# Chapter 7

# Some observations on the role of noise in the inversion of seismic and gravity data

As it has already been pointed out in the rest of the document geophysical inverse problem may be reduced to determining the set(s) of parameters that generate the minimum misfit between some observed and calculated data. At least four kind of errors may limit the accuracy of any inversion process: instrumental noise, inappropriate parameterisation, inexact forward modelling and inability of the inversion routine to detect the global minimum in the search space. While the last point has been the object of a large amount of investigation both in the literature and in this research, the effect of the first three sources of errors is often overlooked. It is very difficult, and often impossible, to discriminate between instrumental noise, inappropriate parameterisation and inexact forward modelling but their overall effect is to prevent the forward model to accurately mimic the physics of the problem under analysis and accordingly to reproduce the observed data.

In this chapter I report on a series of computer experiments I carried out to understand the effect of such errors in the overall inversion process in two different geophysical problems, seismic refraction tomography and inversion of gravity data. In order to have a clear control on the phenomena under analysis a series of synthetic experiments have been performed. In such test random noise

is added to the synthetic data in order to create the mismatch between observed and calculated data that would be present in a real experiment.

I found that in seismic refraction problems the effect of such errors is to prevent the inversion routine obtaining the 'correct' model parameter set. The phenomenon does not depend on the ability of the inversion routine to detect the global minimum in the solution space. This is due to the fact that under the effect of noise the 'correct' parameter set no longer corresponds to the global minimum. Accordingly, this problem goes beyond the choice of the 'best' inversion routine and it deals directly with the limitation imposed by the physics under analysis. Methods to identify the areas of the solution more sensitive to such errors are proposed that can be practically used on field data to assess the reliability of a geological reconstruction. However, better results may be achieved only by including further information in the problem addressed.

I also found that the effect of errors is much less marked in the inversion of gravity data. Accordingly, information from gravity data sets may be added to the inversion of seismic refraction data in order to limit the effect of noise. A method to perform a staged joint inversion of seismic and gravity data is discussed which proved to be successful in two inversion experiments characterised by different refraction geometry. Also, it is shown that the computational effort of such joint inversion involves only a small increase over the single inversion of seismic data.

The content of this chapter has been recently submitted to GEOPHYSICAL JOURNAL INTERNATIONAL and it is currently under review.

## 7.1 Introduction

In order to introduce the ideas presented in this chapter is probably useful to first briefly review the basic concepts in the inversion process. In a geophysical inverse problem we are seeking to determine the main features of the subsurface geological structure from measurements (seismic signals, gravity, magnetic, electrical data, etc.) related to the physical properties of the earth. In this process the unknown geological structure is usually modelled by a number of regularly shaped blocks whose shape and/or position and/or physical properties are the model parameters

to be determined.

For most geophysical problems no method is currently available to directly reconstruct the unknown geology from field measurements, i.e., no explicit mathematical formula is available to directly manipulate the observed data and give information about the unknown geology as output. Rather, a geophysical problem is solved by iteratively perturbing the model parameters in a trial and error fashion, until a parameter set(s) satisfactorily reproduces the observed data. Such a process is referred to as inversion. In order to perform inversion ways to calculate the expected physical response of different parameter sets are needed. This expected physical response is then compared to the observed data. This is obtained by the use of forward algorithms, whereby the physical process under analysis (seismic wave propagation, generation of magnetic anomalies, etc..) is modelled by the use of standard continuum equations.

Given these assumptions, the actual mathematical problem of reconstructing the unknown geology becomes a search, in a multi-dimensional parameter or solution space, for the set(s) of model parameters which give a global minimum to some function of the difference between the model response and the observed data, often referred to as solution surface.

Solving this problem is achieved by the use of inversion algorithms which basically direct the search of the solution space for the global minimum by iterative use of the forward model, i.e. by iteratively sampling different areas of the solution space.

We must be aware that errors are introduced at each stage of this process. At least four sources of errors may be recognised. First, real data contain instrumental noise due to the acquisition system. A source of errors closely connected with this problem is the accurate detection of first arrivals in seismic traces that has been addressed in Chapter 2. Secondly, the model parameterisation is inherently inaccurate. This is because the complexity of a real geological structure can not be reproduced by subdividing the region under analysis into a number of blocks for which regular shapes and some kind of homogeneity are assumed. As an example the behaviour of the real earth in a seismic experiment can only be modelled approximately by homogeneous, isotropic and elastic solid over the dimension of a

model block. Also, we must remember that the earth is a 3-D body which in many experiments is treated as a 2-D section. Similarly, approximations are present in the equations employed in the forward modelling, from which a third source of errors arise. For example, in the seismic refraction experiments described in this study, the seismic energy is supposed to travel along paths approximated by straight line segments. This often fails to model the more complex behaviour of seismic wave propagation in complex media. The possible inability the inversion routine to correctly find the global minimum in the search space represents the fourth source of errors. This has been extensively addressed in Chapter 5 and Chapter 6. Note that while this error is directly connected with the inversion algorithm, the other three error sources directly affect the forward model and their effect in the overall inversion process is more subtle.

While great effort is currently put into understanding the inner mechanics of different inversion algorithms the key role played by the forward model in the overall inversion process is often overlooked. Firstly, the shape of the solution surface depends, among other factors, on the physics of the problem, that is contained in the forward model. Secondly, and most important, if the forward model is not accurate (i.e., does not adequately model the physics of the problem, as discussed above) the result of the inversion procedure will not reconstruct the correct parameter configuration and consequently will not accurately describe the unknown geology.

Usually the efficiency of an inversion procedure is first tested with synthetic data. In this case the same algorithm is used as data generator and as forward routine in the inversion and errors are not present in the process. This can be seen as if the inversion procedure was supplied with a forward routine able to perfectly model the physics of the problem and accordingly to perform a perfect mapping from the parameter space to the solution space (i.e., the correct parameter configuration corresponds to the global minimum in the solution space and has zero misfit). Afterwards, when sufficient mathematical insight into the problem has been acquired, the inversion of real data is attempted. In this case all the sources of errors listed above are present which affect the efficiency of the inversion process.

It is very difficult to discriminate between the effects of instrumental noise, in-

appropriate parameterisation and inexact forward modelling. Such discrimination is beyond the scope of this chapter. In the rest of the discussion I will use the term 'noise' for the overall combination of these errors, i.e., I define as 'noise' everything that is not modelled, or is not accurately modelled, by the forward model.

The aim of this chapter is threefold. The first goal is to obtain a better understanding of the effect of noise in the inversion process. This is attempted on two different geophysical problems, seismic refraction tomography and inversion of gravity data. The presence of inaccuracies in the process are modelled by adding random noise to a set of synthetic data and their effect is discussed by analysing the modification they impose on the solution shapes in both problems.

Since an exhaustive analysis of a solution space in a high-dimensional space in beyond current computation possibilities, the second goal of this study is to propose practical tools to identify the areas of the solution more sensitive to ambiguity and/or errors in the inversion process.

The last goal of this research is to propose a method to improve the results of the inversion of seismic and/or gravity data in presence of noise. The two inverse problems here presented show very different features. Gravity inverse problems are characterised by large ambiguity (i.e., a large number of solutions can be found that fit the data equally well) but are not very sensitive to the presence of noise in the data. Seismic refraction inverse problems are less ambiguous but show a stronger sensitivity to noise. These opposite behaviours suggest that improvements in the inversion of such data sets could be obtained by combining the two processes. The advantages of such joint inversion and its potentiality is shown on two synthetic problems in which noise is included in order to model inaccuracies in the forward modelling that could arise in real experiments.

## 7.2 Seismic refraction tomography

### 7.2.1 Previous analysis

Seismic refraction tomography experiments have been performed and extensively described in Chapter 5. Here some of the results are summarised and further analysed in order to obtain a better understanding of the problematic inherent in

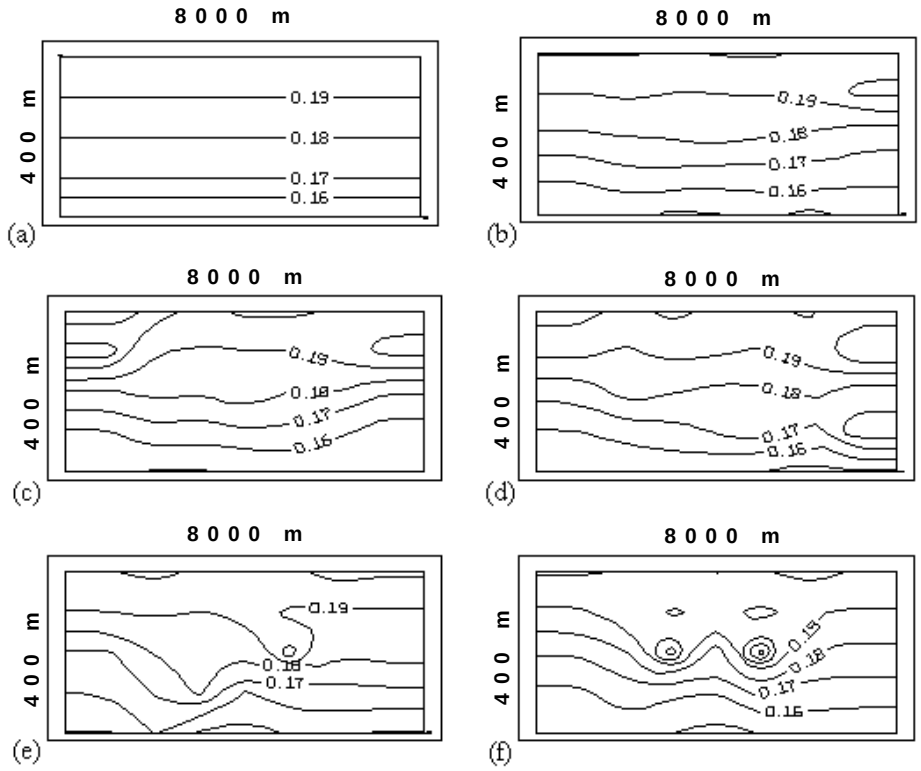


Figure 7.1: Synthetic test simulating a horizontally layered model with linearly varying slowness in the vertical direction. The synthetic image used to generate the data set is presented in Figure 7.1a. Figure 7.1b-f show the results from 5 runs with a Genetic Algorithm initialised with different random populations. Figure 7.1b-d show a satisfactory resemblance with the synthetic image. In Figure 7.1e-f errors are present that are positive in sign.

such an inverse problem.

Figure 7.1 shows one of the test cases discussed in Chapter 5. Figure 7.1a represents the horizontally layered model with linearly varying slowness in the vertical direction that was used to generate the synthetic data set to invert. As it has already been shown in Chapter 5, from a geological perspective the results in Figure 7.1b-d show a satisfactory resemblance with the synthetic image. The main feature in the model, i.e. the horizontal layering, is well reconstructed. Departures from the synthetic model are present only on the lateral areas of the domain, i.e. in areas badly sampled by ray-tracing routine. However, in Figure 7.1e-f errors are present that are indicated by anomalously large values of slowness at some

grid nodes. In a geological interpretation these errors could suggest the presence of small anomalous bodies in the domain. Clearly understanding the sources for such erroneous reconstructions is crucial in assessing the method potential for exploration purposes. It can be noticed that when anomalous reconstructions are contained in my Genetic Algorithm solutions, these are always located in the central area of the domain just above the refractor. In Chapter 5 it was showed that by improving the Genetic Algorithm solutions with the use of local optimisers (in particular the such errors are removed and solutions very close to the global minimum are obtained. However, it has also been shown that improvements to the Genetic Algorithm solution through the use of local optimisers can be obtained only by a very small decrease in the error misfit. On noisy data sets the refinements on the error misfit required by the local optimiser may fall well below the limitations imposed by the presence of noise. In such circumstances the local optimisation of the Genetic Algorithm solution will not be advantageous. Accordingly, understanding the reason behind the presence of such errors and their spatial location in the solution is crucial for the following discussion. This will now be addressed.

### 7.2.2 Discussion on error sources

The accuracy with which the slowness value may be determined at a node position depends on the ray coverage in that area. The more rays that cross a certain region the more weight the value at such node has on the data misfit, the more sensitive the inversion procedure is to slowness value at such a node. A numerical measure of such weight may be represented by the amount of variation in the error misfit as a function of a small perturbation in the slowness value at such node. Since the actual derivative of the error misfit in regard with the slowness values can not be analytically obtained with the use of the approximate ray-tracing routine employed in this study, I calculated the variation of the error misfit for a small slowness perturbation in both signs and averaged the two values at each node location, e.g. at a particular node with slowness $S$ I calculated the misfit for $S \pm \epsilon$ and averaged the misfit. Thus a small average indicates the inversion is not very sensitive to the errors in the slowness at this node. The averages so obtained are then plotted at each node location and in the rest of the chapter I will refer to

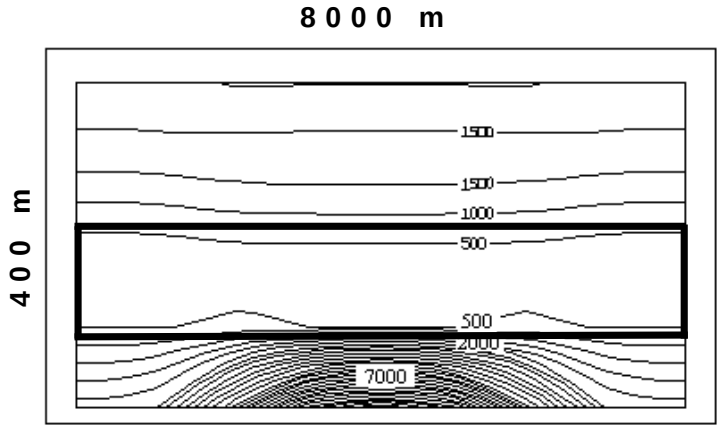


Figure 7.2: Sensitivity plot for the synthetic test in Figure 7.1a, i.e., the average misfit error as a function of a small slowness perturbations in both sign is calculated and plotted at each node location. As we can see a vast minimum is present in correspondence with the lower-central area of the domain.

such diagram as 'sensitivity plot'.

The sensitivity plot for correct solution to the inversion problem presented in Figure 7.1 is shown in Figure 7.2. Notice that the absolute values in the diagram depend on a number of factors, such as the number of seismic rays modelled, the physical dimensions of the calculation domain, the magnitude of the slowness perturbation etc. However, it is not the absolute values but their relative variation that is relevant to this discussion. The sensitivity plot presents a large low in the central area of the picture, just above the refractor. This area is highlighted by the rectangular box in the figure. The consequence of the presence of this low is that an inaccurate reconstruction of slowness values at nodes in this part of the image does not affect the error misfit to a large extent. This creates a flat area in the solution space difficult to be resolved by the Genetic Algorithm. High values in the sensitivity plot can be found close to the surface where rays run along the surface or dive just below it between closely spaced source-receiver couples. The large number of rays running along the refractor are responsible for the 'high peak' present at the bottom of the picture. Accordingly, the inversion procedure is very sensitive to slowness values in these areas where correct values should appear very early in the inversion.

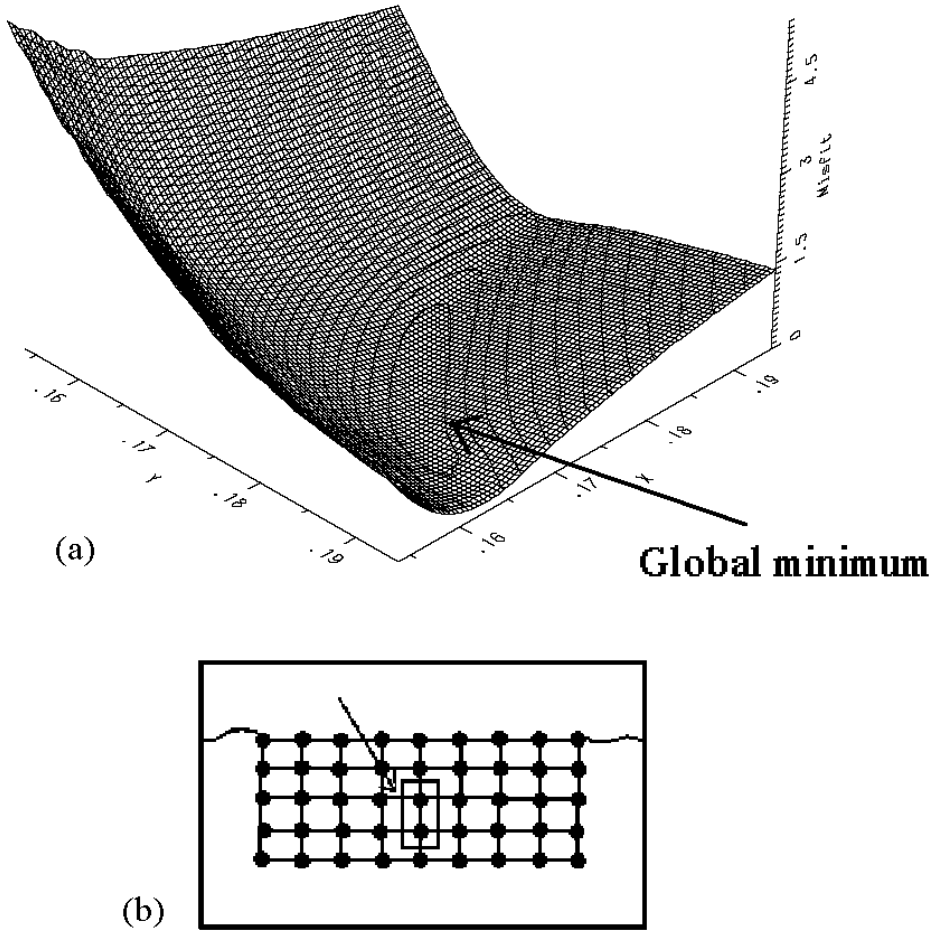


Figure 7.3: (a) Solution surface for the noise-free seismic data set in the 2 dimensional subspace correspondent to the two nodes highlighted in Figure b. A curved valley converging towards a single global minimum can be seen. Also notice the different slope in the two flanks of the valley.

Also, as already mentioned in the previous section, the nodes in error in the solutions in Figure 7.1 are always characterised by slowness values larger than the corresponding values in the synthetic image. To understand why this occurs consider the two nodes enclosed in the rectangle in Figure 7.3b. While the slowness in the rest of the domain is kept fixed the slowness values at such nodes are varied. The error misfit calculated as the sum of the squared errors between the synthetic and the observed traveltimes due to such perturbed values is plotted in the 3-D image in Figure 7.3a. This represents the solution surface for the small two dimensional subspace corresponding to the nodes under analysis. Notice that the

two nodes are located inside the area delimited by the rectangular box in Figure 7.2. Such misfit is obviously zero at the global minimum, i.e., for slowness values equal to the ones in the synthetic image. Such values are 0.17 s/km for the X axis and 0.18 s/km for the Y axis. The surface numerical values may vary for different nodes combinations but the pattern remains similar for the nodes corresponding to central region in Figure 7.2.

Figure 7.3a shows the curved shape of the valley converging towards a single global minimum. More important, it can be noticed that the valley flanks have very different slope, i.e., for slowness values smaller that the correct ones the error misfit varies sensibly, while an equivalent perturbation in the opposite direction creates a much smaller error misfit. Accordingly, if a node assumes a slowness values larger than the correct one the perturbation in the error misfit may not be large enough to sensibly affect the inversion procedure.

### 7.2.3 Effect of the presence noise in the data

From the previous analysis we can expect the central area corresponding to the low values in Figure 7.2 to be also the most sensitive to the presence of noise in the data. Here the effect of such noise is analysed. A random noise whose maximum amplitude is 2.5% of the traveltime was added to the same synthetic data set used in the experiments shown in the paragraph 'Previous analysis'. The resulting noisy data set has been inverted with the use of the Genetic Algorithm with the same implementation described above. Figure 7.4 shows the Genetic Algorithm solution for such inversion. The quality is obviously poorer that the best one of the previous noise-free example (Figure 7.1b), but the position of the refractor is still well reconstructed and the errors are still located only in the middle part of the picture.

Figure 7.6 shows the result from the local optimisation with the SIMPLEX algorithm of this image. The reconstruction of the synthetic image is even poorer than before, despite the reduction in the error misfit. We could suspect that the local search has been attracted in a local minimum farther from the global one, but further analysis shows that this is not the case. The table in Figure 7.5 shows the error misfit for the Genetic Algorithm solution, the local optimiser solution

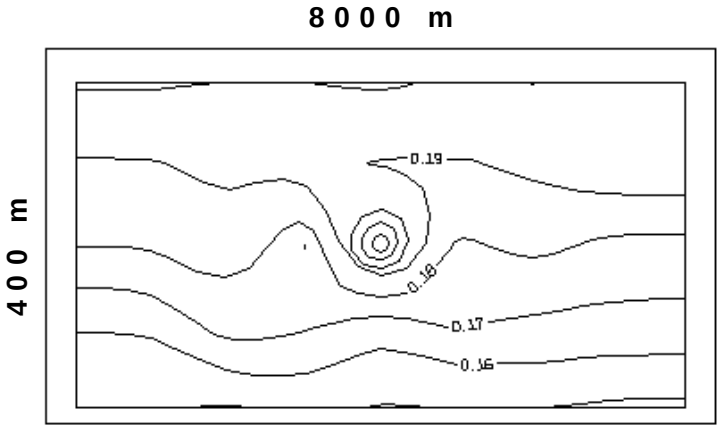


Figure 7.4: Genetic Algorithm solution from the inversion of the seismic noisy data set. The image should be compared with the ones in Figure 7.1. Errors are present in the central part of the image and are positive in sign.

| Image | Synth. data | Noisy data |
|---|---|---|
| Synth. Model | 0 | 48742 |
| GAs sol. | 2376 | 41897 |
| SIMPLEX sol. | 3143 | 39391 |

Figure 7.5: Error misfit for the Genetic Algorithm solution, the local optimiser solution and for the correct image in respect both to the original synthetic data set and to the noisy data set.

and the synthetic image in respect both to the synthetic noise-free data set and to the noisy data set. For the synthetic noise-free data set the synthetic solution has obviously zero misfit and the Genetic Algorithms solution has an error larger than the local optimiser one. For the noisy data set, the Genetic Algorithm solution has an error misfit larger than the local search output, as expected. However, the synthetic image shows an error misfit larger than both the Genetic Algorithm and local optimiser solutions. This suggests that the presence of noise in the data created a distortion in the solution space: the correct image does not correspond to the global minimum in the new solution space.

A further demonstration of this conjecture can be found in Figure 7.7. Here I have repeated the calculation described in deriving Figure 7.3, this time using the

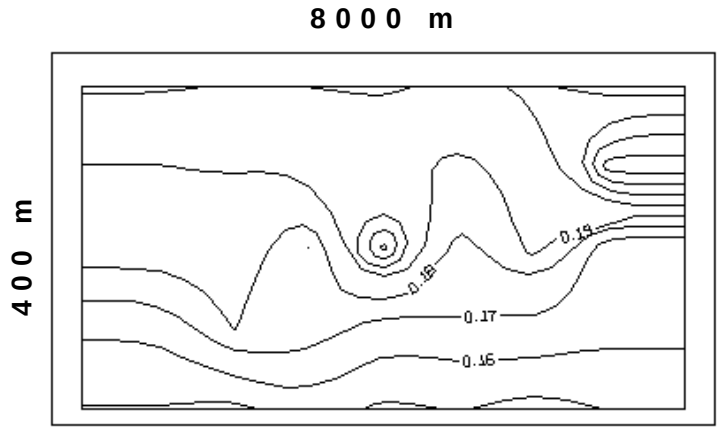


Figure 7.6: Local optimisation of the Genetic Algorithm solution in Figure 7.4 with the SIMPLEX algorithm. Despite the error misfit has been reduced, the resemblance with the synthetic model in Figure 7.1a is worse than that of the Genetic Algorithm solution. Errors are still located in the central part of the picture.

noisy data set. It can be noticed that the presence of noise has'covered' the global minimum valley shown in Figure 7.3a, and the low steep flanks corresponding to the slowness values larger than the synthetic ones, and it has substituted them with a flat region. Here the global minimum of Figure 7.3a is no longer present, rather two minima in different locations can be seen. Clearly it is now impossible for the search algorithm to recover the correct image in the domain, because it no longer corresponds to the global minimum. Notice that this phenomenon is independent of the inversion strategy employed, rather it is due to the physics of the problem.

Also notice that this analysis is not dependent on the sources of errors simulated either; this may be due to errors in the acquisition systems and in the first arrival picking stage as well as in the inadequate problem parameterisation and inaccurate ray-tracing routine. Effort should be put into the accurate detection of first arrivals, forward modelling and adequate parameterisation in order to limit the level of inaccuracies in the process. However, none of these sources of errors may be completely eliminated in field data inversions. When the level of such inaccuracies is estimated to sensibly affect the quality of the result another strategy to invert noisy data sets must be adopted.

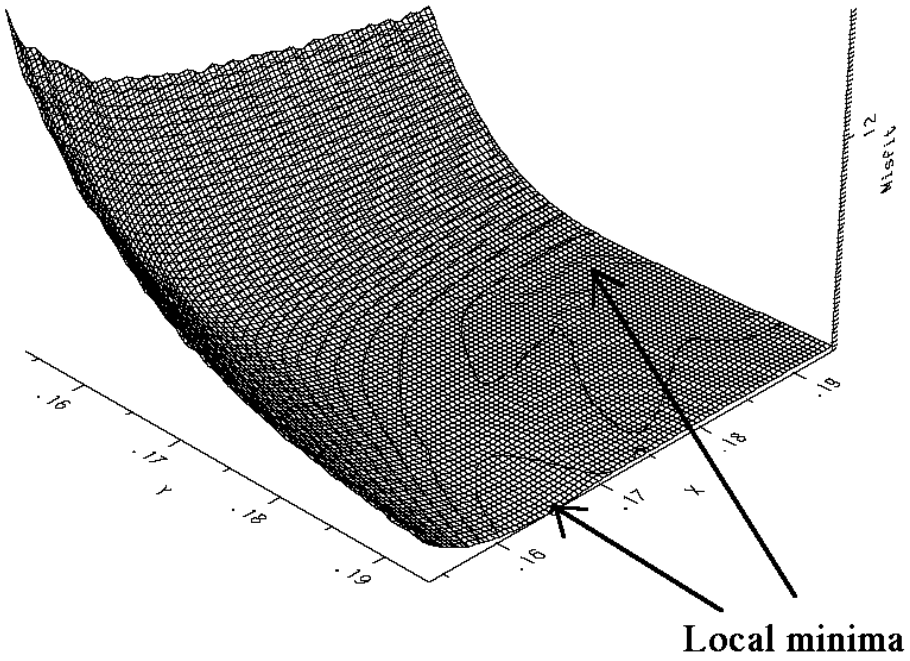


Figure 7.7: Solution surface for the noisy seismic data set in the 2 dimensional subspace correspondent to the two nodes in Figure 7.3b. It should be compared with Figure 7.3a. The curved valley is no longer present and it has been substituted by a flat region where two isolated minima are present. Such minima are in different locations in respect with the global minimum in the noise-free case.

## 7.3 Inversion of gravity data

### 7.3.1 Problem description

A procedure to invert gravity profiles with the same model discretisation adopted for the seismic refraction tomography test has been implemented. Accordingly, the density field is approximated with the same 9 x 5 grid whose spacing is 1000 m in the horizontal direction and 100 m in the vertical direction as used in the seismic experiment. Synthetic gravity measurements have been modelled at 40 stations, regularly spaced along the surface.

### 7.3.2 Forward model

One of the aim of this chapter is to show that the joint inversion of seismic and gravity data may improve the results obtained by the separate inversions of noisy data sets. In order to show this a gravity forward algorithm able to employ the same parameterisation as in the seismic experiment as well as a relationship to convert the velocity values into densities are required. These two problems have been already addressed in the literature. Some of the ideas presented in previous studies have been used in this research and references to published works will be given in the rest of the discussion. All that I need here is a process which generates reasonable synthetic data sets. The formulae in equation 7.3.2 below constitute such a process and form a known forward solution.

In this study the following relationship between seismic velocity and density has been used:

$$\rho = \begin{cases} 0.5\nu - 0.53 & \text{if } \nu > 6.2k/m \\ 0.18\nu + 1.64 & \text{if } \nu \leq 6.2k/m \end{cases}$$

where $\rho$ is the density and $\nu$ in the seismic velocity.

The relationship has been obtained by fitting with two straight lines the mean velocity-density conversion values reported in [3]. Such tabulated value are in turn extrapolated from laboratory p-waves velocities and densities measurements, presented in [6]. Analysis for in situ rock properties are also presented in [4]. Discussions about the validity of such conversion may be found in [3] where the author points out that the previous relationship should merely be considered as an average one, and departures should be expected for most rock materials. Similar conclusions can be found in [4] where is suggested that appropriate parameters for the conversion equation should be estimated locally. Accordingly, appropriate laboratory measurements on rock materials collected in the area under analysis should be performed in case the inversion strategy here presented was applied to field data.

In order to adopt an approximation similar to the one used for the seismic test, each cell in the 9 x 5 calculation grid is divided into a number of smaller subcells. For each subcell a constant density value is determined by a weighted average of the densities at the four grid nodes at the cell corners, with weights inversely

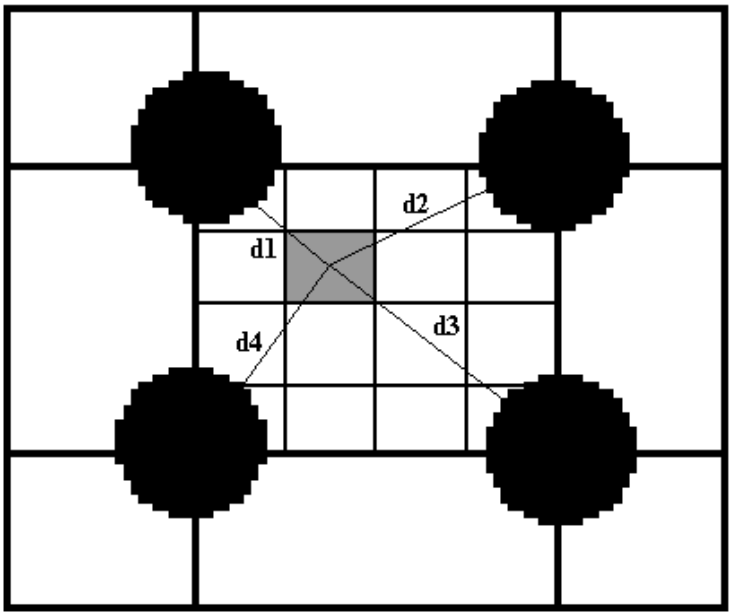


Figure 7.8: Schematic description of the model parameterisation adopted for the gravity forward calculation. Seismic slownesses are defined at the grid nodes (black filled circles). They are converted into velocities and then into densities with the use of the equations described in the text. At this point the densities for the rectangular blocks (subcells) inside the cell is defined by averaging the densities at the cell corners.

proportional to their distance from the centre of the subcell. Such procedure is graphically described in Figure 7.8. In this study the cell have been divided into 16 subcells. Numerical tests showed that further division into a larger number of subcells would have heavily increased the computation time without any sensible improvements in the calculation accuracy.

### 7.3.3 Discussion of source errors

The effect of noise in gravity data has been studied following the same steps employed in the analysis the seismic refraction data.

Figure 7.9 shows the sensitivity plot for the synthetic image presented in Figure 7.1a in the case of synthetic noise-free gravity data. This plot should be compared with that presented in Figure 7.2 for the seismic case. The plot shows a smaller range of values than in the seismic test (I recall here that the relative variation

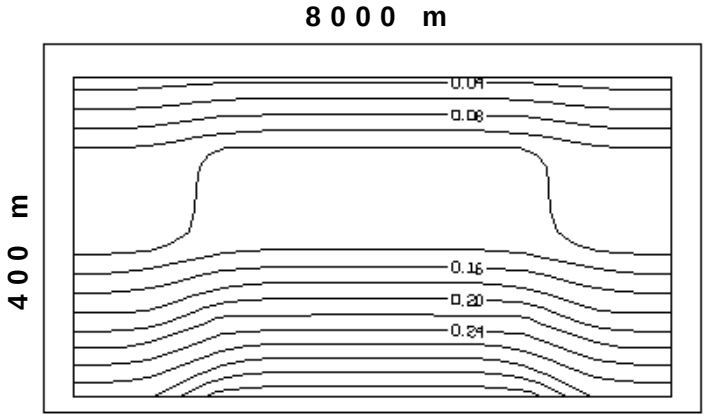


Figure 7.9: Sensitivity plot corresponding to the synthetic image in Figure 7.1a for the gravity noise-free data set. As we can see the low values are now located close to the surface nodes and the central part of the image is well constrained by the data.

and not the absolute values of the plot are relevant to this analysis). The highest values in the plot are still immediately above the refractor. Low values can be found corresponding to the surface nodes. This in due in part to the low density of the superficial layer but mainly to the fact that the surface nodes influence the density on the first row of cells only, while lower nodes influence two row of cells, the one below and the one just above. However, the plot in Figure 7.9 is quite homogeneous. Furthermore, no low values are located in the middle part of the picture, i.e., in the area less defined by the seismic analysis. This suggests that gravity data could be used to add information to the seismic inversion in this part of the solution.

In Figure 7.10 the error misfit as a function of the two nodes shown in Figure 7.3b is presented for the noise-free gravity data set. This image should be compared with Figure 7.3a for the seismic case. Since a conversion between slowness and density has been used, and in order to facilitate the comparison, the axis in the graph are still given in slowness units. The inherent ambiguity is clearly shown by the curved long minimum in the middle part of the picture. The shape of this minimum agrees well with analogous reconstructions of ambiguity domains presented in [1] and [7]. In Chapter 6 I showed that Genetic Algorithms could be successfully used to address such ambiguity in problems characterised by simple

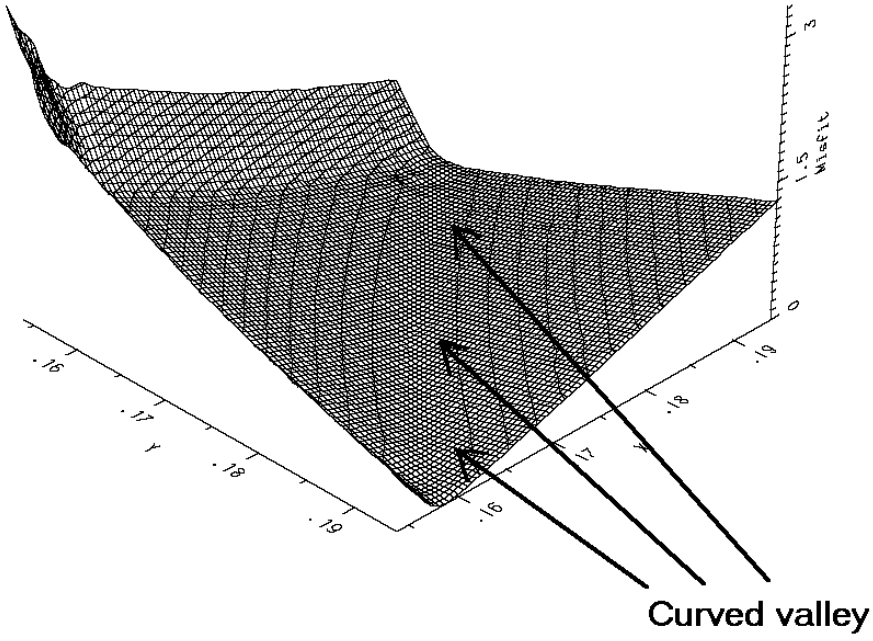


Figure 7.10: Solution surface for the noise-free gravity data set in the 2 dimensional subspace correspondent to the two nodes in Figure 7.3b. The typical potential field quadratic ambiguity is clearly present. Notice that, unlike the seismic case, both the flanks of the valley are quite steep.

geometry with a single geological contact in the vertical direction. In the configuration adopted in this experiment, in which the computation domain is described by 5 nodes in the vertical direction (i.e., more than one geological contact can be modelled in the vertical direction) the ambiguity problem may be too large to be currently tackled with the use of Genetic Algorithms. Accordingly, the inversion of gravity data is not attempted in this study.

However, the analysis of the solution surfaces in the gravity problem with and without noise shows two interesting features. First, unlike in the seismic example, in Figure 7.10 the two valley flanks have similar slope. In addition, in Figure 7.11 the error misfit is presented for the gravity data set to which a random noise at a 2.5% level as used for the seismic analysis has been added. As expected, the valley is less steep and the minimum less evident, but the overall shape close resembles

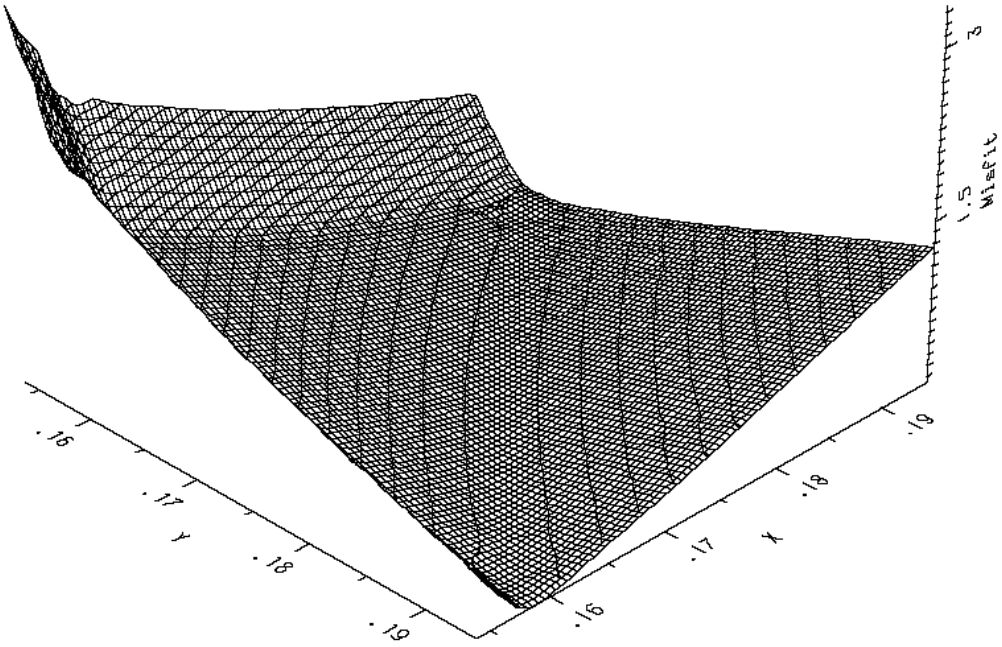


Figure 7.11: Solution surface for the noisy gravity data set in the 2 dimensional subspace correspondent to the two nodes in Figure 7.3b. This image should be compared with Figure 7.10. Notice that no major deformations in the solution surface shape have been imposed by noise.

the noise-free case. Most important, the minimum is still located in the same area as in the noise-free case. Clearly, the inversion of gravity data is less sensitive to the presence of noise than the seismic one. This suggests that gravity data could be used to constrain the inversion of noisy seismic data sets. Accordingly, the potentiality of the joint inversion of seismic and gravity noisy data sets is now analysed.

## 7.4 Joint inversion of seismic and gravity data

Different procedures may be employed when information from multiple data sets are combined in the same inversion process. One possibility is to sequentially invert the different data sets with the output from one inversion being used as the starting model for the inversion of the next data set. This has the advantage not to require any weighting between the different data set, but it has the disadvantage

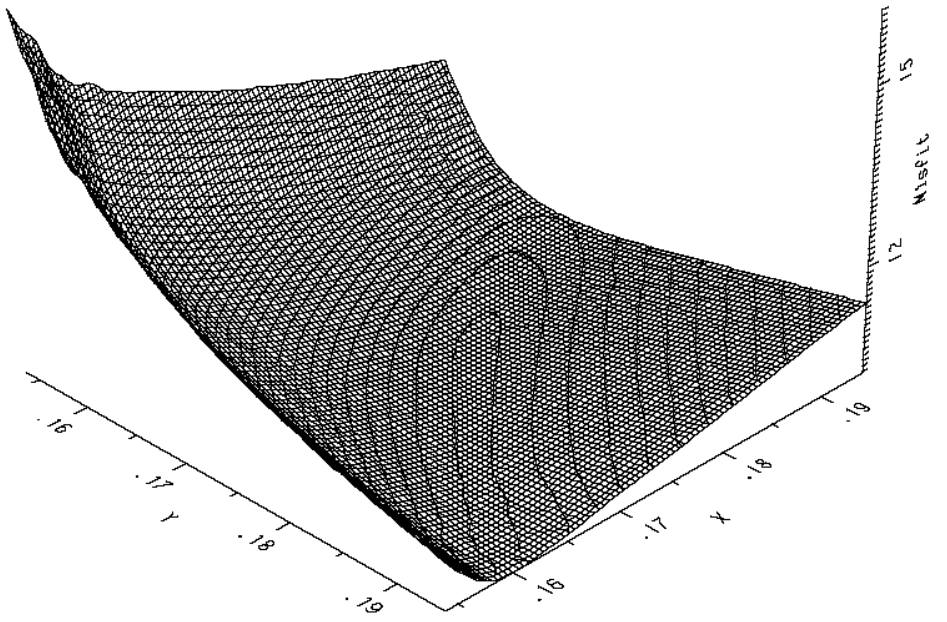


Figure 7.12: Solution surface for the noisy seismic + gravity data set in the 2 dimensional subspace correspondent to the two nodes in Figure 7.3b. This image should be compared with both Figure 7.11 and Figure 7.7. The typical curved potential field ambiguity in still present in the middle of the valley. However, its dimension is reduced compared to the gravity case. Also, the global minimum location corresponds to the one in the noise-free seismic case and no other minima are present.

that it gives less control on the actual fit of both the data sets. Another alternative is to invert the data sets simultaneously. In this case the objective function to be optimised is represented by some combination of the individual data set misfits so that information from the multiple data sets are used contemporaneously. There is a useful discussion of some of the alternatives in [5].

I first consider the effect of combining the seismic and gravity data sets. A typical choice for the misfit function is:

$$Misfit = W_{seis}\frac{\sum_1^{nrays} |\frac{T_{obs}-T_{synth}}{T_{synth}}|^p}{nrays} + W_{grav}\frac{\sum_1^{nstat} |\frac{G_{obs}-G_{synth}}{G_{synth}}|^p}{nstat} \qquad (7.1)$$

where $T$ is the traveltime, $G$ the gravity measurement, $nrays$ the number of rays in the seismic model, $nstat$ the number of stations in the gravity model, $W_{seis}$

and $W_{grav}$ are the weights given to the seismic and gravity misfit, respectively. In my calculation I used $p = 1.25$. Further discussion about the different norms to be used in goodness-of-fit criteria can be found in [2]).

In order to give the seismic data set more control on the inversion and to limit the effect of the ambiguity inherent in gravity inversions I used:

$$W_{seis} = 10\, W_{grav} \tag{7.2}$$

A similar method is used in [5].

With this misfit function I repeated my discussion using the two nodes from Figure 7.3b. The result is displayed in Figure 7.12. There is no single isolated global minimum and the curved ambiguity area is evident in the central part of the image. Note that the vast flat area shown in the solution space from the inversion of seismic noisy data in Figure 7.7 is no longer present. It is also important to observe that the ambiguity area is smaller than the one present on the noise-free gravity case in Figure 7.10 and its position approximately coincides with the location of the global minimum on the seismic noise-free case in Figure 7.3a. Furthermore, no local minima are present. Accordingly, the search algorithm is expected to be able to locate the global minimum in the search domain, or at least to stop very close to it.

I will now jointly invert the same seismic and gravity data sets. A staged procedure to perform this task has been implemented. This procedure can be seen as a compromise between the sequential and the combined inversions described earlier. In order to reduce ambiguity problems and the computation effort in the overall process seismic data only are inverted in the first stage. Then, using the sensitivity plot, I can establish which nodes in the solution are most likely subjected to errors. These nodes are selected and subjected to an additional inversion stage, while the remaining nodes are kept fixed. In the second stage seismic and gravity data are simultaneously inverted and the inversion is carried out again with the use of Genetic Algorithms in order to avoid the risk to get trapped in local minima far from the global solution.

The procedure is illustrated in the optimisation of the image in Figure 7.4. As it has already been shown, this is the result from the inversion with the Genetic

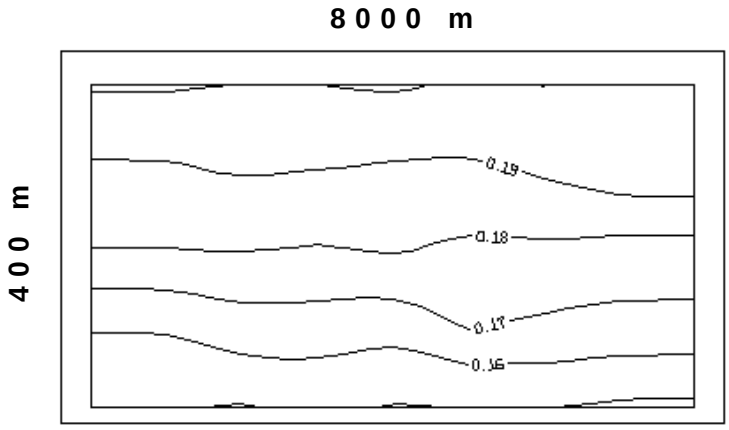


Figure 7.13: Solution from the two staged joint inversion of seismic and gravity noisy data for the model in Figure 7.1a. The errors in the central part of the image have been removed and now the resemblance with the synthetic image is satisfactory.

Algorithm of the seismic noisy data set, ie., the result from the first stage of the process. As already stated, errors are contained in the lower-central part of the image. Following the information contained in the sensitivity plot inFigure 7.2, the second stage of the process is carried out by inverting only the nodes contained in the box in Figure 7.2, while the remaining nodes are kept fixed. The result is shown in Figure 7.13. The low velocity anomaly in the centre of the image has been removed and now the resemblance with the synthetic model is satisfactory.

Some comments on the implementation of the second stage of the process need to be made. First, the nodes under analysis are re-initialised by giving thema random value. This guarantees that the inversion process avoids to get trapped in local minima very far from the global one. Second, in this stage the inversion is performed in a much smaller dimensional space and consequently its computational effort is much smaller that the one in the first stage. I ran such stage with half the Genetic Algorithm population and for one third of the generations usedin the first stage. This resulted in approximately one sixth the number of function evaluations. Accordingly, even taking into account that the computational effort of the forward calculation was almost doubled, due to the calculation of the gravity anomaly, the global cost of the second stage was no more that one third of the one in the first stage. This shows that the improvement in the final image is obtained

at a reasonable cost.

## 7.5 Further test

The staged procedure previously described has been also tested against a model characterised by a step refractor, presented in Figure 7.14a. The same 2.5% level of random noise used in the previous experiments has been added to the synthetic first arrival traveltimes and the data set obtained has been inverted with Genetic Algorithm. The results is shown in Figure 7.14b. The major features of the synthetic image, i.e. the step refractor geometry and the overall layering overlaying it, are well recovered but two low velocities anomaly are located just above the refractor. Notice that the location and sign of such anomalies agree with the ones in the previous test.

Figure 7.14c shows the sensitivity plot for the image in Figure 7.14b. Two areas characterised by low values are present and highlighted by the rectangular boxes. Notice the close spatial relationship of the two low velocity anomalies and the location of the rectangular boxes. It is particularly important to notice that this sensitivity plot has been obtained in a different way from the sensitivity plots in Figure 7.2 and Figure 7.9. Firstly, in this case the plot has been obtained by calculating the variation of the error misfit for the Genetic Algorithm output, while the previous images were obtained for the synthetic model. Secondly, here the noisy data set has been used while previously the noise-free data set was employed. This shows that such analysis is not strongly dependent on its location in the solution space (provided we are relatively close to the global minimum) and that it can be successfully performed on noisy data sets, i.e., that it can be of applicative use.

The nodes contained in the boxes, corresponding to the areas more sensitive to the presence of noise, are then selected and inverted in the second stage, while the other nodes are kept fixed. The result from the final inversion is shown in Figure 7.14d. It can be seen that the resemblance with the synthetic image is quite satisfactory.

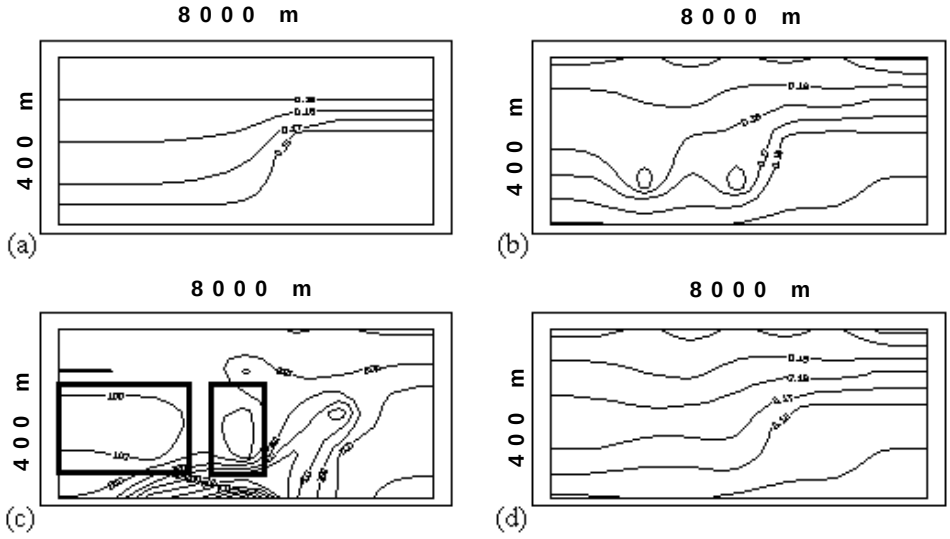


Figure 7.14: Application of the two staged joint inversion of seismic and gravity data to a synthetic model simulating a step refractor (a). In Figure b the solution from the inversion by Genetic Algorithm of seismic data only (first stage) is shown. Figure c shows the detection of the nodes most sensitive to the presence of noise (black boxes). In Figure d the solution from the second stage of the process is presented.

## 7.6 Discussion

After presenting the results at this chapter a fewcomments on the results of the inversion of the Nevoria data set discussed in Chapter 5 and on the accuracy of the ray-tracing routine at Chapter 3 are worthwhile making.

It has been shown that the presence of noise on seismic data can create a deformation on the search space such to mislead the inversion procedure towards a wrong solution. The entity of this effect is a function of the level of noise in the data as well as of the velocity contrast in the solution. Also, in presence of strong velocity contrast the relatively small velocity anomalies imposed by the presence of noise will influence to a lesser extent the geological interpretation of the final solution.

A detailed study to assess the maximumlevel of noise able to affect thequality of a seismic reconstruction for different geological settings would be a natural development of this study but could not be performed in this research because of

time constrains.

In the Nevoria experiment at Chapter 5 Low errors in the solutions, represented by low velocity anomalies, are present only in the large 255 dimensional inversion in Figure 5.17. No low velocity anomalies are present in the solution from the inversion performed by subdividing the overall domain into 7 small 45-dimensional subdomains presented in Figure 5.16. I propose the following interpretation for this phenomenon:

- in the Nevoria experiment much stronger velocity contrasts are present compared to the synthetic tests discussed in this chapter. This suggests that is higher level of noise would be required to produce errors in the velocity field reconstruction;

- even given the last point, this shows that the ray-tracing routine employed in this study gives a satisfactory approximation of the ray propagation in complex media. Clearly this result could not have been achieved if major errors were inherent in the forward algorithm. This point has been already addressed in Chapter 3;

- the low velocity anomalies in the 255 dimensional inversion of the Nevoria data set presented in Figure 5.17 are most likely due to the inability of the Genetic Algorithm to deal with a too large search space, rather than to the presence of noise. Accordingly, these errors should be related to similar errors present in the inversion of synthetic noise-free data with Genetic Algorithm without pseudo subspace method (see Figure 5.1c).

## 7.7 Conclusions

In this work the effect of inaccuracies in the forward calculation in seismic tomography inversion problems have been analysed. It has been shown that errors in the traveltime can create a distortion in the solution space such to mislead the inversion procedure towards a wrong solution.

This problem in independent of the efficiency of the inversion procedure. In presence of such distortion, even an extremely efficient inversion routine, able to

exactly detect the global minimum in the solution space, would not be able to reconstruct the correct parameter configuration, because it no longer corresponds to the global minimum.

When an inversion process applied to real data apparently fails it is hard to establish if this is due to the inability of the inversion routine to locate the global minimum or to inaccuracies in the forward model.

This analysis put emphasis on the importance of accurate forward modelling and appropriate parameterisation. Accordingly, once the problem has been extensively studied in synthetic tests cases and a proper inversion algorithm has been found to deal with the mathematical aspects of the problems, effort should be put into the selection of efficient forward model that can accurately mimic the physics of the process under analysis, in order to enable the procedure to work well on field data.

The inversion of gravity data has been shown to be less sensitive to the presence of such inaccuracies. Errors analogous to the one used in the seismic case do not create major distortions in the solution space.

Accordingly, in this chapter I show that the joint inversion of seismic and gravity data offers great potentiality in the treatment of noisy data sets. Seismic data limit the inherent ambiguity in potential field inversion, while gravity data limit the effect of noise in the seismic inversion.

# Bibliography


[1] M. Al-Chalabi. Interpretation of gravity anomalies by non-linear optimization. *Geophysical Prospecting*, 20:1–16, 1971.

[2] M. Al-Chalabi. When least-squares squares least. *Geophysical Prospecting*, 40:359–378, 1992.

[3] P. Barton. The relationship between seismic velocity and density in the continental crust - a useful constraint? *Geophysical Journal of the Royal Astronomic Society*, 87:195–208, 1986.

[4] G. N. Gogonenkov. *Seismic prospecting for sedimentary formations.* A.A. Balkema, 1990.

[5] L. R. Lines, A. K. Schultz, and S. Treitel. Cooperative inversion of geophysical data. *Geophysics*, 53:8–20, 1988.

[6] J. W. Ludwig, J. E. Nafe, and C. L. Drake. Seismic refraction. In *The sea*, pages 53–84. Maxwell and Wiley, 1970.

[7] D. W. Vasco, L. R. Johnson, and E. L. Majer. Ensemble inference in geophysical inverse problems. *Geophysical Journal International*, 115:711–728, 1993.