

Introduction of Genetic Algorithms for geophysical applications

Gaspar Monsalve

Genetic Algorithms (GA)

- ✓ Strategy to fit a model to a data set using forward modeling.
- ✓ It uses the principles of genetics / evolution, such as natural selection, in order to find the most appropriate model for the set of data.
- ✓ The selection is based on some measurement of data misfit.
- ✓ It facilitates fast convergence and prevents the user from purely random tests.

Advantages of Genetic Algorithms

- ✓ They efficiently search the model space, so they are more likely (than local optimization techniques) to converge toward a global minima.
- ✓ There is no need of linearization of the problem.
- ✓ There is no need to compute partial derivatives.
- ✓ More probable models are sampled more frequently than less probable ones.

Basic operations in the genetic algorithms

- ✓ Given a data set, the evolution starts with a population of randomly selected individuals (models). The initial population should cover the entire range of possible solutions.
- ✓ Subsequent generations are constructed by the action of three reproduction operators:
 - Selection (based on misfit)
 - Cross-over (swapping information)
 - Mutation (Changes in individual genes)

Selection

- ✓ A filter to select the model that best fit the data is commonly used.
- ✓ After defining a probability of reproduction (which can be calculated as a function of misfit), the strongest (fittest) individuals are selected.
- ✓ A good balance should be found in the choice of the “amount of pressure” of the selection algorithm. “High pressure”: fast convergence and risk of a local minimum. “Low pressure”: Very slow convergence.

Reproduction

- ✓ Pairs of “parent” solutions are selected to generate “child” solutions with many of the characteristics of their “parents”.
- ✓ Since only the best individuals are selected for breeding, the average fitness of the second generation will increase.
- ✓ Reproduction includes cross over (high probability) and mutation (low probability).
- ✓ There is usually one offspring per “parent”, so that the population size remains constant.
- ✓ Sometimes individuals can be ranked according to their fitness; the number of offspring will be proportional to the rank position.

Cross-over

- ✓ Exchange (swapping) of genes between “parent” models in order to produce “child” models (next generation).
- ✓ Controls the degree of mixing and sharing of information.
- ✓ It is performed only by those that pass the fitness test.
- ✓ Cross-over favors stronger models and ensures that diversity is kept.
- ✓ “Parents” stronger than their offsprings can be included in the next generation instead of the “child”.

Cross-over methods

Models have string representations (binary, decimal, etc..)

Single point cross-over



Two point cross-over



Cut and splice cross-over



Uniform cross-over

Individual bits in the string are compared between the two parents. The bits are swapped with a fixed probability

Mutation

- ✓ Making random changes in genes.
- ✓ It is made in order to maintain some degree of randomness in the population (helps to avoid local minima).
- ✓ The probability of mutation should be kept low in order to prevent excess of randomness.

Sequence of steps

1. Input: A generation of individuals.
2. Calculation of misfit and selection.
3. Reproduction (cross-over + mutation) =
New population
4. Go to 1. with a new generation

When to stop? (termination criteria)

- ✓ Misfit (a minimum criteria)
- ✓ Number of generations (maximum reached)
- ✓ Allocated budget reached (time, money, etc.)
- ✓ Solution reached a plateau
- ✓ Combinations of previous criteria

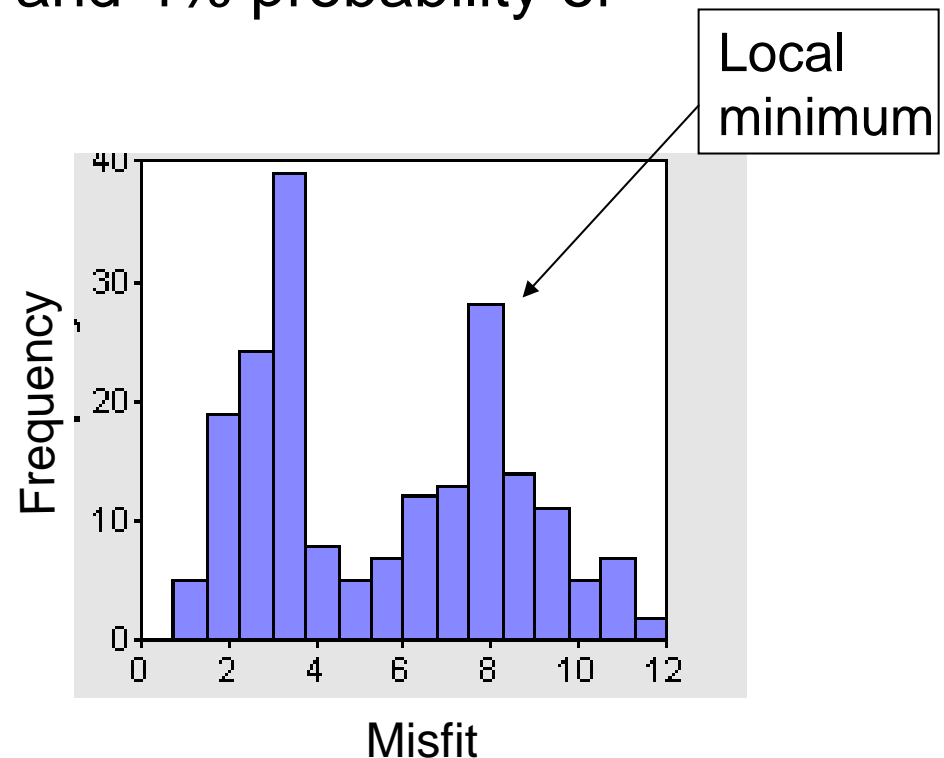
Disadvantages of genetic algorithms

- ✓ They show a very fast initial convergence, followed by progressive slower improvements (sometimes is good to combine it with a local optimization method).
- ✓ In presence of lots of noise, convergence is difficult and the local optimization technique might be useless.
- ✓ Models with many parameters are computationally expensive.
- ✓ Sometimes not particularly good models are better than the rest of the population and cause premature convergence to local minima.
- ✓ The fitness of all the models may be similar, so convergence is slow

Examples: Hypocenter location

(Billings, Kennett and Sambridge, GJI, 1994)

- ✓ Population of hypocenters is generated randomly.
- ✓ Subsequent generations are constructed by the action of the reproduction operators with 90% probability of cross-over and 4% probability of mutation.
- ✓ Each earthquake is located 1000 times. They expect the locations to be tightly clustered about a single region in the parameter space, and with similar misfits.



Examples: Hypocenter location

(Billings, Kennett and Sambridge, GJI, 1994)

- ✓ Parameter correlations in the hypocenter location problem:
 - Strong correlation between depth and origin time.
 - Weak correlation between epicenter and depth.
- ✓ The genetic algorithm search is done in latitude-longitude.
- ✓ After finding an epicenter, they fix lat.-lon., and for each depth they calculate the optimum origin time.

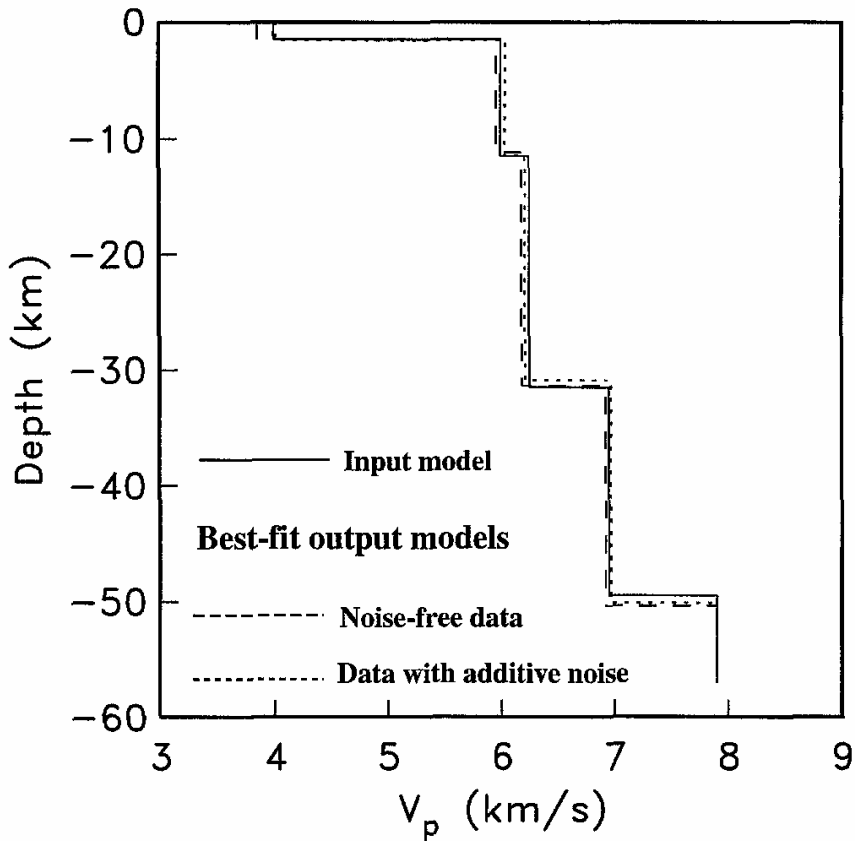
Examples: Velocity structure using waveform fitting

(Bhattacharyya et al., BSSA, 1999, Sambridge and Drijkoningen, 1992, GJI)

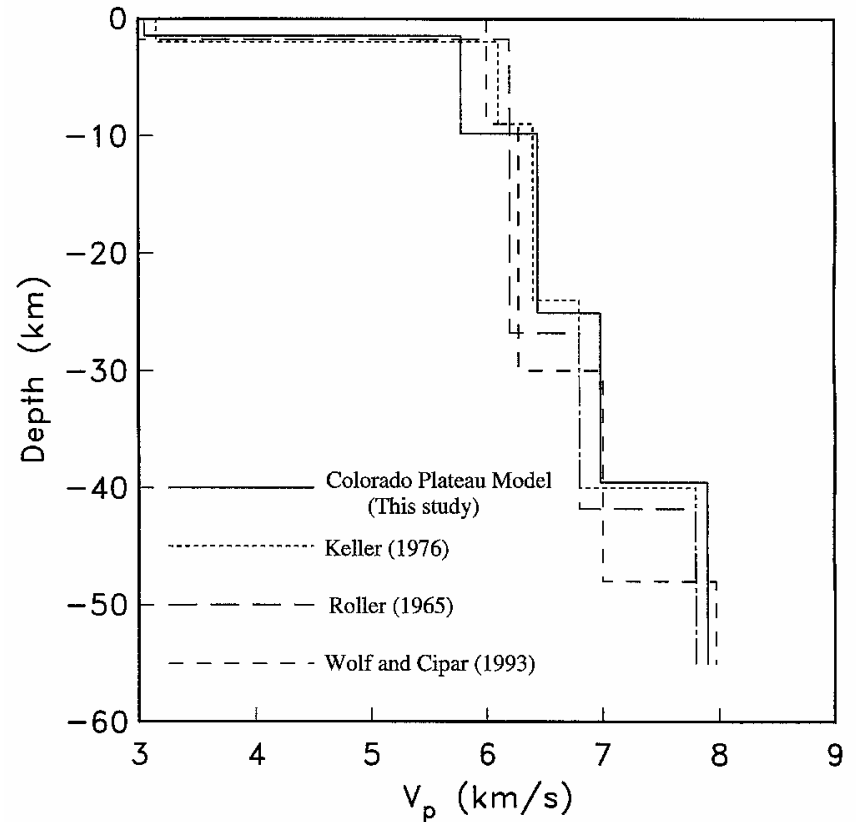
- ✓ They use 1-D velocity models and a known seismic source, so the number of parameters in each model can be reasonably small.
- ✓ Fitness (the cost function) is a function of the cross-correlation coefficient.
- ✓ Tests with synthetic data are useful in order to check for convergence.

Examples: Velocity structure using waveform fitting

(Bhattacharyya et al., BSSA, 1999)



Test with synthetic data



Results with real data

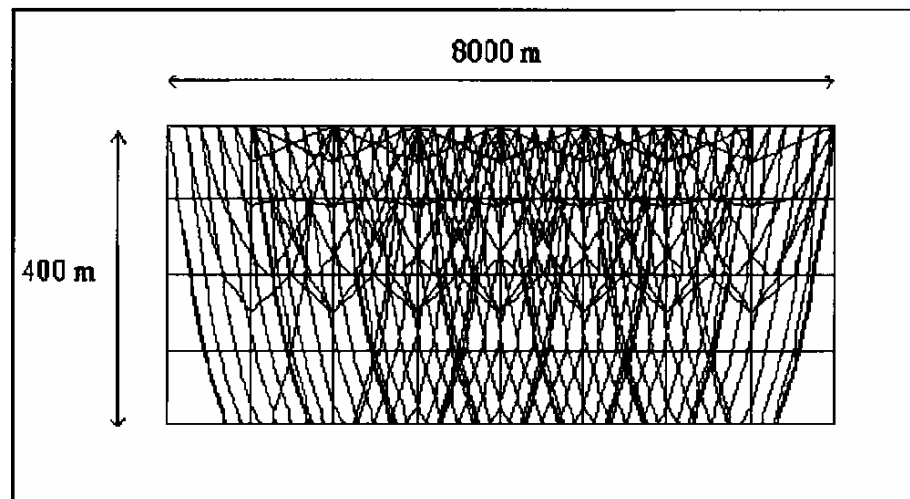
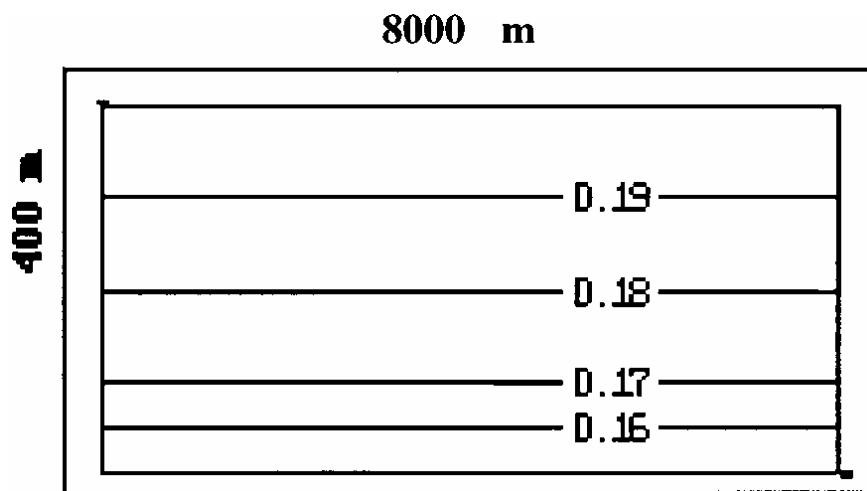
Examples: Seismic Refraction data

(Boschetti et al., 1996, Geophysics, Jin and Madariaga, 1993, G.R.L.)

- ✓ They want to solve for velocity structure in a grid of nodes.
- ✓ They produce synthetic data: waveforms and/or arrival times and/or migrated data.
- ✓ Population: P-wave velocity values in a grid of nodes. Each member of the population can be represented in binary or decimal form.
- ✓ High dimensionality might be a problem: Subdivisions of the model space might be necessary (pseudospace methods).

Examples: Seismic Refraction data

(Boschetti et al., 1996, Geophysics)

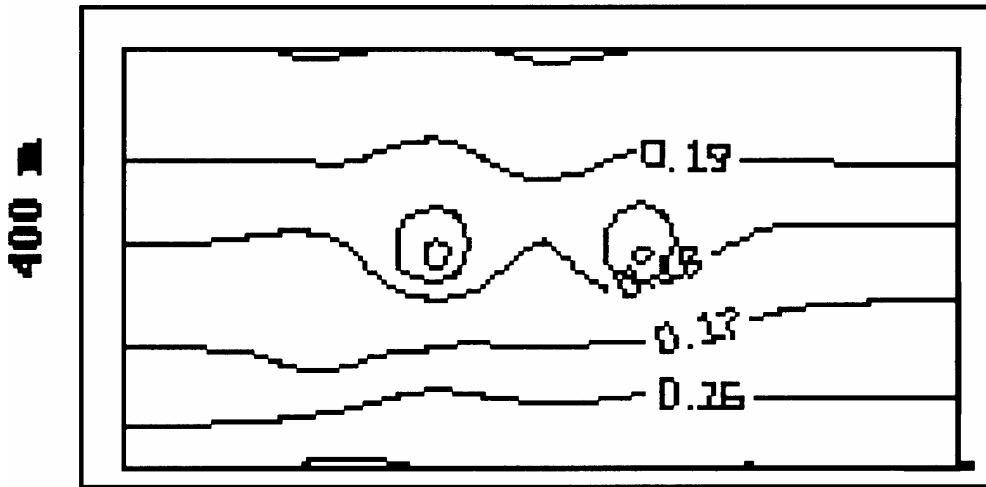


Synthetic data for slowness

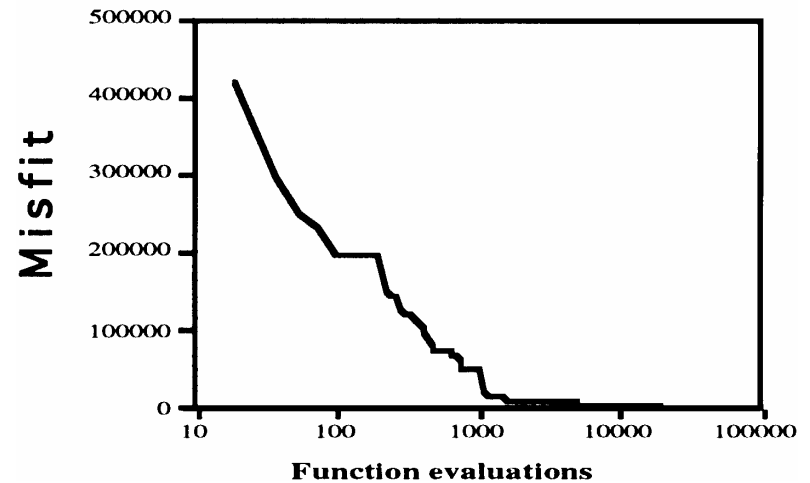
Examples: Seismic Refraction data

(Boschetti et al., 1996, Geophysics)

8000 m



Linear Normalisation Sel.
Convergence curve

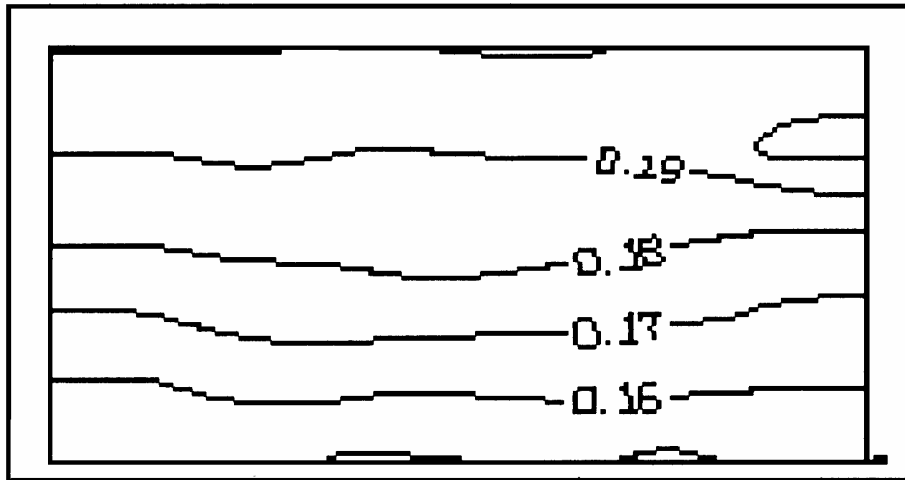


Results of the regular Genetic Algorithm

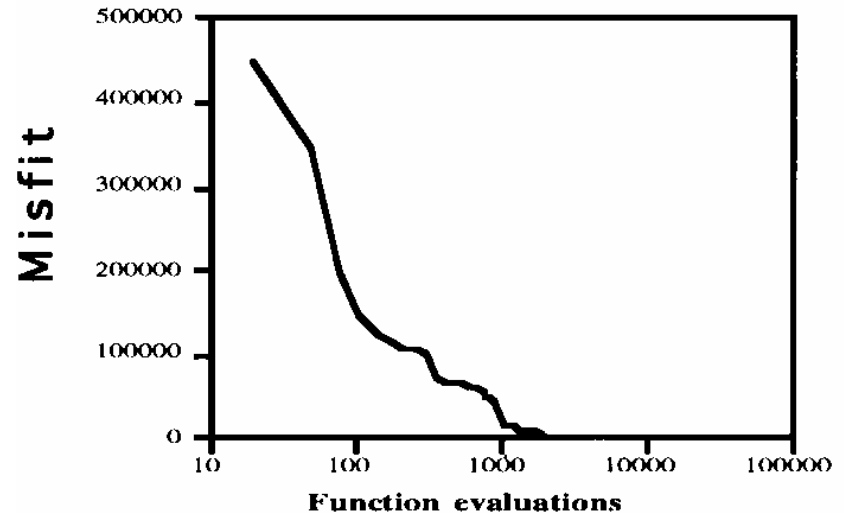
Examples: Seismic Refraction data

(Boschetti et al., 1996, Geophysics)

8000 m



Linear Norm. Sel. with Pseudo Subspace Met.
Convergence Curve

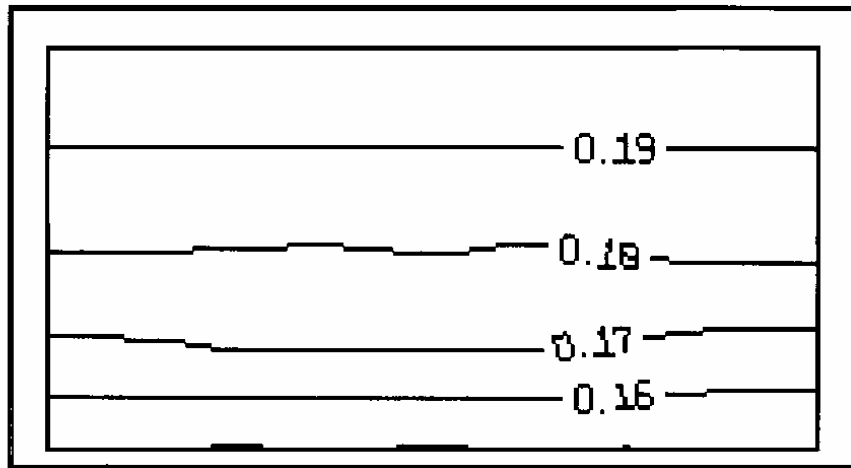


Genetic algorithm + pseudospace method

Examples: Seismic Refraction data

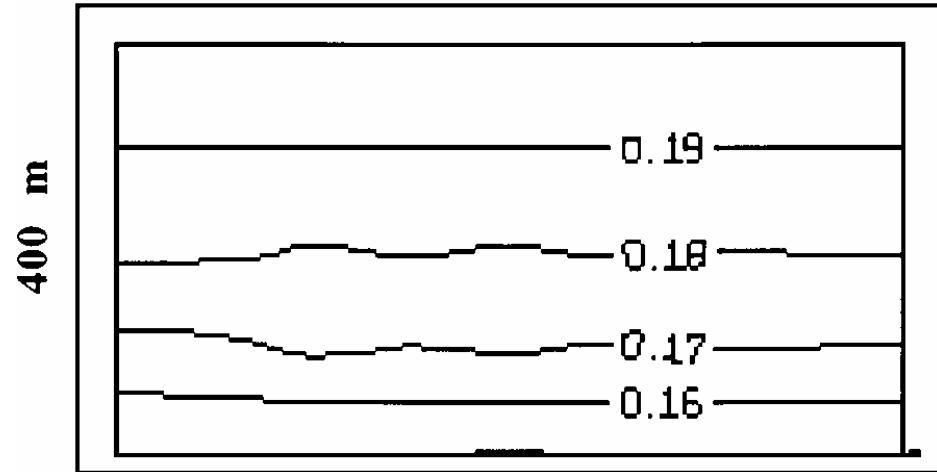
(Boschetti et al., 1996, Geophysics)

8000 m



100 generations of GA +
local search

8000 m



150 generations of GA +
local search

Conclusion

- The main advantage of the GA is the ability to converge to a global minimum.
- It is useful for geophysical applications when combined with local optimization methods.