A New Inversion Procedure for Spectral Analysis of Surface Waves
Using a Genetic Algorithm

by Shahram Pezeshk and Morteza Zarrabi

Abstract A new inversion procedure for spectral analysis of surface waves (SASW) using a genetic algorithm (GA) is presented. The inversion process proposed in this study starts by running a forward solution for the Rayleigh dispersion equation, with sets of random inputs, to find the theoretical phase velocities. Then, it continues by finding new and better sets of inputs through processes that mimic natural mating, selection, and mutation in each generation. The goal of the GA is to find the best match between the theoretical and the experimental dispersion curves. Therefore, with each new generation there is a better agreement between the calculated output theoretical dispersion curve and the input experimental dispersion curve. To start the procedure, two options are available, either requesting the GA-based optimization process to obtain shear-wave velocities and thicknesses for each layer, or providing the thicknesses and requesting the optimization process to obtain the best set of shear-wave velocities. The GA part of the procedure is fast, stable, and accurate, with several advantages compared to the traditional methods. The strength and accuracy of the proposed procedure are presented through two example problems. We show that (1) the inversion process using a GA results in a good agreement between the theoretical and experimental dispersion curves, and (2) the shear-wave velocity profiles obtained from the approach presented in this study and a downhole seismic survey show a good level of agreement.

Introduction

Shear-wave velocity or shear modulus at low strain is an important input parameter in soil dynamic analyses. For example, the shear-wave velocity profile is required data for running computer programs such as SHAKES1 (Idriss and Sun, 1992) for estimation of the site-dependent seismic amplification factor of a site (Crammer et al., 2002; Pezeshk and Liu, 2001; Borcherdt, 1994; Kramer, 1996; Pezeshk et al., 1998; Evans and Pezeshk, 1998; Electric Power Research Institute [EPRI], 1993).

One of the procedures that has gained popularity in recent years in estimating elastic properties and layer thicknesses of soil profiles is the spectral analysis of surface waves (SASW). SASW is a nondestructive soil characterization method based on surface-wave propagation. When the velocity and frequency of a wave are not independent, the wave is dispersive. This dispersive behavior of the surface waves in nonuniform materials is used to obtain the shear-wave velocity profile. Hooke (1985), Park et al. (1999), Sato et al. (1997), Liu et al. (2000), and Lowe (2001) are among investigators who used the dispersive behavior of surface waves for soil shear-wave velocity profiling. In general, the SASW procedure consists of measuring the surface-wave velocity at various frequencies to obtain the dispersion curve for a given site. Good overviews of the SASW procedure can be found in Rix et al. (2002), Hebeler (2001), Matthews et al. (1996), Rix et al. (1991), and Nazarian and Stokoe (1986). Browa et al. (2000) performed an evaluation of the SASW method in which they concluded that the comparison of the shear-wave velocity profile from downhole seismic and SASW testing is "generally good." Furthermore, comparisons of the SASW results with crosshole tests by Nazarian and Stokoe (1984, 1986) showed a good level of accuracy for the SASW procedure. The advantage that the SASW method is independent of boreholes categorizes it as a noninvasive method. This makes SASW practically and economically a great tool for engineers. Also, compared with other noninvasive methods (e.g., reflection and refraction), it may provide good resolution and more flexibility for near-surface media (Foti, 2000). Recent developments in signal generation and signal receiving technologies along with the use of advanced computers and software have had positive effects on the development of SASW.

The general SASW procedure that has been used by several investigators is as follows.

1. Wave generation of a principal vertical ground motion
using either impulsive (hammer) or continuous (shaker) sources.
2. Measurement of the returned signals by geophones or accelerometers that have been placed on the ground in a specific configuration.
3. Recording the received signals by spectrum analyzers, seismographs, or a computer controlled by a program using special software.
4. Spectral analysis of the recorded time series data that results in the development of a dispersion curve; that is, the curve of variation of phase velocity (Rayleigh-wave velocity) with frequency (wavelength).
5. Inversion of dispersion curves to estimate the shear-wave velocity profiles.

The inversion process of the dispersion curve has been the focus of many studies during the last two decades. Hebler (2001), Thomson (1950), Haskell (1953), Nazarian (1984), Horike (1983), Tsun (1999), and Yang (1998), among others, have worked on this inversion technique. Other than the traditional gradient-based inversion methods, these are alternative inversion processes such as simulated annealing (Liu and Luke, 2004; Kolar, 2000; Sharma and Kalikow, 1998) and genetic algorithms (GAs).

GAs have recently been used for inversion procedures by several investigators to identify the Earth's vertical cross section structure, such as the recent work by Chang et al. (2004). They used GAs to model crustal structure in southern Korea. Bischetti et al. (1997) presented a GAs that simultaneously generated a large number of different solutions to several potential field inverse problems. They discussed the effectiveness and flexibility of the GA method for a range of different potential field inverse problems, both in 2D and 3D, on synthetic and field data.

The inversion method used in this study is a new approach to the inversion process using GAs. GAs are optimization and search techniques that simulate the evolution process used by Mother Nature, which is based upon Darwin's survival of the fittest idea. The GA process starts by using a forward method to find a theoretical dispersion curve, and then it continues by optimizing and fitting the theoretical dispersion curve to the experimental dispersion curve. When the best fit between the theoretical and experimental dispersion curves is found, the final result of the optimization will be a shear-wave velocity profile for the site.

Inversion

Inversion is the last and most important part of utilizing SASW in estimation of shear-wave velocity profiles. The inversion process in this study consists of two major steps. Step 1 is to use forward theory, as opposed to inverse theory, to find a theoretical dispersion curve. Step 2 is to use a GA to adjust the theoretical dispersion curve to fit to the experimental dispersion curve and eventually obtain the shear-wave velocity profile.

The forward problem in this study is to solve the Rayleigh dispersion equation (equation 1) for dispersion data or phase velocities (Lai and Rix, 1998):

\[ F_{D} \left[ \hat{\alpha}(\gamma), G(\gamma), \rho(\gamma), v_s, \omega \right] = 0 \]  

where \( \hat{\alpha}(\gamma) \) and \( G(\gamma) \) denote Lame's elastic moduli as functions of depth \( \gamma \), \( \rho(\gamma) \) denotes mass density as a function of depth \( \gamma \), \( k \) denotes the wave number, and \( \omega \) denotes the frequency of excitation. Solving equation (1) will result in a theoretical Rayleigh phase velocity profile.

In the process of the inversion, the values of Poisson's ratio and mass density are hardly changed from their initial values, since the influence of these parameters on the calculated phase velocities is of secondary importance for reasonable initial estimates (Rix et al., 1991). Therefore we can rewrite equation (1) for the purposes of this study as:

\[ F_{D}[V_{s}, H] = 0 \]  

where \( V_{s} \) is an \( n \times 1 \) vector of shear-wave velocities, \( n \) is the number of layers including the half-space, and \( H \) is an \( (n-1) \times 1 \) vector of thicknesses excluding the half-space. The other parameters involved in equation (1) are considered to be constant and are not to be varied in the inversion process.

A computer program developed by Rix and Lai (1999) based on research by Lai and Rix (1995) and Hisada (1994) is used to solve the forward problem expressed in equation (1) or equation (2). The computer program has been developed in a MATLAB environment, which has the capabilities of surface-wave testing using wave propagation theory and signal processing. The output of the program is the modal phase velocities of the Rayleigh waves. A typical plot of the calculated phase velocities by the program versus frequency (\( \omega \)) is illustrated in Figure 1. The data presented in Figure 1 are the fundamental mode phase velocities generated for a

![Figure 1](https://example.com/figure1.png)

Figure 1. A plot of the obtained theoretical dispersion curve using the computer program developed by Rix and Lai (1999).
A New Inversion Procedure for Spectral Analysis of Surface Waves Using a Genetic Algorithm

Typical soil under a set of specific frequencies. The frequencies vary from 3.75 Hz to 100 Hz as follows: (1) \( f = 3.75 \text{ Hz to } 1.5 \text{ Hz with } \Delta f = 0.625 \text{ Hz; (2) } f = 1.5 \text{ Hz to } 0.25 \text{ Hz with } \Delta f = 0.125 \text{ Hz; and (3) } f = 0.25 \text{ Hz to } 0.1 \text{ Hz with } \Delta f = 0.0625 \text{ Hz.} \)

Figure 1 is a typical dispersion curve that demonstrates the dispersive behavior of the Rayleigh waves. Dispersion curves basically show the dependence of the propagation velocity of Rayleigh waves on the frequency or wavelengths of the waves in a layered medium. However, in a half-space medium, the phase velocity is independent of the frequency of the Rayleigh waves.

Genetic Algorithm

The main objectives of this study is to use a genetic algorithm (GA) to estimate shear-wave velocity profiles using data obtained from experimental dispersion curves. A GA is an optimization and search technique that simulates the natural evolution process. GAs are global search methods based on a stochastic approach, which rely on strategy of survival of the best fit (Holland, 1975). The results obtained in an inversion process using GA methodology are considered more dependable (Goldberg, 1989; Pezeshki and Camp, 2002) because:

- The GA approach is independent of initial information, so there is no need to determine a set of initial design parameters.
- GA methods are not gradient-based methodologies; they use objective function information and a probabilistic transition scheme with no use of gradient information.
- GAs do not utilize the variables themselves; instead they use a coding set of variables.
- GAs do not improve a single solution; instead they work on a population of possible solutions.

A GA is based on the mathematical modeling of the mechanism of a genetic evolution strategy (Holland, 1975; Goldberg, 1989; Pezeshki and Camp, 2002). GAs do not rely on the specific relationship between the objective function and the boundary conditions (Pezeshki and Camp, 2002).

All GAs can basically be characterized as follows:

1. They work on a population of problem variables, which are usually created randomly. Variables are grouped in variable sets; each is called a string and composed of a series of characters that defines a possible solution for the problem. Characters in each string are typically binary numbers, which are evaluated after decoding to real or integer numbers to represent the values of the discrete problem variables for a particular solution.
2. The performance of the problem variables, as described by the objective function and the constraints, is represented by the fitness of each string. A mathematical expression, called a fitness function, calculates a value for a solution of the objective function. The fitter solution gets the higher value and the ones that violate the objective function and constraints are penalized. Therefore, like what happens in nature, the fittest and best solutions will survive and get the chance to be a parent of the next generation.
3. In a crossover procedure, two selected parents reproduce the next generation. The procedure first divides the selected parent strings into segments, and then some of the segments of a parent string are exchanged with the corresponding segment of another parent string. One-point (Goldberg, 1989), multiple point, and uniform crossover (Camp et al., 1998; Pezeshki et al., 2000) are among the several crossover patterns. The one-point crossover employed by Goldberg in his Simple Genetic Algorithm (SGA) divides each selected parent set into two parts and then interchanges the second-string parts to generate two new strings (Goldberg, 1989).
4. The mutation operation, which acts as an insurance policy (Goldberg, 1989), guarantees diversity in the generated populations. This is usually done by flipping (0 to 1 or vice versa) a randomly selected bit in the selected binary string to create a mutated string. Mutation prevents a fixed model of solutions from being transferred to the next generation. It allows for the possibility of generating children with nonexisting features from both parent strings.

For this study, we adapted and modified the backbone GA routines from the Genetic Algorithm TOOLBOX for use with MATLAB developed by researchers at the University of Sheffield, Department of Automatic Control and Systems Engineering (Chippertfield et al., 2003). A flowchart of how the inversion process is performed using a GA approach is presented in Figure 2.

Inverting the Experimental Dispersion Curve Using a GA

The objective of this research is to present the procedure used for inverting the experimental dispersion curve using a GA. The procedure seeks to find the best combination of the stiffnesses of the soil profile and their corresponding shear-wave velocities to minimize the difference between the experimental dispersion curve (target) and the theoretical dispersion curve. The theoretical phase velocities (theoretical dispersion curve) are obtained by solving the Rayleigh dispersion equation (equation 2) for each generation. The deviation of the results from the target is measured by the mean of the square of error between the square root of the sum of the squares of the theoretical dispersion curve and target (see Fig. 3).

The optimization problem is formulated as the minimization of the error function \( E \) between the theoretical dispersion curve and the experimental dispersion curve. The error function is defined as:
\[ E = \min \left( \left\| \mathbf{V}^{\text{experimental}} - \mathbf{V}^{\text{theoretical}} \right\|_2 \right)^{1/2} \]  

where \( \mathbf{V}^{\text{experimental}} \) is an \( n_f \times 1 \) vector of experimental phase velocities, \( \mathbf{V}^{\text{theoretical}} \) is an \( n_f \times 1 \) vector of theoretical phase velocities obtained by solving the forward problem \( \mathbf{V}^{\text{theoretical}} = f(V, H) \), \( n_f \) is the number of frequencies, and \( \| . \|_2 \) is the Euclidean norm. The optimization problem presented in equation (3) is also subjected to the following constraints:

\[ V_{l,\text{min}} \leq V_l \leq V_{l,\text{max}} \]  

and

\[ H_{l,\text{min}} \leq H \leq H_{l,\text{max}} \]  

where \( V_{l,\text{min}} \) and \( V_{l,\text{max}} \) are vectors of the lower and upper bound assigned to each layer's shear-wave velocity, respectively; and \( H_{l,\text{min}} \) and \( H_{l,\text{max}} \) are vectors of the lower and upper bound on the thickness of each layer, respectively.

The forward problem \( f(V, H) \) solution may result in dispersion curves that correspond to several modes of propagation (Guceri and Woods, 1992; Rix et al., 1992). In general, surface waves consist of the summation of many modes of propagation. However, the fundamental mode usually dominates when the source is located on the surface. In this study, we used the fundamental mode of propagation in the inversion process. Furthermore, in the proposed GA-based inversion process, we can either determine shear-wave velocities for a given set of layer thicknesses, if known, or determine both shear-wave velocities and layer thicknesses simultaneously. One of the advantages of using GAs is that no derivatives need be calculated, and, as a result, many typical numerical problems that exist with traditional procedures are eliminated.

**Examples**

To illustrate the strengths of the proposed procedure, two examples are presented. Through these examples, the stability of the GA in adapting itself to match a given target, which in this study is the experimental phase velocities, is illustrated. The first example is the application of the proposed procedure to determine a set of thicknesses and shear-wave velocities corresponding to the best match of the theoretical and experimental phase velocities. The second example problem, on the other hand, is to estimate the shear-wave velocity profile that results in a good match between the theoretical and experimental phase velocities for a site in Paris, Tennessee. In both examples, the experimental phase velocities, which are determined from field measurements, are known and given.
Example 1

The design variables for this problem consist of 22 unknowns, 11 thicknesses and 11 shear-wave velocities. A shear-wave velocity of 755 m/sec is assumed for the half-space. The target input data consist of the ordinates of experimental phase velocities at 61 frequencies. A genetic search of a population of 25 over 200 generations with a crossover ratio of 60% and a mutation probability of 1% was utilized. Figure 4 shows the experimental phase velocities and the best obtained theoretical phase velocities at these 61 frequencies. The convergence history is shown in Figure 5. From Figure 5, it can be observed that the GA-based procedure practically converged in 75 generations. The algorithm selected 11 thicknesses and 11 shear-wave velocities, as presented in Table 1 and Figure 6.

It is to be noted that the solution to this problem is not unique. Several different solutions consisting of different combinations of thicknesses and shear-wave velocities were obtained by running the GA-based inversion process that resulted in a very good agreement between the theoretical and the experimental phase velocities.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Thickness (m)</th>
<th>Shear Wave Velocity (m/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>187</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>397</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>466</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>233</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>280</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>494</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>531</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>722</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>372</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>713</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>366</td>
</tr>
<tr>
<td>Half-space</td>
<td></td>
<td>755</td>
</tr>
</tbody>
</table>

Table 1: Thicknesses and Shear-wave Velocities Obtained by the GA for Example 1

![Graph showing phase velocity vs. frequency](image1)

**Figure 4.** Phase velocities of the experimental result and the phase velocities of the converged solution for example 1.

![Convergence history graph](image2)

**Figure 5.** Convergence history of example 1.

![Shear wave velocity profile](image3)

**Figure 6.** Shear-wave velocity profile obtained for example 1.
Example 2

In this example, a comparison of the shear-wave velocity profiles obtained from a site in Paris, Tennessee, using the proposed GA-based inversion procedure and a downhole seismic survey is presented. The shear-wave velocity profile used for comparison was obtained from a downhole seismic survey method performed by Ponzseik et al. (1998). The experimental dispersion curve for this site was determined using a multistation SASW procedure. The shear-wave velocities are then calculated using the proposed procedure. The procedure is similar to that of example 1. In this example there is a population of 60 over 40 generations, and the only parameter that is modified by the program is the shear-wave velocity. Therefore, in each new generation the input shear-wave velocities will be adapted in a way that the output theoretical phase velocities better match the input experimental phase velocities. As can be observed from Figure 7, there is a good agreement between the final output theoretical and the input experimental phase velocities for the Paris site.

The convergence history of example 2 is shown in Figure 8. It can be observed that the problem practically converges in about 25 generations. The final selection of 15 shear-wave velocities by the algorithm is illustrated in Figure 9. In addition, the shear-wave velocities determined from a downhole seismic survey are also plotted as solid squares.

![Graph 1](image1)

**Figure 7.** Phase velocities of the experimental data obtained using the multistation SASW method and the phase velocities of the converged solution for example 2 using the proposed procedure.

![Graph 2](image2)

**Figure 8.** Convergence history of example 2.

![Graph 3](image3)

**Figure 9.** Shear-wave velocity profiles obtained for example 2. The solid line denotes the surface-wave test and solid squares denote the downhole seismic test.
in Figure 9. From this figure, it can be observed that the SAW inversion procedure using a GA algorithm results in a good estimation of shear-wave velocities at different depths in comparison with the seismic downhole survey. Although the test locations in Paris were very close, the top layers of the sites were not necessarily the same owing to the construction procedures around the borehole location. This could explain the discrepancies in the first few meters.

Conclusions

A new inversion procedure for SAW is presented. The procedure uses a genetic algorithm (GA) to obtain the theoretical dispersion curve that matches the experimental dispersion curve. The input to the GA is either a union of soil thicknesses and shear-wave velocities or just shear-wave velocities as a single individual. Each generation of individuals is modified through the processes that mimic nature's mating, natural selection, and mutation. The process continues until an optimum individual set is obtained. The optimum individual will represent a soil profile with a dispersion curve that best matches the experimental dispersion curve. Since the inversion problem is ill-posed with monotonic solutions, the ultimate result of the optimization process will be a model of soil characteristics. The procedure is stable and accurate with several advantages compared with the traditional optimization methods. The results from the two provided examples show a very good agreement between the calculated output phase velocities and the input experimental phase velocities. The shear-wave velocity profile of example 2 obtained from this study agrees well with the results of a downhole seismic test.

Acknowledgments

The authors would like to thank Professor Zhao-Xia Rui of the Georgia Institute of Technology for providing us with his computer programs and helpful suggestions. This work was partially funded by the Tennessee Department of Transportation.

References

on the Website of the Department of Automatic Control and Systems Engineering, University of Sheffield.
method (theoretical aspects), Research Report 417-2. Center for Transportation Research, the University of Texas at Austin, Texas.


Rix, G. J. and C. G. Lo (1999). MATLAB forward solution, Free Computer Programs on the Website of Civil Engineering Department, Georgia Institute of Technology.


Rix, G. J., H. Stokoe II, and J. M. Roest (1993). Experimental study of factors affecting the spectral analysis of surface wave method, Research Report 1213-5, Center for Transportation Research, the University of Austin, Texas.


Manuscript received 25 July 2004.