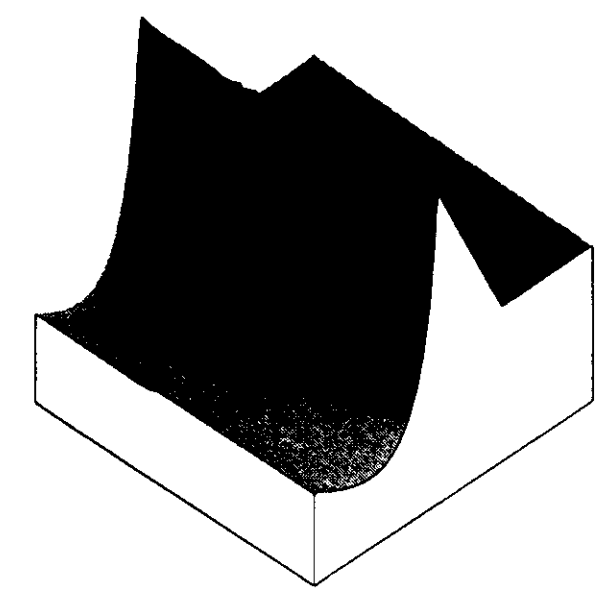
MODEL2: RMS DATA MISFIT (ms)			
Mesh	Spacing (km)	PAR	FIXED
6x10	0.2	1.18 (21)	2.05 (14)
7x12	0.15	0.83 (23)	1.78 (14)
9x17	0.1	0.81 (19)	0.80 (18)
11x21	0.08	0.71 (13)	0.80 (12)

Table 4:

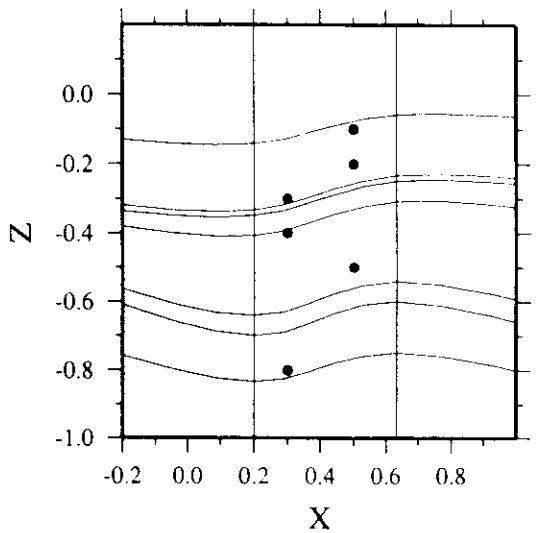
GRID POINTS AND CURVILINEAR GRID LINES

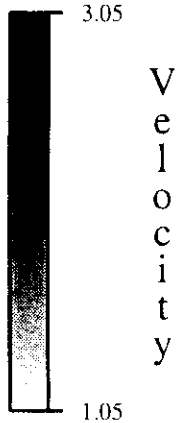


CUBIC B-SPLINES PARAMETRIC SURFACE

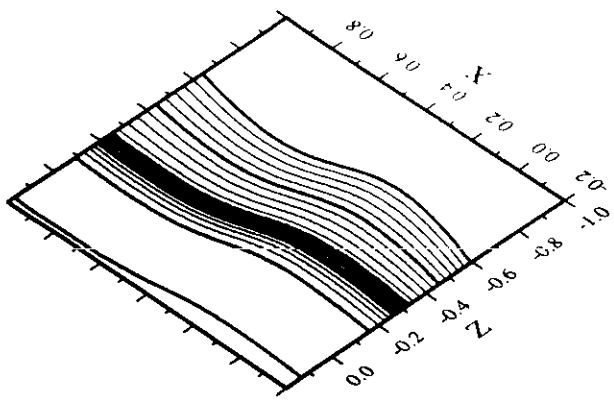


b.

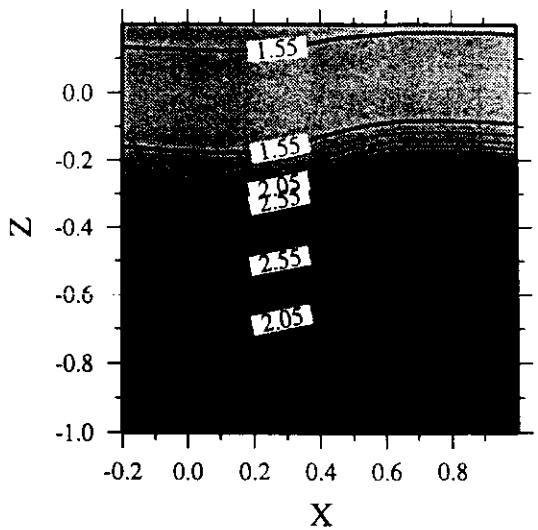




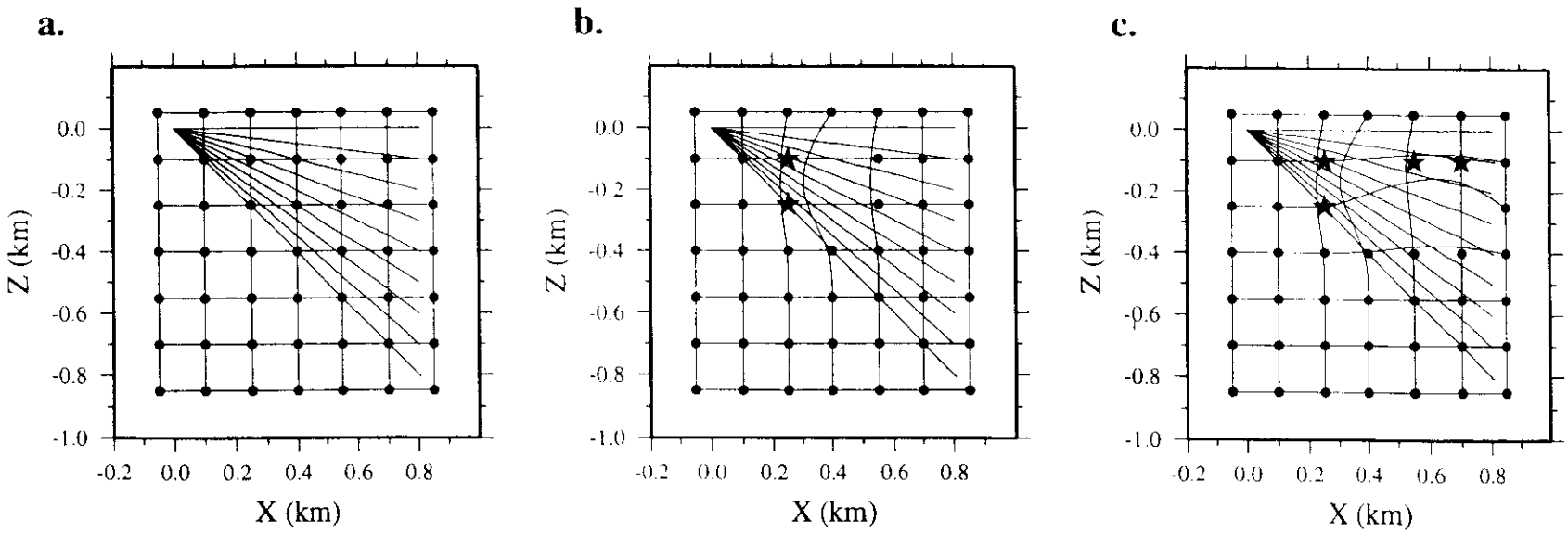
a.



c.

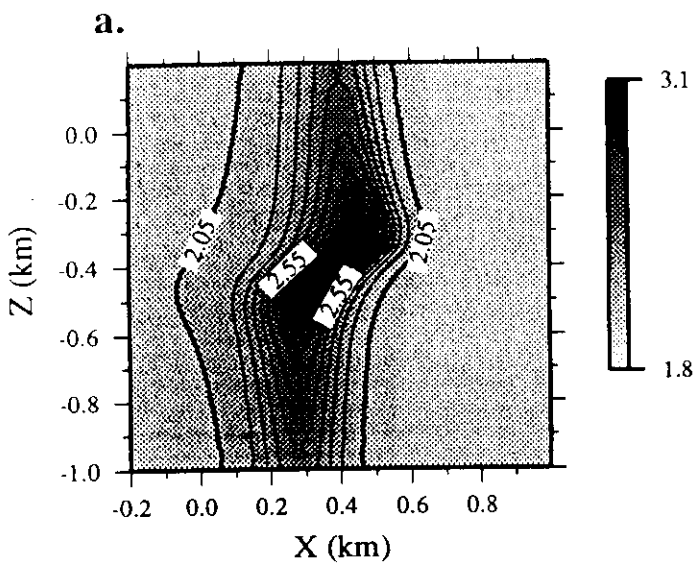


RAY-TRACING IN CURVILINEAR COORDINATES

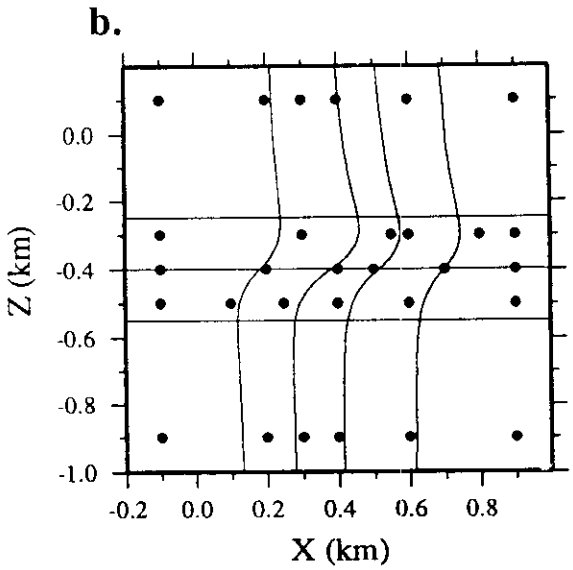


MODEL1: Settings

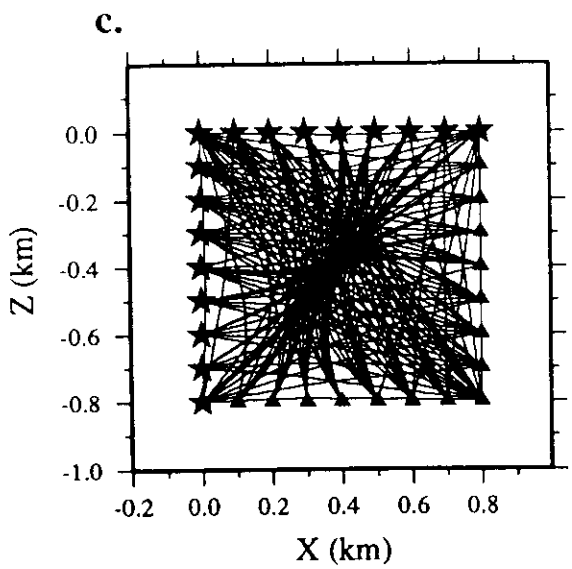
True Velocity Model



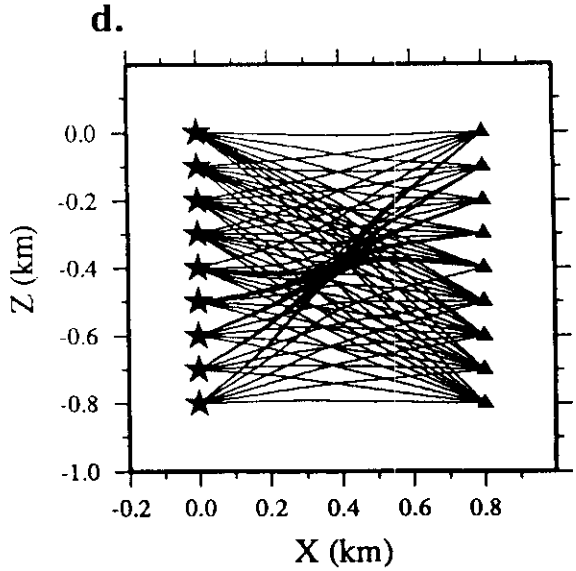
Grid points & Curvilinear grid lines



Medical Ray Coverage

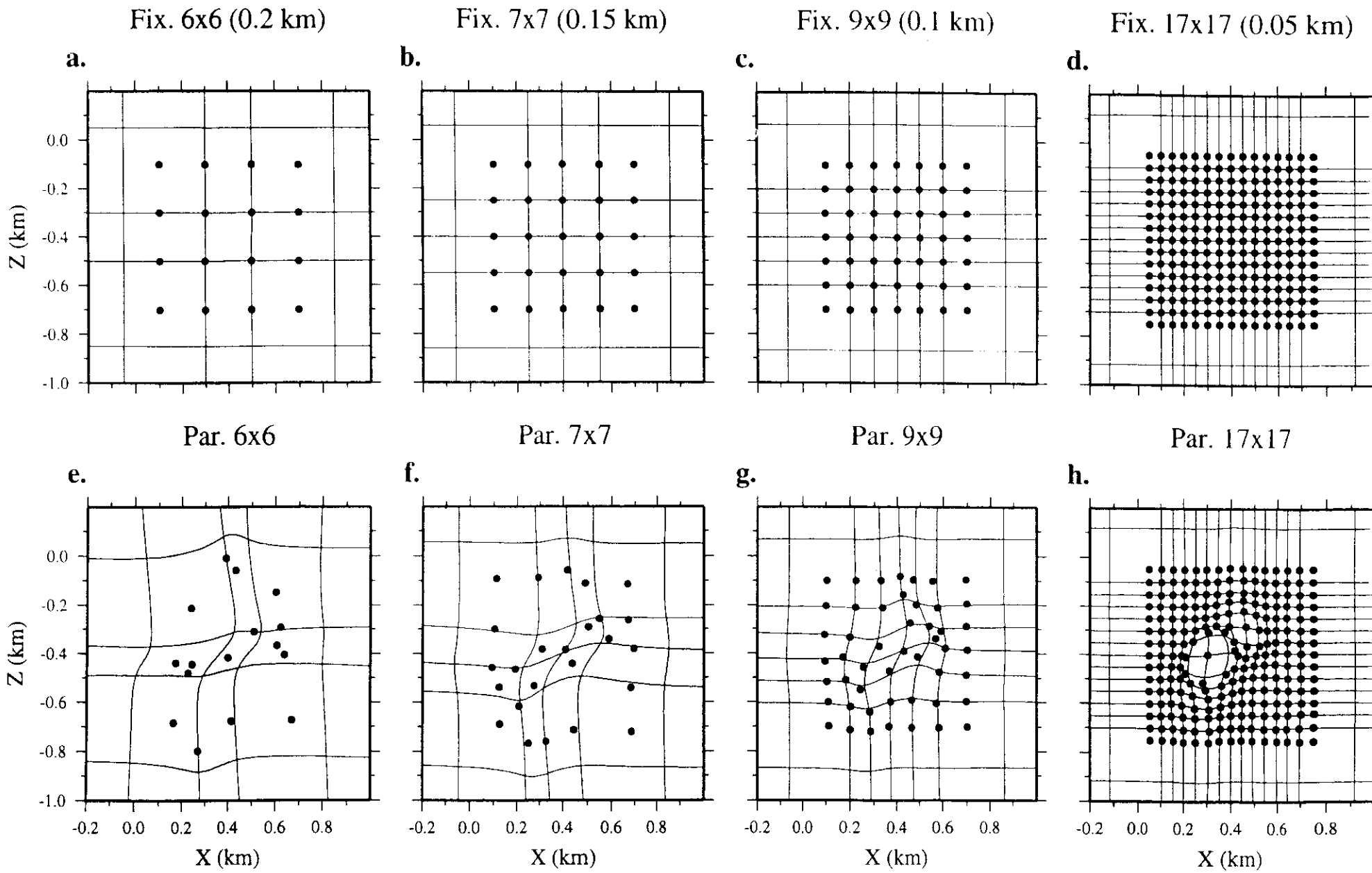


Cross-Hole Ray Coverage

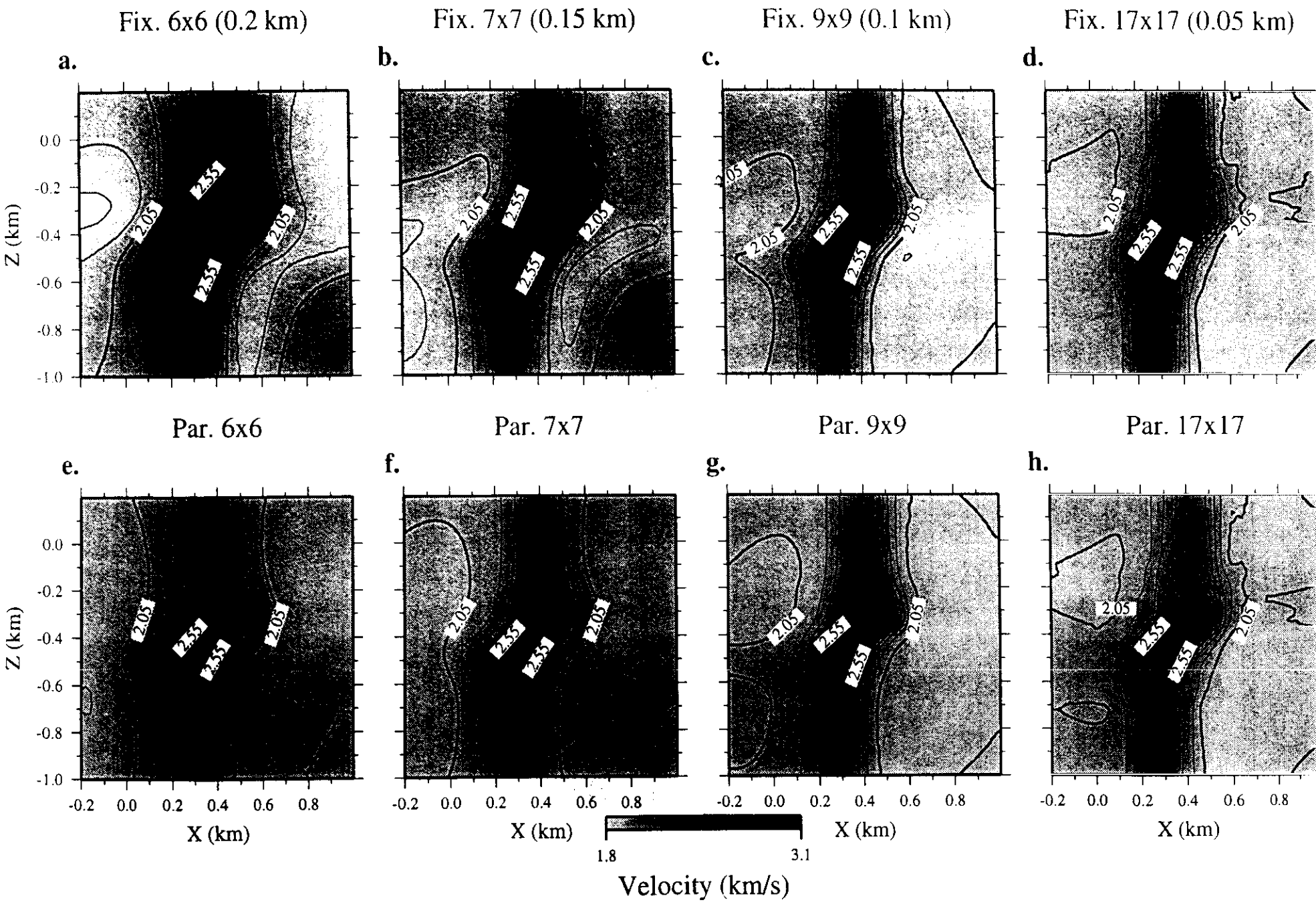


igure 4 - [illegible]

MODEL1: Grid Geometry for Medical Inversions



MODEL1: Medical Inversions



MODEL1: Medical Inversions (absolute percentual diff.)

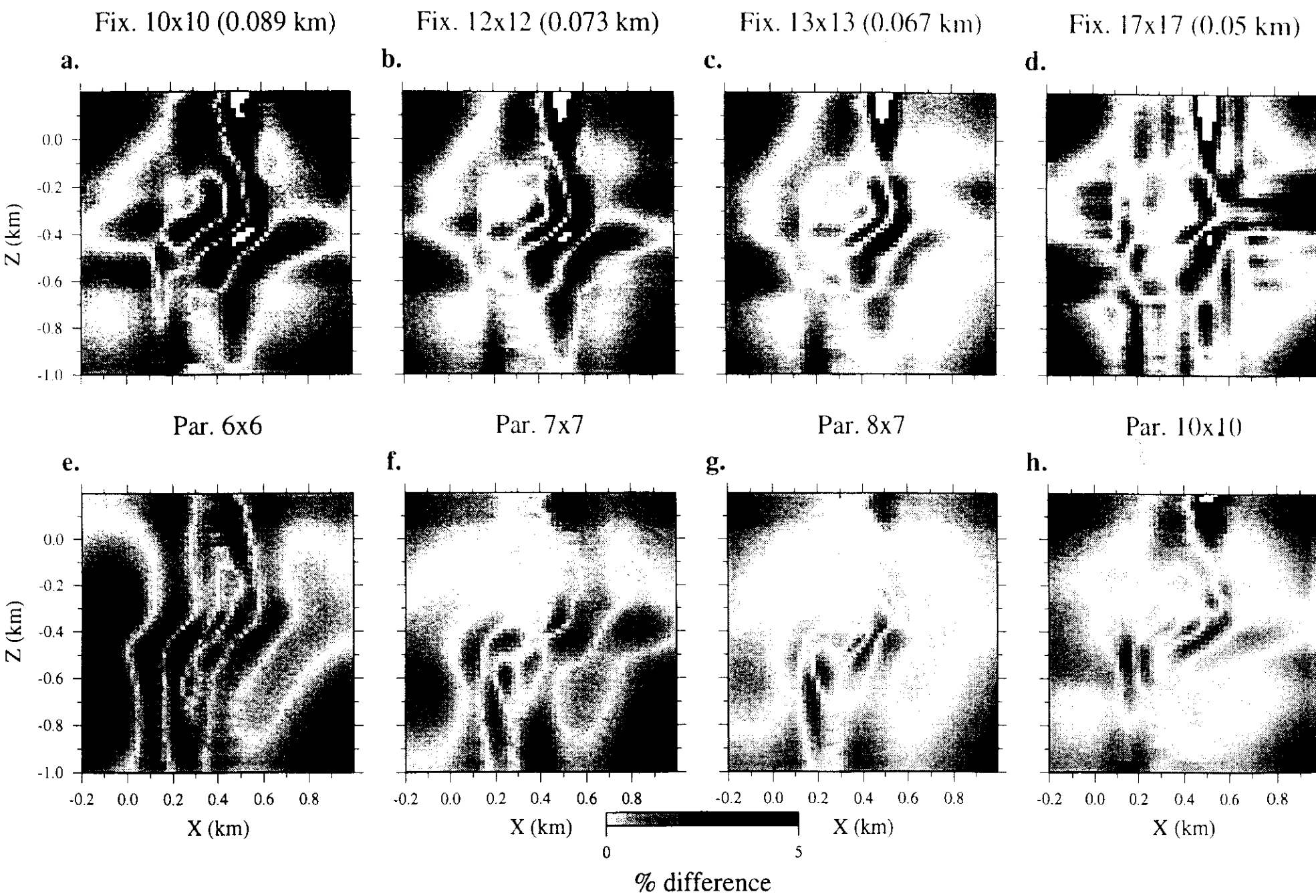
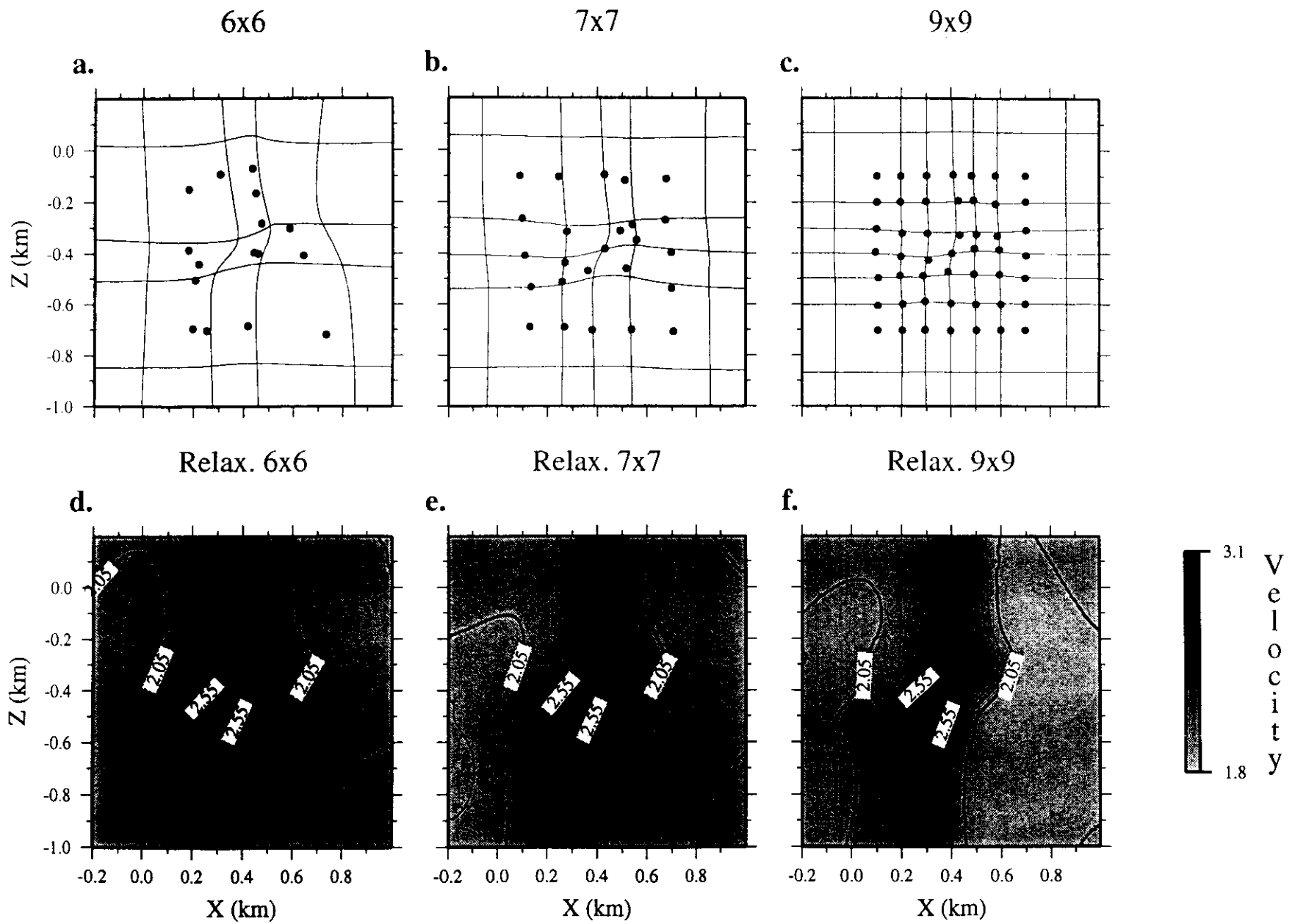


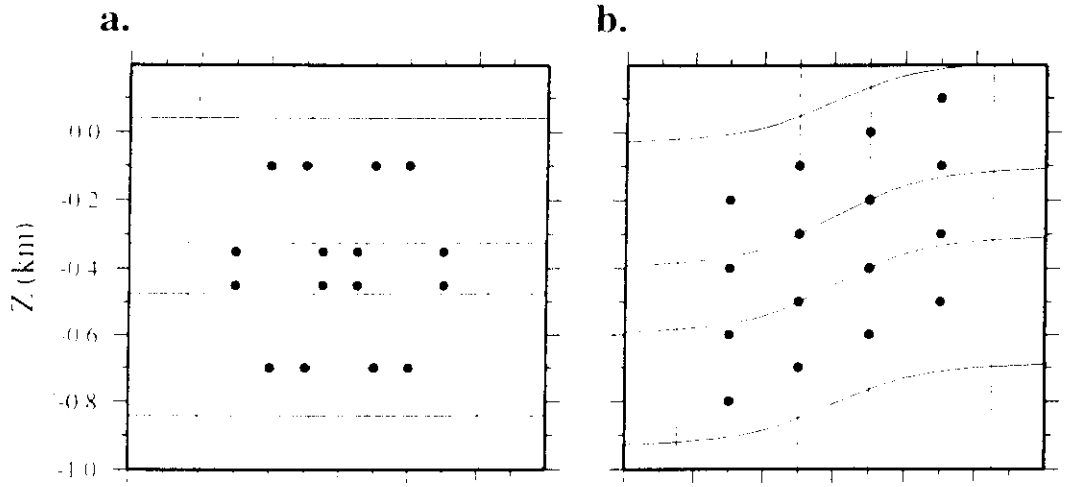
Fig. 7

MODEL1: relaxed initial fixed-grid for Medical Inversions

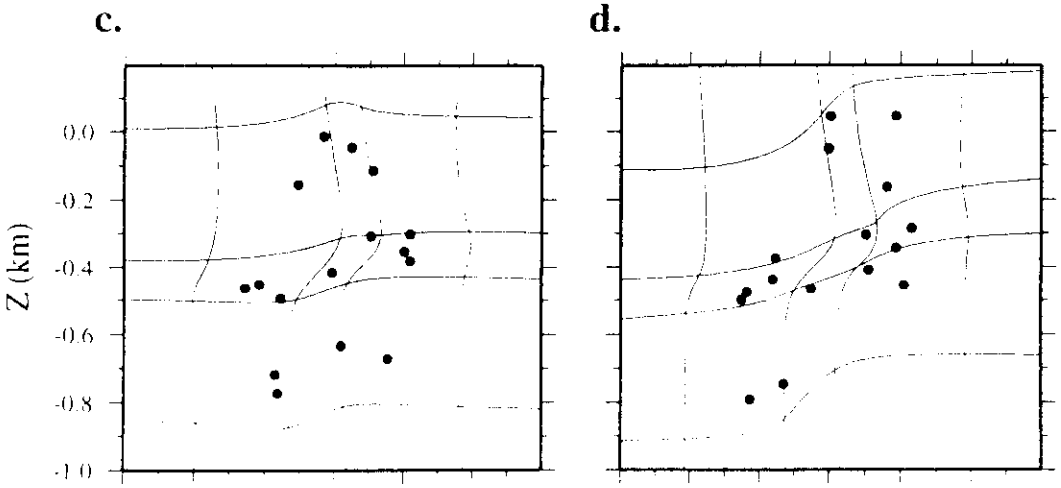


MODEL1: Medical Inversions, 6x6

INITIAL MESH



FINAL MESH



FINAL VELOCITY MODEL

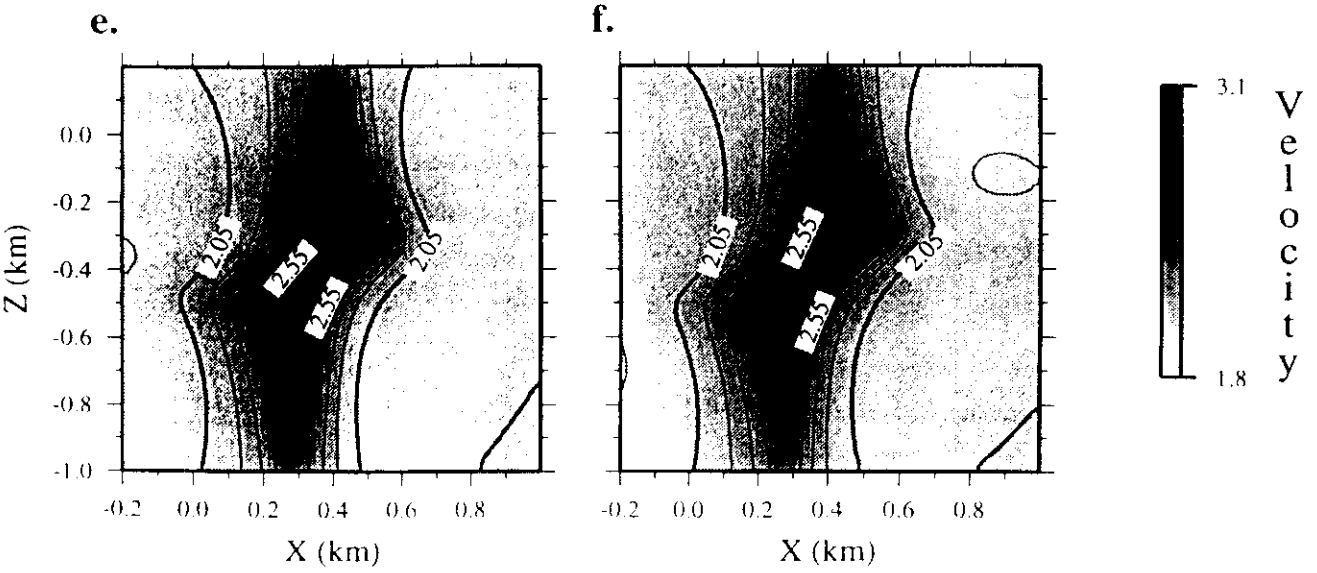
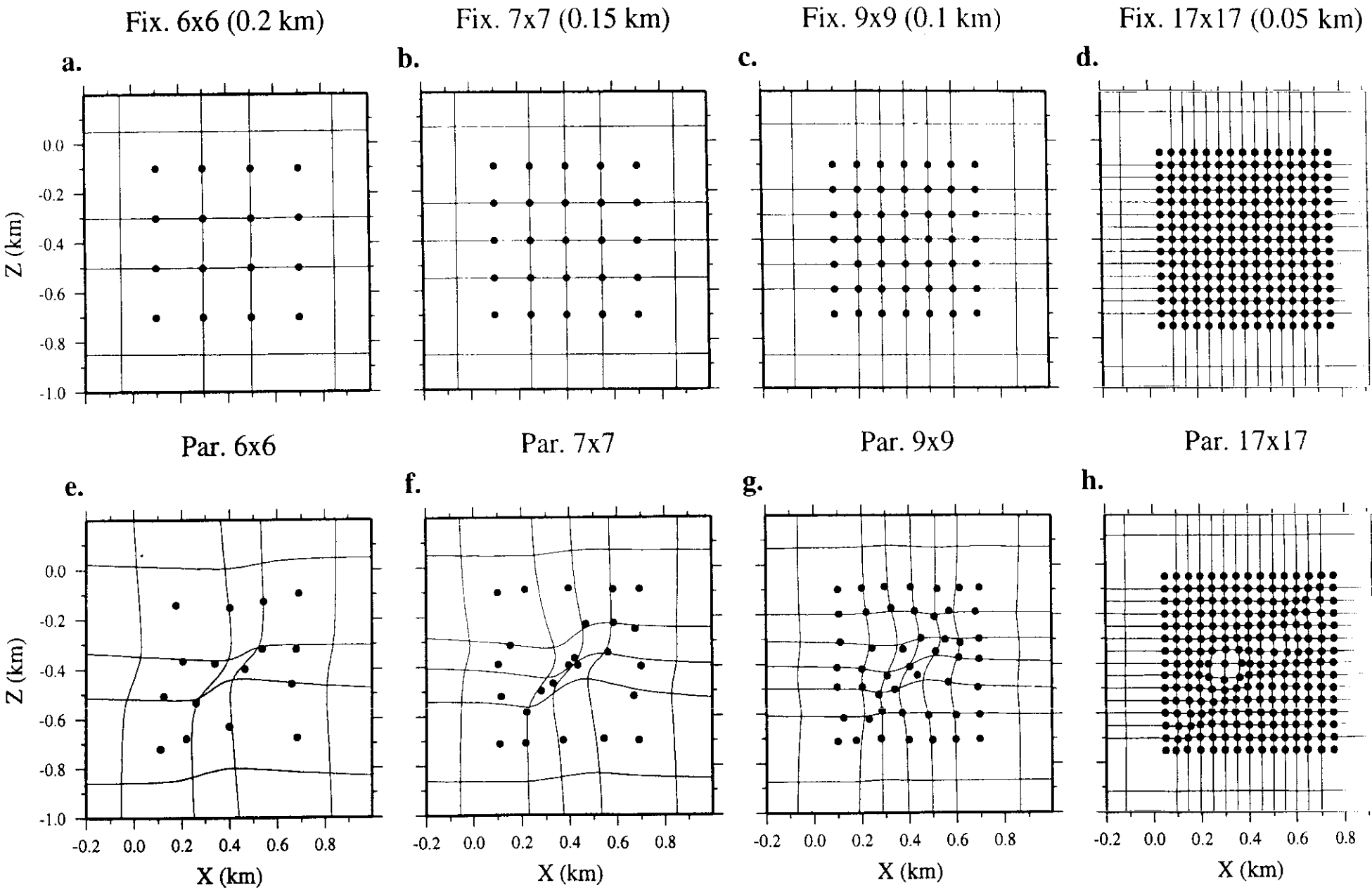
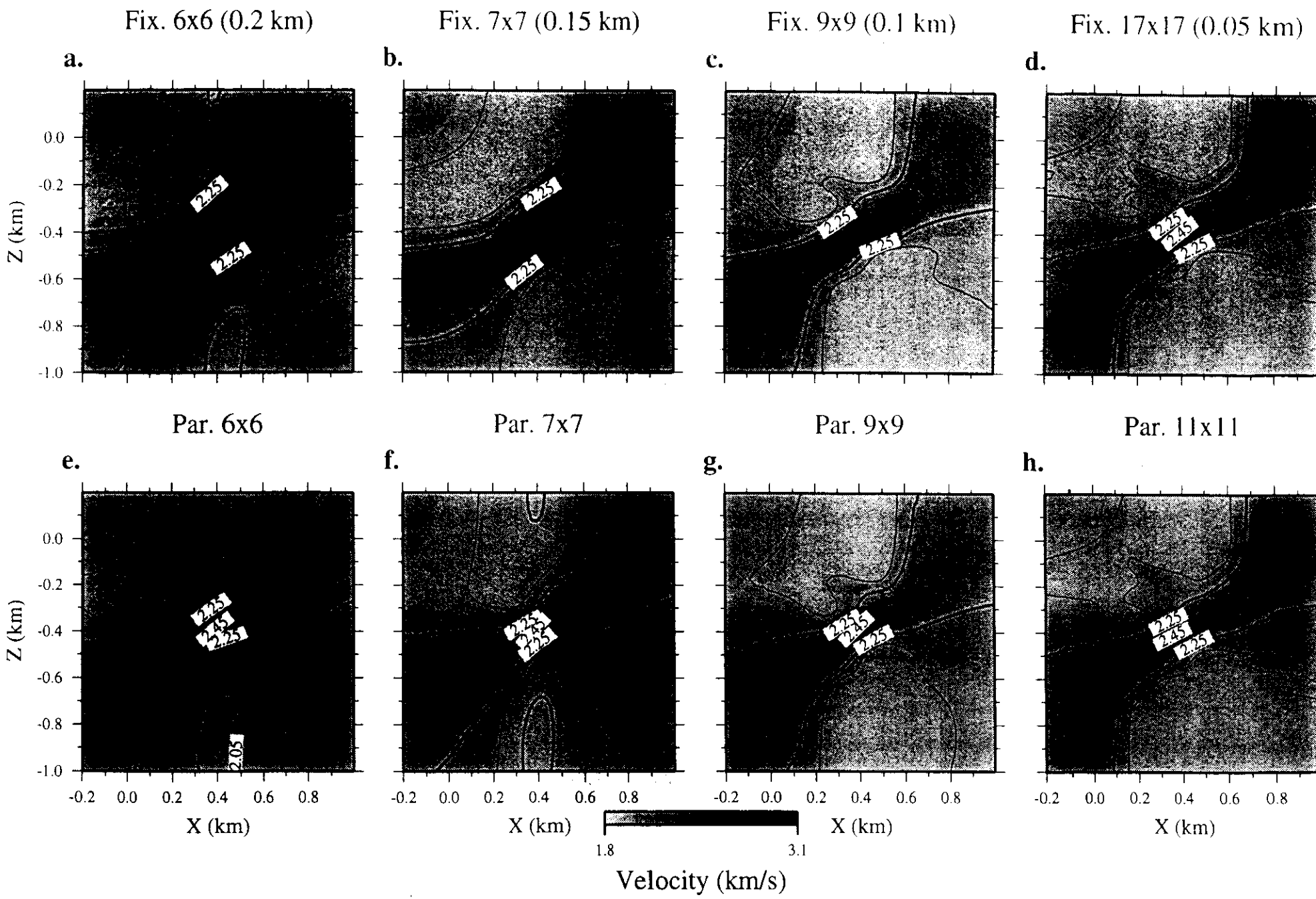


Fig 9

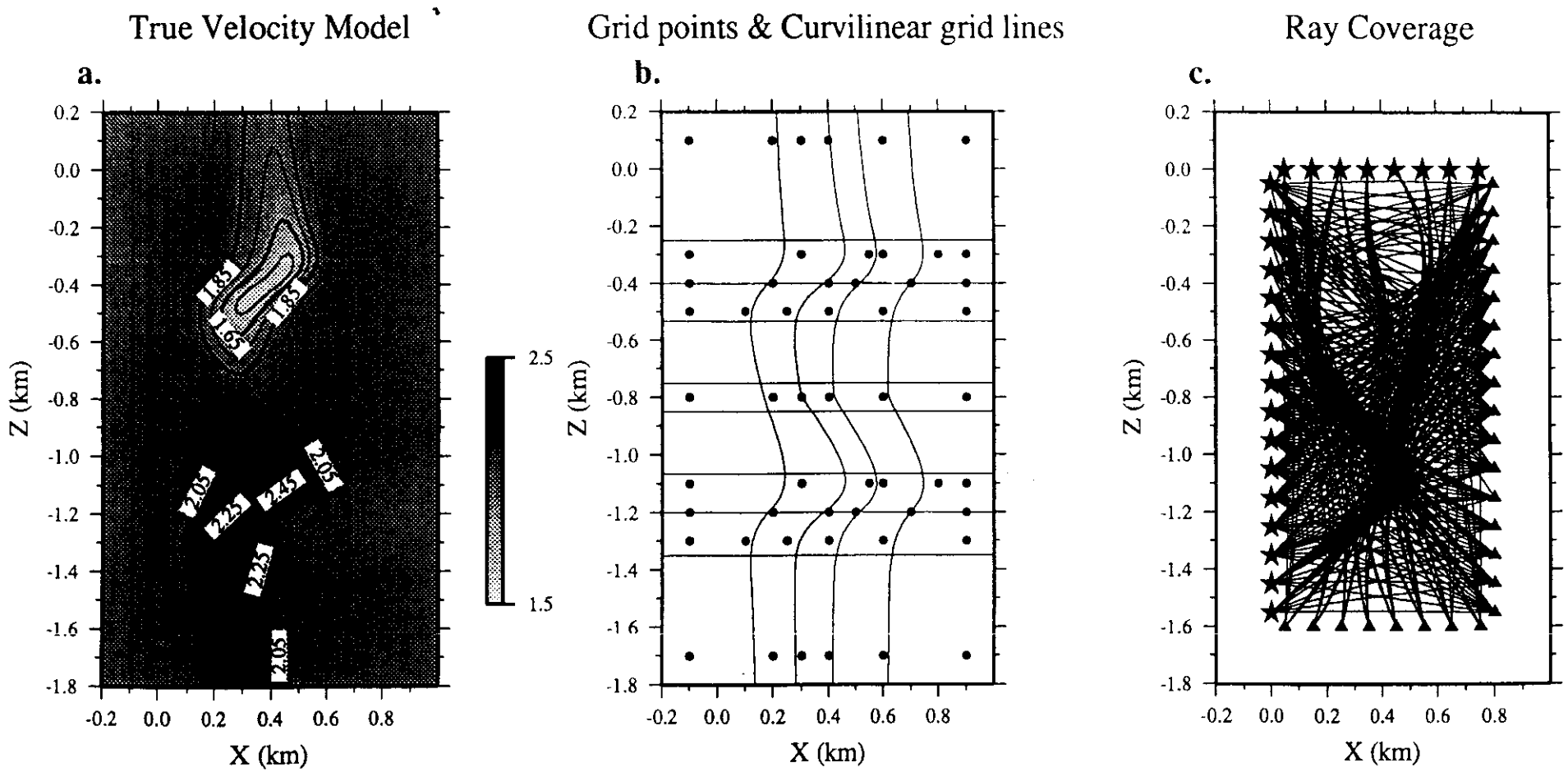
MODEL1: Grid Geometry for Cross-Hole Inversions



MODEL1: Cross-Hole Inversions



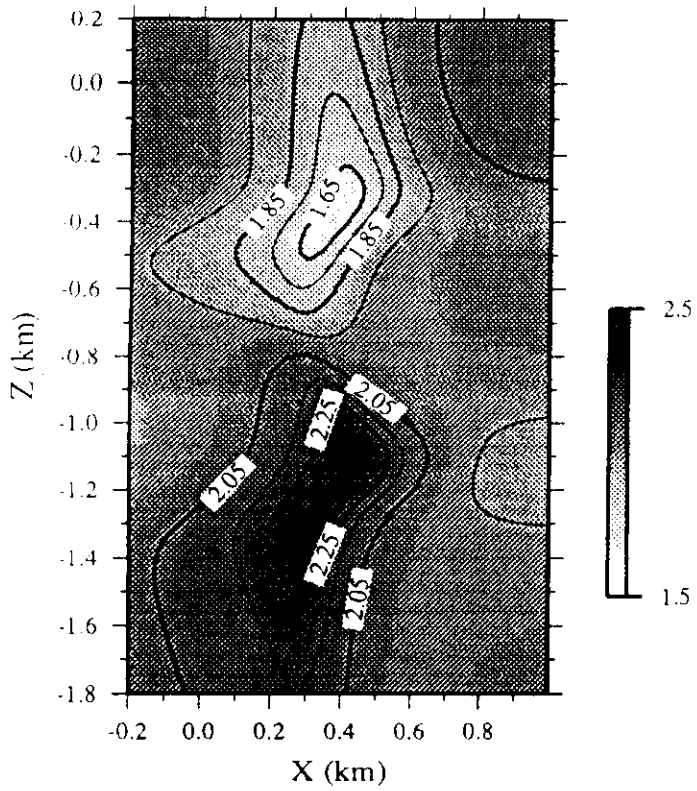
MODEL2: Settings



MODEL2: Adaptive Grid Results

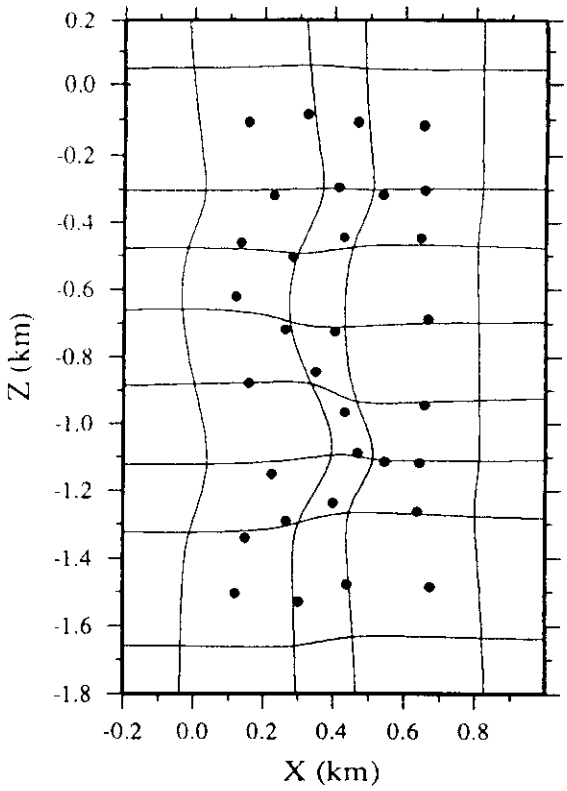
6x10 (0.2 km)

a.



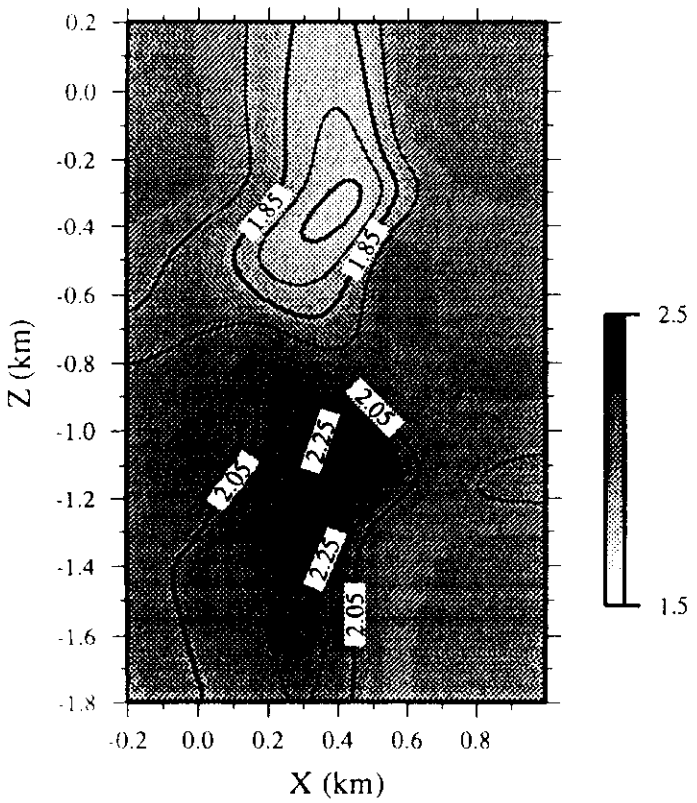
Grid points & Curvilinear grid lines

b.



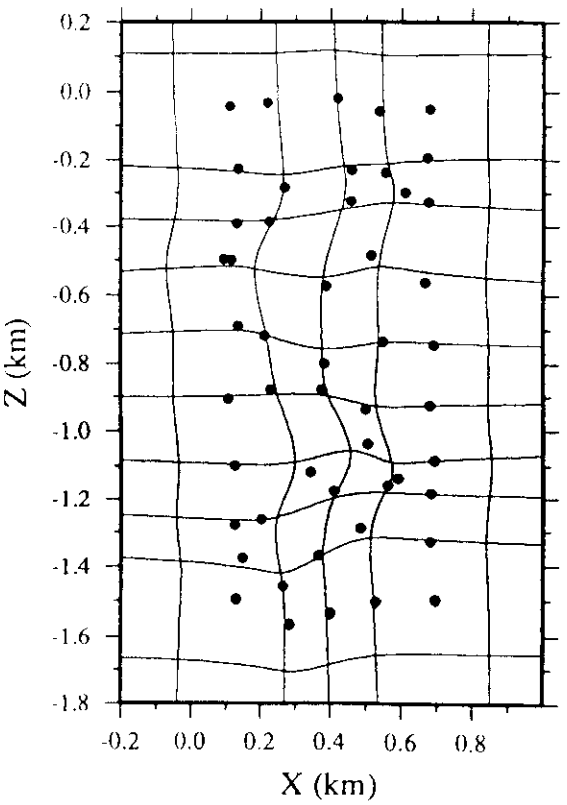
7x12 (0.15 km)

c.



Grid points & Curvilinear grid lines

d.

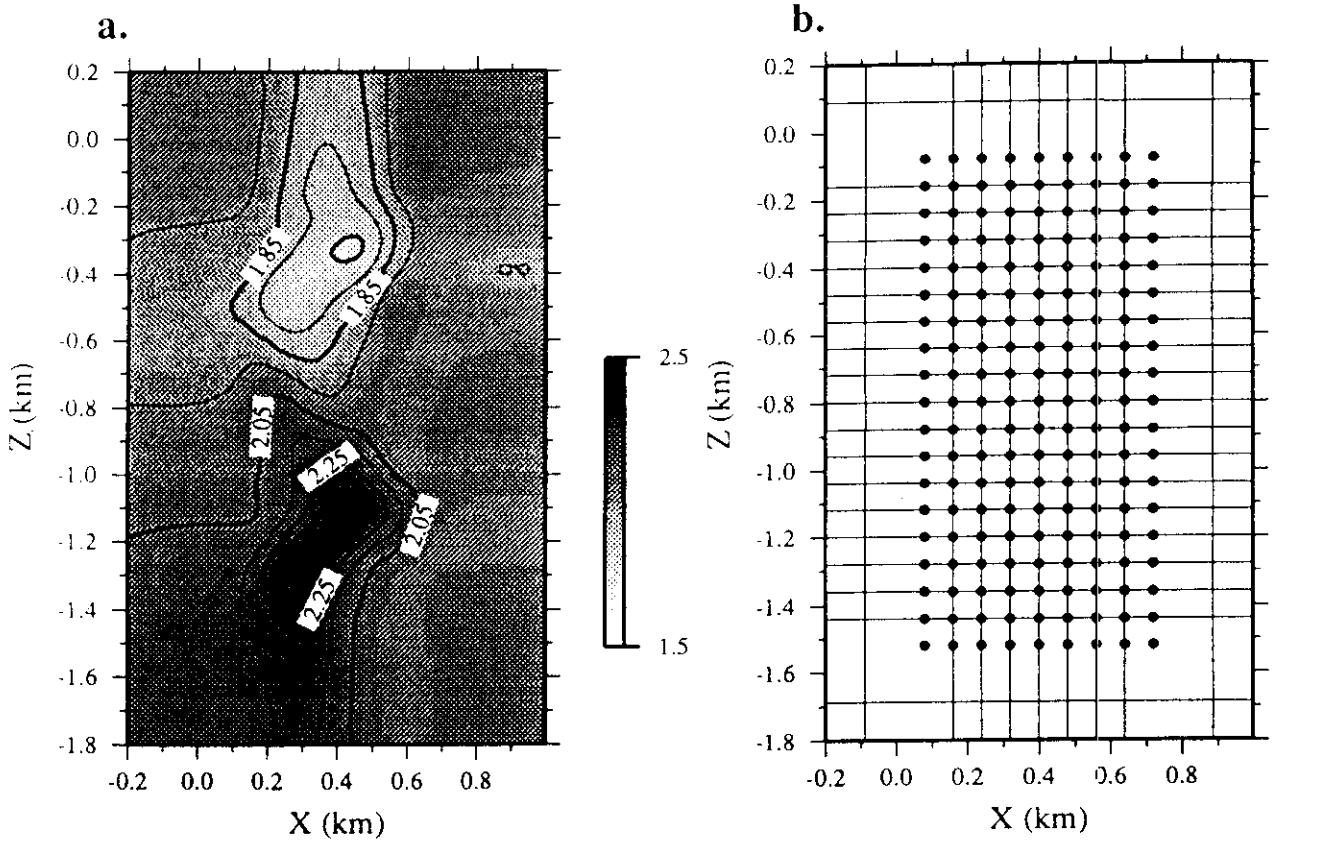


13

MODEL2: Fixed Grid Results

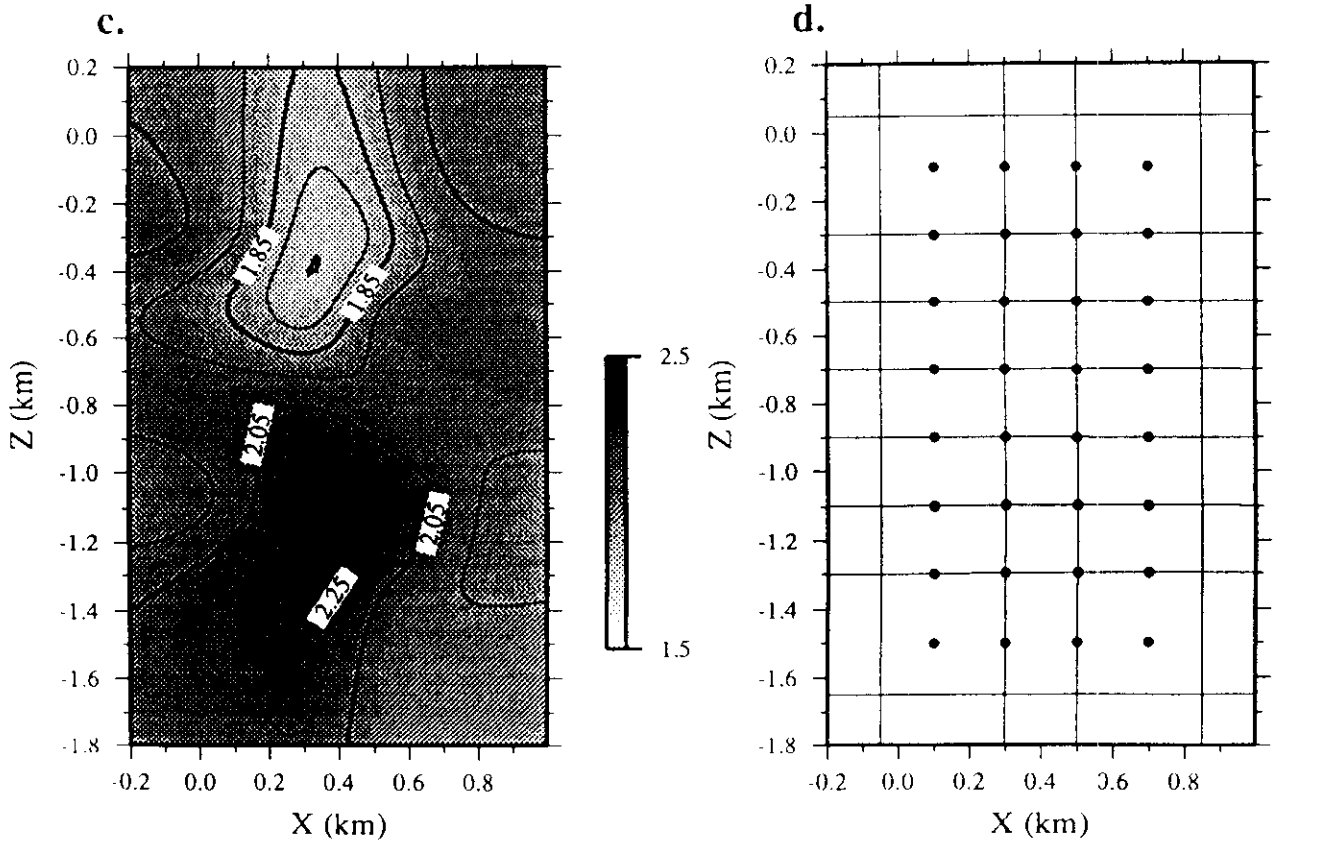
11x21 (0.08 km)

Grid points & Curvilinear grid lines



6x10 (0.20 km)

Grid points & Curvilinear grid lines



MODEL1: Medical Inversions, Noise level test