Non-linear stochastic inversion of 2D gravity data using evolution strategy (ES)

Khadije Ghasemi[1], Seyed-Hani Motavalli-Anbaran[2*], Gilda Karimi[3]

[1] *M. Sc. Student, Institute of Geophysics, University of Tehran, Tehran, Iran*
[2] *Assistant Professor, Institute of Geophysics, University of Tehran, Tehran, Iran.*
[3] *Assistant Professor, Department of Biological Sciences, Kharazmi University, Tehran, Iran*

(Received: 06 November 2017, Accepted: 08 January 2018)

Abstract

In the current work, a 2D non-linear inverse problem of gravity data is solved using the evolution strategies (ES) to find the thickness of a sedimentary layer in a deep-water situation where a thick sedimentary layer usually exists. Such problems are widely encountered in the early stages of petroleum explorations where potential field data are used to find an initial estimate of the basin geometry. However, the gravity data are the non-unique problem, and classical deterministic inversions can only offer a single approximation of the solution. Conversely, ES and evolutionary algorithms in general due to their random nature can offer a range of solutions that all fit the data within the acceptable threshold. The inverse problem is formulated as a single objective unconstrained numerical optimization. In evolutionary algorithms, the random nature of the search follows the same rules of the Darwinian biological evolution. Hence, the search is not exhaustive and requires less computational resources compared to Monte Carlo methods. Herein, first, the algorithm is introduced, and a classical synthetic problem is formed and successfully solved in the presence of white Gaussian noise. Then, a two-layered synthetic oceanic crust is formed. ES is successfully tested for solving the formulated under-determined inverse problem of estimating both the base of the sedimentary layer and the crust (i.e., Moho boundary). Finally, using the proposed method, the thickness of the sedimentary layer of the Caspian basin is found along an E-W profile crossing the Caspian Sea. The results have a good agreement with the previous estimates by deep seismic sounding method. The proposed method could be of particular interest because, in deep-water situations, the high water content of sediments, and the expected large thickness of the sediments among other factors make the use of reflective seismic methods unfeasible.

Keywords: non-linear inversion of gravity, evolution strategy, Caspian Sea, deep-water sediments

[*]Corresponding author: Motavalli@ut.ac.ir

1 Introduction

The non-linear inverse gravity problem is a classical non-linear inversion with more than several decades of history (Silva et al., 2014). The problem consists of estimating the depth of an arbitrary sub-surface layer from the gravity data (i.e., Bott, 1965). It is usually assumed that the density of earth material increases with depth (Hinze et al., 2013). Assuming that the layers are perfectly lied on each other with homogeneous boundaries, that is, the thicknesses of each layer is constant, then the gravity anomalies would not show any features. However, in reality, such situation is not true and structural features as well as tectonic mechanisms alter the layers, hence making it possible to use gravity data. In finding the thickness of the sedimentary layer or the boundary of the basement, it is assumed that the underlying crystalline crust is composed of denser material with respect to the sediments/sedimentary rocks. Therefore, a negative Bouguer anomaly is indicative of variations in the thickness of sediments. In other words, the variations of the Bouguer plate can be indicative of the thickness of the sediments. Classical deterministic optimization methods are historically applied to solve the non-linear gravity inverse problem (e.g., Zeyen and Pous, 1993, Čuma et al., 2012). These local search methods use a starting point (i.e., the thickness of the Bouguer plate) and improve it slightly to find the solution. The improvement of the solution in each iteration involves computing the derivatives of the forward problem with respect to each model parameter, which in case of non-linear problems, means computing a Jacobian matrix (Silva et al., 2014). Consequently, in problems with a large number of model parameters, the optimization process becomes highly expensive in terms of memory consumption (Jamasb et al., 2017). Such inverse problem is an ill-posed one due to the non-uniqueness of the gravity problem, and occasionally under-determination of the problem, i.e., when the number of the model parameters is greater than the number of observables (Zhdanov, 2002). However, classical deterministic optimization methods are inherently faster than other methods.

One of the main objectives of geophysical inversion is to find a model with respect to the geophysical observation. Herein, the geophysical inversion means finding an optimal value for a function with several variables. The function that should be minimized, i.e., cost function, shows the difference (or similarity) between the observed and the calculated data. Both methods of local and global optimizations are used to estimate the characteristics of the model with respect to geophysical data (e.g., Sen and Stoffa, 2013). In local deterministic search methods, if the starting point of the solution is in the vicinity of a local minimum, the algorithm will likely be stuck in that local minimum since the perturbation of the solution at each iteration is based on the derivative of the cost function with respect to the model parameters. To the contrary, global optimization methods, including ES, use the cost function itself. Most of global optimization algorithms inherently are stochastic and use the global information to update their current position. There are successful methods in global optimization based on the stochastic component that can escape from local optimums and overcome on premature stagnation (e.g., simulated annealing methods: Mojica and Bassrei, 2015). A famous class of global optimization methods are evolutionary algorithm, among which, the evolution strategy (ES) is profoundly successful in real-valued numerical optimizations (Bäck and Schwefel, 1993). An advantage of the global optimization methods, including ES, is that they can be successful when there is not an analytical function for the derivative of the cost

function or it is costly to calculate the derivatives (Sen and Stoffa, 2013).

There are three steps in the interpretation of geophysical data (e.g., Menke, 1989):

- Parameterization of model: finding the minimal set of model parameters that determine the properties of the model.
- Forward modeling: finding the physical rules that prepare the tool to compute the theoretical data for model parameters
- Inversion modeling: reconstruction of model parameters from the measured set (i.e., observations).

Inversion is defined as an activity that seeks to develop the model parameters from the observed data. Indeed, one can never reconstruct the true model from the data; in fact, the reconstructed model is an approximation that provides the best fit of calculated data to the observed ones. Moreover, the inverse problem is usually an ill-posed one due to the following reasons (Zhdanov, 2002):

1) Because of the non-uniqueness of physical process, several models fit the data. For example, gravity is measured on the surface of the earth to estimate the mass distribution within the earth. According to the Gaussian theorem, various density distributions inside the sphere can produce similar gravity fields on the surface. Accordingly, by just using gravity data, one cannot determine the mass/density distribution inside the earth. Therefore, additional information in the form of a priori assumptions or constraints, e.g., from seismic observations, are necessary to find a meaningful and realistic model.

2) The real model is usually a continuous function of spatial coordinates, and thus has an infinite degree of freedom. In real world, the inversion problems have limited amount of data. This makes an underdetermined inverse.

3) Observation is always corrupted with noise, which makes it impossible to find a true model, because when the calculated data completely fit the observed ones, the noise is also being fitted.

Global optimization search methods are especially suitable for cost functions with several local minima. The most popular ones are methods that use a random generator at every step, which can be divided into guided and non-guided classes. Non-guided methods are usually referred to as Monte Carlo methods, in which the main idea is to randomly perform an exhaustive search in the model space to find the best solution. Because of the heavy calculation burden, they are most suitable for solving (non-linear) problems with a small number of parameters (Sambridge and Mosegaard, 2002). The second group of random methods is guided methods the most popular of which are the neural network (Krasnopolsky and Schiller, 2003) and evolutionary algorithms (Back, 1996). The focus is on the methods that optimize parameters by using evolution strategy, an algorithm that belongs to the latter class. These methods utilize some biological evolutionary principles to find the best solution for the certain problem. In doing so, they manage to look only at some parts of the search space intuitively, which immensely reduces the computational costs (Snopek, 2005)

2 ES theory

Herein, evolution is used in a purely biological context wherein a system changes by the selection and the variation from one generation to the next one. It is the Darwinian concept about evolutionary changes that have become the main idea for evolutionary algorithm.

From the conventional view, the evolutionary algorithm is an algorithm that simulates Darwinian evolution system that is as follows:

1. One or several populations are used for limited sources.

2. These populations change dynamically by birth and death.

3. A concept of fitness that reflects the ability of individual for survival and reproduction.

4. A concept of various reproduction: offspring is too similar to their parents, but they are not same.

In short, the principles of Darwinian evolution indicate that the fitness of species has improved on average over the generations. A simulation of the evolution based on a candidate solution set that optimizes fitness by cost function improves the fitness in average and thus guides the simulated population to the solution (Blum et al., 2012).

Evolution algorithms could be divided into three main types of search:

- Evolution strategy (Rechenberg, 1989; Schwefel, 1995)
- Genetic algorithm (Goldberg, 1989)
- Evolution programming (Fogel et al., 1991)

They are based on three mechanisms that translate to evolution operators:

- Recombination
- Mutation
- Selection

There is a fundamental difference between ES and GA (genetic algorithm) in the type of operators and the presentation of the population. In ES, real values are used to display the optimization parameters instead of the binary string. In addition, ES serves mutation operator as the main operator, contrary to GA in which the recombination serves as the main operator.

The algorithm consists of several basic cycles. On the one hand, the selection operator attempts to reduce the diversity of the population, but on the other hand the mutation operator and recombination (variation) operator try to increase the diversity of the population. This fact leads to convergence rate and quality of the solution.

Algorithm 1 Evolutionary Algorithm

1: initialize solutions $\mathbf{x}_1, \ldots, \mathbf{x}_\mu \in \mathcal{P}$
2: **repeat**
3: **for** $i = 1$ **to** λ **do**
4: select ρ parents from $\mathcal{P}$
5: create $\mathbf{x}_i$ by recombination
6: mutate $\mathbf{x}_i$
7: evaluate $\mathbf{x}_i \rightarrow f(\mathbf{x}_i)$
8: add $\mathbf{x}_i$ to $\mathcal{P}'$
9: **end for**
10: select μ parents from $\mathcal{P}' \rightarrow \mathcal{P}$
11: **until** termination condition

Figure 1. Pseudocode of a general evolutionary algorithm.

Algorithm 1 (Figure 1) demonstrates the pseudocode of a general evolutionary algorithm.

The optimization process consists of three main steps:

1) ρ parents are selected by the recombination operator of which λ new solutions are combined.

2) Random changes (e.g., noise) are added to the preliminary candidate solution by the mutation operator. The goodness of the solutions in terms of minimizing the cost function is called fitness. The fitness of the new offspring solution is therefore evaluated using the cost function f. An offspring population P′ is formed by putting together all the individuals of a generation.

3) Finally, μ individuals are selected which constitute the novel parental population P of the next generation.

The process is repeated until a satisfying criterion is achieved. Typical satisfying criteria are defined using the value of cost function or via a maximum number of generations.

In the following, a short survey of evolutionary operators are given and variants of evolution strategies that have been successfully utilized in numerical optimizations are outlined (Kramer, 2014).

Within the context of ES algorithm, a set of individuals compete for survival in each generation (i.e., iteration). Each individual is in fact considered as a point in the search space which is of dimension M, where M is the number of model parameters (i.e., the dimension of model space). In other words, each individual is an Mx1 vector which could be a potential solution to the inverse problem. Selection operator in this algorithm is used in two stages. Parent Selection: when individuals are chosen as parents to participate in the crossover. In ES, this selection is not biased by fitness values and parents are drawn randomly with uniform distribution from the population of individuals. It means that everyone is selected with the equal probability to create the next generation. The next time the selection operator is used, the Survivor Selection is called. After the creation of the offspring (children) by mutation and recombination operators, survivors or the individuals for the next generation are selected. There are two approaches: (μ, λ) selection and (μ+λ) selection. In both variants, the selection is deterministic and rank-based. The individuals are ranked according to the fitness values, and the best individuals are selected for the next generation. In the first case, survivors are only selected from offspring that means the parents die in every generation even if they have better fitness values. In the latter case, survivors are selected from the union of parents and offspring. In other words, the parents do not die in any generation and can survive if they have a better fitness value. This form of selection is called elitist selection.

In the biological system, the recombination, also known as the cross over, combine the genetic material of two parents. Some of the evolutionary algorithms incorporate knowledge from two or more individuals using of recombination operator to create the offspring.

If recombination is used, the most common theme in ES is to combine two randomly selected parents to form one offspring. For the parent vectors x and y, one child z is created:

- intermediary recombination:
  $z_i = (x_i - y_i)/2$
- discrete recombination: $z_i = x_i \; or \; y_i$ chosen randomly

Discrete recombination is usually used for the variables representing candidate solution, and intermediary recombination is used for the search of strategy part. The procedure is repeated λ times in order to produce λ offspring.

Mutation is realized by adding a random number Δx_i to each of the values x_i:

$$x_i' = x_i + \Delta x_i \; . \qquad (1)$$

The delta terms are usually drawn from a zero-centered normal distribution N(0,σ) with mean μ=0 and standard deviation σ (Figure 2). In this case, probability density function is:

$$P(\Delta x_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(\Delta x_i - \mu)^2}{2\sigma}} \; . \qquad (2)$$

Regarding the mutation operator, there are two choices to make:

- Whether to use constant step size σ during the entire evolution process or to use a (self-adapted) variable step size.
- Whether to use the same step size σ for all variables x_i or to use an individual σ_i for each variable x_i.

Rechenberg (1989) has stated that ideally, of the mutations should be 'successful' (fitter than parents are) and it is used rule to adapt the step size or to control σ. This adapts step size σ every k generation with respect to measured

proportion ρ_s of successful mutation ($0.8 \le c \le 1$) (Beyer, 2013):

$$\sigma = \frac{\sigma}{c} \quad if\ p_s > 1/5$$
$$\sigma = \sigma \cdot c \quad if\ p_s < 1/5 . \qquad (3)$$
$$\sigma = \sigma \quad if\ p_s = 1/5$$

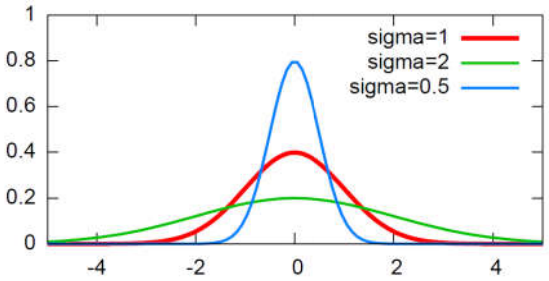


Figure 2. Normal distribution with mean μ=0 and standard deviation σ.

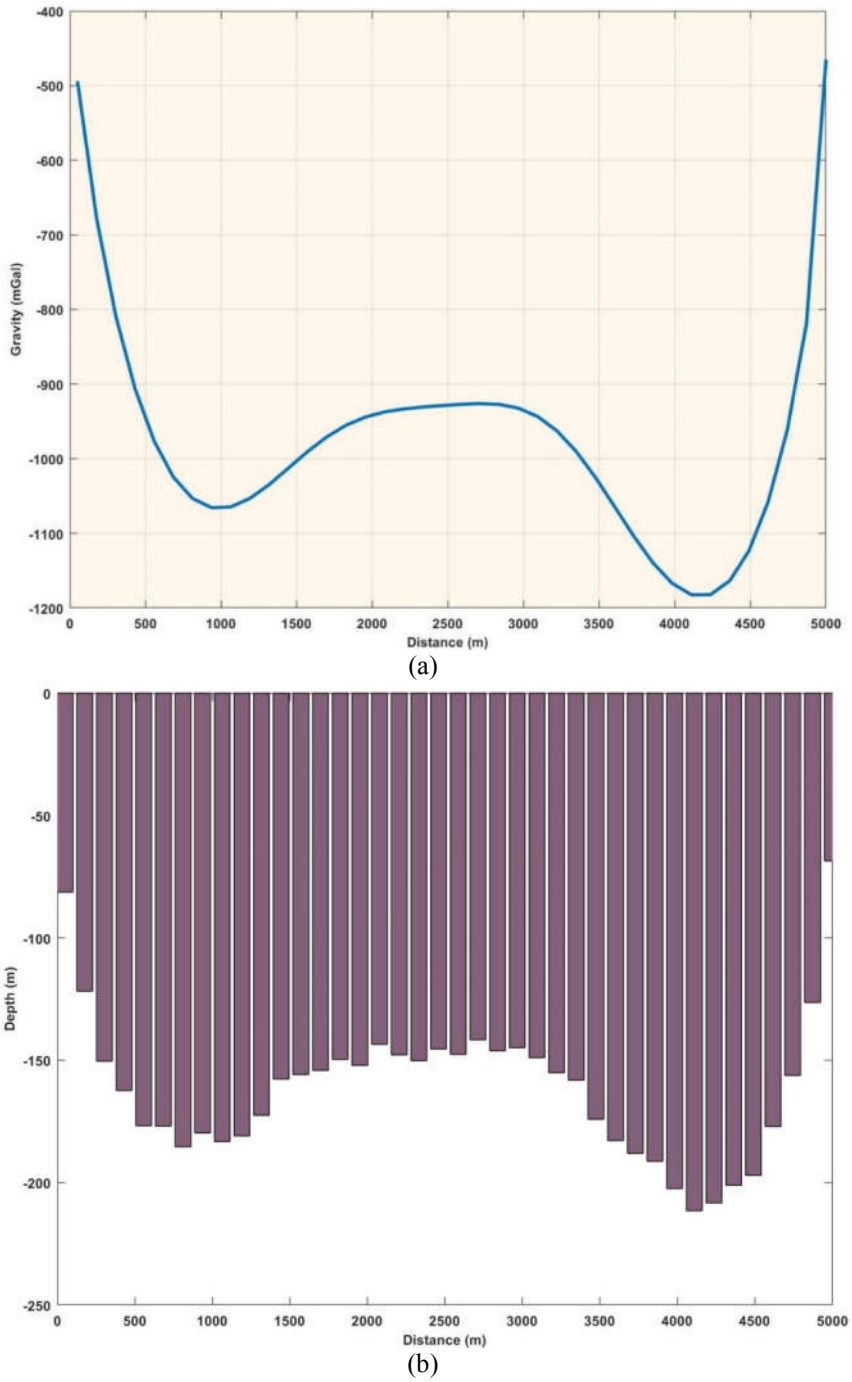


Figure 3. (a) Calculated gravity effect in the 50 observed points; (b) sedimentary basin modeling in 2D by prisms.

3 Forward problem

Several approximations are considered in the assumption of the density contrast for modeling purposes. In this paper, for the first synthetic model, it is supposed that density contrast is constant. In the synthetic two-layered oceanic crust model, a linear density gradient within both the sedimentary layer and the underlying crust is considered wherein the density contrast decreases with depth. At last, for the unconsolidated sediments of the South Caspian Basin, it is assumed that the vertical density gradient follows the hyperbolic law of Silva et al. (2014).

In the first case, the underground space is divided into a set of right rectangular prisms of constant density and variable thickness (Figure 3). Although, the numbers of prisms and observation points can be chosen independently, herein one prism is defined for each data point for the sake of simplicity. The gravity effect of each prism is calculated as (Nagy et al., 2000):

$$\begin{aligned}&\Delta g(\rho) = G\rho_0 \,|||\, x\ln\left(y+r\right)\\&+y\ln\left(x+r\right)-z\arctan(\frac{xy}{zr})\Big|_{x_1}^{x_2}\Big|_{y_1}^{y_2}\Big|_{z_1}^{z_2},\\&r=\sqrt{x^2+y^2+z^2}\end{aligned} \quad (4)$$

where, G is the gravitational constant. For the calculated data, the gravity effect of each prism is found using Equation (4) and then is summed for all prisms.

For the other cases, the modeling of the entire basin is done using polygons with an arbitrary number of vertices and variable density contrast.

In the latter case, the gravity anomaly can be computed by:

$$\Delta g_{zz} = G\sum_{i=1}^{n}\left(\begin{array}{l}2\left(A_z+\gamma_z\times z_{obs}\right)\left(\begin{array}{l}\dfrac{\left(tan^{-1}\left(\frac{z_i}{x_i}\right)-tan^{-1}\left(\frac{z_{i+1}}{x_{i+1}}\right)\right)\left(x_{i+1}-x_i\right)\left(x_iz_{i+1}-x_{i+1}z_i\right)}{\left(x_{i+1}-x_i\right)^2+\left(z_{i+1}-z_i\right)^2}+\\ \dfrac{1}{2}\dfrac{\left(x_iz_{i+1}-x_{i+1}z_i\right)\left(z_{i+1}-z_i\right)log\left(\frac{x_{i+1}^{\,2}+z_{i+1}^{\,2}}{x_i^{\,2}+z_i^{\,2}}\right)}{\left(x_{i+1}-x_i\right)^2+\left(z_{i+1}-z_i\right)^2}\end{array}\right)\\ -\gamma_z\dfrac{\left(\left(x_{i+1}-x_i\right)^2-\left(z_{i+1}-z_i\right)^2\right)\left(x_iz_{i+1}-x_{i+1}z_i\right)^2\left(tan^{-1}\left(\frac{z_i}{x_i}\right)-tan^{-1}\left(\frac{z_{i+1}}{x_{i+1}}\right)\right)}{\left(\left(x_{i+1}-x_i\right)^2+\left(z_{i+1}-z_i\right)^2\right)^2}\\ -\gamma_z\dfrac{\left(x_{i+1}-x_i\right)\left(z_{i+1}-z_i\right)\left(x_iz_{i+1}-x_{i+1}z_i\right)^2log\left(\frac{x_{i+1}^{\,2}+z_{i+1}^{\,2}}{x_i^{\,2}+z_i^{\,2}}\right)}{\left(\left(x_{i+1}-x_i\right)^2+\left(z_{i+1}-z_i\right)^2\right)^2}\\ +\gamma_z\dfrac{\left(z_{i+1}-z_i\right)^2\left(x_iz_{i+1}-x_{i+1}z_i\right)}{\left(x_{i+1}-x_i\right)^2+\left(z_{i+1}-z_i\right)^2}\end{array}\right), \quad (5)$$

$$A_z=\rho_{up}-\gamma_z\times z_{min} \quad , \quad A_x=\rho_{left}-\gamma_x\times x_{left}$$

where i and $i+1$ are the points of mass vertices, n is total number of vertices, G is the universal gravitational constant, z_{obs} is the height coordinate of observation point, x_{obs} is the horizontal coordinate of observation point, ρ_{down} and ρ_{up} are density of highest and lowest vertices of model respectively, ρ_{right} and ρ_{left} are the density of least and most horizontal coordinates of the model respectively, z_{down} and z_{up} are the least and most coordinate height of model vertices, and x_{right} and x_{left} are the least and most horizontal coordinate of model vertices.

4 Inverse problem

Solving the inverse problem is equivalent to find a set of depths that minimizes the cost function:

$$c^g(z) = \frac{1}{N}\left\| g^{obs} - \Delta g^*(z) \right\|_2^2 , \quad (6)$$

where Δg^{obs} is the observed anomaly vector and Δg^* is the calculated anomaly for the model z. As was mentioned, the gravity inverse problem expressed above is ill-posed. Therefore, determination of the solution by minimizing the misfit is too unstable (Pallero et al., 2015). In the case of inverting gravity data for depths, this instability causes the reconstructed models to have zigzag shapes. To reduce the ill-posed nature of the inversion problem and the set of solution to those that fit with a priori information, different types of constraints are usually used (e.g.,Martins et al., 2011; Barbosa et al., 1997). These constraints include utilizing smoothing conditions (relative) or determining the depth value by borehole information (absolute). For this purpose, it requires some regularization parameters to establish the balance between the terms of the cost function (prediction error, regularity, primary condition).

Relative constraints include imposing the smoothing condition to approximate decrease basement relief topography. Smoothing is introduced by regularization term (Snopek, 2005):

$$c^r(z) = \frac{1}{M-1}\sum_{j=1}^{M-1}(z_j - z_{j-1})^2 , \quad (7)$$

or more general:

$$c^r(z) = \frac{1}{L}\left\| Rz \right\|_2^2 , \quad (8)$$

where R is a $L*M$ weighting matrix for regularity condition that applies to the parameter model. The above expression consists of Tikhonov's regularization of order 1, 2 to impose the regularity condition on the first and the second order derivatives of the model.

The final cost function with the regularization term is:

$$c(z) = c^g(z) + \mu^r c^r(z) , \quad (9)$$

where μ^r is the regularization parameter that determines the balance between different terms in the cost function (Barbosa et al., 1997), which is set here by trial and error. The parameter is first set to 0 and then slightly is increased until the solution becomes stable (the zigzag shape disappears). However, there are some automatic methods to determine the regularization parameter, though they are mainly developed in the context of linear inverse problems (Vatankhah et al., 2015). Recently, there have been some attempts to extend the regularization parameter estimation methods to non-linear inverse problems (Mead and Hammerquist, 2013; Farquharson and Oldenburg, 2004).

5 Synthetic examples

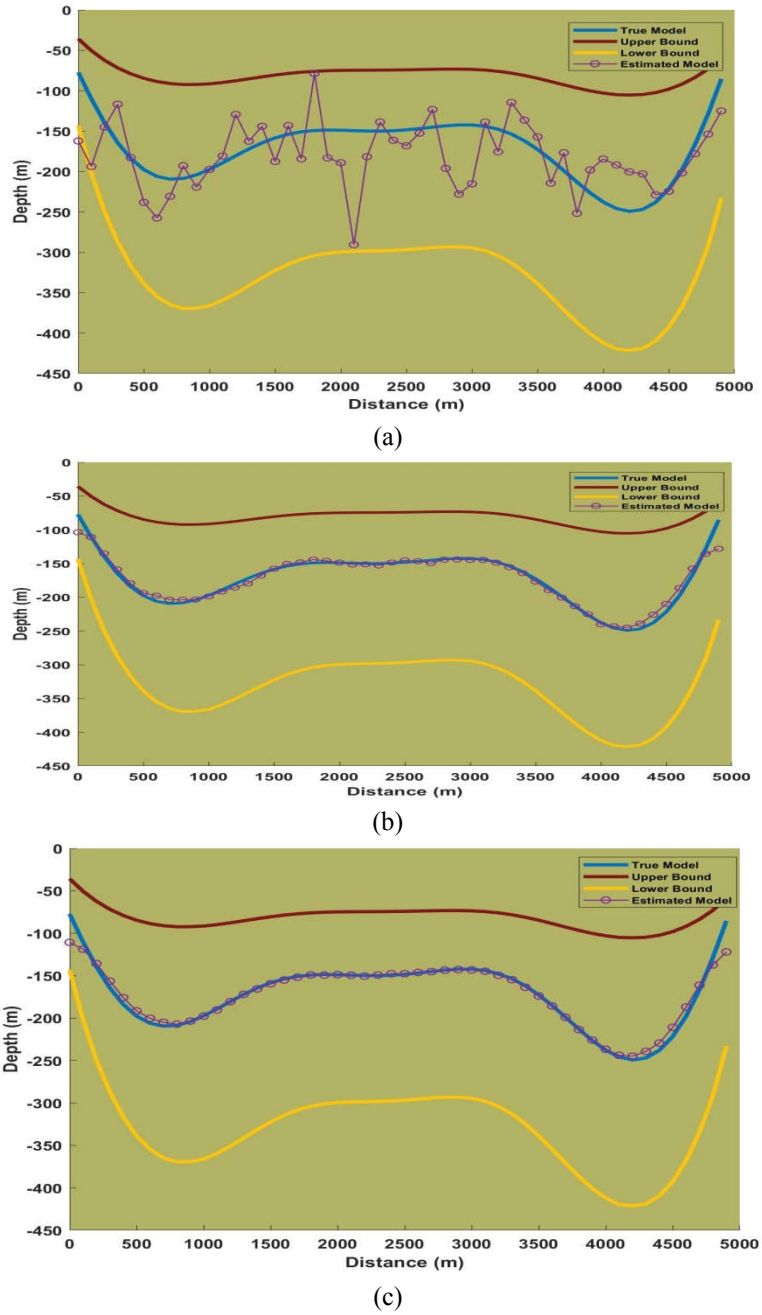
The first synthetic model is a 2D section of a sedimentary basin with a length of 5 km. The basin is discretized and modeled via 40 rectangular prisms with the known density. The maximum depth of prisms reaches 200 m, while the depth of the shallow prisms located on the left and right, is 50 m. The 40 observed points are placed with equal spacing along the

profile. The density contrast between the sediment and the basement is set as $\rho = 150\frac{kg}{m^3}$ for all prisms. The gravity effect in 2D resulting from the change in depth of prisms is calculated using Equation (5) at the observation points on the same height of prism's upper side. Figure 3 demonstrates the sedimentary basin modeling in 2D by prisms and their corresponding calculated gravity effects.

The result of inversion is illustrated in Figures 4 and 5. Upper and low bound of search space is generated by Equation (10) (Pallero et al., 2015):

$$z_j^0 = \frac{\Delta g_j^{obj}}{2\pi G \Delta\rho}, \qquad (10)$$

where z_j^0 is the initial depth for the j^{th} prism, Δg_j^{obj} is the observed gravity anomaly at the center of the prism, G is the universal gravitational constant, and $\Delta\rho$ is designated density contrast.

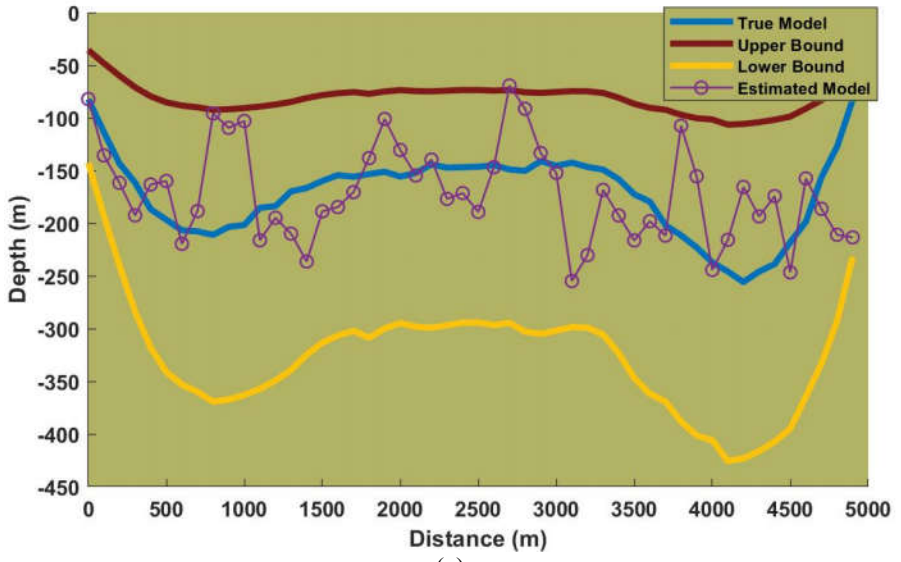


(a)

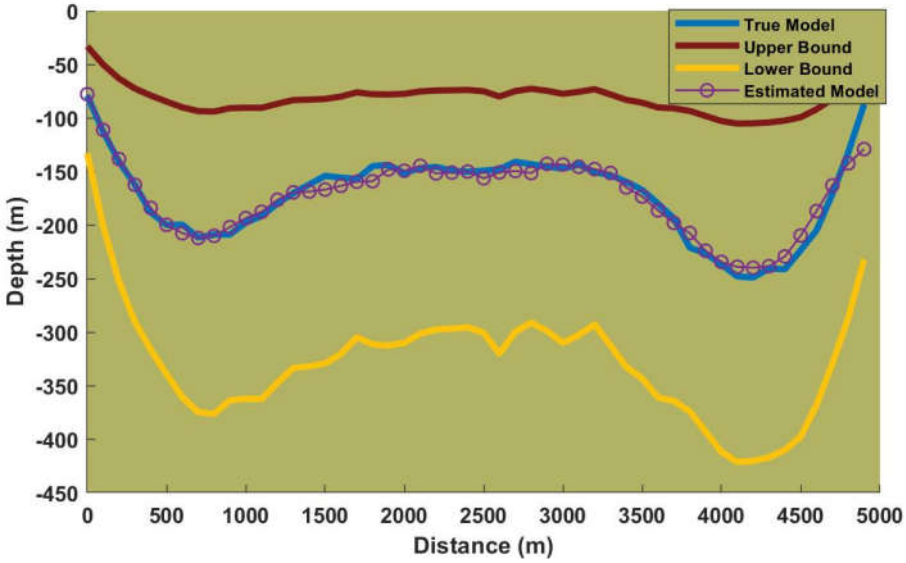
(b)

(c)

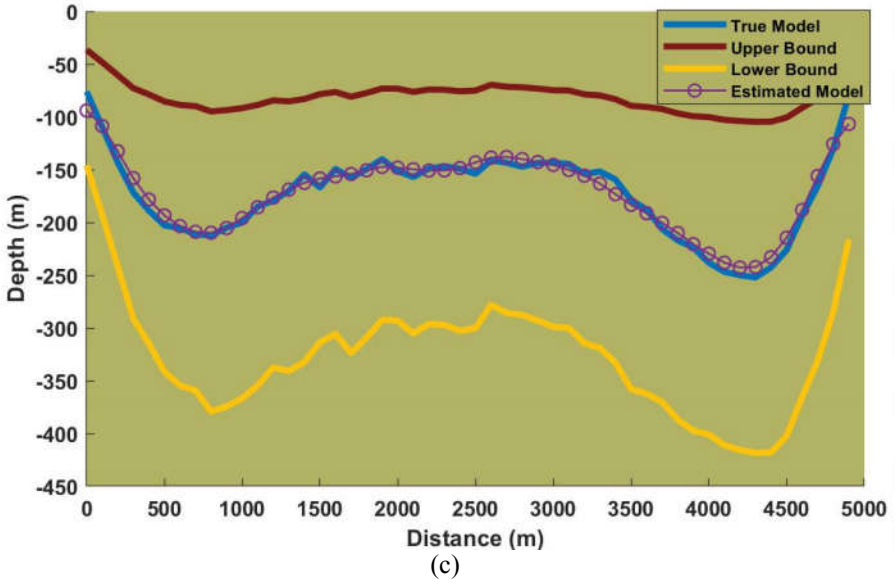
Figure 4. The result of data inversion without the presence of noise, (a) generation 10, (b) generation 70, (c) generation 150.



(a)



(b)



(c)

Figure 5. The result of data inversion with the presence of 3% White Gaussian noise, (a) generation 10, (b) generation 70, (c) generation 150.

Factors $k_{min} < 1.0$ and $k_{max} > 1.0$ are applied for approximate model z_j^0 in order to find the bounds of search space (Pallero et al., 2015).

$$z \in \left[k_{min} . z^0 \ , k_{max} . z^0 \right], \qquad (11)$$

where z^0 is of the same dimension as the main unknown of the inverse problem, z.

These bounds are demonstrated by crimson and yellow lines for upper and lower bounds, respectively. In this inversion, relative constraints, smoothing in norm 2 are used. The parent population is one individual, population of offspring is 30 and cost function is optimized over 150 generations. Offspring are produced by the mutation operator. The first generation is randomly selected from a normal distribution function with a width corresponding to the search space defined by the previous equations. Then, the parent and offspring cost function is compared based on final misfit at each generation, and the solution is optimized by choosing the best of them as next-generation parents. The optimization process consists of exploration and exploitation of solution with and without the presence of 3% White Gaussian noise at generation 10, 80 and 150 is illustrated in Figures 4 and 5, respectively.

The second synthetic model is a two-layered model consisting of a sedimentary layer and an underlying crust. This model has the similar structure to crust model of Makran zone in the southeast of Iran that is simulated by synthetic data with/without 3% White Gaussian noise (see Entezar-Saadatet al., 2017). Herein, a linear density gradient is assumed for the density contrasts within both the sedimentary layer and the underlying crust. The length of the profile is 750 km and the density contrast varies from -800 kg/m^3 to -500 kg/m^3 with increasing depth at the sedimentary layer. For the underlying layer, this value reaches to -250 kg/m^3 at the bottom. The forward problem is calculated by Equation (5) and the gravity effect of two-layered is shown in Figure 6.

The purpose of the inversion is to find the depths to both the base of the sedimentary layer and the crust-mantle boundary (Moho) simultaneously. The result of gravity inversion data with and without the presence of noise are displayed in Figures 7 and 8, respectively.

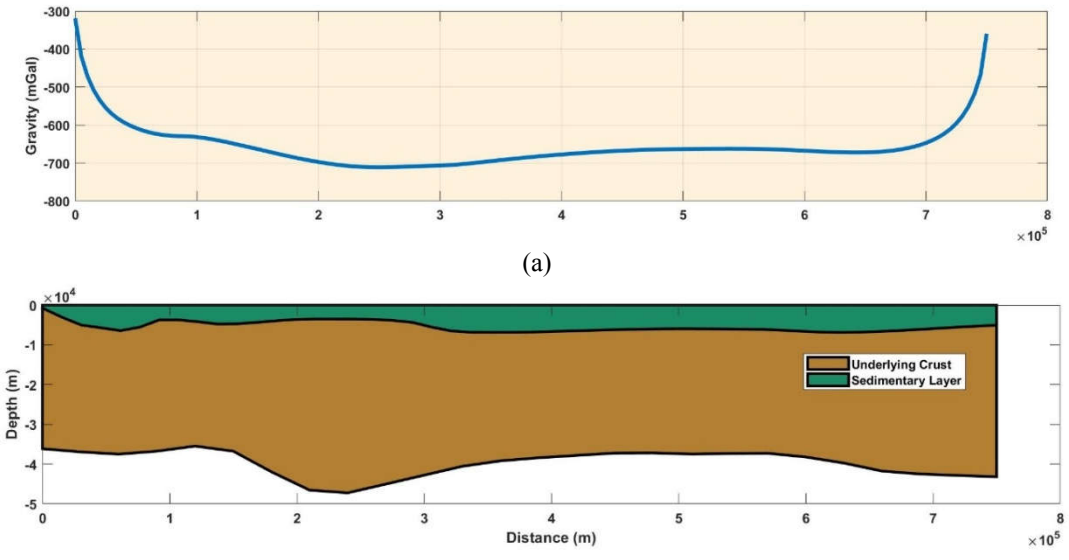


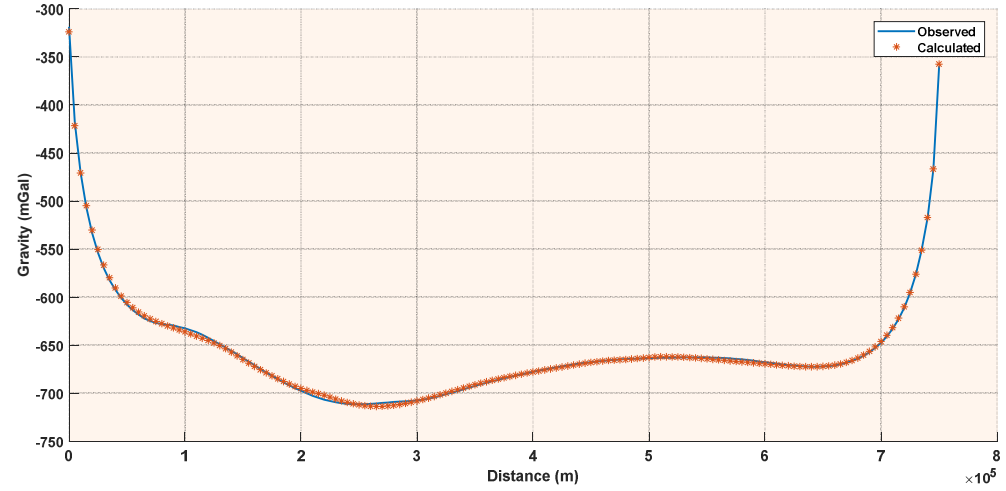
Figure 6. (a) The gravity effect of two layered crustal is calculated. (b) location of crust-mantle boundary and sedimentary basin.

6 Real example: South Caspian basin

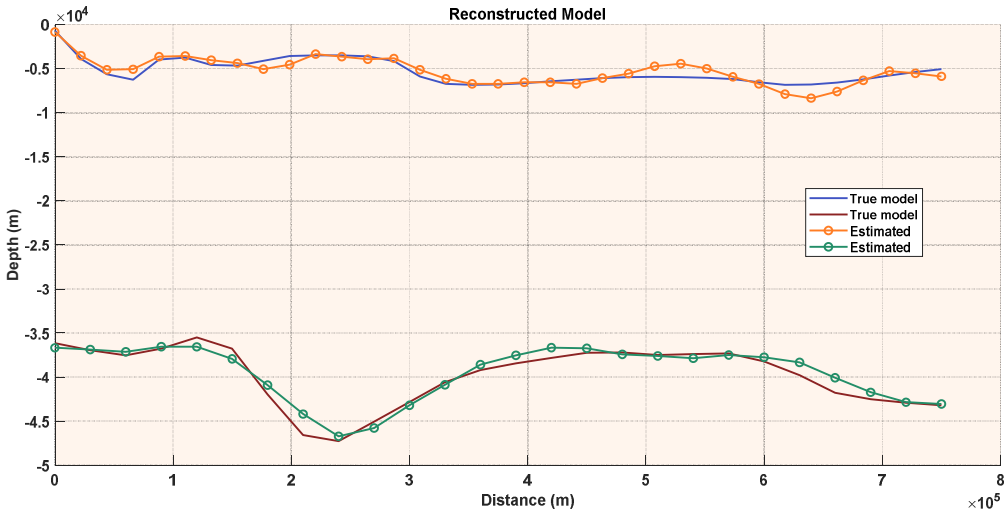
This section presents the application of ES to real gravity data in the South Caspian basin.

The most important properties of the crustal model of Jackson et al. (2002) are: (1) a 20 km variation in the thickness of the sedimentary section across the basin; (2) a lack of a 'granitic' ($v\rho$ = 5.8-6.5kms^{-1}) layer under the central part of the basin; and (3) a thinning of the crust under the basin by 15-20 km compared with its margins.

The distribution of observed points, stacked data, and the free air anomaly are presented in Figure 9. The free air data are extracted from BGI (http://bgi.omp.obs-mip.fr).

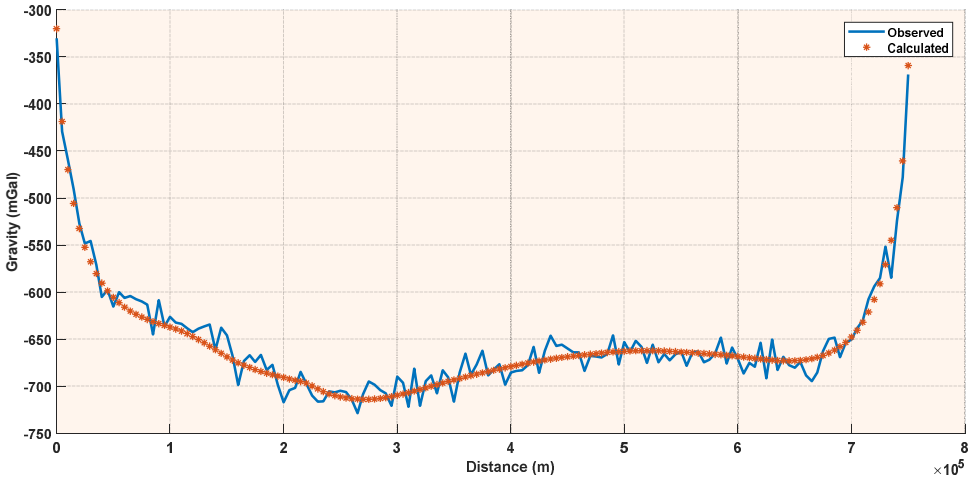


(a)

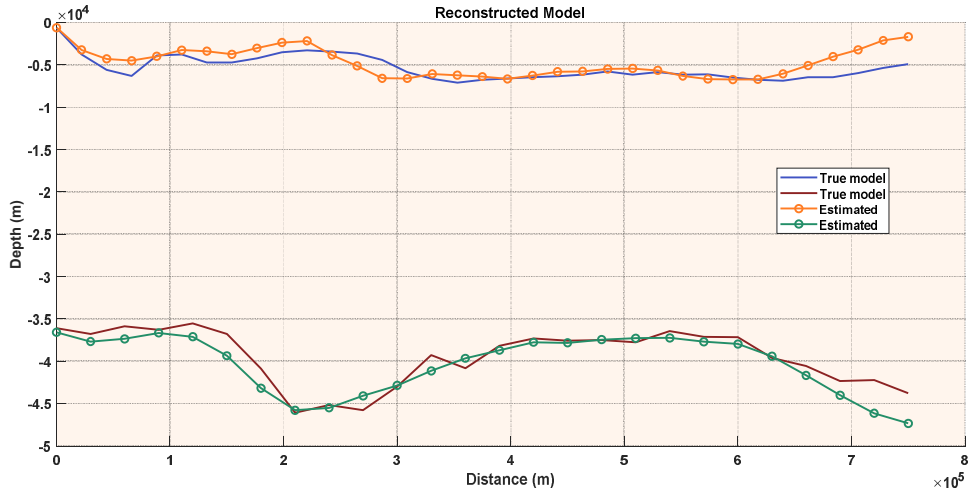


(b)

Figure 7. The result of data without noise. (a) observed gravity and calculated gravity. (b) comparison between real model and reconstructed model by inversion.



(a)



(b)

Figure 8. The result of data with noise. (a) observed gravity and calculated gravity. (b) comparison between real model and reconstructed model by inversion.

The study area is an E-W profile crossing the South Caspian Sea Basin with a length of 400 km and sampling space of 20 km. The profile is intentionally coinciding with the profile examined by Jackson et al. (2002) to check the ability of ES in solving the non-linear gravity inverse problem and compare its results with seismic methods. The profile extends from each side to a distance of 200 km to prevent the artificial edge effects. Because the study region is off-shore, free-air anomalies are used. The aim of the inversion is to acquire the thickness of the sediments. The evolution strategy parameters that are used are $\mu = 1$, $\lambda = 30$. The performance of 150 runs of the elitist variant of ES was investigated. As the evaluation function, the final misfit between the observed and calculated gravity field was used. In the inverse process, the density contrast at the surface is considered to be -1000 kg/m^3. It is also assumed that the density contrast decreases with depth in the form of a hyperbolic law (Silva et al., 2014):

$$\Delta\rho = \Delta\rho_0 \frac{\beta^2}{(z+\beta)^2}, \qquad (12)$$

where β is a factor controlling the decay of the density contrast with the depth and its value is considered to be 45 in the calculations to achieve an exponential

decay in the density contrast. Figure 10 shows the result of inversion and misfit between the observed and calculated data. Besides, the results of Jackson et al. (2002) are plotted for comparison with the estimated thickness.

According to Figure 10, the results match the results of deep seismic sounding. However, the obtained thickness for sediments is less than the results of Jackson et al. (2002) on the right side of the profile. One of the causing factors of this issue can be the data stacking process. It is well apparent in the data (Figure 9) that a decreasing N-S trend exists in the right side of the profile that would reduce the amplitude of the stacked data compared to its original value at the exact location coinciding with the one from Jackson et al. (2002). After data stacking, the data are averaged and therefore the estimated thickness of the sediments is an average along the N-S direction.

Based on the results presented, the evolution strategy algorithm is able to estimate the thickness of sediments well.

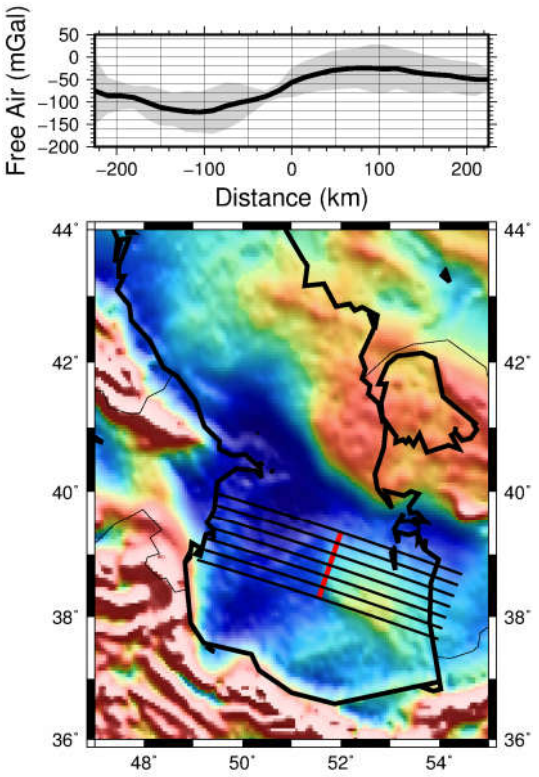


Figure 9. (Up) Stacked data shown with the dark line and the upper and lower bounds of each data shown with gray margin. (Down) distribution of observation points and data profile

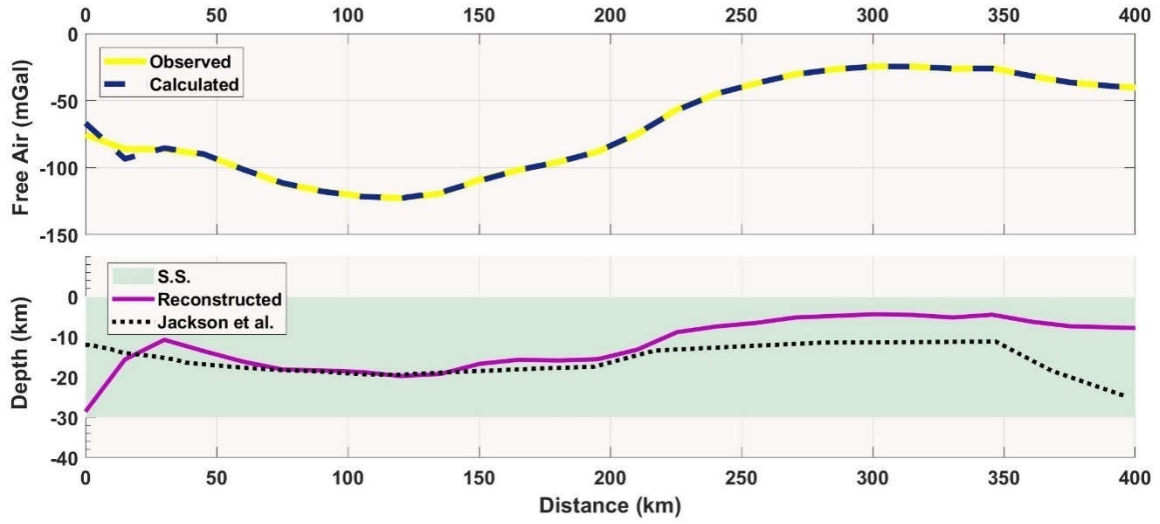


Figure 10. Result of inversion to estimate the depth of sediment layer in Caspian basin. (Up) fitness of calculated data and observation data. (Down) estimated thickness and comparison with Jackson's result.

7 Conclusions

Evolution Strategy is used in this study, which is an evolutionary global optimization search method to solve the non-linear gravity inverse. It should be noted that any algorithm based on the idea of the evolutionary strategy does not necessarily find an optimal solution for the problem. The advantage of these methods is that they find an acceptable solution at an acceptable speed. Another important feature of all these methods that work randomly is that one "never" obtains the results similar to the start model. It can be considered as an advantage of the evolutionary strategy. By running the inversion several times, one could find a wide range of possible solutions for the problem.

An algorithm based on the evolution strategy does not need a mathematical specific characteristic for cost function. Regarding non-uniqueness of the gravity problem, the random nature of this algorithm can be useful. It is possible for a few runs not to converge. This problem can be solved by performing repeated runs. However, on the one hand, the random nature of the algorithm allows for the possibility of exploring a wide range of possible solutions. On the other hand, it is almost necessary for any stochastic algorithm to repeat its runs anyways because the variability of the solutions should be known. Considering the simplicity and small computational cost of ES, this problem is not a severe one. The result of one-layer and two-layer synthetic models, with and without the presence of noise, could validate the accuracy and speed of ES algorithm as well as the possibility of its implementation in the gravity domain. In addition, by applying the method on the real data in South Caspian basin, its strength is confirmed through comparison with previous results of deep seismic sounding.

References

Back, T., 1996, Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms: Oxford University Press.

Bäck, T., and Schwefel, H. P., 1993, An overview of evolutionary algorithms for parameter optimization: Evolutionary computation, **1**(1), 1-23.

Barbosa, V. C. F., Silva, J. B., and Medeiros, W. E., 1997, Gravity inversion of basement relief using approximate equality constraints on depths: Geophysics, **62**(6), 1745-1757.

Beyer, H. G., 2013,The Theory of Evolution Strategies: Springer Science and Business Media.

Blum, C., Chiong, R., Clerc, M., De Jong, K., Michalewicz, Z., Neri, F., and Weise, T. , 2012, Evolutionary optimization, in Variants of Evolutionary Algorithms for Real-World Applications, edited, 1-29: Springer.

Bott, M., 1965, The upper mantle beneath Iceland: Geophysical Journal International, **9**(2-3), 275-277.

Čuma, M., Wilson, G. A., and Zhdanov, M. S. , 2012, Large-scale 3D inversion of potential field data: Geophysical Prospecting, **60**(6), 1186-1199.

Entezar-Saadat, V., Motavalli-Anbaran, S.H., and Zeyen, H., 2017, Lithospheric structure of the Eastern Iranian plateau from integrated geophysical modeling, A transect from Makran to the Turan platform: Journal of Asian Earth Sciences, **138**, 357-366.

Farquharson, C. G., and Oldenburg, D. W., 2004, A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems: Geophysical Journal International, **156**(3), 411-425.

Fogel, D. B., Fogel, L. J., and Atmar, J. W., 1991, Meta-evolutionary programming, paper presented at Signals, systems and computers: Conference record of the twenty-fifth asilomar conference on, IEEE.

Goldberg, D. E., 1989, Genetic algorithms in search, optimization, and machine learning. Reading: Addison-Wesley.

Hinze, W. J., Von Frese, R. R., and Saad, A. H., 2013, Gravity and magnetic exploration: Principles, practices, and applications, Cambridge University Press.

Jackson, J., Priestley, K., Allen, M., and Berberian, M. , 2002, Active tectonics of the south Caspian basin: Geophysical Journal International, **148**(2), 214-245.

Jamasb, A., Motavalli-Anbaran, S. H., and Zeyen, H., 2017, Non-linear stochastic inversion

of gravity data via quantum-behaved particle swarm optimisation: application to Eurasia–Arabia collision zone (Zagros, Iran): Geophysica Prospecting, **65**(S1), 274-294.

Kramer, O., 2014, A Brief Introduction to Continuous Evolutionary Optimization: Springer.

Krasnopolsky, V. M., and Schiller, H., 2003, Some neural network applications in environmental sciences. Part I: forward and inverse problems in geophysical remote measurements: Neural Networks, **16**(3), 321-334.

Martins, C. M., Lima, W. A., Barbosa, V. C., and Silva, J. B., 2011, Total variation regularization for depth-to-basement estimate, Part 1—Mathematical details and applications: Geophysics, **76**(1), I1-I12.

Mead, J., and Hammerquist C., 2013, χ^2 Tests for the Choice of the Regularization Parameter in Nonlinear Inverse Problems: SIAM Journal on Matrix Analysis and Applications, **34**(3), 1213-1230.

Menke, W., 1989, Discrete Inverse Theory.

Mojica, O., and Bassrei, A., 2015, Application of the Generalized Simulated Annealing Algorithm to the Solution of 2D Gravity Inversion of Basement Relief: 3rd Latin American Geosciences Student Conference.

Nagy, D., Papp, G., and Benedek, J., 2000, The gravitational potential and its derivatives for the prism: Journal of Geodesy, **74**(7-8), 552-560.

Pallero, J., Fernández-Martínez, J. L., Bonvalot, S., and Fudym, O., 2015, Gravity inversion and uncertainty assessment of basement relief via Particle Swarm Optimization: Journal of Applied Geophysics, **116**, 180-191.

Rechenberg, I., 1989, Evolution strategy: Nature's way of optimization, in Optimization: Methods and applications, possibilities and limitations, edited, 106-126: Springer.

Sambridge, M., and Mosegaard, K., 2002, Monte Carlo methods in geophysical inverse problems: Reviews of Geophysics, **40**(3).

Sen, M. K., and Stoffa P. L., 2013, Global Optimization Methods in Geophysical Inversion: Cambridge University Press.

Silva, J. B., Santos, D. F., and Gomes, K. P., 2014, Fast gravity inversion of basement relief: Geophysics, **79**(5), G79-G91.

Silva, J. B., Medeiros, W. E., and Barbosa, V. C., 2001, Potential-field inversion: Choosing the appropriate technique to solve a geologic problem: Geophysics, **66**(2), 511-520.

Silva, J. S., and Tenreyro, S., 2006, The log of gravity: The Review of Economics and statistics, **88**(4), 641-658.

Snopek, K., 2005, Inversion of gravity data with application to density modeling of the Hellenic subduction zone: Ph. D. Thesis, University of Bochum.

Vatankhah, S., Ardestani, V. E., and Renaut, R. A., 2015, Application of the χ2 principle and unbiased predictive risk estimator for determining the regularization parameter in 3D focusing gravity inversion: Geophysical Journal International, **200**(1), 265-277.

Zeyen, H. and Pous, J., 1993, 3D joint inversion of magnetic and gravimetric data with a priori information: Geophysical Journal International, **112**(2), 244-256.

Zhdanov, M. S., 2002, Geophysical Inverse Theory and Regularization Problems: Elsevier.