Journal of Computational Acoustics,

BROADBAND GEO–ACOUSTIC INVERSION FROM SPARSE DATA USING GENETIC ALGORITHMS

MARTIN SIDERIUS
Saclant Undersea Research Centre
Viale S. Bartolomeo 400, 19138 La Spezia, Italy

PETER GERSTOFT
Marine Physical Laboratory, Scripps Institute of Oceanography
University of California, San Diego
San Diego, CA 92093-0701

PETER NIELSEN
Saclant Undersea Research Centre
Viale S. Bartolomeo 400, 19138 La Spezia, Italy

Received (to be inserted
Revised by Publisher)

Matched field acoustic inversion techniques were investigated using data from sparsely populated vertical receiving arrays and a single receiver element. The purpose of considering sparse data sets is to investigate, under simulation, the feasibility of reducing the number and complexity of acoustic measurements needed for geo–acoustic inversion. The entire bandwidth of the 1997 Geo–Acoustic Inversion Workshop provided data (25–500 Hz) was used which compensates for lacking spatial information when a limited number of receivers is considered. Forward model PROSIM and inversion code SAGA were applied to benchmark cases: **sd**, **wa**, and **n**. The inversion results generally showed good agreement with the ground truth for full arrays, sparse arrays and single receiver elements. With a simple two layer model, an effective sound speed profile in the bottom was determined which produced a good fit between model and observed pressure fields for the multi-layer case **n**.

1. Introduction

Acoustic signals propagating in shallow water can have many interactions with the ocean bottom resulting in a strong dependency between received signals down range and bottom type. In recent years, matched field processing based inversion methods have been developed which infer properties of the bottom from acoustic measurements[1−7].In matched field inversions, a large number of bottom types are modeled until suitable agreement is found between the observed and modeled fields. Matched field methods generally find an average or effective bottom type which, in the model, will reproduce the observed acoustic field.

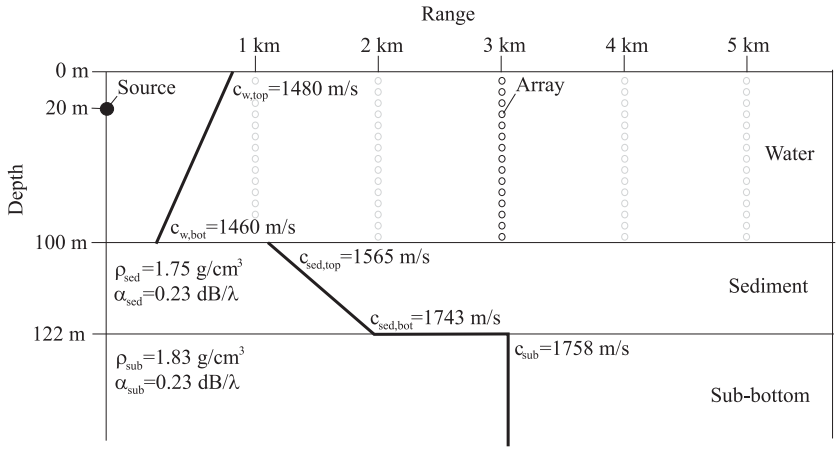

Figure 1: The **sd** and **wa** environment from the Geo-Acoustic Inversion Workshop[12] shown here for the parameters corresponding to ground truth for the **sda** case.

When the acoustic data are sparsely sampled spatially it has been suggested that the information contained in the broadband signal can be used for matched field processing[8–10]. Recently, a manual inversion was performed based on matching time series [11].

The 1997 Geo-Acoustic Inversion Workshop[12] was organized to bring the matched field inversion scientific community together to compare the performance of various algorithms. Acoustic data sets were simulated using the SAFARI code[13] and provided to the participants of the workshop who were not informed of the exact bottom properties used to generate the data. Six environments were used and for each of these, acoustic data were generated for three different realizations of environmental parameters. Limits were given for each of the environmental properties to constrain the search parameters and search space. In all cases considered here the environment consisted of a sediment layer overlying a infinite half-space, as shown in Fig. 1. For the **sd** and **wa** cases the unknown parameters and their bounds are listed in Table 1. In case **n** there was an unknown number of sediment layers and in each layer sound speed and density were unknown. In the inversion a simple two layer bottom model was used in determining an effective environment.

2. Propagation model

An efficient global inversion of environmental parameters is highly dependent on the performance of the applied propagation model. Several thousand forward modeling analyses are often required to obtain a successful inversion and confidence in the results as shown in the test cases treated here. Furthermore, if broad-band signals are considered for inversion purposes the traditional brute-force method of finding the acoustic pressure frequency-by-frequency is far too computationally intensive.

A layered normal-mode model PROSIM based on the newly developed model ORCA[15,16] has been included into the global inversion package SAGA[14]. This normal-mode model is based on dividing the acoustic environment into stratified layers with depth dependent

Table 1: Parameter search bound for the **sd** and **wa** case. Each parameter was discretized into 128 values.

Parameter	Lower bound	Upper bound
Sediment (case **sd** and **wa**)		
Thickness (m)	10	50
Sound speed top (m/s)	1500	1600
Sound speed bottom (m/s)	1550	1750
Density (g/cm^3)	1.4	1.85
Bottom (case **sd** and **wa**)		
Sound speed (m/s)	1600	1800
Density (g/cm^3)	1.6	2.0
Geometric (case **wa**)		
Water depth (m)	100	120
Source depth (m)	10	30
Source range (m)	0	400

properties assuming a $1/c^2$-linear sound speed profile. The acoustic field within each layer is described analytically by using a sum of Airy functions. An eigenvalue equation is established considering continuity of the acoustic field and its derivative across layer interfaces, and by fulfilling the boundary conditions at the sea surface and the bottom of the waveguide. This equation is solved for the eigenvalues of the specific problem. Only the eigenvalues on the real axis are considered here to increase the computation speed, and the attenuation in the sediment layers and sub-bottom is treated by a perturbation theory. The eigenvalues are found by an efficient search algorithm, which only requires about 3.5 iterations to determine each individual mode[17].

The computation speed is shown to be linearly dependent on frequency for single frequency analysis by using the layered model, whereas the computation speed depends quadratically on frequency using the traditional normal-mode models based on numerical integration of mode functions. However, the computation speed increases as the number of layers in the model increases, whereas the speed of the traditional models is almost independent of the depth discretization of the environment. For certain shallow water environments the layered model is shown to be more than 100 times faster at higher frequencies than traditional models[15,17].

In the ORCA broad-band model, calculations are performed by dividing the frequency band of interest into sub-bands of 15 Hz. At the limits of the sub-bands the eigenvalues and mode functions are determined exactly as mentioned above. Furthermore, at the subband limits the first and second order derivative of the wavenumbers with respect to frequency are calculated analytically. The real part of the wavenumbers is interpolated within the sub-bands by applying a 5th-order polynomial[17]. The eigenvalue equation is applied to the interpolated wavenumbers to evaluate the accuracy of the interpolation. If the interpolation of the wavenumbers has been performed successfully the mode functions are interpolated

within the frequency subband using a cubic fit. The coefficients of the cubic polynomial are found by the value of the mode functions and their analytical derivatives with respect to frequency at the subband limits. The attenuation is interpolated linearly within the frequency subband. The above procedure is repeated until the acoustic pressure is established for the entire frequency band of interest, and the technique becomes more efficient at finer frequency sampling of the acoustic pressure.

The real-axis version of ORCA embedded in SAGA has been modified and is called PROSIM, which is a broad-band, range-dependent propagation model based on the adiabatic approximation. The range-dependent acoustic environment is divided into range-independent segments by assuming the eigenvalues and mode functions vary linearly as a function of range within each segment.

The transfer function of shallow-water, range-independent waveguides covering the frequency band from 10 to 10000 Hz involving more than 500 modes and 8000 frequency components can be obtained by using PROSIM in less than 2 minutes on an DEC ALPHA 600-5/266 workstation. This performance is obtained partly by utilizing a discretization of the frequency band of interest into sub-bands of different sizes. The frequency sub-bands are larger at higher frequencies and decrease towards the lower frequencies. For a weak range-dependent environment including 22 modes, 1093 frequency components and 128 range-segments, PROSIM is 35 times faster than the coupled-mode model C-SNAP[18]. Excellent agreement between the two models was found by comparing the received time series for this environment. This indicates that PROSIM is well suited for broadband inversion.

3. Matched field Inversion

3.1. *Objective function*

Two objective functions[14] were used in the analysis of the benchmark problems. In the case of a full array of hydrophones the objective function is the incoherent sum of the Bartlett power in depth for each frequency as given by Eq. (3.1). That is, a correlation is computed in depth between observed and modeled fields and the powers are averaged over frequency to determine the objective function. Both observed and modeled data are complex pressure field vectors referred to here as p and q, respectively. The coherent objective function in depth is given by,

$$\phi_{\mathrm{d}} = 1 - \frac{1}{N_{\mathrm{range}} N_{\mathrm{freq}}} \sum_{k=1}^{N_{\mathrm{range}}} \sum_{i=1}^{N_{\mathrm{freq}}} \frac{|\sum_{j=1}^{N_{\mathrm{depth}}} p_{ijk} q_{ijk}^{*}|^2}{\sum_{j=1}^{N_{\mathrm{depth}}} |p_{ijk}|^2 \sum_{j=1}^{N_{\mathrm{depth}}} |q_{ijk}|^2} \,, \tag{3.1}$$

where $*$ indicates the complex conjugation operation, N_{depth} is the number of receivers in depth, N_{freq} is the number of frequencies considered and N_{range} is the number of receivers in range. From Eq. (3.1), the addition of data in range and frequency are summed in an incoherent fashion (sum of Bartlett powers). When processing coherent in depth several different choices for objective function have been proposed[19].

For sparse arrays, correlation in depth may not be meaningful and the correlation is

performed in the frequency domain which, is followed by a summation of the power in the spatial domain. The objective function is given by,

$$\phi_{\rm f} = 1 - \frac{1}{N_{\rm range} N_{\rm depth}} \sum_{k=1}^{N_{\rm range}} \sum_{j=1}^{N_{\rm depth}} \frac{|\sum_{i=1}^{N_{\rm freq}} p_{ijk} q_{ijk}^*|^2}{\sum_{i=1}^{N_{\rm freq}} |p_{ijk}|^2 \sum_{i=1}^{N_{\rm freq}} |q_{ijk}|^2} \,. \tag{3.2}$$

A classical matched filter is performed by correlating in the time domain and the above expression in frequency domain gives a nearly identical result, as discussed further in the Appendix, where also other choices for objective functions are presented.

3.2. *Genetic algorithms*

The global optimization method genetic algorithms (GA) is used for the optimization. The basic principle of GA is simple: From all possible parameter vectors, an initial population of members is randomly selected. The "fitness" of each member is computed on the basis of the value of the objective function. Based on the fitness of the members a set of "parents" are selected and through a randomization a set of "children" is produced. These children replace the least fit of the original population and the process iterates to develop an overall more fit population. In geo-acoustic inversion, the parameter vectors are the environmental properties such as sediment and sub-bottom sound speed and a population consists of one realization of environmental properties. A more detailed description of genetic algorithms and their application to parameter estimation is given by Gerstoft[5].

3.3. *Convergence*

Three indicators are used to determine the quality of the estimated environmental parameters. Due to the non-uniqueness of the inversion it is not guaranteed that *the* correct solution is found even when all the three criteria are satisfied. The three considered criteria are as follows:

a) Value of objective function. For a good match the objective function should approach a certain value. In particular, for the present inversion, where the parameterization is known *a priori*, it is known that $\phi_{\rm d} \approx 0$ (or $\phi_{\rm f} \approx 0$) for a good match.

b) Plotting of the observed and modeled data with the best match. A visual comparison of the observed and modeled data can often identify problems in the inversion. Often the same data, as used in the inversion, are used when comparing the match; but also data that have not been used in the inversion could be used.

c) A posteriori distributions. The purpose of the inversion is to determine an optimal set of environmental parameters, and thus it is important to have an indication of how well each parameter has been determined. Based on the samples obtained during the inversion, statistics of the convergence for each parameter are computed. Using a Bayesian framework this can be interpreted as a Monte Carlo integration of the likelihood function[20]. However, often the likelihood function is not available, and then a practical weighting of the objective function is performed to give an estimate of the *a posteriori* distributions[5,21]. Due to this *ad hoc* weighting, the *a posteriori* probability should be interpreted with care. The

Table 2: Case **sd**: Results for full arrays at all ranges 1–5 km and for a single hydrophone at range 5 km and depth 50 m.

Parameter	Case	Full Arrays	One Hydrophone	Ground Truth
Sediment thickness (m)	a	23.2	17.9	22.5
	b	41.5	47.8	38.0
	c	31.1	33.6	30.7
Sediment speed top (m/s)	a	1566.1	1566.1	1565.1
	b	1600.0	1600.0	1599.3
	c	1531.5	1531.5	1530.4
Sediment speed bottom (m/s)	a	1748.4	1699.6	1743.1
	b	1669.7	1680.7	1658.5
	c	1605.1	1613.0	1604.2
Sediment density (g/cm^3)	a	1.78	1.77	1.76
	b	1.62	1.65	1.64
	c	1.53	1.53	1.50
Bottom sound speed (m/s)	a	1759.1	1751.2	1757.7
	b	1699.2	1710.2	1707.5
	c	1694.5	1694.5	1689.0
Bottom density (g/cm^3)	a	1.84	1.87	1.83
	b	1.67	1.82	1.87
	c	1.81	1.65	1.70

relative importance of the parameters in the same inversion is precise, but the comparison of inversion results based on different data or approaches should be interpreted with care.

4. Results for the benchmark problems

Inversions were performed on the observed data from cases **sd**, **wa** and **n**. In all cases, a total of 20,000 forward models are computed with PROSIM for the full frequency band from 25–500 Hz in 5 Hz increments. The inversion time was from 40 minutes to 3 hours on a DEC ALPHA 600-5/266. The observed data were provided at ranges of 1, 2, 3, 4 and 5 km and along an array at depths from 1 to 100 m in increments of 1 m. Several combinations of receiver ranges and depths as well as number of receivers used will be described in the following sub-sections.

4.1. *Benchmark case:* **sd**

Case **sd** had 6 unknown parameters: Upper and lower sediment sound speed, sediment thickness, sediment density, bottom sound speed and bottom density. Acoustic data were given for three sets of different environmental parameters, i.e. three realizations of the environment, in this case denoted as **sda**, **sdb** and **sdc**. An inversion was performed on each realization using the data from all observed data points in range and depth. That is, data were used from five vertical arrays with 1 m depth sampling and at ranges from the

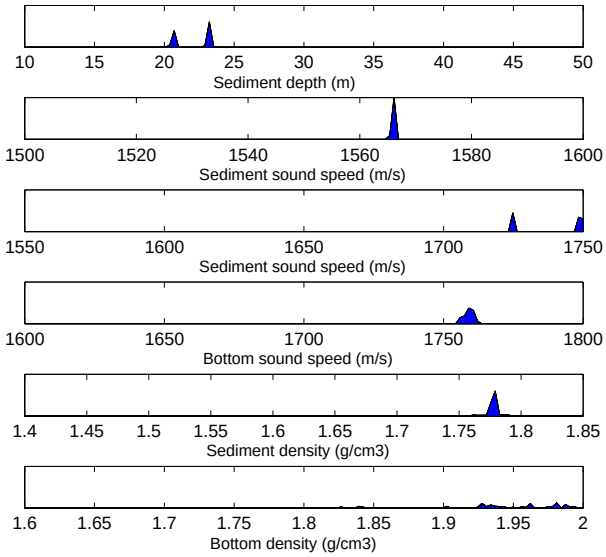

Figure 2: The *a posteriori* distributions for each of the inverted environmental parameters for case **sda**.

source of 1, 2, 3, 4 and 5 km. In a second inversion, only data from a single hydrophone at range 5 km and depth 50 m is included. Table 2 shows the values found by the inversion using data from all the vertical arrays, the single hydrophone results and the ground truth for **sda**, **sdb** and **sdc**.

The **sd** results show good agreement between the inverted environmental parameters and ground truth even for the single hydrophone measurement. In general there was little sensitivity to the density parameter in both the sediment and sub-bottom and the sub-bottom sound speed appeared to be less sensitive than the sound speeds in the sediment. The sensitivity of parameters, or, the degree of confidence given to each parameter found through the inversion process can be summarized using the *a posteriori* distributions as shown for case **sda** in Fig. 2. From the distributions, there are features that appear throughout all test cases including **wa** and **n**. The bottom density is not well determined, and there is a correlation between the sound speed in the sediment at the sediment-bottom interface and the sediment depth (thickness). The 2-D marginal *a posteriori* distributions are given in Fig. 3 as a function of sediment thickness and lower sediment sound speed as produced by SAGA. The correlation is apparent from the figure and this indicates that these two parameters might more efficiently be expressed in terms of a single parameter (possibly searching for a sound speed gradient instead).

4.2. *Benchmark case:* **wa**

Case **wa** has nine unknown parameters. The first six are the same as for **sd** with additional unknowns: Water depth, source range and source depth. Again, there were three realizations of the environment given, **waa**, **wab** and **wac**. Two inversions were made, one using

Table 3: Case wa: Results for full array and single point measurement at depth 50 m.

Parameter	Case	Full Arrays	One Hydrophone	Ground Truth
Water depth (m)	a	115.4	115.0	115.3
	b	107.2	106.9	106.7
	c	120.0	120.0	119.9
Sediment thickness (m)	a	27.6	26.7	27.1
	b	14.7	10.9	31.8
	c	26.4	36.1	29.0
Sediment speed top (m/s)	a	1518.1	1509.4	1516.2
	b	1597.6	1598.4	1584.3
	c	1549.6	1575.6	1565.7
Sediment speed bottom (m/s)	a	1576.8	1562.6	1573.2
	b	1625.6	1642.9	1722.7
	c	1624.0	1587.8	1591.8
Sediment density (g/cm^3)	a	1.50	1.73	1.54
	b	1.77	1.51	1.80
	c	1.78	1.63	1.68
Bottom sound speed (m/s)	a	1749.6	1726.0	1751.3
	b	1699.2	1782.7	1779.0
	c	1647.2	1784.3	1707.1
Bottom density (g/cm^3)	a	1.66	1.96	1.85
	b	1.81	1.73	1.86
	c	1.91	1.81	1.88
Source range (m)	a	5226.8	5220.5	5220.0
	b	5148.0	5103.9	5105.0
	c	5311.8	5289.8	5290.0
Source depth (m)	a	26.2	26.2	26.4
	b	26.7	26.7	28.2
	c	28.4	28.3	28.9

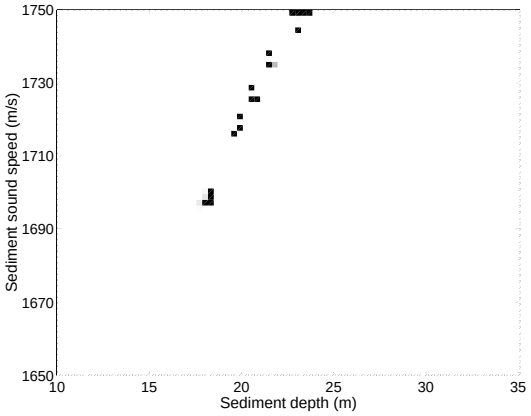

Figure 3: Correlation between the sediment thickness and lower sediment sound speed based on the *a posteriori* distributions for case **sda**. The multiple peaks show the ambiguity in determining each of these parameters.

observed data from the full array and a second using data from a single hydrophone at depth 50 m. The full band of frequencies (FB) from 25-500 Hz in 5 Hz increments was used for both the full array and single hydrophone. Table 3 shows the values found by inversion and the ground truth for the **wa** cases. As for case **sd**, both the full array and the single hydrophone inversion results are very close to the ground truth.

To investigate the relationship between the number of receivers used, inversions were performed on case **waa** using 1, 2, 4, 8, 32, 50 and 100 receivers. In the case of one receiver the location was in the middle of the water column. For more receivers the spacing was equidistant in depth and covering the main part of the water column. A new inversion was performed for each receiver configuration as the direction of the genetic algorithm search will vary in each case. The results are shown in Fig. 4 where the solid line denotes ground truth, the circles denote results obtained using a sparse array and Eq. (3.2) and the "$\times$" denotes the full array using Eq. (3.1). In all cases, the full frequency band was used.

In most cases, Fig. 4 shows the number of receivers used was of little consequence in predicting the environmental parameter value. The inversion results are reasonable even with a single hydrophone.

In addition to the number of receiver elements, the positions in range and depth as well as the optimal frequencies for inversion should be taken into consideration. A full sensitivity study is beyond the scope of this paper but, for comparison, an inversion was performed for case **waa** using observed data from the array closest to the sound source (1 km) using both the full band of frequencies (FB) and using the low frequencies only (LF) 25-200 Hz in 1 Hz increments. In Table 4 the difference is given as a percent of the ground truth value for the 5 km array and 1 km array for both the full band and low frequencies only.

As the present environment is *truly* range independent it is expected that most information, and thus best results, about the bottom parameters is obtained at low frequencies using receivers close to the source, as more energy penetrates into the bottom [22]. The

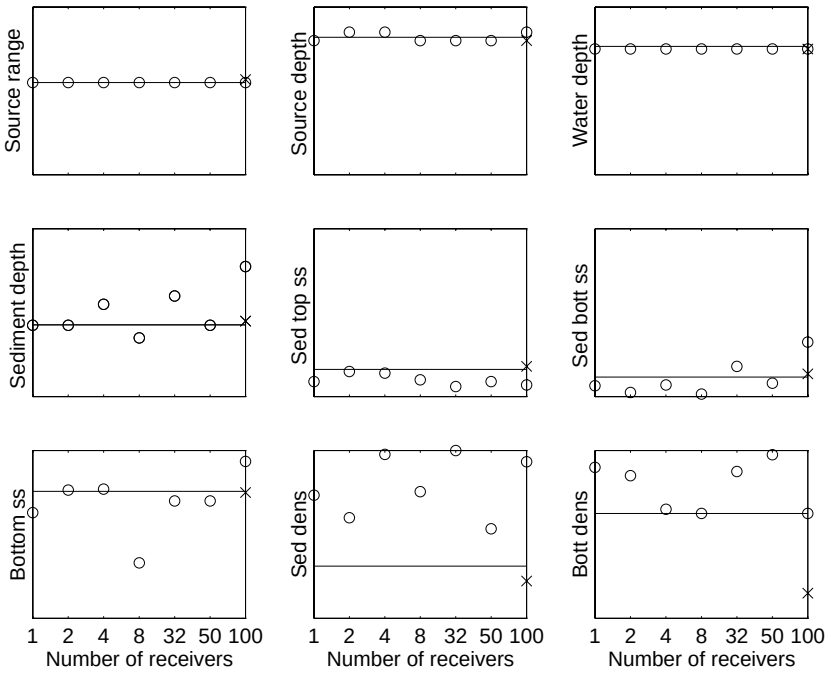

Figure 4: Inversion results for various number of receivers (case **waa**). Solid line denotes ground truth, circles denote sparse array results using Eq. (3.2) and the "×" denotes the full array using Eq. (3.1). The y-axis is scaled to the bounds on the parameters as given in Sect. 1.

Table 4: Inversion Errors (percent) for case **waa**. The percentage is relative to ground truth. For the full array ϕ_{d} is used as objective function, whereas ϕ_{f} is used for the single hydrophone. The boldface numbers indicate the best estimate for a particular parameter.

	full array				one phone (50 m)
Parameter	5 km FB	5 km LF	1 km FB	1 km LF	1 km LF
Water Depth	**0.1**	1.0	**0.1**	0.3	0.2
Sediment thickness	**1.8**	5.2	**1.8**	**1.8**	4.4
Sediment speed top	**0.1**	0.8	0.2	**0.1**	0.7
Sediment speed bot.	0.2	1.1	0.3	**0.1**	**0.1**
Sediment density	2.6	**0.6**	6.4	5.8	**0.6**
Bottom sound speed	**0.1**	0.8	0.2	0.3	0.4
Bottom density	10.3	**1.6**	7.0	2.7	4.8
Source range (offset)	3.1	28.9	**0.2**	4.5	**0.2**
Source depth	0.8	**0**	0.4	**0**	1.0
Mean error	2.1	4.4	1.9	1.7	1.4

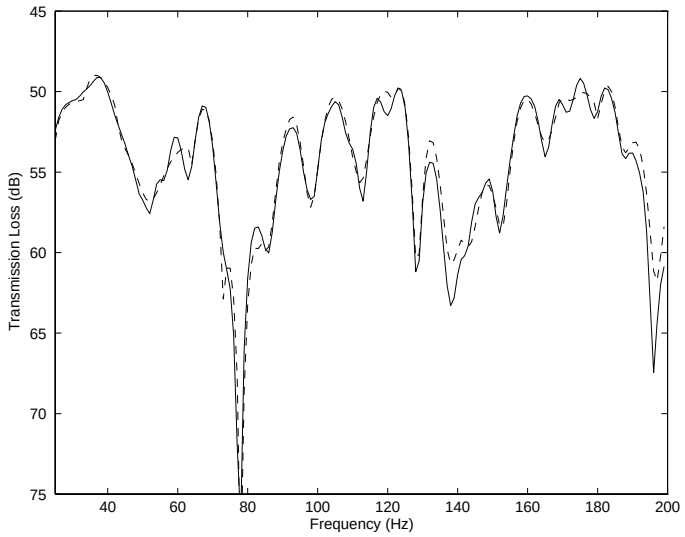

Figure 5: Transmission loss for the **waa**-LF case for one hydrophone at 1 km. The full line is the observed data and dotted line the modeled data with the best environment. The agreement is very good.

results in Table 4 imply that for both objective functions using only the low frequencies and array position at 1 km gives the best results in terms of percentage error. Individual parameters may be estimated better using other configurations of array position and frequency content. This can be observed by noting the boldface numbers in the table indicating the best estimate for a particular parameter. In general a thorough analysis would be required to determine optimal frequencies and array position for inversion.

The best obtained fit for the **waa**-LF case for one hydrophone at 1 km is shown both in frequency (Fig. 5) and time domains (Fig. 6). The transmission loss is produced directly from SAGA, whereas the time series first requires a Fourier Transform of the data. The dissimilarities between observed (benchmark) and modeled data is due to not finding the exact environment and to differences between forward models PROSIM and SAFARI. Based on simulations, it is expected that a significant part of the deviation in Fig. 5 is due to only considering the real axis eigenvalues. It should be noted that in the matched field approach used here, only the relative phase across all frequencies is determined. To compare the time series, the average phase difference between the observed and computed data was determined. The time series have been convolved with a Ricker wavelet (second derivative of the Gaussian function) with a center frequency of 80 Hz.

The two objective functions ϕ_{d} given by Eq. (3.1) and ϕ_{f} given by Eq. (3.2) produce slightly different results. Again, considering case **waa**, the error in the inversion results using ϕ_{d} and ϕ_{f} is compared. Table 5 shows the percent error for each parameter and the mean error over all parameters for full arrays at 5 km and using the full band of frequencies. In general, ϕ_{d} performs better than ϕ_{f} as it produces less error when all parameters are

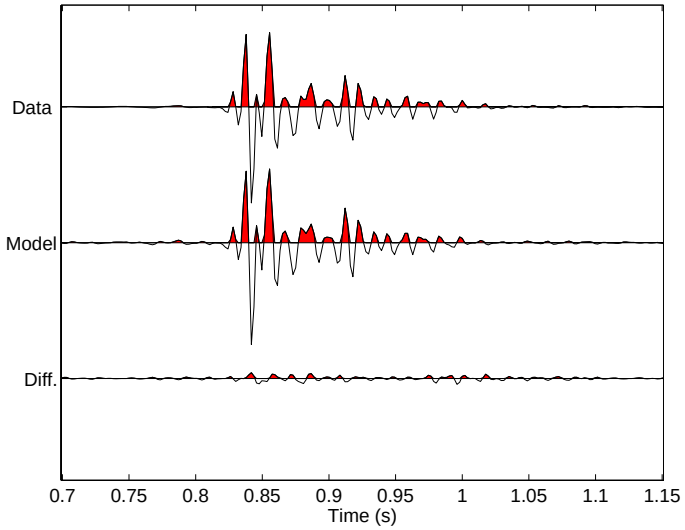

Figure 6: Time series for the **waa**-LF case for one hydrophone at 1 km. The observed data, the modeled data and the difference between the two time series are shown.

Table 5: Inversion Errors (percent) **waa**

Parameter	ϕ_{d}	ϕ_{f}
Water Depth	0.1	0.3
Sediment thickness	1.8	49.8
Sediment speed top	0.1	0.6
Sediment speed bottom	0.2	2.6
Sediment density	2.6	18.2
Bottom sound speed	0.1	2.1
Bottom density	10.3	0.0
Source range	3.1	0.2
Source depth	0.8	0.4
Mean error	2.1	8.2

considered. However, ϕ_{f} outperforms ϕ_{d} in estimating the source position and the bottom density. The improved prediction of the bottom density using ϕ_{f} is most likely a chance result in that the confidence level in the density predictions were low as was indicated in the *a posteriori* distributions for case **sda** in Fig. 2. These results comparing ϕ_{d} with ϕ_{f} can also be seen from Fig. 4.

4.3. *Benchmark case:* **n**

The final case considered was **n** which was the same as case **sd** except rather than having two layers, **n** has an unknown number of bottom layers. A simple approach was used and the multi-layer problem was solved as if only two layers exist. The top layer in the two layer model is allowed to have a sound speed gradient and a representative sound speed profile was determined. The sound speed profile match is shown for the three realizations of test case **n** in Fig. 7 for both the full array and the single receiver. It is seen that although the simplified model does not capture the complex shape of the sound speed profile, it fits the curve well in a least squared sense. The single receiver performed well, giving a reasonable fit to the sound speed profile. The simple two layer model predicts the sediment sound speed well at the top, and seems to approximately determine the sound speed gradient. The inverted bottom sound speed value appears to be an average value across all the layers.

5. Discussion and conclusions

For the benchmark cases considered here, the inversion results agree well with ground-truth. The computational efficiency of PROSIM, the broadband, normal mode forward model made it possible to consider the entire frequency content of the observed data between 25 and 500 Hz. The genetic algorithms package SAGA produced good inversion results from 20,000 forward models for each set of observed data.

Two objective functions have been used, either a coherent summation in depth or a coherent summation in frequency. In the Appendix it is discussed how the latter could equally well be performed in the time domain. The coherent summation in frequency has the advantage that only one hydrophone need be used. In practice a few phones are probably required.

The single hydrophone inversions produced results that agreed well with ground truth. The results in both Fig. 4 and Table 4 indicate that for the benchmark cases, one receiver is sufficient if a coherent processor in frequency is used. The broadband single hydrophone inversion results demonstrate that a few receivers may be sufficient to determine ocean bottom properties.

The multi-layer case **n** provided simulated acoustic data in a more realistic environment. In the inversion a simple two layer model provided a good estimate of the bottom parameters for both the full array and the single hydrophone. This approach is practical since it is impossible to solve for all the acoustic properties that make up the complex structure of the ocean bottom.

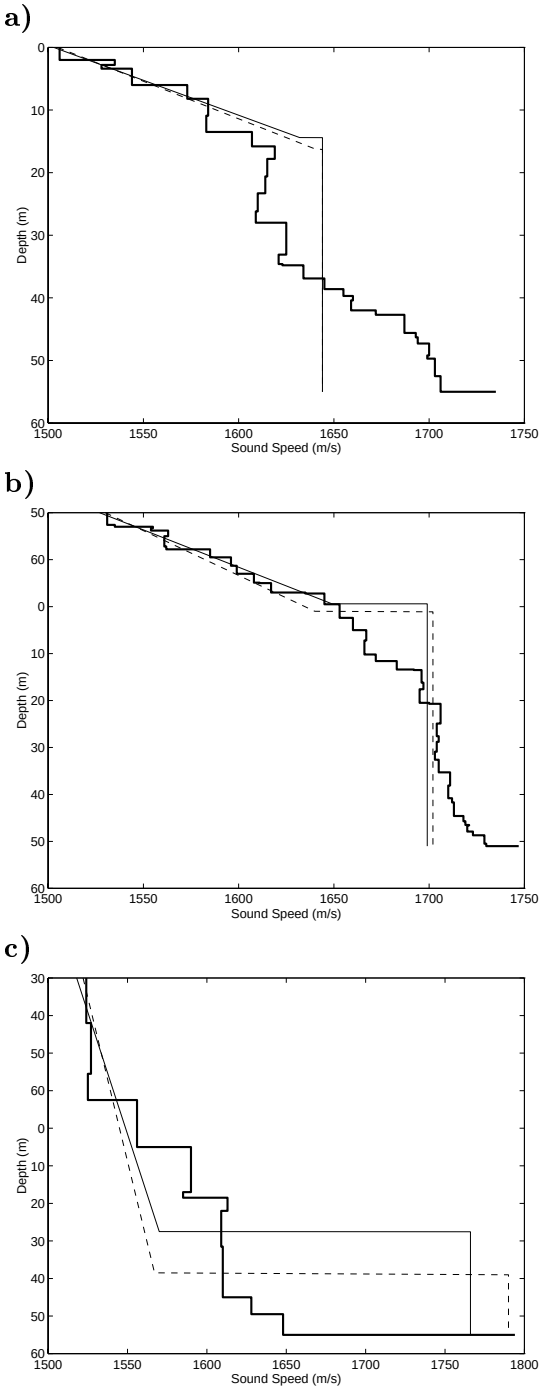

Figure 7: Sound speeds for ground truth (heavy solid line), full array (thin solid line) and single receiver (dashed line) for the multi-layer case **n**. Depth zero is the water-sediment interface. a) case **na** b) case **nb** c) case **nc**.

Acknowledgments

The development of the PROSIM normal mode code was partly funded by the EC under the MAST-III project PROSIM, contract number MAS3-CT95-0008. Christoph Mecklenbräuker has supplied invaluable comments to the Appendix.

Appendix A

Objective functions in time and frequency domain

In this Appendix it is shown that an objective function based on simple correlation in the time domain can be seen as similar to the coherent summation over frequency. For simplicity only a single observed $p(t)$ and modeled $q(t, \mathbf{m})$ time series as functions of time t are used in the beginning. The unknown environmental parameters are included in the model vector $\mathbf{m}$. In the time domain the correlation objective function to be optimized is given by

$$\Phi(\mathbf{m}, \tau) = [\int_{-\infty}^{\infty} p(t+\tau) q(t, \mathbf{m}) \mathrm{d}t]^2 \ . \tag{A.1}$$

where τ is the time delay between the two signals. In order to obtain a perfect match from Eq. (A.1) the absolute timing of the synthetic and measured waveforms must be identical. This requires that both the source and the receiver are monitored. In general the absolute timing of the transmitted source signal and the received signal at the array is not available and thus the absolute arrival time of the two signals is unknown. One way of treating this problem is by searching for a unknown time delay τ. Clearly, for cases where this time delay is known it should be set to zero.

Using the definitions of Fourier transform the following expressions are obtained

$$\begin{aligned} \Phi(\mathbf{m}, \tau) &= \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \left(p(\omega) e^{i\omega\tau} e^{i\omega t} \right) \mathrm{d}\omega q(t, \mathbf{m}) \right\} \mathrm{d}t \right]^2 \\ &= \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ p(\omega) e^{i\omega\tau} \int_{-\infty}^{\infty} \left(q(t, \mathbf{m}) e^{i\omega t} \right) \mathrm{d}t \right\} \mathrm{d}\omega \right]^2 \\ &= \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} p(\omega) q^*(\omega, \mathbf{m}) e^{i\omega\tau} \mathrm{d}\omega \right]^2 \ , \end{aligned} \tag{A.2}$$

where * is complex conjugate. In order to only optimize for the shape of the signal we normalize Eqs. (A.1–A.2) with the power of the two signals $\int_{-\infty}^{\infty} p^2(t) \mathrm{d}t$ and $\int_{-\infty}^{\infty} q^2(t, \mathbf{m}) \mathrm{d}t$. By use of Parsevals theorem this can be done equivalently in the time or frequency domain. In the following the factor $1/2\pi$ will be neglected, as it is just a constant which is unimportant for the optimization. Also the objective function Φ is modified to a function ϕ so that it becomes a minimization problem with an absolute minimum of zero and the maximum is arbitrary.

For data on a vertical array $p_j, j = 1 \ldots N_{\mathrm{depth}}$ (N_{depth} is the number of hydrophones in depth) there are two ways to do the summation: either

$$\phi_1^{\mathrm{t}}(\mathbf{m}, \tau) = 1 - \frac{1}{N_{\mathrm{depth}}} \sum_{j=1}^{N_{\mathrm{depth}}} \frac{\left[\int_{-\infty}^{\infty} p_j(t+\tau) q_j(t, \mathbf{m}) \mathrm{d}t\right]^2}{\int_{-\infty}^{\infty} p_j^2(t) \mathrm{d}t \ \int_{-\infty}^{\infty} q_j^2(t, \mathbf{m}) \mathrm{d}t} \tag{A.3}$$

or

$$\phi_2^{\mathrm{t}}(\mathbf{m},\tau) = \sum_{j=1}^{N_{\mathrm{depth}}} \int_{-\infty}^{\infty} p_j^2(t)\mathrm{d}t - \frac{\left[\sum_{j=1}^{N_{\mathrm{depth}}} \int_{-\infty}^{\infty} p_j(t+\tau) q_j(t,\mathbf{m})\mathrm{d}t\right]^2}{\sum_{j=1}^{N_{\mathrm{depth}}} \int_{-\infty}^{\infty} q_j^2(t,\mathbf{m})\mathrm{d}t} \ . \tag{A.4}$$

In the above equations the same time delay is common to all hydrophones. There are two differences between ϕ_1 and ϕ_2: The main difference is that the numerator of ϕ_1 is squared before the summation, whereas in the numerator of ϕ_2 the summation is done first. Thus in ϕ_2 there are cross terms between the time series, whereas ϕ_1 does not contain cross terms. Also, ϕ_1 gives equal emphasis to the received signal on each phone, whereas ϕ_2 weights each phone according to its energy. If some of the hydrophones have little signal, either it could be malfunctioning or it could be placed at a null, then ϕ_2 is preferred. However, if the phones have different unknown gain then ϕ_1 should be used.

In the frequency domain the above equations become

$$\phi_1^{\mathrm{f}}(\mathbf{m},\tau) = 1 - \frac{1}{N_{\mathrm{depth}}} \sum_{j=1}^{N_{\mathrm{depth}}} \frac{\left| \int_{-\infty}^{\infty} p_j(\omega) q_j^*(\omega,\mathbf{m}) e^{i\omega\tau}\mathrm{d}\omega \right|^2}{\int_{-\infty}^{\infty} |p_j(\omega)|^2 \mathrm{d}\omega \ \int_{-\infty}^{\infty} |q_j(\omega,\mathbf{m})|^2 \mathrm{d}\omega} \tag{A.5}$$

and

$$\phi_2^{\mathrm{f}}(\mathbf{m},\tau) = \sum_{j=1}^{N_{\mathrm{depth}}} \int_{-\infty}^{\infty} |p_j(\omega)|^2 \mathrm{d}\omega - \frac{\left| \sum_{j=1}^{N_{\mathrm{depth}}} \int_{-\infty}^{\infty} p_j(\omega) q_j^*(\omega,\mathbf{m}) e^{i\omega\tau}\mathrm{d}\omega \right|^2}{\sum_{j=1}^{N_{\mathrm{depth}}} \int_{-\infty}^{\infty} |q_j(\omega,\mathbf{m})|^2 \mathrm{d}\omega} \ . \tag{A.6}$$

For discrete frequencies it is seen that Eq. (A.5) corresponds to Eq. (3.2).

For real data the objective functions are best computed in terms of the ensemble covariance matrix across both frequency and depth. This is computed by first ordering all the observations across N frequencies and N_{depth} depths into one vector:

$$\mathbf{g} = [p_1(\omega_1), \ldots, p_1(\omega_N), p_2(\omega_1), \ldots, p_{N_{\mathrm{depth}}}(\omega_N)]^T. \tag{A.7}$$

Based on N_{obs} observations of this vector the ensemble covariance matrix is formed,

$$\mathbf{C} = \frac{1}{N_{\mathrm{obs}}} \sum_{n}^{N_{\mathrm{obs}}} \mathbf{g}_n \mathbf{g}_n^{\dagger}, \tag{A.8}$$

where † is the complex transpose. This covariance matrix contains elements corresponding to all frequencies and phones. It might thus be quite large and it might be more practical to just use the eigenvector corresponding to the largest eigenvalue of the covariance matrix and then base the processing on Eq. (A.6).

Similar to Eq. (A.7), the synthetic data is ordered into one vector given by

$$\mathbf{b} = [q_1(\omega_1)e^{i\omega_1\tau}, \ldots, q_1(\omega_N)e^{i\omega_N\tau}, q_2(\omega_1)e^{i\omega_1\tau}, \ldots, q_{N_{\mathrm{depth}}}(\omega_N)e^{i\omega_N\tau}]^T, \tag{A.9}$$

where also the phase shift due to the unknown arrival time is incorporated.

An objective function corresponding to Eq. (A.6) now becomes

$$\phi_3(\mathbf{m}, \tau) = \mathrm{tr}\mathbf{C} - \frac{\mathbf{b}^\dagger \mathbf{C}\mathbf{b}}{\mathbf{b}^\dagger \mathbf{b}} \ , \tag{A.10}$$

where tr is the trace of the covariance matrix $\mathbf{C}$.

For synthetic data the time delay is known and thus it is not necessary to search for the time delay τ, but for real data without precise synchronization of source and receiver it is treated as a parameter to be optimized. For noise free data it is not necessary to form a covariance matrix, and thus Eq. (A.5) or Eq. (A.6) can be used directly.

References

1. M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," J. Acoust. Soc. Am. **90** (1991) 1910–1922.
2. M. D. Collins, W. A. Kuperman and H. Schmidt, "Nonlinear inversion for ocean bottom properties," J. Acoust. Soc. Am. **92** (1992) 2770–2783.
3. C. E. Lindsay and N. R. Chapman, "Matched field inversion for geophysical parameters using adaptive simulated annealing," IEEE J. Oceanic Eng. **18** (1993) 224–231.
4. S. E. Dosso, M. L. Yeremy, J. M. Ozard and N. R. Chapman, "Estimation of ocean bottom properties by matched-field inversion of acoustic field data," IEEE J. Oceanic Eng. **18** (1993) 232–239.
5. P. Gerstoft, "Inversion of seismo-acoustic data using genetic algorithms and *a posteriori* probability distributions," J. Acoust. Soc. Am. **95** (1994) 770–782.
6. A. Tolstoy, *Matched Field Processing for Underwater Acoustics* World Scientific Pub. Co. River Edge, New Jersey 1993
7. J. P. Hermand and P. Gerstoft, "Inversion of broadband multi-tone acoustic data from the Yellow Shark summer experiments," IEEE J. Oceanic Eng. **21** (1996) 324-346.
8. N. L. Frazer and P. I. Pecholcs, "Single-Hydrophone localization," J. Acoust. Soc. Am. **88** (1990) 995–1002.
9. J. P. Hermand and R. J. Soukup, "Broadband inversion experiment Yellow Shark '95: Modelling the transfer function of a shallow-water environment with range-dependent soft clay bottom," *Third European conference on underwater acoustics*, edited by J.S. Papadakis, 887–892 (Crete University Press, Crete, 1996).
10. Z-H. Michalopoulou and M.B. Porter and J.P. Ianniello, "Broadband source localization in the Gulf of Mexico," J. of Computational Acoustics **4** (1996) 361–370.
11. D.P. Knobles and R.A. Koch, "A time series analysis of sound propagation in a strongly multipath shallow water environment with an adiabatic normal mode approach," IEEE J. Oceanic Eng. **21** (1996) 1–13.
12. A. Tolstoy and R. Chapman, "The matched-field processing benchmark problems," Journal of Computational Acoustics, this issue (1998).
13. H. Schmidt, "SAFARI: Seismo-acoustic fast field algorithm for range independent environments. User's guide," SACLANT Undersea Research Centre, SR-113, La Spezia, Italy, (1987).
14. P. Gerstoft,"SAGA Users guide 2.0, an inversion software package," SACLANT Undersea Research Centre, SM–333, La Spezia, Italy, (1997).
15. S.J. Levinson, E.K. Evans, R.A. Koch, S.K. Mitchell, and C.V. Sheppard, "An efficient and robust method for underwater acoustic normal-mode computation," J. Acoust. Soc. Am. **97** (1995) 1576-1585.

16. E. K. Westwood, C. T. Tindle, and N. R. Chapman, "A normal mode model for acousto-elastic ocean environments," J. Acoust. Soc. Am. **100** (1996) 3631–3645.
17. E.K. Westwood, "Improved narrow band and broadband normal mode algorithms for fluid ocean environments," Presentation at the 131st Meeting of the Acoustical Society of America, Paper 3aUW7, May 1996.
18. C. M. Ferla, M.B. Porter and F.B. Jensen, "C-SNAP: Coupled SACLANTCEN normal mode propagation loss model", SACLANTCEN Undersea Research Centre, SM-274, La Spezia, Italy (1993).
19. G. Haralabus and P. Gerstoft, "Stability of parameter estimates using multi-frequency inversion techniques," SACLANT Undersea Research Centre, SM–333, La Spezia, Italy, (1996).
20. P. Gerstoft and C.F. Mecklenbräuker, "Ocean acoustic inversion with estimation of a posteriori probability distributions", submitted to J. Acoust. Soc. Am. (May 1997).
21. P. Gerstoft, "Global Inversion by genetic algorithms for both source position and environmental parameters," J. of Computational Acoustics **2** (1994) 251–266.
22. P. Ratilal, P. Gerstoft, J.T. Goh "Subspace approach to inversion by genetic algorithms involving multiple frequencies," J. of Computational Acoustics. This issue (1998).