Proceedings of the 2nd EAA International Symposium on Hydroacoustics
24-27 May 1999, Gdańsk-Jurata
POLAND

Sonar Modelling

M. João Rendas(1), *Georges Bienvenu*(2)

(1) Laboratoire I3S, CNRS - Univ. Nice Sophia Antipolis, France
rendas@i3s.unice.fr
(2) Thomson Marconi Sonar, BP. 157, 06903 Sophia Antipolis Cedex, France.
bienvenu@tms.thomson-csf.fr

Extension of the operational range of underwater location techniques has required incorporation of reliable acoustic propagation models, that closely predict oceans' behavior. While opening new possibilities, such as the use of focusing techniques in parallel with existing localization methods, this evolution also resulted in increased sensitivity to the accuracy of the models used. We show how a performance analysis tool based on information theoretic measures of model resemblance can be used in this context in two distinct ways. The first enables the analysis of the sensitivity to the propagation model, and thus identifying the most critical model parameters, for which accurate estimates must be provided. In the second, we discuss how a priori *knowledge of the model structure can be exploited to improve performance of focusing/localization methods.*

1. Introduction

Source localization uses parametric models of the underwater propagation to invert the acoustic field received from a distant source. The correctness of these models is of paramount importance, these systems being particularly sensitive to errors on the parameters related to the description of the sound speed profile.

We introduced [10] a generic tool for global performance analysis based on the Kullback-Leibler divergence, which captures the observability characteristics of parametric estimation problems. Let $\theta \in \Theta$ be the vector of parameters that we want to estimate, taking values in some open subset $\Theta \subset \mathbf{R}^n$. The ambiguity function defined in [10] is a map

$$\mathcal{A}(\theta : \theta_0) : \Theta \times \Theta \rightarrow [0, 1]$$

which describes for all possibles pairs $(\theta, \theta_0) \in \Theta \times \Theta$ the ability of differentiating between the two values of the parameters. More precisely, we know that the probability of error in the associated binary test (that tests the hypothesis θ against θ_0) tends to zero exponentially with a multiple of $\mathcal{A}(\theta : \theta_0) - 1$, when the true value of the parameter is θ_0.

The ambiguity function takes values close to zero when no confusion between the two parameter values exists, while values close to one flag potentially ambiguous pairs of parameter values. It can be shown that the classical Woodward ambiguity is a particular case of this function, when the radar problem is considered [10]. The definition of this ambiguity function is grounded on the geometry of the statistical model on which parameter estimation is based, and $\mathcal{A}(\theta : \theta_0)$ is in fact a (directed) "distance" between the points in the statistical manifold corresponding to the pair (θ, θ_0).

In this paper we discuss two practical uses of this generic tool in the context of underwater localization problems. First, $\mathcal{A}(\theta, \theta_0)$ can be generalized to define a sensitivity measure evaluating the increase in dispersion and eventual introduction of biases that can be expected due to use of wrong *a priori* physical models:

$$\mathcal{S}(\theta : \theta_a | \hat{\gamma}, \gamma_a).$$

This function is by definition the ambiguity between the true source location θ_a and a distinct location θ, when the propagation model is tuned for parameters $\hat{\gamma}$, while the true ocean parameters are γ_a.

In focusing techniques, one estimates the physical parameters of the propagation model along with the source localization. This leads to multivariate nonlinear optimization problems, were the maximum of a relevant score function (dependent on the whole set of parameter values being estimated) is searched. Often, the use of focusing techniques lead to ambiguous situations, were distinct combinations of the source and ocean parameters produce similar observed signals. As a second application of our tool, and noting that $\mathcal{A}(\theta : \theta_a)$ is also the mean (expected) value of the observed score, we discuss how this information can be used to deal with secondary peaks of the likelihood function.

The paper is organized as follows. In Section 2 the sensitivity measure is reviewed, the problem of characterizing the errors in the ocean parameters is addressed, and a simple example showing the result of applying this analysis to a deep water scenario for a tomography problem is presented. In Section 3 we discuss the utilization of the ambiguity function to improve the performance of parametric estimation problems under highly ambiguous conditions. We discuss the use of Genetic Algorithms in this context, and propose a modification of the usual generalized maximum likelihood procedure that is shown, in a simple example, to yield better performance.

2. Sensitivity Measure

Statistically motivated parametric estimation methods are based on knowledge of a family of conditional pdf's, that describes the dependency of the observed data (r) distribution on the parameter of interest (γ): $\mathcal{G}_\gamma = \{p(r|\gamma), \gamma \in \Gamma\}$.

When $\mathcal{G}_\gamma$ perfectly reflects the behavior of the physical environment, the ambiguity function defined in [10] can be used to predict the global performance of the parametric estimation mechanism, identifying possible large errors or configurations of poor observability of the parameter γ. Motivated by the close relationship between Maximum Likelihood estimators and the Kullback-Leibler directed divergence between probability density functions (pdf's) for exponential distributions, that function defines ambiguity between two values $\gamma_0, \gamma \in \Gamma$ as a normalized version of the Kullback-Leibler distance between the corresponding elements of $\mathcal{G}_\gamma$.

We are now interested in quantifying the impact of imperfect world knowledge, i.e., the fact that the true world's behavior is described by a given family of distributions $\mathcal{G}^0_\gamma = \{p^0(r|\gamma), \gamma \in \Gamma\}$, while a distinct model is used in the fitting operation done at the receiver site: $\mathcal{G}_\gamma \neq \mathcal{G}^0_\gamma$. The observed data r is governed by a single member of $\mathcal{G}^0_\gamma$, that we denote by $p^0 = p^0(r|\gamma_0)$. The ability to correctly predict γ_0 using r under model mismatch can be analyzed studying a normalized version of the Kullback-Leibler distance between p^0 and the elements of $\mathcal{G}_\gamma$ [4] :

$$\mathcal{S}(\gamma) \triangleq 1 - \frac{I(p^0 : p(r|\gamma))}{I_{sup}(p^0)} \qquad (1)$$

In this equation $I(p : q) = \mathrm{E}_p\{\ln(p/q)\}$ is the Kullback-Leibler directed divergence between pdf's p and q, where $\mathrm{E}_p\{\cdot\}$ is expectation with respect to the pdf p, and $I_{sup}(p^0)$ is an upper bound on $I(p^0 : p(r|\gamma))$. This sensitivity index is proposed in [10], where its relation to optimal estimation procedures is discussed.

The normalization in (1) is suitable to analyze the performance of a single method. However, when comparing several methods, it is more convenient to work directly with the Kullback directed divergence $I(p^0 : p(r|\gamma)$. In this way, we can not only compare the global performance, but also local performance, making use of the relation between the gradient of the Kullback divergence and the Fisher information matrix (see [4]). Ideally, $I(p^0 : p(r|\gamma))$ is zero for $\gamma = \gamma_0$, and has large values for all $\gamma \neq \gamma_0$. Modeling errors induce systematic biases in the estimation procedure, that are flagged by the fact that $I(p^0 : p(r|\gamma)$ has its minimum at an erroneous value $\gamma^* \neq \gamma_0$.

2.1. Important Biases

Instead of trying to describe exactly the evolution of the information about the environmental parameters as observation interval increases, we alternatively evaluate the impact of those errors that are expected to occur more often. The sensitivity measure eq. (1) yields exactly this information. To analyze coherent receivers, as the ones assumed in this paper, requires the computation of the Kullback directed divergence between the data records themselves. Invoking a large observation interval assumption, we use the asymptotic characterization of the directed divergence between stationary Gaussian processes derived in [8], that expresses $I(p^0 : p(r|\gamma))$ in terms of the power spectral densities of the observations:

$$\begin{aligned}\overline{I}(p^0 : p(r|\gamma)) &= \lim_{T\to\infty} \frac{1}{T} I(p^0 : p(r|\gamma)) \qquad (2)\\ &= \frac{1}{2}\int \left[\mathrm{tr}\left[S^0(\lambda)S_\gamma(\lambda)^{-1}\right] - K - \ln\frac{|S^0(\lambda)|}{|S_\gamma(\lambda)|}\right] d\lambda.\end{aligned}$$

In the previous equation, $S_\gamma(\lambda)$ denotes the spectral density power matrix of the observed vector under $p(r|\gamma)$, and $S^0(\lambda)$ its value under p^0.

2.2. Example

We present below a simple example of application of function to a tomography problem. The model used at receiver assumes that the ocean is horizontally stratified, with two distinct homogeneous layers. In the superficial layer (up to depth $y_{duct} = 914$m), the sound speed decreases linearly with depth (with rate $g_0 =$

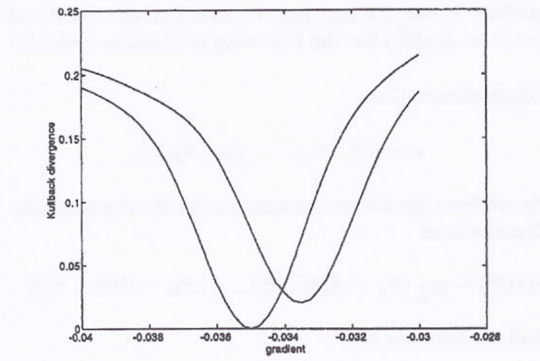


Figure 1: Kullback divergence, gradient estimation with wrong array depth.

$-.035ms^{-1}$), increasing linearly from y_{duct} until the ocean bottom. The ocean boundaries are perfectly flat, with reflection coefficients depending on the grazing angle. The duct and ocean depths (equal to 914 m and 4 Km, respectively) and the sound speed gradient in the deep layer ($g_1 = .013ms^{-1}$) are assumed to be perfectly known at the receiver, as well as all other parameters relating to receiver array (geometry, localization, gain). For all surfaces shown, the distance between the tomography source and the receiving array is 6 Km, and source and receiver immersion are $200m$. The array is vertical, linear, uniform, with $K = 30$ sensors, and sensor spacing is half wavelength at the higher frequency of analysis. The source signal spectrum is flat in the band $[3.5, 4.5]$ KHz.

Figure 1 shows the ambiguity function (Kullback divergence) for estimation of the velocity gradient in the first layer, under a small error on the value of the antenna immersion (receiver is using, as a perfectly known value, a depth of 190 m, when antenna is actually placed at 200 m). We can see that this small mismatch in antenna depth results in a small bias in the estimation of g_0: $\hat{g}_0 \simeq -.0335s^{-1}$ instead of the true value of $g_0 = -.035$. The dashed curve in Figure 1 illustrates the performance of this tomography experiment under no mismatches, i.e., for a correct value of antenna immersion. We can see that aside the introduction of a small bias, the variance of the estimate (related to the second derivative of this curve) is not significantly affected.

3. Using global strcture

As we said previously, use of focusing techniques, by increasing the dimensionality of the vector of parameters being estimated (the number of degrees of freedom of the model that is being fitted to the data) can create potentially ambiguous situations, the score surface being optimized exhibiting several secondary maxima. We discuss now how the global performance analysis tool presented before can be used in this situation. The key observation is that, for inverse problems in non-homogeneous mediums such as the ocean, the shape of the function being optimized depends on the true value of the unknown parameters being sought. The global shape of the ambiguity surface provides, in this manner, information about the true parameter values.

According to this observation, we should be able to "sample" the entire score surface, i.e., to get information concerning its topology, to be able to compare its geometry with the one predicted by $\mathcal{A}(\theta : \theta_0)$. Most numerical optimization techniques use only, at each iteration, the value of the score function at a *single point*, and thus cannot provide information about its global structure. However, a recent stochastic optimization tool, genetic algorithms, provides the basic mechanisms for collectively evolving a set of points in the search space for general non-linear and multivariate optimization problems [3]. They have been recently used for underwater acoustics applications, showing their interest in connection with matched field techniques [2, 5].

For parameter estimation problems, the availability, at each iteration of a genetic algorithm, of a collection of samples of the ambiguity surface of the problem, enables the determination of the correlation between the *observed* ambiguity surface (at the sampled points) and the *predicted* ambiguity surface. The consideration of this information allows early detection of secondary extrema (which yield an ambiguity surface which does not correlate well with the observed one) and thus contributes to speed the convergence of the algorithm to the global optimal values. The global shape of the ambiguity surface provides, in this manner, information about the true parameter values, and can thus be used to tailor the genetic operators and fitness criteria. Two major modifications are proposed:

- *in the evaluation/selection operator:* while canonical genetic algorithms for parameter optimization individually evaluate each element of the population, we present a novel evaluation procedure, which uses the values of the objective function on the whole population for evaluating each individual. Our method provides better rejection of secondary maxima, leading in this way to an increased efficiency of GA's;

- *in the reproduction operators:* while these are homogeneous in canonical GA's, i.e., independent of the sampled values of the objective function, we propose to adjust them using the information provided by the sampled points and the *a priori* model. In this way, the generation of new individuals is directed to regions that have higher probability of corresponding to the global optimum.

3.1. Collective evaluation

Let θ denote the complete set of unknown parameters in the source localization problem (which may include, for instance, parameters describing the radiated signal spectrum or environmental parameters for source focalization [1]). Let $\mathcal{A}(\theta : \theta_0)$ denote the problem ambiguity function, describing the resemblance of the probabilistic models corresponding to different parameter values.

Assume that at iteration k the population of the GA is the set of N_k points

$$P_k = \{\theta_1^k, \ldots, \theta_N^k\},$$

and let the function being optimized be $\mathcal{F}(\theta)$. Generation of the next population is based on the evaluation of $\mathcal{F}$ at all individuals of P_k:

$$\mathcal{F}^k = \{\mathcal{F}(\theta_1^k), \ldots, \mathcal{F}(\theta_N^k)\}.$$

Present GA algorithms determine the fitness of each individual θ_i^k, Φ_i^k, as a scaled positive version of $\mathcal{F}(\theta_i^k)$, $\Phi_i^k = \Phi(\mathcal{F}(\theta_i^k))$, and probabilistically select them for "reproducing" in the next generation, using their relative fitness:

$$p_s(\theta_i^k) = \frac{\Phi_i^k}{\sum_{j=1}^N \Phi_j^k}.$$

Assume now that the objective function is multi-modal, with important secondary lobes. The individuals falling in the secondary lobes of the objective function will be, with high probability (proportional to the importance of the secondary lobes), selected along with those falling in the main lobe of the ambiguity function. If a significant percentage of the population happens to fall in secondary lobes, or if their fitness is higher than those of the individuals inside the main lobe, selection of the next generation on the basis only of $\mathcal{F}^k$ can direct the search towards the wrong regions in parameter space.

We propose a new definition of *collective fitness*, which makes the fitness of each individual θ_i^k dependent also on the value of the objective function for the rest of the population, i.e., on the entire vector $\mathcal{F}^k$ and not just on the value of $\mathcal{F}(\theta_i^k)$.

The new definition of the evaluation function for each individual θ_n^k is a measure of the *matching* between the sampled points and the predicted ambiguity surface. It can, for instance, be based on the likelihood function $L(\mathbf{r}|\theta_n^k)$ for the following estimation problem:

Given observations

$$\mathbf{r} = \{(\theta_1, \mathcal{F}_1), \ldots, (\theta_N, \mathcal{F}_N))\},$$

drawn from the known parametric family of probability distributions

$$p(\mathbf{r}|\theta^o) \sim g_{\theta_n^k}\left(\mathcal{F}_1 - \mathcal{A}(\theta_1 : \theta_n^k), \ldots, \mathcal{F}_N - \mathcal{A}(\theta_N : \theta_n^k)\right),$$

find an estimate of θ^o.

In the above problem, θ_n^k plays the role of the true value of the parameter being estimated, g_{θ^o} is a known distribution, parameterized by θ_n^k which describes the statistical deviation of the observed ambiguity from the predicted ambiguity, $\mathcal{A}(\theta : \theta_n^k)$, assuming that θ_n^k is the true value of the parameter. The above formulation assumes that the observed values of the objective function are distributed around those predicted by the ambiguity surface, with dispersion that may depend on the actual parameter value.

Using the overall shape of the observed ambiguity surface for selecting the next generation of GA's may effectively eliminate secondary extrema of the objective function. The following discussion explains the rational behind our approach. If the location of the secondary lobes predicted by statistical analysis of the problem does not match the observed local extrema of $\mathcal{F}$, then they must correspond to spurious extrema. For concreteness, assume that θ_{i_1} is the individual of the population to which it corresponds the largest value of the objective function:

$$\mathcal{F}(\theta_{i_1}) > \mathcal{F}(\theta_j), \forall j \neq i_1.$$

and let θ_{i_2} be the second best individual. Assume also that the ambiguity function for the problem is not symmetrical. Let $\mathcal{A}(\theta : \theta_{i_1})$ be the predicted ambiguity surface when the true parameter value is equal to θ_{i_1}, and $\mathcal{A}(\theta : \theta_{i_2})$ be the corresponding surface when $\theta^o = \theta_{i_2}$. If

$$\mathcal{A}(\theta_{i_2} : \theta_{i_1}) \simeq 0$$

but

$$\mathcal{A}(\theta_{i_1} : \theta_{i_2}) \simeq 1$$

one can conclude that θ_{i_1} is a secondary minimum, since it yields an ambiguity function that does not match the observed one. Implementation of this general procedure in the context of matched field source localization is described in section 4 of the paper.

3.2. Targeted reproduction/mutation

The second way in which we propose to use *a priori* statistical information about the ambiguity structure of the problem concerns the definition of the stochastic operators that map one population into the next one:

$$\begin{array}{rccc} R: & \Theta^N & \rightarrow & \Theta^N \\ & P_k & \rightsquigarrow & P_{k+1} \end{array}.$$

R is traditionally a pre-defined isotropic probabilistic operator. In standard GA's, R does not depend neither on the iteration index k, nor on the characteristics of the population P_k to which they are applied.

We propose the use of generating operators that are built using the position of the extrema of the *collective fitness* defined above. More precisely, we claim that the new population should be obtained by sampling the following mixture distribution:

$$\begin{aligned} p(\theta^{k+1}|\mathbf{r}^k) &= \sum_{i=1}^{N} p(\mathbf{r}^k|\theta_i^k)\mathcal{A}'(\theta:\theta_i^k) \\ \mathbf{r}^k &= \{\mathbf{r}_i^k\}_{i=1}^N, \quad \mathbf{r}_i^k = (\theta_i^k, \mathcal{F}(\theta_k^i)), \end{aligned}$$

where $\mathcal{A}'(\theta : \theta_i^k)$ is an unit-area version of the ambiguity surface, assuming that θ_i^k is the true value of the parameter.

3.3. Application to source localization

We apply in this section the proposed technique to a source localization problem. In the example presented, we considered the localization of a source in a channel with a bilinear velocity profile, exhibiting multipath. The axis of the duct occurs at a depth of 914 m, for a sound velocity of v_{min} = 1480 m/s, and the gradient of sound speed is $g_0 = -0335$ and $g_1 = -.013$ s^{-1} in the upper an lower layers, respectively. The receiving antenna is a vertical uniform array with 6 sensors and inter-element spacing of 15 m. The receiver searches for the maximum of the Bartlett spectrum,

$$\mathcal{F}(\theta) = \frac{1}{n_{freq}} \sum_{f=f_{min}}^{f_{max}} h_\theta^H C_f h_\theta \,,$$

where C_f is the cross-spectral density matrix at frequency f, and h_θ is a normalized vector describing the nominal acoustic field received for parameters in vector θ. In our simulations, two frequencies f = 50 Hz and f = 60 Hz were used. We considered the use of focalization, by estimating the minimum sound velocity v_{min} along with the source range and depth, (R, D).

The ambiguity structure of the function being optimized can be computed by noting that the cross-spectral matrices C_f have the following structure:

$$C = S\, h_{\theta_0} h_{\theta_0}^H + \Sigma,$$

were θ_0 denotes the true values of the parameters being estimated, S is the signal power and Σ is the noise component matrix. For incoherent noise, asymptotically, $\Sigma \rightarrow \sigma^2 I$, yielding the following model for the observed Bartlett spectrum, dependent on the true parameter values:

$$\mathcal{F}(\theta : \theta_0) = S\mathcal{A}(\theta : \theta_0) + \sigma^2, \tag{3}$$

where we defined the generalized ambiguity

$$\mathcal{A}(\theta : \theta_0) = \left| h_\theta^H h_{\theta_0} \right|^2.$$

Model (3) predicts an observed surface that is a scaled version of the ambiguity surface, plus an unknown constant noise term σ^2.

At each iteration, k, for each element of the population θ_n^k (i.e., for each set of estimates) we compute the generalized angle between the vector of the observed spectral samples,

$$\mathcal{F}^k = [\mathcal{F}(\theta_1^k) \ldots \mathcal{F}(\theta_N^k)]^T$$

and the ambiguity surface predicted for θ_n^k:

$$A(\theta_n^k) = [\mathcal{A}(\theta_1^k : \theta_n^k) \ldots \mathcal{A}(\theta_N^k : \theta_n^k)]^T.$$

The measure of correlation between the observed and predicted surfaces is based on equation (3), which can be written in vector form as

$$\mathcal{F}^k = SA(\theta_n^k) + \sigma^2 = \left[\ A(\theta_n^k) \quad \mathbf{1}\ \right] \left[\begin{array}{c} S \\ \sigma^2 \end{array} \right].$$

Defining

$$M = \left[\ A(\theta_n^k) \quad \mathbf{1}\ \right],$$

it is known that the best estimates of the unknown parameters are

$$\left[\begin{array}{c} \hat{S} \\ \hat{\sigma}^2 \end{array} \right] = M(M^TM)^{-1}M^TS_k = \Pi_M S_k\,,$$

where Π_M is the orthogonal projection matrix in the space spanned by the columns of M, yielding the following correlation between the corresponding estimated S_k and the observed vector:

$$C(\theta_k) = \frac{||\Pi_M \mathcal{F}^k||^2}{||\mathcal{F}^k||^2}.$$

This correlation measure has been used to evaluate each individual in each iteration of the genetic algorithm. Use of $C(\cdot)$ alone is not sufficient, since equally high values (close to unit) can be reached either when θ_n^k is a good estimate, or when all the population is sampling the lower regions of the ambiguity surface. For this reason, the evaluation function combines the values of the Bartlett spectrum at each individual with the normalized correlation :

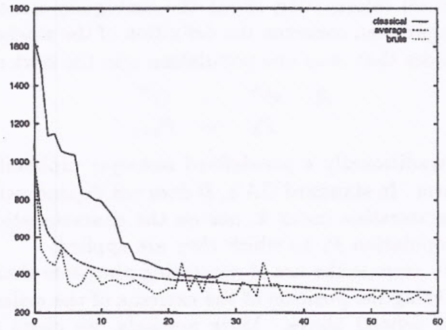
$$\mathcal{E}(\theta_n^k) = \mathcal{F}(\theta_n^k)C(\theta_n^k). \tag{4}$$

Using the product of the two indexes guarantees that high values of $\mathcal{E}$ must correspond both to a large value of $\mathcal{F}$ and to a good agreement of the sampled values to those predicted by the *a priori* model.

Comparison on the basis of the collective fitness instead of isolated scores requires a good sampling of all the dynamic range of the score function. this problem may be solved by evolving in parallel two populations. The first, which we call the *elite* must sample the regions of space with high values of the fitness functions, and provide the potential optimal solutions of the optimization problem. This population is selected according to the ellitist colective criteria presented before. The second population is evolved with a distinct criteria, which is to provide representative values of the score function, and is selected using a representativeness criteria, which measures the amount of individuals that are close to it (present similar values of $(\theta, \mathcal{F}(\theta))$).

We applied this technique to a simple problem where the range of an acoustic source in a two-layer medium is estimated along with one parameter of the velocity profile. Figure 2 compares the mean-square error of the range estimates (over 30 runs of the Genetic Algorithm) of the standard selection procedure (solid line) and of selection using the collective fitness. The population size is 10, and 500 snapshots are used to form the spectral density matrix. Since the fitness value of an individual depends on the entire population, which is randomly replaced at each algorithm iteration, we plot in the figure the error for the best individual of the population (dotted line) and as well as the of the best one according to a smoothed (in time) version of its fitness function (dashed line). As we can see, the modified genetic algorithm is able to find a better solution for this problem than the standard algorithm, which is just using the isolated value of the score function. This result can be justified by the fact that under unwanted parameters the simultaneous optimization of the likelihood function in both wanted and unwanted parameters. i.e., the generalized maximum likelihood, the statistic optimality properties of maximum likelihood no longer hold. In this case, as our example shows, the complete set of score values is indeed providing additional information, which is effectively used by our optimization algorithm. This suggests that for problems of parameter estimation under unwanted parameters, the global geometry of the statistical model being sought must be taken into account in the definition of a suitable score function.

Figure 2: Evolution of mean-square error (averaged over 30 runs).



References

[1] Micahel Collins and W. A. Kuperman. Focalisation: environmental focusing and source localisation. *J. Acoust. Soc. Amer.*, 90:1410–1422, 1991.

[2] Peter Gerstoff. Inversion of seismoacoustics data using genetic algorithms and a posteriori probability distributions. *J. Acoustic Soc. America*, 95:770–782, 1994.

[3] J. H. Holland. *Adaptation in natural and Artificial Systems*. AnnArbor: The University of Michigan Press, 1975.

[4] Solomon Kullback. *Information Theory and Statistics*. Peter Smith, 1978.

[5] H. Martynov Z.-H. Mochalopoulou and M. Porter. Simulated annealing and genetic algorithms for broadband source localization. In J. S. Papadikis, editor, *Proceedings of the 3rd European Conference on Underwater Acoustics*, pages 409–414. Found. for Research and Tech.–Hellas, Inst. of Appl. and Comp. Math., 1979.

[6] José M. F. Moura and M. João Rendas and Georges Bienvenu Environmental limits to source localization. In *Int. Conf. on Acoustic, Speech and Signal Processing95, Detroit, USA*, May 1995.

[7] José M. F. Moura et M.João Rendas. Ambiguity function in sonar. In *Proceedings of the 1992 ASILOMAR Conference, Monterey, CA, USA*, October 1992.

[8] M. S. Pinsker. *Information and Information Stability of Random Variables and Processes*. Holden-Day, Inc., 1964.

[9] M. João D. Rendas and José M. F. Moura. Ambiguity analysis in source localization with unknown signals. In *Int. Conf. on Acoustic, Speech and Signal Processing91, Toronto, Canada*, May 1991.

[10] M. João D. Rendas and José M. F. Moura. Ambiguity analysis in source location. IEEE Trans. Signal Processing, Vol. 46, No. 2, pp:294-305, February 1998.