p-ISSN 2089 – 0133 e-ISSN 2477-6416
URL : https://jurnal.uns.ac.id/ijap/article/view/77921

DOI : https://doi.org/10.13057/ijap.v13i2.77921

A WEIGHTED AVERAGE OF MULTIPLE INVERSIONS OF RAYLEIGH WAVE DISPERSION CURVE USING PARTICLE SWARM OPTIMIZATION FOR GEOTECHNICAL SITE CHARACTERIZATION

Jamhir Safani*[1], Rezki Wirawan[1], Al Rubayn[1], Mohd Nawawi[2], Toshifumi Matsuoka[3]

[1] Dept. of Geophysical Engineering, Faculty of Mathematics and Natural Sciences, Universitas Halu Oleo, Indonesia
[2] Department of Geophysics, School of Physics, Universiti Sains Malaysia, Malaysia
[3] Fukada Geological Institute, Tokyo 113-0021, Japan
*jamhir.safani@uho.ac.id

Received 15-08-2023, Revised 21-10-2023, Accepted 23-10-2023
Available Online 23-10-2023, Published Regularly October 2023

ABSTRACT

Shear wave velocity is an important parameter in geotechnical engineering for studying liquefaction, finding bedrock for the basement of a building, and figuring out the presence of subsurface cavities. This study aims to develop and evaluate the accuracy of the multiple inversions by the Particle Swarm Optimization (MI-PSO) algorithm with a weighted average solution. This algorithm is applied to Rayleigh wave dispersion data for geotechnical site characterization. Two synthetic models, the HVL model and the complex model (i.e., a combination of models with LVL and HVL characteristics), are used to conduct algorithm tests. These synthetic models replicate subsurface characteristics that are frequently encountered in geotechnical cases. Synthetic data tests show that the MI-PSO algorithm with a weighted average solution works excellently. The MI-PSO technique with a weighted average solution resolves the model better than the conventional average solution. When applied to two field data sets, the MI-PSO algorithm with a weighted average solution can delineate target models that are consistent with the qualitative interpretation based on the observed dispersion curve characteristics.

Keywords: MI-PSO; weighted average; dispersion curve, Rayleigh wave, geotechnical site

Cite this as: Safani, J., Wirawan, R., Rubayn, A., Nawawi, M., & Matsuoka, M. 2023. A Weighted Average of Multiple Inversion of Rayleigh Wave Dispersion Curve Using Particle Swarm Optimization for Geotechnical Site Characterization. *IJAP: Indonesian Journal of Applied Physics*, *13*(2), 347-361. doi: https://doi.org/10.13057/ijap.v13i2.77921

INTRODUCTION

Shear wave velocity is a crucial parameter in geotechnical engineering for studying liquefaction, finding bedrock for the basement of a building, and figuring out the presence of subsurface cavities[1]. Several methods have been developed and utilized to estimate the shear wave velocity profile, including spatial autocorrelation (SPAC)[2], spectral analysis of surface waves (SASW)[3], and multichannel analysis of surface waves (MASW)[4].

In recent decades, the MASW has been commonly used to estimate subsurface shear wave velocity (*Vs*) profiles. Rayleigh wave dispersion describes the characteristic of Rayleigh wave phase velocity that changes with frequency[3]. The MASW utilizes this dispersion property to determine the shear wave velocity (*Vs*) profile[4]. The MASW method generally

consists of three stages: surface wave data acquisition, dispersion curve extraction, and dispersion curve inversion[5]. From the obtained dispersion curve, a profile of the shear wave velocity (*Vs*) was derived by inverting the Rayleigh wave dispersion curve. An inversion of the observed Rayleigh wave dispersion curve is carried out to obtain a shear wave velocity (*Vs*) profile[6].

Inversion of the surface wave dispersion curve (i.e., Rayleigh waves and Love waves) utilizing the local search method presents optimization challenges due to the nonlinear and multimodal nature of the dispersion[7-8]. Inversion using a linear approach is only effective when there is solid a priori information and a good initial model[9-10]. Due to its nonlinear nature, the inversion of the Rayleigh wave dispersion curve is typically conducted using nonlinear or global optimization techniques. Genetic algorithms[11-12], simulated annealing[13–15], and particle swarm optimization are three common global optimization techniques used for inverting the Rayleigh dispersion curve.

Particle swarm optimization (PSO) is a population-based optimization technique inspired by flocks of birds and fish. Potential model solutions, known as particles in PSO, move around the search space by following the best solution produced at each iteration[16]. The PSO method continues to develop and spawn numerous variants. PSO techniques have been extensively used to solve geophysical problems[17], including inversion of gravity data[18], geoelectric data[19-20], and magnetic data[21]. A number of PSO variants have been used to solve the inversion problem of the Rayleigh wave dispersion curve. Song et al.[8] used a PSO method with a constriction factor and showed that when inverting the Rayleigh wave dispersion curve, the PSO variant demonstrates a very good model solution. Schutte and Groenwold[22] found that the PSO variant with constriction is very effective for solving low-dimensional inversion problems, but for high-dimensional problems, the PSO variant with dynamic inertia reduction and maximum speed limitation performs better. Because inversion of the Rayleigh wave dispersion curve is a high-dimensional inverse issue, the PSO technique with dynamic inertia reduction is used in this study[23].

Global optimization methods like PSO generate different model solutions for each inversion, although they use the same search space. Multiple inversions (MI) can indicate a model solution trend by averaging all solutions. However, model solutions with high misfits and vice versa will affect the accuracy of the final model solution. Thus, this work investigates whether weighted averaging of MI-PSO model solutions improves model solution accuracy compared to standard averaging.

METHOD

Particle Swarm Optimization

Kennedy and Eberhart[24] were the first to introduce particle swarm optimization (PSO). In the PSO technique, candidate solutions to optimization problems are depicted as flocks of particles moving through the search parameter space, with paths determined by the optimal performance of the particles themselves and the particles surrounding them. In the search for the optimal solution to a problem, the path of a particle through the parameter space (which is always changing position) is determined by the equations of motion:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \quad (1)$$

where t and t +1 denote two successive iterations of the algorithm, and $\mathbf{v}_i$ is the velocity vector of the i-th particle in dimension D.

The velocity vector determines how the particle moves along the search space, and it is determined by three factors: first, inertia, which ensures that the direction of motion of the particle does not change drastically; second, the cognitive component, which determines the tendency of the particle to return to the best position it has passed before; and third, the social component, which identifies the tendency of the particle to move to the best position of the entire herd. The velocity of the i-th particle is determined by the Equation:

$$\mathbf{v_i}(t+1) = \mathbf{v_i}(t) + \mathbf{c1}\big(\mathbf{p_i} - \mathbf{x_i}(t)\big)\mathbf{R_1} + \mathbf{c2}\big(\mathbf{g} - \mathbf{x_i}(t)\big)\mathbf{R_2} \qquad (2)$$

where $\mathbf{p_i}$ is the "personal best" of the particles, i.e., the best solution obtained during optimization by a given individual, and **g** is the "global best," i.e., the overall best solutions found by the flock. The acceleration constants **c1** and **c2** are real numbers that are called, respectively, the "cognitive coefficient" and the "social coefficient." $\mathbf{R_1}$ and $\mathbf{R_2}$ are two diagonal matrices of random numbers distributed uniformly between [0, 1]. These two social and cognitive factors influence the velocity change in Equation (2). Overall, the iteration procedure of Equations (1) and (2) is performed until the conditions are satisfied or the maximum iteration limit is reached [16].

Shi and Eberhart [25] added the inertial weight parameter (**w**) to Equation (2) to balance the PSO's local and global search capabilities, resulting in the particle velocity being represented as:

$$\boldsymbol{v_i}(t+1) = \boldsymbol{w}\boldsymbol{v}_i(t) + \boldsymbol{c1}\big(\boldsymbol{p_i} - \boldsymbol{x_i}(t)\big)\mathbf{R_1} + \boldsymbol{c2}\big(\mathbf{g_i} - \boldsymbol{x_i}(t)\big)\mathbf{R_2} \qquad (3)$$

To reduce the size of the search step when a particle's position in the search space changes, the maximum velocity of each particle ($\mathbf{v}^{max}$) is regulated by a velocity limiting coefficient (γ) [25]:

$$\boldsymbol{v^{max}} = \gamma(\boldsymbol{x}_{UB} - \boldsymbol{x}_{LB}) \qquad (4)$$

where $\mathbf{x}_{UB}$ and $\mathbf{x}_{LB}$ are the upper and lower bounds of the search space for each particle, respectively.

Reducing the inertia weight (**w**) as the iteration advances is supplied to modify the process of determining the optimum value from global search to tend to be local search in order to optimize the optimum solution. The inertial damping coefficient (α) with a value between 0 and 1 (or $0<\alpha<1$) is given to lower the initial inertia weight $\mathbf{w}^0$ when the solution obtained does not increase over a range of successive iterations (h)[26]:

$$\text{If } f(\boldsymbol{g}(t) \geq \boldsymbol{g}(t-h) \text{ , then } \boldsymbol{w}_{t+1} = \alpha \boldsymbol{w}_t \qquad (5)$$

Multiple Inversions by PSO (MI-PSO) with a Wighted Model Solution

Inversion of the observed Rayleigh wave dispersion curves is performed to find the best combination of the shear wave velocities (*Vs*) and the corresponding thicknesses (*d*) for each subsurface layer. This is done by minimizing the difference between the measured dispersion curves and the theoretical dispersion curves. The forward modeling method incorporated to calculate the theoretical Rayleigh wave dispersion curve is being adopted from the stiffness matrix method developed by Kausel and Roësset[27] and Olafsdottir[28]. The objective function used to calculate the misfit between the observed and theoretical dispersion curves is expressed as follows:

$$Misfit = \frac{1}{N}\sum_{i=1}^{N}\frac{\sqrt{\left(C_i^{Obs}(f)-C_i^{Calc}(f)\right)^2}}{C_i^{Obs}(f)} \times 100 \quad (6)$$

where $C_i^{Obs}(f)$ and $C_i^{Calc}(f)$ are the observed and the theoretical (calculated) Rayleigh wave phase velocity, respectively, as a function of frequency, and N is the number of the observed frequency points.
Calculation of the similarity index (SI) uses Equation (7):

$$\mathrm{SI} = \left(1 - \frac{\sum_{m=1}^{M}\frac{\left|p_m^{cal}-p_m^{obs}\right|}{p_m^{obs}}}{M}\right) \times 100\% \quad (7)$$

where p_m^{cal} is the model parameter resulting from the inversion and p_m^{obs} is the true model parameter.

Inversion of the Rayleigh wave dispersion curve with the PSO algorithm involves four types of particles, which represent the physical properties of each subsurface layer, namely layer thickness (d), shear wave velocity (V_S), compression wave velocity (V_P), and density (ρ). These four physical parameters have different levels of sensitivity to the Rayleigh wave phase velocity as the response model, where the shear wave velocity (V_S) and layer thickness (d) are the two most sensitive parameters, respectively[29]. Therefore, inversion of the Rayleigh wave dispersion curve only estimates these two parameters, while compression wave velocity (V_P) and density (ρ) are assumed from a priori information[29].

Here, we propose the multiple inversions by particle swarm optimization (MI-PSO) algorithm. This algorithm was developed by adopting a single inversion algorithm by Raha[23], but we added two other steps, namely an inversion loop and a weighted average solution. The algorithm procedure for the proposed MI-PSO is as follows:

1. Determine the upper limit ($\mathbf{x_{UB}}$) and lower limit ($\mathbf{x_{LB}}$) of the particles, namely the upper and lower limits of the shear wave velocity (Vs) and layer thickness (d) combined in the matrices as follows: $\mathbf{x_{UB}} = \left[V_{S_{UB1}} \ldots V_{S_{UBD+1}}\ d_{UB1} \ldots d_{UBD}\right]$, $\mathbf{x_{LB}} = \left[V_{S_{LB1}} \ldots V_{S_{LBD+1}}\ d_{UB1} \ldots\ d_{UBD}\right]$ (where D is the number of finite thickness layers). In addition, the following quantities are also determined: the number of particles per population M, inertial weight (**w**) = 1, inertial damping coefficient (α) = 0.99, personal acceleration coefficient (**c1**) = 2, global acceleration coefficient (**c2**) = 2, velocity limiting coefficient (γ) = 0.05, maximum number of inversions I_{max}, and maximum number of iterations k_{max};
2. Loop the inversion I from 1 to I_{max};
3. Initialize the initial population $\mathbf{x}_i$ randomly with an upper bound $\mathbf{x}_{UB}$ and a lower bound $\mathbf{x}_{LB}$. Determine the initial velocities of all the particles with $\mathbf{v^0} = 0$. Calculate the misfit of each particle. Initialize the best position of each particle ($\mathbf{p}_i$) from the initial position of each particle. Select the best global position (**g**) of the particle with the best misfit;
4. Loop the iteration t from 1 to kmax, where in each iteration, the following steps are carried out:
   - Update the particle velocity using Equation (3) with the value of the maximum velocity (vmax) according to Equation (4);
   - Update the position of each particle using Equation (2);
   - Update the inertial weight (w) with Equation (5) at each iteration;

- Evaluate the position in each iteration if $\mathbf{x}_{\mathrm{i}cost}(t) < \mathbf{p}_{\mathbf{i}cost}$, update $\mathbf{p}_{\mathbf{i}} = \mathbf{x}_{\mathrm{i}}$, if $\mathbf{x}_{\mathrm{i}cost}(t) < \mathbf{g}_{cost}$ update $\mathbf{g} = \mathbf{x}_{\mathrm{i}}(t)$;

5. The model solution obtained in the I-th inversion is the global model solution (g) at t = k_{max}, where the misfit is $\mathbf{g}_{cost}$;
6. Perform averaging of the inversion model solutions with a weighting coefficient (1/misfit) to obtain a global solution with the following Equation:

$$\mathbf{g} = \frac{\sum_{\mathrm{I}=1}^{\mathrm{I}max}\left(\mathbf{g}_{\mathrm{I}} * \frac{1}{misfit_{\mathrm{I}}}\right)}{\sum_{\mathrm{I}=1}^{\mathrm{I}max}\left(\frac{1}{misfit_{\mathrm{I}}}\right)} \tag{7}$$

RESULTS AND DISCUSSIONS

Inversion for Synthetic Data of HVL Model

Multiple inversions (MI) of the Rayleigh wave dispersion curve with PSO begin by utilizing synthetic data. This is intended to test the accuracy and feasibility of the developed MI-PSO algorithm with a weighted average solution. There are two subsurface profile models used in this synthetic test, namely the high-velocity layer (HVL) model, which is characterized by the presence of a hard layer between the soft layers (Figure 1a), and the complex model, which represents the presence of a low-velocity layer (LVL) and HVL simultaneously in the subsurface medium (Figure 1b). These two synthetic models mimic real models that are often found in the field. The existence of a low-velocity layer (LVL) in both synthetic models can be a trap in geotechnical site investigations if it is not identified properly. These two synthetic models refer to Safani (2007) for the complex model and La Hamimu et al. (2011) for the HVL model. In geotechnical engineering, the stiffness of the subsurface layer is represented by the shear wave velocity parameter (*Vs*).

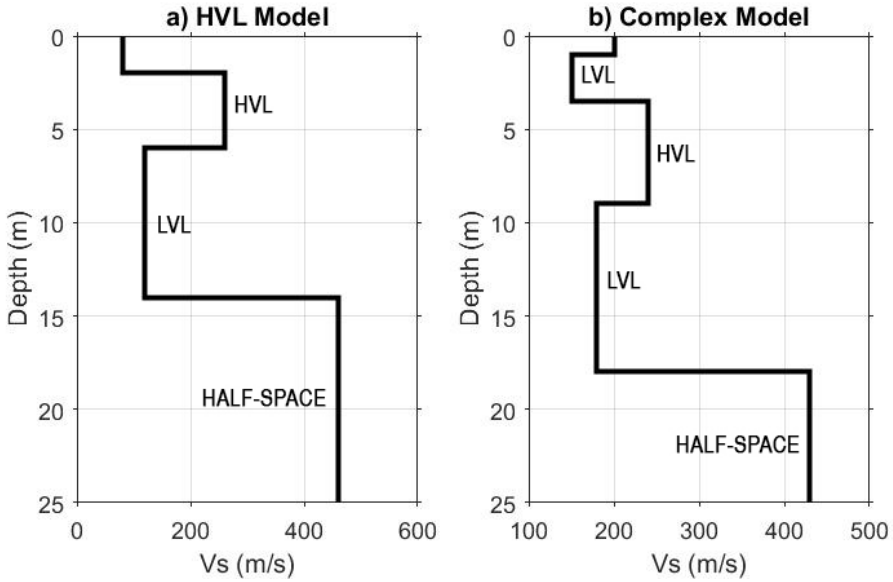

Figure 1. Subsurface profile of *Vs* versus depth: a) for HVL model and b) for complex model

The HVL model shown in Table 1 is characterized by the presence of a hard layer (with *Vs* = 260 m/s) in the second layer, which is flanked by two layers with lower *Vs* values. The inversions of the Rayleigh wave dispersion curve begin with determining the search space by setting the upper and lower bound values of the model space (Table 1).

Table 1. Parameters for HVL model and inverse search spaces

Layers	Parameters for HVL Model				Search Spaces	
	Vs (m/s)	*Vp* (m/s)	ρ (g/cm^3)	*d* (m)	*Vs* (m/s)	*d* (m)
1	80	370	1.8	2	75-150	1-3
2	260	600	1.8	4	150-300	2-6
3	120	700	1.8	8	100-200	4-10
4	460	700	1.8	*Half space*	400-550	*Half space*

The inversions for this first case were carried out 10 times, in which each inversion was done for 500 iterations with 30 particles or populations in the search space. Choosing the number of inversions of 10 times is ideal for seeing trends in changes in inversion results and quite saves computing time. The number of iterations of 500 in inversion for the HVL model is a test to check the convergence characteristics of the proposed algorithm. A small number of particles can lead the model solution to become stuck in local minima because convergence occurs too quickly. This is why, after several trial and error, 30 particles were chosen. The inversion results are shown in Figure 2, which describes the observed versus theoretic dispersion curves (Figure 2a) and 1D profile shear wave velocity (*Vs*) versus depth (Figure 2b) for conventional average model (red dotted line) and weighted average model (green line). Figure 1b shows that the profiles of *Vs* versus depth for the 10 resulting model solutions can very well reconstruct the actual model profile, especially the main target of the HVL pattern, even with slightly varying *Vs* and thickness values.

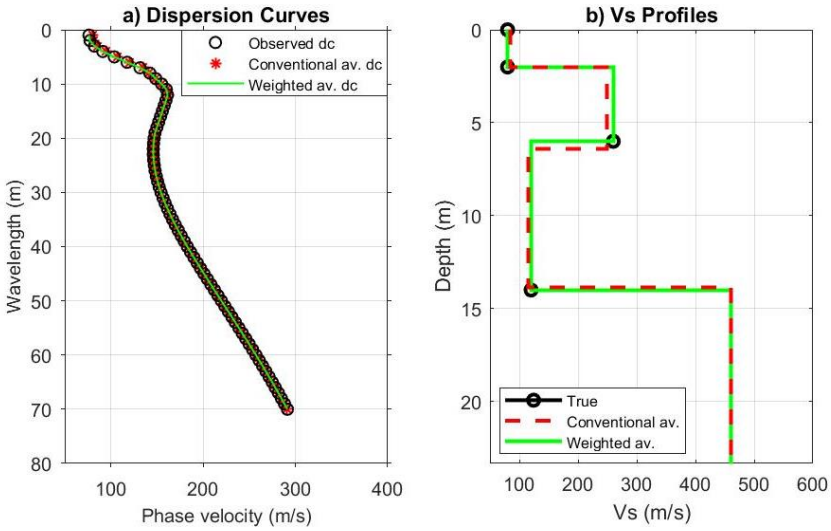

Figure 2. Inversion results for synthetic data of HVL model: a) Dispersion curves, and b) *Vs* profiles

The final model solution of the 10 resulting inversion models is obtained by calculating the conventional average model and the weighted average model (Table 2). In the conventional average model, the model parameters (*Vs* and *d*) of each layer show good results, but there are still some significant discrepancies, especially in the model parameters with a relative error of > 3%. Based on the trend of the resulting model solution, model parameters with a relative error >3% are still acceptable, even though they have a slightly lower accuracy compared to model parameters with a relative error <3%. In the weighted average model, the final model solution obtained from the weighted average is better than the model solution

obtained from the conventional average, where the resulting model profile closely coincides with the actual model (green solid line). The weighted average model's relatively small error also supports this accuracy. The degree to which the model solution obtained resembles the original model is indicated by the similarity index (SI). Both the weighted average model and the conventional average model have a similarity index greater than 95%. This indicates that there is a strong similarity between the two model solutions and the actual model. The weighted average model is marginally better than the traditional average model, as indicated by the difference in SI values of about 4%.

Table 1. Conventional averaging model and weighted average for synthetic data of HVL model

Model Parameters	True Model	Conventional average model			Weighted average model		
		Calculation	Relative error (%)	SI (%)	Calculation	Relative error (%)	SI (%)
Vs_1	80	83.99	4.99		80.00	0.01	
Vs_2	260	249.05	4.21		260.41	0.16	
Vs_3	120	115.29	3.92		119.99	0.01	
Vs_4	460	459.86	0.03	95.48	460.03	0.01	99.96
d_1	2	2.04	1.89		2.00	0.02	
d_2	4	4.38	9.43		4.00	0.05	
d_3	8	7.43	7.15		8.00	0.04	

Figure 3 shows the trend of changes in misfits over iterations for the synthetic data of the HVL model. From the 10 inversions performed, it appears that the misfit values decrease exponentially with iteration. The misfit values for all inversions still show a very random pattern until the 20th iteration, indicating that the inversion process is in the global search stage. At higher iterations, misfit values tend to converge and stabilize. This indicates that the algorithm is starting to be exploitative or local. Even though we set the number of iterations to 500, the misfit values reach a highly convergent pattern around the 100th iteration. In Figure 3, it can also be seen that the 4th inversion has not experienced a significant change in the misfit value since the beginning of the iteration. This might happen because the global minima from the initial model that was generated has been trapped in the local minima. This shows that multiple inversions are important to ensure that the model solution obtained is not stuck at the local minima.

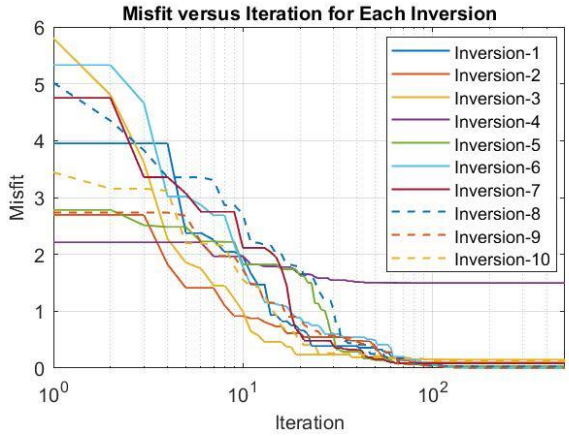

Figure 3. Misfit as a fungtion of iteration for inversion of dispersion curve for HVL model.

Inversion for Synthetic Data of Complex Model

The second synthetic test was performed on the complex model, which is a combination of LVL and HVL characteristics. The nature of the complex model is simply described by the variation pattern of *Vs* values in Table 3. The use of this mixed model is intended to examine the ability of the MI-PSO algorithm to solve more complex inverse problems. Model parameters and search spaces for the complex model can be seen in Table 3.

Table 3. Parameters for complex model and inverse search spaces

Layers	Model Parameters				Search Spaces	
	Vs (m/s)	*Vp* (m/s)	ρ (g/cm^3)	*d* (m)	*Vs* (m/s)	*d* (m)
1	200	420	1.8	1	150-400	0.5-2
2	150	600	1.8	2.5	100-300	1-3
3	240	800	1.8	5.5	150-400	4-6
4	180	1000	1.8	9	150-250	6-10
5	430	1300	1.8	*Half space*	300-500	*Half space*

The same as in the case of the HVL synthetic model, here the inversion is carried out ten times, with the number of iterations in each inversion being 500 for 30 populations. The results of the inversion of the observed dispersion curves are shown in Figure 4. The excellent matching between the observed and theoretical dispersion curves as presented in Figure 4a, indicates the accuracy of the resulting model parameters. This confirms the excellent fit between the true model and all the *Vs* versus depth profiles obtained from the inversion process (Figure 4b). All of the resulting model solutions, namely conventional (red dashed line) and weighted average model (green line), can reconstruct the complex *Vs* versus depth profiles consisting of a combination of models with LVL and HVL characteristics. This shows the excellent ability of the MI-PSO algorithm to solve the inverse problem of the Rayleigh wave dispersion curve for complex subsurface models.

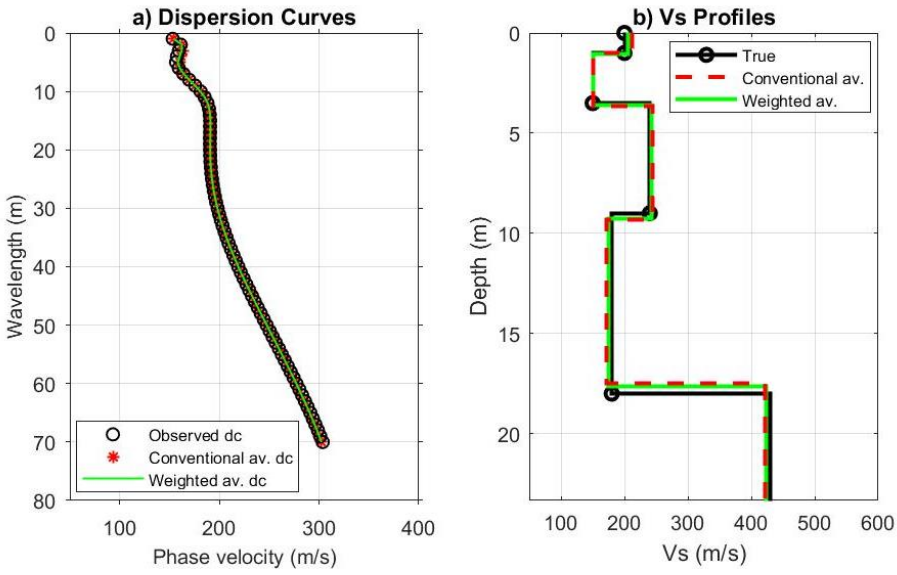

Figure 4. Inversi results for synthetic data of complex model: a) Dispersion curves, and b) *Vs* profiles**.**

The final model solution of *Vs* versus depth obtained from the conventional and weighted averages of all inversion model solutions is shown in Table 4. In general, the relative errors of the weighted average model are less than those of the conventional one. Although the difference is not significant, the similarity index (SI) of the weighted average model is slightly better than the conventional average model, namely 97.09%: 96.15%. These results reaffirm the superiority of the weighted average solution over the conventional one.

The trend of the misfit values obtained during the inversion process for the complex model is generally similar to that for the HVL synthesis model (Figure 5), where the misfit values change significantly around the initial 20 iterations. This shows the nature of the search algorithm, which still tends to be exploratory or global. As in the HVL synthesis model, the misfit values begin to change slowly at higher iterations and towards a convergent trend, especially around the 100th iteration. So, the search algorithm slowly becomes exploitative at around the 100th iteration.

Table 4. Conventional and weigthed average model for the synthetic data of complex model

Model Parameters	True Model	Conventional average model			Weighted average model		
		Calculation	Relative error (%)	SI (%)	Calculation	Relative error (%)	SI (%)
Vs_1	200	212.44	6.22	96.15	204.84	2.42	97.09
Vs_2	150	151.51	1.01		151.16	0.77	
Vs_3	240	244.43	1.85		242.99	1.25	
Vs_4	180	172.37	4.24		174.62	2.99	
Vs_5	430	422.32	1.79		424.05	1.38	
d_1	1	1.02	1.52		1.04	4.32	
d_2	2.5	2.65	6.03		2.58	3.15	
d_3	5.5	5.66	2.95		5.66	2.88	
d_4	9	8.18	9.07		8.36	7.07	

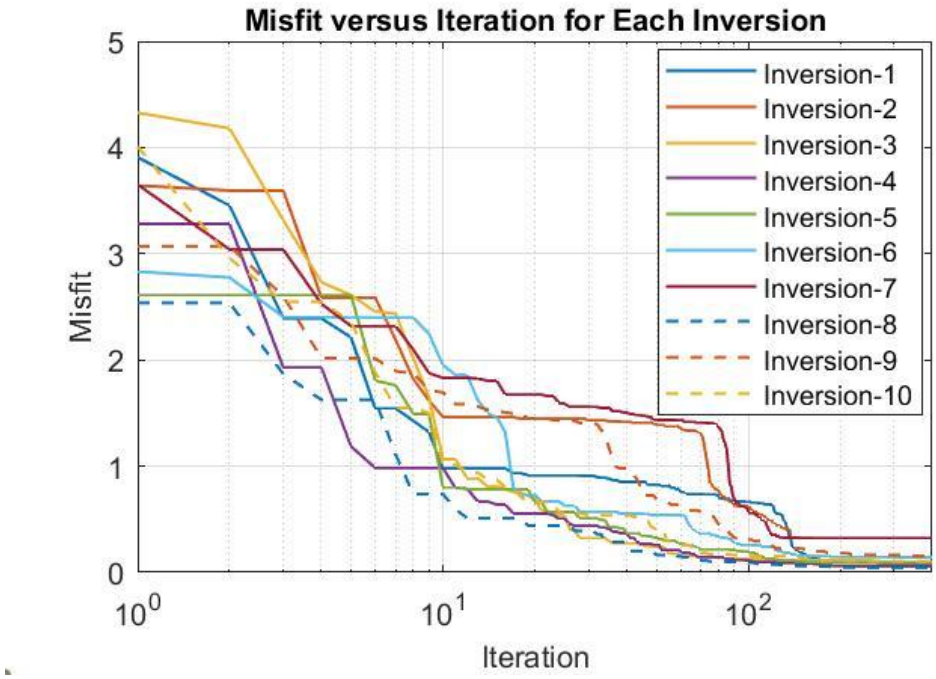

Figure 5. Misfit as a fungtion of iteration for inversion of dispersion curve for complex model.

Inversion for Rayleigh Wave Data from Archeological Site of Lahad Datu (Malaysia)

The black dots in Figure 6a show measured Rayleigh wave dispersion data in Lahad Datu, Malaysia. This site is particularly intriguing since it is an archaeological site that is supposed to be a place of ancient human life. The investigation's goal is to estimate the subsurface structure at the site. The observed dispersion curve shows a trend of a significant decrease in phase velocity over the wavelength range around 2–6 m and a slight change at 16–20 m. Changes in the curve trend with such a profile indicate the presence of two soft layers beneath the surface. This initial assumption will be verified in the inversion process.

The search space used in the inversion of the observed Rayleigh wave dispersion curve is shown in Table 5. It is assumed that the compression wave velocity (*Vp*) and density (ρ) parameters for each layer are known, as presented in Table 5. Because the shear wave velocity (*Vs*) and thickness (*d*) are two parameters that are estimated through the inversion process, each of these parameters is given a search space. The search space is determined based on the characteristics of the existing phase velocity and wavelength.

Table 5. Search space for inversion of dispersion data at the archeological site of Lahad Datu

Layers	Search Spaces			
	Vp (m/s)	ρ (g/cm^3)	*Vs* (m/s)	*d* (m)
1	663	1.8	100-200	0.5-3
2	663	1.8	75-200	0.5-3
3	995	1.8	150-250	2-5
4	995	1.8	120-250	2-10
5	995	1.8	200-300	5-15
6	995	1.8	450-650	*Half space*

The MI-PSO algorithm for inverting the Lahad Datu site dispersion curve is set to 10 inversions, with 600 iterations for each inversion and 50 populations. For this field data, the number of iterations and population for the inversion algorithm are set higher than for synthetic data. This is because field data usually contains a certain amount of noise. Inversion shows that the observed and theoretical dispersion data have a good fit. The best fit of the dispersion curves appears in the wavelength range of 0–40 m (Figure 6a). There is one point where the observed phase velocity does not closely correspond to the theoretical phase velocity, namely at a wavelength of around 100 m. There appears to be a very low underestimate of the calculated phase velocity at the highest wavelength of around 100 m. This basically describes the phase velocity pitfall, which is usually found at the lowest frequency (or highest wavelength) due to over-picking when the dispersion curve extraction is performed [29]. The MI-PSO also displays model solutions, both those calculated individually from each inversion (grey solid lines) and those calculated with the weighted average of the 10 inversions (green solid line in Figure 6b). The model solution from the weighted average is then used to describe the estimated subsurface structure of the Lahad Datu site. The subsurface structure is distributed in five layers plus a half-space consisting of two soft layers in the depth ranges of 1.4–3.0 m and 6.6–12.2 m.

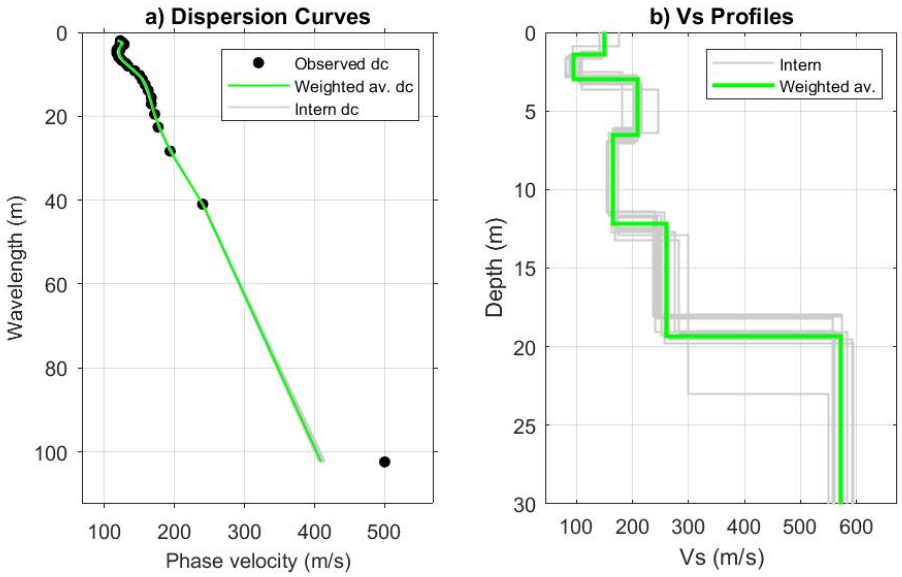

Figure 6. Inversion results for field data in the archeological site of Lahad Datu: a) Dispersion curves, and b) *Vs* profiles.

Tabulation of model parameters for each layer, resulting from individual inversions as well as weighted average model solutions, for Rayleingh wave data at the archeological site of Lahad Datu, is presented in Table 6. The individual results for the 10 inversions show the different parameter values vary with a fairly small misfit value for each inversion (i.e., less than 2%). The lowest misfit value was obtained in the 2nd inversion, and the highest misfit was obtained in the 3rd inversion. The weighted average values for a total of 10 individual inversions indicate the presence of layers with complex characteristics at the Lahad Datu site, representing the presence of a combination of LVL and HVL characteristics. The layers with lower *Vs* are located in the second and fourth layers, with *Vs* values of 93.7 m/s and 164.5 m/s at depths of 1.4–3.0 m and 6.6–12.2 m, respectively. The presence of these two soft layers justifies the qualitative assumptions given earlier.

Table 6. Model solutions for the archaeological site of Lahad Datu

Parameter model	Model for each inverse										Weighted average model
	1	2	3	4	5	6	7	8	9	10	
Vs_1	175.4	148.7	141.1	146.8	149.8	148.8	146.3	149.8	146.7	146.4	150.2
Vs_2	93.5	81.6	109.7	107.3	88.4	87.5	80.6	91.8	104.5	103.7	93.7
Vs_3	181.8	207.4	246.3	210.9	201.7	206.4	213.1	209.6	215.6	202.6	208.8
Vs_4	173.4	165.2	157.6	168.1	162.4	168.5	163.4	153.6	164.6	167.8	164.5
Vs_5	300.0	240.9	251.8	283.0	243.0	263.8	275.2	257.5	236.1	247.5	259.9
Vs_6	549.6	583.2	568.2	572.1	557.6	591.7	560.7	594.3	575.0	574.6	572.8
d_1	0.92	1.55	1.41	1.25	1.44	1.49	1.67	1.42	1.32	1.34	1.39
d_2	1.59	1.19	2.24	2.07	1.35	1.32	1.17	1.48	1.91	1.82	1.57
d_3	4.23	3.42	2.75	3.33	4.02	3.26	3.80	4.04	3.05	3.87	3.60
d_4	6.17	5.25	5.27	6.55	4.97	6.36	6.04	4.54	5.41	5.57	5.61
d_5	10.08	7.67	7.57	5.91	6.47	7.07	6.34	8.35	6.40	5.37	7.15
misfit	1.26	1.11	1.52	1.44	1.14	1.14	1.12	1.22	1.42	1.39	

Inversion for Rayleigh Wave Data from Shiga Site (Japan)

MI-PSO implementation is then carried out on the field data of the Shiga site, Japan. This geotechnical site is a residential area that generally has soft layers near the surface (Safani, 2007). Changes in the Rayleigh wave phase velocity with respect to frequency and wavelength are shown in Figures 7a and 7b, respectively. The phase velocity undergoes a unique changing trend with frequency. In the frequency range of 5 Hz to 20 Hz, the phase velocity decreases with increasing frequency. But at higher frequencies (i.e., > 20 Hz), the phase velocity undergoes at least four segment jumps to higher modes of phase velocities. Such a dispersion curve pattern qualitatively describes the occurrence of wave energy trapping in a near-surface LVL zone [30]. Quantitative estimates of the presence of the LVL will be examined through the inversion of the existing field dispersion curves.

The inversion process begins by setting the search space for shear wave velocity (*Vs*) and layer thickness (*d*), while the compression wave velocity (*Vp*) and density (ρ) parameters are usually assumed from a priori information (Table 7). Setting the *Vp* values in Table 7 does not show a low value in the second layer as seen in the corresponding *Vs* value because the *Vp* parameter usually does not detect the presence of soft layers near the surface (Safani, 2007). The search space settings for parameters *Vs* and *d* are determined based on the trend of the dispersion curve from the field data presented in Figure 7, where the phase velocity at higher frequencies represents shallower layers and at lower frequencies represents deeper layers.

Table 7. Search space for measured data inversion at the Shiga site, Japan

Layers	Search Spaces			
	Vp (m/s)	ρ (g/cm^3)	*Vs* (m/s)	*d* (m)
1	663	1.8	150-300	0.5-3
2	663	1.8	75-200	0.5-5
3	995	1.8	150-450	9-14
4	995	1.8	400-600	3-8
5	995	1.8	500-800	*Half space*

As with the inversion of the previous field data, the inversion of the Rayleigh wave data at the Shiga site was carried out 10 times. The number of iterations and population in each inversion are 600 and 50, respectively. The results of the inversion of the dispersion curve data are presented in Figure 7. In the lowest wavelength range up to 80 m, the theoretical dispersion curve shows a very good agreement with the measured dispersion curve (Figure 7a). However, at the highest wavelengths (i.e., at about 130 m), the theoretical phase velocity shows a very large difference from the corresponding observed phase velocity. As previously described, such a case is referred to as the "pitfalls" of phase velocity, which usually occur at the highest wavelength (or lowest frequency) due to over-picking when dispersion curve extraction is performed. Such a Pitfall is usually solved simply by removing the data from the inversion calculation process. This is necessary to avoid misinterpretation of the results of the This is necessary to avoid misinterpretation of the results of the Rayleigh wave inversion when such a fact is encountered.

The model solution obtained in each inversion is shown with gray solid lines (Figure 7b). There are four layers plus a half-space in the estimation model. Although the ten inversions show diverse model profiles, all of the resulting model solutions show consistency in model profile trends, especially when describing the presence of an LVL in the second layer. The

weighted average profile of the shear wave velocity (*Vs*) versus thickness (*d*) shown by the green solid line in Figure 7b is the final model solution that represents the characteristics of the subsurface layers of the observation area at the Shiga site.

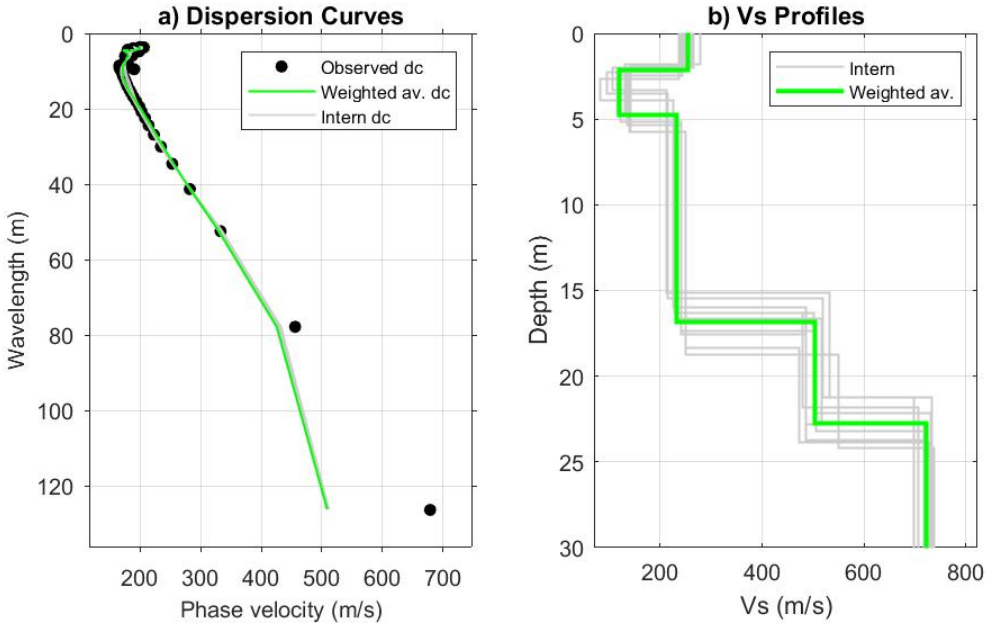

Figure 7. Results of the inversion of the measured data at Shiga site, Japan: a) Dispersion curves, and b) *Vs* profiles.

A more detailed inspection of the model parameter values (or model solutions) generated for each inversion is tabulated in Table 8. The model solution values for each inversion are quite diverse, but all have small misfit values (i.e., less than 3%). The lowest misfit value is achieved in the 8th inversion, and the highest is in the 4th inversion. The calculating results of the weighted average values show that the low-velocity layer (LVL) in the second layer has a shear wave velocity (*Vs*) value of 120.2 m/s with a layer thickness (*d*) of 2.63 m. This LVL is flanked by two harder layers, namely the surface layer (with *Vs* = 255.0 m/s and *d* = 2.13 m) and the third layer (with *Vs* = 233.20 m/s and *d* = 12.08 m).

Table 8. Model solution for the measured data at Shiga site, Japan

Parameter model	Model for each inversion										Weighted average
	1	2	3	4	5	6	7	8	9	10	
Vs_1	237.1	258.6	240.4	247.6	261.6	266.4	257.5	243.0	278.9	252.2	255.0
Vs_2	83.3	135.6	96.2	105.8	140.9	122.2	114.4	122.6	129.9	139.2	120.2
Vs_3	225.2	240.2	215.3	211.5	250.8	228.2	229.9	241.6	231.0	251.0	233.2
Vs_4	487.0	506.9	519.5	532.2	473.6	500.8	480.2	487.0	517.1	550.4	504.7
Vs_5	722.5	723.5	732.0	698.4	727.9	731.7	706.9	723.3	717.3	738.3	722.4
d_1	2.67	2.00	2.28	2.01	1.88	2.02	2.22	2.43	1.81	2.03	2.13
d_2	1.27	3.37	1.22	1.30	3.84	2.73	2.47	2.72	3.04	3.69	2.63
d_3	12.05	12.01	11.93	11.82	12.61	11.55	11.59	12.43	11.79	13.01	12.08
d_4	6.86	5.84	5.81	6.11	5.50	5.86	5.54	6.18	6.05	5.45	5.91
misfit	2.76	2.36	2.84	2.97	2.46	2.25	2.27	2.20	2.27	2.42	

CONCLUSION

The algorithm for multiple inversions using particle swarm optimization (the so-called MI-PSO algorithm) for inversions of Rayleigh wave dispersion curves for geotechnical site characterization has been developed. Algorithm testing is carried out by utilizing two synthetic models that replicate subsurface characteristics that are often found in geotechnical cases, namely the HVL model and the complex model. The complex model represents a combination of models with LVL and HVL characteristics. The test results with synthetic data show that the developed MI-PSO algorithm with a weighted average solution can work with excellent accuracy. The final model solution obtained through the MI-PSO algorithm with a weighted average solution is better than the model solution obtained from the conventional average. The superiority of the weighted average model compared to the conventional one is also indicated by its higher similarity index. All of the resulting model solutions can reconstruct *Vs* versus depth profile patterns very accurately, both for the HVL and the complex models. This shows that the MI-PSO algorithm with a weighted average solution has a very good ability to solve the Rayleigh wave dispersion curve inversion problem, even for complex subsurface models.

Implementation of the MI-PSO algorithm with a weighted average solution on two measured data sets may work excellently and can delineate the target models. The assumption of the existence of a complex model at Lahad Datu's archeological site through a qualitative interpretation of the trend of the dispersion curve can be proven quantitatively through inversion with this algorithm. Likewise, the presence of a near-surface LVL at the Shiga site (Japan) can be estimated and presented in the resulting subsurface profile.

REFERENCES

1 Karray, M. 2010. Shear wave velocity in geotechnical engineering. In *Proceedings of the International Conference on Geotechnical Engineering, Hammamet*, 20-23.

2 Aki, K. 1957. Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors. *Earthq. Res. Inst.*, 415–456.

3 Everett, M. E. 2013. *Near-surface applied geophysics*. Cambridge University Press.

4 Park, C. B., Miller, R. D. & Xia, J. 1999. Multichannel analysis of surface waves. *Geophysics 64*, 800–808.

5 Park, C. B., Miller, R. D. & Xia, J. 2007. Multichannel analysis of surface waves (MASW)—active and passive methods. *Lead. Edge*, *26*, 60–64.

6 Xia, J., Miller, R. D. & Park, C. B. 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics*, *64*, 691–700.

7 Song, X., Gu, H., Zhang, X. & Liu, J. 2008. Pattern search algorithms for nonlinear inversion of high-frequency Rayleigh-wave dispersion curves. *Comput. Geosci.*, *34*, 611–624.

8 Song, X., Tang, L., Lv, X., Fang, H. & Gu, H. 2012. Application of particle swarm optimization to interpret Rayleigh wave dispersion curves. *J. Appl. Geophys.*, *84*, 1–13.

9 Dal Moro, G., Pipan, M. & Gabrielli, P. 2007. Rayleigh wave dispersion curve inversion via genetic algorithms and Marginal Posterior Probability Density estimation. *J. Appl. Geophys.*, *61*, 39–55.

10 Sahraeian, M. S., Ghalandarzadeh, A. & Kavand, A. 2008. Estimation of Shear Wave Velocity By Means of Array Measurement of Microtremors Using Genetic Algorithm Method. *Thr 14th World Conf. Earthq. Eng.*, 1–9.

11 Rubaiyn, A., Priyono, A. & Yudistira, T. 2019. Near-surface S-wave Estimation base on Inversion of Rayleigh Wave Dispersion Curve Using Genetic Algorithm. *IOP Conf. Ser. Earth Environ. Sci.*, *318*, 1–6.

12 Safani, J. & Matsuoka, T. 2013. Soft-geotechnical Zone Determination using Surface-wave for Geotechnical Hazard Mitigation. *Procedia Environ. Sci.*, *17*, 354–360.
13 Beaty, K. S., Schmitt, D. R. & Sacchi, M. 2002. Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure. *Geophys. J. Int.*, *151*, 622–631.
14 Lu, Y., Peng, S., Du, W., Zhang, X., Ma, Z., & Lin, P. 2016. Rayleigh wave inversion using heat-bath simulated annealing algorithm. *Journal of Applied Geophysics*, *134*, 267-280.
15 Wang, Y., Wang, H., Wu, X., Chen, K., Liu, S., & Deng, X. 2021. Near-surface velocity inversion from Rayleigh wave dispersion curves based on a differential evolution simulated annealing algorithm. *Artificial Intelligence in Geosciences*, *2*, 35-46. h
16 Marini, F. & Walczak, B. 2015 Chemometrics and Intelligent Laboratory Systems Particle swarm optimization (PSO). A tutorial. *Chemom. Intell. Lab. Syst.* 149, 153–165.
17 Pace, F., Santilano, A. & Godio, A. 2021. *A Review of Geophysical Modeling Based on Particle Swarm Optimization*. *Surveys in Geophysics*, 42.
18 Pallero, J. L. G., Fernández-Martínez, J. L., Bonvalot, S. & Fudym, O. 2015. Gravity inversion and uncertainty assessment of basement relief via Particle Swarm Optimization. *J. Appl. Geophys.* 116, 180–191.
19 Pekşen, E., Yas, T. & Kıyak, A. 2014. 1-D DC Resistivity Modeling and Interpretation in Anisotropic Media Using Particle Swarm Optimization. *Pure Appl. Geophys.* 171, 2371–2389.
20 Oyeyemi, K. D. *et al.* 2023. Nonlinear inversion of electrical resistivity sounding data for multi-layered 1-D earth model using global particle swarm optimization (GPSO). *Heliyon* 9, e16528.
21 Liu, S., Liang, M. & Hu, X. 2018. Particle swarm optimization inversion of magnetic data: Field examples from iron ore deposits in China. *Geophysics* 83, J43–J59.
22 Schutte, J. F. & Groenwold, A. A. 2005. A study of global optimization using particle swarms. *J. Glob. Optim.* 31, 93–108.
23 Raha, R. 2019. Particle-Swarm-Optimization-using-Matlab. *GitHub repository* https://github.com/ritwikraha/Particle-Swarm-Optimization-using-Matlab.
24 Kennedy, J. & Eberhart, R. 1995. Particle swarm optimization. in *Proceedings of ICNN'95 - International Conference on Neural Networks*, IEEE.
25 Shi, Y. & Eberhart, R. 1998. A Modified particle swarm optimizer. *Proc. IEEE Conf. Evol. Comput. ICEC* 69–73.
26 Fourie, P. C. & Groenwold, A. A. 2002. The particle swarm optimization algorithm in size and shape optimization. *Struct. Multidiscip. Optim.* 23, 259–267.
27 Kausel, E. & Roësset, J. M. Stiffness matrices for layered soils. *Bull. Seismol. Soc. Am.* 71, 1743–1761 (1981).
28 Olafsdottir, E. A. 2019. Multichannel Analysis of Surface Waves for Soil Site Characterization. University of Iceland.
29 Safani, J. 2007. *Surface Wave Dispersion Modelling by Full-Wavefield Reflectivity and Inversion for Shallow Subsurface Imaging*. Doctoral Thesis. Kyoto University.
30 Safani, J., O'Neill, A. & Matsuoka T. 2006. Full SH-wavefield modelling and multiple-mode Love wave inversion. *Exploration Geophysics*, *37*(4), 307-321.