Mathematics

Research article

Posterior analysis of particle swarm optimization results applied to gravity inversion in sedimentary basins

J. L. G. Pallero[1,2,*], **M. Z. Fernández-Muñiz**[2], **J. L. Fernández-Martínez**[2] **and S. Bonvalot**[3]

[1] ETSI en Topografía, Geodesia y Cartografía, Universidad Politécnica de Madrid, Madrid, Spain

[2] Grupo de Problemas Inversos, Optimización y Aprendizaje Automático, Departamento de Matemáticas, Universidad de Oviedo, Oviedo 33007, Spain

[3] Laboratoire GET (Université de Toulouse, CNRS, IRD, CNES), Bureau Gravimétrique International (BGI), Toulouse, France

* **Correspondence:** Email: jlg.pallero@upm.es.

Abstract: As is well known, it is impossible to model reality with its true level of detail. Additionally, it is impossible to make an infinite number of observations, which are always contaminated by noise. These circumstances imply that, in an inverse problem, the misfit of the best estimated model will always be less than that of the true one. Therefore, it is not possible to reconstruct the model that actually generated the collected observations. The best way to express the solution of an inverse problem is as a collection of models that explain the observations at a certain misfit level according to a defined cost function. One of the main advantages of global search methods over local ones is that, in addition to not depending on an initial model, they provide a set of generated models with which statistics can be made. In this paper we present a technique for analyzing the results of any global search method, particularized to the particle swarm optimization algorithm applied to the solution of a two-dimensional gravity inverse problem in sedimentary basins. Starting with the set of generated models, we build the equivalence region of a predefined tolerance which contains the best estimated model, i.e., which involves the estimated global minimum of the cost function. The presented algorithm improves the efficiency of the equivalence region detection compared to our previous works.

Keywords: global optimization methods; particle swarm optimization; inverse problems; uncertainty analysis; gravity inversion
Mathematics Subject Classification: 86A20, 86A22, 90C26

1. Introduction

From a very general perspective, solving an inverse problem involves the integration of methodologies aimed at determining the causes given the effects. This can be exemplified by scenarios such as diagnosing a disease based on symptoms or designing engineering solutions within certain constraints. Common inverse problems in geophysics are the determination of the density distribution in the subsoil using gravity observations at the terrain surface (gravimetric inverse problem [1]), using magnetic observations for mineral prospecting [2], the problem of finding the velocity distribution of a geological medium [3], the seismic reflection inverse problem (which maps the Earth sedimentary crust from transient near-surface recording of echoes, stimulated by explosions or other controlled seismic sources positioned near the surface) [4], the determination of fault parameters using ground deformation [5], or the monitoring of magma volumes and their supply rates in volcanic areas also using ground deformation [6]. Some of these problems are ill-posed, i.e., either the inverse problem does not admit solution, the solution exists but is not unique, or finally, the solution exists and is unique but is unstable, that is, a small change in the data produces a big change in the solution [7, pp. 7-8].

Many inverse problems in geophysics and other fields involve a high number of parameters due to the complexity of the problem and/or the need to model phenomena with high degree of detail. Despite this complexity, physical models are often simplified for ease of formulation and due to the inability to account for all real-world details in a model. There are also problems, such as ones based on potential fields (gravity, magnetics), whose solutions have inherent non-uniqueness (see, for example, [8, pp. 214–217] or [9, pp. 20–22]). Finally, the observed data set is always incomplete and noisy. All these circumstances contribute to the ill-posedness of the problem, leading to many solutions (called *equivalent*), all of them compatible with the observed data set and, in the case of nonlinear inverse problems, probably located in disconnected curvilinear valleys of the cost function topography [10]. It was said by Parker in [11] that "the goal of inverse theory is to determine the parameters from the observations or, in the face of the inevitable limitations of actual measurement, to find out as much as possible about them".

Traditionally, the ill-posed nature of inverse problems has been addressed (selection of best model and its uncertainty estimation) through the combined use of local optimization methods and regularization techniques (see, for example, [12]). Although this approach is fast, it can have, for some nonlinear inverse problems, the drawback of providing a solution that depends strongly on an initial model, resulting in an obtained solution that is not the global minimum of the cost function, but a local one. Also, as important as the global minimum estimation, the uncertainty appraisal is a key task in any inverse problem, but the results based on local optimization methods are, in the case of nonlinear problems, only valid in the neighborhood of the solution that has been found [10].

Global search methods such as genetic algorithms [13], simulated annealing [14] or particle swarm optimization [15] (among others) offer advantages over local optimization methods as they do not rely on an initial model and explore the entire parameter space, thereby enabling estimation of the global minimum of the cost function. After their execution they provide a collection of samples (models) from the parameter space together with their misfit values, from which the user can select the subset of models whose misfit are below a predefined tolerance. The solution of the inverse problem can then be expressed, as it stated by Albert Tarantola in [16], as "the collection of all the models that have not been falsified" (that are compatible with all the a priori information and have a misfit below a predefined

tolerance). This is of particular importance in the case of nonlinear inverse problems, where, using these compatible models, the topography of the cost function lower than the prescribed tolerance can be depicted, making uncertainty appraisal more realistic than using local optimization methods [10].

In some inverse problems, especially in highly dimensional ones, performing posterior analysis and uncertainty appraisal can be challenging. Using only statistics referring to each individual parameter could not be adequate when a global analysis of the entire set of parameters is needed. This is the case with gravity inversion in sedimentary basins, and in this paper we introduce a methodology designed in order to make posterior analysis easily feasible.

2. Cost function landscape in nonlinear inverse problems

In its more general form, a discrete nonlinear inverse problem is presented as

$$\mathbf{F}(\mathbf{m}) = \mathbf{d} + \boldsymbol{\varepsilon}, \tag{2.1}$$

where vector $\mathbf{m}$ belongs to the parameter space $\mathcal{M}$, $\mathbf{m} \in \mathcal{M}$, and is composed of a finite number of parameters, $\mathbf{m} = \{m_1, m_2, \ldots, m_n\}$; vector $\mathbf{d} \in \mathcal{D}$ represents the observed data, $\mathbf{d} = \{d_1, d_2, \ldots, d_l\}$, belonging to the data space $\mathcal{D}$; and $\boldsymbol{\varepsilon}$ is the observational noise, always present. The direct problem functional $\mathbf{F} : \mathcal{M} \to \mathcal{D}$ is considered in this paper as nonlinear.

An inverse problem, linear or not, has many equivalent solutions since they are able to predict the observed data to within some misfit tolerance and are compatible with some prior knowledge. Then, uncertainty estimation in discrete inverse problems involves finding the family M_{tol} of models compatible (below a predefined tolerance value, tol) with the observations and other a priori information, i.e., finding all $\mathbf{m} \in M_{\text{tol}}$ subject to $\|\mathbf{F}(\mathbf{m}) - \mathbf{d}\|_p \leq$ tol. By selecting $p = 2$, a least squares problem is proposed, i.e., the model $\mathbf{m}$ which minimizes the cost function

$$c(\mathbf{m}) = \|\mathbf{F}(\mathbf{m}) - \mathbf{d}\|_2 \tag{2.2}$$

is determined. The cost function (2.2) is in most cases complemented with a regularization term (see [12] for example) in order to stabilize the inversion, but as this does not affect our discussion about the existence of disconnected equivalence regions in the cost function landscape of nonlinear inverse problems, we will ignore it in order to simplify the notation.

One common method for solving a nonlinear inverse problem using local optimization is Newton's method, which implies the linearization of the forward operator $\mathbf{F}$ around an initial model $\mathbf{m}_0$, so

$$\mathbf{F}(\mathbf{m}) = \mathbf{F}(\mathbf{m}_0) + \mathbf{JF}_{\mathbf{m}_0}(\mathbf{m} - \mathbf{m}_0) + o(\|\mathbf{m} - \mathbf{m}_0\|_2), \tag{2.3}$$

where $\mathbf{JF}_{\mathbf{m}_0}$ is the Jacobian of $\mathbf{F}$ in $\mathbf{m}_0$, and $o(\|\mathbf{m} - \mathbf{m}_0\|_2)$ represents the terms of order greater than 1 in the Taylor expansion. Taking into account Eqs (2.1)–(2.3), the solution of a discrete nonlinear inverse problem consists of finding the model $\mathbf{m}$ which minimizes the cost function

$$c(\mathbf{m}) = \|\mathbf{JF}_{\mathbf{m}_0}(\mathbf{m} - \mathbf{m}_0) - \mathbf{d} + \mathbf{F}(\mathbf{m}_0)\|_2. \tag{2.4}$$

As the problem is now linear, well known linear algebra techniques can be employed in order to solve it through an iterative procedure, where the initial model $\mathbf{m}_0$ is updated at each step (see [12]).

Once the model $\mathbf{m}$ that minimizes $c(\mathbf{m})$ is determined, the error analysis can be performed. The boundary of the region $\|\mathbf{JF}_{\mathbf{m}_0}(\mathbf{m} - \mathbf{m}_0) - \mathbf{d} + \mathbf{F}(\mathbf{m}_0)\|_2 \leq$ tol is a hyper-quadric whose axis can be determined in order to provide the uncertainty appraisal (see [10]). However, in the Taylor expansion of the forward operator (Eq (2.3)), the terms of order greater than 1 have been neglected, so this truncation will introduce an error in the approximation of the cost function landscape of the original nonlinear problem. For a deeper understanding of the cost function landscape in linear and nonlinear inverse problems, readers can consult [10, 17].

We will illustrate the preceding statements with a synthetic example. Let a nonlinear function be $y = e^{\alpha x} + 2e^{\beta x}$. For values $(\alpha, \beta) = (10, 1)$, a set of 51 equally spaced values for y have been generated for $x \in [-0.25, 0.25]$ using $\Delta x = 0.01$. Then, the computed y values have been contaminated with white noise of distribution $N(\mu, \sigma^2) = N(0, 5\%^2)$, where 5% means a standard deviation of 5% of the value of y. With the noisy y values, we have applied Newton's method in order to estimate $\mathbf{m}$, and we have obtained the results presented in Figure 1. The upper pictures represent the same solution, but the upper left one shows as background the real nonlinear cost function topography, while the upper right one shows the linearized cost function; the lower picture shows the linearized cost function for other starting guess $\mathbf{m}_0$, as will be explained. The equivalence regions of 10% relative misfit $(100 \cdot \|\mathbf{F}(\mathbf{m}) - \mathbf{d}\|_2 / \|\mathbf{d}\|_2)$ are shown, and it can be seen that in the case of the original nonlinear problem two regions are present (green lines).

However, the estimation using a local optimization method depends for this problem strongly on the location of the initial guess $\mathbf{m}_0$. The result corresponding to the upper right plot in Figure 1 was obtained using $\mathbf{m}_0 = (\alpha_0, \beta_0) = (6, 2)$, and the solution is enclosed in the equivalence region containing the global minimum of the cost function. It can also be seen that the estimated solution (black circle) does not match the true solution (black dot), which is a consequence of the presence of noise in the data (see [10] for details). However, if the initial guess is $\mathbf{m}_0 = (\alpha_0, \beta_0) = (-2, 6)$, as can be seen in the lower picture of Figure 1, the solution is enclosed in the equivalence region of the upper part of the figure, which is a local minimum that is far from the true solution $\mathbf{m} = (\alpha, \beta) = (10, 1)$.

About the cost function topography, we can see in Figure 1 (upper left) its main characteristic in the case of nonlinear problems: isolines of equal misfit are arbitrary bent curves. In linear problems, equal misfit isolines are in the general case hyper-ellipsoids (tending to be elliptic hyper-cylinders for ill-posed problems), but this landscape is totally different in the nonlinear case (see [10, 17] for details). It can be seen (Figure 1, upper left) how the two 10% nonlinear equivalence regions have different sizes and shapes, and the problem solution does not have to be in its center. On the contrary, when a local optimization method is employed, we linearize the problem around the best estimation, and the cost function topography corresponds to a linear problem, as can be seen in Figure 1 (upper right and lower). In this particular example the isolines of equal misfit are ellipses centered in the estimated solution, as we are dealing with a two-dimensional problem. When we apply a local optimization method to a nonlinear inverse problem, two important and general conclusions related to error analysis can be derived:

- Only one equivalence region (around the estimated solution) will be determined, ignoring other possible regions included in the cost function topography. Furthermore, if the initial guess $\mathbf{m}_0$ is far from the region containing the global minimum, the algorithm could be trapped in a local one with its own linearized equivalence region, making both the estimation and the posterior error analysis useless (this is the case with the example in the lower part of Figure 1).

- Although the estimated solution corresponds to the cost function global minimum, the linearized equivalence region is different in shape from the nonlinear one, as can be seen in Figure 1. All models inside an equivalence region of prescribed tolerance must be taken into account, as they can explain the observations (see [16, 18, 19]). However, when we compare the linear and nonlinear regions, we can see that (1) the linearized equivalence region contains models that are outside the true nonlinear equivalence region, and (2) models belonging to the nonlinear equivalence region are located outside the linearized one, so they will be ignored.

These facts are much more pronounced in highly nonlinear problems, so the error analysis based on the linearized cost function topography is valid only in the very near surroundings of the estimated solution, making, in many cases, the uncertainty appraisal useless.
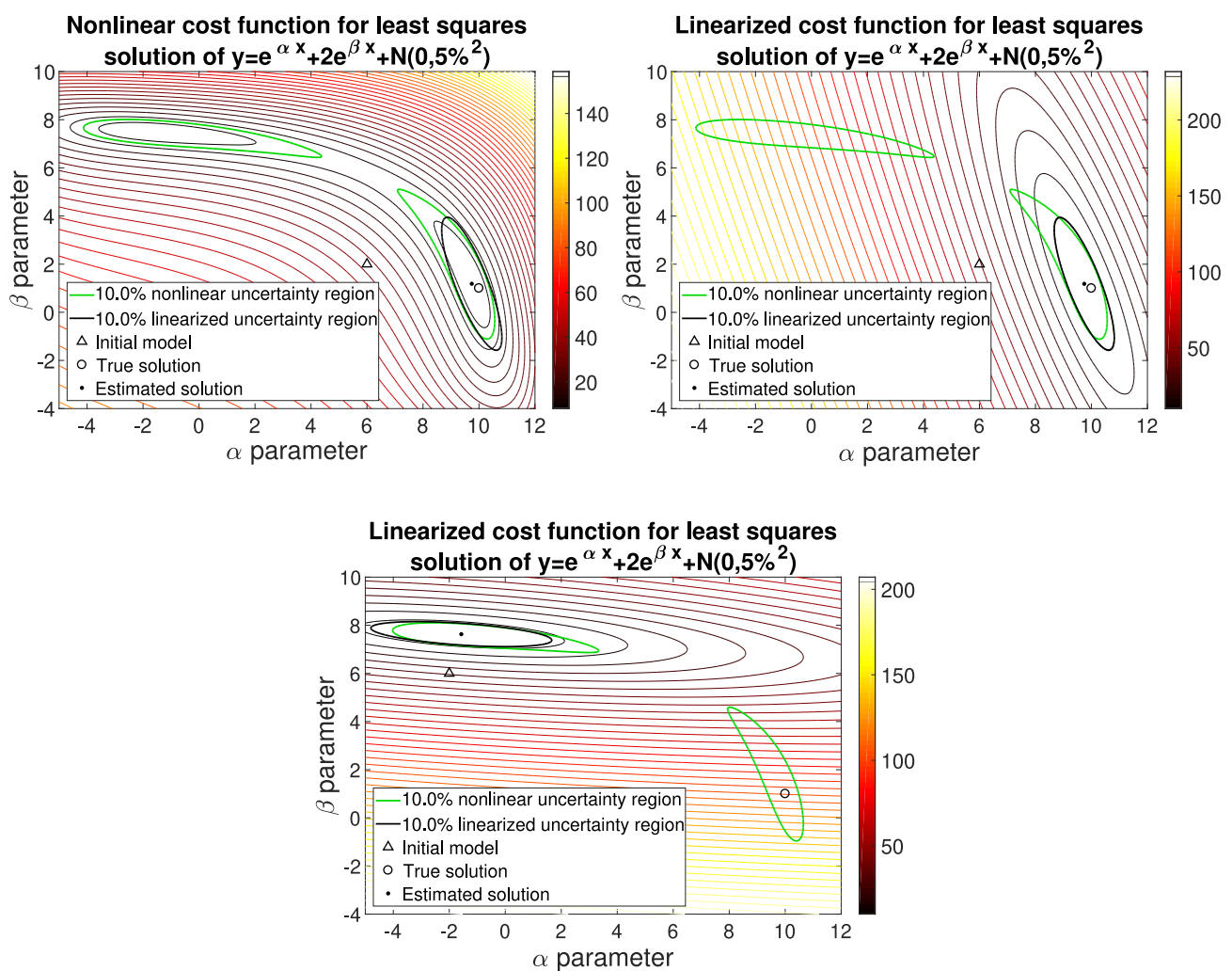
Figure 1. Least squares relative cost function topography ($100 \cdot \|\mathbf{F(m)} - \mathbf{d}\|_2 / \|\mathbf{d}\|_2$) of the synthetic example $y = e^{\alpha x} + 2e^{\beta x} + N(0, 5\%^2)$. Upper left: real nonlinear cost function topography. Upper right: linearized cost function topography for a solution starting with point $\mathbf{m}_0 = (6, 2)$. Lower: linearized cost function topography for a solution starting with point $\mathbf{m}_0 = (-2, 6)$. Color bars represent the relative misfit as percentages.

Global optimization methods can have significant improvements over local optimization ones in certain cases. First of all, it is not necessary to have an initial guess for the solution, as these methods explore the entire parameter space (actually a predefined subset of it, based on the knowledge of the problem), so it is difficult to get trapped in cost function local minima. This is excellent for nonlinear inverse problems and their complex cost function topographies. Also, these methods are based on the evaluation of the forward problem, so neither linearization nor big systems of equations and their factorization are needed. On the contrary, when the problem is highly dimensional, and/or the forward problem is computationally expensive, global optimization methods can be slow. Some techniques can be applied in order to mitigate this drawback [20], although not always is possible.

Nevertheless, it is in the uncertainty analysis where global optimization methods have their principal strengths. During the execution, a set of models is generated covering the parameter space following the rules of each method (genetic algorithms, simulated annealing, particle swarm optimization, etc.), i.e., the cost function topography is sampled following certain rules and focusing on the equivalence regions containing the minima. Then, by selecting the models with a misfit below a predefined tolerance, the real nonlinear equivalence regions can be depicted. In the following sections, we present a methodology designed in order to select the corresponding models for a correct uncertainty representation in a gravity inversion method applied to sedimentary basins.

## 3. Two-dimensional gravity inversion in sedimentary basins

In [21], GRAVPSO2D, a software for the solution of the two-dimensional gravity inversion problem in sedimentary basins based on the particle swarm optimization (PSO) algorithm, was presented (some other studies have been developed in recent years about this topic, such as [22–25], for example). The 2D approximation can be used for gravity when a dimension of an anomalous body is much larger than the other two, which is common in sedimentary basins, where the depth is generally less than the horizontal dimensions (Nettleton suggested in [26, p. 206] that a 2× or 3× factor is enough). The basin is then modeled as a set of rectangles [27–30], and after the prescription of density values, the rectangles' thickness (i.e., the sediment depths) are determined. The observations are gravity anomalies derived from gravity observations at the terrain surface level.

Figure 2 shows a schematic representation of the model. The basin, as stated in the previous paragraph, is composed of a set of M rectangles with known density contrast (it can be constant or variable in depth), whose lower sides represent the interface between the basement and the sediments, and their upper sides are coincident with the terrain surface (which does not need to be planar as in the example of Figure 2). The observed points form a set of N elements located at the terrain surface level. The gravitational anomaly Δg generated by a set of M rectangles over each observation point is

$$\Delta g_i = \sum_{j=1}^{M} F\left(\Delta\rho_j, \mathbf{r}_{ij}, z_j\right), \ i = 1, 2, \ldots N, \tag{3.1}$$

where $\Delta\rho_j$ is the density contrast ($\rho_{\text{sediments}} - \rho_{\text{basement}}$) of each rectangle j, $\mathbf{r}_{ij}$ is the top side position of each rectangle j with respect each point i, and z_j is the z lower side coordinate of each rectangle j (the problem unknowns). The expression for $F\left(\Delta\rho_j, \mathbf{r}_{ij}, z_j\right)$ can be seen in [31, pp. 524-525] or [32] and is nonlinear with respect to r and z. Equation (3.1) can be extended with the inclusion of a term in order to take into account a regional trend in the gravity anomaly, and, in order to reduce the non-uniqueness

of the problem, several kind of constraints can be applied to the prisms, such as fixed depths (data from boreholes, for example) [21]. In this paper we work with Eq (3.1) for the sake of simplicity, since the inclusion of other terms does not affect the discussion. Also, a filtering step is needed in order to produce geologically meaningful results (see [21] for details).
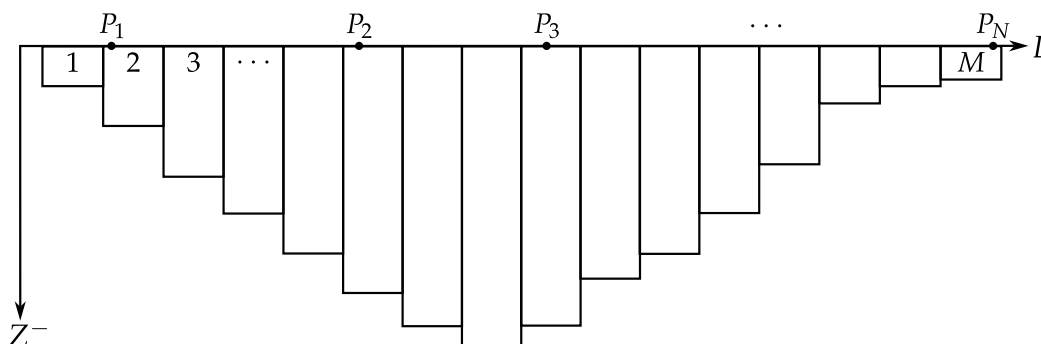
Figure 2. Two-dimensional model of a sedimentary basin as a set of rectangles.

The first step before the PSO execution is the selection of the search space for the parameters. In our case, this means definitions of upper and lower limits for the prisms' z coordinates in Figure 2. During the execution of PSO, a set of models is generated, all enclosed inside the search space. Among these models, the one with the lowest misfit (Eq (2.2)) is considered as the global minimum of the cost function. Also we have at our disposal all the other models, so posterior statistics can be made.

Figure 3 shows a synthetic example where a basin is modeled as a set of 80 rectangles of 1000 m horizontal width, with upper sides at height 0 m, and density contrast $\Delta\rho = -500\,\mathrm{kg\,m^{-3}}$. Their true depths are shown as the solid black line in Figure 3 (left). The horizontal positions of a set of 70 points at height 0 m were randomly generated, and gravity anomaly induced by the basin was computed in each one. These observations were contaminated with random white noise of distribution $N(0, 0.75^2)$ mGal. In order to recover the true model, a population of 400 models with density contrast $\Delta\rho = -500\,\mathrm{kg\,m^{-3}}$ was generated inside the search bounds, selected wide enough to contain all geologically plausible models (red dotted lines in Figure 3, left), and 200 iterations were performed, so a total of 80 000 models were created using the CP-PSO member in its cloud version, i.e., no parameter tuning must be configured as each particle has its own PSO parameters, automatically chosen [33].

The inversion took a time of about 35 s to be completed,[*] and the best generated model can be shown in Figure 3 (left) as a blue line and has a relative misfit of 4.19%, which is an estimation of the global minimum of the cost function. This misfit value is smaller than that corresponding to the true model (black line, labeled as *reference model* in the figure, misfit of 5.44%), which is an effect of the presence of $N(0, 0.75^2)$ mGal noise in data [10]. Still, we are interested in models with a relative misfit below a predefined tolerance, 10% for example, so all of them are extracted from the whole set of 80 000 generated models. In our case the number of models belonging to the equivalence regions with relative misfit below 10% is 51 612, and Figure 3 (right) shows the statistics that can be made for each element (rectangle number 54 as an example in this case). With the empirical cumulative density

---

[*]The computations were performed on a laptop equipped with an Intel Core i7-4800MQ CPU with four cores at 2.70 GHz, running Debian GNU/Linux and MATLAB R2017a. The function for the direct problem solution was written in C and compiled using OpenMP, so all four cores of the CPU were used. After several runs we can conclude that the time spent in the uncertainty analysis is about 10% of the total time, showing that the forward modelling step is the bottleneck of any global search algorithm.

function, ECDF, (upper right panel) the probability of a depth range can be easily computed, while the histogram (lower right panel), which actually is another way to represent the ECDF, allows us to analyze in a more efficient way the frequencies of different depths. In this example we can see in the histogram the depth of the best model together with the depths of the mean, median, and mode models for rectangle 54. Such models do not have to be coincident, and, as can be seen, the best model does not have to be composed by the ensemble of most frequent depths in each rectangle.

The range of variation for each rectangle depth is shown in Figure 3 (left) as yellow bars. These bars could be seen as the limits of an equiprobable range, but this is not true in the general case, as is shown in the lower right panel of Figure 3. So, ECDFs and histograms must be carefully inspected in order to provide a detailed behavior of the problem solution, revealing the importance of both as support to the global view.
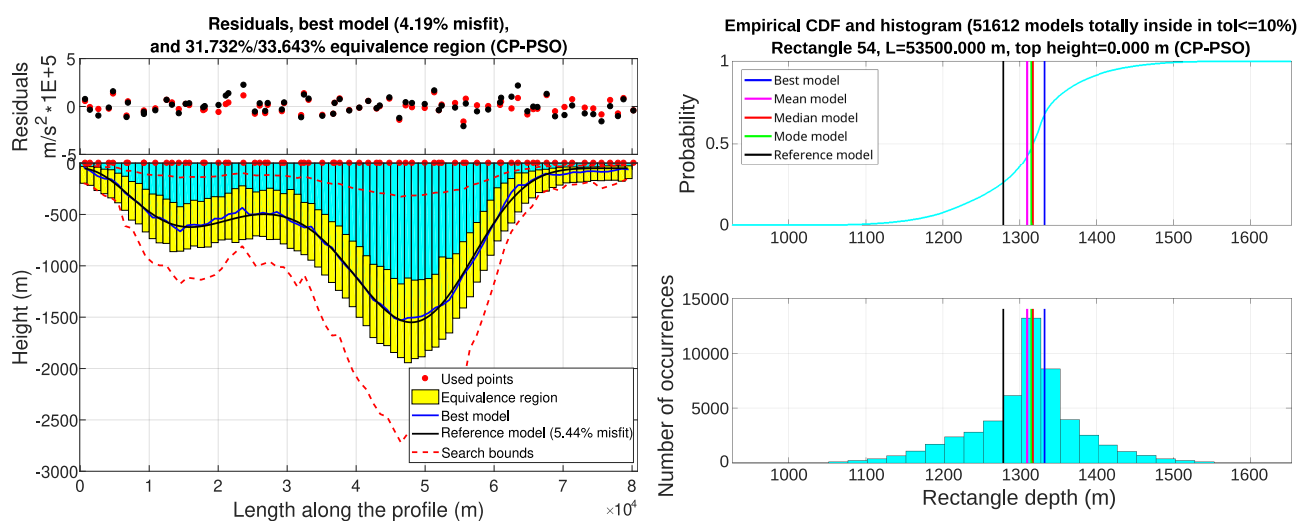
Figure 3. Synthetic example of a sedimentary basin. Left: general overview of the solution containing the best estimated model together with the true model and residuals (red dots for the best and black dots for the true model), and the uncertainty region composed by all generated models with relative misfit below 10% (51 612). Right: empirical cumulative density function and histogram for depth of rectangle number 54 in the profile.

The uncertainty analysis for this kind of problem is sometimes absent in the literature or, in most cases, restricted to the comparison of the inversion results with other information such as previous models, data from boreholes, and/or other geological information. In other cases (see, for example, [22, 23]) it is performed by computing the standard deviation for each rectangle, taking into account all models with misfit below the prescribed tolerance, i.e., using the range of variation depicted in Figure 3 (left). On the other hand, [25] makes the uncertainty assessment via principal component analysis, building the 2D cost function topography corresponding to the two principal components of the solutions. Other authors make cross-plots between different unknowns in order to see their correlation, such as in [34], although in problems with a high number of parameters it may not be practical due to the high number of combinations between them (in problems involving few unknowns it can be a valid option, such as in [35]).

4. Determining the equivalence region containing the best model

As was stated in the previous section, yellow bars in Figure 3 (left) show the limits between each rectangle depth varies for all the models below 10% relative misfit tolerance. Nevertheless, this set can contain models belonging to several disjoint equivalence regions, as we are dealing with a nonlinear problem. This can be proven by the fact that in our example, the models composed of the ensembles of all upper and lower limits of the yellow bars have relative misfits of 31.73% and 33.64%, greater than the tolerance of 10%. However, in order to depict correctly an estimation of the equivalence region of 10% relative misfit where the global minimum is contained, we must extract only the models belonging to this region. We have developed an algorithm in order to do so, based on the information provided by the empirical cumulative density functions previously computed for each rectangle.

The algorithm works on the ECDF functions computed for each rectangle using the models below the predefined tolerance (51 612 models in our exercise). It can be summarized in the following steps (we search for the upper bound depths, but in the case of lower bound the framework is the same except for minor changes in some comparisons):

(1) Select all models of relative misfit around the objective equivalence region: for example, between 9% and 11% if we are looking for the 10% region. From these models, select those completely above the best estimated model.

(2) Compute the Euclidean distance from each of the previous selected models to the estimated best model and select a small set (10 models, for example) of the closest ones. Compute the mean of such models in order to obtain a first approximation to the upper bound of the equivalence region, `upper_bound`. The misfit of this model will not be exactly the searched value. We can use the closest model as first approximation, but experience demostrated that the algorithm is more efficient by using the described way.

(3) Extract from ECDFs the probability of depth for each rectangle in the previous upper bound candidate model.

(4) Add a small probability increment, `IncProb` (experience has shown that 0.01 is a good value), to the extracted probabilities and interpolate in the ECDFs their corresponding depths. We have a new `upper_bound` candidate. Compute its new misfit.

(5) Start an infinite loop while the misfit of the upper bound is different than the misfit of the equivalence region to search (in practice, a small tolerance must be employed, as it is almost impossible to get an upper bound misfit identical to any prescribed value).

  (a) Update the value of `IncProb` to `IncProb/1.25` (1.25 is a good value, again estimated by experience, after performing a lot of inversions in very different environments and using very different numbers of observations and basin models) in order to gradually get closer to algorithm convergence. If `IncProb` takes a very small value (10^{-8} for example), its original value is restored.

  (b) If the misfit of `upper_bound` is lower than the equivalence region to search, we add again the `IncProb` value to the candidate model. If it is greater, we subtract the `IncProb` value. We interpolate new depth values with the updated probability values. Increments and decrements of probability do not mean increments or decrements of the same values in depth, as each rectangle has its own empirical cumulative distribution function.

  (c) The `upper_bound` model must have for each rectangle a depth value lower than that

corresponding to the same rectangle in the best model, so if any of them is deeper, the value of the best model is assigned.

(d) Compute the new misfit of `upper_bound`.

It must be stressed that taking different values than the presented in the previous algorithm description for the `IncProb` parameter and its updating factor does not mean that the algorithm fails, but rather that it will execute less efficiently. These values, as was explained, come from experience and must be understood as suboptimal ones. The pseudocode presented in Algorithm 1 shows the described steps in order to determine the depths of the upper bound.


Algorithm 1: Upper bound of equivalence region containing the best model.

Data: equivalent models, ECDFs, equivalence region to search

Extract first candidate for upper bound and the probabilities of their depths;

Add probability increment and compute new upper bound depths;

Compute misfit for new depths;

while *misfit of upper bound different from misfit of the equivalence region to search* **do**

    Update probability increment;

    if *misfit of upper bound lower than misfit of the equivalence region to search* **then**

        Add probability increment;

    else

        Subtract probability increment;

    end

    Compute new depths with the new probability values;

    Check depths according to best model;

    Compute new misfit for the updated upper bound model;

end

Figure 4 shows the same synthetic example as in Figure 3 after applying the described algorithm. We can see (left panel) how the 10% equivalence region is now narrower than the one in the previous example, as it corresponds to only one region, the one where the best model is included. Their upper and lower bounds have now a relative misfit of 10%, and any other model inside these bounds has a misfit lower than this value. From the original 51 612 models with relative misfit lower than 10%, only 16 194 are totally inside the bounds, which means that the remaining 35 418 models belong to other equivalent regions containing local minima of the cost function, as they are not fully contained in the bounds (this is a proof of the excellent exploratory character of the CP-PSO family member employed in the inversion). In the lower right panel of Figure 4, we can see the histogram of rectangle 54, where we can appreciate that the range in depths has diminished from approximately $[925\,\text{m}, 1650\,\text{m}]$ in Figure 3 to approximately $[1220\,\text{m}, 1400\,\text{m}]$ (if we take into account the 80% most frequent depths, the ranges change from $[1190\,\text{m}, 1450\,\text{m}]$ to $[1290\,\text{m}, 1340\,\text{m}]$). Also, the histogram is not as symmetrical as the one in Figure 3, showing that in nonlinear inverse problems the minimum of the cost function is not placed in general at the geometrical center of the equivalence region (see also Figure 1).
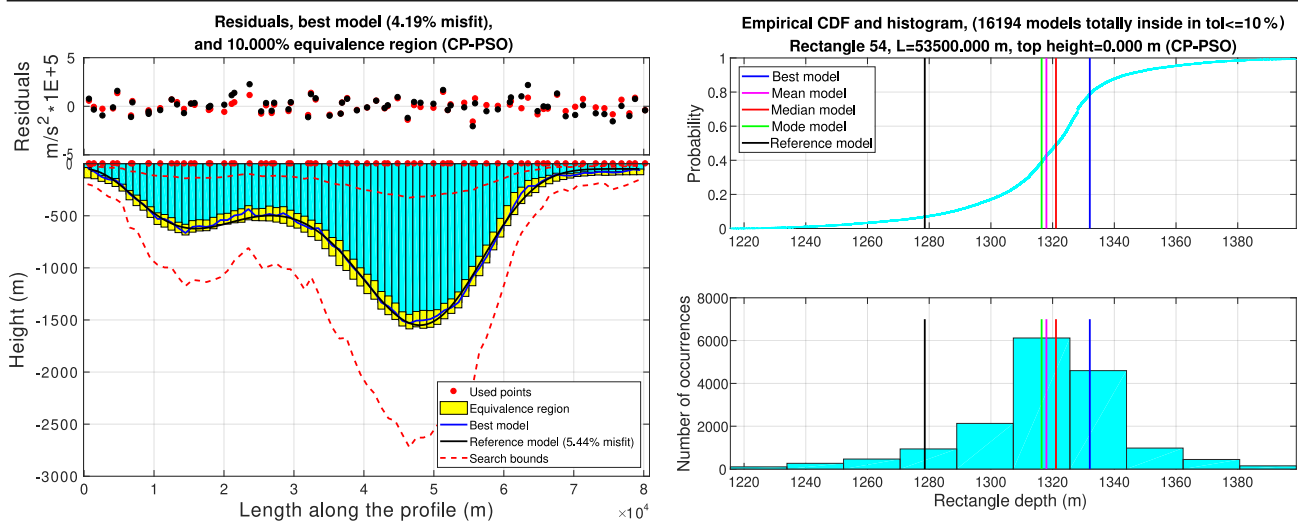
Figure 4. Synthetic example of a sedimentary basin. Left: general overview of the solution containing the best estimated model together with the true model and residuals (red dots for the best and black dots for the true model), and the equivalence region of 10% relative misfit, which encloses entirely 16 194 of the 51 612 models below the 10% tolerance. Right: empirical cumulative density function and histogram for depth of rectangle number 54.

5. Real example: profile inversion in the Argelès-Gazost basin

In this section, we apply our methodology to a gravity profile in the Argelès-Gazost basin, an ancient glacial valley located in the French Pyrenees that is now occupied by the Gave de Pau River. In this area, formed by a glacier that belonged to a larger ice tongue overlying this Pyrenean region in the Quaternary period, various gravity surveys and inversions have been carried out in order to determine the geometry of the basement relief [36–38]. A 3D inversion using PSO was also carried out previously in [39], and in this work we have selected a profile crossing the deepest area in order to check the presented methodology, as well as to compare the 3D and 2D approaches.

The Argelès-Gazost basin is a flat and elongated valley located at an elevation of about 450 m above mean sea level, filled by Quaternary deposits and surrounded by limestones and flysches [38, 40]. It has a length of about 7 km along a NNW direction and a width varying between 1 km and 2 km (see Figure 5, left). Its deepest area is located in the southern part, between the towns of Argelès-Gazost and Pierrefitte-Nestalas. In this area, there are 9 gravity observations almost aligned forming a profile about 1600 m long, which we have used in the inversion (Figure 5, left, red line). The profile is located in such a way that the longest dimension of the basin is perpendicular to it, so this situation is optimal for a 2D inversion. In order to perform the inversion as close as possible to [39], we have used the same value for density contrast, $\Delta\rho = -600\,\mathrm{kg\,m^{-3}}$ (also the same value as the one employed by Moussirou in [36] and Perrouty in [37, 38]). As in [39], a regional trend for the gravity anomalies was modeled using a plane, and we have applied the same model for our profile in this experiment in order to obtain the same residual anomaly as for the 3D inversion. The basin was modeled as a set of 12 rectangles of 150 m width, and the CP-PSO family member was employed in the inversion using a population of 600 models and 400 iterations, so a total of 240 000 models was generated.
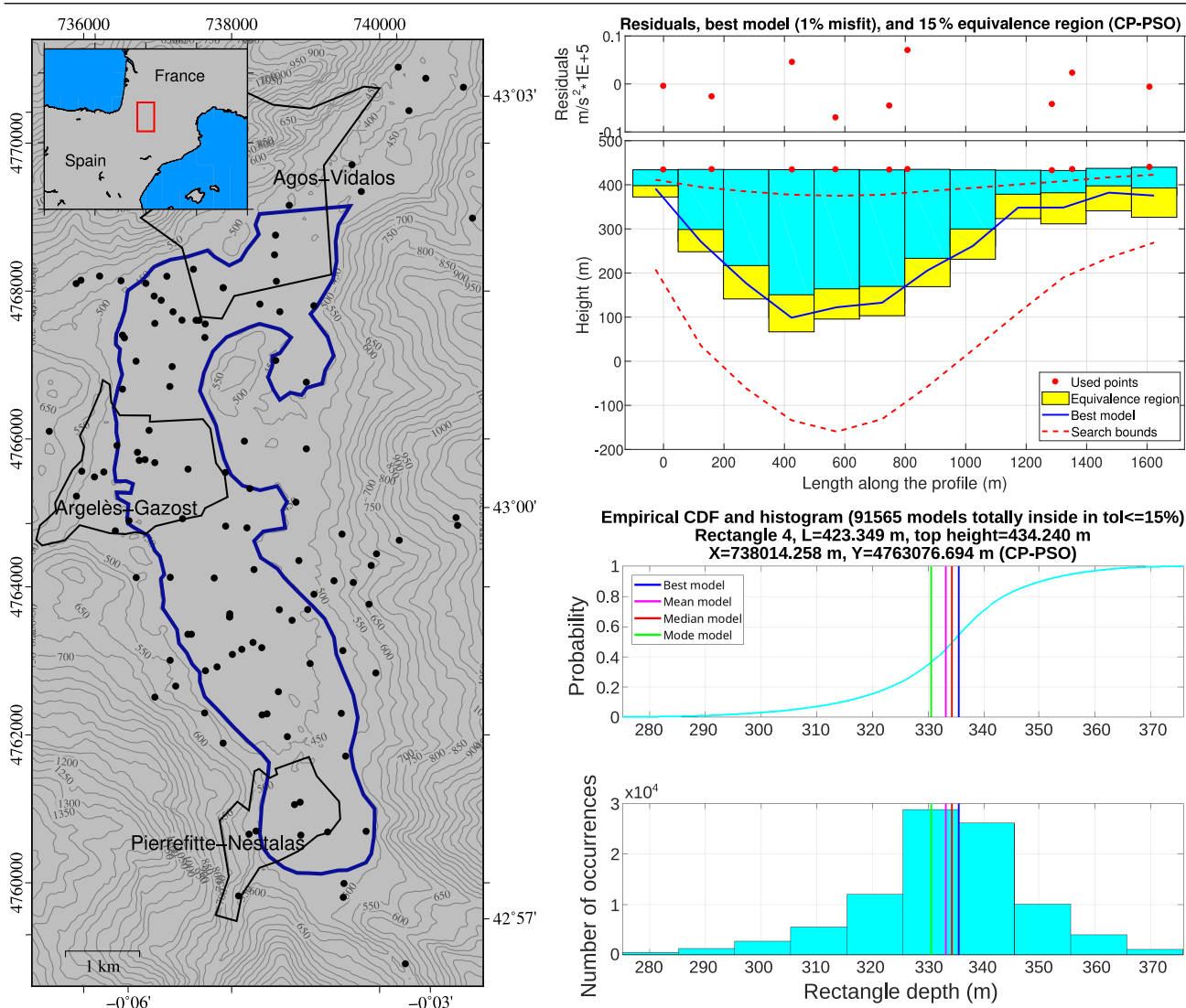
Figure 5. Left: Argelès-Gazost basin localization and gravity observations (black dots) employed in [39]. In this study a profile of 9 points crossing the deepest area (red line), between Argelès-Gazost and Pierrefitte-Nestalas, is used in the inversion. Right: best model and 15% equivalence region (upper), and ECDF and histogram for rectangle number 4 (lower).

Figure 5 (right) shows the results of the inversion for the profile. The best model reached a relative misfit of 1.00%, which is lower than the misfit obtained in [39] for the 3D inversion (9.72%). This was expected, as the 3D inversion comprised the whole basin, and our profile is small and has a very good concentration of observations. About the deepest point in the profile, it corresponds to rectangle number 4 and reaches a depth of 335 m for the best model with a range of [280 m, 370 m] for the 15% equivalence region (the range [315 m, 350 m] comprises 80% of models). These values are almost the same as those obtained in the 3D inversion, where the maximum estimated depth was 333 m with a range of [288 m, 379 m] for the same 15% equivalence region. A total of 91 565 models are inside the 15% equivalence region containing the best model, indicating the good explorative character of the inversion, as the remaining 148 435 models explored the rest of the parameter space.

Moussirou obtained a maximum depth value of 300 m in [36] using a 2.5D modeling based on [41], while Perrouty in [38] obtained a value of 232 m using VPmg inversion (*Virtual Prism magnetic and gravity inversion*), [42]. The value of 300 m by Moussirou lies in our depth range estimation for the 15% equivalence region, although the depth computed by Perrouty is outside this region. This is not due to the inversion method but due to the regional trend elimination in the gravity signal, as in [38] a 3rd degree surface was applied to model the trend, absorbing part of the residual Bouguer signal. As a consequence, the amplitude of the residual anomaly obtained in [38] reaches only −4.2 mGal, while in this work and in [39], where a plane was set to model the regional trend, the amplitude reaches −5.98 mGal, reflecting a deeper sediments-basement interface.

6. Conclusions

Uncertainty analysis is a crucial task in any type of inverse problem. Due to the nature of the cost function topography in nonlinear problems, where multiple minima exist (in cases involving potential fields, there are infinitely many global minima) and the equivalence regions have arbitrarily curved shapes, local optimization methods based on the linearization of the forward operator lack the ability to produce precise uncertainty analysis of the results. Global optimization methods such as particle swarm optimization offer an excellent alternative to local ones, as they explore the entire parameter space and are not constrained by any initial model or by the linearization of the problem. They provide a collection of models, which is valuable information that must be carefully examined to perform uncertainty analysis.

These models generally belong to various equivalence regions of the cost function, each of them with its local minimum, but our interest lies in the one containing the global minimum. Conducting statistics with the entire set of equivalent solutions can lead to erroneous results as models from different equivalence regions may be mixed. Therefore, a methodology for selecting the correct models, regardless of the inverse problem being addressed, must be developed. We have proposed an algorithm tailored for a gravity inverse problem to perform this selection based on the analysis of the empirical cumulative distribution function (ECDF) of each parameter. The proposed algorithm works for the gravity inversion problem in sedimentary basins and improves the efficiency of equivalence region detection compared to our previous version of the software. However, this does not imply that it is suitable for other problems, which must be carefully analyzed to develop the most appropriate procedure. We also emphasized the importance of analyzing not only the overall view of the solution but also the ECDF and the histogram of each individual parameter to illustrate its particular distribution.

Finally, a real example was presented where a profile in the Argelès-Gazost basin was inverted, analyzed using the proposed methodology, and compared with previous results. The 2D inversion produces a result fully compatible with the 3D inversion previously performed, demonstrating that, under suitable conditions, the 2D approach to the gravity inverse problem applied to sedimentary basins is as feasible as 3D inversion.

The proposed algorithm for uncertainty analysis in the two-dimensional gravity inverse problem applied to sedimentary basins has been implemented in versions ≥ 2 of GravPSO2D, which can be downloaded, together with a detailed 54-page user manual, from the web page of the Bureau Gravimétrique International: `https://bgi.obs-mip.fr/catalogue/?uuid=9ce23226-bb75-462e-bd60-5997a768a359`.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Conflict of interest

Prof. M. Z. Fernández-Muñiz is the Guest Editor of special issue "New insights of the Application of Inverse Problems and Machine Learning in Science and Technology" for AIMS Mathematics. Prof. M. Z. Fernández-Muñiz was not involved in the editorial review and the decision to publish this article.
The authors declare no conflicts of interest.

References

1. A. G. Camacho, P. Vajda, J. Fernández, GROWTH-23: an integrated code for inversion of complete Bouguer gravity anomaly or temporal gravity changes, *Comput. Geosci.*, **182** (2023), 105495. https://doi.org/10.1016/j.cageo.2023.105495

2. S. Liu, M. Liang, X. Hu, Particle swarm optimization inversion of magnetic data: field examples from iron ore deposits in China, *Geophysics*, **83** (2018), J43–J59. https://doi.org/10.1190/geo2017-0456.1

3. A. G. Yeguas, J. Almendros, R. Abella, J. M. Ibáñez, Quantitative analysis of seismic wave propagation anomalies in azimuth and apparent slowness at Deception Island volcano (Antarctica) using seismic arrays, *Geophys. J. Int.*, **184** (2011), 801–815. https://doi.org/10.1111/j.1365-246X.2010.04864.x

4. W. W. Symes, The seismic reflection inverse problem, *Inverse Probl.*, **25** (2009), 123008. https://doi.org/10.1088/0266-5611/25/12/123008

5. O. Bar, G. Even-Tzur, Fault-model parameters estimation using a feature-voting technique: Dead Sea fault as a case study, *Comput. Geosci.*, **179** (2023), 105422. https://doi.org/10.1016/j.cageo.2023.105422

6. M. Á. Benito-Saz, F. Sigmundsson, M. Charco, A. Hooper, M. Parks, Magma flow rates and temporal evolution of the 2012–2014 post-eruptive intrusions at El Hierro, Canary Islands, *J. Geophys. Res.: Solid Earth*, **124** (2019), 12576–12592. https://doi.org/10.1029/2019JB018219

7. A. N. Tikhonov, V. Y. Arsenin, *Solutions of ill-posed problems*, Washington, DC: V. H. Winston and Sons, 1977.

8. R. J. Blakely, *Potential theory in gravity and magnetic applications*, Cambridge University Press, 1996.

9. M. S. Zhdanov, *Inverse theory and applications in geophysics*, 2 Eds., Oxford: Elsevier, 2015.

10. J. L. Fernández-Martínez, J. L. G. Pallero, Z. Fernández-Muñiz, L. M. Pedruelo-González, The effect of noise and Tikhonov's regularization in inverse problems. Part II: the nonlinear case, *J. Appl. Geophys.*, **108** (2014), 186–193. https://doi.org/10.1016/j.jappgeo.2014.05.005

11. R. L. Parker, Understanding inverse theory, *Ann. Rev. Earth Planet. Sci.*, **5** (1977), 35–64. https://doi.org/10.1146/annurev.ea.05.050177.000343

12. R. C. Aster, B. Borchers, C. H. Thurber, *Parameter estimation and inverse problems*, 2 Eds., New York: Academic Press, 2012.

13. S. Katoch, S. S. Chauhan, V. Kumar, A review on genetic algorithm: past, present, and future, *Multimed. Tools Appl.*, **80** (2020), 8091–8126. https://doi.org/10.1007/s11042-020-10139-6

14. S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by simulated annealing, *Science*, **220** (1983), 671–680. https://doi.org/10.1126/science.220.4598.671

15. J. Kennedy, R. Eberhart, Particle swarm optimization, *Proceedings of ICNN'95–International Conference on Neural Networks*, 1995, 1942–1948. https://doi.org/10.1109/ICNN.1995.488968

16. A. Tarantola, Popper, Bayes and the inverse problem, *Nature Phys.*, **2** (2006), 492–494. https://doi.org/10.1038/nphys375

17. J. L. Fernández-Martínez, M. Z. Fernández-Muñiz, M. J. Tompkins, On the topography of the cost functional in linear and nonlinear inverse problems, *Geophysics*, **77** (2012), W1–W15. https://doi.org/10.1190/geo2011-0341.1

18. J. A. Scales, R. Snieder, To Bayes or not to Bayes, *Geophysics*, **62** (1997), 1045–1046. https://doi.org/10.1190/1.6241045.1

19. J. A. Scales, R. Snieder, The anatomy of inverse problems, *Geophysics*, **65** (2000), 1708–1710. https://doi.org/10.1190/geo2000-0001.1

20. Z. Fernández-Muñiz, J. L. G. Pallero, J. L. Fernández-Martínez, Anomaly shape inversion via model reduction and PSO, *Comput. Geosci.*, **140** (2020), 104492. https://doi.org/10.1016/j.cageo.2020.104492

21. J. L. G. Pallero, J. L. Fernández-Martínez, Z. Fernández-Muñiz, S. Bonvalot, G. Gabalda, T. Nalpas, GᴿᴀᴠPSO2D: a Matlab package for 2D gravity inversion in sedimentary basins using the Particle Swarm Optimization algorithm, *Comput. Geosci.*, **146** (2021), 104653. https://doi.org/10.1016/j.cageo.2020.104653

22. A. Jamasb, S. H. Motavalli-Anbaran, H. Zeyen, Non-linear stochastic inversion of gravity data via quantum-behaved particle swarm optimisation: application to Eurasia-Arabia collision zone (Zagros, Iran), *Geophys. Prospect.*, **65** (2017), 274–294. https://doi.org/10.1111/1365-2478.12558

23. A. Jamasb, S. H. Motavalli-Anbaran, K. Ghasemi, A novel hybrid algorithm of particle swarm optimization and evolution strategies for geophysical non-linear inverse problems, *Pure Appl. Geophys.*, **176** (2019), 1601–1613. https://doi.org/10.1007/s00024-018-2059-7

24. S. Loni, M. Mehramuz, Gravity field inversion using improved particle swarm optimization (IPSO) for estimation of sedimentary basin basement depth, *Contrib. Geophys. Geod.*, **50** (2020), 303–323.

25. A. Roy, C. P. Dubey, M. Prasad, Gravity inversion of basement relief using particle swarm optimization by automated parameter selection of Fourier coefficients, *Comput. Geosci.*, **156** (2021), 104875. https://doi.org/10.1016/j.cageo.2021.104875

26. L. L. Nettleton, *Gravity and magnetics in oil prospecting*, International Series in the Earth and Planetary Sciences, McGraw-Hill, 1976.

27. M. H. P. Bott, The use of rapid digital computing methods for direct gravity interpretation of sedimentary basins, *Geophys. J. R. Astron. Soc.*, **3** (1960), 63–67. https://doi.org/10.1111/j.1365-246X.1960.tb00065.x

28. V. Chakravarthi, Gravity interpretation of nonoutcropping sedimentary basins in which the density contrast decreases parabolically with depth, *Pure Appl. Geophys.*, **145** (1995), 327–335. https://doi.org/10.1007/BF00880274

29. J. B. C. Silva, W. A. Teixeira, V. C. F. Barbosa, Gravity data as a tool for landfill study, *Environ. Geol.*, **57** (2009), 749–757. https://doi.org/10.1007/s00254-008-1353-6

30. J. B. C. Silva, A. S. Oliveira, V. C. F. Barbosa, Gravity inversion of 2D basement relief using entropic regularization, *Geophysics*, **75** (2010), I29–I35. https://doi.org/10.1190/1.3374358

31. M. B. Dobrin, C. H. Savit, *Introduction to geophysical prospecting*, 4 Eds., Geology Series, McGraw-Hill, 1988.

32. V. C. F. Barbosa, J. B. C. Silva, Generalized compact gravity inversion, *Geophysics*, **95** (1994), 57–68. https://doi.org/10.1190/1.1443534

33. J. L. Fernández-Martínez, E. García-Gonzalo, The PSO family: deduction, stochastic analysis and comparison, *Swarm Intell.*, **3** (2009), 245–273. https://doi.org/10.1007/s11721-009-0034-8

34. O. F. Ladino, A. Bassrei, A hybrid fast 3D inversion algorithm of gravity data for basement relief definition, *SEG Technical Program Expanded Abstracts 2016*, 2016, 1521–1525. https://doi.org/10.1190/segam2016-13840764.1

35. A. Singh, A. Biswas, Application of global particle swarm optimization for inversion of residual gravity anomalies over geological bodies with idealized geometries, *Nat. Resour. Res.*, **25** (2016), 297–314. https://doi.org/10.1007/s11053-015-9285-9

36. B. Moussirou, *Quantification du remplissage quaternaire des vallées glaciaires des Pyrénées par la méthode gravimétrique*, MS. Thesis, Université Toulouse III–Paul Sabatier, 2013.

37. S. Perrouty, *Mesures géophysiques du remplissage sédimentaire Quaternaire dans les vallées de Bagnères de Bigorre et d'Argelès-Gazost (Hautes-Pyrénées)*, MS. Thesis, Université Toulouse III–Paul Sabatier, 2008.

38. S. Perrouty, B. Moussirou, J. Martinod, S. Bonvalot, S. Carretier, G. Gabalda, et al., Geometry of two glacial valleys in the northern Pyrenees estimated using gravity data, *C. R. Geosci.*, **347** (2015), 13–23. https://doi.org/10.1016/j.crte.2015.01.002

39. J. L. G. Pallero, J. L. Fernández-Martínez, S. Bonvalot, O. Fudym, 3D gravity inversion and uncertainty assessment of basement relief via particle swarm optimization, *J. Appl. Geophys.*, **139** (2017), 338–350. https://doi.org/10.1016/j.jappgeo.2017.02.004

40. M. H. Alimen, *Le quaternaire des Pyrénées de la Bigorre*, Imprimerie nationale, 1964.

41. J. W. Cady, Calculation of gravity and magnetic anomalies of finite length-right polygonal prisms, *Geophysics*, **45** (1980), 1507–1512. https://doi.org/10.1190/1.1441045

42. P. K. Fullagar, N. A. Hughes, J. Paine, Drilling-constrained 3D gravity interpretation, *Explor. Geophys.*, **31** (2000), 17–23. https://doi.org/10.1071/EG00017

AIMS Press