Anomaly shape inversion via model reduction and PSO

Z. Fernández-Muñiz, J.L.G. Pallero, J.L. Fernández-Martínez

ARTICLE INFO

KEYWORDS:
Gravimetry
Anomaly detection
Inversion
RR-PSO
Uncertainty analysis

ABSTRACT

Most of the geophysical inverse problems in geophysical exploration consist in detecting, locating and outlining the shape of geophysical anomalous bodies imbedded into a quasi-homogeneous background by analyzing their effect in the geophysical signature. The inversion algorithm that is currently used creates a very fine mesh in the model space to approximate the shapes and the values of the anomalous bodies and the geophysical structure of the geological background. This approach results in discrete inverse problems with a huge uncertainty space, and the common way of stabilizing the inversion consists in introducing a reference model (through prior information) to define the set of correctness of geophysical models. We present a different way of dealing with the high underdetermined character of this kind of problems, consisting in solving the inverse problem using a low dimensional parameterization that provides an approximate solution of the anomaly via Particle Swarm Optimization (PSO). This methodology has been designed for anomaly detection in geological set-ups that correspond with this kind of problem. We show its application to synthetic and real cases in gravimetric inversion performing at the same time uncertainty analysis of the solution. We have compared two different parameterizations for the geophysical anomalies (polygons and ellipses), showing that we have obtained similar results. This methodology outperforms the common least squares method with regularization.

1. Introduction

Geophysical inverse problems serve inform about the location and internal structure of the anomalous bodies taking into account their effect in the observed data. One the main challenges in inversion is to quantify the uncertainty of the solution, that is, the non-uniqueness of the solution (Snieder and Trampert, 1999; Scales and Snieder, 2000), which is due to different sources, such as noise in the measurements data, limited amount of data, physics limitations, modelling hypothesis, parameterization and numerical approximations among others. Deterministic and probabilistic approaches are used to deal with uncertainty (see for instance Menke, 2012; Aster et al., 2012; Tarantola and Valette, 1982). Geophysical prior models inferred by geological studies are also used to constraint the range of the plausible models and also to inform about the optimum design of the acquisition survey. A second difficulty comes from the noise contained in the observed data that deforms the shape of the elongated flat valleys where the equivalent geophysical models are located (Fernández-Martínez et al., 2012a). Depending on the type and the level of noise, the nonlinear equivalence region for a given error tolerance might be reduced in size, or even disappear, and the geophysical solution that is found by the optimization process will be shifted from the most plausible solution that would generate the observed data in absence of noise (Fernández-Martínez et al., 2014a,b). In real practice it is impossible to know the true solution of the inverse problem, but for sure it does not coincide with the solution that has been found. The way that the uncertainty problem is solved consists in sampling different equivalent models with a relative error (data mismatch) higher than the minimum mismatch. The noise has the effect of shifting the plausible models to higher misfit regions. This fact explains why to sample in broader zones of the equivalence region. Conversely, the classical way that the problem is solved is through local optimization methods with regularization to stabilize the inversion through the use of good-quality prior information: reference model and regularity conditions. This numerical method has some drawbacks if the reference model is incorrect or the regularity that it is imposed is not correct, leading the inversion to a different family of equivalent solutions that differ from the real geology. Besides, it has to be understood that the regularization techniques do not provoke the uncertainty of the solution to vanish, and this type of analysis has yet to be correctly performed (Fernández-Martínez et al., 2013). This is not usually the case in practice and the...
practitioners present solutions without their corresponding uncertainty assessment. Different solutions are proposed to deal with an incorrect choice of the regularity conditions. For instance, Zhdanov and Lin (2017) introduced an inversion method to exploit the sharp contrasts of the density between the host media and anomalous targets in the inversion of gravity data. Besides, joint inversions of different types of geophysical and petrophysical data are also used to reduce the uncertainty in the inversion (see Kamm et al., 2015).

A simple way of constraining the uncertainty is by reducing the complexity of the model parameterization, provided that the geophysical model provides enough numerical precision in the data prediction, and solving the inverse problem as a sampling problem in the reduced model parameterization. We show the application of this methodology in gravimetric exploration in a synthetic case and also in real dataset acquired in an evaporitic geological setup to detect the presence of possible cavities. For that purpose, we have used the RR-PSO algorithm (Fernández-Martínez and García-Gonzalo, 2012) which is an exploratory member of the PSO family. The use of PSO and other global optimization techniques in geophysical inversion is not new and has been employed in several fields, sometimes combined with model reduction techniques (see for instance Monterro Santos et al., 2016; Fernández-Martínez et al., 2012b; Soltani-Mohammadi et al., 2016; Mu et al., 2015; Pullier et al., 2015, 2017; Desmarais and Spiteri, 2017, 2018; Ekinci et al., 2020) has recently presented a good summary about the use of global optimization techniques and other metaheuristics in potential field data inversions. Besides, the use of simple geometrical bodies to match gravimetric anomalies combined with PSO and other global optimization techniques has been used by Shaw and Srivastava (2007), Toushmalani (2013) or Roshan and Singh (2017), Trivedi et al. (2020).

The main novelties presented in this paper are: 1. The possibility of reducing the high underdetermined character of the anomaly shape inversion problem by model parameterization using polygons and elliptical bodies 2. Inversion and uncertainty analysis of the geophysical model parameters-background density and location and densities of the anomalies via an exploratory member of the PSO family (RR-PSO). As other members of the PSO family, RR-PSO is a free-parameter tuning version: the inertia and local and global accelerations of the different particles of the swarm are automatically chosen based on stochastic stability analysis of the particle trajectories. 3. The exploratory character of RR-PSO allows sampling the set of plausible solutions in a sampling while optimizing mode that is much faster than random sampling.

In conclusion this paper shows a robust methodology to perform anomaly shape inversion with their corresponding in micro gravimetric surveys with its corresponding uncertainty analysis.

2. Problem statement

The problem consists in obtaining the location and the density of one or more dense bodies when the observed gravity data is contaminated by white noise.

The vertical component of gravity measured at the station \( s^k \) produced by a continuous density distribution \( \rho(x,z) \) is (Zhdanov, 2015):

\[
 g_s = \int_{\Omega} \rho(x,z) dxdz \int_{-\infty}^{+\infty} \frac{z-z^k}{\| r-r^k \|^3} dy^k, \quad k = 1, ..., m \tag{1}
\]

where \( r^k = (x^k, z^k) \) represents the position vector of the measure point \( s^k \) with respect to the considered reference origin, \( r = (x,z) \) stands for the position in the dense body, and \( \Omega \) is the geophysical domain. This formulation is valid for the 2D case, assuming a cylindrical symmetry in the y-axis in 3D. It can be easily adapted to the 3D inverse problem by changing the bounds on \( y \) for the geophysical model.

Taking into account that

\[
\int_{-\infty}^{+\infty} \frac{1}{\| r-r^k \|} dy^k = \frac{2}{\pi} \left( \frac{z}{\| r-r^k \|} \right)^{\frac{1}{2}}
\]

the continuous forward problem can be stated as follows

\[
g_s = F_p(\rho(x,z)), \quad k = 1, ..., m
\]

where

\[
F_p = 2G \int_{\Omega} \frac{z-z^k}{(x-x^k)^2 + (z-z^k)^2} dxdz, \quad k = 1, ..., m
\]

is the continuous forward operator at the station point \( s^k \) and \( \rho(x,z) \) is the continuous density distribution of the considered dense body. Then, from (2)-(4) we have

\[
g_s = 2G \int_{\Omega} \frac{z-z^k}{(x-x^k)^2 + (z-z^k)^2} \rho(x,z) dxdz, \quad k = 1, ..., m.
\]

The discrete forward problem consists in predicting the gravity data originated by a dense body discretized by a mesh composed of rectangular cells, with sides parallel to the coordinate axes \( x, z \) at the gravity measure stations located on the surface of the terrain. The discrete inverse problem consists in adopting a discretization of the geophysical domain to solve the set of equation (5) numerically. The most common discretization is in rectangular cells of constant density, adopting piecewise functions set, that is, a set of functions:

\[
\phi_l(x,z) = \begin{cases} 
1 & \text{if } (x,z) \in \text{cell}_l, \quad l = 1, ..., n, \\
0 & \text{otherwise}
\end{cases}
\]

where \( n_r n_z \) being \( n_r, n_z \) the numbers of cells of the mesh in the \( x \) and \( z \) directions.

In order to obtain a good prediction of the gravity data (Bouguer’s anomaly) via the solution of the forward problem a fine mesh is constructed, therefore the number of mesh nodes (\( n_r n_z \)) is large. The straightforward consequence is the increasing of the number model parameters and the high underdetermined character of the corresponding inverse problem. Adopting this set of basis functions, we have:

\[
\rho(x,z) = \sum_{l=1}^{n_r n_z} \rho_l \phi_l(x,z),
\]

where \( l \) denotes the index of the corresponding cell of the mesh. Taking into account equations (1) and (7), we have

\[
g_s = 2G \int_{\Omega} \frac{z-z^k}{(x-x^k)^2 + (z-z^k)^2} \sum_{l=1}^{n_r n_z} \rho_l \phi_l(x,z) dxdz, \quad k = 1, ..., m,
\]

which can be written as follows

\[
g_s = G \sum_{l=1}^{n_r n_z} \rho_l F_{s^k}
\]

where

\[
F_{s^k} = 2G \int_{\Omega} \frac{z-z^k}{(x-x^k)^2 + (z-z^k)^2} dxdz.
\]

Supposing that the coordinates of the \( l \)-th cell of the mesh are: \( x_l, x_{l+1} \times z_l, z_{l+1} \), \( F_{s^k} \) can be calculated as follows (Barbosa and Silva, 1999):
\[ F_{in} = 2 \int_{t_i}^{t_f} \int_{t_i}^{t_f} \frac{z - z_t}{(x - x_t)^2 + (z - z_t)^2} dxdz = \]
\[ = A \log \frac{A^2 + D^2}{A^2 + C^2} - B \log \frac{B^2 + D^2}{B^2 + C^2} + 2D \left( \arctan \frac{A}{D} - \arctan \frac{B}{D} \right) \]
\[ + 2C \left( \arctan \frac{A}{C} - \arctan \frac{B}{C} \right). \]  
(11)

with \( A = z_h - x_i, \ B = z_h - x_{i+1}, \ C = z_h - z_i, \ D = z_h - z_{i+1}, \) and \((x_i, z_i)\) the coordinates of the station point \(s_i\). Finally, the discrete linear inverse problem to be solved is:

\[ g_{inv} = G F p, \]
(13)

where \( F \in M_{m\times n}(\mathbb{R}) \) and each element \( F_j \) contains the response of the basis function associated to the \( j \)-th cell at the measurement point \( s_j \), \( p = (p_1, p_2, \ldots, p_m) \) is the discrete 2D density, and \( G = 6.67 \times 10^{-11} \text{m}^3 \text{kg}^{-1} \text{s}^{-2} \). Besides, to provide the predictions in mGal, \( g_{inv} \) is usually much smaller than \( g_{Fp} \). Due to the number of stations \( m \) is usually much smaller than the number of parameters \( n \), the problem stated in (13) is under-determined (\( m < n \)) and rank deficient if \( \text{rank}(F) < \text{min}(m, n) \). Therefore, a different parameterization of the model space has to be used to avoid the high underdetermined character of this linear inverse problem, since the null-space of the forward operator has dimension \( m-\text{rank}(F) \). In this paper we used two different forward models: 1. The Barbosa and Silva formulas [11–12] which is valid for a mesh solution of the gravimetric problem. 2. Talwani’s formula that is explained in section 4.1, and computes the gravimetric anomaly produced by a polygonal body. We will show that both approaches provide similar results, but the second approach is faster.

2.1. RR-PSO algorithm

Particle Swarm Optimization (PSO) algorithm is stochastic evolutionary computation technique inspired by the behavior of some individuals in nature [Kennedy and Eberhart, 1995]. This technique minimizes iteratively the cost function in an optimization problem, taking into account the mists of the different particles of the swarm and their previous error track. Different variants of this algorithm can be obtained by modifying how the velocities and the acceleration of the different particles are described. In this paper we used RR-PSO (Fernández-Martínez and García-Gonzalo, 2012).

This algorithm was obtained by adopting a regressive discretization in acceleration and velocity in the PSO continuous model, which results in an optimal balance between exploration and exploitation abilities. The RR-PSO algorithm in continuous time is:

\[ v(i + \Delta t) = v(i) + \psi_1 \Delta t [g(i) - x(i)] + \psi_2 \Delta t [(i) - x(i)], \]
\[ x(i + \Delta t) = x(i) + v(i + \Delta t) \Delta t, \]
\[ x(0) = x_0, \]
\[ v(0) = v_0, \]
(14)

where \((i, \Delta t) \in R^2, v \) is the velocity, \( x(t) \) is the current position, \( l(t) \) is the previous best position, \( g(t) \) is the global best position, \( \psi_1, \psi_2 \) are the random global and local accelerations, \( \varphi \) is the total mean acceleration, and \( \omega \) is a real constant called inertia weight. In formula (14), \( x(t) \) describes the trajectories of any particle in the swarm in continuous time. In discrete time the RR-PSO algorithm for the \( i \)-particle in the swarm is as follows:

\[ v_{i+1} = v_i + \psi_1 \Delta t [g_i - x_i] + \psi_2 \Delta t [l_i - x_i] + 2 \omega \Delta t \varphi [x_i - x^0], \]
\[ x_{i+1} = x_i + v_{i+1} \Delta t, \]
\[ v_i = 0. \]  
(15)

In this case the particles are expressed using the rectangular parameterization for the anomalies (described in section 3) and \( N_{\text{ave}} \) represents the number of gravimetric anomalies.

The space of admissible geophysical models, \( M \), is defined as:

\[ l_i \leq x_i \leq u_i, \quad 1 \leq j \leq n, \quad 1 \leq i \leq N_{\text{ave}}, \]

where \( l_i, u_i \) are the lower and upper limits for the \( j \)-th coordinate of each geophysical model, \( n \) is the number of parameters in the inverse problem, and \( N_{\text{ave}} \) is the number of particles of the swarm size with an initial velocities zero. The initial population is evaluated solving \( N_{\text{ave}} \) forward problems, and the global best and the previous best of each particle are determined. The algorithm iterates till the maximum number iterations is finished. The exploratory character of RR-PSO allows performing an approximated sampling of the nonlinear equivalence region of the model parameters, as defined in section 3. In this paper we used the cloud version of RR-PSO, whose main advantage is that no parameter tuning (inertia, local and global accelerations) is needed, since each particle in the swarm has its own PSO parameters that are randomly selected from a set of PSO parameters that are located in the neighborhood of the upper limit of their second-order stability regions. Particularly, in the case of RR-PSO the optimum parameter sets are located along the line \( \varphi = 3(\omega - 3/2) \), mainly for inertia values \( \omega > 2 \) (see Fernández-Martínez and García-Gonzalo, 2012). Further details for the implementation of this algorithm for geophysical inversion and uncertainty analysis in gravimetric inversion of sedimentary basins can be also consulted in Pallero et al. (2015, 2017).

3. Model reduction via anomaly shape reparameterization

An alternative to reduce the indeterminacy is dimensionality reduction via model reparameterization, describing the anomalous bodies by polygons or ellipses, instead of discretizing the domain into a large number of cells. This idea drastically decreases the number of model parameters at the expense of converting the inverse problem to nonlinear, since the forward model depends nonlinearly in the geophysical parameters that describe the value and the location of the density anomalies.

The simplest parameterization that can be adopted in the case of an homogeneous medium with an anomalous body imbedded on it is \( [\rho_a, x_a, x_l, x_r, x_z] \), where \( \rho_a \) represents the density of the homogeneous background, \( x_a, x_r, x_z \) the minimum and maximum initial and final \( x \)-coordinates for the anomalous body and \( x_z \) the corresponding \( z \)-coordinates. In order to locate this anomalous body, a 2D grid with \( n_x \) cells in the \( x \)-direction and \( n_z \) cells in the \( z \)-direction has to be constructed, and the inverse problem solved. This parameterization allows reducing the dimension from \( n_x n_z \) to \( 5 n_z + 1 \): one parameter for the background density \( \rho_b \) and five to parameterize each of them, anomalous bodies. The given parameterization with the model parameters represents the current position of a swarm particle in (15).

In the present case, the synthetic model that we built for numerical experimentation has a background density and it is composed by two anomalous bodies with the following parameters:

- Background density: \( \rho_b = 2720 \text{ kg/m}^3 \).
- Anomalous body 1: \( [\rho_a, x_l, x_r, x_z, x_z] = [3500, 161, 299, 69, 139] \).
Fig. 1. Synthetic model composed of two anomalous bodies embedded into a homogeneous background. The rectangles represent the search space used to find these anomalous bodies via PSO.

Search space:
Density background \( \rho_b \)
Anomaly \([\rho_a, x_1, y_1, z_1, z_1']\)

Swarm of geophysical models
\([\rho_b, \rho_a, x_1, y_1, z_1, z_1']\)

Barbosa
Talwani
Grid generation

g prediction
\(g^{pred} = F\rho\)

Fig. 2. Rectangular parameterizations. Flowchart for the inversion via Barbosa-Silva's and Talwani's formulas.

- Anomalous body 2: \([\rho_a, x_1, y_1, z_1, z_1'] = [1000, 571, 801, 99, 171]\).

Note that the anomalous body 1 has a density of 3500 which is higher than the background density, while in the case of the anomalous body 2 the density is much lower (1000).

Fig. 1 shows a sketch of this synthetic model. It is necessary to differentiate between the parameters of the anomalous bodies (indicated in yellow and purple in Fig. 1) and the lower and upper limits of the search space to locate these bodies (indicated in blue and brown). To generate the synthetic data (Bouguer’s anomaly), we have used a mesh to discretize the density model composed of 500 cells in the \(x\) direction and 100 cells in the \(z\) direction. Therefore, the \(F\) matrix of the forward operator, calculated according to formulas (11)-(12), has a size of \(100 \times 500\), to determine the gravimetric anomaly produced by the 50000 cells of the density model in each of the 20 measurement stations placed between coordinates 200 and 800 in the \(x\)-direction. The observed data \(g^{obs}\) generated with this synthetic model (formulas 8 and 9) is perturbed with a white Gaussian noise of zero mean and standard deviation of 5% of the mean of total anomaly (0.52 mgal in this particular case).

For the inverse problem we have adopted a rectangular parameterization. We have used the following search space:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background density</td>
<td>[2500]</td>
<td>[3000]</td>
</tr>
<tr>
<td>Anomalous body 1</td>
<td>[2000, 100, 200, 40, 90]</td>
<td>[4000, 250, 400, 99, 180]</td>
</tr>
<tr>
<td>Anomalous body 2</td>
<td>[500, 400, 750, 60, 120]</td>
<td>[2000, 800, 900, 120, 190]</td>
</tr>
</tbody>
</table>

Fig. 2 shows the flowchart for the rectangular parameterizations via the Barbosa and Silva’s formulas and Talwani’s. The main difference between both approaches is that Talwani’s formula (explained in section 4.1) does not need a grid generation to compute \(F\) and make the prediction according to (11)-(13). In the case of Barbosa and Silva’s formulas, although the inverse problem is solved using the rectangle parameterizations for the anomalies, there is a previous step of a grid generation corresponding to this model in order to use formula (13). In this case we have used Barbosa and Silva’s formulas to compute the Bouguer’s anomalies produced by these two bodies. The comparison of both approaches will be done in the real case presented in section 4.

RR-PSO is used as an approximated posterior sampler, to analyze the set of equivalent models that fit the observed data \(g^{obs} \in \mathbb{R}^m\) with the same misfit tolerance \(E_{tol}\):

\[
M_{tol} = \left\{ m \in \mathbb{R}^m : \frac{\|Fm - \mathbf{g}^{obs}\|_2}{\|\mathbf{g}^{obs}\|_2} \leq E_{tol} \right\}.
\] (16)

In this case we have considered a misfit tolerance \(E_{tol} = 10\%\) for the
Fig. 3. (A) Observed and predicted gravity anomalies from the best inverted model. (B) Convergence curve of RR-PSO.

posterior analysis. Fig. 3A shows the observed data (g_{obs}) and the predictions (g_{pred}) obtained with the best inverted model found by RR-PSO using a maximum number of 50 iterations and a swarm size of 50 particles.

Fig. 3B shows the convergence curve, that is, the variation of the relative misfit error with the iterations. Although the relative error starts at 31.2%, the region of lower misfits (relative error lower than 5%) is attained at iteration 10 and after iteration 36 it remains below 3%. This situation can be described as overfitting since the data mismatch is lower than the level of noise added to the observed data. This is one of the dangers when the only target is minimizing the data misfit, without performing the corresponding uncertainty analysis. It can be observed that after 5 iterations most of the models are located within the region of relative misfit of 10%, where the posterior analysis is performed.

Fig. 4A shows the true density model and the best inverted model shown both in matrix form. The best inverted model was:
Fig. 5. Posterior histograms of the two dense bodies' inversion: densities and anomalies locations.

- Inverted Background density: \( \rho_b = 2819 \text{ kg/m}^3 \).
- Inverted Anomalous body 1: \( [\rho_a, x_1, y_1, z_1, \sigma] = [3013, 127.1, 303.3, 45.3, 147.3] \).
- Inverted Anomalous body 2: \( [\rho_a, x_2, y_2, z_2, \sigma] = [1686, 593.2, 812.9, 104, 163.7] \).

Obviously, due to effect of noise in data (Fernández-Martínez et al., 2014a,b) the best inverted does not coincide with the synthetic model that has been used to generate the data. The inverted density value of the second anomalous body (1686 kg/m\(^3\)) is higher than the true value (1000 kg/m\(^3\)). The position of this anomalous body is better outlined than the one of the high anomalous dense body. Its inverted value (3013 kg/m\(^3\)) is smaller than the true value (3500 kg/m\(^3\)). Nevertheless, this methodology serves to detect quite accurately the presence of both anomalous bodies. RR-PSO sampled 1433 equivalent models in the \( M_{\text{rel}} \) region of 10% relative misfit. Based on these models we have estimated inverted the median model \( m_{\text{med}} \) and the interquartile range (IQR) for the uncertainty analysis of the solution (see Fernández-Martínez et al., 2017). Fig. 4B shows both fields. It can be observed that the median model delineates better the location of both anomalous bodies. In this case, the median model provides the following values for the density background and the density and location parameters for the two anomalous bodies:
bodies, deduced from the posterior sampling. It can be observed that the gravimetric inverse problem in the reduced parameterization.

It is possible to observe that in some cases, the behavior is coherent in all the profiles: located between 200 and 600 m in the direction and centered around 600 m. These anomalous bodies are marked by black arrows. For all profiles we have calculated the median model and the IQR of the sampled models.

Table 1 and Fig. 8 show the inverted anomalies (Ai) for these profiles. The number of anomalous bodies varies between 9 anomalies for profile P1, 8 anomalies for profiles P3 and P8, and 6–7 anomalies for the profiles located in the central part of the survey (P2, and P4 to P7). Fig. 8 shows the different profiles the median and IQR models obtained with the equivalent models sampled by RR-PSO in the region of equivalence of 20% and IQR. The lowest gravimetric anomalies are the anomaly 7 in profiles P3 and P5, the anomaly 4 in profile P6 and the anomaly 8 in profile P1. The lowest density anomaly found was 1500 $\text{kg/m}^3$ in profiles P3 and P5 and the highest 3200 $\text{kg/m}^3$ in profiles P3 to P6 and P8. The positions of the low density anomalous bodies are also very coherent in all the profiles: located between 200 and 600 m in the x direction and centered around 600 m. These anomalous bodies are shallower than 80 m in depth, but in profile P6 they go deeper (closer to 99 m). The inverted density of the background varies between 2600 $\text{kg/m}^3$ (P3, P5, P7 and P6) and 2633 $\text{kg/m}^3$ (P1). The density values obtained are very different for separated profiles for only 50 m. The highest density anomalous body in the center of the profiles shows different depths. These results seem to be in agreement with the following results:

- Median density background: $\rho_b = 2828 \text{ kg/m}^3$.
- Median Anomalous body 1: $[\rho_a, x_i, y_i, z_i] = [3015, 123.3, 302.6, 41.7, 147.4]$.
- Median Anomalous body 2: $[\rho_a, x_i, y_i, z_i] = [1698, 592.4, 816.3, 3015, 123.3, 302.6, 104.9, 163.8]$.

Particularly, the shape of the low IQRs fits very nicely the shape of the anomalous bodies. The IQR grows where these models are uncertain, that is, when a posterior exhibits a high variability. An alternative view for the uncertainty assessment of the solution consists in analyzing the posterior marginal probability distributions of the model parameters. Fig. 5 shows the posterior histograms corresponding of the two dense bodies, deduced from the posterior sampling. It can be observed that the dispersion of these parameters are different and correspond to the degree of indetermination of the model parameters induced by the observed data. It is possible to observe that in some cases, the behavior is close to Gaussian, while in other cases the histograms exhibit a multimodal behavior that corresponds to the nonlinear character of the gravimetric inverse problem in the reduced parameterization.

Fig. 6. Real study case. (A) Residual and (B) predicted anomalies from the best model (mGal) with profiles.

Fig. 7. Observed and predicted gravimetry profiles (in mGal), from the best inverted model, using rectangular bodies. Black arrows mark the approximated location of the anomalous bodies in each profile.

4. Application to a micro-gravimetric survey

Finally, the methodology has been applied to the interpretation of a micro-gravimetric survey in a geological set up composed of Cretaceous rocks (slicts, muds, sandstones, granodiorite, conglomerates and limestones) and Triassic evaporites (gypsum and halite). Data were collected using a Scintrex CG3 gravimeter with a sensitivity of 5 $\mu$Gal at about 700 stations at 10–30 m intervals along the profiles and 50 m distance between the profiles. Station altitude was measured by a Trimble Dual Frequency GPS with about 1 cm accuracy in the horizontal and vertical coordinates. Fig. 6 shows the observed and predicted Bouguer’s anomalies via the inversion of the profiles P1 to P8.

The aim of this micro-gravimetric survey was the detection of plausible cavities (low dense anomalous bodies) in the salt to avoid leakages of water in a future dam construction to improve the geotechnical analysis of the site. Fig. 7 shows the observed and the predicted data of the gravimetric profiles (P1 to P8) obtained from the best inverted model. The x coordinates are relative to each profile and are the distances taken along these profiles. As it can be observed, the predicted data matches perfectly the observed. The relative misfit is in all cases lower than 0.5% as shown later in Table 3. The positions of the anomalous bodies are marked by black arrows. For all profiles we have calculated the median model and the IQR of the sampled models.
The analytical formula is as follows:

$$g^{\text{nm}}(x_i, z_i) = 2G\Delta \rho \sum_{l=1}^{N_{\text{vert}}} \frac{\beta_i}{1 + \alpha^2} \left[ \log \frac{\Delta z_i + \Delta z_{i+1}}{\sqrt{\Delta x_i + \Delta x_{i+1}}} - \alpha \left( \frac{\Delta z_{i+1}}{\Delta x_{i+1}} - \frac{\Delta z_i}{\Delta x_i} \right) \right].$$

with

$$\Delta z_i = z_i - z_{i-1}, \quad \Delta x_i = x_i - x_{i-1},$$

and

$$\alpha = \frac{x_{i+1} - x_i}{z_{i+1} - z_i}, \quad \beta_i = x_i - \alpha_i z_i,$$

$$\text{Table 1}
Rectangular parameterizations. Inverted anomalies from the best model for different profiles using rectangular bodies.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Profile 1</th>
<th>Profile 2</th>
<th>Profile 3</th>
<th>Profile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>[1643, 26.6, 73, 45.1, 30.5]</td>
<td>[2094, 11, 100, 78, 21.9]</td>
<td>[2860, 39, 90, 246, 30.1]</td>
<td>[2985, 4, 41.2, 15.9, 56.8]</td>
</tr>
<tr>
<td>A2</td>
<td>[2902, 141, 259.2, 8.7, 30.2]</td>
<td>[2800, 120, 294.3, 11.2, 39.9]</td>
<td>[2770, 108, 120, 11.3, 68.8]</td>
<td>[2795, 64.1, 112.2, 16.8, 62.5]</td>
</tr>
<tr>
<td>A3</td>
<td>[1940, 260, 325.1, 13, 32.3]</td>
<td>[1651, 262, 393.4, 15.8, 40.4]</td>
<td>[1992, 139.1, 140.9, 17, 20.1]</td>
<td>[2099, 1546.6, 180, 19.9, 67.3]</td>
</tr>
<tr>
<td>A4</td>
<td>[3091, 396.2, 427.9, 11.7, 44.8]</td>
<td>[3167, 466.3, 533.6, 9.1, 79.7]</td>
<td>[2818, 1844.2, 226.8, 5, 70]</td>
<td>[3010, 1891, 243.2, 25.1, 54.4]</td>
</tr>
<tr>
<td>A5</td>
<td>[2744, 434.7, 460.7, 17.3, 29.9]</td>
<td>[2959, 607.5, 679.5, 5.7, 43.8]</td>
<td>[1899, 2993, 401.5, 21.8, 70.2]</td>
<td>[1898, 2848, 408.7, 24.6, 70.5]</td>
</tr>
<tr>
<td>A6</td>
<td>[3197, 472.8, 510, 8.3, 89.7]</td>
<td>[3100, 766.6, 897.6, 8.5, 79]</td>
<td>[3200, 463.6, 559.9, 10.4, 79.7]</td>
<td>[3200, 477.7, 557.6, 5.5, 88]</td>
</tr>
<tr>
<td>A7</td>
<td>[3048, 527.3, 599.2, 6.7, 90]</td>
<td>[1500, 597.2, 667, 6.8, 66]</td>
<td>[3100, 750, 900, 5.2, 100]</td>
<td></td>
</tr>
<tr>
<td>A8</td>
<td>[1500, 552.7, 567.3, 14.3, 81.1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A9</td>
<td>[3186, 712.7, 800, 7.8, 89.8]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RMSE and relative misfit error values obtained by the Talwani's and Barbosa and Silva's formulas for all the profiles. Bold faces denote the smaller relative misfit in each case.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Profile 1</th>
<th>Profile 2</th>
<th>Profile 3</th>
<th>Profile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>[3093, 4.3, 46.7, 5, 30.1]</td>
<td>[2200, 11, 27.2, 5.1, 59.9]</td>
<td>[2861, 3.2, 10.4, 60]</td>
<td>[3200, 1.3, 30, 5.2, 59.8]</td>
</tr>
<tr>
<td>A2</td>
<td>[2796, 72.6, 99.9, 19.9, 45.4]</td>
<td>[1866, 68.8, 75.4, 123, 33]</td>
<td>[1659, 60, 87.7, 20, 29]</td>
<td>[1886, 30, 51, 17.2, 28]</td>
</tr>
<tr>
<td>A3</td>
<td>[2099, 156.2, 160, 18.9, 40]</td>
<td>[2989, 140.1, 230, 18.3, 58.5]</td>
<td>[2980, 168.6, 220.6, 16.1, 69.6]</td>
<td>[3200, 128.9, 220, 15.6, 37.2]</td>
</tr>
<tr>
<td>A4</td>
<td>[3200, 207.9, 210, 5, 70]</td>
<td>[1500, 300, 300.9, 20, 96.3]</td>
<td>[2025, 2948, 329, 20, 71.8]</td>
<td>[1859, 2999, 3029.9, 10.1, 70]</td>
</tr>
<tr>
<td>A5</td>
<td>[2091, 308.4, 409.4, 19.9, 71.5]</td>
<td>[1872, 380, 391.7, 17.9, 98.8]</td>
<td>[2099, 367.9, 403.2, 18.1, 70.8]</td>
<td>[2887, 3536, 390.1, 10.1, 31.1]</td>
</tr>
<tr>
<td>A6</td>
<td>[3200, 476.4, 543.5, 5.1, 88.8]</td>
<td>[3200, 461.9, 524.9, 5, 97.9]</td>
<td>[3119, 4157, 525, 5, 63.8]</td>
<td>[1631, 4028, 421.4, 9, 59.8]</td>
</tr>
<tr>
<td>A7</td>
<td>[1500, 552.9, 643.9, 7.4, 79.7]</td>
<td>[1610, 541.5, 644.4, 29.3, 79.3]</td>
<td>[1790, 555.7, 641.5, 12.2, 68.3]</td>
<td>[3200, 462.1, 532.4, 10, 79.3]</td>
</tr>
<tr>
<td>A8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

geology of the area. The order position of each anomalous body is read from left to right in each profile. To obtain these results we have used rectangular parameterizations, Barbosa and Silva's formulas, and the number of swarm particles used by RR-PSO was 50 and 100 the total number of iterations.

4.1. Forward solution via Talwani’s formula

In this section we compare the results of the inversion to those obtained via Talwani’s formula (Grant and West, 1965) which describes the gravity attraction of an n-sided polygon in a measure point. The analytical formula is as follows:

$$g^{\text{nm}}(x_i, z_i) = 2G\Delta \rho \sum_{l=1}^{N_{\text{vert}}} \frac{\beta_i}{1 + \alpha^2} \left[ \log \frac{\Delta z_i + \Delta z_{i+1}}{\sqrt{\Delta x_i + \Delta x_{i+1}}} - \alpha \left( \frac{\Delta z_{i+1}}{\Delta x_{i+1}} - \frac{\Delta z_i}{\Delta x_i} \right) \right].$$

(17)
Fig. 8. Median models and IQR obtained from the profiles P1 to P8.

Fig. 9. Profile 3. (A) Observed and (B) predicted gravity anomalies from the best model obtained by Barbosa’s and Talwani’s formulas.
where $G$ is the gravitational constant, $\Delta \rho = \rho_a - \rho_b$ is contrast density between anomaly and the host rock, and $x_i$ and $z_i$ are coordinates of the vertices of the polygon. This formula does not need a grid construction to perform the gravity prediction.

Fig. 9 shows the comparison to model the gravity anomaly corresponding to profile P3. The Talwani’s formula predicted with a misfit error of 4.59%, while the formula of Barbosa did it with 5.86%, although the appearance is very similar, and because they were obtained in two different simulations. Similar results can be obtained for the other profiles. Table 2 shows the percentage misfit error and RMSE values obtained by the Talwani’s and Barbosa and Silva’s formulas for all the profiles. The results show the same order of magnitude in all the cases. In profiles P1-P2-P6 and P8, Barbosa and Silva’s showed the smallest error (in bold faces) while in the rest (P3-P4-P5 and P7) Talwani’s formula provided the optimum result with lower computational cost. Therefore, the procedure shown in this paper for anomaly detection in microgravity surveys can be used by Talwani’s formula, which is simpler and much faster to implement.

4.2. Model parameterization by elliptical bodies

Perhaps, a more realistic representation of dense anomalies is the case of oriented ellipses, which are modelled by 6 parameters: the density of the anomalous body, $\rho_a$, the coordinates of the center $(x_c, z_c)$, the lengths of its semiaxes $a$ and $b$, and its orientation, as given by the angle $\theta$ formed by the semimajor axis with the horizontal. The aim of this section is to show that the results are coherent with the ones obtained with rectangles. The flowchart in this case is shown in Fig. 10.

The inversion uses Talwani’s formula to make the forward predictions corresponding to the elliptical bodies. For that purpose, the ellipses have been discretized by a polygonal line with a fixed number of points (50 in this case) and the gravity attraction is calculated via the Talwani’s formula. Barbosa and Silva’s formula could be also used by creating a grid and finding the cells that are discretized to the different anomalies to create the density model. Nevertheless, in this case this approach has not been implemented since Talwani’s formula is simpler and provides enough accuracy, as shown in the previous section.

Fig. 11 shows the median inverted models and IQR for the different profiles and for models with less than 20% misfit error. It can be observed that in profiles P1, P2 and P3 there is a high density anomaly located close to the $800 \text{ m}$ on the $x$ direction which is not visible on the other profiles. Most important low density anomalies are situated between 250m-400m and 580m-650m on the $x$ direction, as it can be observed in profiles P2 to P7. The results obtained with this method are similar to those shown in Fig. 8. Fig. 12 shows the error iteration curves for each profile and both model parameterizations (rectangles and ellipses). It can be observed that the misfit error is in both parameterizations less than 35% being smaller for all profiles in the case of ellipses, which is less than 5% for profiles P1, P2 and P6 to P8. For both parameterizations, the misfit error is smaller than 10% when the number of iterations is greater than 30.

Table 3 shows the number of iterations needed by the algorithm to reach the indicated misfit error region (5.5% and 20%) and also the iteration when the best model was found, comparing rectangular and elliptical parameterizations. It can be observed that although in most of the cases the number of iterations needed to reach the region of lower misfits (5.5%) is smaller in the case of the ellipse, the best model is found sooner in the case of rectangular parameterizations. Nevertheless, the order of magnitude is the same.

Finally, an interpretation of the anomalous bodies can be made by projecting the locations of the inverted anomalies on the $x$-$y$ map. Fig. 13 shows the situation and extension of the most important anomalies. The low density anomaly bodies A1 and A3 present density values of $\rho_1 = 1874 \pm 175 \text{ kg/m}^3$ and $\rho_2 = 1701 \pm 224 \text{ kg/m}^3$, with a coefficient of variation of 9.4% and 13.2% respectively. Likewise, the density values of anomalous bodies A2 and A4 are $\rho_2 = 3175 \pm 44 \text{ kg/m}^3$ and $\rho_4 = 3129 \pm 50 \text{ kg/m}^3$, which coefficients of variation are 1.4% and 1.6% respectively. Anomaly A4 is not monitored on profiles P4 to P8 because they are shorter in length. It can be observed that the lower percentages of the variation correspond to the high density anomalies. This floor plan representation, although approximated, delineates very well the geotechnical problems that might be present in the construction of a dam northeast of the P8 profile, since some subsurface connection might exist through A2 and A4 provoking water leak.

5. Conclusions

In this paper we have presented a methodology to perform anomaly shape inversion in microgravimetry which is based on model reduction by reparameterization and posterior sampling via PSO. This problem is associated to the exploration of ore bodies imbedded into a quasi-homogeneous background and also the detection of cavities via different geophysical methods (gravimetry, magnetic, electric). The state of the art consists in creating a very fine mesh to approximate the shapes and the values of these anomalies, inverting at the same time the geophysical structure of the geological background. This approach results in discrete inverse problems with a huge uncertainty space (null space for linear inverse problems) composed of the set of geophysical models that fit the observed data within the same error bounds. These models are located in flat curvilinear valleys of the cost function topography. The common way of stabilizing the inversion and picking a final solution consists in introducing a reference model (through the prior information) to define the set of correctness of geophysical models (Tikhonov and Arsenin, 1977). This approach has some drawbacks if the reference model was chosen in a different basin where the most plausible solution of the inverse problem is located, that is, if both models
Fig. 11. Median models and IQR obtained from the profiles P1 to P8 using ellipses.

Fig. 12. (A) Iterations misfit error (in %) from the profiles P1 to P8 using rectangles and (B) ellipses.

have completely different structures. Then, the inversion can lead to a wrong solution. We present a different way of dealing with the high underdetermined character of this kind of problems, consisting in first solving the inverse problem using a low dimensional parameterization that provides a numerical approximation of the anomaly via Particle Swarm Optimization (PSO). We show its application to a synthetic and real case in gravimetry performing at the same time uncertainty analysis of the solution. We have also compared the solutions obtained
discretizing the geophysical kernel (Barbosa and Silva, 1994) and by using Talwani’s formula (Grant and West, 1965), showing that both formulas provide very accurate results. Besides, the results are similar using rectangular and elliptical parameterizations to describe the density anomalies. This methodology outperforms the common least squares method with regularization that needs the use of prior information to stabilize the inversion. It can be easily adapted to other geophysical techniques by changing the forward problem operator, and also to 3D parameterizations.

Computer code availability

A software is freely available at https://github.com/jlfmuniovi/Grav2DPSO.

Author contributions

J.L.F.M. designed the methodology. Z.F.M. and J.L.F.M. analyzed the data and wrote the paper; Z.F.M., J.L.F.M. and J.L.G.P designed and developed the computer programs.

Declaration of competing interest

Authors have not any conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cageo.2020.104492.

References

Tsugutani, T., 2013. From Bayes to Tarantola: new insights to understand uncertainty in inverse problems. J. Geophys. 82 (6), G101–G114.
Tsugutani, T., 2013. From Bayes to Tarantola: new insights to understand uncertainty in inverse problems. J. Geophys. 82 (6), G101–G114.
Tsugutani, T., 2013. From Bayes to Tarantola: new insights to understand uncertainty in inverse problems. J. Geophys. 82 (6), G101–G114.

Fig. 13. (A) Bouguer’s anomaly and (B) interpretation of the density anomalies projected onto the surface.