Particle Swarm Optimization of 2D Magnetotelluric data

GEOPHYSICS, VOL. 84, NO. 3 (MAY-JUNE 2019); P. E125–E141, 22 FIGS., 6 TABLES.
10.1190/GEO2018-0166.1

Particle swarm optimization of 2D magnetotelluric data

Francesca Pace[1], Alessandro Santilano[2], and Alberto Godio[1]

ABSTRACT

We implement the particle swarm optimization (PSO) algorithm for the two-dimensional (2D) magnetotelluric (MT) inverse problem. We first validate PSO on two synthetic models of different complexity and then apply it to an MT benchmark for real-field data, the COPROD2 data set (Canada). We pay particular attention to the selection of the PSO input parameters to properly address the complexity of the 2D MT inverse problem. We enhance the stability and convergence of the solution of the geophysical problem by applying the hierarchical PSO with time-varying acceleration coefficients (HPSO-TVAC). Moreover, we parallelize the code to reduce the computation time because PSO is a computationally demanding global search algorithm. The inverse problem was solved for the synthetic data both by giving a priori information at the beginning and by using a random initialization. The a priori information was given to a small number of particles as the initial position within the search space of solutions, so that the swarming behavior was only slightly influenced. We have demonstrated that there is no need for the a priori initialization to obtain robust 2D models because the results are largely comparable with the results from randomly initialized PSO. The optimization of the COPROD2 data set provides a resistivity model of the earth in line with results from previous interpretations. Our results suggest that the 2D MT inverse problem can be successfully addressed by means of computational swarm intelligence.

INTRODUCTION

The interpretation of geophysical data requires the solution of the inverse problem, which is, in most cases, nonlinear and ill posed. During the past three decades, global search algorithms as inversion methods have become of growing interest because the probabilistic or evolutionary approach has been adopted to find the optimum solution, which is affected by nonuniqueness. The most important global search algorithms generally used for the inversion of geophysical data are simulated annealing (SA), the genetic algorithm (GA) (Sen and Stoffa, 2013), the ant colony algorithm (ACO) (Yuan et al., 2009), and particle swarm optimization (PSO) (Shaw and Srivastava, 2007).

The inversion of magnetotelluric (MT) data is usually based on algorithms such as Occam, nonlinear conjugate gradient (NLCG), and Gauss-Newton (GN), which are now widely recognized as milestones among two-dimensional (2D) and three-dimensional (3D) MT inversion codes (Avdeev, 2005; Boerner, 2010; Siripunvaraporn, 2012; Ghaedrahmati et al., 2014; Newman, 2014). Even if they ensure convergence in few iterations, they are all based on the local search principle; consequently, the final solution depends on the initial assumption of the starting model. If a homogeneous half-space is adopted as a starting model, some trials have to be done to define the most appropriate value of the electrical resistivity to start with, depending on the data set and inversion code (Miensopust et al., 2013). Otherwise, the inversion should be initially constrained by an a priori model that can resolve the nonuniqueness of the solution by using information from well-log data (Yan et al., 2017a), seismic data (Yan et al., 2017b), MT data (Santilano, 2017), or other geophysical methods. However, if the a priori knowledge is unreliable or unavailable, the initial guess can create a bias in the final result and interpretation (Dong and Jones, 2018). Global search methods have also recently become of pivotal importance in MT, with the essential advantage that the inversion is independent from the starting model. Another advantage of metaheuristic methods, such as GA, SA, and PSO, is that they are theoretically able to find, as the final solution, the global minimum of a function without being trapped in one of several local minima. The solution of the MT inverse problem using SA was first explored by Dosso and Oldenburg (1991), whereas Everett and Schultz (1993) and then Pérez-Flores and Schultz (2002) focus on GA. There are also some recent works adopting the Markov-chain-Monte-Carlo method to solve the 1D MT inversion: Xiang et al. (2018), Conway et al. (2018), and Brodie and Jiang (2018). So

Manuscript received by the Editor 27 February 2018; revised manuscript received 6 September 2018; published ahead of production 25 January 2019; published online 12 March 2019.
[1]Politecnico di Torino, Department of Environment, Land and Infrastructure Engineering (DIATI), Torino, Italy. E-mail: francesca.pace@polito.it (corresponding author); alberto.godio@polito.it.
[2]Institute of Geosciences and Earth Resources-National Research Council (IGG-CNR), Pisa, Italy. E-mail: alessandro.santilano@igg.cnr.it.

far, PSO has been applied by Shaw and Srivastava (2007), Godio and Santilano (2018), and Santilano et al. (2018), although they investigate only the 1D MT inverse problem. PSO application to other geophysical problems includes the interpretation of vertical electrical sounding (Fernández Martínez et al., 2010a), gravity (Darisma et al., 2017), and multitransient electromagnetic data (Olalekan and Di, 2017). The main limitation of these works is that they analyze either 1D field data or oversimplistic 2D synthetic models. Other scientific applications of the PSO algorithm are artificial neural networks, biomedical engineering, hydrogeology, electronics, electromagnetics, power systems, robotics, and signal processing (Poli, 2008; Adhan and Bansal, 2017 and references therein).

The present paper focuses on the implementation of the PSO algorithm for the 2D MT inverse problem. A preliminary application of this method to MT and audio-MT synthetic data has been presented in Pace et al. (2017). The novelty of this paper concerns the validation of the method on two MT synthetic models of different complexity and, for the first time to the authors' knowledge, the application to real-field data, the COPROD2 data set (Jones, 1993a). This data set was made available to the electromagnetic induction scientific community with the aim of comparing different techniques for 2D MT inversion (Jones, 1993b). Since several inversion solutions have been made available so far, COPROD2 represents an interesting (and challenging) field data set for the application of our method. We started from the PSO code of Ebbesen et al. (2012), but then we modified that generic MATLAB code for our specific purpose. The efficiency of the PSO algorithm was improved by applying the principle of hierarchical PSO with time-varying acceleration coefficients (HPSO-TVAC) (Ratnaweera et al., 2004). Previous works on PSO applied to the geophysical inverse problem have always considered constant values for the social and cognitive accelerations of particles (Shaw and Srivastava, 2007; Godio and Santilano, 2018; Santilano et al., 2018). However, this assumption is not adequate for the 2D inverse problem due to its high dimensionality and complex searching behavior. We carried out a detailed sensitivity analysis to find the most appropriate values of the time-varying accelerations: Their iterative variation improved the convergence speed of the algorithm and prevented the solution from being trapped in some local minima. The tuning of the social and cognitive accelerations of the particles was hence crucial to finally achieve the convergence of the solution. In addition, a new parallelized version of the code was developed with the aim of overcoming the time-consuming nature of PSO, which is computationally demanding, like the other global search algorithms. We modified the released PSO code to be run on a high-performance computing (HPC) cluster.

The ensuing sections are organized as follows: First, we explain the way the swarm intelligence principle is applied to the geophysical inverse problem and, specifically, to the 2D MT inverse problem. Then, the first synthetic model is adopted to show the calibration of the main tuning parameters of PSO, the accelerations, and the population size. Once the input arguments are set, the two different synthetic models of MT data are optimized and the obtained results are illustrated. After the method validation with the synthetic models, the PSO algorithm is applied to the COPROD2 data set, and the final resistivity model is compared with results reported in the literature and using well-established algorithms. Finally, the computational improvements are outlined.

PSO APPLICATION TO 2D MT INVERSE PROBLEM

The PSO algorithm is a population-based algorithm that simulates the self-organizing behavior of species living in groups, such as flocks of birds or schools of fish. The way they share knowledge to search for food or find the best reciprocal distance in motion fascinated Kennedy and Eberhart (1995) so strongly that they proposed applying this evolutionary approach to the optimization of nonlinear problems. Simple interactions between individuals yield a complex collective behavior, meaning that each individual is able to adapt and derive new and coherent behaviors in case of changes in the external environment. The most striking feature of this method is that every particle has a memory component that rules its behavior. This is influenced by the cognitive knowledge of the particle and the experience of its neighbors, whose leadership can be emulated. Pivotal references for computational swarm intelligence are Kennedy et al. (2001) and Engelbrecht (2007).

The basic concept of PSO application to geophysics is that each particle of the swarm represents a possible solution of the MT inverse problem, that is an electrical-resistivity model. Since the solution of the problem is affected by nonuniqueness, the search space of solutions needs to be fully explored to find the best model, which fits the observed data. This need is fulfilled by the adaptive and swarming behavior of the particles. During the optimization process, iteration after iteration, the particles "fly" within the search space, bounded between a minimum and a maximum resistivity value. At the end of the swarming, the optimized solution is identified. Readers can find a detailed description of the application of PSO to the geophysical 1D inverse problem in Fernández Martínez et al. (2010a, 2010b) for VES and self-potential methods and Santilano et al. (2018) for MT.

Since the implementation of the PSO algorithm for the 2D MT problem required a high number of particles forming the swarm and numerous iterations to achieve convergence, the standard release of the code for MATLAB appeared to need some modifications. Several variations of the PSO algorithm have been proposed to accelerate convergence and avoid a solution trapped in local minima (Zhan et al., 2009). The PSO variant that showed improved outcomes, with respect to the standard PSO, was the HPSO-TVAC (Ratnaweera et al., 2004). This method takes the social and cognitive behavior of particles into account to enhance the solution convergence and stability.

In the 2D MT problem, the particle of the swarm represents a resistivity model, which is a vector whose elements are the resistivity values of the 2D mesh cells. Each particle of the swarm changes its position $\mathbf{x}$ within the search space by means of the velocity vector $\mathbf{v}$. The vectors $\mathbf{x}$ and $\mathbf{v}$ are updated iteration by iteration according to

$$\mathbf{v}_i^{k+1} = \omega^k \mathbf{v}_i^k + \alpha_1^k \gamma_1 (\mathbf{P}_i - \mathbf{x}_i^k) + \alpha_2^k \gamma_2 (\mathbf{G} - \mathbf{x}_i^k), \quad (1)$$

$$\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \mathbf{v}_i^{k+1}, \quad (2)$$

where $i = [1, \ldots, N]$, N is the number of particles forming the swarm; k is the current iteration number; $\mathbf{x}_i^k$ and $\mathbf{v}_i^k$ are the current vectors of position and velocity of the ith particle, respectively; ω^k is the inertia weight that linearly decreases from 0.9 (first iteration) to 0.4 (last iteration) to balance the momentum remembered from the previous iteration (Shi and Eberhart, 1998); α_1^k is the cognitive acceleration toward the best particle position $\mathbf{P}$, also called the "local best"; α_2^k is the social acceleration toward the best global position $\mathbf{G}$ (or the "global best") found by the group leader; and γ_1 and $\gamma_2 \in [0, 1]$ are the uniformly distributed random values that provide

stochastic perturbation. At the beginning ($k = 0$), the velocity vector ($\mathbf{v}_i^0$) is zero and the position vector ($\mathbf{x}_i^0$) is randomly initialized. Then ($k > 0$), the particle velocity ($\mathbf{v}_i^k$) changes according to three terms: cognitive memory α_1^k, social attraction α_2^k, and inertia component ω^k. Finally, the particle position $\mathbf{x}_i^k$ is updated. Although all of the previous PSO applications to the inverse problem adopted constant accelerations (Fernández Martínez et al., 2010a; Godio and Santilano, 2018; Santilano et al., 2018), we set both acceleration coefficients to vary at each iteration, according to the HPSO-TVAC approach. At the beginning of the HPSO-TVAC optimization, α_1 was larger than α_2, then they linearly reversed. In this way, at the start the diversity of the swarm ensured the search space exploration (high α_1^k), and, at the end, the exploitation of the best regions and the convergence toward the best solution were enabled (high α_2^k). Hence, the resulting adaptive behavior was enhanced. In more detail, the cognitive and social accelerations changed according to, respectively,

$$\alpha_1^k = \alpha_1^{\max} - (\alpha_1^{\max} - \alpha_1^{\min})\left(\frac{k-1}{\max(k)-1}\right), \tag{3}$$

$$\alpha_2^k = \alpha_2^{\min} + (\alpha_2^{\max} - \alpha_2^{\min})\left(\frac{k-1}{\max(k)-1}\right), \tag{4}$$

where α^k is the acceleration value at the current iteration k; $\alpha_1^{\max}$ and $\alpha_2^{\max}$ are the maximum acceleration values for the cognitive and social accelerations, respectively; $\alpha_1^{\min}$ and $\alpha_2^{\min}$ are the minimum acceleration values for the cognitive and social accelerations, respectively; and $\max(k)$ is the maximum number of iterations set for the optimization (Engelbrecht, 2007, p. 313). So, at the first iteration ($k = 1$), $\alpha_1^{k=1} = \alpha_1^{\max}$ and $\alpha_2^{k=1} = \alpha_2^{\min}$, whereas, at the last iteration ($k = \max(k)$), $\alpha_1^{k=\max(k)} = \alpha_1^{\min}$ and $\alpha_2^{k=\max(k)} = \alpha_2^{\max}$. In addition, the accelerations were chosen to satisfy the stability solution conditions (Perez and Behdinan, 2007):

$$\alpha_1 + \alpha_2 < 4, \tag{5}$$

$$\frac{\alpha_1 + \alpha_2}{2} - 1 < \omega < 1. \tag{6}$$

The values of the accelerations influenced the way the particles explored the model space and changed their trajectory with respect to the local and global bests. A thorough sensitivity analysis on the PSO control parameters, accelerations and inertia, can be found in Ratnaweera et al. (2004) and Fernández Martínez et al. (2010a, 2010b). These works tested several benchmark functions and identified the best ranges of these values, ensuring the convergence and stability of the PSO algorithm. Starting from their results, and obeying equations 5 and 6, we performed some tests to assess the influence of several acceleration val-

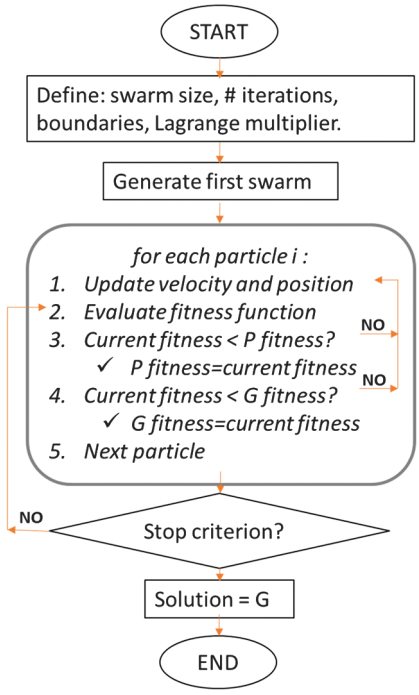

Figure 1. The PSO algorithm flowchart: **P** is the local best solution and **G** is the global best solution.

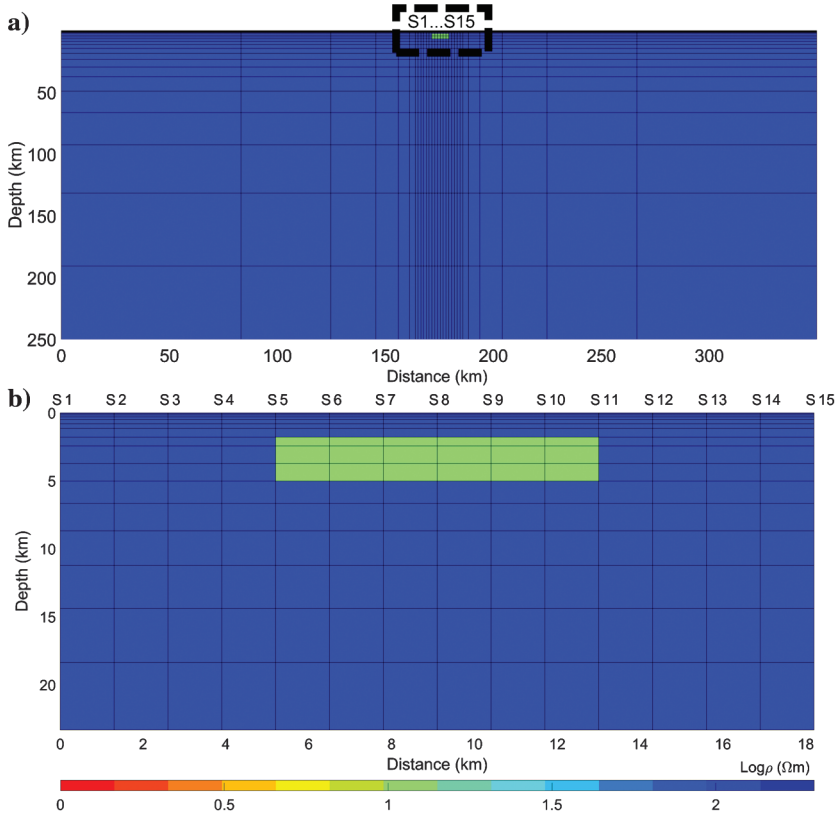

Figure 2. Synthetic model 1: (a) The 2D mesh is discretized into 33 layers and a total of 957 grid cells. The labels S1, ..., S15 indicate the location of the 15 MT stations. The dashed area is shown in (b): A 10 ohm-m conductive body is hosted in a 100 ohm-m medium.

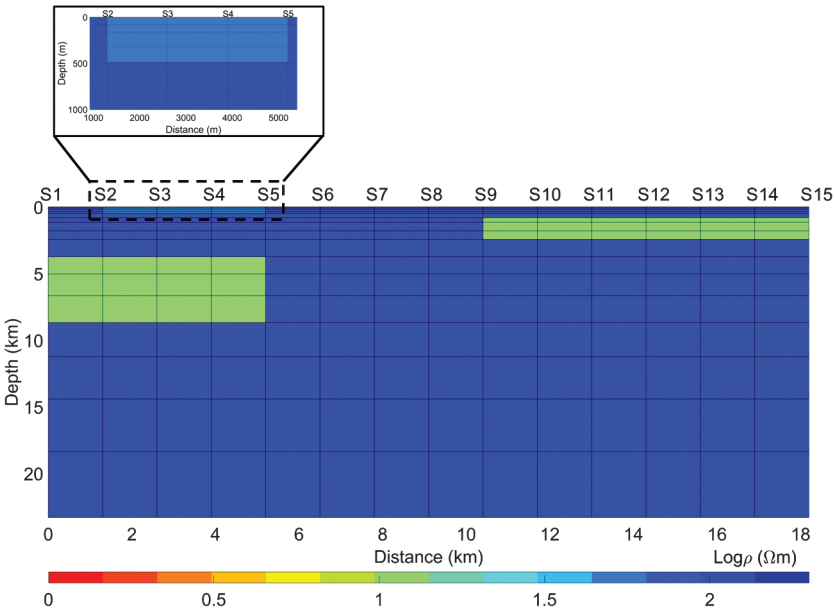

Figure 3. Synthetic model 2: Two 10 ohm-m deep anomalies and one superficial 50 ohm-m body are embedded in a 100 ohm-m host medium. The labels S1, S2, . . . , S15 indicate the 15 MT stations. The magnified box on the top shows the 50 ohm-m body below S2–S5.

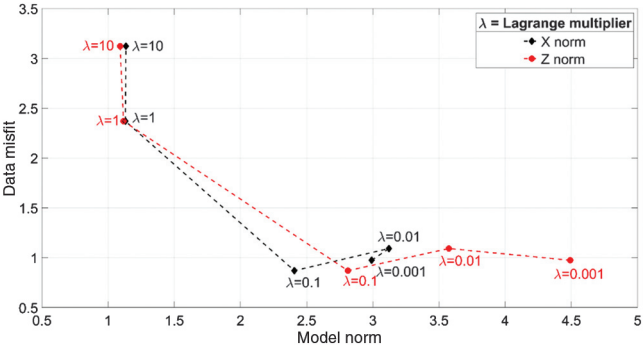

Figure 4. The L-curve response for synthetic model 1 along the horizontal (black diamonds) and vertical (red circles) directions. The trade-off between the data misfit and the model norm indicates the best Lagrange multiplier λ equal to 0.1.

Table 1. Synthetic data from example 1 were adopted to perform the calibration of the cognitive acceleration α_1 and social acceleration α_2 starting from different values at the first iteration ($k = 1$).

		$\alpha_2^{k=1} = 0.25$	$\alpha_2^{k=1} = 0.5$	$\alpha_2^{k=1} = 0.75$
$\alpha_1^{k=1} = 2.75$	rms	0.91	2.17	2.18
	$F(\mathbf{m})$	1.34	2.88	3.03
$\alpha_1^{k=1} = 2$	rms	0.88	0.86	0.91
	$F(\mathbf{m})$	1.47	1.37	1.33
$\alpha_1^{k=1} = 1.5$	rms	1.11	1.01	0.87
	$F(\mathbf{m})$	1.73	1.52	1.4444

Note: The final values of the rms error and fitness function $F(\mathbf{m})$ are listed for each test.

ues on the solution of the MT inverse problem. For cognitive and social accelerations, we adopted three different maximum values, $\alpha_1^{\max}$ and $\alpha_2^{\max}$ equal to 1.5, 2, and 2.75, and three different minimum values, $\alpha_1^{\min}$ and $\alpha_2^{\min}$ equal to 0.25, 0.5, and 0.75. This sensitivity analysis was applied to synthetic model 1, and the results are presented in the next section.

The fitness function

The final goal of the optimization process is the minimization of the fitness or objective function. The particle with the lowest fitness value is awarded with the global best position **G** and is going to attract neighbors depending on the social acceleration α_2. The fitness function that we adopted was the same as that of Everett and Schultz (1993) for the calculation of the data misfit, whereas the Occam-like regularization was added as proposed by deGroot-Hedlin and Constable (1990). Therefore, for 2D MT data, the function to be minimized was

$$F(\mathbf{m}) = \left(\frac{1}{M}\left\|\frac{\log(\boldsymbol{\rho}_{\mathbf{a,o}}) - \log(\boldsymbol{\rho}_{\mathbf{a,p}})}{\log(\boldsymbol{\Delta\rho}_{\mathbf{a,o}})}\right\|_2^2 + \frac{1}{M}\left\|\frac{\boldsymbol{\phi}_{\mathbf{o}} - \boldsymbol{\phi}_{\mathbf{p}}}{\boldsymbol{\Delta\phi}_{\mathbf{o}}}\right\|_2^2\right)^{1/2} + \lambda_x\|\partial_x\mathbf{m}\|_2 + \lambda_z\|\partial_z\mathbf{m}\|_2, \tag{7}$$

where $\boldsymbol{\rho}_{\mathbf{a,o}}$ and $\boldsymbol{\rho}_{\mathbf{a,p}}$ are the observed and predicted apparent resistivity, respectively; $\boldsymbol{\Phi}_{\mathbf{o}}$ and $\boldsymbol{\Phi}_{\mathbf{p}}$ are the observed and predicted impedance phases, respectively; $\boldsymbol{\Delta\rho}_{\mathbf{a,o}}$ and $\boldsymbol{\Delta\Phi}_{\mathbf{o}}$ are the errors in observed apparent resistivity and phase, respectively; M is the number of degrees of freedom, i.e., the number of evaluated data; λ_x and λ_z are the Lagrange multipliers in the x- and z-directions, respectively, set as the trade-off between the model and data misfit to regulate the model roughness, and $\partial_x\mathbf{m}$ and $\partial_z\mathbf{m}$ are the first derivatives of the model solution along the x- and z-directions, respectively. The solution **m** is the electrical-resistivity model, i.e., the vector of resistivity values of the 2D domain. This vector has as many elements as the grid cells of the 2D mesh and is represented by the particles of the swarm. At each iteration, the particle that best minimizes the objective function is assumed as the global best solution (**G**), whereas the other particles can be either attracted or driven away looking for other solutions in the search space. At the end of the optimization, the particle with the minimum $F(\mathbf{m})$ is selected as the final solution and most of the other particles converge to it (swarming behavior). Apparent-resistivity values were transformed to their logarithmic values because they can cover different orders of magnitude. The first part in the right side of equation 7 addresses the minimization between the observed data, apparent resistivity ($\boldsymbol{\rho}_{\mathbf{a}}$) and impedance phase ($\boldsymbol{\Phi}$), and predicted data computed by the forward modeling. This calculation of the misfit is defined as the square root of the sum of two squared Euclidean norms because $\boldsymbol{\rho}_{\mathbf{a}}$ and $\boldsymbol{\Phi}$ can have different orders of magnitude and ranges. The forward modeling incorporates the physics of the problem and,

starting from the assumed model $\mathbf{m}$, predicts the responses $\boldsymbol{\rho}_\mathbf{a}$, and $\boldsymbol{\Phi}$ for each particle of the swarm. The remaining part of equation 7 was added to minimize the roughness of the model solution $\mathbf{m}$: In the horizontal and vertical directions, the differencing operator on $\mathbf{m}$ was weighted by the Lagrange multiplier λ. This approach is the Occam-like optimization and has been adopted for the 1D MT problem in Godio and Santilano (2018) and Santilano et al. (2018). In this way, the minimization of the objective function looks for the smoothest model that fits the data, thus ensuring a balance between the data fitting and the roughness of the model. The value of λ was appropriately chosen following the L-curve criterion (Farquharson and Oldenburg, 2004). It consists in finding the optimal trade-off between the misfit of the data and the roughness of the final model (i.e., the model norm) in the horizontal and vertical directions. The synthetic and real models analyzed in this work had their specific optimal value of λ.

PSO input arguments

The PSO algorithm was iterated enough to guarantee as robust a minimization of the fitness function as possible. Previous PSO applications adopted the maximum number of iterations as the unique stopping criterion (Godio and Santilano, 2018; Santilano et al., 2018). However, the number of iterations is problem dependent and its arbitrary choice can lead to either an ending before the solution convergence or unnecessary computation (Engelbrecht, 2007). In this work, we took into account the fitness-function trend during the minimization process. PSO ran as long as the fitness value did not minimize for 80 consecutive iterations or, if this condition was not satisfied, up to a maximum number of 6000 iterations. Another stopping criterion was the minimum root-mean-square (rms) error of the data equal to 1 ($\pm$10% of tolerance), to avoid the fitting of the data below their uncertainty (deGroot-Hedlin and Constable, 1990).

The swarm size, i.e., the number of particles, influences the way that particles distribute over the search space to guarantee the exploration of possible solutions. The swarm size must be sufficiently high to ensure a wide initial coverage of the search space, so that the particles can efficiently explore all of the regions potentially hosting the global minimum. This behavior is missed if the swarm is too small, although giving the advantage of unburdening the computational complexity. An interesting analysis on the relation between the swarm size and the computational complexity can be found in Van den Bergh and Engelbrecht (2001). The number of particles is a problem-dependent parameter, and it is usually set proportional to the number of unknowns, that is, for us, the number of resistivity cells the 2D domain was discretized into. The ratio between the problem unknowns and the number of particles was suggested to be between 8 and 12 times the unknowns by Engelbrecht (2007, p. 241) for GA and Fernández Martínez et al. (2010a) for PSO. Starting from these guidelines, we performed a sensitivity analysis to verify the influence of this ratio on the solution of the MT inverse problem. The number of particles was set at 6, 8, 9, 10, and 12 times the number of unknowns. This analysis was carried out on synthetic model 1, and the results are shown in the next section.

The initialization of the optimization is another essential feature of PSO. At the beginning, the particle distribution within the search space is, by default, completely random and bounded between a minimum and maximum value of resistivity. This range is kept constant during the optimization but can vary from each layer (or group of layers or cells) to another (Godio and Santilano, 2018). The decision of the lower and upper resistivity boundaries is problem dependent and should be coherent with the desired coverage of the

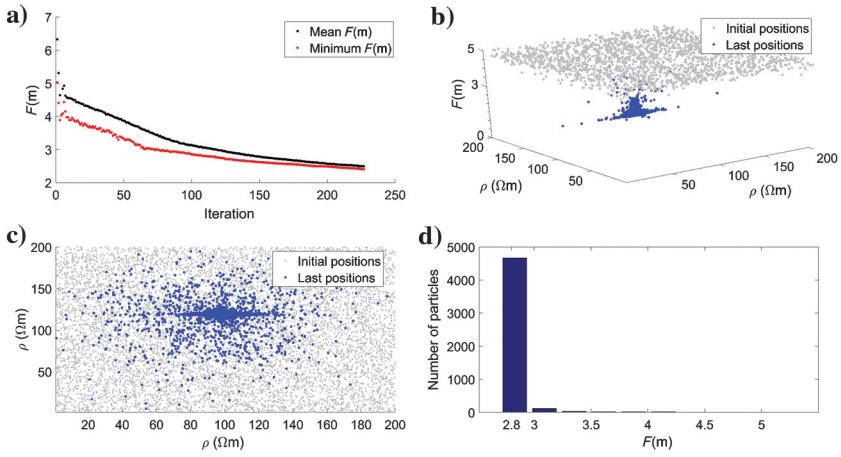

Figure 5. Fitness function $F(\mathbf{m})$ and particle positions at the end of the optimization: (a) Fitness-function value, iteration after iteration, for the best particle (red dots) and the rest of the swarm (black dots), (b) the fitness-function value as a function of the particle positions in the resistivity (ρ) search space, at the first (gray dots) and final (blue dots) iterations, (c) plain view of (b), and (d) final distribution of the fitness-function values among all the particles.

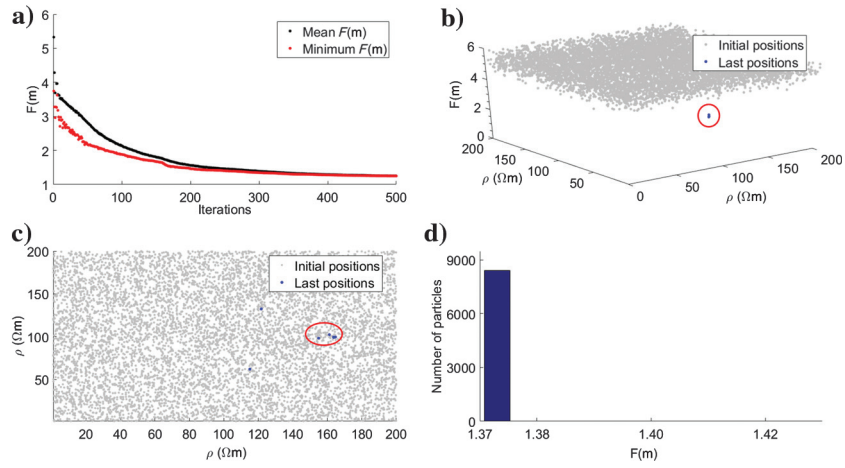

Figure 6. Fitness function $F(\mathbf{m})$ and particle positions at the end of the optimization: (a) Fitness-function value, iteration after iteration, for the best particle (red dots) and the rest of the swarm (black dots), (b) the fitness-function value as a function of the particle positions in the resistivity (ρ) search space, at the first (gray dots) and final (blue dot) iterations, (c) plain view of (b) with all particles converged to the last position (red-circled blue dot), and (d) final distribution of the fitness-function values among all the particles.

search space of solutions. We set the boundaries far larger than the limits of the apparent-resistivity curves. After the random initialization, the adaptive behavior controls the position updating and a stochastic perturbation is guaranteed by γ_1 and γ_2 of equation 1. Local search algorithms usually deploy a starting model (a homogeneous or a priori model) to initialize the geophysical inversion. The a priori information is derived from geologic (well-log) data or other geophysical methods. Although it is possible to use a priori information to partially influence the swarm behavior, the key factor of global search algorithms such as PSO is that they do not require a starting model. To demonstrate this, synthetic data were optimized starting with and without aprioristic information, which was given to the particles in the form of starting positions in the search space. This a priori information was given only to a small amount of particles, 5% of the total, so that the initial position of the rest of the swarm was randomly selected and the swarming nature of PSO was obeyed. We derived the a priori information from the solution PSO gave for the 1D MT inverse problem (Santilano, 2017).

The PSO flow chart is shown in Figure 1. This procedure was repeated three times (or "trials") for each study case, due to the variability on the solution given by the random initialization. In fact, the final solutions coming from different initial random distributions are quite similar but not identical, as shown in Santilano et al. (2018) for 1D MT. The solution with the lowest fitness value was then selected as the final optimized model.

Table 2. Sensitivity analysis on the population size as PSO input argument.

Number of particles	Times the unknowns	rms	Runtime (h)	Iterations
5700	6	0.88	3.47	166
7500	8	0.90	7.17	275
8600	9	0.86	4.60	154
9500	10	0.88	5.82	176
11,500	12	0.87	6.52	165

Note: The number of particles was 6, 8, 9, 10, and 12 times the number of unknowns of the problem (957 grid cells). Results are analyzed in terms of rms errors, total runtime in hours, and the maximum number of iterations reached.

Table 3. Results of PSO applied to the two synthetic models (with and without a priori initialization) and to the COPROD2 data set (without a priori initialization).

Data set	Initialization	rms	Runtime (h)	Iterations
Synthetic model 1	No a priori	0.86	4.6	154
	A priori	0.91	3.17	250
Synthetic model 2	No a priori	0.9	28.8	1674
	A priori	0.99	0.55	53
COPROD2	No a priori	2.42	8	6000

Note: Results are presented in terms of rms error, runtime, and number of iterations performed before the optimization stop. The runtime is in hours and refers approximately to one single trial.

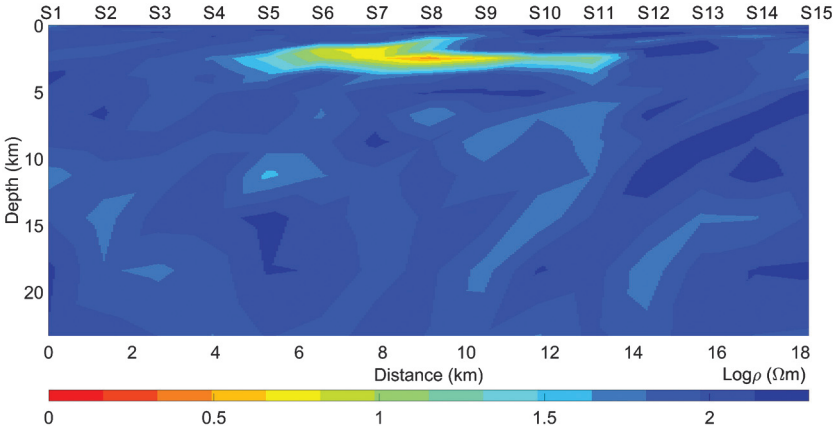

Figure 7. The PSO solution for synthetic model 1, after approximately 150 iterations without a priori initialization for the 8600 particles of the swarm. Lagrange multiplier $\lambda = 0.1$.

Computational aspect

Since the optimization process implied the computation of several model responses, the reliability of the solution was also related to the accuracy of the forward modeling. We adopted the 2D MT forward modeling described in Candansayar (2008). It is based on the finite-difference technique, which solves the complex system of MT equations for transverse-electric (TE) and transverse-magnetic (TM) polarizations. First, the electric and magnetic fields are derived for each mesh node and, finally, the apparent resistivity and impedance phase are calculated. We adopted this forward-modeling code because it is stable, published, and written in MATLAB.

Addressing the 2D problem made the overall computation time consuming due to several factors. The runtime was affected not only by the number of iterations, but also by the population size and the number of unknowns. The number of iterations depends on the complexity of the problem. The population size was related to the number of unknowns, i.e., to the desired resolution of the 2D model. Obviously, an unnecessarily dense mesh grid would have made the computation excessively long. From taking into account all these issues, a heavy computation effort was to be managed. To speed up the computation, we developed and applied the parallel computing option for the PSO algorithm. First, we enabled the option "*UseParallel*," that was potentially provided but not implemented in the standard code. Then, the most overloaded "*for*" loops were set to run as parallel *for* loops, such as, for example, the loop that evaluates the fitness function for each particle. Finally, the PSO algorithm was enabled to run in parallel on the academic cluster by activating the "*Parallel Computing Toolbox*" of MATLAB. All of the simulations were executed on a 24-core node of a HPC cluster for academic research. The CPU model of the single node is an Intel Xeon E5-2680 v3 2.50 GHz (turbo 3.3 GHz) with 128 GB of RAM. The sustained performance of the cluster is 9.7 TFLOPS.

2D OPTIMIZATION OF MT SYNTHETIC DATA

The theoretical MT data sets were computed from two synthetic models depicted in Figures 2 and 3. They covered a 2D domain 350 km long and 250 km deep, to take proper boundary conditions into account for the MT forward modeling (Simpson and Bahr, 2005). Fifteen MT stations were centrally placed in the mesh and reciprocally spaced 1.3 km. The mesh discretization used for the generation of the synthetic data was different from that used for the optimization. Specifically, the latter was slightly coarser than the former due to the computational load given by the thousands of forward-modeling calculations in the PSO algorithm. The mesh size along the horizontal direction has been kept constant between the stations and doubled from the outer stations toward the boundaries. Along the vertical direction, the layer size increased logarithmically with depth. In the case of a priori given, the optimization ran on a subdomain of approximately 400 cells because the a priori information regarded only 15 stations. Without the initial conditioning, the domain extended far away from stations due to the boundary conditions and the number of cells increased up to approximately 900 for synthetic model 1 and approximately 750 for synthetic model 2.

Both synthetic models simulated the presence of one or more electrically conductive features embedded in a resistive body beneath the station sites. In case of different configurations, e.g., prior knowledge of a conductor outside the location of the sites or the ocean nearby, the mesh discretization would have been denser outside the station locations and the mesh boundaries would have been enlarged. Moreover, the presence of the ocean could have been addressed in PSO using a priori information, as previously explained about the initial position of the particles. The first synthetic model is presented in Figure 2. Figure 2a shows the entire 2D mesh, discretized into 33 layers for a total of 957 grid cells. The synthetic model 1 is quite simple and composed of a host medium of 100 ohm-m including a conductive body of 10 ohm-m from 3 to 5 km depth. Figure 2b shows synthetic model 1 as a magnification in the center of the whole mesh.

Figure 3 illustrates synthetic model 2 as a subsection of the true mesh, too. The mesh is discretized into 754 grid cells, and, even so, the model space is not undersampled. The 100 ohm-m resistive medium hosts, from the bottom up, a 10 ohm-m body 4–9 km deep on the left side of the mesh, a 10 ohm-m body 1–2.5 km deep on the right side of the mesh, and a 50 ohm-m body 0–500 m deep under stations S2–S5. The forward modeling that created the synthetic data considered 26 frequency values between 10^{-2} and 10^{3} Hz. Synthetic data were corrupted with uncorrelated Gaussian noise of 10%. This noise corresponded to $\mathbf{\Delta\rho_{a,o}}$ and $\mathbf{\Delta\Phi_a}$ matrices in equation 7, which are the normalization terms of the data misfit. At each kth iteration, the noise influenced the forward calculated response of

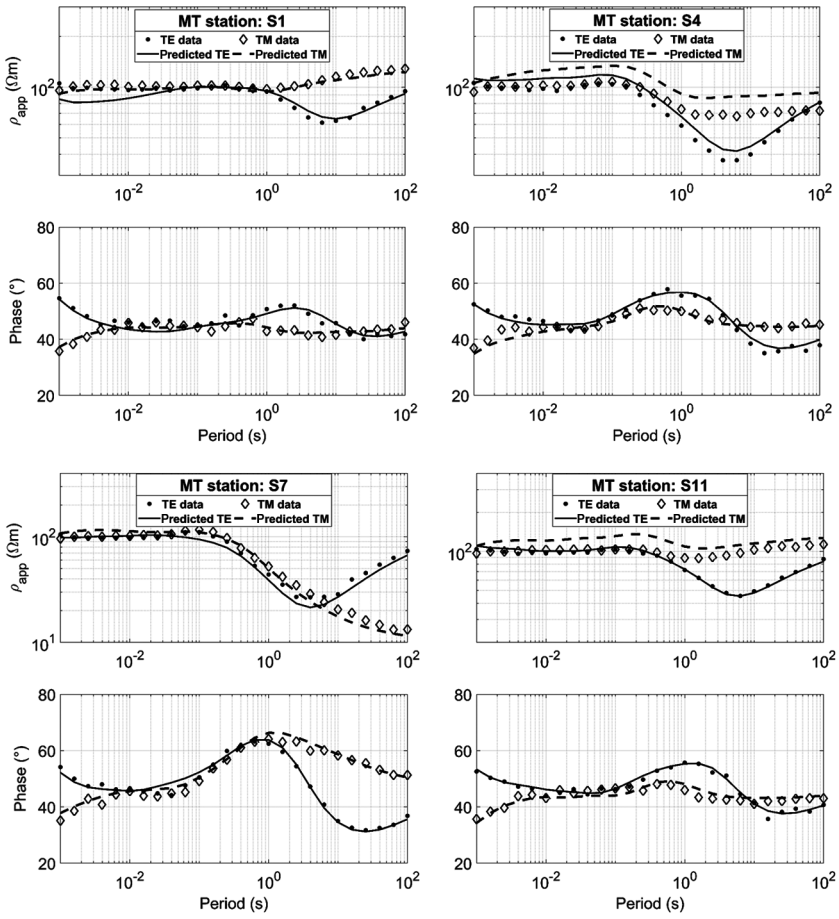

Figure 8. Fitting curves between data of synthetic model 1 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for TE and TM polarizations. The selected MT stations are S1, S4, S7, and S11. The synthetic data are marked as dots for TE and diamonds for TM, whereas the PSO-predicted data are plotted as solid lines for TE and dashed lines for TM. The optimization was randomly initialized.

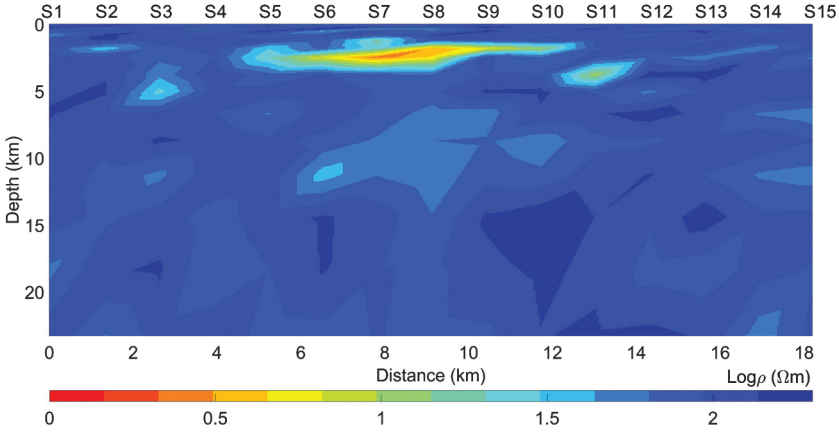

Figure 9. The PSO solution for synthetic model 1 using a swarm size of only 5700 particles (six times the unknowns), after approximately 160 iterations and without a priori initialization. Lagrange multiplier $\lambda = 0.1$.

the corresponding kth model $\mathbf{m}$. The optimization process was constrained by upper and lower resistivity boundaries equal to 200 and 1 ohm-m, respectively.

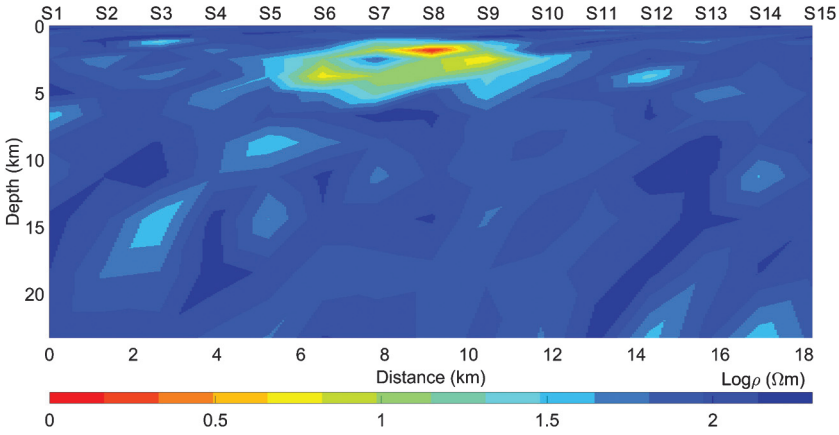

Figure 10. The PSO solution for synthetic model 1, after approximately 250 iterations and with a priori information given to 5% of the particles. Lagrange multiplier $\lambda = 0.1$.

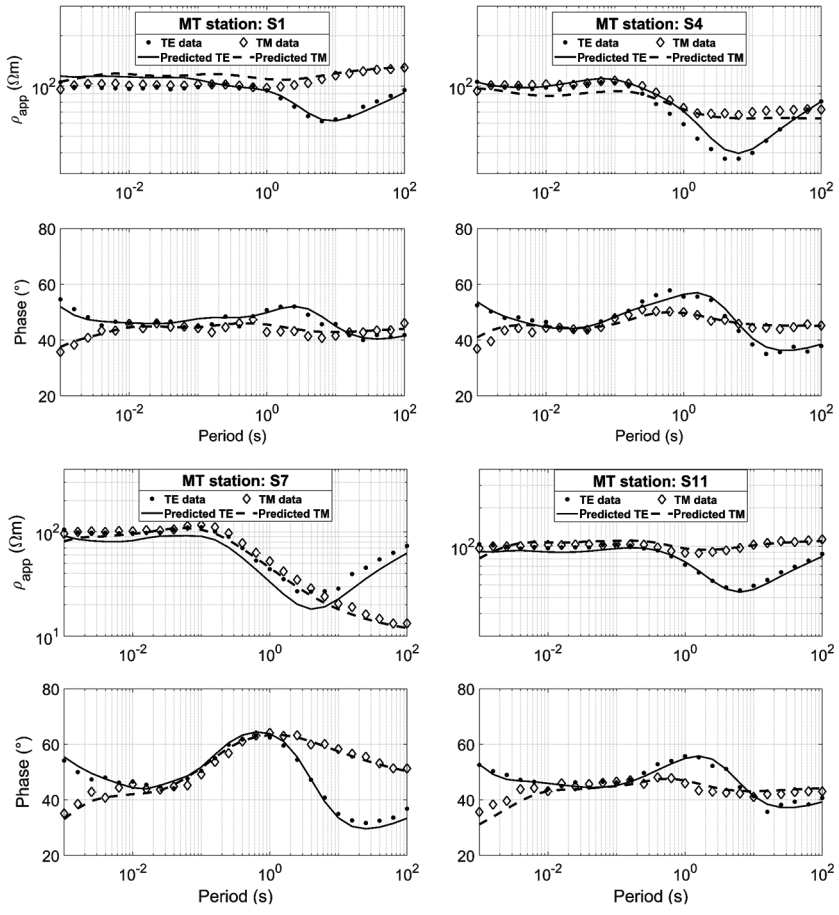

Figure 11. Fitting curves between data of synthetic model 1 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for TE and TM polarizations. The selected MT stations are S1, S4, S7, and S11. The synthetic data are marked as dots for TE and diamonds for TM, whereas the PSO-predicted data are plotted as solid lines for TE and dashed lines for TM. The optimization was initialized with a priori information.

This section is divided into two parts. In the first subsection, the Lagrange multiplier of synthetic model 1 is identified and, then, synthetic model 1 is adopted as a study case to calibrate two input arguments of PSO, the accelerations and the population size; in the second subsection, the final resistivity models are presented for synthetic models 1 and 2.

Calibration of the PSO input arguments

The sensitivity analysis on the Lagrange multiplier was carried out on synthetic model 1 using benchmark values for the accelerations and the population size. These values were chosen as benchmarks for the best convergence of the solution after Ratnaweera et al. (2004). The calibration of the accelerations and population size is presented in the next paragraph because it has significance if the most appropriate Lagrange multiplier is adopted. As a benchmark, the cognitive acceleration α_1 linearly decreased from $\alpha_1^{\max} = 2$ to $\alpha_1^{\min} = 0.5$ and the social acceleration α_2 linearly increased from $\alpha_2^{\min} = 0.5$ to $\alpha_2^{\max} = 2$. The benchmark population size was approximately nine times the number of unknowns, i.e., given 957 cells, 8600 particles. To retrieve the optimal value of the Lagrange multiplier λ, we performed a sensitivity analysis on five different values in the range between 0.001 and 10. λ_x and λ_z were contextually analyzed with the same value and the optimal value was chosen as the point of maximum curvature in the plot of data misfit versus model norm. Figure 4 shows the data misfit of synthetic model 1 with respect to the model roughness along the horizontal (black diamonds) and vertical (red circles) directions. The best trade-off value was equal to 0.1 for λ_x and λ_z.

The sensitivity analysis on the cognitive and social accelerations was carried out for synthetic model 1 once its optimal value of the Lagrange multiplier was identified. For this calibration, the population size was fixed to the aforementioned benchmark value of 8600 particles and its sensitivity analysis is shown in the next paragraph. We chose three different values for the maximum cognitive acceleration, $\alpha_1^{\max} = 2.75, 2, 1.5$, and three different values for the social acceleration, $\alpha_2^{\min} = 0.25,\ 0.5,\ 0.75$. These values were selected on the basis of the existing literature and equations 5 and 6. The solution reliability was evaluated via some parameters of the optimization process, such as the first stopping criterion achieved, the solution clustering, and the trend of the fitness function at each iteration. The simulations ran until one of the three stopping criteria was first fulfilled, i.e., when the fitness function did not significantly decrease and (almost) all of the particles converged to a unique position in the search space or solutions. Table 1 lists the rms errors and the fitness-function values ($F(\mathbf{m})$) at

the end of the optimization of each test. Our results are largely consistent with the acceleration values pointed out in Ratnaweera et al. (2004) for other applications. The tests using $\alpha_1^{k=1} = 2.75$ with $\alpha_2^{k=1} = 0.5$ and with $\alpha_2^{k=1} = 0.75$ ended before that the rms was equal to 1 because the fitness function did not decrease for 80 consecutive iterations. These values prevented an effective minimization, as shown in Figure 5, which summarizes the optimization performance using $\alpha_1^{k=1} = 2.75$ and $\alpha_2^{k=1} = 0.5$. The four subplots show, in order: Figure 5a — the fitness-function values of the best particle (red dots) and the mean values of the rest of the swarm (black dots) from the first to the final iteration; Figure 5b — the fitness-function values of the whole swarm as a function of the particle positions in two representative dimensions of the search space, i.e., the first two cells of the 2D grid, at the first (gray dots) and final (blue dots) iterations; Figure 5c — the positions, i.e., the resistivity values, of the particles in the first two cells of the 2D grid at the first and final iterations (gray dots and blue dots, respectively); and Figure 5d — the histogram containing the distribution of the fitness-function values at the last iteration among all of the particles (8600 in this case). Figure 5 reveals that the optimization did not end in a convergence state because at the last iteration the minimum $F(\mathbf{m})$ was not reached by the totality of the particles (Figure 5a and 5d) and the distribution of the particles in the search space was still scattered (the blue dots in Figure 5b and 5c). The other tests in Table 1 show an optimal convergence, rms errors equal to approximately 1, and the minimized $F(\mathbf{m})$ between 1.33 and 1.73. Figure 6 plots the optimization performance using $\alpha_1^{k=1} = 2$ and $\alpha_2^{k=1} = 0.5$. The minimization of the fitness function was more effective than that of Figure 5 because all of the particles converged toward a unique position (the blue dots in Figure 6b and 6c) with the same fitness-function value corresponding to the peak in Figure 6d. It is evident that, iteration by iteration, particles converged from an initial scattered distribution to a unique position following the best particle leadership. In this way, the fitness-function value dropped and the histogram developed a unique peak. This sensitivity analysis confirmed $\alpha_1^{\max} = 2$, $\alpha_1^{\min} = 0.5$, $\alpha_2^{\min} = 0.5$, and $\alpha_2^{\max} = 2$ as optimal acceleration values for a robust minimization of the fitness function. These accelerations were applied to the optimization of the other MT data sets of this work.

The synthetic model 1 was also a study case for the sensitivity analysis on the population size, to assess the influence of the number of particles on the solution and the runtime. The tests were performed using five different values, chosen as multiples of the number of unknowns (957): 5700, 7500, 8600, 9500, and 11,500 particles, i.e., 6, 8, 9, 10, and 12 times the unknowns. The accelerations and Lagrange multiplier were set as explained before for the corresponding sensitivity analyses. The results are shown in Table 2. All of the tests reached the minimum rms of approximately 0.9, but with different numbers of iterations and runtimes because of the different initial distributions of the particles in the search space of solutions. The test using the multiple 8 gave the worst result because the solution was found after the biggest runtime and the highest number of iterations. Differently, the multiple 9 gave the best result, with the minimum number of iterations and the second shortest runtime. The ratio of 9, i.e., 8600 particles for synthetic example 1, ensured the most effective convergence and exploration of the solution space, so that it was adopted for the other tests presented in this work.

Results from two synthetic examples

In this subsection, we present the results of PSO applied to the two synthetic examples depicted in Figures 2 and 3. The optimization of MT data from the synthetic models was performed adopting the optimal values for the Lagrange multipliers, accelerations, and population size reported in the previous subsection. The results regarding synthetic model 1 are presented in this order: the model obtained without external conditioning of the PSO initialization, the model resulting from a poorly populated swarm, and, finally, the result after the PSO initialization with a priori information given

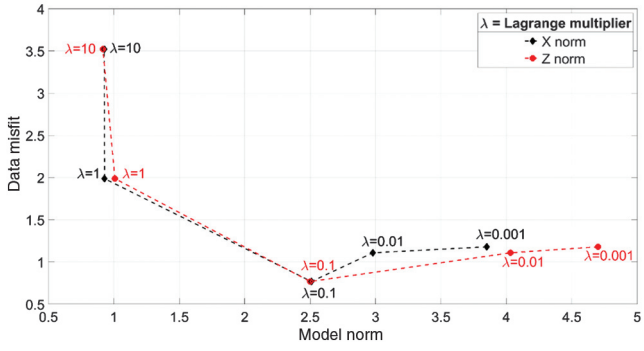

Figure 12. The L-curve response for synthetic model 2 along the horizontal (black diamonds) and vertical (red circles) directions. The trade-off between the data misfit and the model norm indicates the best Lagrange multiplier λ equal to 0.1.

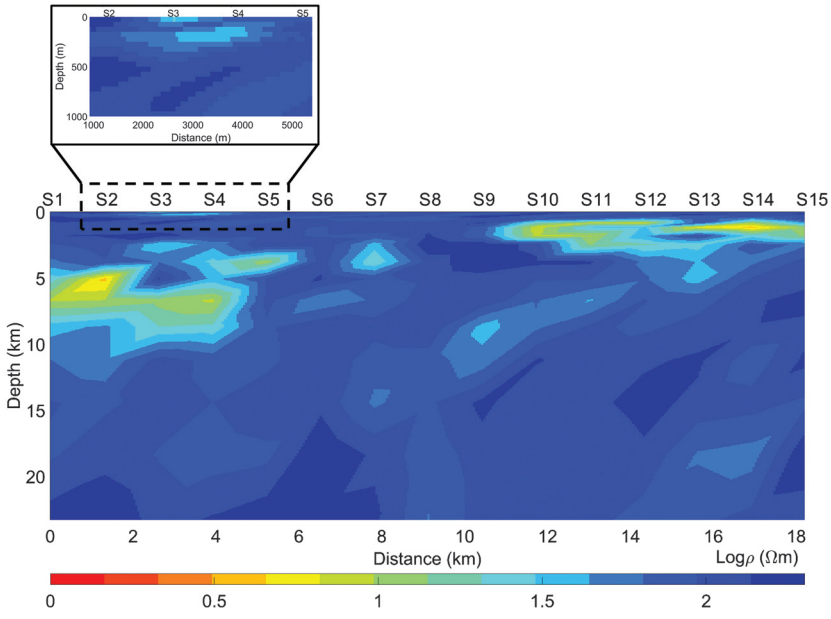

Figure 13. The PSO solution for synthetic model 2, after 1674 iterations and without a priori initialization. Lagrange multiplier $\lambda = 0.1$.

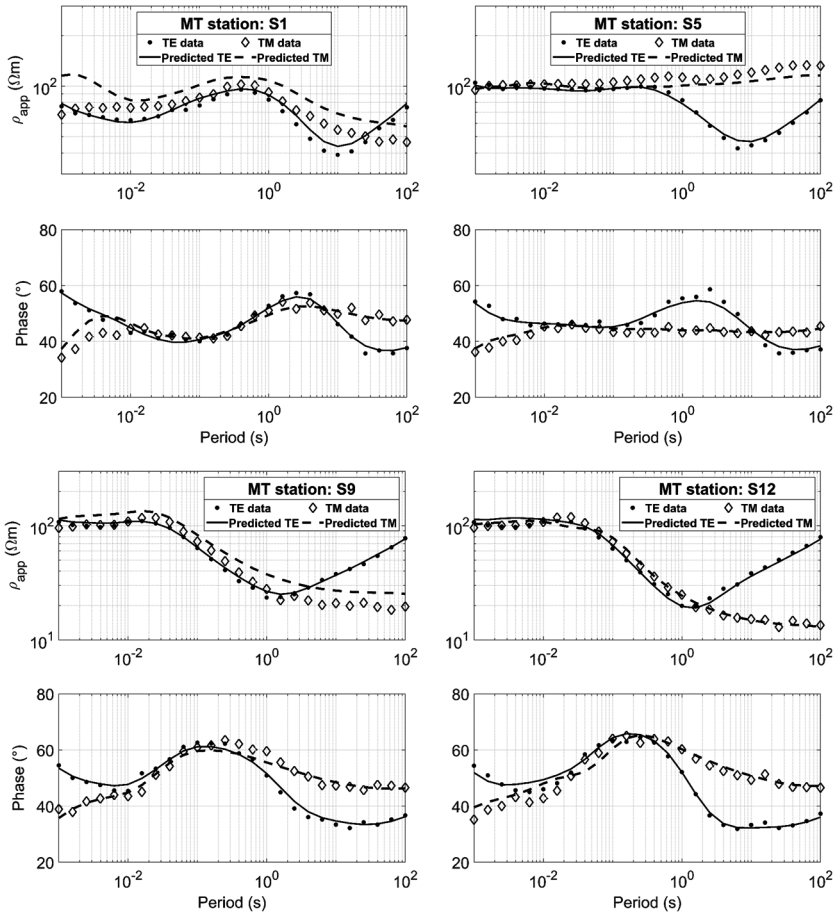

Figure 14. Fitting curves between data of synthetic model 2 and calculated data for the apparent resistivity (ρ_{app}) and impedance phase for TE and TM polarizations. The selected MT stations are S1, S5, S9, and S12. The synthetic data are marked as dots for TE and diamonds for TM, whereas the PSO-predicted data are plotted as solid lines for TE and dashed lines for TM. The optimization was randomly initialized.

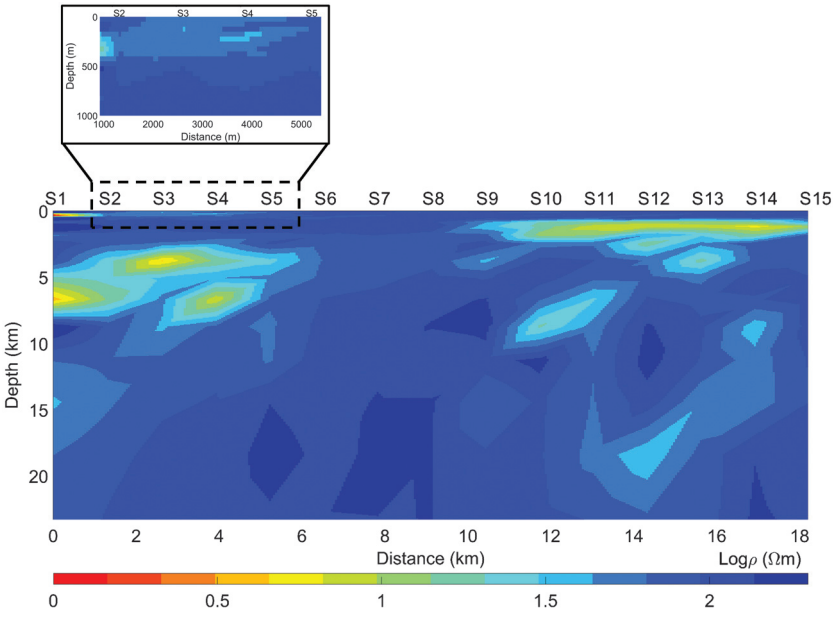

Figure 15. The PSO solution for synthetic model 2, after 53 iterations and with a priori information given to 5% of the particles. Lagrange multiplier $\lambda = 0.1$.

as the starting model. The resistivity model obtained without a priori information is shown in Figure 7. After approximately 150 iterations and 4.6 h, the rms error stabilized around the final value of 0.86, whereas $F(\mathbf{m})$ was 1.37. These values are listed in Table 3, but they can also be found in Table 1 for $\alpha_1^{k=1} = 2$ and $\alpha_2^{k=1} = 0.5$, and in Table 2 for 8600 particles. The resistivity model in Figure 7 was largely comparable with the true model in Figure 2 because the conductive anomaly was correctly detected in size and resistivity. Figure 8 plots the fitting curves between the synthetic and calculated data for apparent resistivity (ρ_{app}) and phase, both TE and TM polarizations. The synthetic data are marked as dots for TE and diamonds for TM, whereas the PSO-predicted data are plotted as solid lines for TE and dashed lines for TM. Four stations were selected for their different positions in relation to the lateral discontinuities: S1, S4, S7, and S11. They show an example of poor (S4 ρ_{app}), average (S11), and good fit (S1 and S7). Considering the high number of unknowns and the wide range of variation of ρ_{app}, it could be said that these curves are clearly similar to each other as also proved by the low rms error.

The influence of the population size on the optimization process is presented in Table 2. Figure 9 shows the effect of a poorly populated swarm on the final resistivity model. This result followed from a population size of 5700 particles, i.e., six times the unknowns. The result was similar to the true model in Figure 2 because the conductive body was identified. However, the output was not completely appreciable due to some lateral conductive artifacts that broke in the homogeneous 100 ohm-m background. As expected, this outcome was the consequence of an ineffective initial distribution of the particles in the search space of solutions, and possibly, of the missing of the global minimum.

The a priori information used to initialize the optimization came from the PSO solutions of the 1D inverse problem for the 15 stations of synthetic model 1. Only 5% of the particles were initially influenced with this solution. After 250 iterations, the rms error reached the minimum threshold, with a corresponding fitness-function value of 1.4. The final resistivity model is shown in Figure 10 and is comparable with the original of Figure 2 because the conductive anomaly was adequately identified. Figure 11 plots the fitting curves of the selected stations. The PSO-predicted responses were distinctly consistent with the synthetic data, and the difference with the curves of Figure 8 was negligible, except for the slight improvement for ρ_{app} of S4 and S11. Table 3 lists the details regarding rms, runtime (in hours), and the total number of iterations.

The optimization of synthetic data from synthetic model 2 (the true model in Figure 3) was performed after the calibration of the input arguments, the accelerations, and population size. The identification of the optimal Lagrange multiplier λ for synthetic model 2 was inferred from the L-curve response presented in Figure 12. It refers to the data-misfit trend as a function of the horizontal (black diamonds) and vertical (red circles) roughness for synthetic model 2. The best tradeoff value was 0.1 for λ_x and λ_z.

The best solution for synthetic model 2 without a priori initialization is illustrated in Figure 13. All the three low-resistivity bodies were accurately positioned as can be seen from the magnified panel. After 1674 iterations, the minimum fitness-function value was 1.3 and the RMS error was 0.9. Figure 14 plots the comparison between the synthetic and the calculated data for apparent resistivity (ρ_{app}) and phase (TE and TM). Stations S1, S5, S9, and S12 were selected as being representative for their poor (S1), average (S5 and S9), and good fit (S12). Taking the complexity of this synthetic example into proper account, the curves are in good agreement. The runtime is reported in Table 3.

The a priori information was set as previously explained. Once 5% of the particles were initially influenced, convergence was reached only after 53 iterations with a minimum fitness-function value of 1.5. Figure 15 displays the final output. The 10 ohm-m lateral bodies were correctly imaged, whereas the superficial 50 ohm-m body was scarcely identified. A distinct difference between the models with and without a priori is indeed the superficial body, as can be seen in the magnified panels of Figures 13 and 15. Figure 16 graphically demonstrates the low rms error of 0.99, which is reported in Table 3. The fitting curves of Figure 16 show a good agreement between synthetic and predicted data and no significant improvements compared with Figure 14.

2D OPTIMIZATION OF MT FIELD DATA

The COPROD2 data set

The COPROD2 data set collects long-period MT measurements along a profile of 35 sites crossing a 2D geoelectrical structure in Saskatchewan and Manitoba, Canada (Jones and Savage, 1986). The name stands for "Comparison of One-dimensional PROfiles from MT Data," whereas the "2" refers to the two-dimensionality different from the 1D data set called "COPROD." The most appreciable advantages of this data set are the following: A wide period bandwidth (from 2.6×10^{-3} s to 1.8×10^{3} s), low impedance errors (<2%), previous static shift correction, and the possibility of comparing different models from well-established inversion algorithms (Jones, 1993b). Our aim is to apply the PSO algorithm to detect deep electromagnetic anomalies, whereas any geologic interpretation is beyond the scope of this paper.

Since responses at low periods (below 10 s) have been widely recognized as 1D, original data were selected from 10.67 to 910.2 s (deGroot-Hedlin and Constable, 1993). As proposed in the aforementioned studies, a subset of 20 MT stations, from the 8th to the 27th of the original line, was chosen to focus only on the center of the 400 km east–west profile. This selection was also adopted because these 20 sites have the same number of acquisition frequencies (14) within the considered interval. The errors on the data were kept as original for TE and TM apparent resistivity and phase. Figure 17 plots MT observations and error bars for stations 12, 13, and 14, chosen as representative of the subset. The maximum observed error is 1.2 for TE apparent resistivity (on logarithmic scale) and 6.92° for TM phase.

The 2D model was divided into 10 layers, from 1.8 to 60.5 km deep, and the thickness of each layer increased logarithmically with depth. Along the horizontal direction, the mesh was approximately 200 km long and subdivided into 34 bricks, one for each station plus others as boundary conditions. The total number of cells was 340. Since some structures of the region are known to be highly conductive, the lower boundary of the problem was set equal to 0.1 ohm-m. Literature references also state that superficial sediments are far more conductive than the resistive basement. For this reason, the upper boundary of resistivity was chosen to be different between the upper and underlying layers. We observed that a search

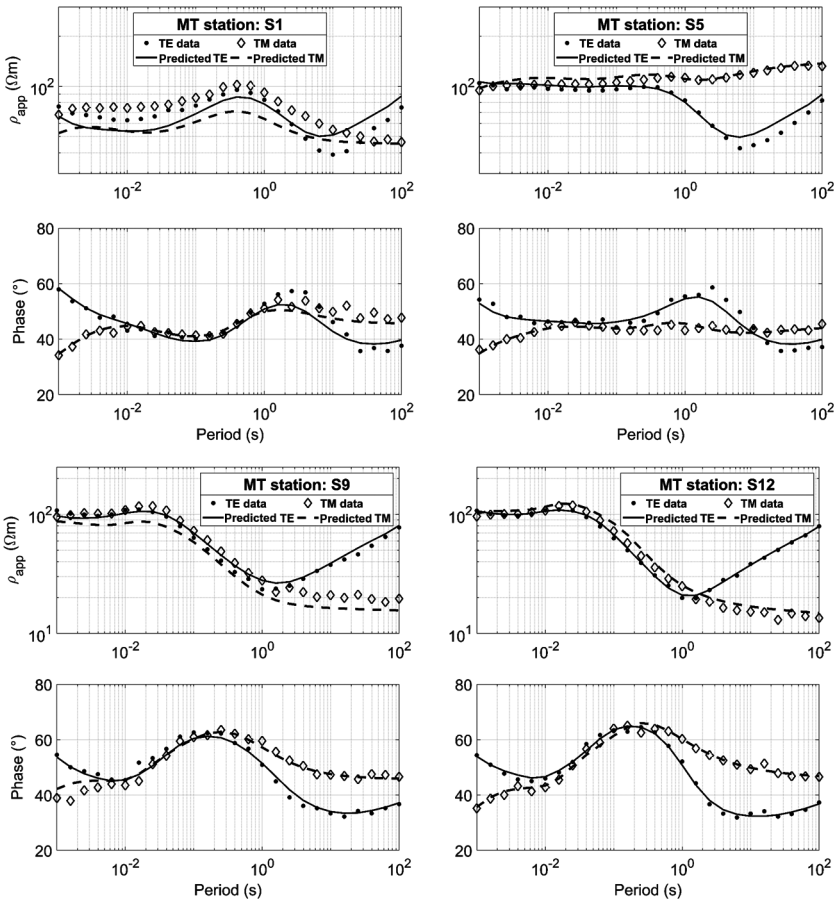

Figure 16. Fitting curves between data of synthetic model 2 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for the TE and TM polarizations. The selected MT stations are S1, S5, S9, and S12. The synthetic data are marked as dots for TE and diamonds for TM, whereas the PSO-predicted data are plotted as solid lines for TE and dashed lines for TM. The optimization was initialized with a priori information.

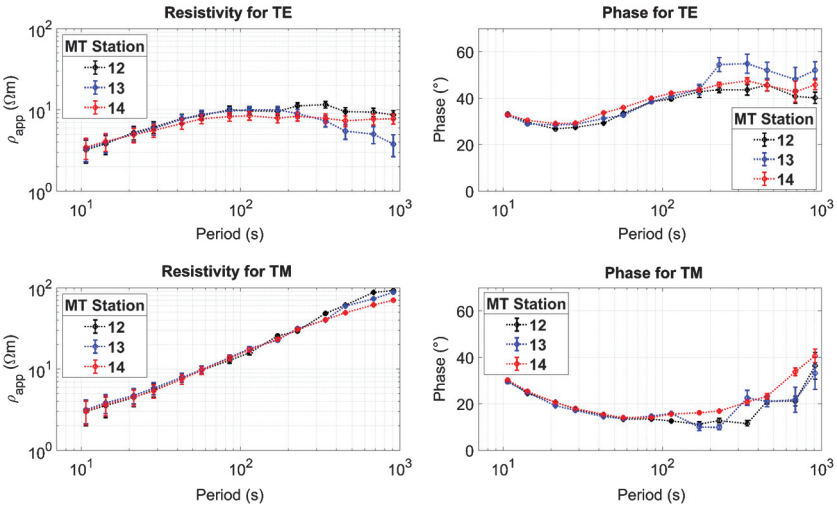

Figure 17. The MT responses and error bars for TE and TM modes of three representative stations (12, 13, and 14) of the COPROD2 data set. They show the high quality of the data. The ρ_{app} stands for the apparent resistivity.

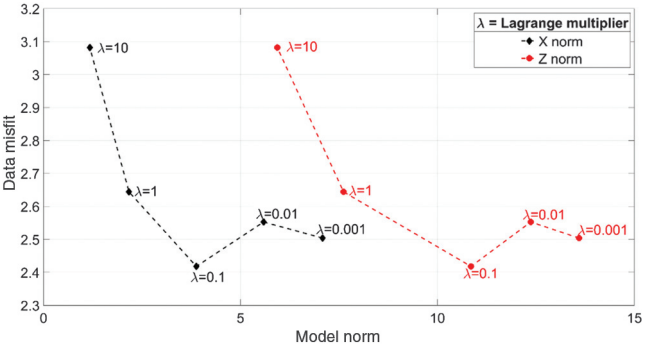

Figure 18. The L-curve response for COPROD2 data along the horizontal (black diamonds) and vertical (red circles) directions. The trade-off between data misfit and model norm indicates the best Lagrange multiplier λ equal to 0.1.

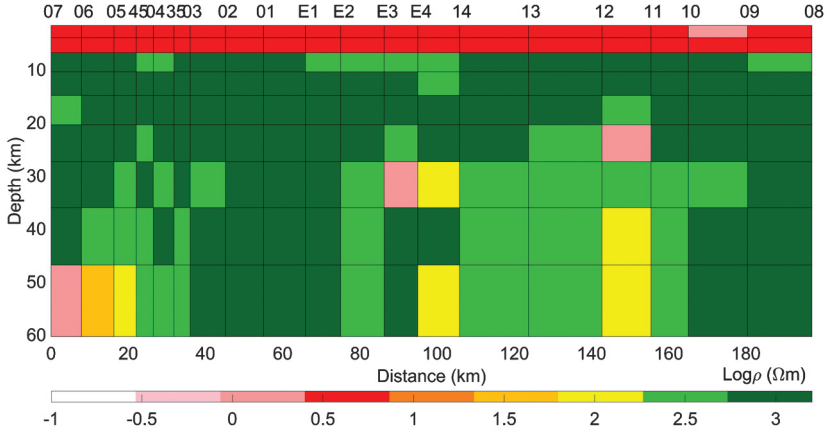

Figure 19. Resistivity model of COPROD2 data from PSO computation, after 6000 iterations. Lagrange multiplier $\lambda = 0.1$.

space too large for the upper layers would have driven the solution toward no convergence and erroneous local minima. In detail, the first two superficial layers, namely, up to 5 km deep, had 10 ohm-m as an upper boundary, whereas, the layers below had 1000 ohm-m. The population size was equal to 2500 particles, proportional to the number of cells. The Lagrange multiplier λ was chosen after a sensitivity analysis on five different values in the range between 0.001 and 10. The value that coincides with the point of maximum curvature in the plot of data misfit versus model norm is 0.1, as shown in Figure 18. A priori information was not given: The optimization started with a completely random initialization.

The final model from COPROD2 data was computed after 6000 iterations and is depicted in Figure 19. The shallow conductive structure was extensively identified, whereas, at a depth from 5 km to the bottom, the background resistivity was predominantly 1000 ohm-m. The most significant feature of this model was represented by the low-resistivity anomalies below stations E3-E4 and 12-11 at approximately 20–35 km of depth. Our output is well comparable with the ones represented in Figure 20 and are reported by Jones (1993b). For ease of comparison, the color scale and the name of the stations of Figure 19 were plotted as the original ones in Figure 20. There was good agreement between our model and those called "degroot-2" (deGroot-Hedlin and Constable, 1993), "rasmussen" (Rasmussen, 1993), "wu" (Wu et al., 1993), and "uchida" (Uchida, 1993): Low-resistivity anomalies were identified in the same regions. The similarity can be explained by the same approach adopted in the Occam's inversion using the smoothing parameter, except for "wu", which used a different approach. In detail, the most evident similarity was the conductive region in the first 5–7 km of depth. Another similarity regarded the 30 km deep conductor below the stations from E2 to 14 and its extension at greater depths. A further correspondence was the interruption of the 1000 ohm-m structure below the stations from 13 to 11 at approximately 20 km of depth. Apart from the "wu" model, all of the analyzed results presented a low-resistivity region (approximately 100 ohm-m) in the westernmost part of the model at a depth greater than 30 km.

Figure 21 plots the apparent resistivity (ρ_{app}) and phase at selected periods for the 20 stations in the horizontal axis. The observed data are marked with dots for TE and diamonds for TM, and the predicted responses are plotted with a solid line for TE and a dashed line for TM. The rms error is reported in Table 3. The final fitness-function value was 26.6.

Computation time

The computationally demanding nature of the PSO algorithm was actually expected due to the high number of (1) iterations, (2) population size, and (3) cells assembling the mesh. The standard release of the code was not effective in addressing the 2D inverse problem; therefore, we applied some modifications to develop a parallelized

version of PSO. The tests on an HPC cluster proved that, when 24 cores were adopted, the runtime saving was more than 80% with respect to the use of four cores (see Figure 22). A test using the nonparallelized release of the code (one single worker) would have been unfeasible in terms of machine working load. Figure 22 shows the dramatic speedup of PSO computation for a reference simulation of 150 iterations and 10,000 particles: The black lines indicate the running duration in hours, whereas the blue indicate lines the total speedup in percentage. The parallel environment "shared" (the dotted lines) exploited workers of the same node, whereas "orte" (the dashed lines) referred to workers from different machines of the cluster. It could be seen that "shared" was a bit faster than "orte", especially at high numbers of cores.

The total runtime of PSO computations is reported in Table 3 for each study case. These values refer to one single trial, whereas a total of three trials were performed. Runtimes are not directly comparable to each other because the stopping criterion was met after different numbers of iterations, i.e., less than 1600 iterations for the synthetic models and 6000 for the real data set. The optimization of synthetic examples stopped because the minimum rms error was achieved, whereas the optimization of the real data stopped because the fitness function did not minimize for 80 iterations. The synthetic examples were optimized in a few iterations, but the runtime was longer than that of the COPROD2 data set due to the higher number of layers and particles. In fact, the different runtimes are explained by the low number of layers the COPROD2 model is discretized (i.e., about one-third of that of the synthetic models). The optimization of synthetic data 1 and 2 without a priori differed in the number of iterations needed, but the runtimes are quite similar if the proportion between the iterations is taken into account. The optimization of synthetic data 1 and 2 had shorter runtimes when a priori initialization was given.

DISCUSSION

Our tests on synthetic data demonstrate the reliability of PSO in solving the 2D inverse problem for MT data sets.

The choice of the most appropriate values of accelerations and population size was crucial for obtaining valid models. The initial sensitivity analysis on the PSO input arguments was essential to identify the most appropriate tuning coefficients, which effectively minimized the objective function and enhanced the solution convergence. The calibration of the social and cognitive accelerations led to the optimal values of $\alpha_1^{max} = 2$, $\alpha_1^{min} = 0.5$, $\alpha_2^{min} = 0.5$, and $\alpha_2^{max} = 2$. Our findings are hence in agreement with Ratnaweera et al. (2004). We demonstrated that the population size was directly proportional to the total runtime, but at the same time, a poorly populated swarm negatively influenced the model. The reason for this was that the search space was ineffectively covered by the initial random distribution of the particles. We showed that the best ratio between the number of unknowns and the number of particles

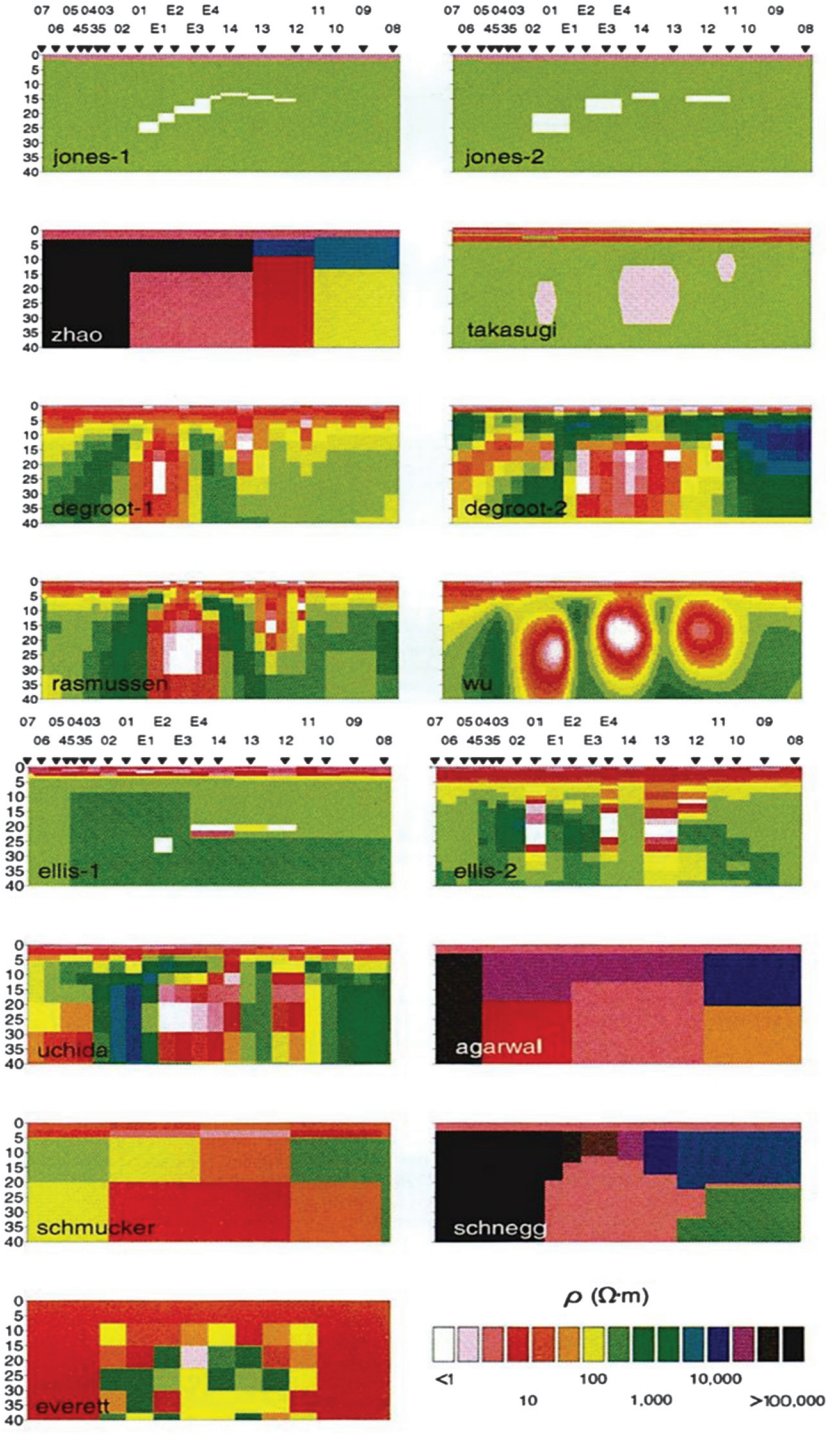

Figure 20. Reference models of COPROD2 data from Jones (1993b). The 20 stations are sorted and named as in Figure 19. The color scale for the resistivity (ρ) is consistent with Figure 19: white ($\rho < 1$ ohm-m), pink ($\rho = 1$ ohm-m), red ($\rho = 10$ ohm-m), yellow ($\rho = 100$ ohm-m), and green ($\rho = 1000$ ohm-m).

was 9. This outcome is significant for our high-dimensional problem because, so far, the literature has suggested increasing the number of particles up to 12 times the number of unknowns. The conclusion of our analysis slightly modifies this ratio, with the advantage of avoiding extra computational load.

We showed that the application of PSO did not require an initial assumption about the solution (i.e., a priori information). At the same time, we introduced a novel and valid tool to potentially communicate external or additional information to the swarm, in terms of the initial position of particles within the search space. Our findings showed that, if the geologic or geophysical information is reliable (e.g., from wells, seismic reflectors, and so on), it influences the behavior of particles at the beginning of the optimization. This kind of initialization resulted in shorter runtimes because the swarm did not waste time searching for local minima, which were already given from the beginning (see Table 3). On the other hand, using default random initialization, the results of synthetic models proved that there was no requirement for a priori initialization because the final resistivity model was perfectly comparable with the original synthetic model. Moreover, there was high solution quality, despite two factors: Gaussian noise disturbing the data and the presence of equivalent solutions in the MT inverse problem. In fact, the conductive anomalies embedded in the resistive host medium were accurately identified in terms of size and resistivity values. The rms errors were around 1 and, interestingly quite similar with and without a priori initialization. Another element confirming the PSO independence from the starting model is the comparison of the fitting curves in Figures 8 and 11 for synthetic model 1, and Figures 14 and 16 for synthetic model 2. These plots proved that there were not substantial differences between calculated responses with and without a priori, even considering the stations above the lateral discontinuities (S4 and S11 for synthetic model 1 and S5 and S9 for synthetic model 2).

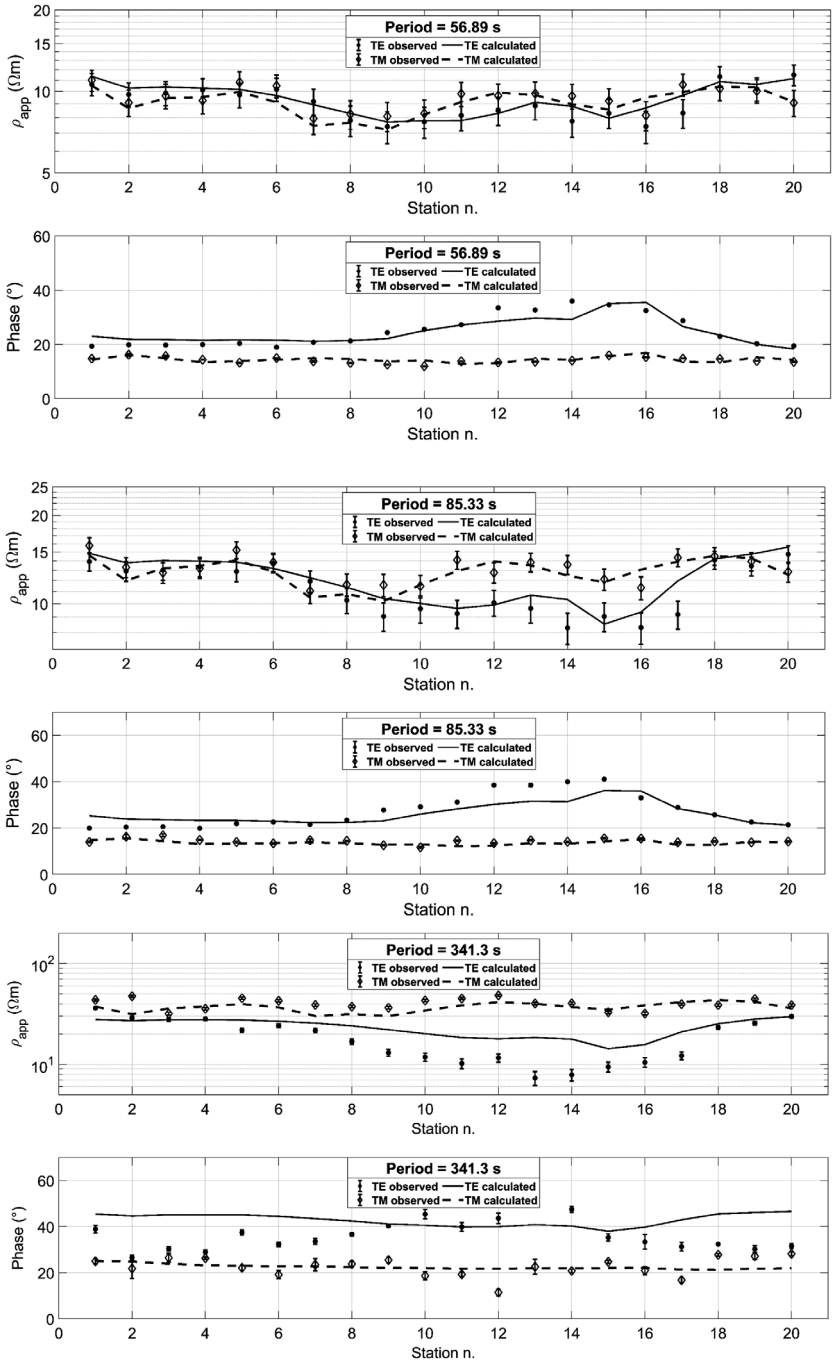

Figure 21. Fitting curves between observed apparent resistivity (ρ_{app}) and phase, and predicted responses at selected periods: 56.9, 85.3, and 341.3 s. Observed data include error bars and are marked with dots for TE and diamonds for TM. Calculated responses are plotted with a solid line for TE and a dashed line for TM. The optimization was randomly initialized.

A significant result arose from the application of PSO to real 2D data, the COPROD2 data set. As regards the problem settings, the uppermost layers of the 2D mesh had different boundary conditions with respect to the underlying layers, due to the complexity of the investigated area. Many applications of global search algorithms to geophysics have considered different resistivity boundaries between one layer (or group of layers) and another, so that each unknown of the problem can independently be bounded within its search space (Godio and Santilano, 2018). The setting of the boundary conditions is not a trivial step for the deterministic inversion either because it implies full comprehension of the problem and some insight into the possible solution. For the optimization of the COPROD2 data set, preliminary information from the geology of the area facilitated the definition of a wide interval of resistivity values within which the solution could be searched. We applied this approach to enhance the convergence of the solution. This option distinguished the solution space of upper layers from that of deep layers. The final model had a mean resistivity of 6 ohm-m in the two upper layers, thus confirming the presence of

superficial sediments. At a depth of approximately 25 km, a conductive region breaks the 1000 ohm-m background, with a minimum value of 1.2 ohm-m. The final rms error was slightly bigger than that of the synthetic examples and was negatively affected by the mismatch of data at long periods (Figure 21). This was unexpected, given the satisfactory behavior of the fitting curves of the synthetic examples in Figures 8 and 14. It would have been interesting to quantitatively compare our result with the model obtained by Everett and Schultz (1993) using GA, which is also a global search algorithm. Unfortunately, their rms misfit of 1.48 is not directly comparable with our value of 2.42 because there were substantial differences in our methods, such as period range, number of stations, mesh discretization, and stopping criteria for iterations. Interestingly, the adoption of Occam-like optimization may provide a more effective solution of the resistivity distribution with respect to the GA. It has also been proved in the literature that PSO ensures a higher convergence with respect to the other global search algorithms (Yuan et al., 2009; Fernández Martínez et al., 2010a). The application to field data represents a new, encouraging approach for their optimization by means of computational swarm intelligence.

Regarding the optimizations run without a priori initialization, the runtimes of the different data sets were not straightforwardly comparable due to the different number of iterations required to achieve convergence. The COPROD2 optimization needed 6000 iterations to stop at rms = 2.42, whereas the synthetic examples reached rms = 1 in fewer iterations but taking a runtime proportionally longer than that for the real-data optimization. This is mainly explained by the high level of mesh discretization for the synthetic models, approximately 800 cells, compared with approximately 340 cells for the domain chosen for real-data interpretation. The more unknowns, the greater the swarm size and hence the computation time. The computation speedup was obtained by introducing the parallelization of the code. The tests performed on an HPC cluster pointed out the capability of our version of PSO to speed up the computation by more than four times with respect to running it on a simple machine of four cores. The decrease of the runtime allowed us to efficiently perform several trials of the optimization process, starting from different random distributions of the swarm. The final fitness-function value of the synthetic examples was lower than that of the real data due to the peculiarity of the data sets. The choice of the optimal Lagrangian multiplier may be seen as a computational cost because the sensitivity analysis of different values of λ was performed. This analysis could represent a slight limitation of the presented method because PSO ran for each investigated value of λ. However, once the balance was found, we were able to deploy the model with the adequate level of smoothing.

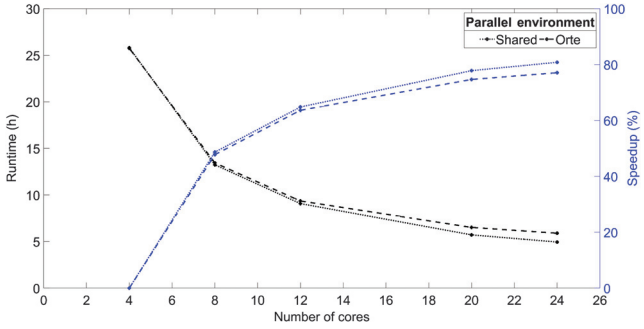

Figure 22. Black curves show the computation times in hours (left ordinate axis) as a function of the number of cores exploited for a reference PSO simulation of 150 iterations with a 10,000-particle swarm. The right ordinate axis and the blue curves refer to the total runtime speedup with exploited cores increasing. The dotted lines refer to "shared" parallel environment (workers of a single node) and the dashed lines refer to "orte" (workers of different nodes).

Although we reduced the computation time, it remained not comparable with that of deterministic algorithms. PSO applied to the 2D inverse problem is relatively time consuming if clusters cannot be exploited and densely discretized meshes are adopted. However, the parallelization of the code has the potential of making the PSO computation more manageable. Moreover, high computing capacity is nowadays within everyone's reach and global search algorithms, despite the skeptical view of the past, can now be considered worthy of attention. We do not see the long runtime as a scientific barrier for the application of PSO to high-dimensional geophysical problems. The computational load was balanced by the advantages of this metaheuristic method, namely, independence from the choice of the starting model and the solution driven by the evolutionary approach.

CONCLUSION

The particle swarm optimization (PSO) algorithm has proven to be a valid method to solve the two-dimensional (2D) inverse problem for magnetotelluric (MT) data, for synthetic and field (COPROD2) data sets. This work extended the application of PSO to MT inversion from the one-dimensional (1D) problem, already visited in the literature, to the 2D problem. The stochastic nature of PSO and the combination of exploration and exploitation behaviors played a key role in finding the optimized solution within the search space, which was composed of all the possible solutions of resistivity models.

The standard release of the code was easily implemented for our specific application. We observed striking improvements moving from standard PSO to hierarchical PSO with time-varying acceleration coefficients (HPSO-TVAC). This issue has not been addressed in previous research on PSO applied to geophysics, but it was crucial in the optimization of 2D MT data. In fact, thanks to time-varying acceleration coefficients, the optimization ended with true convergence and stability. The complexity of the 2D problem had a direct influence on the computation time, which we reduced with the parallelization of the code. Running PSO on an high performance computing (HPC) cluster resulted in runtime savings of approximately 80%.

We carried out a detailed sensitivity analysis on some input parameters of the PSO algorithm due to their direct influence on the stability and convergence of the solution. The social and cognitive accelerations and the population size were investigated to retrieve their optimal values and analyze their effect on the final resistivity models and total runtime. We first applied PSO to 2D MT synthetic data, to validate the method. The initialization of the optimization was purely random by default, but we also tried to influence it with a starting model derived from PSO solutions of the 1D problem. In this case, the optimization was externally but not totally influenced because only a small portion of the swarm was initialized. We proved that a priori information as the starting model can be avoided. The resistivity models that did not receive the a priori initialization were in line with the original synthetic models. Then, PSO was applied to the field data set COPROD2, with a random initialization. The optimization of COPROD2 data produced

a valid resistivity model, largely comparable with results from existing research.

The most important conclusions of this work are that PSO can be successfully applied to the 2D MT inverse problem and the a priori starting model is not required for the achievement of valid models. Our results are encouraging enough to extend the application of evolutionary algorithms to other geophysical inverse problems, bearing in mind that the high dimensionality of the problem implies runtimes longer than those of local search methods. Future work will investigate a further speedup of the PSO computation, as well as its application to other MT real data sets. A possible direction of future studies should consider the comparison between local and global search methods for the 2D MT inversion, as well as the challenge of the 3D inverse problem.

ACKNOWLEDGMENTS

The authors thank the editors and the four reviewers for their valuable and helpful comments, which resulted in a vastly improved paper. Computational resources were provided by hpc@polito (http://hpc.polito.it).

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

REFERENCES

Adhan, S., and P. Bansal, 2017, Applications and Variants of Particle Swarm Optimization: A Review: International Journal of Electronics, Electrical and Computational System, **6**, no. 6, 215–223.

Avdeev, D. B., 2005, Three-dimensional electromagnetic modelling and inversion from theory to application: Surveys in Geophysics, **26**, 767–799, doi: 10.1007/s10712-005-1836-x.

Boerner, R. U., 2010, Numerical modelling in geo-electromagnetics: Advances and challenges: Surveys in Geophysics, **31**, 225–245, doi: 10.1007/s10712-009-9087-x.

Brodie, R. C., and W. Jiang, 2018, Trans-dimensional Monte Carlo inversion of short period magnetotelluric data for cover thickness estimation: ASEG, Extended Abstracts, 1–7, doi: 10.1071/ASEG2018abT5_1F.

Candansayar, M. E., 2008, Two-dimensional inversion of magnetotelluric data with consecutive use of conjugate gradient and least-squares solution with singular value decomposition algorithms: Geophysical Prospecting, **56**, 141–157, doi: 10.1111/j.1365-2478.2007.00668.x.

Conway, D., J. Simpson, Y. Didana, J. Rugari, and G. Heinson, 2018, Probabilistic magnetotelluric inversion with adaptive regularisation using the No-U-Turns sampler: Pure and Applied Geophysics, **175**, 2881–2894, doi: 10.1007/s00024-018-1870-5.

Darisma, D., U. Said, and W. Srigutomo, 2017, 2D gravity inversion using particle swarm optimization method: 23rd European Meeting of Environmental and Engineering Geophysics, EAGE, Extended Abstracts, We 23P2 16, doi: 10.3997/2214-4609.201702117.

deGroot-Hedlin, C., and S. Constable, 1990, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data: Geophysics, **55**, 1613–1624, doi: 10.1190/1.1442813.

deGroot-Hedlin, C., and S. Constable, 1993, Occam's inversion and the North American central plains electrical anomaly: Journal of Geomagnetism and Geoelectricity, **45**, 985–999, doi: 10.5636/jgg.45.985.

Dong, H., and A. G. Jones, 2018, On the influences of random starting/prior models in three-dimensional magnetotelluric inversion: 24th EM Induction Workshop, Abstract, 258.

Dosso, S., and D. Oldenburg, 1991, Magnetotelluric appraisal using simulated annealing: Geophysical Journal International, **106**, 379–385, doi: 10.1111/j.1365-246X.1991.tb03899.x.

Ebbesen, S., P. Kiwitz, and L. Guzzella, 2012, A generic particle swarm optimization matlab function: American Control Conference, IEEE, Extended Abstracts, 1519–1524, doi: 10.1109/ACC.2012.6314697.

Engelbrecht, A. P., 2007, Computational intelligence: An introduction: John Wiley and Sons Ltd.

Everett, M., and A. Schultz, 1993, Two-dimensional nonlinear magnetotelluric inversion using a genetic algorithm: Journal of Geomagnetism and Geoelectricity, **45**, 1013–1026, doi: 10.5636/jgg.45.1013.

Farquharson, C. G., and D. W. Oldenburg, 2004, A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems: Geophysical Journal International, **156**, 411–425, doi: 10.1111/j.1365-246X.2004.02190.x.

Fernández Martínez, J. L., E. García Gonzalo, J. P. Fernández Álvarez, H. A. Kuzma, and C. O. Menéndez Pérez, 2010a, PSO: A powerful algorithm to solve geophysical inverse problems: Application to a 1D-DC resistivity case: Journal of Applied Geophysics, **71**, 13–25, doi: 10.1016/j.jappgeo.2010.02.001.

Fernández Martínez, J. L., E. García-Gonzalo, and V. Naudet, 2010b, Particle swarm optimization applied to solving and appraising the streaming-potential inverse problem: Geophysics, **75**, no. 4, WA3–WA15, doi: 10.1190/1.3460842.

Ghaedrahmati, R., A. Moradzadeh, N. Fathianpour, and S. K. Lee, 2014, Investigating 2-D MT inversion codes using real field data: Arabian Journal of Geosciences, **7**, 2315–2328, doi: 10.1007/s12517-013-0869-6.

Godio, A., and A. Santilano, 2018, On the optimization of electromagnetic geophysical data: Application of the PSO algorithm: Journal of Applied Geophysics, **148**, 163–174, doi: 10.1016/j.jappgeo.2017.11.016.

Jones, A. G., 1993a, COPROD2, http://www.complete-mt-solutions.com/mtnet/data/coprod2/coprod2.html, accessed 2 October 2017.

Jones, A. G., 1993b, The COPROD2 dataset: Tectonic setting, recorded MT data, and comparison of models: Journal of Geomagnetism and Geoelectricity, **45**, 933–955, doi: 10.5636/jgg.45.933.

Jones, A. G., and P. J. Savage, 1986, North American Central Plains conductivity anomaly goes east: Geophysical Research Letters, **13**, 685–688, doi: 10.1029/GL013i007p00685.

Kennedy, J., and R. Eberhart, 1995, Particle swarm optimization: International Conference on Neural Networks IV, IEEE, Extended Abstracts, 1942–1948, doi: 10.1109/ICNN.1995.488968.

Kennedy, J., R. Eberhart, and Y. H. Shi, 2001, Swarm intelligence: Morgan Kaufmann Publishers.

Miensopust, M. P., P. Queralt, and A. G. Jones, and , 3D MT modellers, 2013, Magnetotelluric 3-D inversion — A review of two successful workshops on forward and inversion code testing and comparison: Geophysical Journal International, **193**, 1216–1238, doi: 10.1093/gji/ggt066.

Newman, G., 2014, A review of high-performance computational strategies for modeling and imaging of electromagnetic induction data: Surveys in Geophysics, **35**, 85–100, doi: 10.1007/s10712-013-9260-0.

Olalekan, F., and Q. Di, 2017, Particle swarm optimization method for 1D and 2D MTEM data inversion: 87th Annual International Meeting, SEG, Expanded Abstracts, 1219–1224, doi: 10.1190/segam2017-17671977.1.

Pace, F., A. Santilano, and A. Godio, 2017, Particle swarm optimization of electromagnetic data with parallel computing in the 2D Case: 23rd European Meeting of Environmental and Engineering Geophysics, EAGE, Extended Abstracts, Tu 23 B04, doi: 10.3997/2214-4609.201702021.

Perez, R. E., and K. Behdinan, 2007, Particle swarm approach for structural design optimization: Computers and Structures, **85**, 1579–1588, doi: 10.1016/j.compstruc.2006.10.013.

Pérez-Flores, M. A., and A. Schultz, 2002, Application of 2-D inversion with genetic algorithms to magnetotelluric data from geothermal areas: Earth, Planets and Space, **54**, 607–616, doi: 10.1186/BF03353049.

Poli, R., 2008, Analysis of the Publications on the Applications of Particle Swarm Optimisation: Journal of Artificial Evolution and Applications, **2008**, 1–10, doi: 10.1155/2008/685175.

Rasmussen, T. M., 1993, Two-dimensional Occam model of COPROD2 data — First order description of resolution and variance: Journal of Geomagnetism and Geoelectricity, **45**, 1027–1037, doi: 10.5636/jgg.45.1027.

Ratnaweera, A., S. K. Halgamuge, and H. C. Watson, 2004, Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients: IEEE Transactions on Evolutionary Computation, **8**, 240–255, doi: 10.1109/TEVC.2004.826071.

Santilano, A., 2017, Deep geothermal exploration by means of electromagnetic methods: New insights from the Larderello geothermal field (Italy): Ph.D. thesis, Politecnico di Torino.

Santilano, A., A. Godio, and A. Manzella, 2018, Particle swarm optimization for simultaneous analysis of magnetotelluric and time-domain electromagnetic data: Geophysics, **83**, no. 3, E151–E159, doi: 10.1190/geo2017-0261.1.

Sen, M. K., and P. L. Stoffa, 2013, Global optimization methods in geophysical inversion: Cambridge University Press.

Shaw, R., and S. Srivastava, 2007, Particle swarm optimization: A new tool to invert geophysical data: Geophysics, **72**, no. 2, F75–F83, doi: 10.1190/1.2432481.

Shi, Y., and R. Eberhart, 1998, A modified particle swarm optimizer: IEEE International Conference on Evolutionary Computation, Extended Abstracts, 69–73, doi: 10.1109/ICEC.1998.699146.

Simpson, F., and K. Bahr, 2005, Practical magnetotellurics: Cambridge University Press.
Siripunvaraporn, W., 2012, Three-dimensional magnetotelluric inversion: An introductory guide for developers and users: Surveys in Geophysics, **33**, 5–27, doi: 10.1007/s10712-011-9122-6.
Uchida, T., 1993, Inversion of COPROD2 magnetotelluric data by use of ABIC minimization method: Journal of Geomagnetism and Geoelectricity, **45**, 1063–1071, doi: 10.5636/jgg.45.1063.
Van den Bergh, F., and A. P. Engelbrecht, 2001, Effects of swarm size on cooperative particle swarm optimisers: Proceedings of the Genetic Evolutionary Computation Conference (GECCO-2001), 892–899.
Wu, N., J. R. Booker, and J. T. Smith, 1993, Rapid two-dimensional inversion of COPROD2 data: Journal of Geomagnetism and Geoelectricity, **45**, 1073–1087, doi: 10.5636/jgg.45.1073.
Xiang, E., R. Guo, S. E. Dosso, J. Liu, H. Dong, and Z. Ren, 2018, Efficient hierarchical trans-dimensional Bayesian inversion of magnetotelluric data: Geophysical Journal International, **213**, 1751–1767, doi: 10.1093/gji/ggy071.
Yan, P., M. A. Garcia Juanatey, T. Kalscheuer, C. Juhlin, P. Hedin, A. Savvaidis, H. Lorenz, and J. Kück, 2017a, A magnetotelluric investigation of the Scandinavian Caledonides in western Jämtland, Sweden, using the COSC borehole logs as prior information: Geophysical Journal International, **208**, 1465–1489, doi: 10.1093/gji/ggw457.
Yan, P., T. Kalscheuer, P. Hedin, and M. A. Garcia Juanatey, 2017b, Two-dimensional magnetotelluric inversion using reflection seismic data as constraints and application in the COSC project: Geophysical Research Letters, **44**, 3554–3563, doi: 10.1002/2017GL072953.
Yuan, S., S. Wang, and N. Tian, 2009, Swarm intelligence optimization and its application in geophysical data inversion: Applied Geophysics, **6**, 166–174, doi: 10.1007/s11770-009-0018-x.
Zhan, Z.-H., J. Zhang, Y. Li, and H. S.-H. Chung, 2009, Adaptive particle swarm optimization: IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), **39**, 1362–1381, doi: 10.1109/TSMCB.2009.2015956.