Constrained Traveltime Inversion Using Reflected Waves of Surface Shot Gathers

Jianghai Xia and Richard D. Miller, Kansas Geological Survey*

SUMMARY

We formulated an approach to estimate parameters of a layered earth model such as interval velocity, reflector depth, and dip by inverting reflected traveltime with constraints on dipping directions and depth to reflectors. We applied the "shooting method" (Lines, Bourgeois and Covey, 1984) to calculate modeled reflection time. We derived an explicit expression of the Jacobian matrix. The efficiency of computation is achieved by inverting model parameters layer by layer from shallow to deep, which is equivalent to solving a diagonal matrix. We used the Levenberg-Marquardt (L-M) method (Marquardt, 1963) and the singular value decomposition *(SVD)* technique (Lanczos, 1961; Golub and Reinsch, 1970) to solve the least-squares problem. Two kinds of criteria, *RMS* error and the maximum deviation, were introduced to keep track of the convergence of iterations to reduce the possibility of rejecting an acceptable modified model. Two error-free synthetic examples, performed on a *PC* computer, showed the approach was stable and fast.

INTRODUCTION

The traveltime inversion method is not new to geophysics since much work has been carried out to investigate the internal structure of the earth from surface data or vertical seismic profiles (Johnson and Gilbert, 1972; Crosson, 1976; Gjoystdal and Ursin, 1981; Lines, Bourgeois and Covey, 1984; and Lee, 1990). The generalized linear inversion theory is applied to traveltime inversion (Backus and Gilbert, 1970; Aki and Lee, 1976).

A layer-stripping method that estimates parameters of the shallow layer before deeper layers has been applied to invert reflection times of three-dimensional *(3-D)* surface seismic data and offset vertical seismic profile *(VSP)* data. Gjoystdal and Ursin (1981) estimated interval velocities and reflection interfaces from reflection times in *3-D* surface seismic data using this method. Lines et al. (1984) applied the method to *VSP* data to estimate dips of the layers in the vicinity of a well location with known interval velocities and reflector depths.

As opposed to the layer-stripping method, Lee (1990) developed a method to invert all layer parameters simultaneously and show that the unconstrained inversion problem is equivalent to a set of inverse problems with reduced-size matrices. Inversion of the reduced-size matrix is numerically accurate and computationally efficient.

The only way to reduce the degree of nonuniqueness of inversion is to constrain an earth model as tightly as possible. This paper implements the traveltime inversion method with constraints on dipping directions of interfaces and depth to reflectors. The constraints obviously reduce the number of possible solutions of the inverse problem.

The Jacobian matrix is calculated by explicit expressions of derivatives of traveltime with respect to the constrained layer parameters in this paper to improve computational efficiency and accuracy. We invert the parameters of a layered model, layer by layer, from shallow to deep. With this treatment in the inverse procedure, the lower-triangular matrix turns out to be a diagonal matrix equivalently. Numerically, a diagonal matrix is more stable and more easily solved than a lower-triangular matrix. The *L-M* method is employed to solve the least-squares problem. The *SVD* technique is used to speed up iterations.

METHOD

The formulation of the approach consists of five parts: a forward ray-tracing method; a constrained layered earth model; an objective function and partial derivatives of traveltime functions; a least-squares solution; and a solution by the *L-M* method and *SVD* technique.

1. Forward ray-tracing method

A ray-tracing method used in calculating modeled arrival time is termed the "shooting method" (Lines, Bourgeois and Covey, 1984). The source point and the receiver point *(x, z)* are specified for a two-dimensional *(2-O)* layered medium. A ray emerges at a given take off angle ϕ and is traced through the layered medium by using Snell's law as shown in Figure 1:

$$\sin \gamma_i / V_i = \text{constant}, \quad (i=1,2,\ldots,N), \tag{1}$$

where γ_i is the angle of incidence and *Vi* is the seismic wave velocity for the ith layer. The ray is reflected by a point (X_N, Z_N) then back to the surface at some point (x_0, z). The difference between the receiver location *(x, z)* and the intersection point of the ray with the surface (x_0, z) is computed to give a value of $\Delta x = x - x_0$. If the absolute value of Δx is not less than some specified value, then ϕ is increased by $\Delta\phi$ and ray tracing is recomputed until $\Delta x < \varepsilon$, where ε is some specified tolerance for the ray intersection with the surface. This value of ϕ is used as an initial guess for ravs traveling to an adjacent receiver. For each ray, the traveltime through each layer is computed and accumulated to give a total traveltime for the kth receiver. For a ray passing through *N* layers (see Figure l), the traveltime for the kth receiver can be written as

$$T_N^k = \sum_{i=1}^{N} \left(s_i Q_{N,i}^k + s_i \tilde{Q}_{N,i}^k \right), \tag{2}$$

where s_i is the slowness of the ith medium,

$$Q_{N,i}^k = \left[\left(\Delta X_{N,i}^k \right)^2 + \left(\Delta Z_{N,i}^k \right)^2 \right]^{1/2},$$

$$\Delta X_{N,i}^k = X_{N,i}^k - X_{N,i-1}^k, \quad \Delta Z_{N,i}^k = Z_{N,i}^k - Z_{N,i-1}^k,$$

$\left(X_{N,i}^k, Z_{N,i}^k \right)$ are the coordinates of the intersection of the down-going ray with ith interface which will be reflected by the Nth interface and received by the kth receiver, $Q_{N,i}^k$ represents the length of the ray segment between the *(i-1)* and ith interface (of medium *i)* for the ray reflected from Nth interface and received at the kth receiver, and tilde indicates the quantity corresponding to the upgoing raypath, denoted as *U* below for notational simplicity. We are using mathematic notation as close as possible to Lee's (1990).

For the layered earth model, the dipping interface is defined using slope a_i and intercept depth b_i by

$$Z_{N,i}^k = a_i X_{N,i}^k + b_i. \tag{3}$$

In a given layered earth model (known the slowness s_i, slope a_i and intercept b_i), the modeled traveltime can be calculated by equations (1) - (3).

2. Constrained layered earth model

In the real world, some information about the dipping directions and the intercept depth of layers may be available from geology or other geophysical data. This information should be used as constraints in the inversion procedure. In this paper, two kinds of constraints are considered in the inversion. First, the sign of the dipping direction of interface is kept the same during iterations, and second, the intercept depth of a shallow layer is smaller than that of the deeper layer, which is $b_i < b_{i+1}$. To use the *L-M* method to solve the constrained problem,- we redefine equation (3) as

$$Z_{N,i}^k = Sign(dip) q_i^2 X_{N,i}^k + \left(d_i^2 + L_{i-1} \right), \tag{4}$$

where *Sign(dip)* is 1 for a dipping downward interface, -1 for dipping upward, and 0 for horizontal, q_i^2 is the absolute value of the slope of the ith interface, d_i^2 is the intercept thickness of the ith layer, and L_{i-1} is the intercept depth to the $(i\text{-}1)$th interface. The parameters to solve are s_i, q_i and d_i rather than s_i, a_i and b_i. The sign of the dip will remain unchanged while solving for layer parameters. The constraint on the intercept depth is satisfied in the inverse procedure when equation (4) is used. The constrained inverse problem also becomes an unconstrained problem if interfaces are defined by equation (4).

3. Objective function and partial derivatives of traveltime function

We define the objective function in the inverse procedure as

$$E=\sum_{i=1}^{N}\sum_{k=1}^{M}\left[O_i^k-T_i^k\left(\vec{P}^{n+1}\right)\right]^2, \qquad (5)$$

where M is the number of receivers, N is the number of interfaces, O_i^k is the observed traveltime from the ith interface received by the kth receiver, T_i^k is the modeled traveltime based on equation (2), and $\vec{P}^{n+1}$ is the $(n+1)$th estimation of the parameter vector with dimension L $(=3N)$.

The linear approximation of the nonlinear traveltime function (equation 2) is performed by a Taylor-series expansion. Thus, the traveltime function could be approximated as

$$T_i^k\left(\vec{P}^{n+1}\right)=T_i^k\left(\vec{P}^{n}\right)+\sum_{j=1}^{L}\left(\frac{\partial T_i^k}{\partial P_j}\right)_{\vec{P}=\vec{P}^n}\delta P_j, \qquad (6)$$

where δP_j is the jth component of the modification vector $\delta\vec{P}$, which will be added to the nth estimation of the parameter vector $\vec{P}^n$ to form the $(n+1)$th estimation $\vec{P}^{n+1}$, which is defined by

$$P_j^{n+1}=P_j^n+\delta P_j, \qquad (j=1,2,\ldots,L). \qquad (7)$$

Substituting T_i^k ($\vec{P}^{n+1}$) in equation (5) by equation (6), the solution of δP_j is given by minimizing the objective function E.

For the constrained' layer model, we derive the partial derivatives of the traveltime function as follows:

$$\frac{\partial T_i^k}{\partial s_j}=Q_{i,j}^k+U,$$

$$\frac{\partial T_i^k}{\partial q_j}=\left[s_j\left(Q_{i,j}^k\right)^{-1}\Delta Z_{i,j}^k X_{i,j}^k-s_{j+1}\left(Q_{i,j+1}^k\right)^{-1}\Delta Z_{i,j+1}^k X_{i,j}^k\right]2Sign\left(dip_j\right.$$

$$\frac{\partial T_i^k}{\partial d_j}=2s_j\left(Q_{i,j}^k\right)^{-1}\Delta Z_{i,j}^k d_j+U. \qquad (8)$$

4. Least-squares solution for $\delta\vec{P}$

Let equation (6) be equal to the observations $\vec{O}$, and written in matrix form, it turns out to be

$$G\delta\vec{P}=\left(\vec{O}-\vec{T}^n\right)=\delta\vec{m}, \qquad (9)$$

where G is the Jacobian matrix with NM ($=N\times M$) rows and $L(=3N)$ columns, $\vec{O}$ is the NM dimensional column vector of the observations, $\vec{T}^n$ is the NM dimensional column vector of the modeled traveltime at the nth iteration, and $\delta\vec{P}$ is the modification vector with dimension L and will be added to the parameter vector $\vec{P}^n$. The least-square solution for $\delta\vec{P}$, which should reduce the objective function E, can be written as

$$\delta\vec{P}=\left(G^TG\right)^{-1}G^T\left(\vec{O}-\vec{T}^n\right), \qquad (10)$$

where the matrix G^T is the transposed matrix of G, and $\left(G^TG\right)^{-1}$ is the inverse of the matrix G^TG.

Lee (1990) points out that G is in a lower-triangular form for the unconstrained layer model. According to equation (4), the property that G is in lower-triangular form still holds in the constrained problem. G can also be written as

$$G=\begin{bmatrix} G_{11} & & & 0 \\ G_{12} & G_{22} & & \\ \vdots & \vdots & & \\ G_{1N} & G_{2N} & \cdots & G_{NN} \end{bmatrix}, \qquad (11)$$

where G_{ji} are themselves matrices defined as

$$G_{ji}=\begin{bmatrix} \dfrac{\partial T_i^1}{\partial s_j} & \dfrac{\partial T_i^1}{\partial q_j} & \dfrac{\partial T_i^1}{\partial d_j} \\ & \vdots & \\ \dfrac{\partial T_i^M}{\partial s_j} & \dfrac{\partial T_i^M}{\partial q_j} & \dfrac{\partial T_i^M}{\partial d_j} \end{bmatrix}.$$

G_{ji} consists of the partial derivatives of the traveltime reflected from the ith interface, which is defined by equation (2), with respect to the jth layer parameters. The constrained inverse problem, therefore, is equivalent to a set of inverse problems with reduced-size matrices, which could be expected to balance the complication caused by transforming the constrained problem into an unconstrained problem.

Equation (10) can be written as follows:

$$\delta\vec{P}_1=G_{11}^{-1}\delta\vec{m}_1,$$

$$\delta\vec{P}_2=G_{22}^{-1}\left(\delta\vec{m}_2-G_{12}\delta\vec{P}_1\right), \qquad (12)$$

$$\delta\vec{P}_N=G_{NN}^{-1}\left(\delta\vec{m}_N-\sum_{i=1}^{N-1}G_{iN}\delta\vec{P}_i\right),$$

where $\delta\vec{P}_i$ is a three-component column vector that holds the modification to the parameters of the ith layer,

$$\delta\vec{P}_i=Col\left(\delta s_i,\delta q_i,\delta d_i\right),$$

and $\delta\vec{m}_i$ is an M dimensional vector that holds the deviations between the observed and calculated traveltime from the ith layer,

$$\delta\vec{m}_i=Col\left(O_i^1-T_i^1,O_i^2-T_i^2,\cdots,O_i^M-T_i^M\right).$$

$\delta\vec{P}$ will be added to $\vec{P}^n$ (equation 7) to obtain the $(n+1)$th estimate of the layer parameters. Equation (12) allows us to estimate parameters layer by layer.

We find an interesting fact about equation (12). If the initial parameter estimates for layer one happen to be the actual correct values for the layer, the modifications for layer one, $\delta\vec{P}_1$, will be zero. In this case the modifications for layer two are only dependent on the matrix G_{22} ; if the initial estimates for layer one and layer two happen to be the actual correct values of the parameters, the modifications for layer three are only dependent on the matrix G_{33} ; and so on. This fact suggests that if the parameters are estimated layer by layer from shallow to deep, the Jacobian matrix (11) is reduced to a diagonal matrix equivalent. This property makes estimating parameters of the layered model very easy and fast.

The least-squares solution is usually unstable, especially when data contain errors. The *L-M* method and *SVD* technique can be used to overcome the problem of numerical instability.

5. The *L-M* method and the *SVD* technique

It is well known that the *L-M* method produces quadratic convergence with the advantage of numerical stability in iterations compared to the least-squares method. The *L-M* method is chosen here to calculate the general inverse of G_{11}^{-1}, G_{22}^{-1}, $\cdots$, G_{NN}^{-1}, which can be shown as

$$G_{ii}^{-1}=\left(G_{ii}^TG_{ii}+\theta_i^2I\right)^{-1}G_{ii}^T, \qquad (i=1,2,\ldots,N), \qquad (13)$$

where I is the unit matrix, and θ_i^2 is called as a damping factor. Actually, the direction determined by the *L-M* method is the interpolation between the directions determined by the steepest descent method and the least-squares method.

To stabilize the iterative process the *SVD* technique, which is considered to have a high numerical stability, is introduced to find the general inverse of G_{ii}. Matrix G_{ii} in equation (13) can be decomposed by the *SVD* technique to the three component matrices. Combining the properties of each of the component parts of the *SVD*, the general inverse of G_{ii} determined by the L-M method has the form of

$$G_{ii}^{-1} = V_{ii}\left\{\left(\Lambda_{ii}^2 + \theta_i^2 I\right)^{-1}\Lambda_{ii}\right\}U_{ii}^T, \qquad (i=1,2,...,N). \qquad (14)$$

If equation (13) is used during the iterative process, the program calculates each inverse matrix of $G_{ii}^T G_{ii} + \theta_i^2 I$ for each damping factor θ_i^2. This means that each successful iteration $(E(n+1)<E(n)$, n is the iteration number) may require more than one calculation of the general inverse of G_{ii}. Equation (14) shows that *SVD* of matrix G_{ii} is needed only once for each successful iteration. A lot of computational time can undoubtedly be saved.

COMPUTATIONAL CONSIDERATIONS

THE DAMPING FACTOR. First, fast process speed is expected, so the small damping factor, say, $\theta^2 = 10^{-5}$ (the subscript i in the damping factor is omitted for notational simplicity because damping factors will be same for all G_{ii}) will be chosen. If the objective function E is reduced, a smaller damping factor will be chosen ($\theta_{n+1}^2 = \theta_n^2 / 10.0$, n is the iteration number), which makes the next step reduction of the objective function as much as possible. In this situation, the direction of the vector $\delta\vec{P}$ is closer to that determined by the least-squares method; otherwise, a bigger damping factor would be chosen ($\theta_{n+1}^2 = \theta_n^2 \times 10.0$), which may reduce the objective function. In this case, the direction of the vector $\delta\vec{P}$ is closer to that determined by the steepest descent method. When the damping factor is equal to 10^5, the program treats the point (in parameter space) as the minimum point or the best estimate of the parameters and the iterative process will automatically be terminated if the objective function is not reduced (which is a criterion to stop the program other than when the objective function E reaches some tolerance).

CRITERIA OF CONVERGENCE. The traveltime function (2) is actually a function of the model parameters (L = 3N) and the intersections (X_i, Z_i) between rays and interfaces (see Figure 1). The intersections are determined by Snell's law after the model parameters are given. For inverse procedures, a hidden assumption requires that the intersections not be changed too much after model parameters are updated. This assumption is valid only when the model parameters are near a minimum point of the objective function (5); otherwise, the objective function may not be reduced even though the direction of modification vector reaches the negative gradient. To make the inverse procedure smooth by controling it, we introduce another criterion, the maximum deviation between the observation and modeled traveltime. Therefore, successful iterations require the rms error *rms* $E = \frac{1}{M}\sqrt{\sum_{k=1}^{M}\left[O_i^k - T_i^k\right]^2}$, $(i=1,2,...N)$, or the maximum deviation, $MAXD = \max_{1\le k\ge M} | O_i^k - T_i^k |$, to be reduced.

SYNTHETIC EXAMPLES

The first example is from Lee (1990). Table 1 shows the model parameters, the initial estimates, and the final estimates. Figure 2 shows the traveltime calculated from these models. Twenty receivers are evenly spread 25 m apart. The maximum offset is 500 m. The initial rms error and the maximum deviation are 1,625 ms (milliseconds) and 2,215 ms, respectively. They are reduced to 0.05 ms and 0.11 ms after 21 iterations, respectively. It takes only 8.45 seconds of CPU time on a 486-based *PC*. Both final estimates of the model parameters and the final traveltime curve exactly match the actual model parameters and the true traveltime curve, respectively.

Table 2 shows the parameters of the second example. Figure 3 shows the traveltime calculated from these models. Twenty receivers are evenly spread 25 m apart. The maximum offset is 500 m. The initial rms error and the maximum deviation are 160 ms and 289 ms, respectively. They are reduced to 0.003 ms and 0.004 ms after 16 iterations, respectively. It takes only 6.09 seconds of CPU time on the same *PC*.

The final estimates are dependent on the initial estimates. We tried different initial estimates for the second example. Most of them converged to the true model parameters. To invert real data one may try different initial estimates, then compare the final estimates and the value of the objective function and determine the most reasonable final estimates.

CONCLUSIONS

We have presented an approach to invert traveltime of reflected waves by the L-M method and the *SVD* technique based on the constrained layered model. The approach reduces the degree of the nonuniqueness and makes solutions compatible with the geologic constraints. But the approach does not guarantee that the globe minimum point of the objective function can be determined. To invert real data, one may try different initial estimates to determine the most reasonable final estimates.

The stability and efficiency of the approach are shown by two synthetic examples. The efficiency of computation is achieved by inverting model parameters layer by layer from shallow to deep, which is equivalent to solving a diagonal matrix.

ACKNOWLEDGEMENTS

The authors appreciate the efforts of Mary Brohammer in manuscript preparation.

REFERENCES

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

Backus, G., and Gilbert, F., 1970, Uniqueness in the inversion of inaccurate gross Earth data: Phil. Trans. Roy. Soc., London, **266a,** 123-192.

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

Gjoystdal, H., and Ursin, B., 1981, Inversion of reflection times in the three-dimensions: Geophysics, 46, 972-983.

Golub, G.H., and Reinsch, C., 1970, Singular value decomposition and least-square solution, Num, Math., 14, p. 403-420.

Johnson, L.E., and Gilbert, F., 1972, Inversion and interference for teleseismic ray data: Methods in computational physics, New York, Academic Press. v. 12, 231-266.

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

Lee, M.W., 1990, Traveltime inversion using transmitted waves of offset VSP data: Geophysics, 55, 1089-1097.

Lines, L.R., Bourgeois, A., and Covey, J.D., 1984, Traveltime inversion of offset vertical seismic profiles-A feasibility study: Geophysics, 49, 250-264.

Marquardt, D.W., 1963, An algorithm for least squares estimation of nonlinear parameters: Jour. Soc. Indus. Appl. Math., 2, 431-441.

Table 1. The model parameters, the initial estimates, and the final estimates of the first example. Traveltime curve is shown in Figure 2.

	Layer Number	Velocity (m/s)	Dip (slope)	Depth (m)
Model Parameters	1	1500.0	0.176	1500.0
	2	2000.0	0.087	2000.0
	3	2500.0	-0.035	3000.0
	4	4200.0	-0.123	4500.0
Initial Estimates	1	1000.0	0.010	1600.0
	2	1500.0	0.010	2200.0
	3	2000.0	-0.010	2900.0
	4	2500.0	-0.010	4800.0
Final Estimates	1	1500.0	0.176	1500.0
	2	1999.8	0.087	2000.0
	3	2419.3	-0.033	2967.5
	4	4318.2	-0.119	4508.4

Table 2. The model parameters, the initial estimates, and the final estimates of the second example. Traveltime curve is shown in Figure 3.

	Layer Number	Velocity (m/s)	Dip (slope)	Depth (m)
Model Parameters	1	800.0	0.200	200.0
	2	900.0	0.200	500.0
	3	1000.0	0.200	1000.0
	4	1200.0	0.200	1500.0
	5	1500.0	0.200	2000.0
Initial Estimates	1	600.0	0.050	100.0
	2	700.0	0.050	400.0
	3	800.0	0.050	700.0
	4	1000.0	0.050	1200.0
	5	1300.0	0.050	1800.0
Final Estimates	1	800.0	0.200	200.0
	2	900.0	0.200	500.0
	3	1000.0	0.200	1000.0
	4	1200.0	0.200	1500.0
	5	1500.0	0.200	2000.0

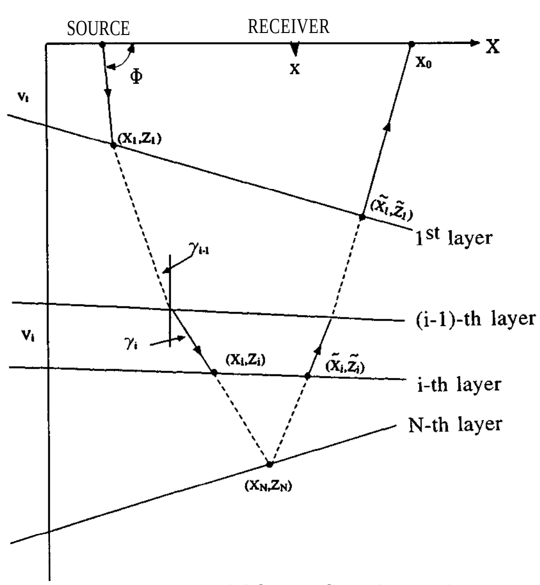

Figure 1. Model for a reflected arrival ray tracing, modified from Lee (1990).

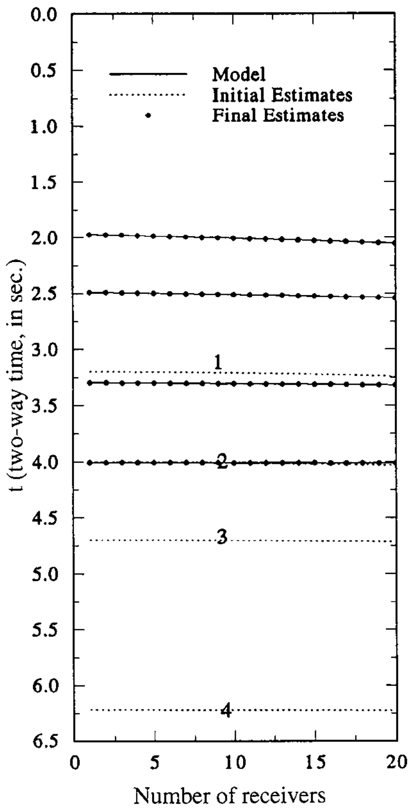

Figure 2. The reflected trveltime based on the models shown in Table 1. The receiver interval is 25 m.

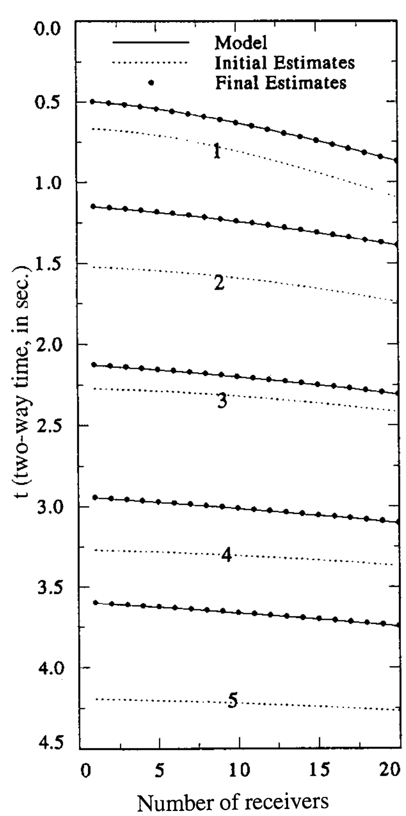

Figure 3. The refelcted traveltime based on the models shown in Table 2. The receiver interval is 25 m.