SIMULATED ANNEALING DETERMINATION OF SHEAR WAVE TRAVEL TIME

by

C.H. Cheng

Earth Resources Laboratory
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

and

Jack Foley

Geophysical Laboratory
Hanscom AFB, MA 01731-5000

ABSTRACT

The method of simulated annealing is introduced to obtain relative moveouts between different depths from an iso-offset section. This method has been shown to be more consistent than conventional picks based on peaks, troughs, or zero crossings especially in situations where the signal-to-noise ratio is low or the wavelet is emergent. This method also provides a means of quantifying the relative confidence in each pick over the entire depth of the well. The method has been applied to the data obtained by the ARCO shear wave logging tool and compared favorably with more conventional estimates of shear wave slowness and was shown to be robust, even in areas of weak arrivals.

INTRODUCTION

Direct shear wave logging has become an increasingly important tool in exploration and production geophysics. Shear wave logging tools use dipole radiation in the borehole to generate flexural waves in the formation. The fundamental mode of the flexural wave is analogous to the Stoneley wave in conventional acoustic logging. It is an interface wave. However, it is highly dispersive, and shear wave log traces are characterized by a long ringy signal. In soft sediments, the high attenuation associated with the formation shear wave also attenuates the flexural wave. Furthermore, the flexural wave is best excited at relatively low frequencies (around 2 to 3 kHz, as compared to 5 to 10 kHz for

conventional full waveform logs, Schmitt, 1988). This combination of low frequency and high attenuation results in flexural wave arrivals that are often emergent, making the exact travel time selection difficult to accurately measure manually or with interactive computer graphics software. Furthermore, a number of existing or prototype shear wave tools have only two or at most four receivers, making array processing difficult (Willis, 1983).

In essence, we are back to the original problem of shear wave velocity determination from full waveform logs that we encountered a decade ago, except that this time around the problem is more severe because of the lower frequency content and higher attenuation. In conventional monopole full waveform logging, the shear wave only exists in the "hard" formation where the shear wave velocity is higher than the acoustic velocity of the borehole fluid. Under such circumstances, the shear wave attenuation tends to be low, and the frequencies of the shear/pseudo-Rayleigh wave packet are usually around 10 kHz at cutoff. Even under these better conditions, Willis (1983) pointed out that the picking of individual arrival times on different traces is prone to error and the results are usually inconsistent from depth to depth.

In this paper we look at a new statistical approach when dealing with the problem of obtaining a shear wave velocity from the shear wave log. The key difference from conventional velocity picking and using the resultant moveout is that instead of using receivers from the same depth we will be using the same source-receiver (iso-offset) combination over a large number of depths. This is done by recasting the phase identification problem as an optimization problem. We design a function which will be maximized when each of the traces on the depth array is time shifted by amounts related to the differential travel time of the flexural wave arrival. In this way, we obtain a relative travel time of the flexural wave arrival over the whole depth. The process is repeated for a second source-receiver pair. The flexural wave moveout between the two receivers can then be determined by cross-correlating the first traces on the two source-receiver pairs. This way, we are guaranteeing consistent picks over the whole depth.

In the next section, we detail the simulated annealing based optimization method for measuring arrival times of shear wave logging data. The real-data application of this technique is presented in the final section. The simulated annealing technique is applied to data from ARCO's shear wave logging tool and the results are compared with automatic arrival time picking based on zero crossings provided by ARCO and also with shear wave velocity determined from cross correlation of the data from the two receivers at individual depths.

SIMULATED ANNEALING METHOD

The application of simulated annealing optimization to the problem of shear wave velocity determination helps to overcome some of the problems which make the travel time picks of low frequency emergent arrivals difficult and unreliable. In particular, the simulated annealing method helps us to quantify the accuracy of the picks by defining a probability associated with each travel time selection. As shown in Foley (1990), this method works well when the signal-to-noise ratio is low and when we ordinarily have difficulty pinpointing emergent arrivals in noisy data. Methods which utilize a probabilistic approach to solve various types of seismic travel time problems have had limited but successful application in the past. Rothman (1986) used the simulated annealing method to solve the problem of station statics corrections in reflection data. Landa et al. (1989) used a generalized simulated annealing technique (Bohachersky et al., 1986) for trace coherency optimization and CMP velocity estimations. Zelt et al. (1987) used a semblance optimization technique to select low amplitude refracted arrivals.

In the shear wave travel time applications, we search for a set of trace shifts that maximizes (optimizes) the total energy content of a group of seismograms windowed to contain the shear/flexural wave arrival. This group of seismograms is made up of records from a single source-receiver offset for a certain depth range.

$$\mathcal{E} = \sum_{i=1}^{N-window} \Big(\sum_{j=1}^{N-traces} d_{i,j}(\tau_j) \Big)^2, \tag{1}$$

where $\mathcal{E}$ is defined as the power of the stacked trace, called the *stack power*, which we want to optimize, $d_{i,j}(\tau_j)$ is the i^{th} point in the j^{th} seismogram and τ is the shift of the j^{th} seismogram determined with this procedure. We use this optimization technique to determine the lags $\tau(i)$ in the data which maximize the stack power of the targeted arrival. Only windowed portions of the full waveforms are used in this process, and these windows are defined for each depth by assuming that the arrivals fall within a predetermined range of travel times. These windows only define the portion of the seismogram in the time in which the targeted arrival is expected and do not shape or transform the data. We will refer to the stack of these windows as the *annealing stack*. We use simulated annealing optimization to efficiently maneuver through this multi-dimensional data space in an iterative process to locate the configuration with the maximum annealing stack power.

Simulated annealing is a Monte Carlo optimization procedure which is designed to globally optimize an objective function that contains local minima (Rothman, 1986). The term Monte Carlo means that at some point in the process a randomly generated number is used, and as we shall see below, this process is a Monte Carlo method because we generate a random number to make selections from probability distributions. The idea of simulated annealing, introduced by Kirkpatrick et al. (1983), is that one can

make an analogy between optimization of a function and the physical process of crystal growth (annealing). In this travel time problem, we wish to align all seismograms recorded from an iso-offset section such that the power of the annealing stack is greatest; this is our objective function. In early iterations we allow traces to be shifted with few constraints. These early iterations are analogous to the high *temperature* state of crystal growth, where a crystal is still mainly liquid and can attain virtually any molecular configuration. Later, the *temperature* of the process is lowered and the range of possible trace shifts becomes limited. Similarly, before a melt solidifies, the motion of individual molecules is not widely constrained, but as the temperature is lowered (slowly) to the Curie point, the range of motion is reduced until the final crystal lattice is developed.

This optimization can also be thought of as the determination of a set of parameter values (trace shifts) which minimizes a function (1/stack power) which contains local minima. This function has an energy state at each possible combination of trace lags, and simulated annealing finds the lowest energy level. Attempting to locate this global minimum with conventional iterative improvement techniques can lead to solutions which rest in local minima (Tarantola, 1987). We use simulated annealing to insure that the global minimum is reached in each case by iteratively optimizing the function and probabilistically allowing steps which *temporarily* have higher energy levels than previous iterations; we carefully allow for uphill motions on the objective function to avoid local minima. An outline of the simulated annealing procedure described below is shown schematically in Figure 1 and follows the procedure outlined by Rothman (1986).

1. The procedure begins by determining the time window for each trace for the targeted arrival. This could be a window around the shear wave arrival picked manually, or based on a multiple of the P-wave arrival time.

2. All trace shifts are initially set to zero.

3. The windowed data are summed to produce the *annealing stack* and the stack power is calculated.

4. One seismogram is removed from the annealing stack and a cross correlation is made between that trace and the sum of the remaining traces, which we refer to as the *partial stack*. This correlation is referred to as the *trace correlation* below.

5. The trace correlation function is converted into a probability distribution using a temperature conversion factor. The time shifts are weighted by the probability function and a selection is made from this weighted distribution. This time pick is the trace lag for the current iteration.

6. The trace lag is applied to the data and the shifted trace is added back into the partial stack.

7. We return to step 4 until each seismogram in the iso-offset section is processed for the current iteration.

8. The temperature parameter is lowered to limit the probability of shifting traces by lags with low trace correlation levels.

9. We return to step 3 for N iterations.

There are a few key points in this procedure:

- At each step in the procedure the data are aligned and summed to produce an annealing stack. This function is a very poor representation of the source wavelet in early iterations, and for the first iteration it is simply a summation of all the prediction windows.

- By the end of the optimization procedure the annealing stack develops into an excellent representation of the source wavelet.

- The trace correlation function is converted into a probability distribution using the relation,

$$PDF(i) = exp(XCOR(i)/T). \tag{2}$$

  Here $PDF(i)$ is the probability distribution, $XCOR(i)$ is the trace correlation function and T is the conversion temperature. The *temperature* parameter controls the conversion from trace correlation to probability distribution. High temperatures (early iterations) produce probability distributions which are more uniform, where low correlation values have relatively high probabilities, and low temperatures (late iterations) produce distributions with fewer and greater spikes; see Figure 2.

- A random selection is made from the PDF to determine the shift to be applied to the trace before it is added back into the partial stack. The random nature of this selection makes this a *Monte Carlo* method.

- With this procedure we probabilistically allow for shifts not associated with the maximum trace correlation when defining the optimal trace alignment. This allows us to avoid local minima.

- The procedure is executed for each station in the annealing stack to complete a single iteration. Many iterations are often required with simulated annealing (Rothman, 1986), but in this work stable trace configurations are generally found in as few as 25 iterations. Implications of these values are discussed below.

IMPLEMENTATION ISSUES

Conversion to Absolute Travel Times

The simulated annealing method is applied separately to each source-receiver separation. After a suite of trace shifts which maximize the stack power are determined, we calculate the relative delay for each station included in the event. The procedure is followed for all receivers. To obtain a shear wave velocity, we need to know the moveout between the receivers at the first depth, since the arrival times are picked relative to the first depth. In our example using ARCO's shear wave tool, there are two receiver positions. In order to compare the results with the automatic picks provided by ARCO, we shifted our relative picks such that the mean difference between the simulated annealing picks and ARCO's picks is minimized.

Conversion Temperature

The conversion from cross correlation to probability distribution is very important in the simulated annealing method (Rothman, 1986). In general there are three different conversion schemes: no annealing ($T = 0$; conventional cross-correlations), steady state annealing ($T = constant$ for all iterations) and temperature varying annealing ($T = T_0 \cdot [1 - cooling\ rate]^{iteration}$). The temperature is a sensitive parameter, and a 10% change in starting temperature can cause the procedure to get trapped in local minima. By implementing a scheme where temperature varies with iteration, a wider range of starting temperatures can be used. We significantly reduce the importance of the starting temperature and therefore automate the process. The tradeoff of this implementation is that the number of iterations necessary for the procedure to converge is increased. We find that in most cases we see convergence within 30-150 iterations. This convergence is quite rapid in contrast to the 3500 iterations Rothman (1986) needed to anneal data for static corrections. The relatively small number of iterations required in this application indicates that global minima can be easily reached with the application of this probabilistic approach to travel time picking and that the problem is not strongly contaminated with local minima.

RESULTS

The results of applying simulated annealing to the shear wave logging data are compared with those from two different picks: direct picks from ARCO and a simple cross-correlation between the two receivers with no interpretation of the cross-correlation function.

Before we go on to discuss the simulated annealing results, it is instructive to examine the shear wave velocity obtained from the more conventional methods and thus have a feel for the accuracies involved. Figure 3 plots the shear wave slowness obtained by direct picks and those from cross-correlation. The two different methods of determining shear wave velocity show a small offset of 1.6 μs/ft but a relatively large standard deviation of 7.6 μs/ft. Since the shear wave tool has a digitization interval of 11 μs and an inter-receiver spacing of 4.5 ft, a mispick of 1 point represents an error of 2.44 μs/ft.

Figures 4 and 5 show the comparisons between the picks provided by ARCO and those obtained by simulated annealing. The simulated annealing was done using 650 traces representing 325 ft. The two receivers were separately iterated. Since simulated annealing provides only a relative moveout between depths in this case, these relative times were converted to absolute times by matching them with the ARCO picks such that the mean difference between them is minimized. Also plotted by the side is the cross correlation peak for each depth.

In general, there is excellent agreement between the simulated annealing picks and those done by more conventional means from ARCO. For the near receiver, the mean difference between the two picks is 4 μsec, less than half a digitization point. The standard deviation is 22 μsec, or two digitization points. For the far receiver, the mean difference between the two picks is even smaller at 1 μsec, but the standard deviation is larger at 36 μsec, a little over 3 digitization points.

Figure 6 shows the comparison of the shear wave slowness as a function of depth obtained by simulated annealing and conventional picks. There is good agreement between the two, with a mean difference of 1.1 μs/ft (or half a digitization point) and a standard deviaton of 8.6 μs/ft (or three and a half points). Figure 7 shows the comparison of the shear wave slowness as a function of depth obtained by simulated annealing and cross correlation between two receivers. There is even better agreement between the two, with a mean difference of 0.3 μs/ft and a standard deviation of 5.8 μs/ft (or two and a half points).

One of the reasons why the standard deviation for the far receiver is larger is the significant disagreement between the simulated annealing picks and the ARCO picks at depths 3710 to 3720 ft. At the same time, the cross correlation peak shows much poorer correlation in the simulated annealing picks. This results in the only place where there is significant disagreement between the simulated annealing shear wave slowness and that from conventional picks. In order to resolve this difference, we have plotted the actual waveforms at 3714.5 ft together with the picks (Figure 8). For the near receiver, the two picks lie right on top of each other (actually they are one digitization point apart, but are too close to resolve at this plotting scale). For the far receiver, it is clear that there is a distortion of the waveform shape of the first arrival, resulting in a double trough right after the first peak. Simulated annealing picks the first trough, while ARCO picks the second trough.

Another way of visualizing what is happening is to examine the iso-offset plots (Figures 9 and 10). Figure 9 is the iso-offset plot of the near receiver seismograms. As can be seen from the figure, the waveforms are consistent and the first arrivals are easy to identify. However, in the iso-offset plot for the far receiver (Figure 10), one can see where the differences between the simulated annealing and the conventional picking algorithm arise. The leading arrivals from 3705 to 3715 ft are emergent, making proper identification of the arrival time difficult. Since simulated annealing uses the cross-correlation between different traces within the arrival window (in this case with a window length of about 100 points or 1100 μsec), it does a much better job in these cases where the exact pick can be ambiguous. Significantly, simple cross correlation between the two receivers results in a shear wave slowness more in line with the simulated annealing results (see Figure 7).

CONCLUSIONS

We have introduced the method of simulated annealing to obtain relative moveouts between different depths from an iso-offset section. This method has been shown to be more consistent than conventional picks based on peaks, troughs, or zero-crossings. Because it performs cross correlation over a wide depth range, it is possible to obtain consistency over the entire well, instead of just over the receiver array. It is also possible to compare the relative reliability of the picks over depth. We have applied this method to a section of shear wave logs and the results are more consistent than conventional picking and cross correlation methods over a section where the first arrivals are emergent.

ACKNOWLEDGEMENTS

We would like to thank Ken Tubman and ARCO Oil and Gas Co. for providing the shear wave logging data and the actual shear wave picks. This work was supported by the Full Waveform Acoustic Logging Consortium at M.I.T.

REFERENCES

Bohachecsky, I., M. Johnson, and M. Stein, Generalized simulated annealing for function optimization, *Technometrics, 28,* 209–217.

Foley, J.E., 1990, *Crustal Structure from Teleseismic Bodywave Data,* Sc.D. Thesis, M.I.T., Cambridge, MA.

Kirkpatrick, S., C. Gelatt, and M. Vecchi, 1983, Optimization by simulated annealing, *Science, 220,* 671–680.

Landa, E., W.B. Beydoun, and A. Tarantola, 1989, Reference velocity model estimation from prestacked waveforms: coherency optimization by simulated annealing, submitted to *Geophysics.*

Rothman, D.H., 1986, Automatic estimation of large residual statics corrections, *Geophysics, 51,* 332–346.

Schmitt, D.P., 1988, Shear wave logging in elastic formations, *J. Acoust. Soc. Am., 84,* 2215–2229.

Tarantola, A., 1987, *Inverse Problem Theory,* Elsevier, New York.

Willis, M.E., 1983, *Seismic Velocity and Attenuation From Full Waveform Acoustic Logs,* Ph.D. Thesis, M.I.T., Cambridge, MA.

Zelt, C.A., J.J. Drew, M.J. Yedlin, and A.M. Ellis, 1987, Picking noisy refraction data using semblance supplemented by a Monte Carlo procedure and spectral balancing, *Bull. Seis. Soc. Am., 77,* 942–957.

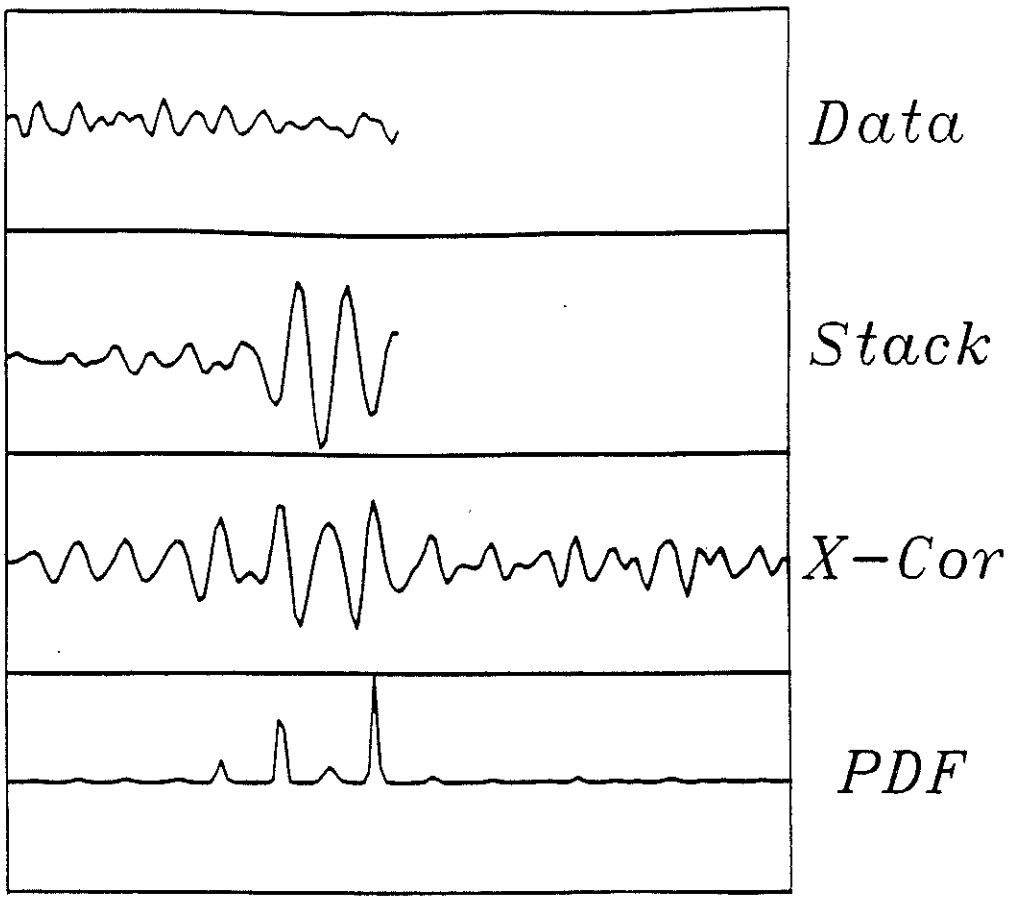

Figure 1: Simulated annealing optimization procedure schematic: A seismogram is windowed to contain the proper phase (top). All other depths are added to create the partial stack (second). A cross correlation is calculated between the windowed seismogram and the partial stack (third) and then converted into a probability distribution (fourth) using the temperature conversion scheme. A random selection is made from the distribution to define the trace shift. The data is shifted and the stack is updated. This procedure is done for each trace defining one iteration.

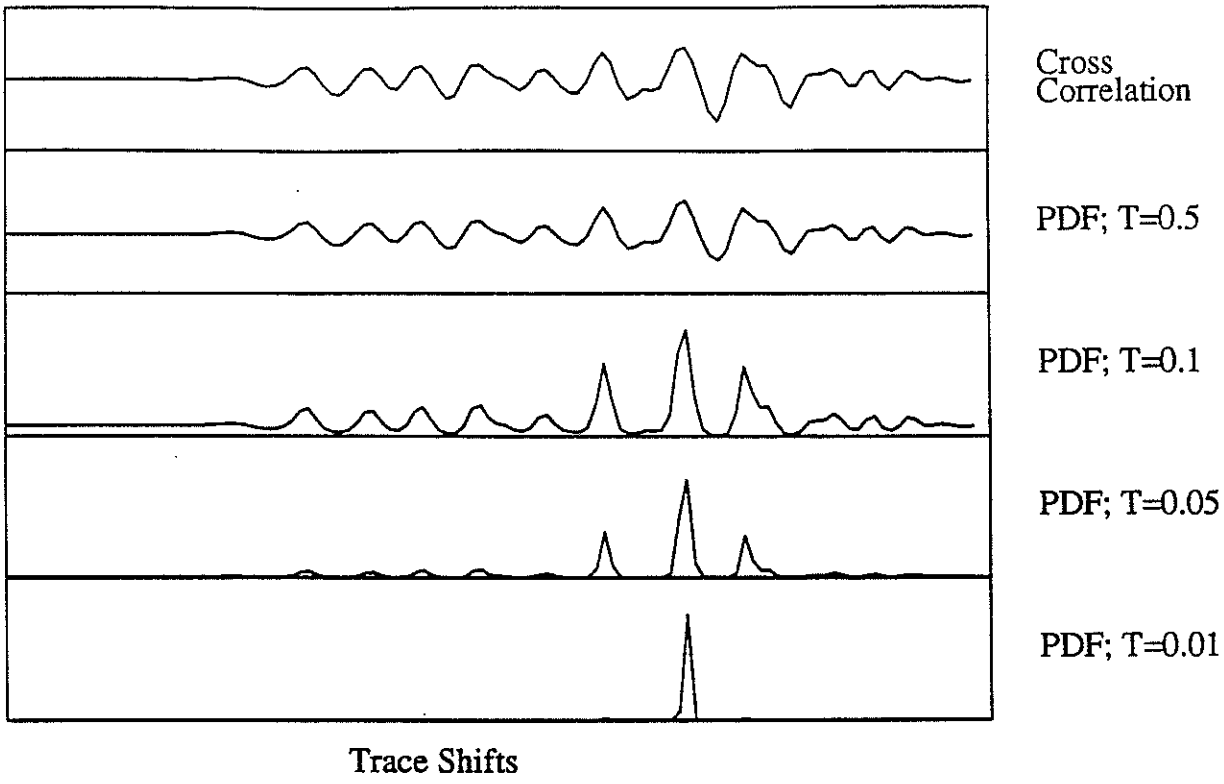

Figure 2: The conversion of cross correlation to probability distribution is dependent on the *temperature* (see text). Top figure shows a typical cross correlation between a prediction window and a partial stack. Below are the probability distributions for *temperatures* of 0.5, 0.1, 0.05, and 0.01. As the *temperature* is lowered, we limit the probability associated with low correlation lags.

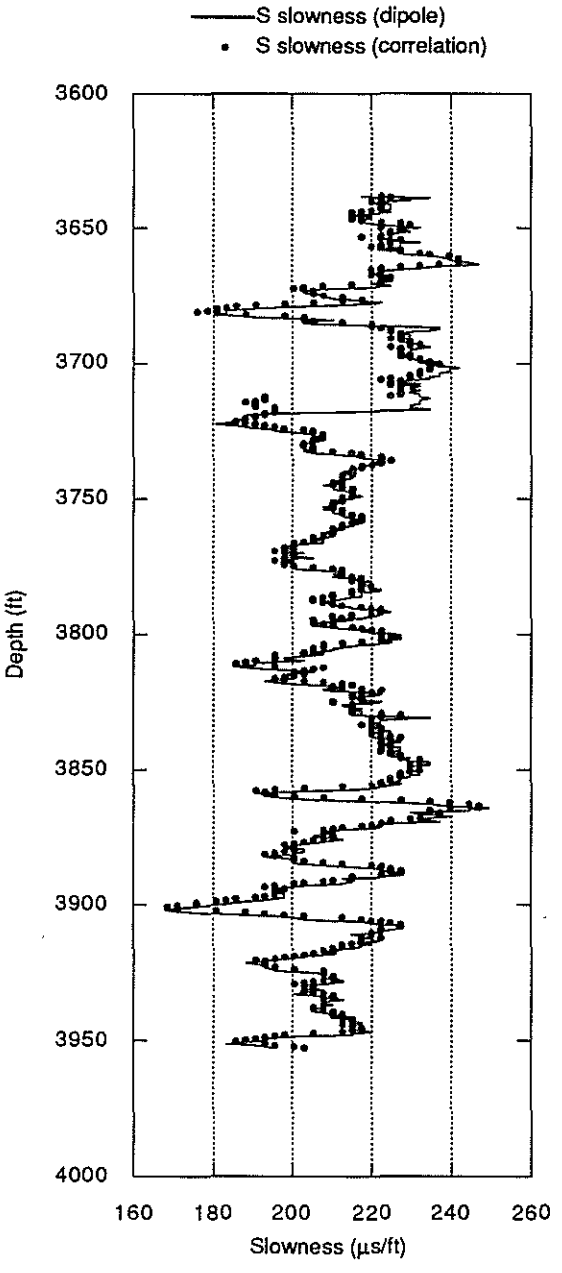

Figure 3: Comparison between shear wave slowness obtained by a conventional picking algorithm used by ARCO and a straightforward cross correlation method which measures the moveout between the two shear wave receivers.

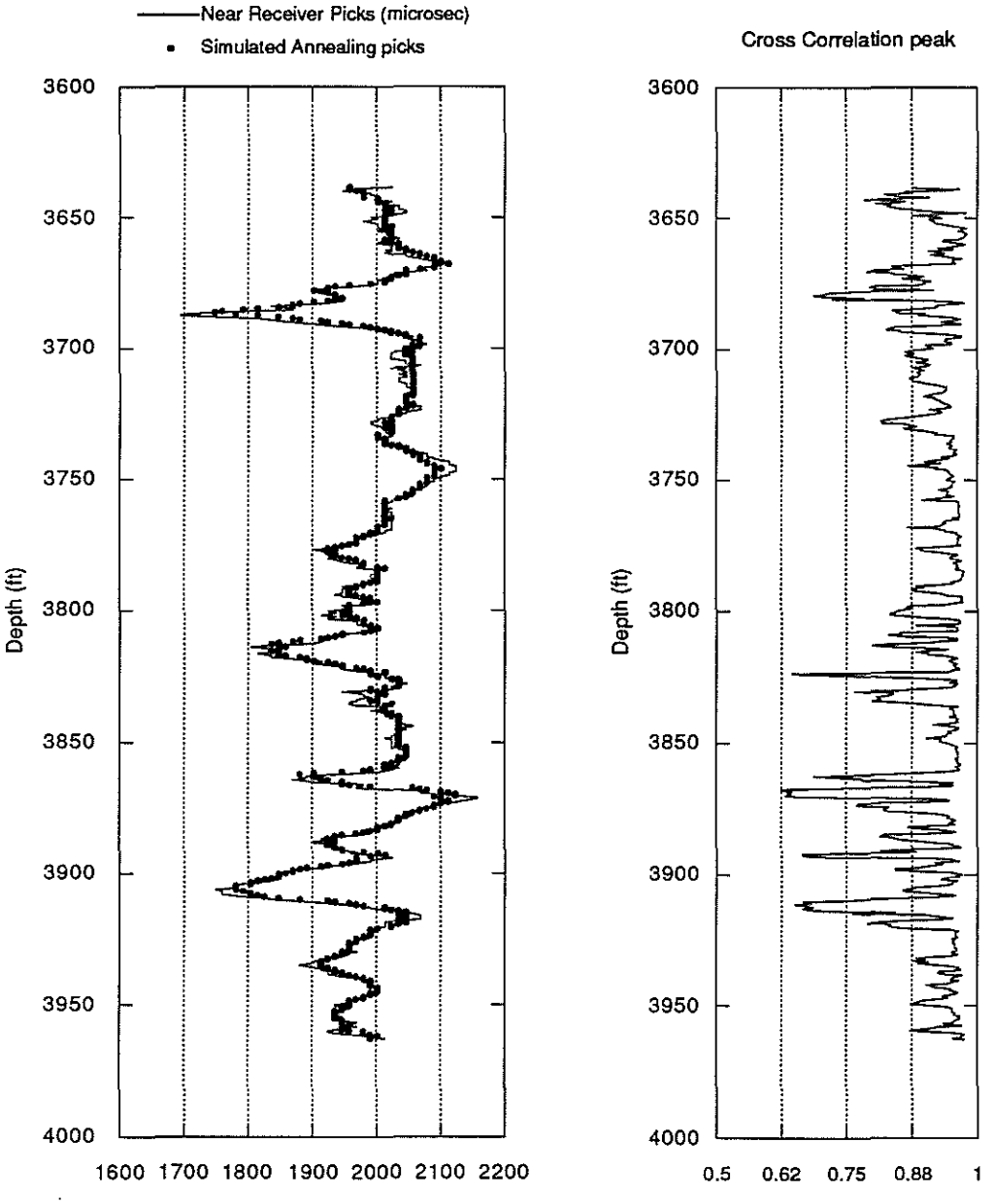

Figure 4: Comparison of arrival time picks for the near receiver from simulated annealing and those done by ARCO using a conventional picking algorithm. Also plotted is the peak cross correlation value for the trace from simulated annealing.

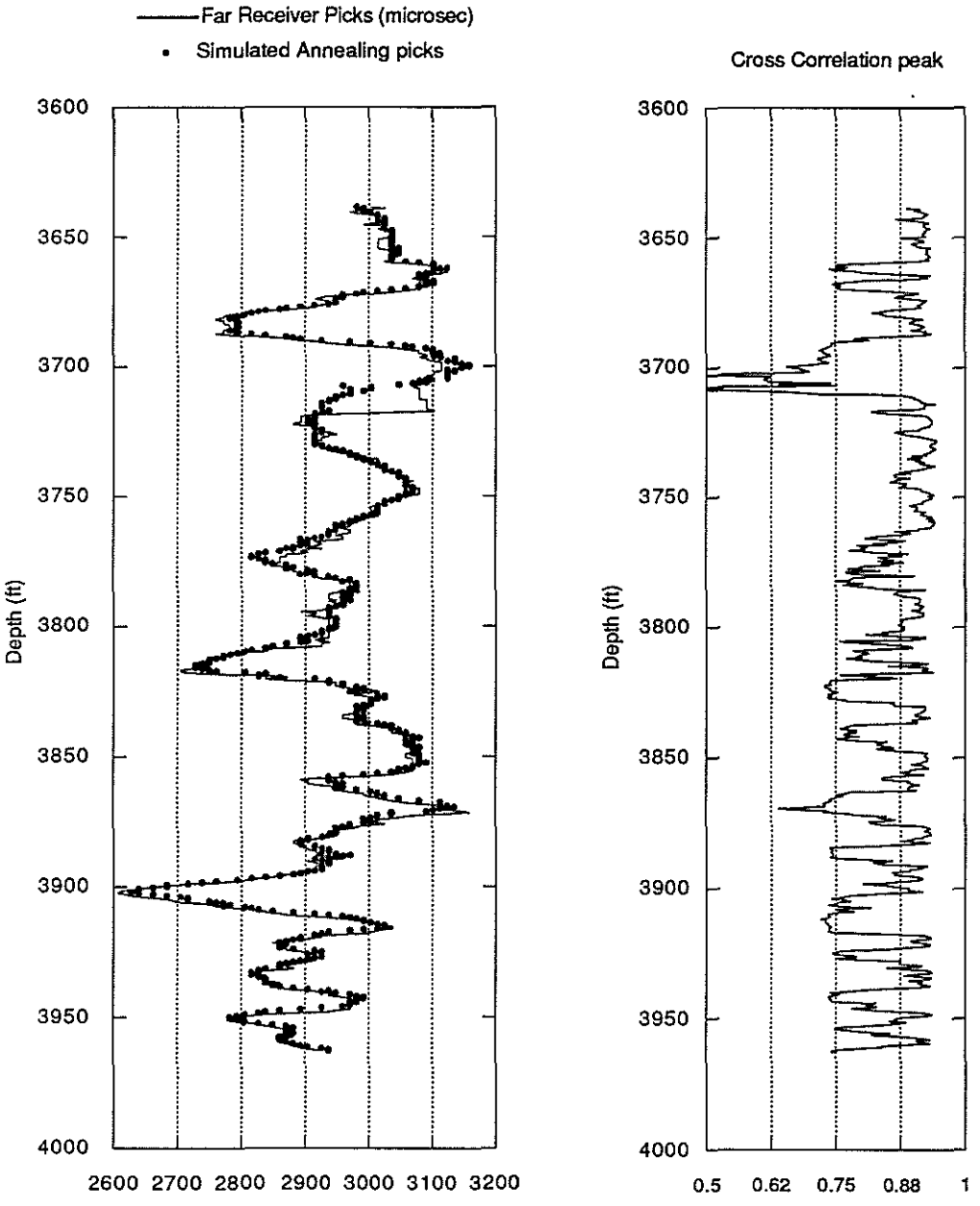

Figure 5: Comparison of arrival time picks for the far receiver from simulated annealing and those done by ARCO using a conventional picking algorithm. Also plotted is the peak cross correlation value for the trace from simulated annealing.

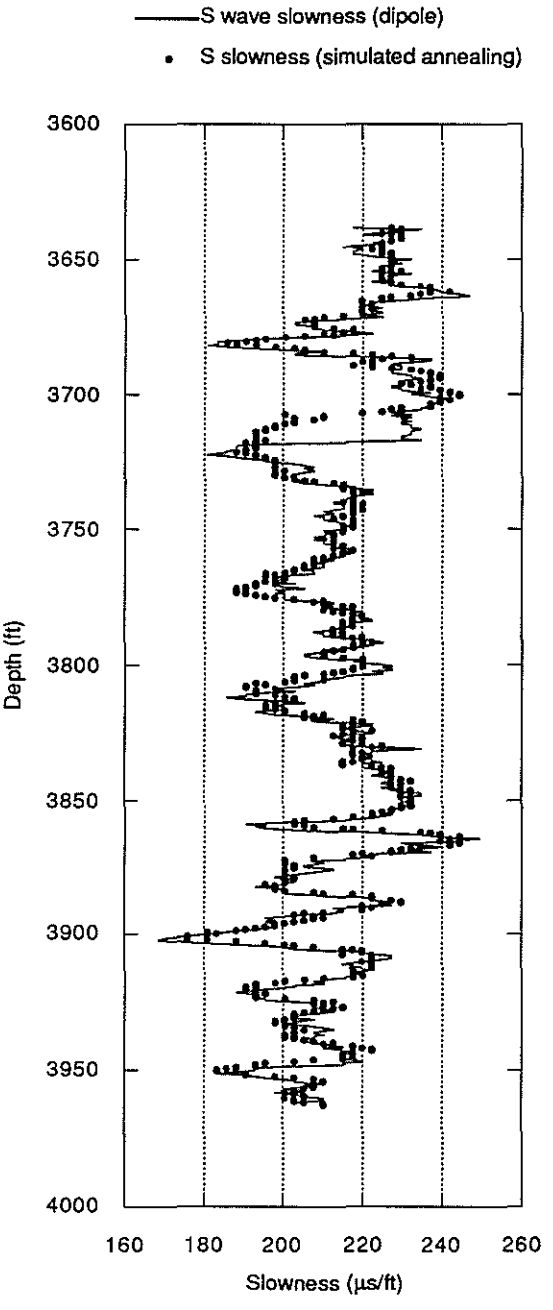

Figure 6: Comparison between shear wave slowness obtained by a conventional picking algorithm used by ARCO and the results from simulated annealing.

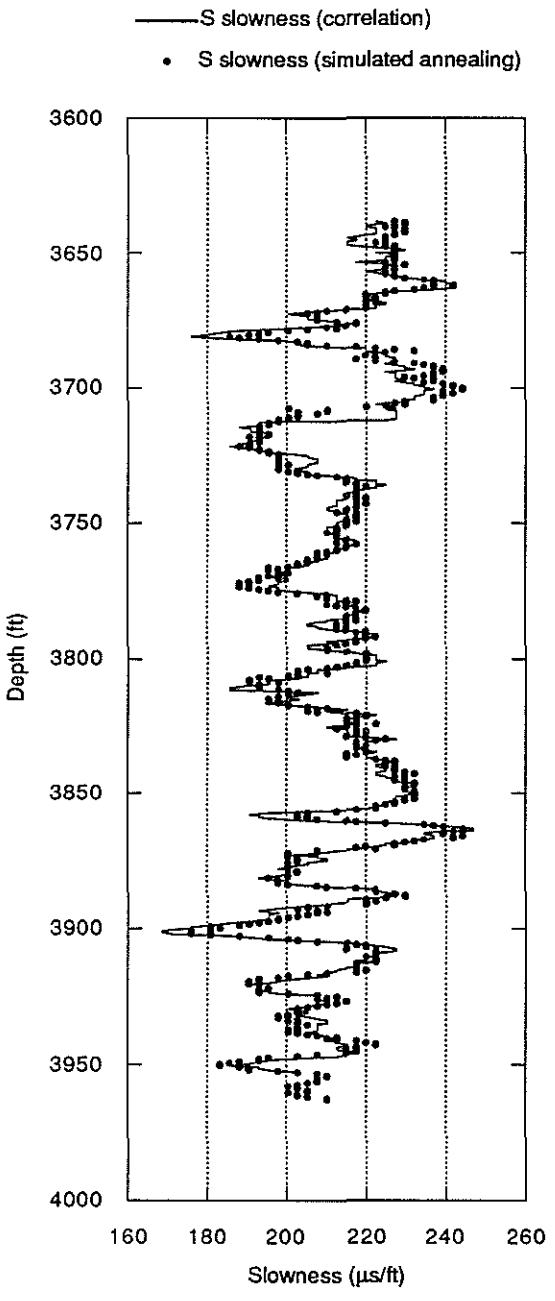

Figure 7: Comparison between shear wave slowness obtained by a conventional cross correlation between the near and far receivers and the results from simulated annealing.

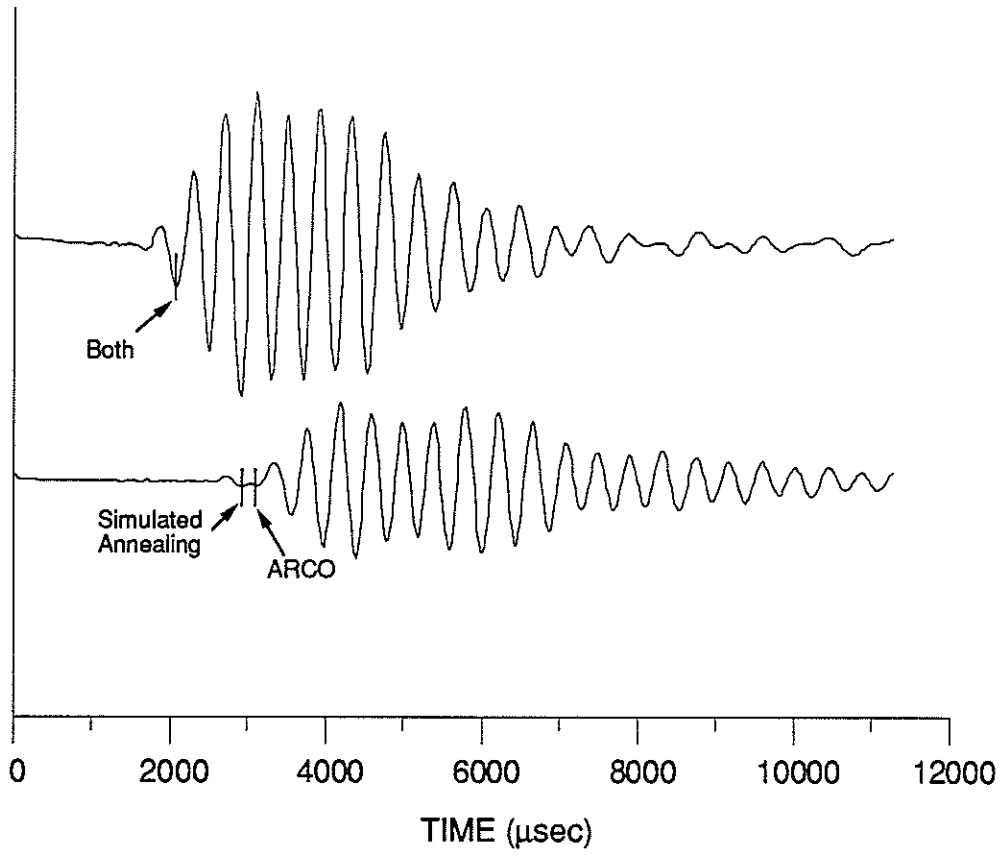

Figure 8: Seismograms for the near and far receivers obtained by the shear wave logging tool at 3714.5 ft. Plotted on top of the seismograms are the arrival picks from ARCO and from the simulated annealing algorithm.

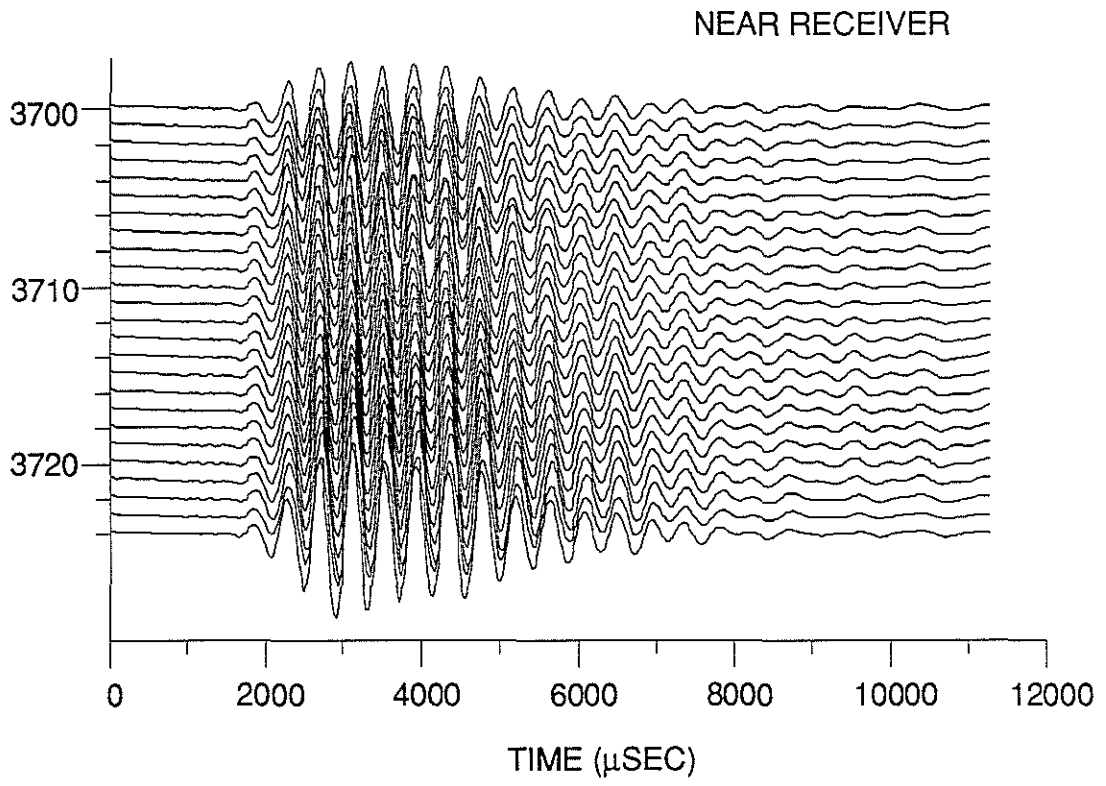

Figure 9: Iso-offset shear wave log seismograms for the near receiver for the depths 3700 to 3724 at one foot intervals.

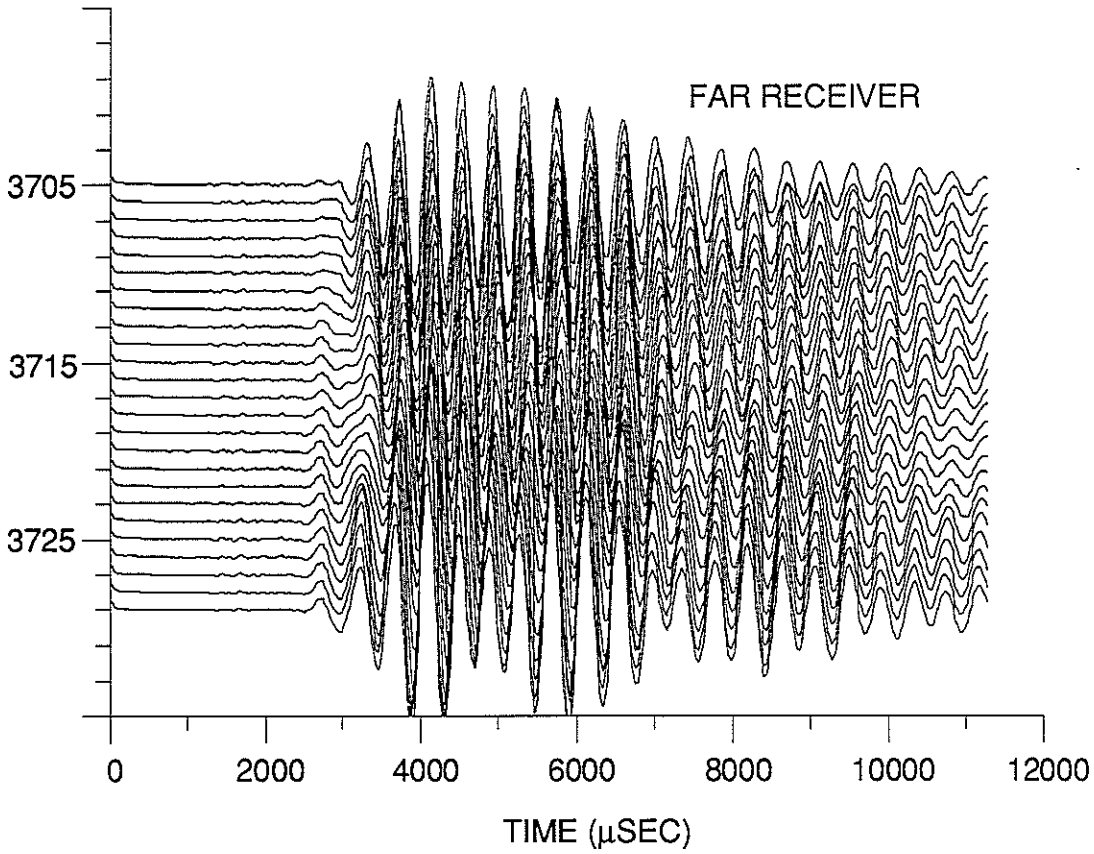

Figure 10: Iso-offset shear wave log seismograms for the far receiver for the depths 3705 to 3729 at one foot intervals.