Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 506553
UNCLASSIFIED	

TITLE

GENETIC ALGORITHM INVERSION OF THE 1997 GEOACOUSTIC INVERSION WORKSHOP TEST
CASE DATA

System Number:

Patron Number:

Requester:

Notes:

Genetic Algorithm Inversion of the 1997 Geoacoustic Inversion Workshop Test Case Data

Garry J. Heard

Defence Research Establishment Atlantic
PO Box 1012 Dartmouth, NS B2Y 3Z7, CANADA

David Hannay and Scott Carr

Jasco Research Ltd.
#102–7143 West Saanich Road, Brentwood Bay, BC V8M 1P7, CANADA

An analysis of the 1997 Geoacoustic Inversion Workshop test case data was carried out by Jasco Research Ltd. under the auspices of a research contract let by the Defence Research Establishment Atlantic. The analysis of the test cases served to benchmark the performance of a Genetic Algorithm (GA) inversion code written by Jasco Research called SAGA_INV. The inversion program made use of Westwood's ORCA propagation model, FORTRAN subroutines, and Interactive Data Language (IDL) by Research Systems Incorporated. The code is designed to run on UNIX (or Linux) workstations, but should be relatively simple to port to other operating systems. SAGA_INV is capable of performing inversions with either Simulated Annealing (SA) or GA optimisation schemes; however, only the GA portion of the code has been benchmarked with the workshop test cases at the present time. Not all of the workshop test cases were processed: this study was concerned only with the CAL, SD, SO, AT, and WA data sets. The CAL data was processed using three different cost functions: i) standard Bartlett processor, ii) a broadband coherent processor, and iii) a transmission loss mismatch function. These processors were applied to three frequency bands: i) 76 frequencies between 25 and 100 Hz, ii) 9 frequencies between 28 and 36 Hz, and iii) 13 frequencies between 44 and 56 Hz. The latter two frequency regimes were intended to simulate 1/3-octave bands centred at 32 and 50 Hz, respectively. Four different receiving arrays were simulated: i) a 1550-m aperture horizontal, bottom mounted array at approximately 1 km range, ii) a similar array at approximately 4.2 km range, iii) a 55-m aperture 12-element vertical array located at 1 km range, and iv) a similar vertical array at 5 km range. In addition to processing the CAL data set, all three realizations of the SD, SO, AT, and WA data sets were also processed; however, only the transmission loss cost function and the two simulated 1/3-octave bands were considered for these test cases.

1.   Introduction

The Geoacoustic Inversion Workshop[1], held at MacDonald-Dettwiler in Richmond BC, provided an opportunity for researchers to try their inversion schemes in a blind test using a number of synthetic data sets of varying complexity and to then compare results. In this paper the inversion results obtained with a new Simulated Annealing (SA) and Genetic Algorithm (GA) optimisation program called SAGA_INV are presented for a number of different cases involving three different cost functions, several processing bandwidths, and simulated vertical and horizontal receiving arrays. The data sets used in the inversions are described elsewhere[2] those that were considered for this paper include: CAL, SD, SO, AT, and WA test cases. Details of the processing applied to each data set will be given in the next section of this paper.

SAGA_INV[3] was developed by Jasco Research Ltd. under contract to Defence Research Establishment Atlantic (DREA). SAGA_INV is capable of using both SA and GA optimisation schemes. All of the inversion results presented in this paper were obtained using the GA routines within SAGA_INV.

Table 2.1. Known parameters and search ranges for each environment considered.

Parameter		CAL	SD	SO	AT	WA
				Environment		
Z_s	[m]	20	20	20	20	10-30
ΔR	[m]	0	0	0	0	0-400
D	[m]	100	100	100	100	100-120
$c(0)$	[m/s]	1480	1480	1480	1480	1480
$c(D)$	[m/s]	1460	1460	1460	1460	1460
h_{sed}	[m]	50	10-50	10-50	10-50	10-50
$c_{sed}(D)$	[m/s]	1550	1500-1600	1500-1600	1500-1600	1500-1600
$c_{sed}(D+h_{sed})$		1700	1550-1750	1550-1750	1550-1750	1550-1750
ρ_{sed}	[g/cc]	1.5	1.4-1.85	1.4-1.85	1.4-1.85	1.4-1.85
α_{sed}	[dB/λ]	0.23	0.23	0.23	0.05-0.5	0.23
c_{hsp}	[m/s]	1800	1600-1800	1600-1800	1600-1800	0.23
ρ_{hsp}	[g/cc]	2.0	1.6-2.0	1.6-2.0	1.6-2.0	1.6-2.0
α_{hsp}	[dB/λ]	0.23	0.23	0.23	0.05-0.5	0.23

Genetic algorithms, like simulated annealing techniques, are what are known as directed search optimisation methods. While SA attempts to find optimum solutions to problems by emulating the annealing process of a cooling metal, which naturally finds a minimum energy state, GA models the natural process of genetic drift, which attempts to find the optimum adaptation of individuals to a given environment. Genetic algorithms include the basic steps of fitness evaluation, selection, reproduction, and mutation. An excellent overview on the subject has been provided by Beasley et al.[4]

The SA portion of SAGA_INV was derived from the GA portion and is therefore somewhat unusual in that discretized binary gene values rather than floating point values represent the state space variables. In other respects the SA portion of the code is quite conventional in that only one gene is perturbed at a time and all genes are perturbed at each temperature step. Gene perturbation sizes are selected from a cubed uniform distribution which favours smaller changes in value over large changes. The Metropolis[5] criterion is used to accept state changes. One benefit of using the GA framework to implement the SA code is that *a posteriori* statistics can be calculated at the completion of the SA inversions. The GA portion of SAGA_INV was considered to be the main part of the program and because of this only the GA routines have been used with the Workshop test cases at present. The GA code allows the user to control the number of independent populations and the number of individuals in the populations. The user also has control of the probabilities of mating and mutation. Individuals for mating are selected probabilistically and the most fit individuals are favoured. All state space variables are represented by fixed length genes and the user may select the number of bits in the variable representation. Uniform crossover masking is used to reduce the importance of gene placement within the chromosome. Processing continues until all populations have converged. Once convergence has been reached *marginal a posteriori* statistics[6] are calculated using stored cost values for each generation.

The inversion package consists of a widget-based GUI (graphical user interface) written in Interactive Data Language (IDL), FORTRAN subroutines, and Westwood's ORCA[7] propagation model. Currently this program suite runs on Linux-based PC systems and Silicon Graphics Indigo workstations. The total time required for an inversion run with 76 frequencies using a 133 MHz Pentium PC varied from 19 to 72 hours. A 1/3-octave inversion typically took 2 to 7 hours.

Seven test cases were generated for use at the workshop. This paper discusses results obtained from inversion of data from five of the test cases: CAL, SD, SO, AT, and WA. The CAL test case was special in that only one realization of the environment was given and all environmental parameters were known. The other test cases were basically similar to the CAL environment except that there were varying numbers and combinations of unknown parameters and that there were three independent realizations of these test cases (see solution sets A, B, and C in Table 4.1). We did not process the multi-layer test case N or the elastic test case EL.

The next section describes the data processing in more detail. Section 3 describes the results obtained with the CAL data set and Section 4 describes the results obtained by 1/3-octave bandwidth processing with the SD, SO, AT, and WA data sets. Finally, conclusions are presented in Section 5.

2. Data Set Processing

The GA inversions each included 64 independent populations with 32 individuals in each population. Eight bits were used in the gene representations providing 256 possibilities for each state variable. Since the minimum number of unknowns was 6 and the maximum was 9, the number of state conditions varied from approximately 2.8×10^{14} to 4.7×10^{21}. Approximately 60000 forward runs of the propagation model ORCA were required for each inversion, which implies that the efficiency of solution is very high as only a very tiny fraction of the possible states were searched.

Table 2.1 details the known values or the *a priori* search ranges for each of the environmental variables in the test cases. Source depth is denoted by Z_s, range uncertainty by ΔR, water depth by D, sound speed by c, layer thickness by h, density by ρ, and attenuation in dB/λ by α. Subscripts *sed* and *hsp* denote the sediment layer and basement half-space respectively.

For each inversion the probability of mating, *pmate*, was fixed at 85% and the probability of mutation at 3%. Mating is accomplished by selecting 32 individuals probabilistically from the fitness ordered set of individuals. The probability that the ith individual from the ordered set is chosen is given by $\exp[a/b \cdot i/nindiv]$, where a and b are the average and best cost function values for the current generation, and *nindiv* is the number of individuals in the population. The scaling factor, a/b, increases the probability of less fit individuals mating when the most fit individual is not significantly fitter than the average individual. This scaling increases gene diversity and helps reduce premature gene convergence. Mating employs the uniform crossover technique[4] which reduces the dependence on gene order within the chromosome. Note that not all of the selected 'parents' undergo mating by gene crossover. A fraction of the parent population 1-*pmate*, 15% in this case, are simply cloned into the following generation. Following the mating process, mutation is carried out on each offspring by complementing one randomly selected bit within the chromosome according to the selected mutation probability.

The cycle of cost evaluation, parent selection, mating, and mutation continues for each population in turn until 95% of individuals within a population share the same chromosome. This is the convergence condition described by Beasley.[4] When all populations have converged, processing stops and the *marginal a posteriori* statistics are computed from the stored costs for every forward model run. The *marginal a posteriori* distribution for parameter *j* was calculated according to

$$\sigma_j(x_j) = \frac{\sum_{v=1}^{N_{obs}} \exp\left[-\phi(\mathbf{m}^v)/T\right]\delta(\mathbf{m}_j^v - x_j)}{\sum_{v=1}^{N_{obs}} \exp\left[-\phi(\mathbf{m}^v)/T\right]},$$

2.1

where N_{obs} is the total number of models tested, $\mathbf{m}^v$ is the vth model vector, $\phi(\mathbf{m}^v)$ is the cost function value for model v, and T is the temperature parameter. The value of σ_j is computed for each possible value of the jth parameter denoted by x. The temperature T was set to the average cost function value obtained during the inversion. This differs from Gerstoft[5] who sets the temperature to the average of the best 50 cost values.

2.1. Cost Functions

The CAL test case was inverted using three different cost functions: a standard Bartlett processor, a broadband coherent processor, and a transmission loss (TL) mismatch function. The Bartlett processor is given by

$$C_B(\mathbf{m}) = 1 - \frac{1}{M}\sum_{k=1}^{M} \frac{\sum_{i=1}^{N}\left|Q_i^*(\mathbf{m},\omega_k)P_i(\omega_k)\right|^2}{\sum_{i=1}^{N}\left|Q_i(\mathbf{m},\omega_k)\right|^2 \sum_{i=1}^{N}\left|P_i(\omega_k)\right|^2},$$

2.2

where N is the number of hydrophones, M is the number of frequencies, $P_i(\omega_k)$ is the complex pressure on the ith hydrophone at frequency k, and $Q_i(\mathbf{m},\omega_k)$ is the corresponding value calculated by the propagation model for input parameter vector $\mathbf{m}$. The coherent processor previously described by Hannay & Chapman[8] was modified to eliminate the search over time uncertainty τ, because absolute timing of the received signals with respect to source transmission was available. The coherent processor is given by

$$C_C(\mathbf{m}) = 1 - \frac{\left[\Re\left\{\sum_{i=1}^{N}\sum_{k=1}^{M} Q_i^*(\mathbf{m},\omega_k)P_i(\omega_k)\right\}\right]^2}{\sum_{i=1}^{N}\sum_{k=1}^{M}\left|Q_i(\mathbf{m},\omega_k)\right|^2 \sum_{i=1}^{N}\sum_{k=1}^{M}\left|P_i(\omega_k)\right|^2},$$

2.3

where $\Re$ represents the real part of the expression in braces. The TL mismatch function is the root-mean-square (rms) difference between the modelled and measured transmission loss averaged over all of the hydrophones and all frequencies

$$C_{TL}(\mathbf{m}) = \sqrt{\frac{\sum_{i=1}^{N}\sum_{k=1}^{M}\left[-20\log\left|Q_i(\mathbf{m},\omega_k)\right| - (-20)\log\left|P_i(\omega_k)\right|\right]^2}{N\cdot M}}.$$

2.4

Note that in the TL mismatch function the difference in TL is calculated in log-space. This has been done intentionally to emphasize the importance of nulls in the TL matching. Tests have been conducted in both linear and log domains, and it appears that the log domain is best suited for use with moderate to high signal-to-noise ratio data.

Only the TL processor was used on the SD, SO, AT, and WA test case data sets. This choice of processor was partly motivated by an inversion study[3] currently being completed by Jasco Research Ltd. for DREA.

2.2. Bandwidths

The CAL test case data was processed with three different bandwidths: 25–100 Hz, 28–36 Hz, and 44–56 Hz. The first band contained 76 frequencies spaced 1 Hz apart, while the two narrower bands contained 9 and 13 frequencies respectively, spaced 1 Hz apart. The narrower bands were an attempt to simulate 1/3-octave bands at 32 and 50 Hz. For ease of reference, the 76-frequency band will be referred to as the broadband data and the two narrower bands will be referred to by the centre frequencies of the 1/3-octave bands they are simulating.

At present the SD, SO, AT, and WA test case data sets have only been processed in the 1/3-octave bands (32 Hz band and 50 Hz band). These 1/3-octave bands were considered to be more typical of the bandwidths available in sea tests.

2.3. Simulated arrays

Each of the five test case data sets were processed using four different simulated arrays: two vertical arrays (VLA), and two horizontal arrays (HLA). The simulated VLA consisted of 12 hydrophones and spanned an aperture of 50 m. The VLA was located at two different ranges from the nominal source position (1 km and 5 km) with the hydrophones spanning the depth interval: 20–75 m. The simulated HLA consisted of 32 hydrophones spaced at 50 m, providing a 1550-m aperture. The HLA was positioned on the sea floor (100 m depth) with the source in endfire aspect at ranges of 50 and 3450 m relative to the first hydrophone in the array. Table 2.2 provides a summary of the array details. For ease of reference, the different arrays will be referred to as array 1 to 4 as denoted in column one of Table 2.2.

Table 2.2 Details of simulated arrays used for inversion.

	Type	Number of Elements	Spacing	Ranges	Depths
Array 1	HLA	32	50 m	50 – 1600 m	100 m
Array 2	HLA	32	50 m	3450 – 5000 m	100 m
Array 3	VLA	12	5 m	1000 m	20 – 75 m
Array 4	VLA	12	5 m	5000 m	20 – 75 m

Table 3.1. Summary of cost function and array type performance with broadband data.

Array	Trans. Loss 1	2	3	4	Std. Bartlett 1	2	3	4	Cross corr. 1	2	3	4
Z_s	●	●	●	○	●	●	●	●	●	○	●	○
D	○	●	●	○	●	●	●	●	●	●	●	●
ΔR	●	·	●	·	●	●	●	○	●	●	●	●
h_{sed}	·	·	○	·	○	●	●	·	●	·	●	·
$c(D)$	·	·	·	·	·	●	○	·	●	·	●	·
$c(D+h_{sed})$	·	·	·	·	·	●	·	·	●	·	●	·
ρ_{sed}	·	·	·	·	·	●	○	·	●	·	●	○
c_{hsp}	·	·	·	·	·	●	○	·	●	·	●	·
ρ_{hsp}	·	·	·	·	·	●	·	·	○	○	·	·

●   Parameter estimated within 25% of search range.

○   Parameter estimated within 50% of search range.

·   Parameter not localized.

3. CAL Data Set Results

The CAL data set was first processed with the broadband data using all three of the cost functions and all four simulated receiving arrays. From inspection of the resulting *marginal a posteriori* density distributions, it is possible to draw some slightly subjective comparisons regarding the variation of the inversion accuracy with cost function and array type. These comparisons are summarized in Table 3.1.

A large solid dot in Table 3.1 indicates that a reasonable value was obtained for the parameter of interest with the corresponding cost function and receiving array. An open circle indicates that a parameter estimate was obtained, but that there is a large uncertainty associated with the estimate. A small dot indicates that the parameter value was not estimated at all. Inspection of Table 3.1 shows that only about half of the cost function-array choices resulted in a useful estimate of at least some parameters. This is somewhat disappointing as a large amount of information was encoded in these 76-Hz bandwidth inversions and noise was not included in the synthetic test cases.

Several general observations can be made from the results in Table 3.1. First, is that with only one exception the arrays (1 & 3) which were closest to the source provided the best estimates of the geoacoustic parameters. The exception to this was array 2 with the Bartlett processor. It is not clear why the more distant array functioned best in this case. This particular solution may simply have been the result of a fortuitous selection of the state space parameters resulting in convergence to the true solution. Second, the coherent processor provided better overall results than either the Bartlett or TL processor. In particular, the results at close range were noticeably improved with the coherent processor. Results at greater ranges were not as successful. This performance decrease with range may indicate the existence of relative phase errors in the propagation models: something that should probably be investigated in further work. Third, the TL processor was the worst performer of the three cost functions investigated. This

Table 3.2. 1/3-Octave band inversion results using CAL test case data. Parameter search ranges were defined by the uncertainty ranges from Table 2.1, for the corresponding environment in column 1.

CAL	Z_s	ΔR	D	h_{sed}	$c_{sed}(D)$	$c_{sed}(D + h_{sed})$	ρ_{sed}	α_{sed}	c_{hsp}	ρ_{hsp}	α_{hsp}
sd	–	–	–	46.4	1548.3	1665.3	1.57	–	1730	1.86	–
at	–	–	–	44.7	1548.1	1683.6	1.58	0.22	1718	1.96	0.31
so	20.4	0.023	–	44.1	1538.5	1681.5	1.55	–	1761	1.81	–
wa	20.8	0.005	102.3	42.1	1556.0	1658.0	1.51	–	1699	1.84	–
True	20	0	100	50	1550	1700	1.5	0.23	1800	2.0	0.23

result is not particularly surprising, but there is considerable value in being able to use TL information in this way. Measurement of TL is very simple and robust. In addition, TL data is available from many past experiments where more detailed information is often not available. Fourth, somewhat unexpectedly, a VLA does not appear to lead to better geoacoustic estimates than an HLA. Of course, the reader needs to keep in mind that the number of phones and aperture of the HLA were considerably greater than those of the VLA in this study. However, the simulated arrays are not unrealistic and when one takes VLA motions into account, bottomed HLA arrays are an attractive option. Finally, source depth, source range, and water depth were the only parameters consistently estimated with useful uncertainties. Geoacoustic parameters were only estimated usefully in 25% of the broadband inversions.

Despite the negative comments of the last paragraph, useful results were obtained from inverting the CAL data set using the 32 and 50 Hz 1/3-octave bands and the TL cost function. In fact, the inversion of array 1 data in the 32 Hz band with the TL cost function outperformed the inversion using the broadband data. The 32 Hz inversion was also superior to the results obtained with the 50 Hz band. Among the possible reasons for the improved performance with the narrower processing bandwidth are that the lower frequencies contain the bulk of the information about the bottom and sub-bottom, or that relative modelling errors across the broadband data were sufficient to reduce the correlations. As previously, the nearer array (#1) produced a better result than the more distant array (#2) and the HLA receivers produced better results than the VLA receivers. Source depth, water depth, and range were again the best determined parameters, but the estimates were not limited to these variables. Table 3.2 summarizes the results obtained inverting the CAL data in 1/3-octave bands. The inversions were performed using four different sets of unknowns as indicated by the ad, at, so, and wa titles in the first column. The numbers presented represent averages of the individual results for the four different arrays and two 1/3-octave bands.

4.   One-Third Octave Results

In this section, the results of applying the TL cost function with the four different receiving arrays to data for the 32 and 50 Hz bands from test cases SD, SO, AT, and WA are discussed. It should be noted here that the AT environment with a bottomed HLA caused the ORCA model to crash. This is probably due to having searched for the mode values at the sea floor where the receiver was located and may also involve the sediment attenuation values which covered a large range (0.05–0.5 dB/λ). The results of the 1/3-octave inversions for each of the separate realizations A, B, and C are shown in Table 4.1 where the true values are given in the bottom part of each block and the values obtained by averaging the values corresponding to the highest peaks in the *a posteriori* distributions for each array and frequency band are given in the upper part of each block.

Table 4.1. Average results from 1/3-octave band inversions and true values.

SD	Z_s	ΔR	D	h_{sed}	$c_{sed}(D)$	$c_{sed}(D+h_{sed})$	ρ_{sed}	α_{sed}	c_{hsp}	ρ_{hsp}	α_{hsp}
A	-	-	-	17	1565	1705	1.75	-	1745	1.84	-
B	-	-	-	43	1589	1685	1.61	-	1690	1.97	-
C	-	-	-	29	1513	1640	1.49	-	1674	1.90	-
True A	20	0	100	22.5	1565.1	1743.1	1.76	0.23	1757.7	1.83	0.23
True B	20	0	100	38.0	1599.3	1658.5	1.64	0.23	1707.5	1.87	0.23
True C	20	0	100	30.7	1530.4	1604.2	1.50	0.23	1689.0	1.70	0.23

AT	Z_s	ΔR	D	h_{sed}	$c_{sed}(D)$	$c_{sed}(D+h_{sed})$	ρ_{sed}	α_{sed}	c_{hsp}	ρ_{hsp}	α_{hsp}
A	-	-	-	41	1542	1609	1.47	0.18	1655	2.00	0.35
B	-	-	-	13	1518	1710	1.67	0.27	1747	1.86	0.35
C	-	-	-	35	1519	1575	1.42	0.09	1625	1.88	0.38
True A	20	0	100	44.6	1546.9	1624.8	1.40	0.12	1676.3	1.78	0.08
True B	20	0	100	13.4	1558.3	1564.6	1.77	0.03	1758.8	1.95	0.20
True C	20	0	100	30.4	1533.9	1557.6	1.43	0.08	1615.2	1.69	0.05

SO	Z_s	ΔR	D	h_{sed}	$c_{sed}(D)$	$c_{sed}(D+h_{sed})$	ρ_{sed}	α_{sed}	c_{hsp}	ρ_{hsp}	α_{hsp}
A	25.0	0.38	-	18.5	1585	1710	1.62	-	1740	1.75	-
B	15.5	0.16	-	30.0	1520	1710	1.50	-	1755	1.80	-
C	28.5	0.04	-	43.0	1534	1655	1.50	-	1685	1.82	-
True A	25.1	0.38	100	22.3	1592.6	1698.7	1.61	0.23	1744.0	1.80	0.23
True B	15.3	0.18	100	36.0	1510.4	1745.4	1.47	0.23	1768.9	1.91	0.23
True C	27.9	0.04	100	46.9	1541.8	1674.2	1.50	0.23	1764.4	1.87	0.23

WA	Z_s	ΔR	D	h_{sed}	$c_{sed}(D)$	$c_{sed}(D+h_{sed})$	ρ_{sed}	α_{sed}	c_{hsp}	ρ_{hsp}	α_{hsp}
A	26	0.20	117	31	1520	1605	1.55	-	1720	1.75	-
B	27	0.01	107	22	1555	1725	1.71	-	1745	1.71	-
C	28	0.24	119	34	1530	1680	1.61	-	1680	1.89	-
True A	26.4	0.22	115.3	27.08	1516.2	1573.2	1.54	0.23	1751,3	1.85	0.23
True B	26.5	0.11	106.7	31.84	1584.3	1722.7	1.80	0.23	1779.0	1.86	0.23
True C	28.2	0.29	119.9	28.95	1565.7	1591.8	1.68	0.23	1707.1	1.88	0.23

The *a posteriori* distributions for these 1/3-octave inversions ranged from well-defined narrow peaks to essentially uniform (with spikes) distributions over the entire search range. It is interesting to note that some of the inversions produced excellent results, while others gave very poor results. It is not readily apparent what distinguishes these cases. Perhaps the difference in behaviour, between similar inversions, occurred because the global minimum does not always lie within the sampled parameter space of the GA populations. If this situation arises early in an inversion then mutation is the only process which can "find" the global minimum, and it is not always successful. It is also interesting to note that several very well-defined distributions were observed that led to incorrect solution states. In these cases, it would appear that there were solutions that were very close in cost value to the true solution and the GA was trapped in the wrong minima.

Some general observations can be made from the GA 1/3-octave results with the TL processor. First, the HLA receivers did better than the VLA receivers. This result was not very surprising since the HLA sees more variability of pressure amplitude along the receiver than the VLA in this situation. Second, the closest arrays (1 & 3) provided better results than the more distant arrays (2 & 4) most of the time; there were some notable exceptions to this. Third, there was not a noticeable change in performance with a change in the number of unknown variables. Searching for more unknowns did not generally result in

poorer quality answers. Finally, source depth, water depth, and range were again the most reliably estimated parameters. Geoacoustic parameters were usefully estimated, but with considerably less success than the localization parameters. Averaging the best values for each parameter for multiple independent inversions (different arrays, different bandwidths, etc.) resulted in reasonable estimates for most of the geoacoustic parameters.

5.   Conclusions

The CAL, SD, SO, AT, and WA test case data sets from the Geoacoustic Inversion Workshop were inverted using a GA optimisation scheme. It was shown that while source localization parameters can be obtained more readily than geoacoustic parameters, it was possible to invert both complex pressure field data and real-valued TL data to obtain estimates of the primary geoacoustic parameters. In the data sets considered, the most important geoacoustic parameters appeared to be water depth, sediment layer thickness, and sediment sound speed. Other geoacoustic parameters may be obtained by inversion with varying degrees of success. The most difficult parameters to obtain with the simulated receivers that were considered in this study appear to be the basement density and the attenuations in both the sediment and basement. In summary, individual inversion attempts may fail to produce useful estimates of geoacoustic parameters; however, repeated efforts with variations in the search space, parameter ranges, cost functions, and sampled data on average yielded good results.

Broadband information was very useful in performing inversions, but in some cases superior results were obtained using limited bandwidths. Exactly why this should be the case is not known, but it may be related to the preferential penetration of lower frequencies into the sub-bottom which then in turn carry most of the sub-bottom information. Useful future work would be to investigate the relative performance of 1/3-octave bands with higher centre frequencies and determine if performance degrades. Additionally, it would be useful to compare the phase differences with frequency in the two propagation models used in this work.

The coherent processor performs best, followed by the standard Bartlett processor. The TL processor discards much of the available information and is the worst performer of the three cost functions examined. However, the calculation of TL is simple and much data of this sort is available. The performance of this processor was sufficient to warrant future attention.

In general, the nearer arrays performed better than more distant arrays. This may be due in part to energy travelling by steep angle paths to the nearer arrays that sample the bottom more completely. The use of HLA receivers also warrants attention as the straightforward deployment of bottomed arrays and the lack of movement in the hydrophones are both desirable features in actual receiver deployments.

References

1   Geoacoustic Inversion Workshop, Benchmark Workshop for Geoacoustic Inversion Methods, 24 - 26 June 1997, Hosted by MacDonald-Dettwiler, 13800 Commerce Parkway, Richmond, BC Canada. Sponsors: Ocean Acoustics and Signal Processing Group, Office of Naval Research. Organizers N. Ross Chapman and A. Tolstoy.

2   N.R. Chapman, A. Tolstoy, and G.A. Brooke, "Workshop97: Benchmarking for geoacoustic inversion in shallow water," Summary report by conference organizers to appear in this volume.

3   David Hannay, "Development of genetic algorithm inversion software, and application to transmission loss data from the Western Bank and Emerald Basin," DREA Contractor Report (under review).

4   D. Beasley, D.R. Bull, and R.R. Martin, "An overview of genetic algorithms Part 1: Fundamentals, and Part 2: Research topics," University Computing, **15**, p. 58–69 and p. 170–181, 1993.

5   W.H. Press et al. "Numerical Recipes in C The Art of Scientific Computing," Cambridge University Press, p. 343–352, 1988.

6   P. Gerstoft, "Global inversion by genetic algorithms for both source position and environmental parameters," J. Comp. Acoust., World Scientific. **2**, 251–266, 1994.

7   Evan K. Westwood, C.T. Tindle, and N.R. Chapman, "A normal mode model for acousto-elastic ocean environments," J. Acoust. Soc. Am., **100**, p. 3631–3645, 1996.

8   D.E. Hannay and N.R. Chapman, "Coherent broadband matched field processing for geoacoustic parameter estimation," *Proc. of the 8th MFP Workshop*, NCCOSC RDT&E Div. (NRaD) TD2932, 12–14 June 1996.