

Earthquake location — genetic algorithms for teleseisms

B.L.N. Kennett^a and M.S. Sambridge^b

^a *Research School of Earth Sciences, Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, Australia*

^b *Institute of Theoretical Geophysics, Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK*

(Received 31 January 1992; accepted 1 July 1992)

ABSTRACT

Kennett, B.L.N. and Sambridge, M.S., 1992. Earthquake location — genetic algorithms for teleseisms. *Phys. Earth Planet. Inter.*, 75: 103–110.

The location of earthquakes requires the estimation of the spatial and temporal components of the hypocentre. This can be achieved by a direct minimisation of a measure of the misfit between observed and calculated travel times, and also slownesses and azimuths if array data are available. An efficient means of carrying out this optimisation procedure is to make use of genetic algorithms. This technique is based on the use of many estimates of the hypocentre location at once and the properties of the cluster of estimated locations in four dimensions are exploited in the course of the optimisation process. Each estimate of the hypocentral location is represented on a local discrete grid by a bit-string and successive iterations generate new bit-strings (and hence location estimates) by operations based on biological analogues. These operations are the replication of the best-fitting bit-strings, the cross-over of information between pairs of bit-strings and the mutation of individual bits in a string. The non-local character of the information on the misfit function carried in the cloud of hypocentral estimates is usually sufficient to prevent the location being trapped in local minima of the misfit surface. Convergence to the global misfit minimum can be achieved with a very limited sampling of the original spatial and temporal grid. No derivatives of the seismic phase information are required and so the technique is easily generalised to three-dimensional velocity models, and can be used with any suitable measure of the quality of an earthquake location by the choice of the misfit criterion between observed and calculated quantities.

1. Introduction

The process of estimating the hypocentre of an earthquake involves a number of stages. Firstly, the picking of the arrival times of different seismic phases and an association with a particular type of propagation process. Secondly, the choice of the measure of misfit between observed and computed arrival times (for three-dimensional (3-D) models this should include an allowance for modelling error). Thirdly, the application of the algorithm by which the measure of misfit is reduced to an acceptable level.

The majority of existing algorithms require the construction of the partial derivatives of the travel time for a particular phase with respect to the spatial location. This is not too difficult for 1-D models in which the seismic velocities vary only with depth for a limited number of phases, but is much less convenient where the wavespeeds vary in three dimensions.

There is considerable merit in developing earthquake location schemes that are not dependent on the differentiation of the measure of the misfit between observed and calculated times, and hence on the partial derivatives of travel times. Such methods can be used with any convenient method of calculating travel times, e.g. the Buland and Chapman (1983) algorithm for radial Earth models, or for 3-D varying situations the

Correspondence to: B.L.N. Kennett, Research School of Earth Sciences, Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, Australia.

finite-difference method of Vidale (1990) or the graph theoretic approach of Moser (1991). Such non-linear algorithms also allow the choice of a measure of the quality of an estimated location that can be based on realistic statistics, and which is not constrained to be some variant of weighted sum of squared residuals.

Sambridge and Kennett (1986) have introduced a method for the location of regional events based on a directed grid search in 4-D hypocentre space in which progressively tightening bounds on the origin time are used to constrain the spatial components of the hypocentre.

For teleseisms, Kennett (1992) has described one class of non-linear optimisation scheme for earthquake location based on the use of a nested grid search in 4-D space. This approach works well even when the initial constraints in the solution are quite weak, e.g. 2° in latitude and longitude, 60 km in depth and 6 s in origin time. However, this procedure requires a large number of travel-time evaluations, particularly in the initial exploration of the model space. With a robust estimate of location misfit based, for example, on an L1 norm, good results can be obtained but it is difficult to be sure that a global minimum of the misfit function is reached. In many location problems the behaviour of the misfit function in 4-D space is rather complex with many local minima.

An alternative approach which is very effective in finding the global minimum of the misfit function with a modest number of travel-time evaluations is based on the use of genetic algorithm techniques. Such genetic algorithms were developed in work on decision making in artificial intelligence research and have been recently applied to a range of geophysical problems by Scales et al. (1991), Sen and Stoffa (1992) and Gallagher et al. (1991). The novel feature of such an algorithm is that it works with many estimates of the hypocentre location at once and uses the properties of the cluster of estimates to drive the optimisation process. Because the information content of the cluster of hypocentre estimates is not confined to the immediate locality of a particular combination of hypocentre parameters, it is possible to escape from local minima of the misfit function on the way to the global minimum.

2. The genetic algorithm

The genetic algorithm is based on working with a group of Q hypocentral estimates simultaneously. Each hypocenter is represented by a bit-string: the initial search domain for each hypocentral parameter is divided into 2^n parts so a location can be described by n bits, and the combination of the representations for each of the parameters gives the requisite bit string. For the example below we used the same initial search domain as in the nested grid search, with 256 points in latitude and longitude (i.e. eight bits each) with a grid spacing of 0.008° , and 128 points in depth and time (seven bits each) with a grid of 0.469 km in depth and 0.0469 s in time. The bit-string for each hypocentre estimate, denoted as \mathbf{h} , is therefore $8 + 8 + 7 + 7 = 30$ bits long, and we have used a set of 24 hypocentre estimates. The initial population of Q hypocentres is created using a Monte-Carlo procedure.

We need to specify a misfit function $C(\mathbf{h})$ which provides a measure of the difference between the observed quantities (arrival times, azimuths and slownesses) and those calculated for the suggested hypocentre \mathbf{h} . A single iteration of a genetic algorithm is based on three stages, these are discussed in some detail in Sambridge and Drijkoningen (1992) and we have attempted to summarise the behaviour in Fig. 1. On the left we have a set of bit-strings representing the current estimates of hypocentre location; on the right the new generation of estimates produced by the action of three basic processes, 'replication', 'cross-over' and 'mutation'.

Replication. From the initial population of Q bit-strings an interim population of Q parents is generated by selecting models from the original group with the likelihood of selection determined by a probability depending on misfit, for the k th model

$$P_r(\mathbf{h}_k) = A \exp[-BC(\mathbf{h}_k)].$$

The net result is that those hypocentral estimates with the smallest current level of misfit can be passed unscathed to the new set. Based on the biological analogue this process is termed 'replication'.

Cross-over. From the parent population of Q bit-strings a new generation of Q strings are derived each of which is derived from a mixing of the bit-strings from two parents. All the Q parents are randomly paired to produce $Q/2$ couples. A cross-over probability P_c is assigned and if a random number between 0 and 1 is less than P_c parts of the two strings are interchanged. If such a 'cross-over' is selected, the location at which the strings are cut is determined randomly, other-

wise the two parent strings are passed unscathed to the next generation.

Mutation. The final process is 'mutation' in which (with a rather low probability P_m) any bit in an individual string is allowed to flip between 0 and 1 so that some degree of local diversity is introduced into the inversion process.

The action of the three types of process is to produce a new generation of Q bit-strings, each

TABLE 1

Marianas event — hypocentre solutions and data

Location		GA	ISC			
Latitude (N)		14.473	14.500			
Longitude (E)		146.755	146.900			
Depth (km)		62.055	33.000			
Origin time (s)		53.988	49.400			
L1 misfit		0.684	1.363			
L2 misfit		0.735	2.766			
Station	Phase		Residual	Delta (deg)	Back azimuth (deg)	
<i>Arrival times</i>						
		Min	s			
GUA	Pn	18	24.20	0.06	2.02	242.8
GUMO	Pn	18	24.20	-0.17	2.03	244.6
WB2	P	24	52.20	-0.11	36.31	200.0
BJI	P	24	57.00	-0.43	36.93	319.5
PKI	P	27	43.60	1.03	58.22	293.6
KKN	P	27	44.50	1.14	58.33	293.9
DMN	P	27	45.80	1.35	58.49	293.6
COL	P	28	38.20	-0.02	66.65	25.0
INK	P	29	17.00	0.70	72.86	22.5
YKA	P	30	4.40	1.00	81.23	27.7
SOD	P	30	35.00	-0.69	87.66	340.4
KJF	P	30	41.00	-1.05	89.00	337.5
SUF	P	30	48.00	-0.59	90.39	336.6
NUR	P	30	56.00	-1.14	92.23	335.2
HFS	P	31	17.30	-0.05	96.66	338.4
<i>Azimuth at array</i>						
			Degrees			
WB2	P		200.000	0.569	36.31	200.0
YKA	P		104.250	0.054	81.23	27.7
<i>Slowness at array</i>						
			s deg ⁻¹			
WB2	P		8.880	0.334	36.31	200.0
YKA	P		5.550	0.222	81.23	27.7

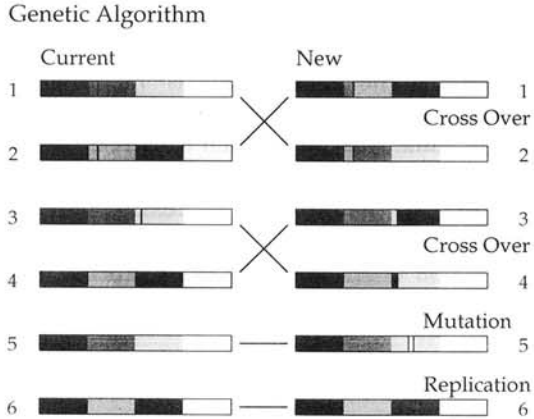


Fig. 1. Schematic representation of the action of the genetic algorithms in modifying one generation of hypocentral estimates represented by bit-strings into a new generation of estimates. The three basic operations are replication, cross-over between pairs of bit-strings and mutation of a single bit. Each of these processes is controlled by a separate probability distribution.

of which corresponds to a new hypocentre estimate within the prescribed representation. Each new estimate is then tested for its level of misfit and then the cycle of the genetic algorithm is repeated and a new generation of hypocentre estimates are created. We have found it advantageous to ensure that the best hypocentre in any generation is automatically passed to the next, since this enhances the rate of convergence of the procedure.

Because a number of different probability distributions enter into the genetic algorithm, optimum performance requires a certain amount of tuning. From a range of empirical tests for many earthquakes with different geographic station distribution, we have found that good performance can be achieved for teleseismic location with a cluster of 24 hypocentral estimates, a probability of 0.9 for cross-over between bit-strings and a probability of 0.04 for mutation of any bit.

3. Application of the genetic algorithm

Sambridge and Gallagher (1992) describe the application of the genetic algorithm in the con-

text of regional events. We will here discuss the application of this approach to teleseismic events, using the recent set of global travel times presented by Kennett and Engdahl (1991).

An advantage of the non-linear approach to location is that we can readily include different types of information. For the example we have used 15 travel-time readings for P taken from the ISC Bulletin for an event in the Marianas for which the azimuthal distribution of station coverage is limited. We have supplemented this travel-time information with simulated azimuth and slowness data for two arrays (WRA in Australia, YKA in Canada) in order to demonstrate that all the possible information for teleseismic location can be readily combined in a single operation. The misfit criterion employed was based on a sum of L1 contributions for each of the travel-time, azimuth and slowness residuals, with normalization by estimated errors. The azimuth and slowness data are assigned the same weight as the individual travel times. The data set used for this example is presented in Table 1, together with the residuals for the final location determined using the genetic algorithm. We also display the ISC solution for the event using only travel-time data, which has a constrained depth of 33 km. Note that the ISC calculation is based on the Jeffreys and Bullen (1940) travel-time tables rather than the iasp91 travel times (Kennett and Engdahl, 1991) used for the genetic algorithm.

The example was run using the genetic algorithm procedure for 54 generations and a total of 1320 hypocentres were tested in the progress to the global minimum of the misfit function out of approximately 10^9 grid locations in the region specified at the outset. Further generations gave no significant improvement and the achievement of the global minimum can be judged by the contour slices through the 4-D misfit function displayed in Fig. 4. For each trial location the travel times and slowness were calculated using the tau-spline representation of the iasp91 travel times, with corrections for ellipticity. The fast access to the travel-time information provided by the tau-spline representation (Buland and Chapman, 1983) facilitates the application of the genetic algorithm.

TABLE 2
Marianas event — location of successive hypocentral estimates

	Latitude N (deg)	Longitude E (deg)	Depth (km)	Time (s)	L1 misfit
ISC solution	14.500	146.900	33.000	49.400	1.363
Reference point for plotting	14.475	146.800	60.000	53.000	—
Monte-Carlo	14.739	146.473	33.236	53.605	1.532
2	14.535	146.818	57.803	52.943	0.808
3	14.535	146.818	57.803	53.510	0.776
4	14.394	146.825	58.276	53.510	0.761
5	14.410	146.810	57.803	53.510	0.752
6	14.425	146.818	57.803	53.510	0.726
7	14.488	146.771	59.693	53.510	0.715
8	14.488	146.755	62.055	53.888	0.694
9	14.473	146.755	62.055	53.888	0.687
Without array azimuth or slowness	14.472	146.747	60.638	53.888	0.654

The genetic algorithm has been able to achieve rapid convergence to a global minimum in this non-linear optimisation problem without extensive sampling of the model space. This arises from the range of information on the character of the misfit function carried within the cluster of Q hypocentral estimates as it traverses the 4-D location space. Both local information derived from the effects of perturbations of low order bits and

global information from the sampling of the many different location estimates are exploited in the progress to a minimum. The application of a fully Monte-Carlo procedure to the same data set led to slow convergence (at least 50 times as many

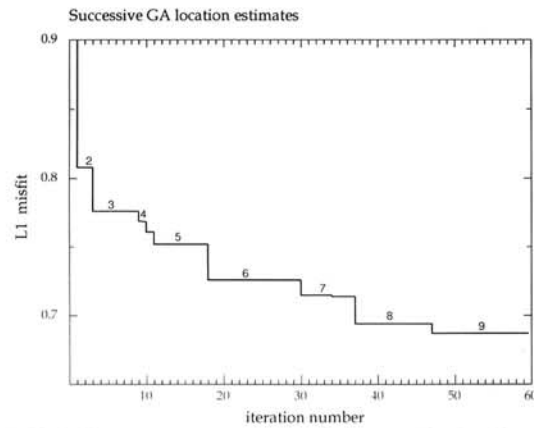


Fig. 2. The convergence pattern of the genetic algorithm using an L1 misfit function for a teleseismic location for an event in the Marianas arc using both travel-time and array data. The figures associated with each plateau in the misfit measure as a function of iteration number refer to the hypocentral estimates plotted in Fig. 3. The plot is terminated once the global minimum is reached.

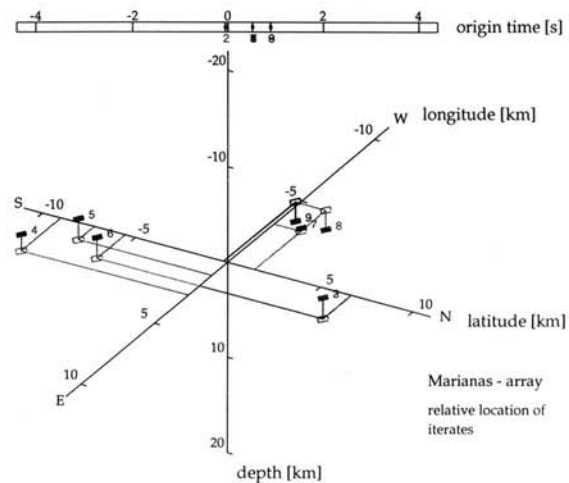


Fig. 3. Four-dimensional representation of the progress of the genetic algorithm in generating the estimate of the location of a teleseismic event in the Marianas arc using both travel-time and array data using an L1 misfit function. The reference location has been chosen to allow the various stages of the location procedure to be conveniently displayed and does not represent the centre of the original grid. The solid symbols indicate the spatial and temporal locations and are marked with key numbers which refer to the misfit behaviour in Fig. 2. The best hypocentral estimate is indicated by '9'.

hypocentral estimates were needed to achieve the same level of misfit).

Figures 2 and 3 show the progress of the location algorithm for the Marianas example by tracking the misfit and the 4-D representation of the current best hypocentral estimate. In Fig. 2, each of the plateaux in the misfit function represent the best fit currently available until the string manipulation induced by the genetic algorithm creates an improved hypocentre. These plateaux in the misfit function represent the locations of local minima in the 4-D behaviour, and we have found that it is necessary to allow relatively high probabilities of mutation of individual bits in order that an escape can be made from such local depressions in the misfit surface. It is unfortunately rather difficult to develop an adequate 4-D display of the properties of the misfit function to show the nature of the local minima.

However, the mapping of the movement of the best hypocentral estimate in four dimensions in

Fig. 3 indicates the way in which the algorithm is working. The reference location used for the centre of the display in Fig. 3 has been chosen to allow a convenient representation of the various hypocentral estimates in a single diagram. The reference point is displaced from the ISC location, which was used as the centre of the 4-D grid, by 2.75 km in latitude, 10.75 km in longitude, 27 km in depth and 3.6 s in time. The sequential estimates for the hypocentral locations for this example are compared in Fig. 3 and Table 2. We find there is a tendency for these estimates to come in clusters, connected by the effect of a mutation step, separated by larger jumps in position induced by the action of the bit-string exchange in the cross-over step.

The successful global minimisation achieved for this example is illustrated by Fig. 4 which shows four contoured slices through the misfit function passing through the solution determined by the genetic algorithm process. The slices show

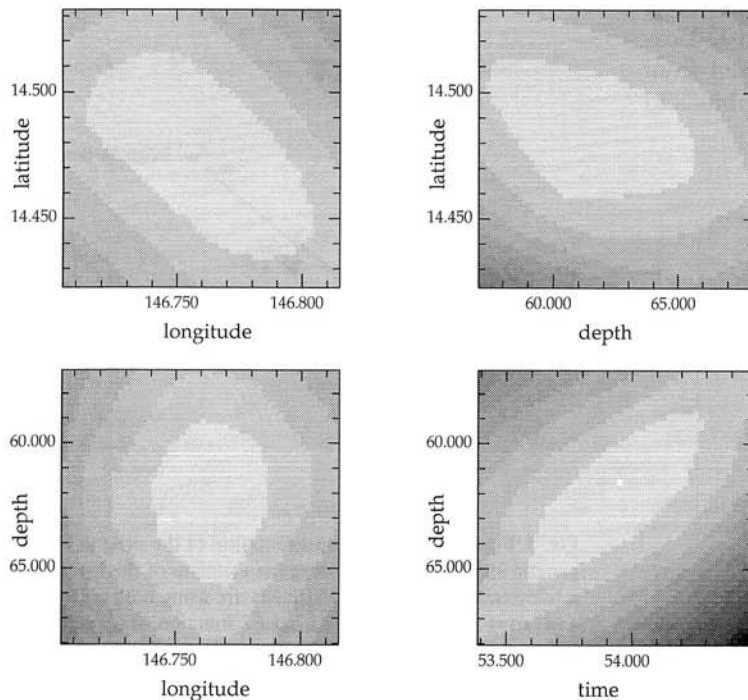


Fig. 4. Representation of the behaviour of the L1 misfit function in the neighbourhood of the hypocentral location determined by the genetic algorithm. Slices through the proposed solution (9 in Fig. 3) are shown in latitude/longitude, latitude/depth, longitude/depth and depth/origin-time. Progressively darker grey tones indicate increasing misfit.

the cross-section in latitude/longitude, latitude/depth, longitude/depth and depth/origin-time. In each case the minimum is centred on the solution determined by the genetic algorithm procedure. As the misfit increases the toning employed becomes darker. The contours of the misfit function can be interpreted in terms of the appropriate probability function as discussed by Sambridge and Kennett (1986). For the L1 measure of misfit used in this example, this would be the exponential distribution. We note that the central region is quite irregular in shape and occupies a zone which is approximately 5 km in each linear dimension and 0.5 s in time. This is a realistic estimate of the likely precision of a hypocentral estimate. However, by making a different choice of location criterion it is often possible to move the position of the estimated hypocentre by rather more than the formal error limits (see, e.g. Kennett, 1992).

In this example it was possible to escape from all the local minima of the misfit function and to reach the global minimum and hence the best attainable location for the given data and choice of misfit function. This does not always occur, and it is certainly possible for the genetic algorithm to get trapped by a deep local minimum. Indeed we have found it necessary to raise the probability of the mutation of an individual bit to a level which leads to the preservation of diversity in the set of hypocentral estimates as the generations progress, rather than the convergence of the cluster of estimators with successive generations demonstrated in the work of Sen and Stoffa (1992) on the application of similar techniques to waveform inversion.

4. Discussion

With this example we have demonstrated that practical non-linear inversion can be achieved for teleseismic problems without the need for any calculation of partial derivatives of travel times. The genetic algorithm procedure exploits the properties of a cluster of many simultaneous estimates for the hypocentral location and is able to

extricate itself from local minima in the misfit function. The procedure is particularly useful when many different types of information are employed because each type of information is treated in the same way. It is thus easy to combine the travel times from different types of seismic phases and array data (azimuth and slowness), which is very helpful for initial determinations of hypocentres using a limited teleseismic network. Further, since there is no need to calculate any derivatives of travel times, the genetic algorithm approach can be readily extended to any three-dimensionally varying velocity model for which two point ray tracing can be performed.

An attractive feature of the genetic algorithms is the ease with which different types of seismic information can be incorporated into the location procedure. In the example we have used only P-wave data but there is no difficulty to include later phase information for which the iasp91 travel-time tables are particularly useful. When later phases are used, there is an additional problem of ensuring that the correct phase identification is applied to individual arrivals. The problem of assigning phase types without knowing the correct location poses a further class of inverse problem for which the genetic algorithms appear to be well suited, and we are currently investigating ways of incorporating phase association into the location procedure.

References

- Buland, R. and Chapman, C.H., 1983. The computation of seismic traveltimes. *Bull. Seismol. Soc. Am.*, 73: 1271–1302.
- Gallagher, K.L., Sambridge, M.S. and Drijkoningen, G.G., 1991. Genetic algorithms: an evolution on Monte Carlo methods in strongly non-linear geophysical optimization problems. *Geophys. Res. Lett.*, 18: 2177–2180.
- Jeffreys, H. and Bullen, K.E., 1940. *Seismological Tables*. British Association for the Advancement of Science, London.
- Kennett, B.L.N., 1992. Locating oceanic earthquakes — the influence of regional models and location criteria. *Geophys. J. Int.*, 108: 848–854.
- Kennett, B.L.N. and Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification. *Geophys. J. Int.*, 105: 429–465.

- Moser, T., 1991. Shortest path calculation of seismic rays. *Geophysics*, 56: 59–67.
- Sambridge, M.S. and Drijkoningen, G., 1992. Genetic algorithms in seismic waveform inversion. *Geophys. J. Int.*, 109: 323–342.
- Sambridge, M.S. and Gallagher, K.L., 1992. Earthquake hypocentre location using genetic algorithms. *Bull. Seis. Soc. Am.*, submitted.
- Sambridge, M.S. and Kennett, B.L.N., 1986. A novel method of hypocentre location. *Geophys. J. R. Astron. Soc.*, 87: 313–331.
- Scales, J.A., Smith, M.L. and Fischer, T.L., 1991. Global optimization methods for highly non-linear inverse problems. *J. Comp. Phys.* (in press).
- Sen, M. and Stoffa, P.L., 1992. Rapid sampling of model space using genetic algorithms. *Geophys. J. Int.*, 108: 281–292.
- Vidale, J.E., 1990. Finite-difference calculation of traveltimes in three dimensions. *Geophysics*, 55: 521–526.