Soil Dynamics and Earthquake Engineering 75 (2015) 147–157

Contents lists available at ScienceDirect

Soil Dynamics and Earthquake Engineering

journal homepage: www.elsevier.com/locate/soildyn

Grey Wolf Optimizer for parameter estimation in surface waves

Xianhai Song *, Li Tang, Sutao Zhao, Xueqiang Zhang, Lei Li, Jianquan Huang, Wei Cai

Institute of Geophysics and Geomatics, Hubei Subsurface Multi-scale Imaging Key Laboratory, China University of Geosciences, 388 Lumo Road, Wuhan, Hubei 430074, China

ARTICLE INFO

Article history:
Received 22 December 2014
Received in revised form
22 March 2015
Accepted 8 April 2015
Available online 29 April 2015

Keywords:
Swarm intelligence
Grey Wolf Optimizer
Rayleigh waves
Surface waves
Dispersion curves

ABSTRACT

This research proposed a novel and powerful surface wave dispersion curve inversion scheme called Grey Wolf Optimizer (GWO) inspired by the particular leadership hierarchy and hunting behavior of grey wolves in nature. The proposed strategy is benchmarked on noise-free, noisy, and field data. For verification, the results of the GWO algorithm are compared to genetic algorithm (GA), the hybrid algorithm (PSOGSA)-the combination of Particle Swarm Optimization (PSO) and Gravitational Search Algorithm (GSA), and gradient-based algorithm. Results from both synthetic and real data demonstrate that GWO applied to surface wave analysis can show a good balance between exploration and exploitation that results in high local optima avoidance and a very fast convergence simultaneously. The great advantages of GWO are that the algorithm is simple, flexible, robust and easy to implement. Also there are fewer control parameters to tune.

1. Introduction

In recent years Rayleigh waves have captured the interest of a constantly increasing number of researchers from different disciplines for a wide range of applications [1–3]. They can be used to obtain near-surface S-wave velocity models [4], to map bedrock [5], to infer subsurface properties in viscoelastic media [6], to determine Q of near-surface materials [7,8], to assess soil liquefaction potential [9], to delineate a shallow fault zone [10], to characterize pavement structure [11,12], to characterize seismic site structure [13], and to perform a joint inversion with refractions [14,15], reflection travel times [16], Love waves [17] or attenuation curves [18]. In these significant applications, utilization of Rayleigh wave dispersive properties is often divided into three procedures: field data acquisition [19–22], reconstruction of dispersion curves [23], and inversion of phase velocities [24–28].

Once Rayleigh wave dispersion curve is properly identified, its inversion is the key point to obtain S-wave velocity profiles [29–31]. A variety of local optimization methods have been developed and widely used to interpret Rayleigh wave data [32–34]. However, inversion of Rayleigh waves is typically a highly nonlinear, multiparameter, and multimodal inversion problem. The objective function for surface wave inversion has massive local optima with the number increasing exponentially with dimension. Consequently, linearized inversion strategies are prone to being trapped by local minima, and their success depends heavily on the choice of the initial model and on the accuracy of partial derivatives. Thus, global optimization methods that can overcome this limitation are particularly attractive for surface wave analysis, such as genetic algorithms [35,36], simulating annealing [37–39], and Monte Carlo [40–42].

Nature has always been an inspiration source for scientists. Mirjalili et al. conceived the idea of mimicking the social leadership hierarchy and hunting behavior of grey wolves into optimization problems and called the resulting technique as Grey Wolf Optimizer (GWO) [43]. GWO, a newcomer among population-based swarm intelligence optimization algorithms, is characterized by several appealing advantages: simplicity, flexibility, derivation-free mechanism, and local optima avoidance. Also, it is easy to implement; and it has fewer control parameters to adjust, and it has a fast convergence characteristic.

First, GWO is fairly simple. It is inspired from the particular leadership hierarchy and hunting behavior of grey wolves in nature. The simplicity allows computer scientists to simulate natural concepts and develop the algorithm more effectively. Moreover, the simplicity assists other scientists to learn the algorithm quickly and apply it to their problems. Second, flexibility refers to the applicability of GWO to different problems without any special changes in the structure of the algorithm. GWO is readily applicable to different problems since it assumes problems as black boxes. Third, GWO has derivation-free mechanisms. In contrast to gradient-based optimization approaches, GWO optimizes problems stochastically. It can be effectively used for addressing problems for which objective functions are non-differentiable, stochastic, or even discontinuous. Finally, GWO has superior abilities to avoid local optima compared to conventional optimization techniques. This makes GWO highly suitable for

* Corresponding author. Tel.: +86 27 67883251; fax: +86 27 67883251.
E-mail address: songxianhaiwcy@sina.com (X. Song).

http://dx.doi.org/10.1016/j.soildyn.2015.04.004
0267-7261/

solving highly nonlinear, multivariable, multimodal function optimization problems.

Mirjalili et al. [43] have recently tested GWO on unimodal, multimodal, fixed-dimension multimodal, and composite functions to benchmark its performance in term of exploration, exploitation, local optima avoidance, and convergence. It has been shown that the GWO algorithm is able to provide very competitive results compared to other well-known meta-heuristics. The GWO algorithm has been successfully applied to three classical engineering design problems and real optical engineering [43]. Song et al. [44] have successfully applied GWO for solving combined economic emission dispatch problems. Emary et al. [45] have used GWO for feature subset selection. Mirjalili [46] has investigated the effectiveness of GWO in training multi-layer perceptions (MLP). Saremi et al. [47] proposed the use of evolutionary population dynamics (EPD) in the GWO algorithm to further enhance its performance. Mirjalili et al. [48] have compared GWO with Multi-Verse Optimizer (MVO). Results from both applications and investigations show that the GWO algorithm has the superior performance not only in terms of exploring the promising regions extensively but also in terms of exploiting the optimum.

Although there are a lot of population-based algorithms in the literature, the operators of algorithms are usually designed to accomplish two phases [48]: exploration versus exploitation. In the former phase, an algorithm should be equipped with mechanisms to explore the search space as extensively as possible. In fact, promising regions of the search space are identified in this phase. In the exploitation phase, however, there should be emphasizes on local search and convergence towards promising areas obtained in the exploration phase. Exploration and exploitation are two conflicting stages with no specific mathematical definition. The majority of population-based algorithms have been tuned adaptively to smoothly transit between exploration and exploitation. For instance, the inertia weight in PSO is mostly decreased linearly from 0.9 to 0.4 in order to emphasize exploitation as iterations increase. However, there is no mechanism for significant abrupt movements in the search space for PSO and this will likely result in the poor performance of PSO. Therefore, finding a good balance between exploration and exploitation when designing an algorithm is challenging. There is no clear rule for an algorithm to realize the most suitable time for transiting from exploration to exploitation due to both unknown shape of search spaces and stochastic nature of population-based algorithms. This is the reason why current multi-solution stochastic optimizers still prone to local optima stagnation. Parameter estimation in surface waves is considered as a challenging problem due to its high nonlinearity and to its multimodality. It has been proven that GWO shows a good balance between exploration and exploitation that results in high local optima avoidance and a very fast convergence simultaneously. Therefore, the high level of exploration and exploitation that may assist GWO to outperform other optimizers in this field motivates our attempts to investigate its efficiencies in parameter estimation in surface waves.

In this study, we demonstrate a GWO application on surface wave data for near-surface S-wave velocity profiles. The proposed procedure is tested on noise-free, noisy, and field data. Furthermore, the results of the GWO algorithm are compared to GA, PSOGSA, and local search algorithm to further verify the performance of GWO. Results from both synthetic and field data demonstrate that GWO has the high level of exploration and exploitation that result in high local optima avoidance and a very fast convergence simultaneously in parameter estimation in surface waves.

2. Grey Wolf Optimizer (GWO)

The social hierarchy and the hunting behavior of grey wolves are mathematically modeled by Mirjalili et al. [43] in order to design GWO.

2.1. Social hierarchy

Grey wolf belongs to Canidae family. Grey wolves are considered as apex predators, meaning that they are at the top of the food chain. Grey wolves mostly prefer to live in a pack. Of particular interest is that they have a strict social dominant hierarchy from alpha, beta, delta, to omega.

In order to mathematically model the social hierarchy of grey wolves when designing GWO, the fittest solution is considered as the alpha (α). Consequently, the second and third best solutions are named as the beta (β) and the delta (δ), respectively. The rest of the candidate solutions are assumed to be the omega (ω). In the GWO algorithm, the hunting (optimization) is guided by α, β, and δ. The ω wolves follow these three wolves.

2.2. Encircling prey

In addition to the social hierarchy of grey wolves described above, group hunting is another interesting social behavior of grey wolves. According to Muro et al. [50], the main phases of grey wolf hunting include: (1) Tracking, chasing, and approaching the prey; (2) Encircling, pursuing, and harassing the prey until it stops moving; (3) Attacking towards the prey. In order to mathematically model encircling behavior, the following equations are proposed [43]:

$$D = |C \cdot X_p(t) - X(t)| \tag{1}$$

$$X(t+1) = X_p(t) - A \cdot D \tag{2}$$

Where t indicates the current iteration; A and C are coefficient vectors; X_p is the position vector of the prey; and X indicates the position vector of a grey wolf. The coefficient vectors $A = a \cdot (2r_1 - 1)$ and $C = 2r_2$. where a is linearly decreased from 2 to 0 over the course of iterations;r_1,r_2 are random values in [0,1]; so A is random values in the interval $[-a, a]$.

2.3. Search for prey (exploration)

Grey wolves mostly search according to the position of the alpha, beta, and delta. They diverge from each other to search for prey and converge to attack prey. In order to mathematically model divergence, A is utilized with random values greater than 1 or less than -1 to oblige the search agent to diverge from the prey. This emphasizes exploration and allows GWO to search globally. That is, $|A| \geq 1$ forces the grey wolves to diverge from the prey to hopefully find a fitter prey.

Another component of GWO that favors exploration is C. The C vector contains random values in [0, 2]. This component provides random weights for prey in order to stochastically emphasize ($C \geq 1$) or deemphasize ($C < 1$) the effect of prey in defining the distance in Eq. (1). This assists GWO to show a more random behavior throughout optimization, favoring exploration and local optima avoidance. It is worth mentioning that C is not linearly decreased in contrast to A. GWO deliberately requires C to provide random values at all times to emphasize exploration/exploitation not only during initial iterations but also final iterations. This component is very helpful in case of local optima stagnation, especially in the final iterations.

2.4. Attacking prey (exploitation)

In order to mathematically model approaching the prey, the value of a is linearly decreased. Thus A is a random value in the interval $[-a, a]$. When random values of A are in [-1,1] ($|A| < 1$), GWO forces the wolves to attack towards the prey.

2.5. Hunting

In order to mathematically simulate the hunting behavior of grey wolves, the first three best solutions obtained so far are saved and oblige the other search agents (including the omegas) to update their positions according to the position of the best search agents. The following formulas are proposed [43]:

$$D_\alpha = |C_1 \cdot X_\alpha - X|, D_\beta = |C_2 \cdot X_\beta - X|, D_\delta = |C_3 \cdot X_\delta - X| \quad (3)$$

$$X_1 = X_\alpha - A_1 \cdot D_\alpha, X_2 = X_\beta - A_2 \cdot D_\beta, X_3 = X_\delta - A_3 \cdot D_\delta \quad (4)$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \quad (5)$$

To sum up, the search process starts with creating a random population of grey wolves (candidate solutions) in the GWO algorithm. Over the course of iterations, alpha, beta, and delta wolves estimate the probable position of the prey. Each candidate solution updates its distance from the prey. The parameter a is decreased from 2 to 0 in order to adaptively emphasize exploration and exploitation, respectively. Candidate solutions tend to diverge from the prey when $|A| \geq 1$ and converge towards the prey when $|A| < 1$. Finally, the GWO algorithm is terminated by the satisfaction of an end criterion.

GWO has only one main parameter (a) to be adjusted. The adaptive values of parameters a and A allow GWO to smoothly transition between exploration and exploitation. Therefore, exploration and exploitation are guaranteed by the adaptive values of a and A. With decreasing A, half of the iterations are devoted to exploration ($|A| \geq 1$) and the other half are dedicated to exploitation ($|A| < 1$).

3. GWO for surface wave analysis

We implement a series of MATLAB tools based on MATLAB 2014 for surface wave analysis. We are developing SWIGWO, a software package for Surface Wave Inversion via Grey Wolf Optimizer.

In the current research, we focus our attention on inversion results of fundamental-mode surface waves for near-surface S-wave velocities and layer thicknesses by fixing P-wave velocities (or Poisson's ratio) and densities to their known values or good estimates. The reduction of the computation effort and the fact

Table 1
Soil properties for three layered models and search space in the inversion of surface waves based on the GWO algorithm.

Model number	Layer number	Parameters				Search space	
		S-wave velocity (m/s)	P-wave velocity (m/s)	Density (g/cm^3)	Thickness (m)	S-wave velocity (m/s)	Thickness (m)
Model A	1	200	663	1.9	4	100–300	2–6
	2	300	995	1.9	2	150–450	1–3
	3	400	1327	1.9	6	200–600	3–9
	4	500	1658	1.9	Half-space	250–750	Half-space
Model B	1	200	663	1.9	4	100–300	2–6
	2	150	765	1.9	2	100–300	1–3
	3	300	1102	1.9	6	150–450	3–9
	4	400	1470	1.9	Half-space	200–600	Half-space
Model C	1	150	498	1.9	4	50–250	2–6
	2	250	829	1.9	2	100–400	1–3
	3	200	841	1.9	6	100–400	3–9
	4	400	1470	1.9	Half-space	200–600	Half-space

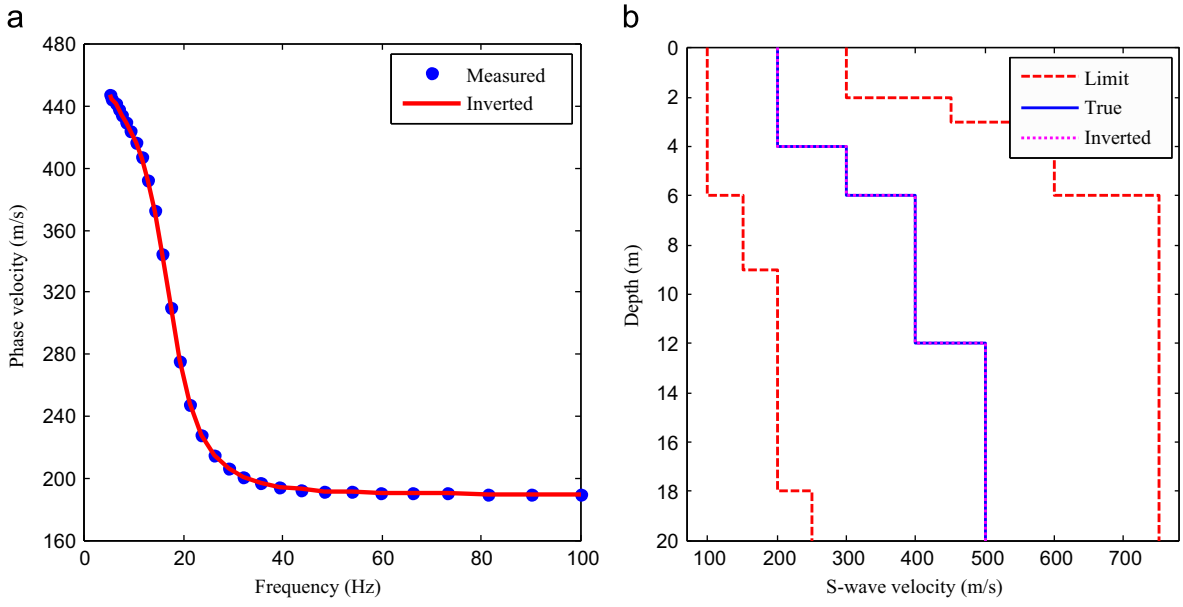

Fig. 1. Inversion results of Model A using the GWO algorithm. (a) Noise-free synthetic data (solid dots) and modeled dispersion curve (solid line). (b) The lower and upper bounds of the search area (dashed lines), True model (solid line) and inverted profile (dotted line).

that S-wave velocities and layer thicknesses are the most important parameters that determine Rayleigh wave propagation in high frequency range [2] are the main reasons for these choices. To simulate more realistic cases where no *a priori* information is available and fully evaluate the performance of GWO in terms of exploration, exploitation, local optima avoidance, and convergence, we use a wider search boundary. By design, the lower and upper limits of the search areas depart 50% or more from their true values in all of the latter tests.

Our inversions of surface wave data are conducted with the following parameter sets: the population size of grey wolves $N=10*D$, where D is the problem dimension. By a few test runs of GWO, maximum iteration number G for convergence is set to be 50 for noise-corrupted synthetic data and real data, and G is set to be 100 for noise-free synthetic data in our inversion procedure. The GWO algorithm was run 20 times on each model to get final results. The procedure will set out to find the global minimum of Mean Square Error (MSE) misfit between the measured and the predicted phase velocities. Forward modeling of Rayleigh wave dispersion curves is based on the fast delta matrix algorithm [51].

4. Synthetic data inversion

To examine the performance of the GWO algorithm described above, three synthetic models are used. These models are designed to simulate situations commonly encountered in shallow engineering site investigations. As shown in Table 1, Model A with S-wave velocities increasing with depth, represents a multilayer geologic structure. Model B with a soft layer trapped between two stiff layers models a real complex pavement structure containing a low velocity layer (LVL). Model C, characterized by a stiff layer

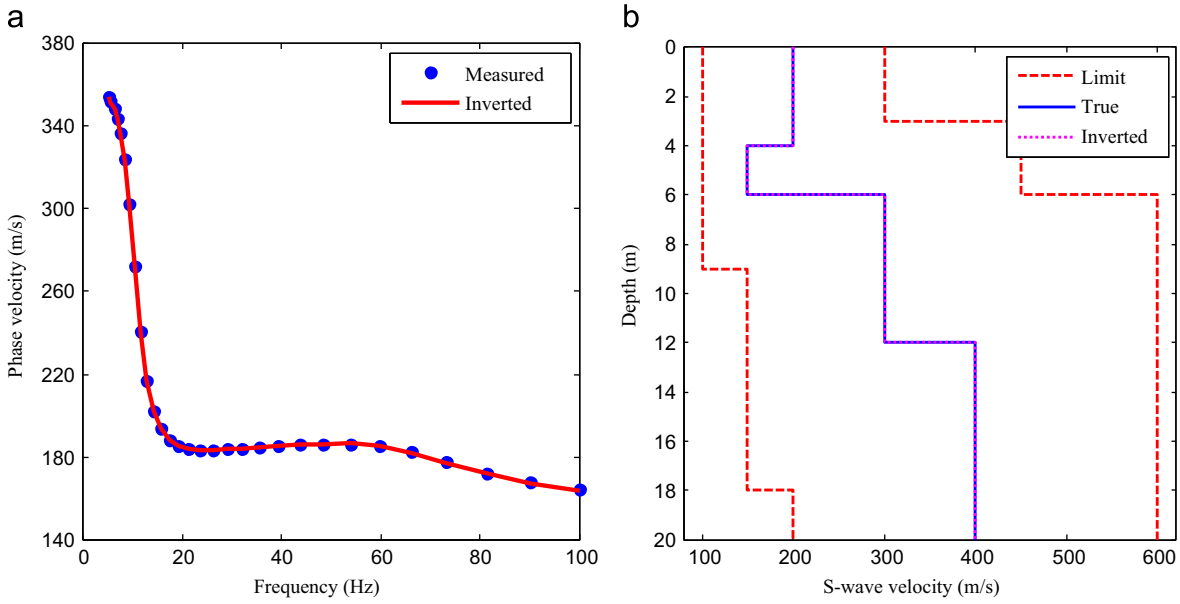

Fig. 2. Inversion results of Model B using the GWO algorithm. (a) Noise-free synthetic data (solid dots) and modeled dispersion curve (solid line). (b) The lower and upper bounds of the search area (dashed lines), True model (solid line) and inverted profile (dotted line).

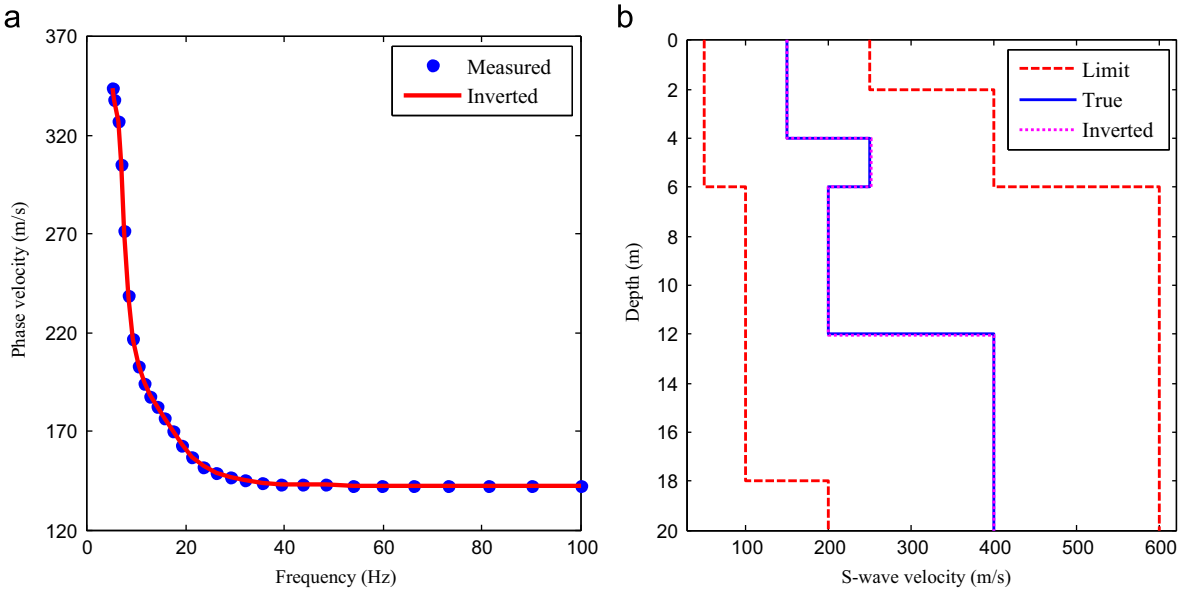

Fig. 3. Inversion results of Model C using the GWO algorithm. (a) Noise-free synthetic data (solid dots) and modeled dispersion curve (solid line). (b) The lower and upper bounds of the search area (dashed lines), True model (solid line) and inverted profile (dotted line).

sandwiched between two soft layers, simulates a complex geotechnical soil profile containing a high velocity layer (HVL).

4.1. Noise-free synthetic data

The performance of GWO on surface wave data is first investigated on noise-free synthetic data. Figs. 1, 2 and 3 demonstrate inversion results of GWO on three noise-free synthetic data sets from Model A, Model B, and Model C, respectively. The search spaces are reported in Table 1, which depart 50% or more from their true values. It can be noted that three true models are quite well recovered from the GWO inversions (dotted lines in Figs. 1b, 2b, and 3b). The overall average errors between the true models and the estimated models from GWO are 0.32%, 0.20%, and 0.28%, respectively. The model responses (solid lines in Figs. 1a, 2a, and 3a) from the GWO-deduced models match the measured phase velocities (solid dots in Figs. 1a, 2a, and 3a) almost perfectly.

4.2. Contaminated synthetic data

In practice picked surface wave phase velocities are inevitably noise. How would the algorithm behave for noise-corrupted data? The performance of the proposed inverse procedure is further tested on contaminated synthetic data. We perturb the true surface wave phase velocities from three multilayer models by the introduction of a 10% of white Gaussian noise.

Inversion results of GWO in this realistic condition are illustrated in Figs. 4, 5, and 6. One can see that GWO is very robust in the presence of noisy data and acceptable solutions can be obtained by GWO inversions. The modeled dispersion curves (dashed lines in Figs. 4a, 5a, and 6a) from the GWO-based

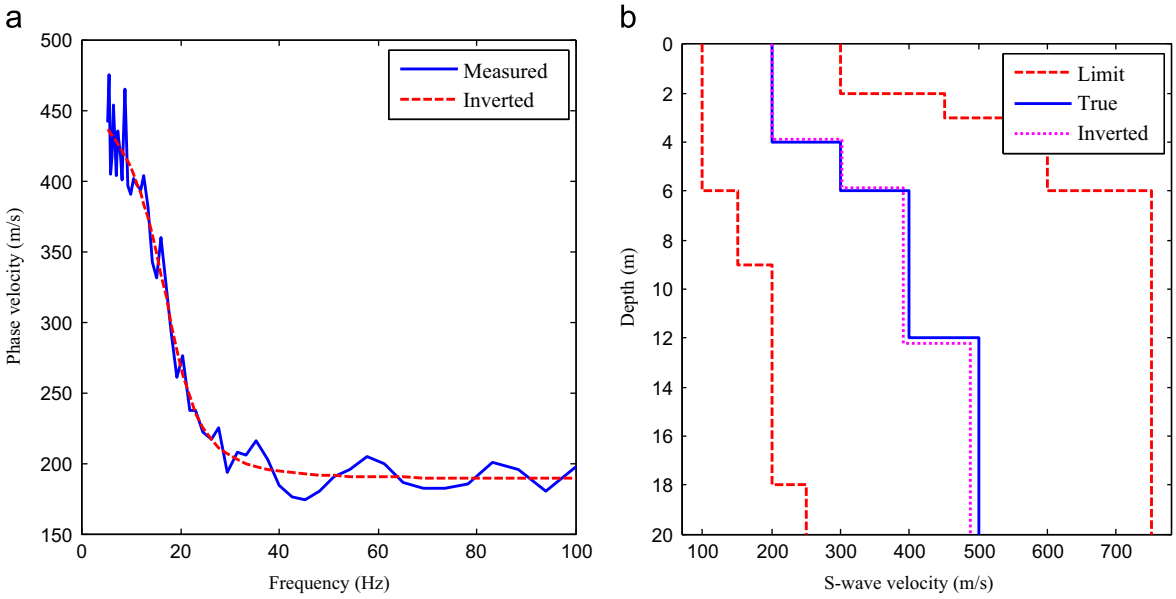

Fig. 4. Inversion results of Model A using the GWO algorithm. (a) Contaminated synthetic data (solid line) and modeled dispersion curve (dashed line). (b) The lower and upper bounds of the search area (dashed lines), True model (solid line) and inverted profile (dotted line).

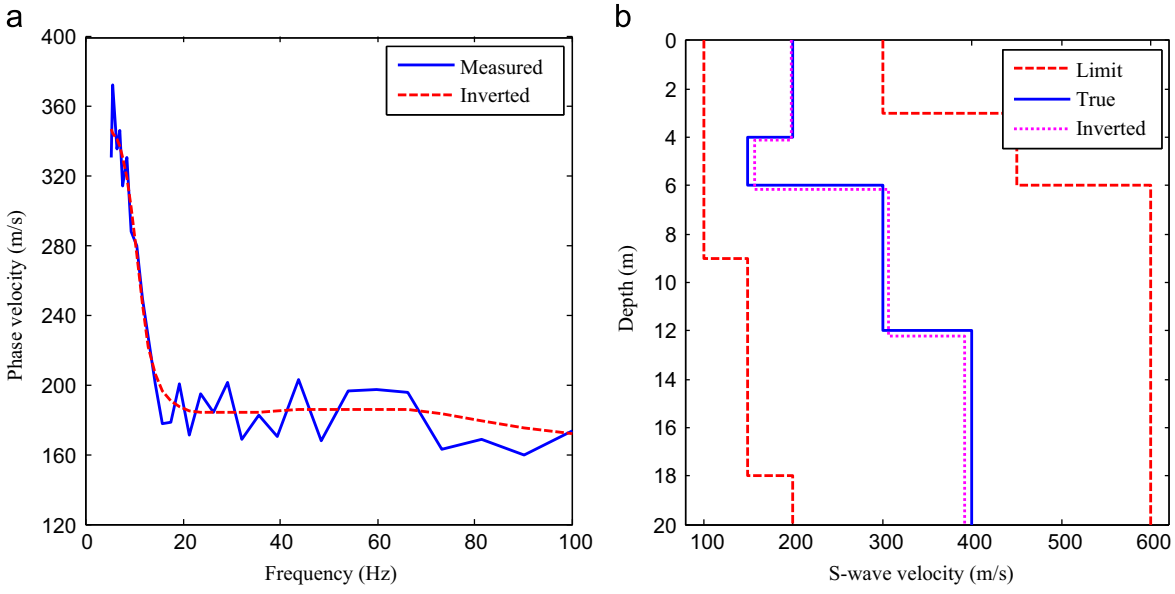

Fig. 5. Inversion results of Model B using the GWO algorithm. (a) Contaminated synthetic data (solid line) and modeled dispersion curve (dashed line). (b) The lower and upper bounds of the search area (dashed lines), True model (solid line) and inverted profile (dotted line).

solutions fit the measured phase velocities (solid lines in Figs. 4a, 5a, and 6a) reasonably well. The overall average error between the true models and the models estimated by GWO are 2.09%, 2.41%, and 2.73%, respectively. These results show the superior performance of GWO in solving parameter estimation in surface waves.

4.3. Convergence behavior analysis

According to GWO, there should be abrupt changes in the movement of search agents over the initial steps of optimization. This assists a meta-heuristic to explore the search space extensively. Then, these changes should be reduced to emphasize exploitation at the end of optimization. In order to investigate the convergence behavior of GWO, the fitness history of GWO iteration using noise-free and noisy synthetic data from Model A is illustrated in Fig. 7, which provides valuable insights into the performance of the proposed inversion scheme in terms of exploration, exploitation, local optima avoidance, and convergence. The convergence curve in Fig. 7a demonstrates a typical characteristics of the GWO algorithm. It shows a very fast initial convergence for the global exploration at the first 50 iterations, and then gradually converge to zero for the local exploitation in the next 50 iterations. As it can be observed in Fig. 7b, the minimum misfit values significantly decrease for the global exploration at the first 25 iterations, and then gradually converge to a similar constant value for the local exploitation at the next 25 iterations.

According to Fig. 7, we can conclude that GWO shows a good balance between global exploration and local exploitation that results in high local optima avoidance and a fast convergence simultaneously. This superior capability is due to the adaptive value of A. As mentioned above, half of the iterations are devoted to exploration ($|A| \geq 1$) and the rest to exploitation ($|A| < 1$). In addition, the C parameter always randomly obliges the search agents to take random steps towards/outwards the prey. This

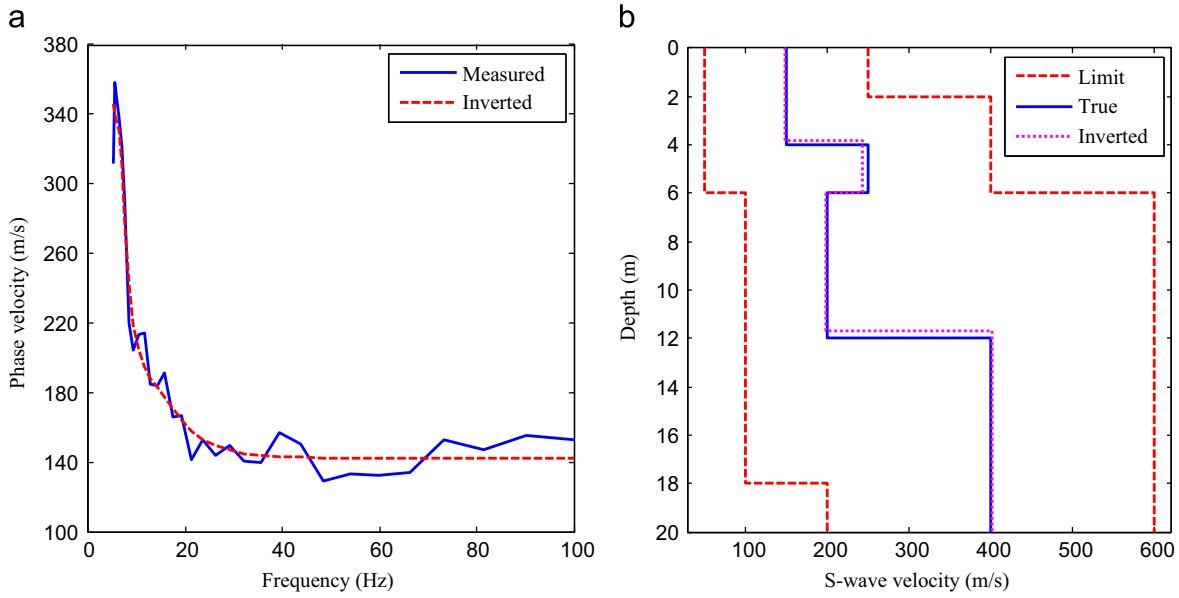

Fig. 6. Inversion results of Model C using the GWO algorithm. (a) Contaminated synthetic data (solid line) and modeled dispersion curve (dashed line). (b) The lower and upper bounds of the search area (dashed lines), True model (solid line) and inverted profile (dotted line).

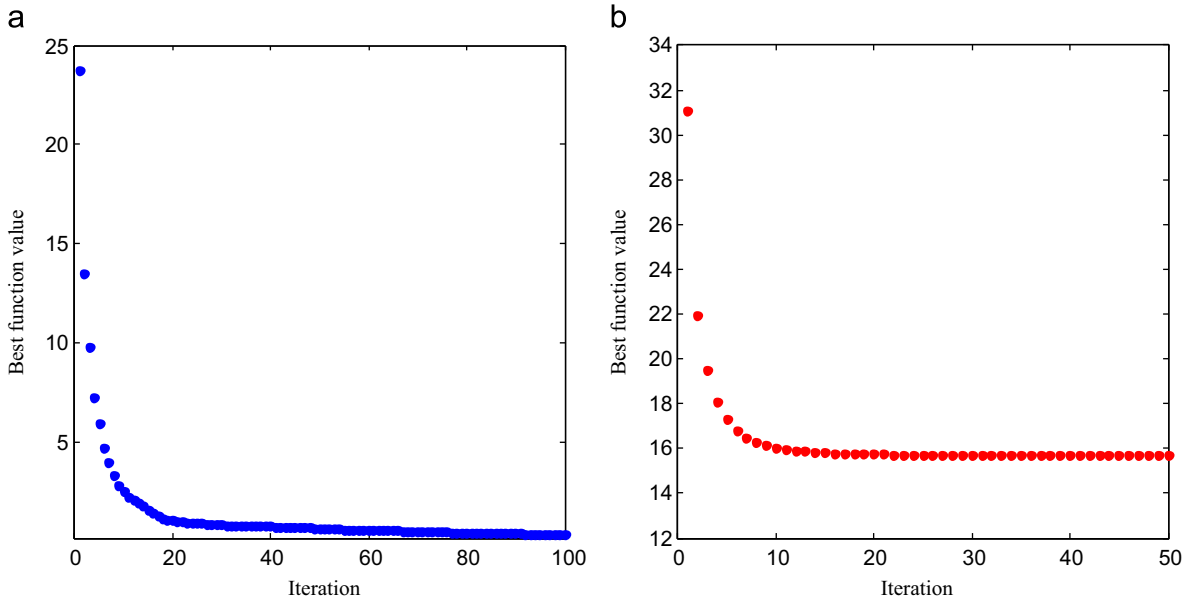

Fig. 7. Fitness behavior of GWO iteration using noise-free (a) and noisy (b) synthetic data from Model A.

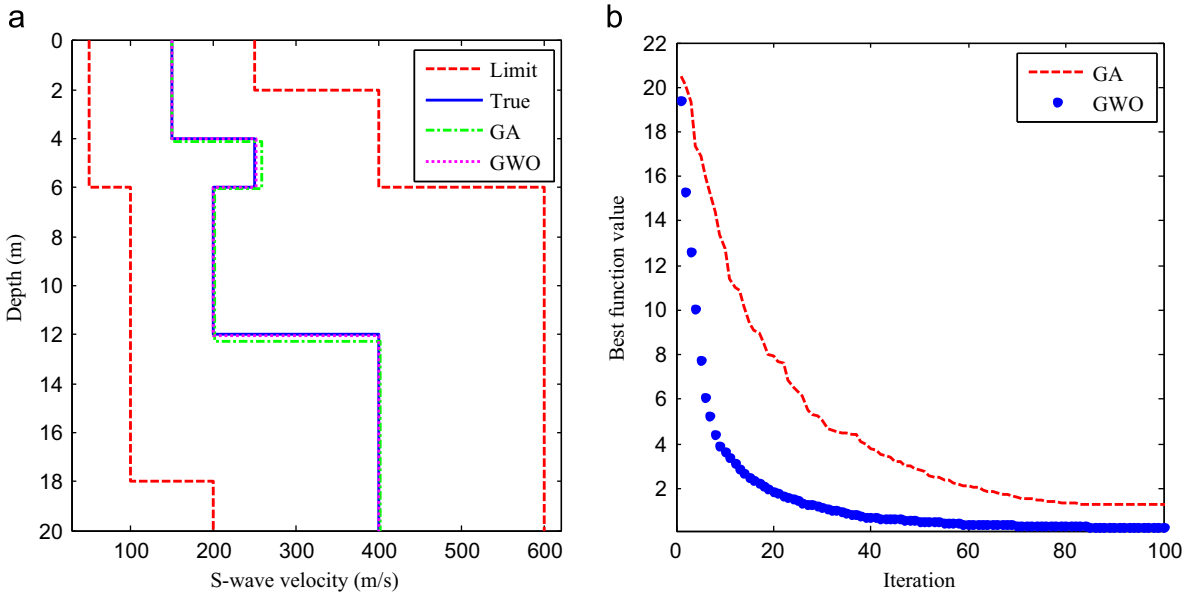

Fig. 8. Comparisons of inverted results between GA and GWO on noise-free synthetic data from Model C. (a) S-wave velocity profiles and the search space. (b) The convergence behavior of GA and GWO iteration.

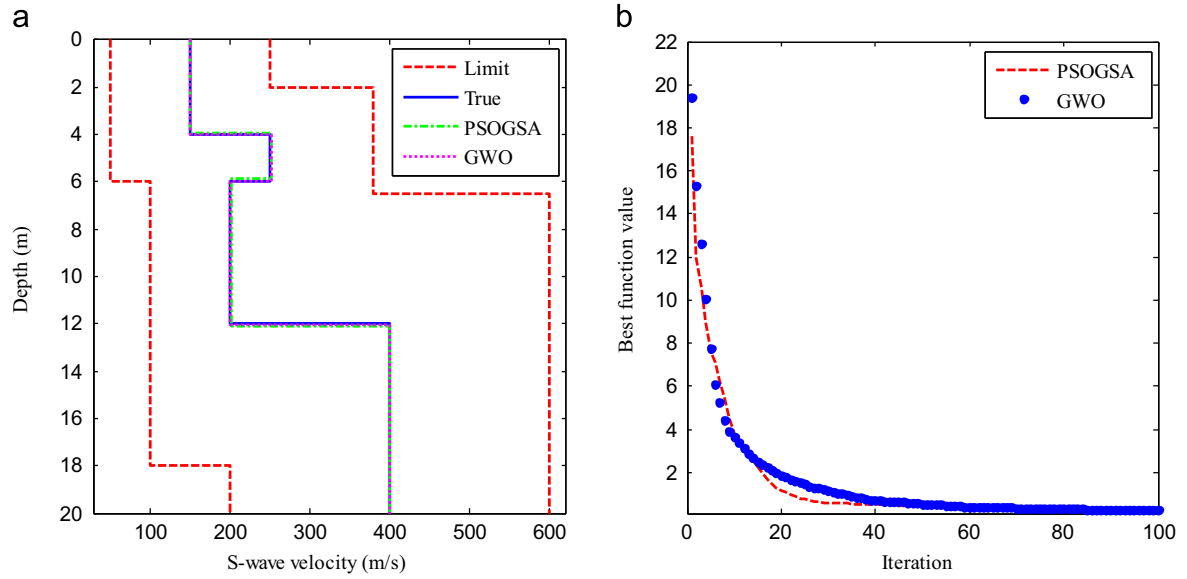

Fig. 9. Comparisons of inverted results between PSOGSA and GWO on noise-free synthetic data from Model C. (a) S-wave velocity profiles and the search space. (b) The convergence behavior of PSOGSA and GWO iteration.

mechanism is very helpful for resolving local optima stagnation even when the GWO algorithm is in the exploitation phase.

5. Comparisons with other algorithms

For verification, the results of the GWO algorithm are compared to genetic algorithm (GA) and the hybrid algorithm (PSOGSA)-the combination of Particle Swarm Optimization (PSO) and Gravitational Search Algorithm (GSA) [49]. The noise-free synthetic data set from Model C as used in GWO inversion has been employed to implement parameter estimation in surface waves by GA and PSOGSA. For comparisons, the same search space, the population size, and maximum iteration/generation number have been utilized in the present study. In GA inversion, the probabilities of selection, crossover, and mutation are set as 0.8, 0.6, and 0.02, respectively. For PSOGSA, we use the following parameter sets: Gravitational constant $G_0 = 1$; Descending coefficient $\alpha = 23$; Weighting function $w = rand()$, where $rand()$ is random number in [0,1]; Weighting factors $c_1 = 0.5$, $c_2 = 1.5$.

Table 2
Statistical results on noise-free synthetic data sets from Model C.

Model number	GWO			GA			PSOGSA		
	AVE	STD	OAE	AVE	STD	OAE	AVE	STD	OAE
Model C	0.241	0.057	0.28%	1.256	0.540	1.96%	0.322	0.111	1.46%

Comparisons of inversion results between GA and GWO are demonstrated in Fig. 8. Comparisons of inverted results between PSOGSA and GWO are shown in Fig. 9. Table 2 reports statistics of inversion results in the form of AVE, STD and OAE for each algorithm. The statistical results that are presented are average

(AVE) and standard deviation (STD) of the obtained best Mean Square Error (MSE) in the last iteration as well as overall average errors (OAE) of estimated model parameters by the algorithms.

Table 2 and Fig. 8 show that GWO outperforms GA in terms of not only the minimum MSE but also the accuracy of solutions. It may be observed in Fig. 8b that the search agents of GWO tend to extensively explore promising regions of the search spaces in the first 50 iterations and exploit the best one in the next 50 iterations. It can be clearly seen in Fig. 8b that GA has a fast initial convergence in the first 70 generations, followed by progressively slower improvements as it approaches the minimum MSE at the next 30 generations.

Obviously, lower average (AVE) and standard deviation (STD) of MSE as well as overall average error (OAE) of solutions in the last iteration (Table 2) indicates the superior performance of GWO. The reason for improved MSE is the high local optima avoidance of this algorithm. According to the mathematical formulation of the GWO algorithm, half of the iterations are devoted to exploration of the search space (when $|A|>1$). This promotes exploration of the search space that leads to finding diverse search areas during optimization. In addition, the C parameter always randomly obliges the search agents to take random steps towards/outwards the prey. This mechanism is very helpful for resolving local optima stagnation even when the GWO algorithm is in the exploitation phase. The reason for the high accuracy of the solutions provided by the GWO algorithm is that this algorithm is equipped with adaptive parameters to smoothly balance exploration and exploitation. Half of the iteration is devoted to exploration (when $|A|>1$) and the rest to exploitation (when $|A|<1$). In addition, the GWO algorithm always saves the three best obtained solutions at any stage of optimization. Consequently, there are always guiding search agents for exploitation of the most promising regions of the search space. In other words, GWO benefits from intrinsic exploitation guides, which also assist this algorithm to provide remarkable results. These mechanisms described above assist GWO to provide very good exploration, exploitation, local optima avoidance, and fast convergence simultaneously [46].

It is also worth discussing the poor performance of GA in this subsection. Generally speaking, GA has been designed based on various mutation mechanisms [46]. Mutation in GA maintains the diversity of population and promotes exploitation, which are the first main reasons for the poor performance of GA. In addition, selection of individuals in this algorithm is done by a deterministic approach. Consequently, the randomness in selecting an individual is less and therefore local optima avoidance is less as well. This is the second reason why GA failed to provide good results in this model. Furthermore, the chromosome of the best individual might also be destroyed by the crossover. This is likely to be the third reason for the poor performance of GA.

The results of Table 2 and Fig. 9 reveal that GWO and PSOGSA have the best performances on Model C in terms of improved MSE. The average and standard deviation of MSE show that the performances of these two algorithms are very close. The reconstruction accuracy of the GWO algorithm, however, is much higher. These results are strong evidences for the efficiencies of GWO in surface wave analysis. The results testify that the GWO algorithm has superior local optima avoidance and fast convergence simultaneously.

The search space of parameter estimation problem in surface waves is usually unknown and very complex with a massive number of local optima. Although gradient-based algorithms are much faster and without extra computational cost, they have the high probability of entrapment in local optima. So GWO is a good option for optimizing this challenging real problem. To further highlight this feature, we utilize gradient-based algorithm in MASW [2] to invert Model C again with a 7-layer model, each thin layer being 2 m thick. Fig. 10 demonstrates inverted results of MASW on noise-free synthetic data from Model C. It can be noted that the inverted model (dashed line in Fig. 10b) is characterized by a general pattern of oscillation between the third thin layer and the sixth thin layer although the inverted dispersion curve (solid line in Fig. 10a) explains the measured dispersion curve (solid dots in Fig. 10a) perfectly. Four maximum relative errors in the deduced S-wave velocity model from the third to the sixth thin layer are 3.56%, 9.33%, 10.73%, and 5.93%, respectively. Once more, these results evidence the merits of the proposed GWO algorithm in parameter estimation in surface waves.

6. Field data inversion

To further investigate and evaluate the applicability and robustness of GWO described above, surface wave data (solid dots in Fig. 11a) acquired in a waste disposal site in NE Italy [35] have been reanalyzed using the GWO approach in the present study. This is essentially characterized by an 18-m-thick unconsolidated sediment sequence lying over a fractured limestone basement. A number of

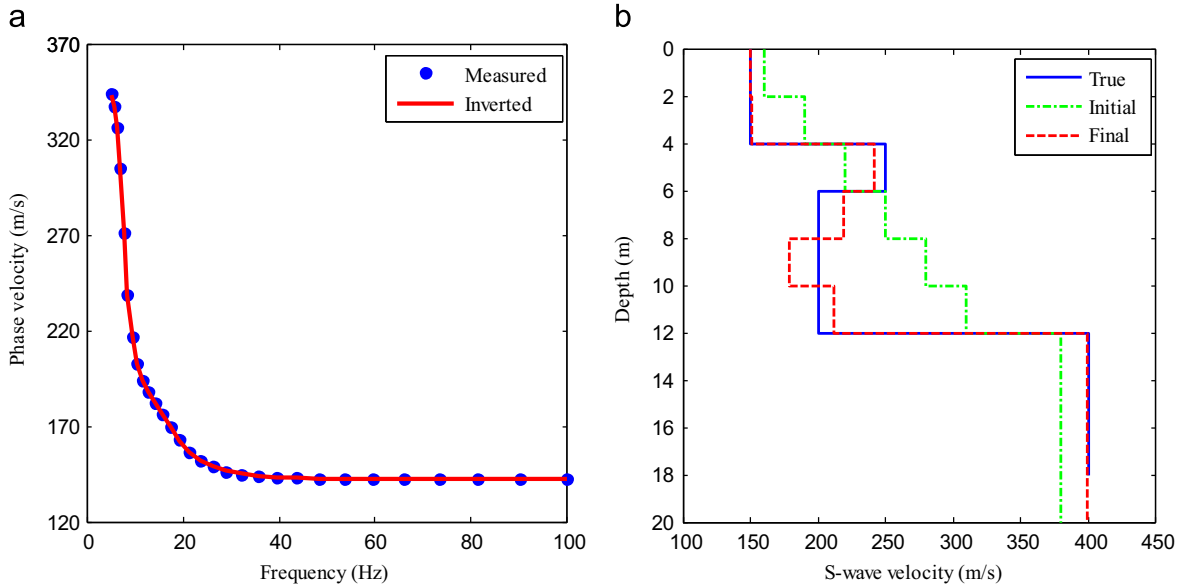

Fig. 10. Inverted results of MASW on noise-free synthetic data from Model C. (a) Measured and inverted dispersion curves. (b) True, Initial, and Final S-wave velocity profiles.

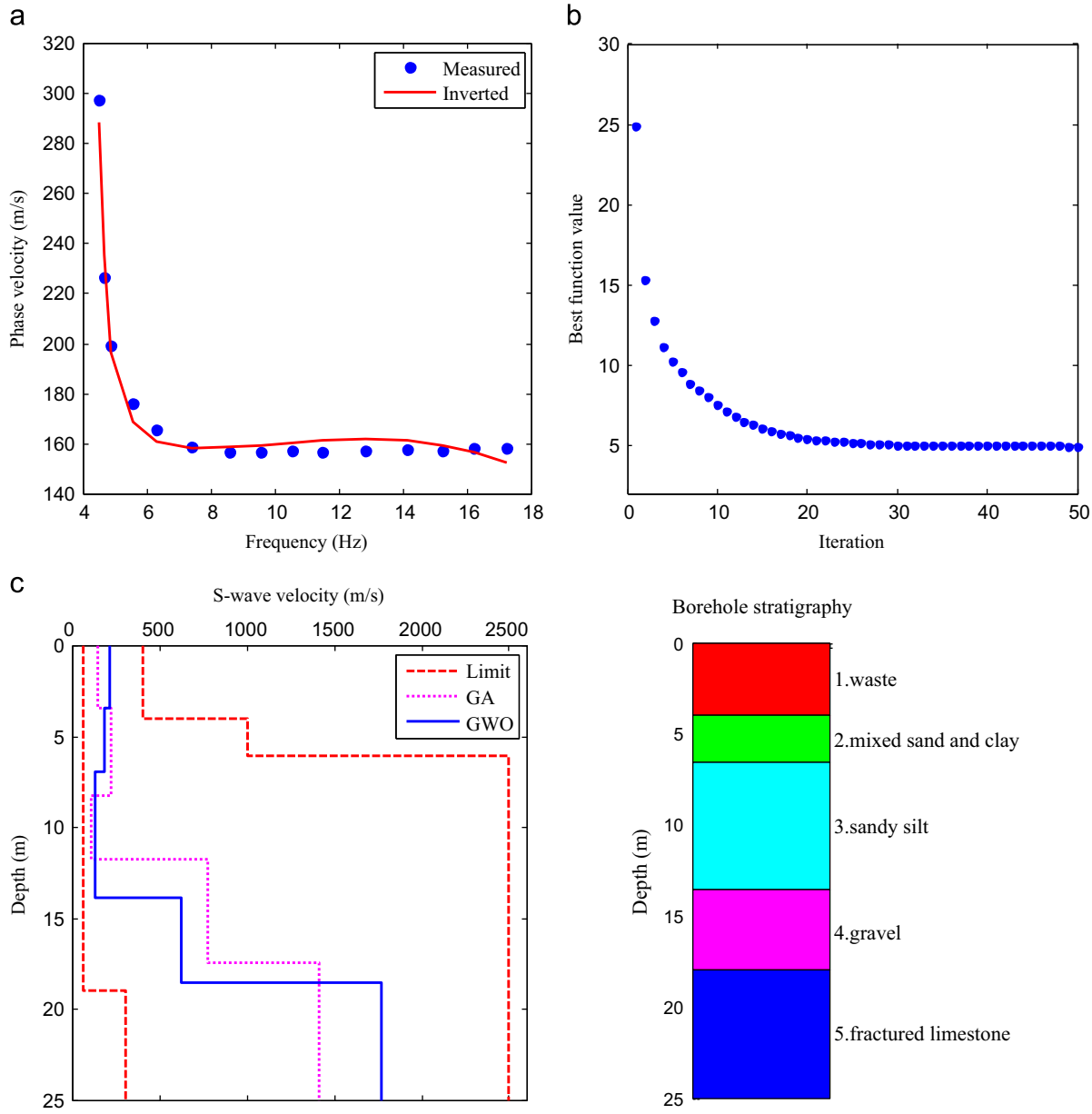

Fig. 11. A real-world example from a waste disposal site in NE Italy. (a) Solid dots and solid line represent measured phase velocities [35] and calculated dispersion curves from the GWO-deduced model, respectively. (b) The convergence process of GWO implementation. (c) The lower and upper bounds of the search space (dashed lines), and inverted S-wave velocity profiles from GA (dotted line) [35] and GWO (solid line). Also shown is the borehole stratigraphic column.

geophysical surveys were conducted and results were compared with borehole data and measurements on samples with the goal of determining the identification power of each methodology and classifying waste typology and extension to plan future remediation acts. Because of the poor geotechnical characteristics of the uppermost layers and in particular the extremely heterogeneous and unconsolidated superficial waste level, seismic records present a limited spectral content and, as far it concerns the P-wave component, a low signal-to-noise ratio strongly dominated by ground roll components [35].

Similar to the inverse strategy of noisy synthetic data sets, we considered S-wave velocities and layer thicknesses as variables while fixing densities and Poisson's ratio according to estimated values commonly adopted for the relevant stratigraphic units (0.45–0.40 for variously silty alluvia, 0.25 for gravel and 0.20 for the limestone bedrock). Search space, Poisson's ratio and densities adopted for nonlinear inversion of real data by the GWO algorithm are reported in Table 3. During the inversion procedure, a 5-layer subsurface structure suggested by Dal Moro et al. [35] is adopted to perform inversion of the observed dispersion curve by the proposed inverse procedure.

Table 3
Field dataset: search space, Poisson's ratio [35] and density adopted for the inversion of surface waves based on the GWO algorithm.

Layer number	S-wave velocity (m/s)	Thickness (m)	Poisson's ratio values	Density (g/cm^3)
1	60–400	1–5	0.45	1.9
2	60–400	1–5	0.45	1.9
3	60–400	2–8	0.40	2.0
4	300–1000	2–8	0.25	2.2
5	900–2500	Half-space	0.20	2.4

Inversion results of GWO for the real example are illustrated in Fig. 11. Similar to Fig. 7b, the misfit values in Fig. 11b significantly decrease for global exploration in the first 25 iterations, and then gradually converge to a similar constant value for local exploitation in the next 25 iterations. Fig. 11c reports the best solutions of the implementation from GA [35] and GWO. As shown in Fig. 11c, the GWO-deduced model (solid line) is consistent with the GA-inverted model (dotted line) for the first three layers of the vertical profile. In particular, the GWO-estimated profile is in fairly good agreement with borehole stratigraphic also shown in Fig. 11c. The depths of the velocity discontinuities between all layers, including the two deepest silt-gravel and gravel-limestone contacts, are well delineated by GWO. The modeled dispersion curve (solid line in Fig. 11a) from the GWO-deduced solution fits the measured phase velocities (solid dots in Fig. 11a) reasonably well.

7. Conclusions

This research proposed a novel and powerful surface wave inversion scheme called Grey Wolf Optimizer (GWO) inspired by grey wolves. The GWO algorithm mimics the social leadership hierarchy and hunting behavior of grey wolves in nature. Four types of grey wolves such as alpha, beta, delta, and omega are employed for simulating the leadership hierarchy. Three main steps of hunting, searching for prey, encircling prey, and attacking towards prey, are implemented to perform parameter estimation in surface waves. The proposed strategy is benchmarked on noise-free, noisy, and field data in terms of exploration, exploitation, local optima avoidance, and convergence. In the inversion procedure, we adopted wider search space boundaries to simulate more realistic cases where no *a priori* information is available and in order to further investigate the performance of the proposed algorithm. The search spaces are approximately 50% off the true values in both synthetic and field data tests. For verification, the results of the GWO algorithm are compared to GA, PSOGSA, and MASW. Results from both synthetic and real data demonstrate that GWO shows a good balance between exploration and exploitation that results in high local optima avoidance and a very fast convergence simultaneously. According to this comprehensive study, the GWO algorithm is highly recommended to be used in parameter estimation in surface waves. The great advantages of GWO are that the algorithm is simple, flexible, robust, and easy to implement. Also there are fewer parameters to tune.

This preliminary research is the first attempt to invert surface wave data for near-surface S-wave velocity profiles by GWO. This work provides quite valuable insights into the performance of GWO for other geophysical inversions as well. For future work, we will use other recent algorithms such as Mind Blast algorithm, Ant Lion Optimizer [52], Ions Motion Algorithm, and Multi-Verse Optimizer [48] to solve the same problem.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (NSFC) (No. 41174113) and the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) (No. CUG130103). The authors greatly appreciate Dr. Seyedali Mirjalili for providing his excellent GWO and PSOGSA MATLAB codes to conduct this study. The authors owe special thanks to two anonymous reviewers for their valuable and constructive suggestions and encouraging comments, which improve the quality of the manuscript significantly.

References

[1] Park CB, Miller RD, Xia J. Multichannel analysis of surface waves. Geophysics 1999;64(3):800–8.
[2] Xia J, Miller RD, Park CB. Estimation of near-surface shear-wave velocity by inversion of Rayleigh wave. Geophysics 1999;64(3):691–700.
[3] Zeng C, Xia J, Miller RD, Tsoflias GP. Application of the multiaxial perfectly matched layer (MPML) to near-surface seismic modeling with Rayleigh waves. Geophysics 2011;76(3):T43–52.
[4] Socco LV, Foti S, Boiero D. Surface-wave analysis for building near-surface velocity models-established approaches and new perspectives. Geophysics 2010;75(5):A83–102.
[5] Miller RD, Xia J, Park CB, Ivanov J. Multichannel analysis of surface waves to map bedrock. Lead. Edge 1999;18:1392–6.
[6] Lai CG, Rix GJ. Solution of the Rayleigh eigenproblem in viscoelastic media. Bull. Seismol. Soc. Am. 2002;92(6):2297–309.
[7] Xia J. Estimation of near-surface shear-wave velocities and quality factors using multichannel analysis of surface-wave methods. J. Appl. Geophys. 2014;103:140–51.
[8] Xia J, Xu Y, Miller RD, Ivanov J. Estimation of near-surface quality factors by constrained inversion of Rayleigh-wave attenuation coefficients. J. Appl. Geophys. 2012;82:137–44.
[9] Lin C-P, Chang C-C, Chang T-S. The use of MASW method in the assessment of soil liquefaction potential. Soil Dyn. Earthq. Eng. 2004;24(9–10):689–98.
[10] Ivanov J, Miller RD, Lacombe P, Johnson CD, Lane Jr. JW. Delineating a shallow fault zone and dipping bedrock strata using multichannal analysis of surface waves with a land streamer. Geophysics 2006;71(5):A39–42.
[11] Ryden N, Park CB. Fast simulated annealing inversion of surface waves on pavement using phase-velocity spectra. Geophysics 2006;71(4):R49–58.
[12] Ryden N, Park CB, Ulriksen P, Miller RD. Multimodal approach to seismic pavement testing. J. Geotech. Geoenviron. Eng. 2004;130(6):636–45.
[13] Foti S, Parolai S, Albarello D, Picozzi M. Application of surface-wave methods for seismic site characterization. Surv. Geophys. 2011;32(6):777–825.
[14] Dal Moro G. V_S and V_P vertical profiling via joint inversion of Rayleigh waves and refraction travel times by means of bi-objective evolutionary algorithm. J. Appl. Geophys. 2008;66(1–2):15–24.
[15] Ivanov J, Miller RD, Xia J, Steeples DW, Park CB. Joint analysis of refractions with surface waves: an inverse solution to the refraction-traveltime problem. Geophysics 2006;71(6):R131–8.
[16] Dal Moro G, Pipan M. Joint inversion of surface wave dispersion curves and reflection travel times via multi-objective evolutionary algorithms. J. Appl. Geophys. 2007;61(1):56–81.
[17] Dal Moro G, Ferigo F. Joint analysis of Rayleigh-and Love-wave dispersion: issues, criteria and improvements. J. Appl. Geophys. 2011;75(3):573–89.
[18] Lai CG, Rix GJ, Foti S, Roma V. Simultaneous measurement and inversion of surface wave dispersion and attenuation curves. Soil Dyn. Earthq. Eng. 2002;22(9–12):923–30.
[19] Lin C-P, Chang T-S. Multi-station analysis of surface wave dispersion. Soil Dyn. Earthq. Eng. 2004;24(11):877–86.
[20] Tian G, Steeples DW, Xia J, Miller RD, Spikes KT, Ralston MD. Multichannel analysis of surface wave method with the autojuggie. Soil Dyn. Earthq. Eng. 2003;23(3):243–7.
[21] Tian G, Steeples DW, Xia J, Spikes KT. Useful resorting in surface wave method with the autojuggie. Geophysics 2003;68(6):1906–8.
[22] Zhang SX, Chan LS, Chen CY, Dai FC, Shen XK, Zhong H. Apparent phase velocities and fundamental-mode phase velocities of Rayleigh waves. Soil Dyn. Earthq. Eng. 2003;23:563–9.
[23] Park CB, Miller RD, Ryden N, Xia J, Ivanov J. Combined use of active and passive surface waves. J. Environ. Eng. Geophys. 2005;10(3):323–34.
[24] Forbriger T. Inversion of shallow-seismic wavefields: I. Wavefield transformation. Geophys. J. Int. 2003;153:719–34.
[25] Forbriger T. Inversion of shallow-seismic wavefields: II. Inferring subsurface properties from wavefield transforms. Geophys. J. Int. 2003;153:735–52.
[26] O'Neill A, Dentith M, List R. Full-waveform P-SV reflectivity inversion of surface waves for shallow engineering applications. Explor. Geophys. 2003;34:158–73.
[27] O'Neill A, Matsuoka T. Dominant higher surface-wave modes and possible inversion pitfalls. J. Environ. Eng. Geophys. 2005;10(2):185–201.
[28] Xia J, Miller RD, Park CB, Tian G. Inversion of high-frequency surface waves with fundamental and higher modes. J. Appl. Geophys. 2003;52(1):45–57.
[29] Cercato M. Global surface wave inversion with model constraints. Geophys. Prospect. 2011;59:210–26.
[30] Zhang SX, Chan LS, Xia J. The selection of field acquisition parameters for dispersion images from multichannel surface wave data. Pure Appl. Geophys. 2004;161:185–201.
[31] Zhang SX, Chan LS. Possible effects of misidentified mode number on Rayleigh wave inversion. J. Appl. Geophys. 2003;53:17–29.
[32] Cercato M. Addressing non-uniqueness in linearized multichannel surface wave inversion. Geophys. Prospect. 2009;57:27–47.
[33] Lai CG, Foti S, Rix GJ. Propagation of data uncertainty in surface wave inversion. J. Environ. Eng. Geophys. 2005;10(2):219–28.
[34] Maraschini M, Ernst F, Foti S, Socco LV. A new misfit function for multimodal inversion of surface waves. Geophysics 2010;75(4):G31–43.
[35] Dal Moro G, Pipan M, Gabrielli P. Rayleigh wave dispersion curve inversion via genetic algorithms and marginal posterior probability density estimation. J. Appl. Geophys. 2007;61(1):39–55.

[36] Yamanaka H. Comparison of the performance of heuristic search methods for phase velocity inversion in the shallow surface wave method. J. Environ. Eng. Geophys. 2005;10(2):163–73.
[37] Beaty KS, Schmitt DR. Repeatability of multimode Rayleigh-wave dispersion studies. Geophys. 2003;68:782–90.
[38] Beaty KS, Schmitt DR, Sacchi M. Simulated annealing inversion of multimode Rayleigh-wave dispersion curves for geological structure. Geophys. J. Int. 2002;151:622–31.
[39] Pei D, Louie JN, Pullammanappallil SK. Application of simulated annealing inversion on high-frequency fundamental-mode Rayleigh wave dispersion curves. Geophysics 2007;72(5):R77–85.
[40] Foti S, Comina C, Boiero D, Socco LV. Non-uniqueness in surface-wave inversion and consequences on seismic site response analyses. Soil Dyn. Earthq. Eng. 2009;29:982–93.
[41] Maraschini M, Foti S. A Monte Carlo multimodal inversion of surface waves. Geophys. J. Int. 2010;182(3):1557–66.
[42] Socco LV, Boiero D. Improved Monte Carlo inversion of surface wave data. Geophys. Prospect. 2008;56:357–71.
[43] Mirjalili S, Mirjalili SM, Lewis A, Grey Wolf Optimizer. Adv. Eng. Softw. 2014;69:46–61.
[44] Song HM, Sulaiman MH, Mohamed MR. An application of Grey Wolf Optimizer for solving combined economic emission dispatch problems. Int. Rev. Model. Simul. 2014;7(5):838–44.
[45] Emary E, Zawbaa HM, Grosan C, Hassenian AE. Feature subset selection approach by Gray Wolf Optimizer. Adv. Intell. Syst. Comput. 2015;334:1–13.
[46] Mirjalili S. How effective is the Grey Wolf Optimizer in training multi-layer perceptrons. Appl. Intell. 2015. http://dx.doi.org/10.1007/s10489-014-0645-7.
[47] Saremi S, Mirjalili SZ, Mirjalili SM. Evolutionary population dynamics and Grey Wolf Optimizer. Neural Comput. Appl. 2015. http://dx.doi.org/10.1007/s00521-014-1806-7.
[48] Mirjalili S, Mirjalili SM, Hatamlou A. Multi-Verse Optimizer: a nature-inspired algorithm for global optimization. Neural Comput. Appl. 2015. http://dx.doi.org/10.1007/s00521-015-1870-7.
[49] Mirjalili S, Mohd Hashim SZ. A new hybrid PSOGSA algorithm for function optimization. In: IEEE international conference on computer and information application (ICCIA 2010), China; 2010. p. 374–7.
[50] Muro C, Escobedo R, Spector L, Coppinger RP. Wolf-pack (Canis lupus) hunting strategies emerge from simple rules in computational simulations. Behav. Process. 2011;88(3):192–7.
[51] Buchen PW, Ben-Hador R. Free-mode surface-wave computations. Geophys. J. Int. 1996;124:869–87.
[52] Mirjalili S. The Ant Lion Optimizer. Adv. Eng. Softw. 2015;83:80–98.