INTERNATIONAL SOCIETY FOR
SOIL MECHANICS AND
GEOTECHNICAL ENGINEERING

SIMSG ISSMGE

This paper was downloaded from the Online Library of the International Society for Soil Mechanics and Geotechnical Engineering (ISSMGE). The library is available here:

https://www.issmge.org/publications/online-library

This is an open-access database that archives thousands of papers published under the Auspices of the ISSMGE and maintained by the Innovation and Development Committee of ISSMGE.

Geotechnical and Geophysical Site Characterisation 5 – Lehane, Acosta-Martínez & Kelly (Eds)
© 2016 Australian Geomechanics Society, Sydney, Australia, ISBN 978-0-9946261-2-7

Inversion of effective phase velocity seismic surface wave data by partial least squares regression

G. Heymann, D.N. Wilke & S. Kok
University of Pretoria, Pretoria, South Africa

ABSTRACT: Inversion of seismic surface wave data is an important stage in the process to obtain the profile of shear and compression wave velocities of a ground profile. Basic inversion techniques often assume that Rayleigh waves are propagated by the fundamental mode. For normally dispersive ground profiles this may be a reasonable assumption, but for inversely dispersive profiles or profiles, with high velocity contrasts, higher modes will also contribute strongly at high frequencies of excitation. Most inversion strategies for such profiles require that the phase velocity be propagated at one dominant mode at any given frequency of interest with a clear discontinuity when the phase velocity moves from one mode to the next. In practice this may sometimes be achieved by employing a large number of geophones, but for complex profiles the number of geophones needed may be impractical. For such profiles, using the number of geophones typical for surface wave testing, the effective phase velocity will be measured where mode superposition occurs with a gradual transition from one mode to the next. Inversion of such data is not possible with conventional techniques, as each mode can not be identified separately and requires the effective phase velocity dispersion data to be inverted directly. This paper proposes a partial least square regression strategy for the inversion of effective phase velocity dispersion data.

1 INTRODUCTION

Seismic surface wave testing has become popular for geotechnical site characterization (eg. Foti et al. 2015, Stokoe et al. 2004). These tests are relatively fast to perform, cost effective and non-invasive. Numerous active surface wave tests have been developed where an active seismic source is used such as a vibrator, sledge hammer or explosive source. Active tests include CSW - continuous surface wave test (e.g. Matthews et al. 1996), SASW - spectral analysis of surface waves (e.g. Stokoe et al. 1994) and MASW - multichannel analysis of surface waves (e.g. Park et al. 1999). In addition a number of passive surface wave techniques have been developed which uses background vibration as seismic source. Passive tests typically target low frequencies and therefore allow characterisation of deep layers and are often used to complement active surface wave tests. Passive tests include SPAC - spatial autocorrelation (Aki 1957) which uses regularly shaped geophone arrays such as L-shaped, T-shaped, circular or triangular and ReMi - Refraction Microtremors (Louie 2001) which uses linear geophone arrays.

The dispersion data lies central to surface wave testing, and it is measured experimentally. Inversion of the dispersion data is required to construct a profile of seismic wave velocity with depth. Most analytical techniques used for inversion of dispersion data assume that the energy is propagated by discrete modes, be it the fundamental mode only, or the fundamental mode together with other discrete higher modes. This is only true for very long geophone arrays placed sufficiently far from the seismic source. For typical active surface wave testing this is often not the case and mode superposition occurs in which case the effective phase velocity is measured. The continuous surface wave test has the advantages of using a vibrator (Heymann 2013), which allows good quality control in the field and a short geophone array. Foti et al. 2015 noted that with a short geophone array there is less risk of insufficient signal to noise ratio, high frequency attenuation and spatial aliasing as well as lateral variations of the soil properties. However when using a short geophone array it is likely that the effective phase velocity will be measured and not discrete modes as often assumed by inversion algorithms. Robust techniques for inversion of effective phase velocity are therefore required.

It is well known that surface wave inversion is ill-posed. Various least squares inversion strategies have been proposed and additional information can be incorporated to alleviate the ill-posed nature of the problem (Rix 2005). Least squares strategies require the minimisation of the least squares error in experimental and modelled response, by changing the soil profile characteristics. Alternatively, the maximum deviation from the reference response can be minimised using a minimum-maximum formulation.

In this study an alternative approach is consider to characterise soil profiles by directly mapping the effective phase velocity to the soil properties. A map is first constructed by conducting a number of independent analyses to compute the effective phase velocity for randomly chosen soil profiles. These runs are then used to correlate information between the effective phase velocity and soil characteristics to obtain a lower dimensional description of the simulated data. Linear regression is applied to the lower dimensional description of the effective phase velocity to the lower dimensional description of the soil characteristics. Once the direct mapping has been done, the inversion of experimental field data is much faster than conventional inversion techniques.

There are practical benefits for fast direct inverse approaches, as opposed to the conventional minimisation approach as statistical quantification of the ill-posed nature of the problem is computationally efficient. This may be used to create technology that allows the sufficiency of the experimental measurements to be statistically evaluated in real-time in the field.

In this study we conduct a simulated experiment to quantify the accuracy, robustness and degree of ill-posedness of the problem as well as the suitability of a partial least squares regression strategy.

2 FORMULATION OF THE PROBLEM

Typical geotechnical experiments measure seismic surface wave phase velocity at specific frequencies. A soil characterisation inverse analysis aims to identify the soil layer characteristics that match a model dispersion curve to the experimentally measured data. The construction of the dispersion curve requires the phase velocity - frequency relationship of every mode, which requires the roots of the secular function to be computed at every experimentally measured frequency.

The dispersion curve is numerically computed by finding the roots of the secular function to give the dispersion modes. The dispersion curve is then computed as the combined participation of the various modes (Rix 2005). To compute the modes, an analysis tool `mat_disperse,` was used to compute the roots of the secular Rayleigh function. The secular Rayleigh function S for waves, requires the computation of roots for a specific wave number k and frequency f.

The secular function S has the following form:

$$S(k,f) = \frac{\det(E(k,f))}{g(k,f)} \quad (1)$$

where E is a 2 by 2 matrix and g is a scalar function. The secular function is a complex function that takes real inputs. The complex modulus can be computed by computing the magnitude of the complex function, often referred to as the absolute value of a complex number. A *Matlab* or *Octave* implementation of the secular function is freely available (Rix 2005), and defined in the user defined function `secular.m`. However, computing the roots may prove challenging especially for inversely dispersive profiles (Wilke et al. 2014). We opted to compute the dispersion modes using Dinver (Wathelet et al. 2004) and `mat_disperse` to compute the effective phase velocity.

A virtual experiment was conducted in which a target soil profile was chosen and the reference soil profile computed using the *Matlab* analysis tool `mat_disperse`. As the root finding can be challenging and often fails to find all the modes or complete modes, an *a priori* analyses was conducted to give N observations. These observations were used to construct radial basis interpolation fields.

A soil profile typically has four independent unknowns namely the shear wave velocity (V_s), compression wave velocity (V_p), density (ρ) and thickness (t). In the virtual experiment it was assumed that each layer only has two unknowns; shear wave velocity, and thickness. The density was assumed to be known and the Poisson's ratio, which relates compression wave velocity to shear wave velocity, was also assumed to be known. The aim of the virtual experiment was to estimate the shear wave velocity profile for three soil layers from the effective phase velocity dispersion curve.

2.1 *Partial least squares regression*

Before the partial least squares regression can be constructed it is necessary to first use known soil characteristics and obtain the effective Rayleigh phase velocities. `Dinver` (Wathelet et al. 2004) was used to calculate the first seven discrete modes and `mat_disperse` to compute the effective phase velocity. The shear wave velocities were taken to range between 200 m/s and 600 m/s and the soil layer thicknesses for the upper two layers were chosen to range between 1 m and 20 m with the third layer infinitely thick. A Latin Hypercube was constructed for 1000 soil profiles $\mathbf{Y}$. These profiles were then analysed in parallel as they are completely independent, to obtain the effective phase velocity $\mathbf{X}$. `Dinver` was only able to successfully compute the dispersion modes for 556 of the 1000 profiles.

The partial least squares regression is constructed as follows:

1. Both **X** and **Y** are normalised as z-scores i.e. the data sets have zero mean and is divided by their standard deviation.

2. The user chooses the number of latent variables (modes) *M*.

3. Set $\mathbf{X_1} = \mathbf{X}$ and $\mathbf{Y_1} = \mathbf{Y}$.

4. Solve the optimisation problem in which the covariance is maximized as follows:

$$\max_{w_i} cov(\mathbf{X}_1\mathbf{w}_1, \mathbf{Y}_1), \text{ subject to } \mathbf{w}_1^{\mathrm{T}}\mathbf{w}_1 = 1 \qquad (2)$$

5. Compute $\mathbf{X}_2 = \mathbf{X}_1 - \mathbf{t}_1\mathbf{p}_1^{\mathrm{T}}$

$$\min_{p} \left\| \mathbf{X}_1 - \mathbf{t}_1\mathbf{p}^{\mathrm{T}} \right\|^2, \text{ with } \mathbf{t}_1 = \frac{\mathbf{X}_1\mathbf{w}_1}{\left\|\mathbf{X}_1\mathbf{w}_1\right\|} \qquad (3)$$

6. Compute $\mathbf{Y}_2 = \mathbf{Y}_1 - \mathbf{t}_1\mathbf{c}_1^{\mathrm{T}}$

$$\min_{c} \left\| \mathbf{Y}_1 - \mathbf{t}_1\mathbf{c}^{\mathrm{T}} \right\|^2, \qquad (4)$$

7. Repeat steps 4 - 6 using $\mathbf{X}_k$ and $\mathbf{Y}_k$ for k = 2, 3,..*M*

8. Construct **T** that consists of $\mathbf{t}_k$, **W** that consists of $\mathbf{w}_k$, **P** that consists of $\mathbf{p}_k$ and **C** that consists of $\mathbf{c}_k$.

9. For linear regression $\mathbf{x}\beta$, the regression coefficients $\beta = \mathbf{W}(\mathbf{T}^{\mathrm{T}}\mathbf{X}\mathbf{W})^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{y}$.

The procedure outlined above is packaged in *Matlab* under the function `plsregress`. This allows for the efficient and convenient construction of inverse maps.

The only unknown parameter is the number of components to use in the partial least squares regression.

Computing the mean difference between the desired profile and the simulated profiles ranks the data points. We used responses that had the lowest difference. The number of points depends on the number of components and we used 4 times the number of components as the number of points used in the partial least squares regression. The reason is that partial least squares regression is a linear map and therefore not very flexible and by focussing the data around the solution a more accurate specialised map is obtained as opposed to an inaccurate generic map.

3 NUMERICAL EXPERIMENT

The soil profile investigated in the numerical experiment is shown in Table 1, with the Rayleigh modes and effective phase velocity depicted in Figure 1. `Dinver` was used to compute the fundamental modes and `mat_disperse` is used to compute the effective phase velocity.

Table 1. Initial profile parameters.

Thickness m	Vs m/s	Vp m/s	Density kg/m^3
5	350	600	1800
10	400	700	1800
∞	450	800	1800

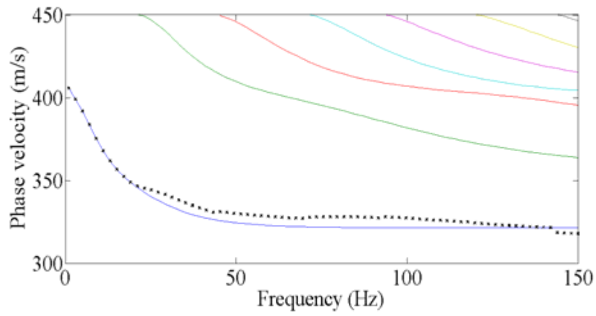

Figure 1. Rayleigh modes and effective phase velocity for the soil profile.

The sampled frequencies were 1; 3; 5; 7; 9; 12; 15; 18; 22; 26; 30; 34; 38; 42; 46; 50; 55; 60; 65; 70; 80; 90; 100; 125; 150 Hz. A short geophone array of five geophones with spacing 0.75m, typical of continuous surface wave (CSW) testing, were used with the geophones located at 1.5; 2.25; 3; 3.75 and 4.5 metres from the source to compute the effective phase velocity.

1000 independent simulations were conducted using Latin Hypercube sampling assuming the shear wave velocity to range between 200 m/s and 600 m/s and the soil layer thicknesses between 1 m and 20 m. `Dinver` was able to solve only 556 of the profiles. Figure 2 shows every tenth response as well as the effective phase velocity we aim to recover as a solid line. In addition Figure 3 depicts the variation of each variable as a box and whisker plot, with a range of one standard deviation and mean of 0.5 together with the solution we aim to recover.

Three simulated experiments were conducted:
1. Recover only shear wave velocity, given the rest of the profile.
2. Recover only thickness, given the rest of the profile.
3. Recover both shear wave velocity and thickness.

The aim of experiment 1 was to recover the shear wave velocity for each layer. Poisson's ratio is assumed to be known to compute the compression wave velocity. The layer thickness as well as the density (1800kg/m^3) was assumed to be known.

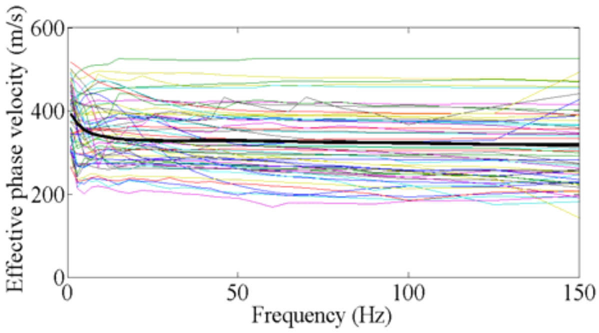

Figure 2. Response samples generated.

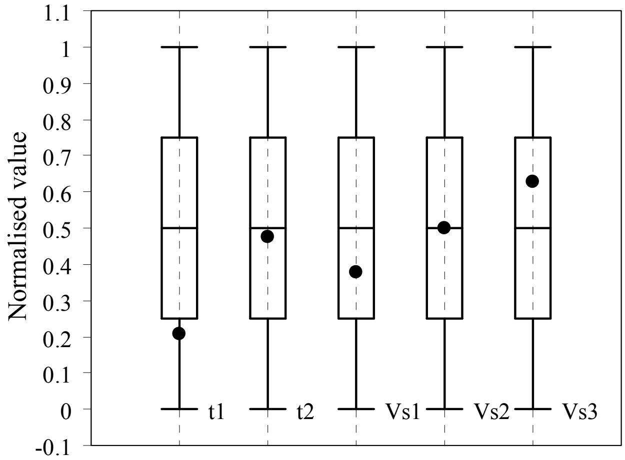

Figure 3. Variation of each variable for the generated samples.

Table 2 shows the shear wave velocity recovered from experiment 1 and Figure 4 shows the effective phase velocities for using 1 to 5 components in the partial least squares regression.

Table 2. Recovered shear wave velocity.

Components	Vs 1 m/s	Vs 2 m/s	Vs∞ m/s
1	348.27	416.82	443.48
2	348.04	403.70	446.88
3	348.50	404.21	451.10
4	349.17	403.21	451.51
5	348.79	403.86	449.82
Solution	350	400	450

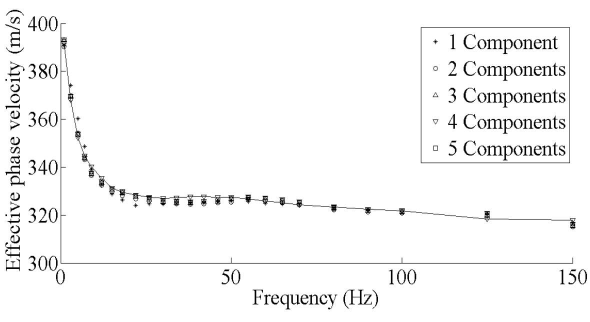

Figure 4. Response from the estimated shear wave velocities.

For experiment 2 all the soil parameters were assumed to be known except the thickness of layer 1 and layer 2. Table 3 shows the thickness recovered and Figure 5 shows the effective phase velocities.

Table 3. Recovered thickness.

Components	Thickness 1 m	Thickness 2 m
1	4.81	9.97
2	4.84	9.92
3	4.83	10.69
4	4.83	10.93
5	4.86	11.10
Solution	5	10

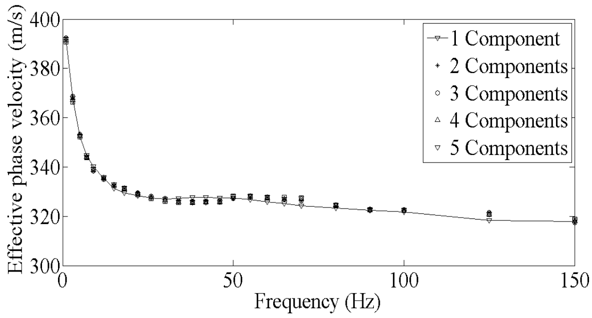

Figure 5. Response from the estimated layer thicknesses.

The aim with virtual experiment 3 was to recover the layer thickness for the top two layers and the shear wave velocity for all three layers. Poisson's ratio was assumed to be known to compute the compression wave velocity and the density was assumed to be known and taken as 1800kg/m^3. Table 4 shows the recovered layer thicknesses and shear wave velocities and Figure 6 show the effective phase velocities using 1 to 5 components in the partial least squares regression

Table 4. Recovered shear wave velocity and thickness.

Comp	Vs 1 m/s	Vs 2 m/s	Vs∞ m/s	Thick 1 m	Thick 2 m
1	350.95	401.62	441.08	5.31	15.92
2	347.32	389.45	443.60	5.71	10.66
3	345.48	414.92	446.53	7.40	11.15
4	344.19	439.04	413.33	4.96	8.04
5	351.74	445.72	430.57	7.36	13.80
Solution	350	400	450	5	10

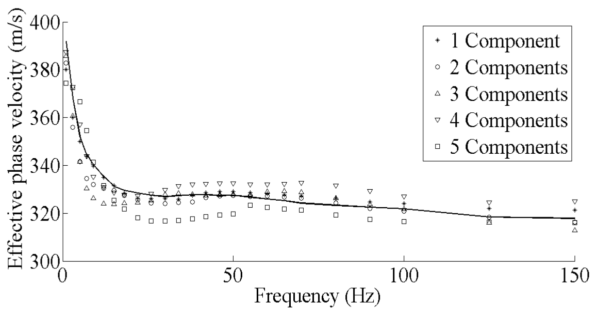

Figure 6. Response from the estimated shear wave velocities and layer thicknesses.

4 DISCUSSION

The main advantage to use the direct mapping inverse approach is the time required for the inversion analysis. Once the map has been constructed, the inversion analysis to find the ground profile which best fits the measured dispersion data, takes about 17 micro seconds using a 2.4 GHz Intel Core i5 processor. Using a conventional inversion analysis approach typically takes a few minutes using the same processor. Such fast inversion analysis may allow technology to be developed which allow inversion to be conducted in the field.

Table 5 shows the results for experiment 1 which attempted to recover the shear wave velocities for the three layer model. The shear wave velocities recovered by the inversion analysis were within approximately 1% of the true value for all layers if more than two components are used for the least squares regression. Figure 4 show that the calculated effective phase velocities closely matched the true effective phase velocity of the ground profile.

Table 5. Error in recovered shear wave velocity.

Components	Vs 1 %	Vs 2 %	Vs∞ %
1	-0.49	4.21	-1.45
2	-0.56	0.93	-0.69
3	-0.43	1.05	0.25
4	-0.24	0.80	0.34
5	-0.34	0.97	-0.04

The error of the recovered layer thicknesses for virtual experiment 2 are shown in Table 6. The layer thickness recovered by the least squares regression for layer 1 was within 4% of the true layer thickness, but for layer 2 it was 11%. This appears to suggest that the accuracy with which layer thickness can be determined is less than the accuracy with which shear wave velocity can be determined. In addition the accuracy appears to be better for shallow layers. No advantage is evident for increasing the number of components of the least squares regression.

Table 6. Error in recovered layer thickness.

Components	Thickness 1 %	Thickness 2 %
1	-3.66	-0.28
2	-3.12	-0.73
3	-3.33	6.97
4	-3.25	9.30
5	-2.65	11.03

Table 7 shows the error when attempting to simultaneously recover both shear wave velocity and layer thickness. It may be observed that the accuracy is worse in both cases compared with attempting to recover only shear wave velocity or layer thickness. The accuracy with which the shear wave velocity was recovered was 11% as opposed to 1% when the layer thicknesses were known. The accuracy with which the layer thicknesses were recovery reduced to approximately 60% compared with 11% when the shear wave velocities were known. Again, no advantage is evident for increasing the number of components of the least squares regression.

Figure 6 shows the effective phase velocities when both the layer thickness and wave velocity is recovered. It indicates that the scatter of the calculated wave velocities are significantly more than when either the layer thickness is known (Figure 4), or shear wave velocity is know (Figure 5).

Table 7. Error in recovered shear wave velocity and thickness.

Comp	Vs 1 %	Vs 2 %	Vs∞ %	Thick 1 %	Thick 2 %
1	0.27	0.41	-1.98	6.27	59.21
2	-0.76	-2.64	-1.42	14.25	6.65
3	-1.29	3.73	-0.77	48.10	11.56
4	-1.66	9.76	-8.15	-0.67	-19.51
5	0.50	11.43	-4.32	47.26	38.09

5 CONCLUSIONS

A number of conclusions may be drawn from the inversion of effective phase velocity seismic surface wave data by partial least squares regression. The technique applies a direct mapping inverse approach and once the map has been constructed the inversion analysis to find the ground profile which best fits measured dispersion data is significantly faster than conventional inversion techniques.

The results suggest that if *a priori* information is available regarding the thickness of the soil layers for a given profile, the shear wave velocity of the layers may be computed with a high degree of accuracy. However, simultaneous inversion of both layer thickness and shear wave velocity reduces the computational accuracy.

The efficiency with which the inverse analysis can be computed using the direct inverse map, will allow the variation of the recovered ground profile parameters to be quantified statistically.

REFERENCES

Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors, Tokyo University, Bulletin of the Earthquake Research Institute, Vol. 25, pp.415-457.

Foti, S., Lai C. G., Rix G. J. and Strobbia, C. (2015). Surface Wave Methods for Near-Surface Site Characterization. CRC Press, Boca Raton, pp. 467.

Heymann, G (2013). Vibratory sources for continuous surface wave testing. Geotechnical and Geophysical Site Characterization (ISC'4), Coutinho & Mayne (eds). Taylor and Francis Group, London, pp.1381-1386.

Louie, J.L., 2001, Faster, Better: Shear-wave velocity to 100 meters depth from refraction microtremor arrays, Bulletin of the Seismological Society of America, Vol. 91, 347-364.
Matthews, M.C., Hope, V.S. and Clayton, C.R.I. (1996). The use of surface waves in the determination of ground stiffness profiles. Proceedings of the Institution of Civil Engineers: Geotechnical Engineering, Vol.119, pp.84-95.
Park, C. B., Miller, R. D., and Xia, J. (1999). Multichannel analysis of surface waves: Geophysics, v. 64, p.800-808.
Rix, G. (2005). Surface Waves in Geomechanics: Direct and Inverse Modelling for Soils and Rocks. Chapter: Near-Surface Site Characterization Using Surface Waves, Volume 481. CISM International Centre for Mechanical Sciences.
Stokoe, K.H., Joh, S. and Woods, R. D. (2004). Some contributions of in situ geophysical measurements to solving geotechnical engineering problems. Proceedings ISC-2 on Geotechnical and Geophysical Site Characterization. Viana da Fonseca & Mayne (eds.), Millpress, Rotterdam, pp.97-132.
Stokoe, K.H., Wright, S.G., Bay, J.A. and Roesset, J.M. (1994). Characterization of geotechnical sites by SASW method. Geophysical Characterization of Sites (ISSMFE Technical Committee no. 10) by R.D. Woods, Oxford and IBH Publications, pp. 15-25.
Wathelet M., Jongmans D. and Ohrnberger M. (2004). Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements. Near Surface Geophysics, pp. 211-221.
Wilke, D.N., Kok, S. and Heymann, G. (2014). Comparison of two inverse strategies to characterize soil profiles, Proceedings of the 4th International Conference on Engineering Optimization (EngOpt2014), pp. 1005-1010, Lisbon, Portugal, 8-11 September, 2014. ISBN: 978-1-138- 02725-1.