

Nonlinear crosshole traveltime tomography using hybridizing particle swarm optimization with linearized least squares algorithm

M.R. EBRAHIMI¹, M.A. RIAHI¹ and M. SANIEE ABADEH²

¹ *Institute of Geophysics, University of Tehran, Iran*

² *Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran*

(Received: 4 May 2018; accepted: 27 August 2018)

ABSTRACT Crosshole tomography problem usually uses linearized (local) techniques. Applying a linearized scheme, the cost function will converge to the global minimum only if a starting model close enough to the global minimum is available. In contrast, global optimization approaches have the ability to find global minimum: even in functions with several local minima regardless of the starting model. However, they are expensive especially for most large dimensional seismic problems. We overcome the limitations of each individual approach by introducing a sequential hybrid algorithm combining particle swarm optimization (PSO) with linearized least squares (LLS). PSO simulates the social behaviour of birds flocking or fishes schooling. It is implemented to obtain the appropriate initial guess for the LLS scheme. To compare the performance of the LLS, PSO, and hybrid algorithms, the methods are tested on the noise-free and noisy synthetic data sets. The resulting tomograms show that the proposed hybrid approach is a more effective inversion tool rather than using the LLS and the PSO. Moreover, a comparison between PSO and simulated annealing approach reveals that PSO provides superior results. Finally, real data from the Seinsfeld site was inverted by the hybrid method and a low-velocity anomaly was detected in the site.

Key words: traveltime tomography, nonlinear inversion, linearized least squares inversion, particle swarm optimization, hybrid optimization.

1. Introduction

One of the restrictions of seismic reflection and refraction methods is that the near-surface weathered layer attenuates the high-frequency components which are indispensable for having the high-resolution components of the Earth interior image. A common practice to overcome this problem is to use crosshole data in which the sources and receivers can be put into separate boreholes (Bregman *et al.*, 1989).

Recent applications of the crosshole velocity model are reported in the literature; for instance, dam leakage detection (Dai *et al.*, 2017), geotechnical and engineering (Butchibabu *et al.*, 2017), mining investigation (Wong, 2000; Greenhalgh *et al.*, 2003), heterogeneity detection and anisotropy estimation (Meng *et al.*, 2017), coal exploration (Rossi *et al.*, 2009), and detection of anomalies with sharp boundaries (Cardarelli, 2008).

Crosshole traveltimes tomography is a nonlinear optimization problem as in many areas the velocity changes by 10-20% or more. Subsequently, this causes significant ray path curvature (Bregman *et al.*, 1989). The tomography problem can be written as:

$$\mathbf{d} = \mathbf{G}(\mathbf{m}) + \mathbf{n} \quad (1)$$

where \mathbf{m} is the model parameter (representing seismic velocity, or slowness, as a function of position), \mathbf{d} expresses the traveltimes data, and \mathbf{G} is a nonlinear operator that predicts traveltimes given a velocity model and survey geometry. The data error (noise) is labeled as \mathbf{n} . We can also write Eq. 1 in a format that explicitly denotes the discrete nature of the data (Aster *et al.*, 2013).

Typically, there are two main approaches to solve the nonlinear traveltimes tomography: 1) gradient-based techniques (local search) which depend on the gradient information of the objective function, e.g. least squares, Levenberg-Marquardt, Occam, conjugate gradient, etc.; 2) global optimization schemes which simulate certain social behaviour, natural physical process or biological phenomenon, e.g. genetic algorithm (Mazzotti *et al.*, 2016), Monte Carlo (Sen and Stoffa, 2013), simulated annealing (Sen and Stoffa, 2013), particle swarm optimization (PSO), etc.

The minimum of the objective function in the gradient-based (local) methods can be moved through local linearization scheme iteratively. The linearized iterative techniques start from a suitably initial (starting) velocity model which has to be determined by the user. The final result of the local methods strongly depends on the starting model (Moret *et al.*, 2006). This approach will converge to the global minimum with greater possibility if the starting model is adjacent to the global minimum of the objective function. In other words, if the starting model is far from the global minimum, the method will converge to a local optimum of the objective function depending on the position of the starting model. Another difficulty in the convergence of the gradient-based techniques is inherent instabilities in calculating numerical partial derivatives of the ray paths. However, a fast computation is the advantage of local search methods (Sen and Stoffa, 2013).

On the contrary, global optimization methods use information about the objective function more globally to update the current model. Besides, these schemes are not much affected by the initial model and then even work with poor initial models. As a result, these algorithms have the ability to find the global minimum of some functions which has several local minima regardless of the starting model and the linearization process. Furthermore, global search techniques do not require partial derivatives; therefore they are able to solve even highly complex nonlinear inverse problems. However, they are expensive especially with large dimensionality that is involved in most seismic inverse problems and the considerable number of forward calculations which is required (Soupios *et al.*, 2011; Sajeve *et al.*, 2017).

Briefly, the main drawbacks of the local and global optimization approaches are the requirement of a starting model and the computational cost, respectively. Consequently, to overcome the disadvantages of each method, a sequential hybrid optimization scheme can be used. In the sequential hybrid method, a model close to the global minimum of the objective function (a suitable initial model) is obtained by the global approach, while the local approach applies in order to achieve the accurate solution close enough to the global minimum of the objective function (Chunduru *et al.*, 1997). Several successful sequential hybrid optimization approaches have been

proposed to solve the geophysical data inverse problems (Chunduru *et al.*, 1997; Mendes, 2009; Tognarelli *et al.*, 2016; Noack and Funke, 2017; Galuzzi *et al.*, 2018).

Göktürkler (2011) and Soupios *et al.* (2011) use sequential hybrid approaches based on the GA/least squares and SA/least squares to invert seismic crosshole data, respectively. Their studies show that the hybrid strategy is a more effective tool to invert crosshole traveltime data rather than local or global technique alone.

In this study, a sequential hybrid tomographic inversion scheme based on PSO method and linearized least squares (LLS) inversion is proposed. PSO is a stochastic global optimization method introduced by Kennedy and Eberhat (1995). It is inspired by social behaviour of fish schooling or bird flocking. PSO, similar to other global approaches, explores the model space globally because it applies a random search to update the model. It has been used for the solution of geophysical problems (Song *et al.*, 2012; Tronicke *et al.*, 2012; Zhe and Hanming, 2013; Paasche and Tronicke, 2014; Guo *et al.*, 2018; Liu *et al.*, 2018).

In this paper, first, a robust and efficient algorithm for solving nonlinear crosshole traveltime tomography is presented which utilizes hybrid of the PSO and LLS algorithms. The proposed tomography algorithm is a combination of a global search method based on the PSO algorithm and a gradient-based (local) approach performed by the LLS algorithm. Since the global techniques are not sensitive to the initial model, the PSO is applied to construct a suitable starting model (background velocity distribution) for the LLS scheme which gives fast optimum searches. Consequently, this technique reduces computation cost and escapes from being stuck into the local minima of the cost functions. Moreover, it obtains a good model even with poor starting solution. To compare the performance of LLS, PSO and sequential hybrid application, the methods are tested, both in sequential hybrid and independently, on the noise-free and noisy synthetic crosshole seismic data sets. Then the PSO runtime and its convergence are compared to the SA approach. Finally, the sequential hybrid method is tested on a real crosshole data set.

2. Linearized least squares

The least squares solution of Eq. 1 is to find an \mathbf{m} that minimizes the sum of the squared differences between calculated and observed traveltimes. If the error in the data is Gaussian, then the least squares of this difference (objective function) will be:

$$\|\mathbf{G}(\mathbf{m}) - \mathbf{d}\|_2^2 \quad (2)$$

where $\|\cdot\|$ denotes the norm 2.

Generally, the result accuracy of Eq. 2 will depend on several factors including: 1) the assumptions made in the forward calculations; 2) the noise in the observed data; 3) the ray geometry dependency on the velocity model; 4) error measurements in the source or receiver positions; 5) the inaccurate first break picking.

Accordingly, the problem is extremely ill-conditioned. To provide extra constraints on the model space, a regularization term is frequently incorporated into the cost function. This term helps to reduce the non-uniqueness of the solution (Tikhonov, 1963). Consequently, the regularized objective function will be:

$$\|G(m) - d\|_2^2 + \alpha^2 \|L_i m\| \quad G \in R^{m \times n}, \quad m \in R^{m \times 1}, \quad d \in R^{m \times 1} \quad (3)$$

where α is the regularization parameter (damping factor). It regulates the tradeoff between the data misfit and the model roughness. We use the L-curve criterion approach to determine the optimum value of α (Hansen, 1992).

Matrix L_p , $i = 0, 1, 2$ is the 2-D roughening operator. It is defined as:

$$L_0 = I, \quad L_1 = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & \dots & -1 & 1 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \end{bmatrix} \quad (4)$$

Selection of the matrix L depends on the prior information about the model parameters. The higher order of L leads to a much smoother model (Aster *et al.*, 2013).

As it was said before, the function $G(m)$ in this research is nonlinear, therefore, in order to use the nonlinear regularized least squares method, linearization should be performed first.

Given a trial model at the k th iteration (m_k), m_{k+1} can be obtained using the Taylor's series approximation about m_k and ignoring second-order terms (Aster *et al.*, 2013; Ghanati *et al.*, 2017):

$$G(m^k + \delta m) \approx G(m^k) + J(m^k) \delta m \quad (5)$$

where J is the linear differential operator (Jacobian) matrix. Mathematically, J is defined as

$$J_{ij}(m^k) = \frac{\partial G_i}{\partial m_j}, \quad j = 1, 2, 3, \dots, N, \quad i = 1, 2, 3, \dots, M \quad (6)$$

where M is the number of traveltime data and N implies the number of model parameters (slowness).

Using Eq. 5, the solution of nonlinear regularized least squares at the $(k+1)$ th iteration is:

$$m^{k+1} = m^k + \Delta m, \quad \Delta m = (J(m^k)^T J(m^k) + \alpha^2 L^T L)^{-1} \{ J(m^k)^T [G(m^k) - d] - \alpha^2 L^T L m^k \} \quad (7)$$

To stabilize the solution of the LLS algorithm, we have also applied the second-difference regularization scheme.

3. Particle swarm optimization

PSO belongs to the swarm intelligence algorithms and it is a population-based search method which imitates the social behaviour of birds or fishes. PSO explores the model space of a cost function by adjusting the trajectories of individual agents. In the PSO these agents are called "particles". The movement of a particle includes both stochastic and deterministic components.

On the other hand, each individual is attracted towards its own best location in history and the position of the current global best of the swarm. Furthermore, it simultaneously has a tendency to move randomly (Kennedy and Eberhart, 1995; Hedeshi and Saniee Abadeh, 2014).

A particle within the swarm can interact with other particles; however their capabilities are relatively restricted by some specific rules. Thus, the behaviour of a particle in a swarm is usually negligible, while their social and collective social behaviour is of profound significance. Each particle is expressed by a vector in the search space which is called the velocity vector and controls the next movement of the agents. All particles update their velocity based on three components including: 1) current velocity; 2) the best position which they have explored up to now; 3) the global best position explored by the swarm (Hedeshi and Saniee Abadeh, 2014; Awange *et al.*, 2018).

In the PSO, the swarm of particles moves through a predefined search space. Each particle p in the swarm, has the following components at time t :

$x_{p,i}(t)$: i th component of the particle position,

$v_{p,i}(t)$: i th component of the particle velocity,

$y_{p,i}(t)$: i th component of the personal best position of particle p ($pbest$),

$y_i(t)$: i th component of the global best position of the swarm ($gbest$).

Suppose that f indicates the objective function which we want to find its global minimum in an N -dimensional space, an update of the $pbest$ of particle p in iteration $t+1$ is (Kiranyaz *et al.*, 2014):

$$y_{p,i}(t+1) = \begin{cases} y_{p,i}(t) & \text{if } f(x_p(t+1)) > f(y_{p,i}(t)) \\ x_{p,i}(t) & \text{else} \end{cases} \quad \forall i \in [1, N] \quad (8)$$

Because $gbest$ is the index of the global best particle, then:

$$\hat{y}(t) = y_{gbest}(t) = \min [y_1(t), \dots, y_S(t)].$$

For each iteration, positional updates for each particle, $p \in [1, S]$ along with each dimensional component, $i \in [1, N]$, are:

$$v_{p,i}(t+1) = w(t)v_{p,i}(t) + c_1 r_{1,i} [y_{p,i}(t) - x_{p,i}(t)] + c_2 r_{2,i} [\hat{y}_i(t) - x_{p,i}(t)] \quad (9)$$

$$x_{p,i}(t+1) = x_{p,i}(t) + v_{p,i}(t+1) \quad (10)$$

where c_1, c_2 are the acceleration coefficients and w is the inertia weight. r_1 and r_2 are uniform distributed random variables in $[0, 1]$ at each iteration.

Notice that, in the Eq. 9, the first term is the memory term so that it expresses the history of the previous information, the second and the third terms are the cognitive and the social parts, respectively. The cognitive part describes the particle's own experience. In the social part, the particle is guided by the best particle towards the global best solution so far achieved. The $gbest$ is the common guide for all particles so that the social part exists in each velocity update equation (Sayed, 2014).

Depending on the problem to be minimized, the stopping criteria for the iteration in the PSO can be as below: 1) repeated until a maximum number of iterations is exceeded, 2) velocity

updates become zero, and 3) an acceptable cost function value is obtained (Kiranyaz *et al.*, 2014).

Swarm size, ω , c_1 , and c_2 control the performance of the PSO algorithm. Various authors have studied the impact of these parameters on the convergence of the PSO (e.g. Trelea, 2003; Jiang *et al.*, 2007; Qian and Li, 2017).

Several PSO features make it so efficient in solving optimization problems in comparison with other global search methods. They are as follows (Rezaee Jordehi and Jasni, 2013; Awange *et al.*, 2018):

- PSO has fewer parameters to be tuned by a user;
- its concepts and coding are very simple;
- its convergence is fast;
- it needs less computational burden;
- PSO's computational behaviour is less affected by the initial models.

The PSO algorithm to obtain background velocity for the LLS inversion is carried out by the following steps:

1: a swarm of particles with random positions were distributed into the model space. Their initial velocities are set to zero;

2: the forward calculation, for each particle, is performed. The cost function is evaluated, and the *pbest* for each particle and *gbest* for swarm are stored. The objective function Φ is (Göktürkler, 2011; Soupios *et al.*, 2011):

$$\Phi = \frac{1}{N} \sum_{j=1}^N (t_j^{obs} - t_j^{cal})^2 \quad (11)$$

where N is the number of traveltimes, t^{cal} and t^{obs} are the calculated and observed traveltimes, respectively and i shows each observation;

3: if the updated model is better than the current *pbest* and *gbest*, then the velocity and position vectors will be updated. If the location of each particle exceeds the boundaries of the defined parameter space, the direction of the velocity vector of the particles will be reversed;

4: steps 2-3 are repeated until a certain stopping criterion is reached. Ultimately, the final *gbest* position is saved as the final solution of the inverse problem.

4. Synthetic data example

In this paper, all the forward modelling and optimization methods are coded and implemented in MATLAB.

We compute the first-arrival times from source to receiver through the finite-difference solution of the eikonal equation (Zelt and Barton, 1998).

The synthetic velocity model between two boreholes along with the source-receiver geometry and the bent ray paths are shown in Fig. 1. The size of this study is 10×20 m. This model displays a number of layers with different velocities and thicknesses including horizontal and dipping interfaces, one low-velocity anomaly in the interval of 12.5 and 14.5 m and two high-velocity anomalies. The high-velocity anomalies are located at the depths of 5.5 to 7.5 m and 8.5 to

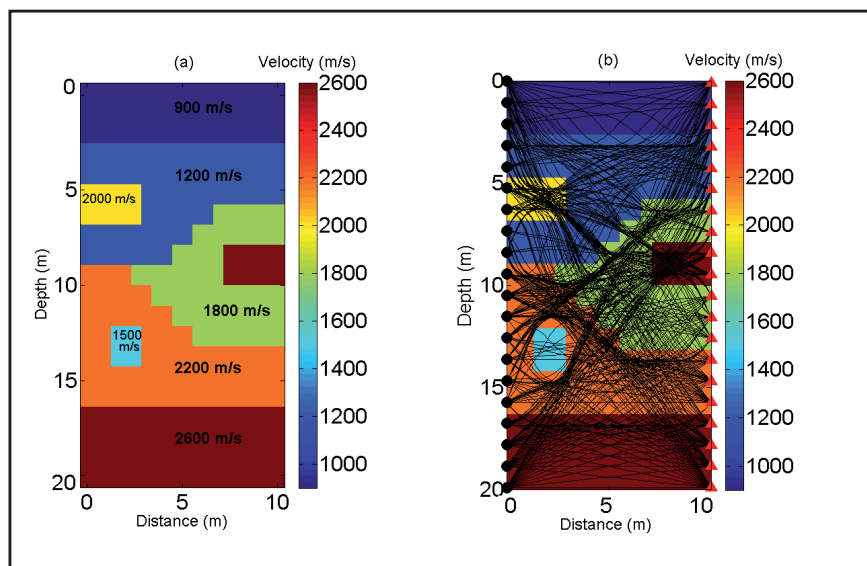


Fig. 1 - Synthetic model (a) and source-receiver geometry (b); Red triangles and black circles indicate receivers and sources, respectively. The bent raypaths are displayed by the black lines.

10.5 m. The velocity contrast between the background and the anomalies presented in this example is more than 20% which is large enough to claim the nonlinear inversion, rather than the linear one. The velocities range from 900 to 2600 m/s.

The test incorporates 20 equally spaced sources (the black filled circles) and 20 equally spaced receivers (red filled triangle), producing 440 ray paths. Both shots and receivers are located between 0.1 and 19.9 m. The straight ray path coverage and source-receiver geometry for this experiment is presented in Fig. 1b. The medium is discretized into 20×10 constant slowness (velocity) grids.

To study the noise effect on the inverted models, the noisy data was generated by adding random noise to the noise-free data set. The added noise has a Gaussian distribution with zero mean and ±0.1 ms standard deviation.

4.1. Parameter selection in the PSO

4.1.1. The inertia weight

The inertia weight (w) brings about a right balance between exploitation and exploration process. It controls the contribution rate of a particle’s previous velocity to its current velocity. There was no concept of inertia weight in the first version of the PSO. It was introduced by Shi and Eberhart (1998) who considered w to be a constant value. They expressed that a small value of w leads to a local search while a large value of w helps a global search. Then, many researchers introduced many strategies for dynamic adjustment of the inertia weight to enhance the capabilities of the PSO convergence (e.g. Shi and Eberhart, 1998, 1999; Singh *et al.*, 2007; Yue-lin and Yu-hong, 2007). In our paper, we set inertia weight linearly decreasing from 0.99 to 0.5 as Eberhart and Shi (2000) have proposed which results in a gradually change from an exploratory mode to an exploitative one.

4.1.2. The maximum velocity

The maximum limit of the particle velocity also affects the performance of the PSO algorithm. If it is too high, then the particles' movement becomes erratically. As a result, without adequate search space exploration, they are rapidly attracted to the global best (*gbest*). Thus, the chance of getting stuck in the false optima decreases, as well as the particles may exceed search space. On the contrary, a too small maximum velocity slows down the algorithm, the particles move is excessively bounded and the algorithm convergence is not guaranteed. Accordingly, a suitable maximum velocity value should be selected. Abido (2001) has recommended the following maximum velocity value in order to guarantee uniform velocity throughout all dimensions:

$$V_{max} = \frac{X_{max} - X_{min}}{N_I} \quad (12)$$

where X_{min} and X_{max} are the minimum and maximum values which particles have reached so far, respectively. N_I is the number of intervals in the d th dimension selected by the user.

In our PSO application, we set $V_{max} = 0.1 * (X_{max} - X_{min})$ which was an appropriate maximum velocity value.

4.1.3. The acceleration coefficients

In the PSO, two stochastic acceleration parameters (c_1, c_2) guide the search towards the best solution. C_1 pulls the particles towards their own best solutions (*pbest*) which keeps the diversity of the population and improves the grouping capability. While c_2 is a convergence part and attracts all particles towards the *gbest*.

When c_1 increases, compared with the c_2 , it pulls the individuals towards their *pbest* and results in excessive wandering of particles within the model space and decreases their convergence to the *gbest* solution. On the other hand, a larger value of c_2 leads individuals to converge to a local optimum (Kennedy and Eberhart, 1995). Consequently, these two parameters play a key role in reaching the optimum solution efficiently and accurately.

In literature, a frequently acceptable setting for most of the optimization problems is $c_1=c_2=2$ (Ozcan and Mohan, 1999; Suganthan, 1999). In our PSO implementation, we vary c_1 and c_2 within reasonable ranges and we arrive at similar conclusions ($c_1=c_2=2$) which is a suitable choice for our tomographic problem.

4.1.4. The swarm size

Too many particles slow down the algorithm, whereas too few particles cause the algorithm to get stuck into the local optima. There is no specific rule for determining the particle size. But generally, the particle size should be increased if the dimension of the problem at hand increases.

We found that 350 particles are the optimal number of particles in this experiment and it is sufficient to explain our data.

4.1.5. The stopping criteria

The PSO algorithm is repeated until a predefined stopping criterion (acceptable limit of the objective function and/or maximum number of iterations) is satisfied and, then, the final global

best position of the swarms is saved as the optimum model. The PSO algorithm in our synthetic example, as the initial model for the LLS inversion, will terminate if the relative error becomes less than 1% on the two successive iterations. Then, we select this velocity model as the initial model for the LLS approach. This means that we start the LLS approach when the starting velocity model is already close enough to the global optimum.

For the pure PSO inversion strategy, the PSO algorithm will stop if the relative error becomes less than 0.1% on the two successive iterations.

4.2. Results

Fig. 2 shows the resulting tomograms from the PSO algorithm after 136 iterations for the noise-free and 149 iterations for the noisy data sets as the initial model for the LLS inversion. To take into account the stochasticity of the processes; we perform 50 runs for each data set and then plot their average convergence. The obtained models somehow display the subsurface anomalies specially the low-velocity zone located at the top and the high-velocity layer located at the bottom of the model. However, these solutions are not precise. Table 1 summarizes the PSO parameters used in the inversion of the synthetic data sets.

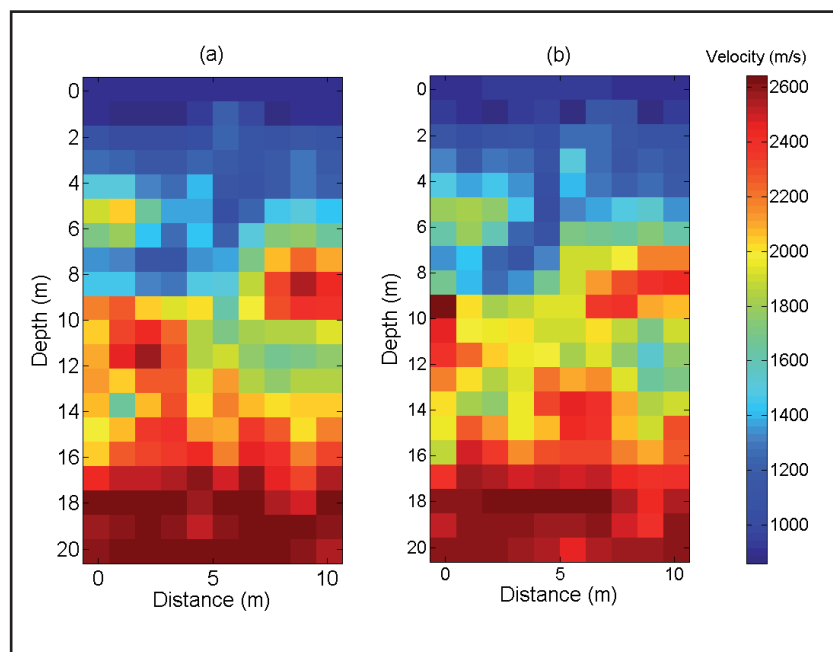


Fig. 2 - The velocity model inverted from the PSO algorithm using the synthetic data: a) the noise-free data after 136 iterations; b) the noisy data set after 149 iterations.

Furthermore, Fig. 3 displays their corresponding convergence curves (the iteration number with respect to the cost function value). The convergence curves gradually decrease for both data sets (noisy and noise-free), as it is shown in the figure. In addition, the errors related to the noise-free data are lower than the noisy data.

Fig. 4 presents the results of the PSO, sequential hybrid and LLS optimization methods for the synthetic model. Figs. 4a and 4d show the results of the LLS optimization. A homogeneous velocity of 1500 ms⁻¹ considered as the starting models for both cases. They were obtained after

Table 1 - The PSO parameters used in optimization of the synthetic model.

Number of model parameters	Particle size	w	c_1	c_2	Initial model
200	350	Linearly decreased from 0.99 to 0.5	2	2	homogeneous velocity (1500 m/s)

Table 2 - The LLS parameters used in optimization of the synthetic model.

Number of model parameters	Number of traveltimes	Type of regularization	Initial model
200	441	Second difference	homogeneous velocity (1500 m/s)

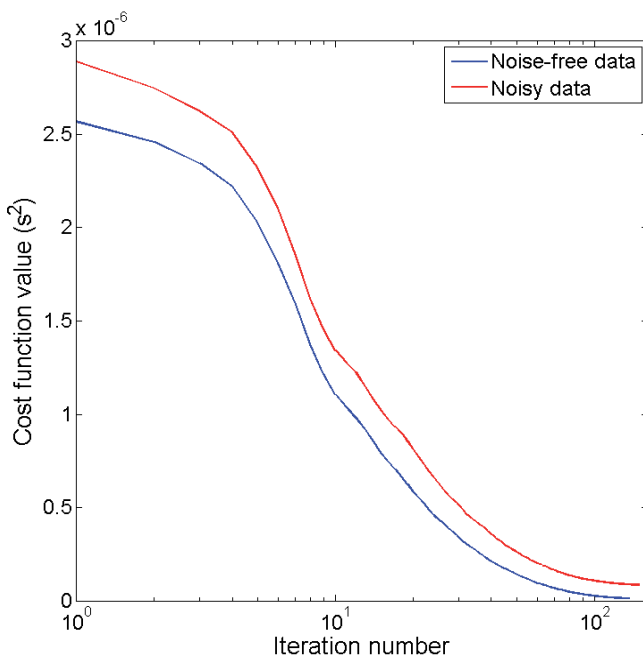


Fig. 3 - Convergence curves of the synthetic data for the PSO inversion. Blue and red curves indicate noise-free and noisy data, respectively.

10 iterations. Table 2 provides the LLS inversion parameters in the both sequential hybrid and local approaches.

The final tomograms obtained by the pure PSO scheme with noise-free data after 385 iterations and noisy data after 402 iterations are illustrated in Figs. 4b and 4e, respectively.

Figs. 4c and 4f show the final models obtained by the sequential hybrid scheme with the noise-free and noisy data, respectively. Their corresponding initial velocity models are shown in Fig. 2. They were obtained by the LLS optimization approach after 6 iterations. It should also be noted that we get Fig. 4 by using the same tomographic grid of Fig. 2 and then oversampled to smoother illustration.

There are two criteria to compare the results of the obtained tomograms for the synthetic model, including RMS errors between the observed and calculated traveltimes and Euclidean distances between the true and the inverted models. A smaller value of RMS errors indicates better fitting between the observed and calculated data. Also, a smaller value of the Euclidean

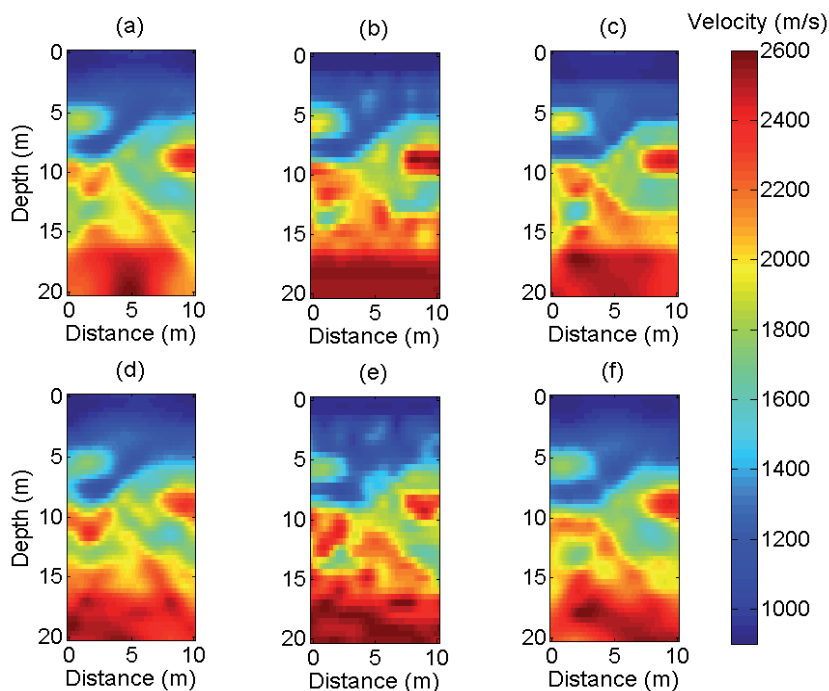


Fig. 4 - Velocity tomograms obtained from noise-free data using a) the LLS algorithm; b) PSO and c) the sequential hybrid optimization. Velocity tomograms obtained from noisy data using d) the LLS algorithm; e) the PSO approach; f) the sequential hybrid optimization scheme.

distance expresses that the inverted model is in the vicinity of the true model. Table 3 summarizes the RMS error and Euclidean distance for the obtained tomograms.

As expected, the results for the sequential hybrid approach are better than the results for the LLS scheme in terms of having smaller RMS errors and Euclidean distances in both noise-free and noisy data sets. Furthermore, the PSO and the sequential hybrid results in the noise-free data are almost the same. However, their comparison in the noisy data demonstrates the superiority of the sequential hybrid algorithm. It is reasonable to conclude that the regularization term in the hybrid approach makes the solution more stable. Also, the RMS errors and the Euclidean distances of the hybrid scheme in the noise-free data are lower than the noisy data. The solution by the hybrid optimization is less affected by the added noise than the LLS and the PSO methods.

It is interesting to notice that PSO, LLS, and sequential hybrid approaches take about 4600, 120, 1500 s of the CPU time, respectively (on a personal computer, CPU: Intel Core2 2.66 GHz, Memory: 4GB). It should also be noted that each iteration involves $20 \times 10 = 200$ forward calculations.

Table 3 – The RMS errors and the Euclidean distances for pure PSO, sequential hybrid and LLS approaches with both noisy and noise-free data sets.

	Noise-free data			Noisy data		
	sequential hybrid	PSO	LLS	sequential hybrid	PSO	LLS
RMS errors (ms)	0.02	0.02	0.036	0.11	0.142	0.13
Euclidean distances (s/km)	0.018	0.019	0.045	0.033	0.097	0.081

4.3. Comparison between PSO and simulated annealing as the initial guess

We compare the results of the PSO algorithm and the simulated annealing in order to evaluate the effectiveness of the algorithm. We only test noise-free data. For a better comparison between the performances of SA and PSO, we use the same search space and number of iterations for SA scheme. Also, 350 iterations per temperature were performed (analogous to 350 particles in the PSO). The SA parameters are summarized in Table 4. To take into account the stochasticity of the processes; we perform 50 runs for each method and then plot their average convergence curves. According to Fig. 5, the PSO model in each iteration performs better than the SA model, because the PSO has lower convergence curve. It means that the PSO solution is more precise. Furthermore, although the SA convergence rate is faster than that of the PSO, the SA solution at the end of 136 iterations is not better than the PSO solution.

Table 4 - The parameters for the SA optimization

Iteration	Initial temperature	Temperature reduction rate (C)
136	0.001	0.99

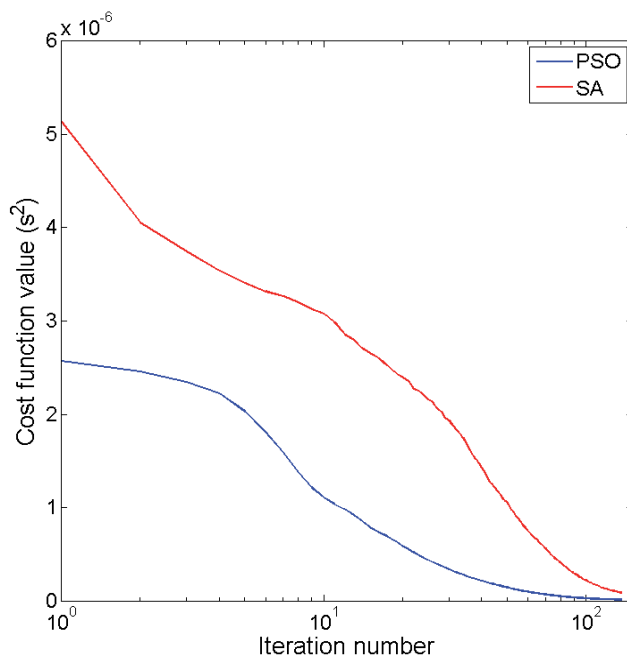


Fig. 5 - Convergence curves of the synthetic noise-free data for the PSO (the blue curve) and the SA (the red curve) schemes after 136 iterations.

Further, the PSO takes about 1500 s to finish whereas the SA takes about 2200 s (~1.5 times slower than the PSO).

The results demonstrate the superiority of the sequential hybrid PSO-LLS approach in comparison with the sequential hybrid SA-LLS scheme (Göktürkler, 2011) in terms of runtime and precise solution. Briefly, after 136 iterations: 1) PSO is faster than SA, it is equivalent to having faster hybrid PSO-LLS than hybrid SA-LLS, too, and 2) PSO gives a model close to the true model as the initial guess for the LLS, but SA does not.

5. Real crosshole data experiment

5.1. Data description

A real crosshole data acquisition was carried out in a site near the Seinsfeld (Rhineland-Palatinate, Germany). The test site is an aquifer and is covered mostly by the Triassic sandstones. The crosshole survey was carried out to identify the subsurface structures between two wells separated by 48 m and extending to the depth of 60 m. To generate reputable pulses with high-frequency content, a down hole sparker manufactured by the Geotomografie GMBH was used in the first well. A 24-channel hydrophone string with a constant spacing of 1 m was used in the second well. P-wave first arrival times are picked for the optimization process from the acquired seismic traces (Kroger *et al.*, 2011). An exemplary shot gather is shown in Fig. 6. As shown in this figure, the field data has high quality and almost all first breaks are easily distinguishable.

We select a region including 33 sources and 24 receivers, resulting in 792 first breaks of P-waves. The string of 24 hydrophones recorded all the shots. The first shot location was at the bottom of the well and other shots are located upwards. The first receiver is also located at the bottom of the well.

The straight raypaths coverage and the source-receiver geometry of the experiment are shown in Fig. 7. The tomographic model was discretized into 384 velocity (slowness) cells (24 cells in horizontal and 16 cells in the vertical direction).

5.2. Applying the hybrid approach to the real data

Fig. 8 shows the velocity model and the corresponding convergence curve of applying the PSO algorithm to the real data after 156 iterations, respectively. A homogeneous velocity of 3000 ms^{-1} is considered as the starting model for the PSO optimization in the real data. The stopping criterion is the same as the synthetic example (Section 4.1.5). As expected, the convergence curve of the PSO gradually decreases. For our real crosshole data, we found that 350 particles are sufficient to explain our data.

The RMS curve for the sequential hybrid inversion of the real data after 7 iterations is shown in Fig. 9a. Referring to Fig. 9a, it is observed that the RMS curve decreases with the iteration at

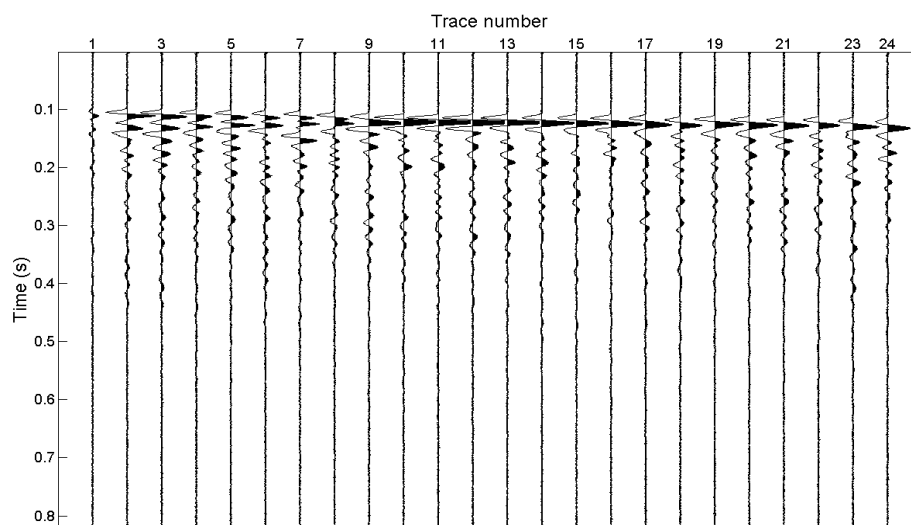


Fig. 6 - An example of the time signal traces recorded on the 24-channel hydrophone array in the Seinsfeld site.

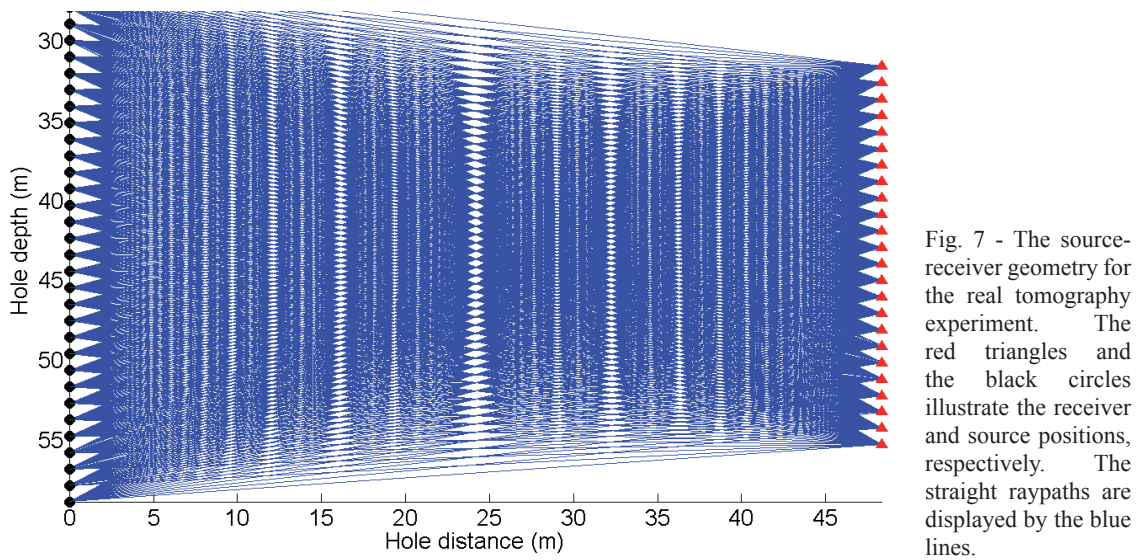


Fig. 7 - The source-receiver geometry for the real tomography experiment. The red triangles and the black circles illustrate the receiver and source positions, respectively. The straight raypaths are displayed by the blue lines.

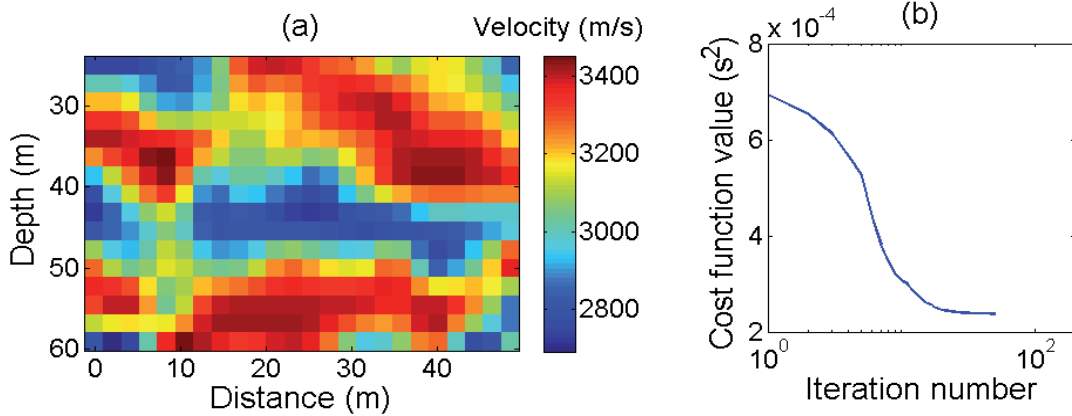


Fig. 8 - Velocity model inverted from the PSO algorithm using the real data after 156 iterations (a) and the corresponding convergence curve (b).

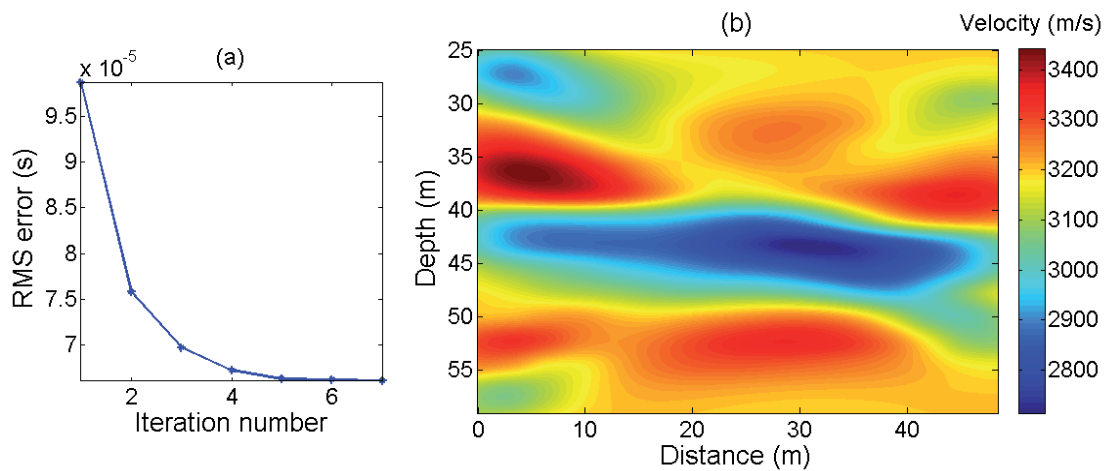


Fig. 9 - RMS curve for 7 iterations of the sequential hybrid inversion (a) and illustration of resulting tomogram from the sequential hybrid optimization at the Seinsfeld (Germany) site test (b). Velocity ranges from 2650 ms⁻¹ to 3450 ms⁻¹.

the first five iterations, then, its rate decreases. Hence, this number of iteration is enough to study the subsurface inhomogeneities and anomalies.

Fig. 9b shows the obtained velocity distribution using the sequential hybrid algorithm. Its initial velocity model is shown in Fig. 8a. It was obtained by the LLS inversion approach with 5 iterations. It should also be noted that we get Fig. 9b by using the same tomographic grid of Fig. 8a and then oversampled to smoother illustration.

The obtained model shows that the velocity between two wells changes from 2650 ms⁻¹ to 3450 ms⁻¹. Furthermore, a low-velocity lens (~2500 ms⁻¹) between 40 and 48 m is detected. This anomaly is in accordance with the acquired EM tomogram carried out in this test site (Gheymasi *et al.*, 2016).

Fig. 10 illustrates a comparison of the four observed and calculated traveltimes from the sequential hybrid approach. This figure shows that there is a good agreement between the observed and calculated traveltimes in all shots.

5.3. Post-inversion model assessment analyses

Any inverted tomogram needs post-inversion model assessment analyses including a combination of appraisal and estimation problems. Estimating covariance of the model parameters is one of the possible procedures for the appraisal approach (Tarantola, 1987; Menke, 2012).

5.3.1. Model covariance matrix

The main diagonal elements of the covariance matrix give an estimate of how the data errors and uncertainties in the assumptions about the model within the inversion process are mapped into the model parameter error. The model covariance matrix (MCM) is given by (Cardarelli, 2003; Menke, 2012):

$$Cov(\mathbf{m}) = \mathbf{J}^g(\mathbf{m})Cov(\mathbf{d})[\mathbf{J}^g(\mathbf{m})]^T \tag{13}$$

where $\mathbf{J}^g = [\mathbf{J}(\mathbf{m})^T \mathbf{W}_d^T \mathbf{W}_d \mathbf{J}(\mathbf{m}) + \alpha^2 \mathbf{L}^T \mathbf{L}]^{-1} \mathbf{J}(\mathbf{m})^T$ is the generalized inverse of the Jacobian matrix. The superscript ‘*T*’ indicates matrix transpose, *Cov(d)* is the covariance of a data vector and *W_d* is the data weighting matrix. We assumed that data errors are uncorrelated, so *W_d* is a

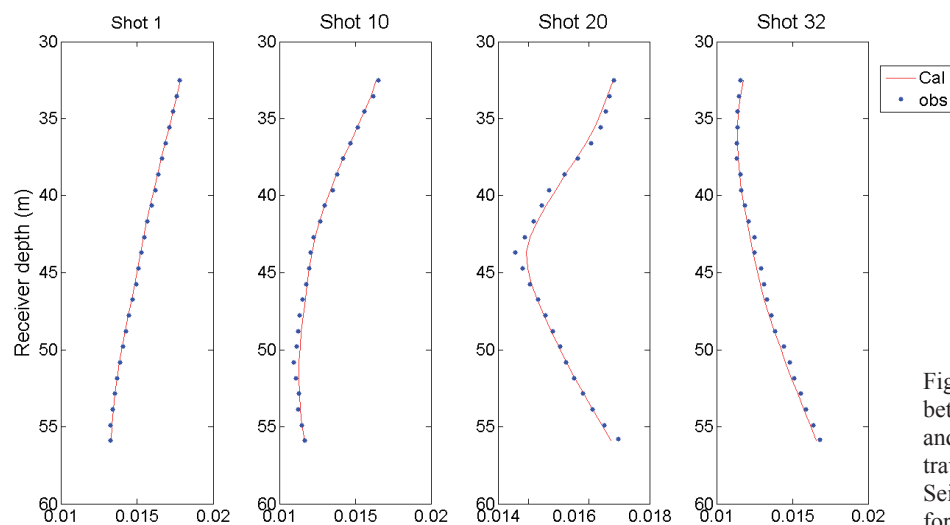


Fig. 10 - Comparison between the calculated and observed traveltimes using the Seinsfeld site data sets for 4 shots.

diagonal matrix with entries that are the reciprocals of the standard deviations of the uncertainties associated with each traveltime. In general, W_d is rough estimate. Hence, the obtained uncertainty by Eq. 13 is an approximation of the true uncertainties.

Moreover, the square root of the main diagonal elements of the covariance matrix gives the error of the model parameter (Menke, 2012; Ghanati *et al.*, 2017):

$$error(m_i) = \sqrt{Cov(m_{ii})} \tag{14}$$

Fig. 11a shows the model covariance matrix for the real data. The diagonal trend in the matrix is clearly visible. In the covariance matrix, the off-diagonal values for any row, display the correlation between the error in the cell corresponding to the diagonal element, and the error in all the other cells.

Fig. 11b shows the slowness errors (Eq. 14) for the tomogram in Fig. 9b. If we ignore the low ray density cells (the blue color zones at the top and the bottom of the figure), all of the slowness

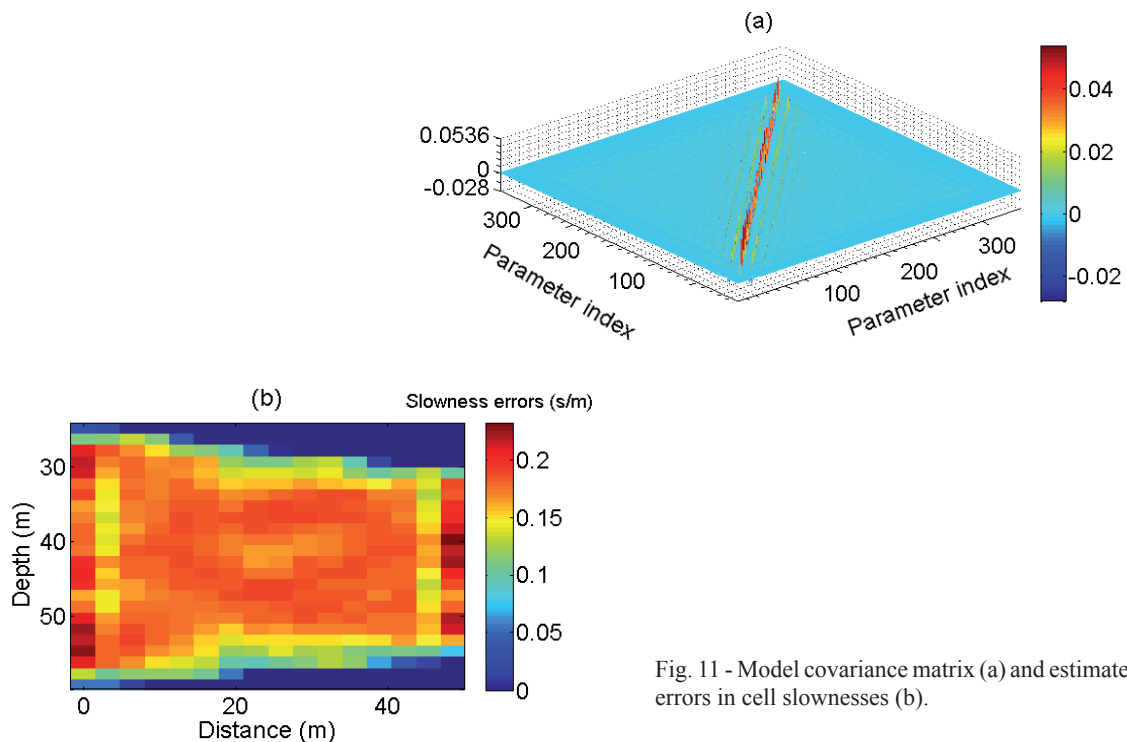


Fig. 11 - Model covariance matrix (a) and estimated errors in cell slownesses (b).

errors will be less than 0.23 s/m and greater than 0.05 s/m. It means that the velocity errors will be between 4 and 20 m/s. These results suggest that the obtained tomogram is relatively insensitive to errors in the traveltimes.

5.3.2. Model resolution matrix

The second tool to evaluate a derived inverse model is determining the model resolution matrix. Model resolution matrix inquires how closely an estimated model is matched to the true model (Yao *et al.*, 1999; Aster *et al.*, 2013). The model resolution matrix becomes:

$$error(m_i) = \sqrt{Cov(m_{ii})} \tag{14}$$

If $R(m)$ becomes an identity matrix, it means that each model parameter is uniquely estimated. Otherwise, the estimates of the model parameters are the weighted averages of the true model parameters. On the other hand, if the diagonal entry of MRM corresponding to a particular grid cell is less than 1, then the model slowness of the cell is also affected by the slownesses of surrounding cells. If the diagonal element is 1, then the model slowness of the cell is affected only by the true slowness of the cell (Moret *et al.*, 2006).

Fig. 12a displays the model resolution matrix for the real data. The diagonal trends in the matrix are visible. Fig 12b displays the diagonal entries of the model resolution matrix. The larger value entries indicate better-resolved grid cells and the lower values correspond to poorly-resolved cells.

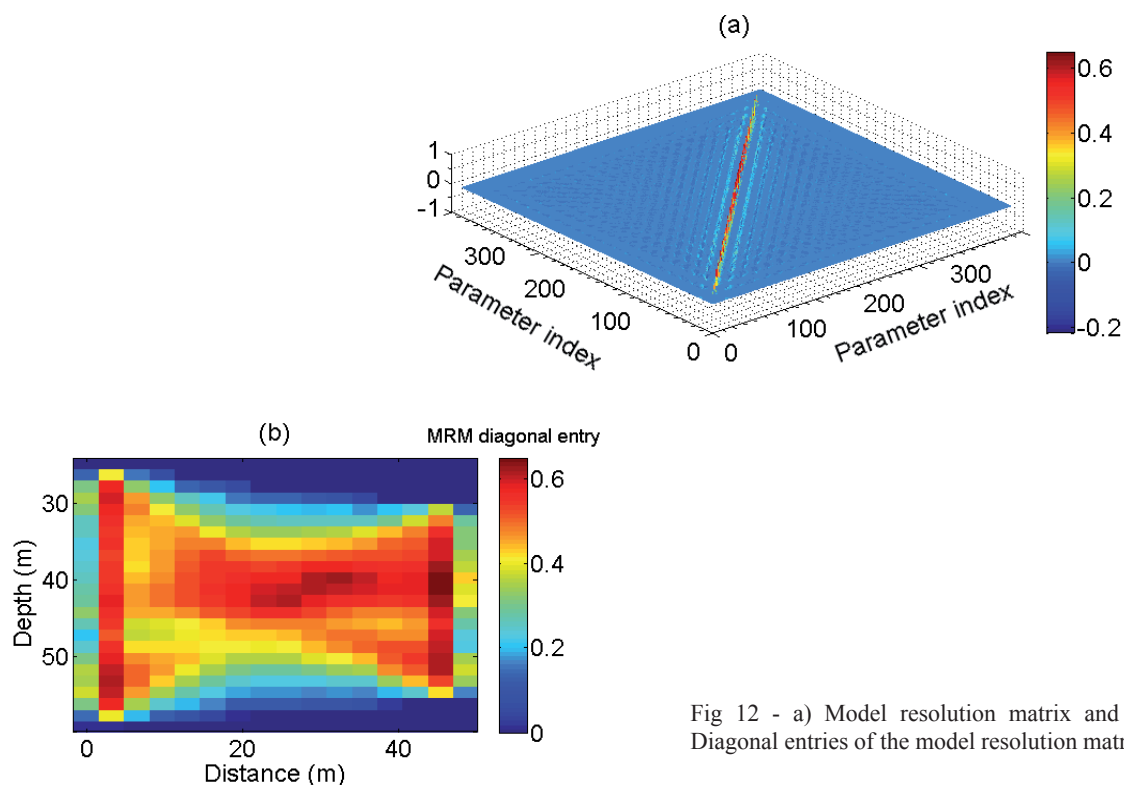


Fig 12 - a) Model resolution matrix and b) Diagonal entries of the model resolution matrix.

6. Conclusions

In this paper, we have suggested a sequential hybrid approach based on a local and a global optimization algorithm for solving nonlinear crosshole traveltime tomography. PSO was selected for the global optimization scheme and the LLS were used for local optimization. The PSO was applied to obtain a good starting velocity model for the LLS scheme.

With the proposed (sequential hybrid) method, the solution can be approached at a faster speed rather than using the PSO alone. Furthermore, this strategy improves the LLS inversion results

by avoiding its shortcomings in the Jacobian partial derivative matrix computation, selecting the starting model, and premature convergence towards local minima.

Applied on both noisy and noise-free data sets, inverting for the synthetic crosshole data indicated that the sequential hybrid optimization approach is more efficient for estimating the velocity model between two wells rather than using the LLS alone. Also, in order to evaluate the results obtained by the PSO algorithm, it is compared with the result from the simulated annealing approach applied on the noise-free data. The results obtained by the PSO are shown to be superior.

Finally, a real data from the Seinsfeld test site was inverted by the sequential hybrid optimization and a low-velocity lens ($\sim 2500 \text{ ms}^{-1}$) between 40 and 48 m was detected. The model covariance matrix of the obtained tomogram show that all of the velocity errors (except low ray density cells) are less than 20 m/s and greater than 4 m/s.

Acknowledgments. The research in this paper was partially supported by Iranian National Science Foundation: INSF under grant no. 96000458. We would like to thank the BGTA Editorial Board and the anonymous reviewers for their very helpful suggestions. The authors also thank prof. Pantelis Soupios for helpful conversations.

References

- Abido M.; 2001: *Particle swarm optimization for multimachine power system stabilizer design*. In: Proc. IEEE Power Engineering Society Transmission and Distribution Conference, vol. 3, pp. 1346-1351, doi:10.1109/PSS.2001.970272.
- Aster R.C., Borchers B. and Thurber C.H.; 2013: *Parameter estimation and inverse problems*. Elsevier, Amsterdam, The Netherlands, 363 pp.
- Awange J.L., Paláncz B., Lewis R.H. and Völgyesi L.; 2018: *Particle swarm optimization*. In: Mathematical Geosciences, Springer, Cham, Switzerland, pp. 167-184.
- Bregman N.D., Bailey R.C. and Chapman C.H.; 1989: *Crosshole seismic tomography*. Geophys., **54**, 200-215, doi:10.1190/1.1442644.
- Butchibabu B., Sandeep N., Sivaram Y.V., Jha P.C. and Khan P.K.; 2017: *Bridge pier foundation evaluation using cross-hole seismic tomographic imaging*. J. Appl. Geophys., **144**, 104-114, doi:10.1016/j.jappgeo.2017.07.008.
- Cardarelli E.; 2003: *Ray Tracing applied to travel time seismic tomography (theory and examples)*. Boll. Geof. Teor. Appl., **44**, 281-305.
- Cardarelli E.; 2008: *Modelling-2D complex anomalies by seismic tomography, (detection and delineation of anomalies with sharp boundaries)*. Boll. Geof. Teor. Appl., **49**, 265-277.
- Chunduru R.K., Sen M.K. and Stoffa P.L.; 1997: *Hybrid optimization methods for geophysical inversion*. Geophys., **62**, 1196-1207, doi:10.1190/1.1444220.
- Dai Q., Lin F., Wang X., Feng D. and Bayless R.C.; 2017: *Detection of concrete dam leakage using an integrated geophysical technique based on flow-field fitting method*. J. Appl. Geophys., **140**, 168-176, doi:10.1016/j.jappgeo.2017.03.013.
- Eberhart R.C. and Shi Y.; 2000: *Comparing inertia weights and constriction factors in particle swarm optimization*. In: Proc. Congress Evolutionary Computation, IEEE, La Jolla, CA, USA, pp. 84-88, doi:10.1109/CEC.2000.870279.
- Galuzzi B.G., Tognarelli A. and Stucchi E.M.; 2018: *A global-local experience of 2D acoustic FWI on a real data set*. In: Extended Abstract 80th EAGE Conference and Exhibition, Copenhagen, Denmark, doi:10.3997/2214-4609.201800887.
- Ghanati R., Ghari H. and Fatehi M.; 2017: *Regularized nonlinear inversion of magnetic anomalies of simple geometric models using Occam's method: an application to the Morvarid iron-apatite deposit in Iran*. Acta Geod. Geophys., **52**, 1-26, doi:10.1007/s40328-017-0193-9.
- Gheymasi H.M., Gholami A., Siahkoobi H.R. and Amini N.; 2016: *Robust total-variation based geophysical inversion using split Bregman and proximity operators*. J. Appl. Geophys., **132**, 242-254, doi:10.1016/j.jappgeo.2016.06.011.
- Göktürkler G.; 2011: *A hybrid approach for tomographic inversion of crosshole seismic first-arrival times*. J. Geophys. Eng., **8**, 99-108, doi:10.1088/1742-2132/8/1/012.

- Greenhalgh S., Zhou B. and Cao S.; 2003: *A crosswell seismic experiment for nickel sulphide exploration*. J. Appl. Geophys., **53**, 77-89, doi:10.1016/S0926-9851(03)00029-6.
- Guo Q., Zhang H., Tian J., Liang L. and Shang Z.; 2018: *A nonlinear multiparameter prestack seismic inversion method based on hybrid optimization approach*. Arabian J. Geosci., **11**, doi:10.1007/s12517-018-3392-y.
- Hansen P.C.; 1992: *Analysis of discrete ill-posed problems by means of the L-curve*. SIAM Rev., **34**, 561-580.
- Hedeshi N.G. and Saniee Abadeh M.; 2014: *Coronary artery disease detection using a fuzzy-boosting PSO approach*. Comput. Intell. Neurosci., **2014**, ID783734, doi:10.1155/2014/783734.
- Jiang M., Luo Y.P. and Yang S.Y.; 2007: *Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm*. Inf. Process. Lett., **102**, 8-16, doi:10.1016/j.ipl.2006.10.005.
- Kennedy J. and Eberhart R.C.; 1995: *Particle swarm optimization*. In: Proc. IEEE International Joint Conference on Neural Networks, Perth, Australia, vol. 4, pp. 1942-1948.
- Kiranyaz S., Ince T. and Gabbouj M.; 2014: *Multidimensional particle swarm optimization for machine learning and pattern recognition*. Springer-Verlag, Berlin Heidelberg, Germany, 349 pp., doi:10.1007/978-3-642-37846-1.
- Kroger B., Fechner T. and Kemna A.; 2011: *Cross-hole measurement and analysis of interfacial seismoelectric signals from shallow sedimentary boundaries*. In: Extended Abstract 73rd EAGE Conference and Exhibition incorporating SPE EUROPEC, doi:10.3997/2214-4609.20149313.
- Liu S., Liang M. and Hu X.; 2018: *Particle swarm optimization inversion of magnetic data: field examples from iron ore deposits in China*. Geophys., **83**, 1-62, doi:10.1190/geo2017-0456.1.
- Mazzotti A., Bienati N., Stucchi E., Tognarelli A., Aleardi M. and Sajeve A.; 2016: *Two-grid genetic algorithm full-waveform inversion*. The Leading Edge, **35**, 1068-1075, doi:10.1190/tle35121068.1.
- Mendes M.; 2009: *A hybrid fast algorithm for first arrivals tomography*. Geophys. Prospect., **57**, 803-809, doi:10.1111/j.1365-2478.2008.00755.x.
- Meng X., Wang S., Tang G., Li J. and Sun C.; 2017: *Stochastic parameter estimation of heterogeneity from crosswell seismic data based on the Monte Carlo radiative transfer theory*. J. Geophys. Eng., **14**, 621-632, doi:10.1088/1742-2140/aa6130.
- Menke W.; 2012: *Geophysical data analysis: discrete inverse theory, 3rd ed., MATLAB edition*. Academic Press, Elsevier Inc., Amsterdam, the Netherlands, 330 pp.
- Moret G.J., Knoll M.D., Barrash W. and Clement W.P.; 2006: *Investigating the stratigraphy of an alluvial aquifer using crosswell seismic traveltime tomography*. Geophys., **71**, B63-B73, doi:10.1190/1.2195487.
- Noack M.M. and Funke S.W.; 2017: *Hybrid genetic deflated Newton method for global optimization*. J. Comput. Appl. Math., **325**, 97-112, doi:10.1016/j.cam.2017.04.047.
- Ozcan E. and Mohan C.K.; 1999: *Particle swarm optimization: surfing the waves*. In: Proc. Congress Evolutionary Computation, Washington, DC, USA, pp. 1939-1944, doi:10.1016/j.cam.2017.04.047.
- Paasche H. and Tronicke J.; 2014: *Nonlinear joint inversion of tomographic data using swarm intelligence*. Geophys., **79**, R133-R149, doi:10.1190/geo2013-0423.1.
- Qian W. and Li. M.; 2017: *Convergence analysis of standard particle swarm optimization algorithm and its improvement*. Soft Comput., **22**, 1-24, doi:10.1007/s00500-017-2615-6.
- Rezaee Jordehi A. and Jasni J.; 2013: *Parameter selection in particle swarm optimization: a survey*. J. Exp. Theor. Artif. Intell., **25**, 527-542, doi:10.1080/0952813X.2013.782348.
- Rossi G., Sardu G., Del Negro E., Sorgo D., Nieto D. and Podda P.; 2009: *Seismic tomography in Monte Sinni (Sardinia, Italy) coal mine*. Boll. Geof. Teor. Appl., **51**, 223-238.
- Sajeve A., Aleardi M. and Mazzotti A.; 2017: *Genetic algorithm full-waveform inversion: uncertainty estimation and validation of the results*. Boll. Geof. Teor. Appl., **58**, 395-414, doi:10.4430/bgta0199.
- Sayed A.H.; 2014: *Adaptation, learning, and optimization over networks*. Found. Trends in Mach. Learn., **7**, 311-801, doi:10.1561/22000000051.
- Sen M.K. and Stoffa P.L.; 2013: *Global optimization methods in geophysical inversion*. Cambridge University Press, Cambridge, UK, 302 pp.
- Shi Y. and Eberhart R.C.; 1998: *Parameter selection in particle swarm optimization*. In: Porto V.W., Saravanan N., Waagen D. and Eiben A.E. (eds), Evolutionary Programming VII. Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg, Germany, vol.1447, pp. 591-600, doi:10.1007/BFb0040810.
- Shi Y. and Eberhart R.C.; 1999: *Empirical study of particle swarm optimization*. In: Proc. Congress Evolutionary Computation, Washington, DC, USA, pp. 1945-1950, doi:10.1109/CEC.1999.785511.

- Singh V.P., Radha T. and Pant M.; 2007: *Particle swarm optimization using Gaussian inertia weight*. In: Proc. Conference Computational Intelligence and Multimedia Applications, Sivakasi, Tamil Nadu, India, pp. 97-102, doi:10.1109/ICCIMA.2007.96.
- Song X., Tang L., Lv X., Iv X., Fang H. and Gu H.; 2012: *Application of particle swarm optimization to interpret Rayleigh wave dispersion curves*. J. Appl. Geophys., **84**, 1-13, doi:10.1016/j.jappgeo.2012.05.011.
- Soupios P., Akca I., Mpogiatzis P., Basokur A.T. and Papazachos C.; 2011: *Applications of hybrid genetic algorithms in seismic tomography*. J. Appl. Geophys., **75**, 479-489, doi:10.1016/j.jappgeo.2011.08.005.
- Suganthan P.N.; 1999: *Particle swarm optimiser with neighbourhood operator*. In: Proc. Congress Evolutionary Computation, Washington, DC, USA, pp. 1958-1962, doi:10.1109/CEC.1999.785514.
- Tarantola A.; 1987: *Inverse problem theory*. Elsevier, Amsterdam, The Netherlands, 358 pp.
- Tikhonov A.N.; 1963: *Solution of incorrectly formulated problems and the regularization method*. Sov. Math. Dokl., **4**, 1035-1038.
- Tognarelli A., Aleardi M. and Mazzotti A.; 2016: *Estimation of a high resolution P-wave velocity model of the seabed layers by means of global and a gradient-based FWI*. In: Extended Abstract, Near Surface Geoscience - Second Applied Shallow Marine Geophysics Conference, Barcelona, Spain, pp. 26-30, doi:10.3997/2214-4609.201602146.
- Trelea I.C.; 2003: *The particle swarm optimization algorithm: convergence analysis and parameter selection*. Inf. Process. Lett., **85**, 317-325, doi:10.1016/S0020-0190(02)00447-7.
- Troncke J., Paasche H. and Böniger U.; 2012: *Crosshole travelttime tomography using particle swarm optimization: A near-surface field example*. Geophys., **77**, R19-R32, doi:10.1190/geo2010-0411.1.
- Wong J.; 2000: *Crosshole seismic imaging for sulfide orebody delineation near Sudbury Ontario Canada*. Geophys., **65**, 1900-1907, doi:10.1190/1.1444874.
- Yao Z.S., Roberts R.G. and Tryggvason A.; 1999: *Calculating resolution and covariance matrices for seismic tomography with the LSQR method*. Geophys. J. Int., **138**, 886-894, doi:10.1046/j.1365-246x.1999.00925.x.
- Yue-lin G. and Yu-hong D.; 2007: *A new particle swarm optimization algorithm with random inertia weight and evolution strategy*. In: Proc. International Conference Computational Intelligence and Security Workshops, Harbin, China, pp. 199-203, doi:10.1109/CISW.2007.4425479.
- Zelt C.A. and Barton P.J.; 1998: *Three dimensional seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin*. J. Geophys. Res., Solid Earth, **103**, 7187-7210, doi:10.1029/97JB03536.
- Zhe Y. and Hanming G.; 2013: *Non-linear prestack seismic inversion with global optimization using an edge-preserving smoothing filter*. Geophys. Prospect., **61**, 747-760, doi:10.1111/1365-2478.12001.

Corresponding author: Mohammad Ali Riahi
Institute of Geophysics, University of Tehran
North Kargar, Tehran, Iran.
Phone: +98 216 1118219; e-mail: mariahi@ut.ac.ir