Chapter 5

Application of Genetic Algorithms to seismic refraction tomography

5.1 Introduction

In this chapter the results from some experiments in the application of Genetic Algorithms to seismic refraction tomography problems are discussed. First the different Genetic Algorithms implementations presented in Chapter 4 are compared on a synthetc tests in order to select the most effective configuration. Then the potentiality of such implementation is analysed on a number of synthetic tests simulating different refractor geometries as well as on physical model and field data.

5.2 Inversion of synthetic data

5.2.1 Synthetic data generation

Synthetic first-arrival travel time data were generated using a line length of 8000 m with a group interval of 200 m and shots spaced at every 500 m. This results in a total of 615 rays/travel times which are inverted to define the velocity structure in the first few hundred meters below the surface. For the purposes of inversion, the slow-

ness distribution in the sub-surface is defined by a 9 x 5 grid whose spacing is 1000 m in the horizontal direction and 100 m in the vertical direction. A linear slowness gradient is assumed between the grid nodes. This has been extensively described in Chapter 3 where details about the ray-tracing routine are given. Thus, the model has a 45-dimensional solution space, formed from the slowness values at each of the grid nodes. This dimensionality is small compared to conventional seismic tomographic inversion but, to my knowledge, represents one of the highest-dimensional Genetic Algorithm applications to seismic data published so far.

5.3 Comparison of different Genetic Algorithms

To explore the effects of the different choices described in Chapter 4, a series of Genetic Algorithms have been compared on a dataset obtained with an horizontally layered synthetic model with linearly varying slowness in the vertical direction (see Figure 5.1a). This model is characterised by slowness values which allow for the occurrence of both diving and refracted rays and a uniform distribution of ray paths (Figure 5.1b).

Three different specific choices were tested; linear normalisation selection, linear normalisation selection with pseudo subspace method and parent selection with pseudo subspace method (see Chapter 4 for details about the processes employed in these different Genetic Algorithm implementations). Each test consisted of five runs with different random seeds, i.e. with different randomly chosen initial populations. For each implementation the best solution found by the algorithm in these five runs, together with the convergence curve is presented (Figure 5.1c-h).

Comparison of Figures 5.1c and 5.1e shows an improvement in the performance of the linear normalisation selection when the pseudo subspace method is used. The essential features of the test model are reproduced with only significant differences occurring at the top right of the model in an area of poor ray coverage. This suggests that the

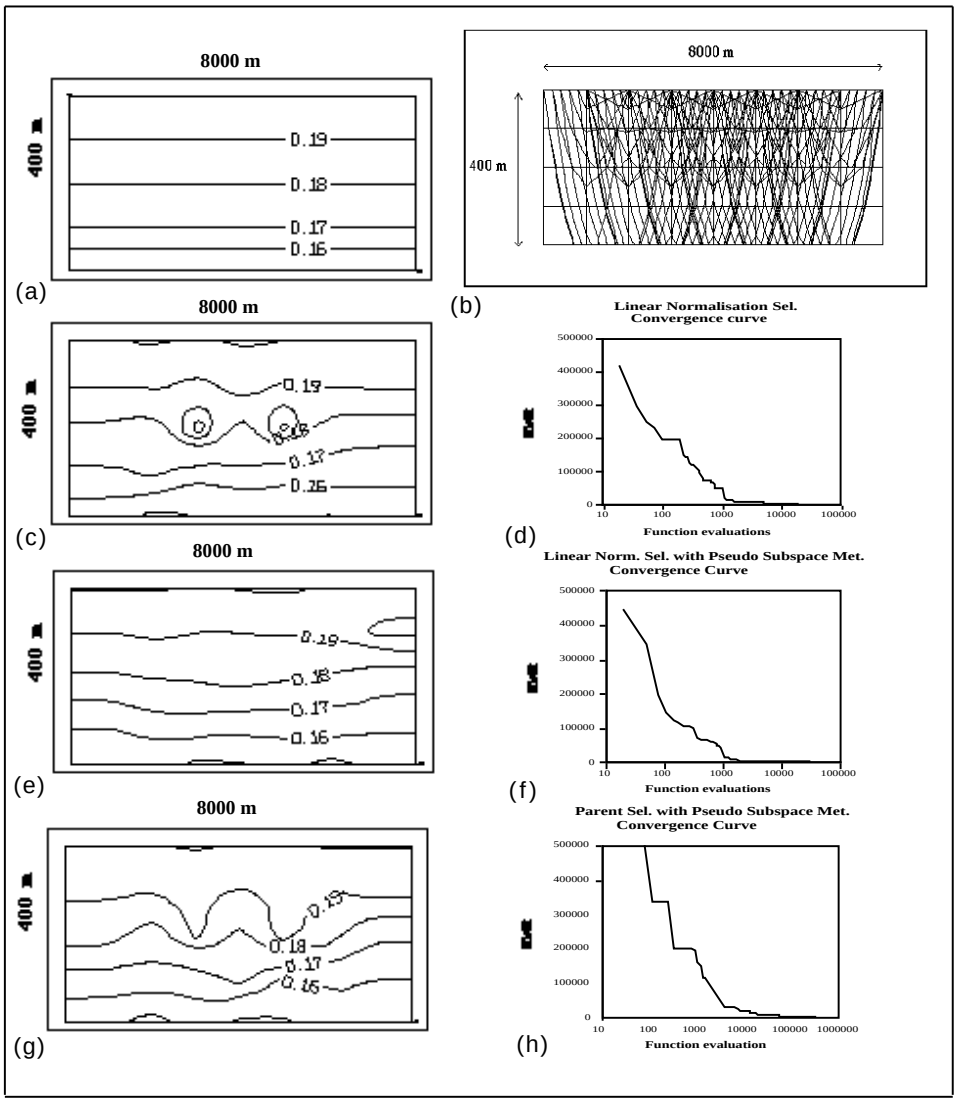

Figure 5.1: Results from the comparison ofdifferent Genetic Algorithm implementations on an horizontally layered synthetic test with linearly varying slowness in the verticaldirection. The synthetic model (a) together with the ray diagram (b) are presented. Both final models and convergence curves are shown for linear normalisation selection Genetic Algorithm (c and d), linear normalisation selection with pseudo subspace method (e and f), parent selection with pseudo subspace method (g and h).

algorithm benefits from the inclusion of the pseudo subspace method.

The quality of the solution obtained with the parent technique (Figure 5.1g), even though it is obtained at a much higher cost in terms of the number of required function evaluations (Figure 5.1h) is not the best of the three solutions. This suggests that efforts to maintain diversity within the population will not necessarily improve the quality of the final solution.

Clearly the Genetic Algorithm with linear normalisation selection including pseudo subspace method outperformed the other implementations. In the initial generations, using a coarse grid, the algorithm still appears to be able to locate the favourable area of the solution space in which to concentrate the more detailed and time consuming subsequent analysis.

5.3.1 Local search method

The convergence curves in Figure 5.1 illustrate a well known characteristic of Genetic Algorithms; they are poor optimisers [3]. They show a very fast initial convergence, followed by progressively slower improvements. In fact this behaviour is common to many optimisation techniques, but is of particular concern in Genetic Algorithms. The form of the curve suggests that the algorithm should be stopped when an approximate solution has been found, because further improvements may be very costly. Improvements to the Genetic Algorithm solution can only develop though crossover or mutation, i.e. random events. Depending on the exact form of the solution space, further improvements to high-fitness solutions can be a rare event.

In the case of the best solution found by the Genetic Algorithm with pseudo subspace method, small errors are present in each parameter. This is due to the fact that the traveltime misfit due to slightly wrong parameters may be reduced by introduced small errors at adjacent nodes. Improvement in the solution at this stage may be obtained only by accurate tuning of most of the parameters at the same time, that cannot be efficiently achieved by a Genetic Algorithm process. However, further improvements to the inversion method can be obtained by combining the Genetic Algorithm with a local optimising method. In the initial stages of the inversion the space-sampling prop-

erties of the Genetic Algorithm is used to direct the search to the region close to the global solution. This solution can be further improved using a local search method such as a hill-climber algorithm.

Other hybrid Genetic Algorithm implementations have already been successfully applied to geophysical problems. A related approach has been used by K. Mathias et al., [7] although in this case no pseudo subspace method was used and local search was applied periodically to more than one solution obtained by the Genetic Algorithm. Also, Sen and Stoffa [10] showed how Genetic Algorithms performance can be greatly enhanced by importing some elements from Simulated Annealing process.

A number of algorithms for the local optimisation of functions in multi-dimensional spaces are described in the literature [5, 8]. The choice of the algorithm to use is problem-specific and often experimental trials are required. I tested four local search methods: the downhill SIMPLEX method, Powells method, the conjugate gradient method [8], and the local search routine from the Hill-climber method described by De La Maza and Yurez [2]. Of these the SIMPLEX method, Powells method and the local search routine from the Hill-climber method do not use gradient information while the conjugate gradient method does. The approximate calculation performed by the ray-tracing routine does not give gradient information. Consequently, the numerical calculation of the required derivatives is time consuming and the conjugate gradient method is considerably slower than the other methods.

After extensive tests the SIMPLEX code was found to be the most reliable and stable algorithm and accordingly was used in the successive tests. Figure 5.2 illustrates the result from the local optimisation of the best individual from the Genetic Algorithm with pseudo subspace method, together with its convergence curve. The SIMPLEX algorithm was able to improve the solution, minimising the squared error up to a level of approximately zero misfit. The solution is almost indistinguishable from the original model (Figure 5.1a). All the layers are well reconstructed both in terms of velocity values and of vertical position, with only a minor departure from the original at the middle left-hand side of the solution. Clearly the two-staged procedure linking the Genetic Algorithms global search with the local optimisation has been very effective.

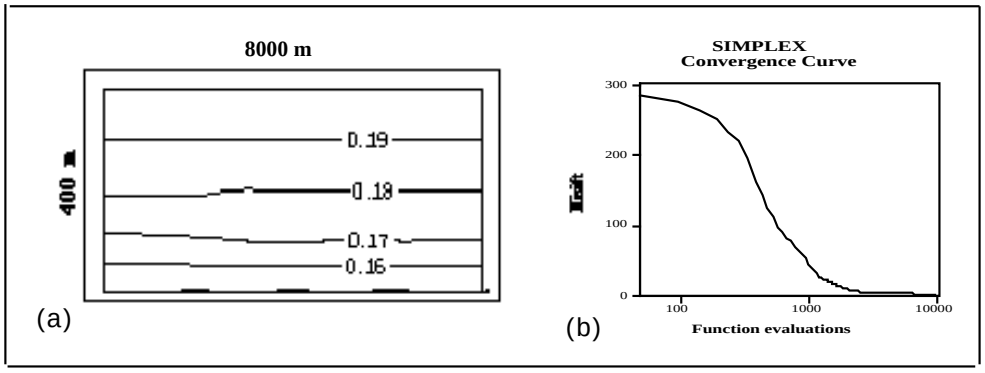

Figure 5.2: Result from the optimisation of the Genetic Algorithm solution with the SIMPLEX algorithm (a) and convergence curve (b).

5.3.2 Efficiency

An inversion process performance is not only evaluated in terms of the accuracy of the solution it can produce, but also in terms of its cost. In case of a Genetic Algorithm this is basically measured in terms of the number of function evaluations (ray-tracing in this case) required to obtain the final image. In general, Genetic Algorithms are considered a relatively expensive method, to be used only in cases where local procedures proved to be unsuitable. Raiche [9] considers this to be particularly true for geophysical applications. However, my results suggest that this is not always the case.

Clearly, an important decision in terms of the efficiency of the two-stage optimisation techniques described above, is when to terminate the Genetic Algorithm and implement the local search. I now implemented the Genetic Algorithm to output the best solution every 50 generations after the final grid configuration has been achieved. The best solutions after 100, 150, 200 and 300 generations for the Genetic Algorithm with linear normalisation selection incorporating the pseudo subspace method are presented in Figure 5.3. These pictures should be compared with the result at the end of the process, i.e. after 400 generations, already shown in Figure 5.1e. Clearly an acceptable solution is obtained after only 100 generations, with only minor improvements occurring up to 300 generations, after which variations are negligible.

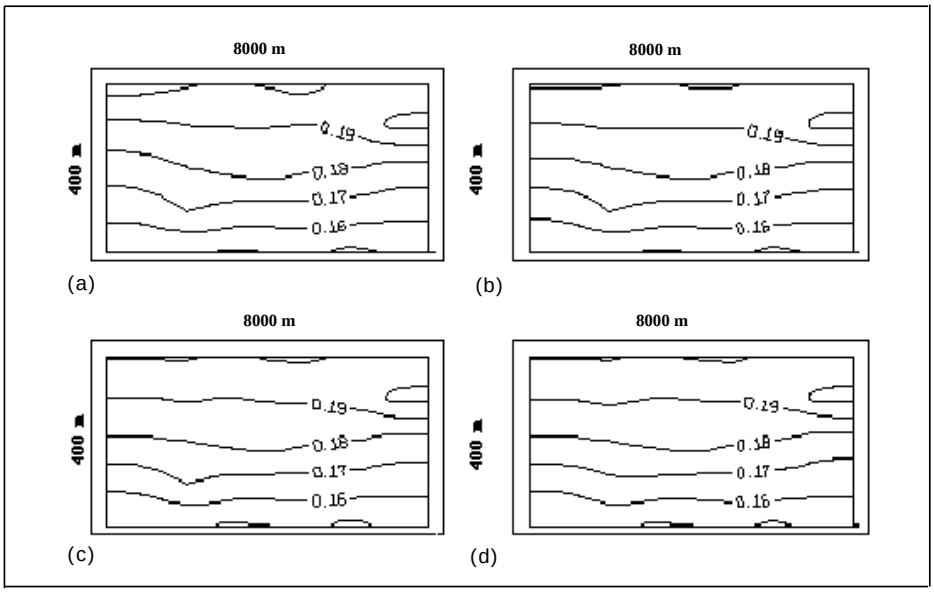

Figure 5.3: Best solutions after 100 (a), 150 (b), 200 (c) and 300 (d) generations for the Genetic Algorithm with pseudo subspace method. An acceptable solution is already obtained after only 100 generations.

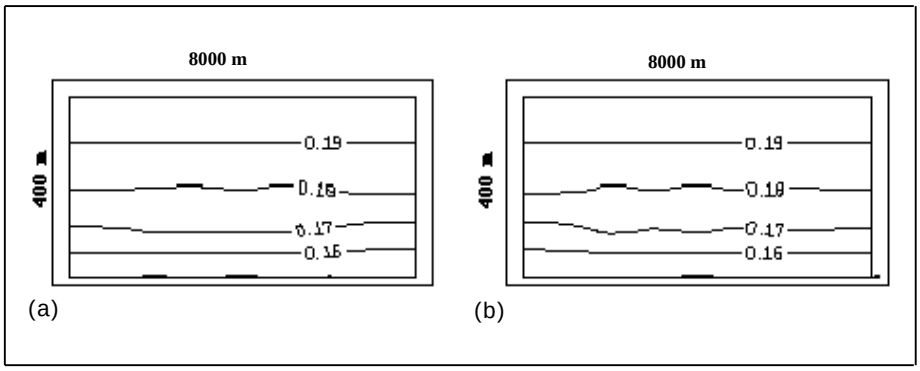

Figure 5.4: Models resulting from local optimisation of the best solutions obtained after 100 (a) and 150 (b) generations of Genetic Algorithm with pseudo subspace method. Only minor differences may be found between the two models.

An even more interesting test is to improve, by local search, such solutions and compare the final results. In Figure 5.4 the results obtained using the local search, initialised with the solutions after 100 and 150 generations, are shown. The differences between the results are minimal, even when compared with the result from the solution locally improved after 400 generations (Figure 5.2a). This shows that Genetic Algorithms have discovered the 'good' valley in the solution space very rapidly. Furthermore, the computation effort, i.e. the number of function evaluations, for the local improvement of the 100, 150 and 400 generations solutions by the SIMPLEX code is almost the same (Figure 5.5). Clearly the space sampling of the Genetic Algorithm after the 100th generation has not been very productive. After a solution located in a good valley has been found there is little advantage in using the Genetic Algorithm for further space sampling because more effective results may be obtained by the local search. It is very difficult to generalise about when to stop the Genetic Algorithm and begin the local search. A number of alternatives are available, for instance, waiting for no improvements to happen for a few generations, or waiting for variations in the misfit to fall below a predetermined value. However, choosing such a threshold can be very difficult and problem specific. Because the Genetic Algorithm process is non-deterministic, it is pos-

Generations	GAs	Optimization
100	6316	10403
150	10814	10620
400	24286	10355

Figure 5.5: The number of function evaluations performed in 100, 150 and 400 generations by the Genetic Algorithm and the number of evaluations required by the local search to optimise these models.

sible to have many generations without any improvement, followed by very rapid improvements as new domains are discovered by crossover or mutation. Research to define effective criteria suitable for controlling the inversion of the seismic data is on-going.

As shown in Figure 5.6, in 100 generations the Genetic Algorithm performs approximately 6000 function evaluations to locate the region in the solution space containing the global minima. This computation effort is comparable to that required by the local search to further improve the solution. Thus, a Genetic Algorithm applied to this problem should be considered good not only in terms of accuracy but also in terms of computation effort. Inversion of the synthetic data employing only the local optimiser starting from a random point in the solution space (i.e., under the same conditions used to initialise the Genetic Algorithm process) were attempted and resulted in the process getting trapped in local minima very far from the global solution. Even in these cases the process required a number of iterations larger than required by the Genetic Algorithm to find an acceptable solution.

5.3.3 Stability

As described above, each implementation of the Genetic Algorithm was tested on five different random seeds (initial populations). The stability of the proposed inversion procedure can be illustrated by comparing the results of the five different tests. Since the Genetic Algorithm process is non-deterministic we expect the solutions to differ. The best solution is illustrated in Figure 5.1e. In Figure 5.6 the four other results obtained after 150 generations using the linear normalisation Genetic Algorithm

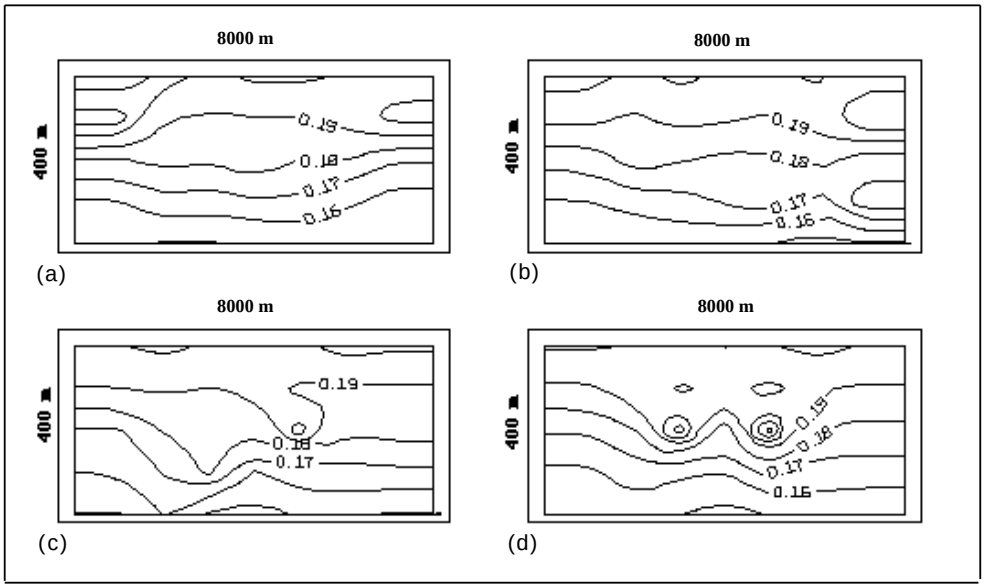

Figure 5.6: Results obtained using different starting populations after 150 generations using Genetic Algorithm with pseudo subspace method. These results should be compared to the synthetic model in Figure 5.1a. Solutions in (a) and (b) are comparable to the synthetic image, however (c) and (d) contain discrepancies.

with pseudo subspace method are presented. In two cases there has been good reconstruction of the layering with minor anomalous features (Figure 5.6a and 5.6b), whilst in the remaining two examples major anomalies are present (Figure 5.6c and 5.6d). However, the results obtained after local optimisation using the SIMPLEX code are in acceptable agreement with the synthetic image in every case (Figure 5.7).

This suggests that in all five tests the Genetic Algorithm with pseudo subspace method has been able to find good valleys in the solution space. All these solutions were sufficiently close to the global minimum for the local optimiser to further reduce the error. All the tests were randomly initialised but converged to the same solution. This suggests that only one global solution is present in the application, i.e., no cases of clear ambiguity are manifested, although obviously the number of tests is small. Given the relative inexpensiveness of the Genetic Algorithm process, as a general rule it might be advisable to perform more than one Genetic Algorithm inversion with different, randomly chosen, initial populations and locally improve the best solutions found. This would help identify where more than one solution to the problem is possible.

Similarly, the solution obtained from the local search does not necessarily have zero misfit. In such a case it is impossible to discriminate if such solution represents a local minimum or a point whose surrounding space topography is so complicated that it cannot to be further improved even by the local search. In one of the tests we performed this has actually been the case. Calculating the function values along the direction connecting a solution obtained from a local search to the synthetic image, showed values to be constantly decreasing even at very small steps. Thus, the local search had failed to complete the exploration of the good valley. This was a lucky trial: not all the possible directions can be examined in a 45-dimensional space, and the existence of curved valleys connecting the local solution to the global one cannot usually be excluded. Thus, the assumption that the output from a local search must be a local minimum should not be taken for granted. However, in the test examples the difference between the local optimiser solution and the global minimum is apparently not large.

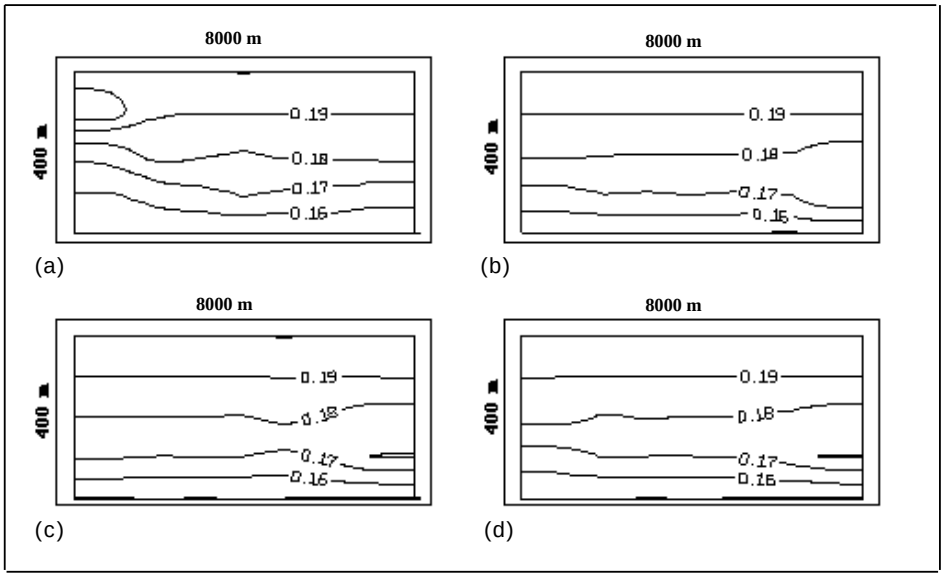

Figure 5.7: Results from local optimisation, using the SIMPLEX code, of the four solutions shown in Figure 5.6 obtained with Genetic Algorithm with pseudo subspace method. All the reconstructions are in good agreement with the synthetic case (Fig. 1a).

5.3.4 Other synthetic tests

The two-stage inversion procedure described above has been further tested on three other synthetic data sets; a shallow horizontal refractor (see Figure 5.8), a refractor incorporating a step (Figure 5.9) and an isolated buried body with anomalous velocity (Figure 5.10). These noise-free data sets allow us to test the method without the complications due to the presence of noise. The same conditions used in the flat horizontal layers inversion example have been applied in all these tests. All the synthetic images have been created with a 9 x 5 slowness grid resulting in a 45-dimensional inversion problem.

Notice that the only 'a priori' information used in these inversions is contained in the Genetic Algorithm implementation, i.e., slowness field parameterisation and slowness value constraints. Since the method here described was developed as part of a project to define the near surface structure of Precambrian rocks in the Western Australian shield (see the field data application at the end of the chapter) the synthetic models have slownesses of around 0.18 s/km. Accordingly, the values of slownesses in the solution set were limited to lay between 0.14 s/km and 0.25 s/km, equivalent to velocities between 7.14 km/s and 4.0 km/s. This information is the same for all the different tests and no specific a priori information or starting model is needed in the individual runs.

The pseudo subspace method process has been implemented in three stages. In the first stage the inversion is performed on a 3 x 2 grid, whose spacing is 4000 m in the horizontal and 400 m in the vertical direction. The node spacing is then halved at each stage. Consequently, in the second stage the inversion is performed on a 5 x 3 grid and in the last stage it reaches the final configuration of 9 x 5 nodes. Each test consisted of five runs with different random seeds and the best solution found by the Genetic Algorithm has been further improved by the local search. Here the results are presented:

Shallow Horizontal Refractor. This configuration is similar to the horizontal flat layer case previously analysed, except that the refractor is raised to the middle of the model and the slowness below it kept constant (Figure 5.8a). This approximates the case of a shallow layer whose velocity increases with depth, overlying

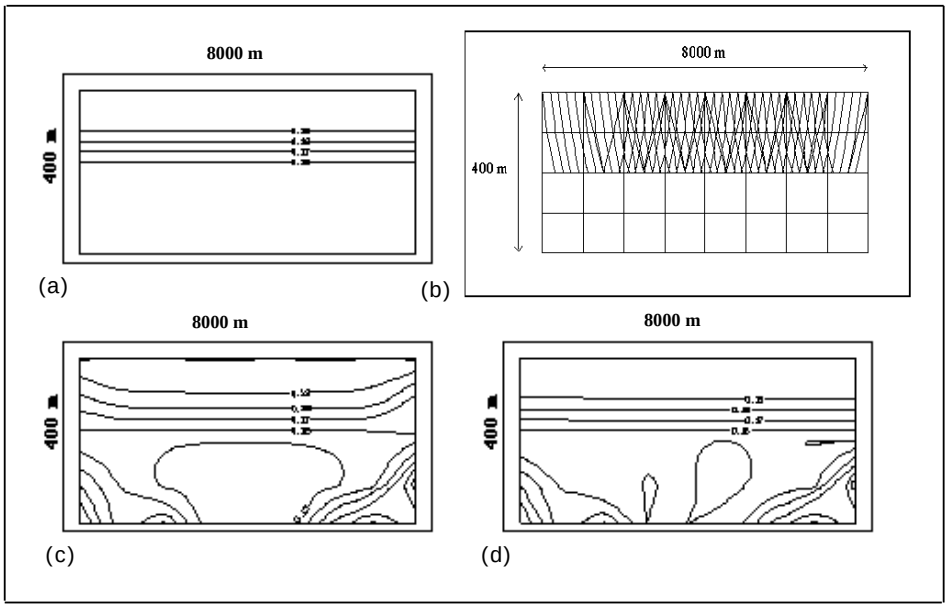

Figure 5.8: Inversion of synthetic data set simulating the presence of a shallow horizontal refractor. a) synthetic model, b) ray diagram, c) result obtained using Genetic Algorithm with pseudo subspace method, d) result obtained after local optimisation.

a layer with a constant and higher velocity. This configuration is particularly interesting because it tests the ability of the inversion process to correctly recover the depth of the refractor. Intuitively it is expected that many possible solutions to the problem exist with increased depth to the refractor being compensated by a higher velocity for the overlying area, and vice-versa. The solution obtained by the Genetic Algorithm is shown in Figure 5.8c and after local optimisation in Figure 5.8d. The match between true model and the final solution is very good with the refractor correctly located in the vertical direction and perfectly horizontal. From the ray-paths diagram (Figure 5.8b) we see that no rays penetrate to the lower part of the model and hence no real solution can be expected from this area. The slowness values are simply artefacts of the Genetic Algorithm. Note further that as the rays are concentrated in the area above the refractor it is not surprising that the slowness in this area is well recovered.

Step refractor. This model is a development of the previous with a step incorporated into the refractor (Figure 5.9a). As it can be seen in Figure 5.9b, this model allows rays to dive through the refractor. Again, the results after the two stages of the inversion procedure are presented in Figure 5.9c (solution from Genetic Algorithm) and 5.9d (from the Genetic Algorithm plus local optimiser). The final solution is accurate in the right-hand and left-hand side of the model, however, small errors are present in the central region. Overall the final model is very close to the original one, with the best fit in areas where there are most rays, especially when these vary in orientation.

Concealed Body. In the final model an isolated body with low velocity is superimposed on a vertical velocity gradient (Figure 5.10a). As shown in Figure 5.10b most of the rays run along the top and the borders of the body. In the resulting solution the outline of the body is almost perfectly reconstructed (Figure 5.10d) and this allows, as a consequence, excellent recovery of the vertical velocity gradient. Note that the result obtained using the Genetic Algorithm (Figure 5.10c) is so close to the starting model that

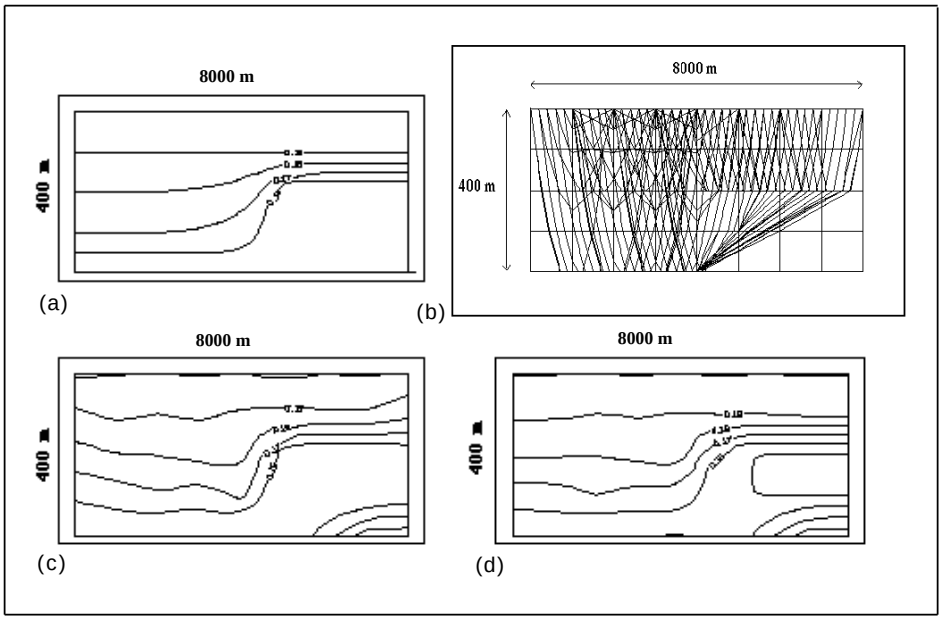

Figure 5.9: Inversion of synthetic data set simulating the presence of a step refractor. a) synthetic model, b) ray diagram, c) result obtained using Genetic Algorithm with pseudo subspace method, d) result obtained after local optimisation.

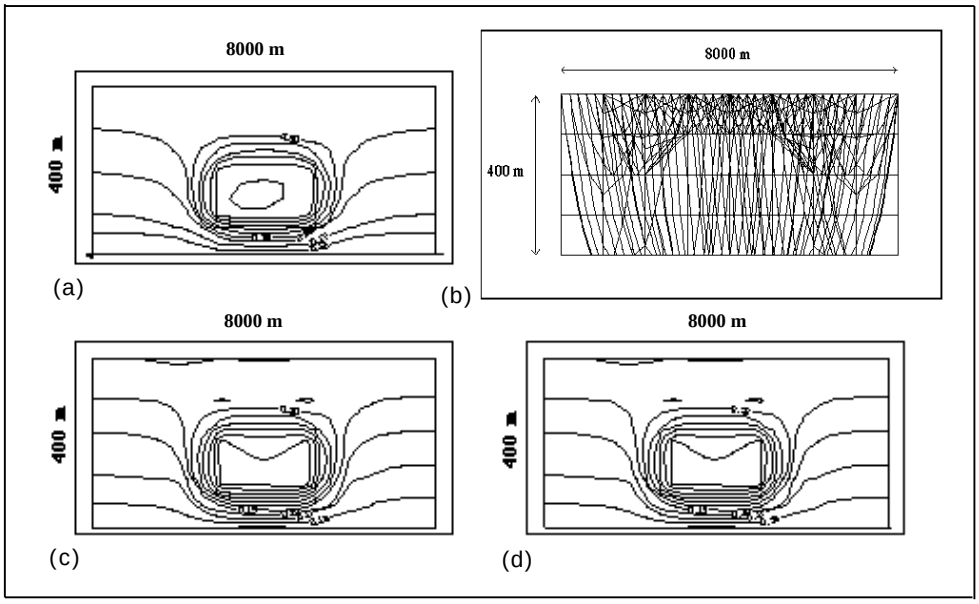

Figure 5.10: Inversion of synthetic data set simulating the presence of a concealed body. a) synthetic model, b) ray diagram, c) result obtained using Genetic Algorithm with pseudo subspace method, d) result obtained after local optimisation.

the local search was not able to further improve it.

5.4 Inversion of model data

From Figure 5.2 it can be noticed that improvements to the Genetic Algorithm solution through the use of the local optimiser can be obtained only by a very small decrease in the error misfit. The success of such process in different synthetic tests is very important from a theoretical perspective because it shows that Genetic Algorithm search with the use of the pseudo subspace method is successful in detecting the valley containing the global minimum in the solution space. However, on model and real data sets, i.e., 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 is obviously useless. This phenomenum is further discussed in Chapter 7 where a detailed study of the influence of noise in the tomographic inversion of refraction data is presented. However, the synthetic tests presented in this study show that the quality of the Genetic Algorithm solution is satisfactory even without local optimisation. Consequently, the performance of the Genetic Algorithm inversion procedure without local optimisation has also been tested using model data. The analysis of the effectiveness of the inversion procedure on such data set characterised by a relatively low level on noise has been considered a useful step towards the application to the real field seismic data set discussed at the end of the chapter.

5.4.1 Experiment configuration

A model data were collected in the Physical Modelling Laboratory at Curtin University of Technology. The model consisted of a block of Plexiglass/Perspex having a P-wave velocity of 2670 m/s. Its top surface models a fault with a dip of approximately 60 degrees (Figure 5.11). The block was immersed in water (velocity 1470 m/s) which acted as a low velocity layer overlying the faulted refractor. Notice that in this experiment both the layers are characterised by constant velocity. Accordingly, a ray-tracing algorithm suitable for constant velocity media has been used. The original algorithms from [1] was implemented and used in this case. This is the algorithm that was modified in order to allow for varying slowness inside the cells, and it has been described in Chapter 3.

Data were collected along a single profile perpendicular to the fault. The profile is 20 cm long, and the experiment has been scaled at $1 : 100,000$ so as to model a 20 km long seismic survey. At this scale the water layer is 1500 m thick on the upthrown side of the fault and 3100 m thick on the downthrown block.

Forty shots, with a scaled spacing of 500 m, were modelled. Data were recorded at 291 receiver positions spaced at a scaled interval of 50 m. The entire spread was progressively moved from left to right across the model, as shown in Figure 5.11, i.e. from the shallower water

layer to the thicker. Due to geometry problems the source was kept at the left end of the spread during the entire survey and consequently the number of receivers decreased at each shot. Accordingly, a larger number of rays have been collected on the right side of the calculation domain. A split-spread configuration may have yielded better results.

Two piezoelectric transducers with 1 MHz resonant frequency were used as both sources and receivers to acquire the data. The transducers were 1.3 *cm* wide and were inclined at approximately 30 degrees to enhance the signal to noise ratio of critically refracted arrivals. This inclination, together with the transducer width being one order of magnitude larger that the receiver spacing caused the effective receiver position to vary across the transducer surface, due to the different incidence angles of direct, diffracted and critically refracted arrivals. To address this problem the ray-tracing program was run with two different receiver positions for each receiving transducer location. For direct arrivals the assumed position was on the transducer margin closest to the source, while for refracted arrivals the position was assumed to coincide with the centre of the transducer. Arrival times were then compared and the earlier arrival used in the inversion process. Tests with synthetic data showed that this method correctly reproduced the first arrival data.

5.4.2 Experimental results

Before attempting the inversion of the model data a further inversion of a synthetic data set was performed. The synthetic data set was created using a synthetic model simulating the geometry and seismic velocities of the physical model (Figure 5.12a). The error-free synthetic data so obtained allowed to test the potentiality of the inversion algorithm in this problem from a mathematical perspective and to compare it with the result from the actual model data. In all the following examples the inversion has been performed using the pseudo subspace method in three stages. In the first stage there were two cells in the vertical direction, one representing the body and one the layer underneath the domain. The cell underneath the domain is required to allow the rays to be refracted before reaching the surface. Notice that in this case the domain is discretised by constant velocity cells, as already mentioned.

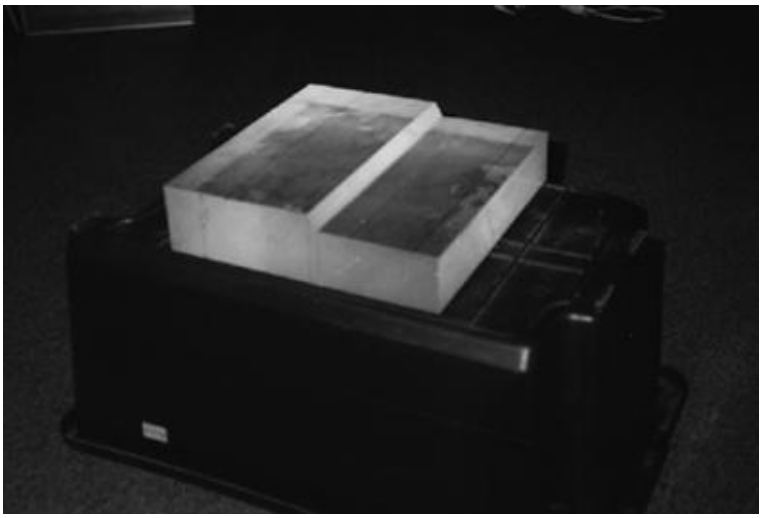

Figure 5.11: The model used in the experiment at the Physical Modelling Laboratory at Curtin University of Technology. It consists in a block of Plexiglass/Perspex characterised by a seismic velocity of 2670 m/s. The model has been immersed in water so as to simulate the presence of a low velocity layer overlying a higher velocity refractor.

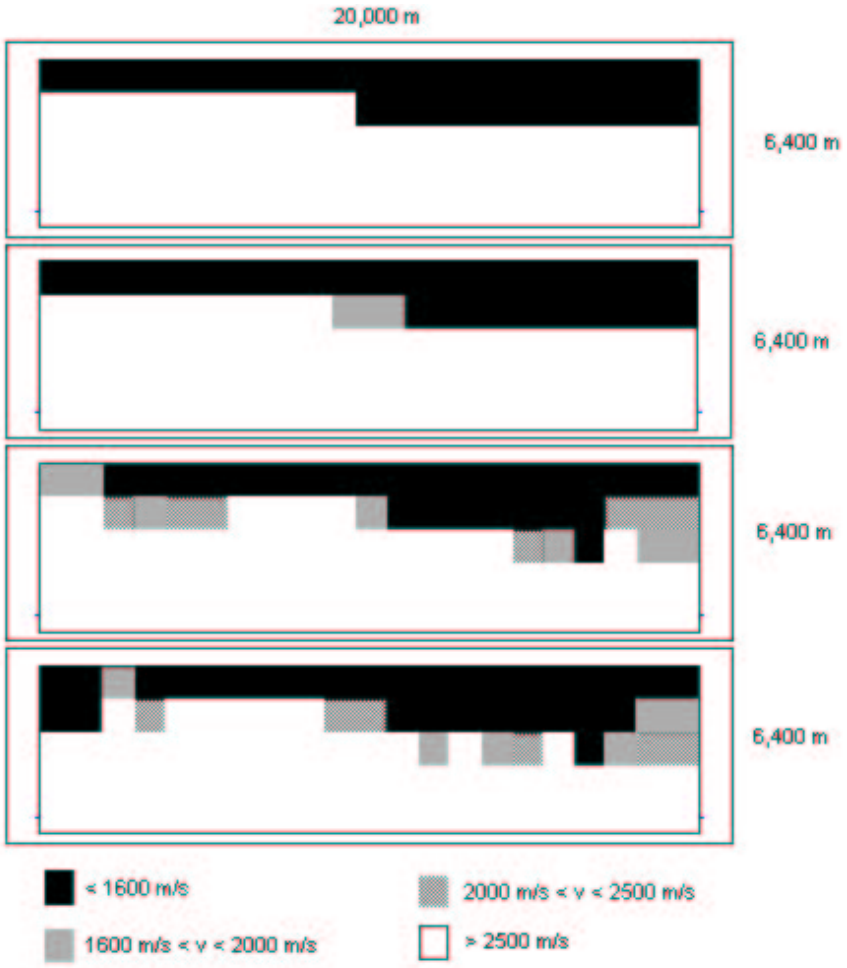

Figure 5.12: (a) Synthetic model reproducing the geometry and velocities of the physical model used in the experiment. (b) Result of the inversion of synthetic data for a final grid configuration of 9 x 5 nodes. Errors can be found in correspondence of the fault due to the relatively coarse horizontal grid spacing. (c) Result of the inversion of synthetic data in the case a final grid configuration of 21 x 5 nodes. The main features of the original image, i.e. the fault, refractor position and the lower velocity layer overlaying it have been recovered. (d) Result form the inversion of the real data set for a final configuration of 21 x 5 nodes. As for the synthetic example, the variation in the thickness of the low velocity layer is clear, as well as the presence of the fault in the middle of the picture.

In the second stage the cells dimension is halved resulting in three vertical cells (two for the body and one for the bottom). Eventually in the last stage the final configuration with five cells is reached. The number of cells in the horizontal direction varied from one experiment to the other, as will be shown, but the halving process is analogous.

Due to the statistical character of Genetic Algorithms, all the tests were run five times and only the best results are shown in the diagrams. However, the quality of the results obtained from the different runs is comparable.

Figure 5.12b shows the result of inverting the synthetic data with a final grid configuration of 9 x 5 nodes. As already mentioned, because of the use of first arrivals only, no rays sample the cells in the bottom part of the domain. Accordingly, no information from this area may be collected and the Genetic Algorithm result from this part of the grid is totally random. For the sake of clarity, this area has been set to a constant velocity, in this case equal to the Plexiglass/Perspex P-wave velocity. This is justified by the fact that the Plexiglass/Perspex P-wave velocity must be considered a high velocity limit for the blocks in this area: if the velocity was higher the rays would propagate in such area before reaching the surface.

The result is particularly satisfactory. Errors can be found in correspondence of the fault, however this is due to the relatively coarse horizontal grid spacing. Notice the similarities between this test and the ones already shown in Figure 5.8 and 5.9. In this case both the the vertical position and the step in the refractor have been satisfactorly reconstructed.

Figure 5.12c shows the result of using a final grid configuration of 21 x 5 nodes. The ray diagram for this case is presented in Figure 5.13. The diagram shows that the ray-tracing routine accurately models the rays diffracted at the top of the step. A large shadow area occurs adjacent to the fault in the downthrown block, and the ray coverage is poor close to the borders of the domain. Accordingly, the tomographic reconstruction in such areas should not be considered reliable. Nevertheless, the main features of the original image have been recovered: it is possible to recognise the refractor approximately in the correct position, the fault and a generally lower velocity layer overlaying it. The image is less well defined than the previous one. This because the larger

dimensional space (105 dimensions in this case) hinders both the forward and the inverse process. In the forward process, the larger number of cells decreases the ray density per cell, reducing the definition with which the image can be reconstructed and resulting in ambiguity problems. This is particularly true for the top right area of the domain, where the low velocity layer is thicker. In the inversion process, the geometric dimension of the calculation domain increases exponentially with the number of inversion parameters and a much larger population should be used in the Genetic Algorithm. However, this would result in the calculation time becoming impratically long on the current hardware. The reconstruction of this image required approximately 15,000 function evalutations. Better results could probably be obtained with a larger population and the implementation of parallel Genetic Algorithms on appropriate hardware could probably offer a valid alternative for large optimisation problems.

Since this result has been obtained with synthetic, error-free data, the quality of the image in Figure 5.12c is unlikely to be bettered by any inversion of the model data.

Having acquired some knowledge of the problem complexities the inversion of the model data set was tackled. The result is shown in Figure 5.12d. In this case a final grid configuration of 21 x 5 nodes was specified. Again the main features of the model have been recovered. As in the synthetic example, the variation in the thickness of the low velocity layer is clear, as well as the presence of the fault in the middle of the picture. The refractor runs mostly along the correct position. The quality of the result is comparable to the one obtained with synthetic data for the same grid configuration. Accordingly, I can state that the maximum possible resolution has been approached.

5.5 Application to field data

Following the promising results obtained with both synthetic and model data the Genetic Algorithm presented in this study has also been applied to the inversion of a real data set in order to test its efficiency on practical situations.

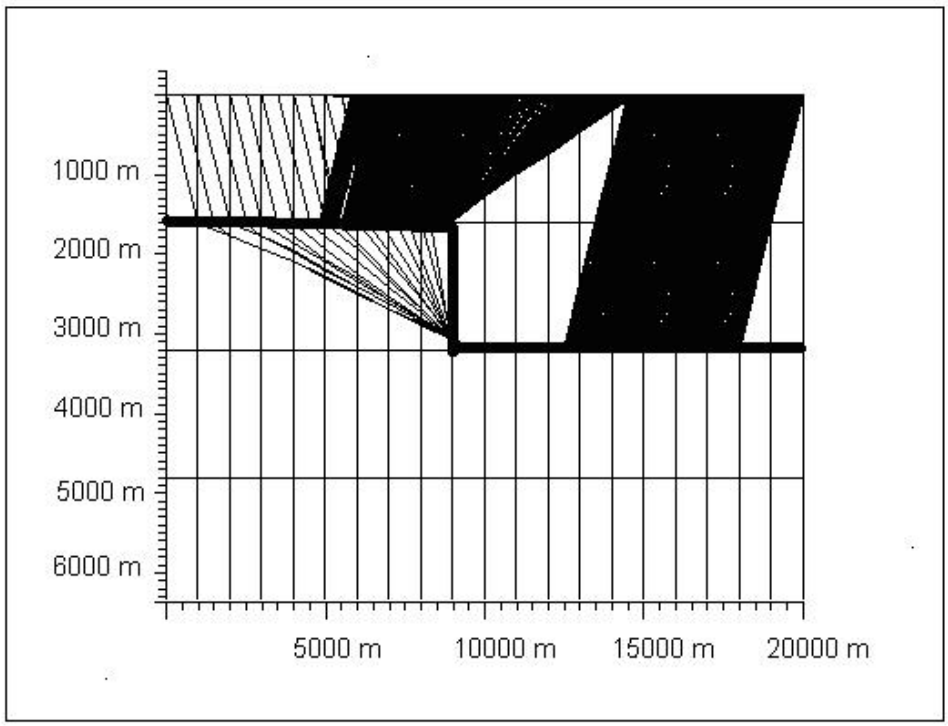

Figure 5.13: Ray-diagram for the synthetic model with 21 x 5 nodes. The diffracted rays are clearly modeled at the top of the step. A large shadow in correspondence of the step is shown as well as the poor ray coverage close to the borders of the domain. These phenomena are responsible for the lack of definition in such areas in the final reconstruction of the image.

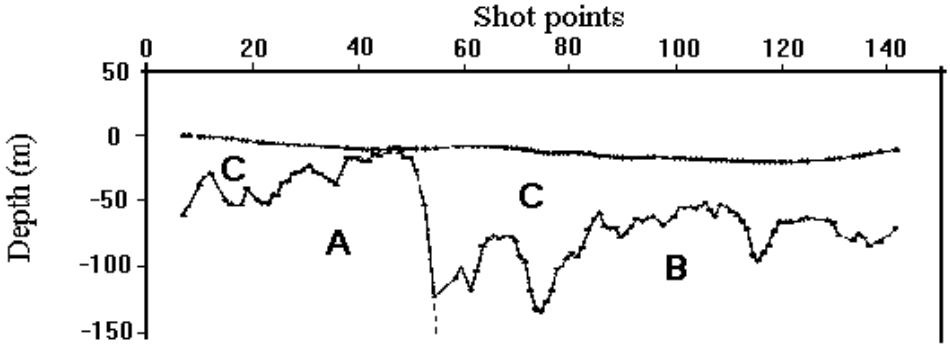

Figure 5.14: Result of the analysis of the real data set recorded close to the Nevoria Gold Mine, Southern Cross, Western Australia, with the plus-minus method. A schematic description of the geology of the area (obtained through lithologic analysis of RAB chips) is also given: a weathered layer (c) overlays a basement formed by greenstones (a) and granitoids (b) divided by an almost vertical contact.

5.5.1 The seismic data set

Refraction data from a seismic survey near the Nevoria Gold Mine, Southern Cross, Western Australia, have been used. The area is mainly characterised by greenstones and granitoids overlain by a thick weathered cover (regolith). Such weathered profiles are of great exploration interest and they have recently been the target of various geophysical studies in order to deduce which is the most viable method to map their base and internal structure. For this purpose the seismic refraction method efficiency has been tested. During that survey, shots were fired and data recorded at 151 stations 25 m apart. The profile lies across an almost vertical contact between greenstone and granitoids that reaches the surface near shot n. 45 in Figure 5.11. The first arrivals had been previously analysed with the plus-minus method [6] in a study performed, in collaboration, by the University of Western Australia and Curtin University in Perth, Western Australia (see [4]). The results are presented in Figure 5.14.

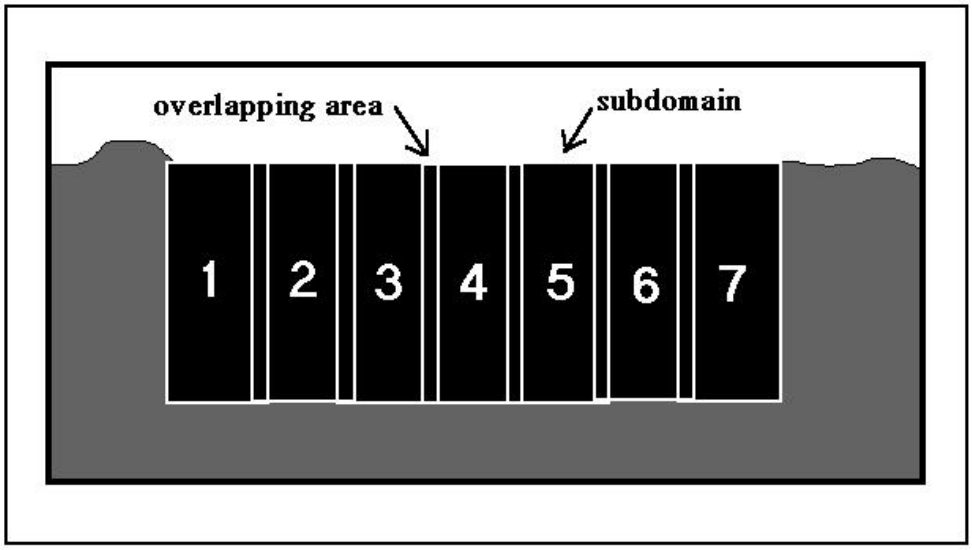

Figure 5.15: Schematic description of the process performed to subdivide the search domain into 7 small subdomains. Notice that such subdomains overlap at the lateral borders in order to disregard the results obtained in areas of poor ray coverage.

5.5.2 Experimental results

The same data set has then been inverted with the use of a Genetic Algorithm with pseudo subspace method. The main aim in this experiment was to test the ability of the algorithm to detect the refractor position. The inversion has been carried out on a domain 3750 m long and 160 m deep, with a resolution of 75 m in the horizontal and 40 m in the vertical direction, resulting in a 51 x 5 nodes grid. In order to reduce the problem dimensionality the domain has been divided into 7 small subdomains (9 x 5 nodes), each overlapping for 2 nodes at the lateral borders. Figure 5.15 helps to describe the process. In this way the problem has been reduced to the same dimensionality as the synthetic tests above. Once all the 7 subdomains were inverted, the nodes at the extreme of the single subdomains were disregarded, due to the low ray density in such areas and the remaining nodes linked together in order to obtain the global 51 x 5 node solution.

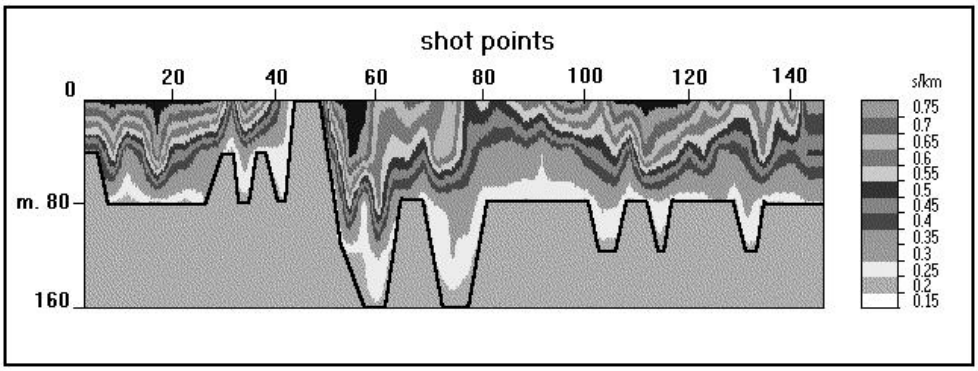

Figure 5.16: Inversion of the real data set with Genetic Algorithm with pseudo subspace method. The image has been obtained by subdividing the search domain into 7 small subdomains and by linking the individual results. The steep fault and the undulations in the refractor position agree well with the previous analysis obtained with the plus-minus method presented in Figure 5.14.

This solution is shown in Figure 5.16. As already mentioned above for the synthetic and model tests the result below the refractor is not reliable due to the limited amount of rays diving in such area. In order to facilitate the interpretation, the position of the refractor has been marked by a thick line and the slowness value below such line have been muted (notice that the refractor position is discretised at the nodes location). Such operation has been performed with the use of the ray-tracing routine: each node slowness value is altered by a small amount both in negative and positive sign, if the error misfit does not undergo any change it means that no rays dive in area close to the node and such node should be muted.

The result in Figure 5.16 agrees well with the previous analysis obtained by the more traditional method presented in Figure 5.14. Notice that a detailed comparison of the two pictures is not possible due to both the approximations involved in the plus-minus method and the coarse parameterisation in the Genetic Algorithm inversion. Nevertheless, the two images are characterised by the same main features: the steep contact towards the left part of the picture as well as the acute undulation in the refractor position in the centre and the smoother

undulation at the right-hand side of the picture. Some disagreement are present immediately to the left- hand side of the area where the refractor reaches the surface (shots 35-40 in Figure 5.16). This may be due to the coarse parameterisation in the Genetic Algorithm inversion.

Eventually, an attempt to invert the overall domain in a single process has been performed. The pseudo subspace method has been implemented in such a way to divide the search into 3 stages of different dimensionality. In the first stage the search is performed on a grid with 13 nodes in the horizontal direction and two in the vertical. Then the dimensionality is increased to 25 x 3 and eventually to the final configuration of 49 x 5 nodes, approximating the 51 x 5 dimensionality of the final image in Figure 5.16. Such result is presented in Figure 5.17. It can be noticed that the solution lacks of definition in the reconstruction of the refractor position. The solution space is too large to be accurately searched and some details can not be resolved. Furthermore, errors are left in some parts of the picture. The deep undulation in the middle part of the image is substituted by a flat refractor with a low velocity anomaly over it. This kind of error was also present in the synthetic test inverted with Genetic Algorithm without subspace search (see Figure 5.1c) and is due to details that can not be resolved in too large search domains. Clearly the effort to keep the problem dimensionality as low as possible by subdividing the search domain into small subdomains have been particularly beneficial. Notice that this has been achieved even at the cost of reducing the amount of information available. By dividing the domain into a number of small subdoimans, not only the problem dimensionality but also the amount of rays to be inverted has been reduced because only the rays contained in the individual subdomains were used and the ones crossing the subdomains were discarded. However, even in this way better results have been obtained compared to the inversion performed on the overall domain with all the information available.

Nevertheless, it should be noticed that the solution in Figure 5.17 is not totally unsatisfactory: the contact between greenstones and granitoids is well recovered as well as reasonably good in the reconstruction of the average refractor position. Also, the stratification overlying the refractor resembles the one presented in Figure 5.16. This suggests that problems whose dimensionality is larger that the 9 x 5 nodes synthetic

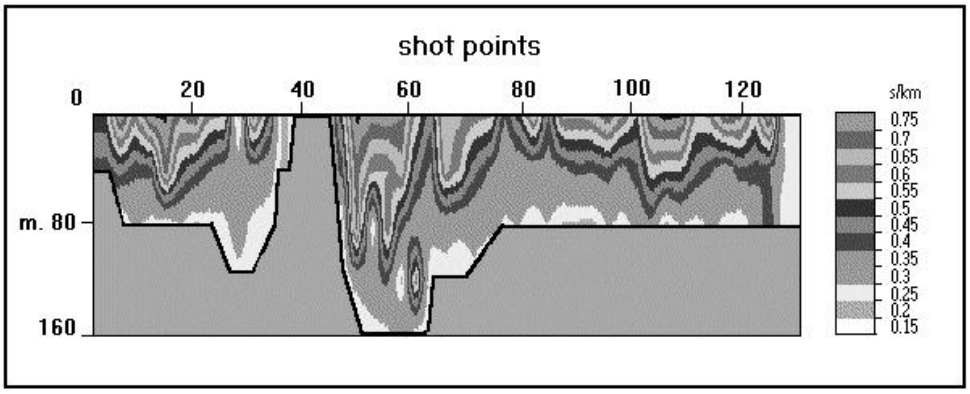

Figure 5.17: Inversion of the real data set with Genetic Algorithm with pseudo subspace method without subdivision of the search domain. The solution should be compared with Figure 5.16. Errors are present in the refractor position and in the presence of a low velocity anomaly in the centre of the picture.

tests presented above should not be considered completely beyond the potentiality of the method. This confirms the results obtained with the physical model data described above in which tomographic problems up to 105 dimensions were satisfactory inverted.

5.6 Conclusions

Geophysical problems tend to have larger dimensionality than most of the optimisation problems Genetic Algorithms have been traditionally applied to. This has meant that Genetic Algorithm are not usually applied to large complex geophysical problem. The experiments here presented show that Genetic Algorithms process strongly benefits of the inclusion of a pseudo subspace method, whereby the complexity and dimensionality of a problem is progressively increased during the inversion. Such implementation allows to tackle larger dimensional problems and to quickly locate the region of the solution space containing the global minimum in refraction tomography problems and proved to be robust and to require a limited amount of 'a priori' information.

The experiments presented in this chapter show also that any further possibility to reduce the problem dimensionality, such as subdividing the search into small subdomains, should be pursued.

The method potentiality can not be yet compared with traditional seismic processing techniques in very large problems and it should be considered a useful tool to obtain a relatively fast and accurate preliminary analysis. However, the good results obtained in the reconstruction of both the refractor position and the slowness field on synthetic and field data is particularly promising. In addition, the process proved to be effective, in terms of computation effort.

Bibliography

[1] Asakawa and Kawanaka. Seismic ray tracing using linear traveltime interpolation. *Geophysical Prospecting*, 41:99–111, 1993.

[2] M. De la Maza and D. Yuret. Dymanic hill climbing. *AI Expert*, March:26–31, 1994.

[3] K. DeJong. Genetic algorithms are not function optimizers. In *Fundations of Genetic Algorithms 2*, pages 5–18. Morgan Kaufmann Publishers, 1993.

[4] M. C. Dentith, M. C. Jones, and A. Trench. Exploration for gold-bearing iron formation in the burbidge area of the southern cross greenstone belt, w.a. *Exploration Geophysics*, 23:111–116, 1992.

[5] L. Dixon and G. Szego. *Towards global optimization.* North-Holland Publishing Company, 1978.

[6] J. G. Hagedoorn. The plus-minus method of interpolating seimsic refraction sections. *Geophysical Prospecting*, 7:158–182, 1959.

[7] K. Mathias, D. Whitley, C. Stork, and T. Kusuma. Staged hybrid genetic algorithm search for seismic data imaging. In *ICEC 1994*, pages 356–361. IEEE, 1993.

[8] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery. *Numerical recipes.* Cambridge University Press, 1992.

[9] A. Raiche. Progess, problems, and challenges. *Surveys in Geophysics*, 15:159–207, 1994.

[10] M. K. Sen and P. L. Stoffa. Rapid sampling of model space using genetic algorithms: examples from seimsic waveform inversion. *Geophysical Journal International*, 108:281–292, 1992.