Maximum-likelihood passive localization using mode filtering

Melvin J. Hinich
Applied Research Laboratories, The University of Texas at Austin, Austin, Texas 78713

Edmund J. Sullivan[a)]
SACLANT Undersea Research Centre, Viale San Bartolomeo 400, 19026 La Spezia, Italy

(Received 13 June 1988; accepted for publication 15 August 1988)

The maximum-likelihood estimator for passive range and depth estimation of an acoustic point source in a shallow-water waveguide is presented. The data from a vertical array of hydrophones are passed through a modal filter, the output of which is the set of complex modal amplitudes associated with the normal-mode model of acoustic propagation. The range and depth estimates are then found by a maximum-likelihood estimation procedure that uses these modal amplitudes as inputs. This technique is compared to the matched-field procedure and is shown to have better signal-to-noise and sidelobe behavior for a given scenario. Results are given for both synthetic and real data. The results with the real data demonstrate the importance of the mode-filtering property of the maximum-likelihood estimator presented in this work.

PACS numbers: 43.60.Gk, 43.30.Jx, 43.30.Wi

INTRODUCTION

The idea of including environmental information within the sonar signal processing scheme has been put forth on several occasions. In the case of shallow-water passive localization, the problem has become tractable due to the existence of sophisticated propagation models coupled with the advent of modern computer technology. In 1973, Hinich[1] used the normal-mode model of propagation to develop the maximum-likelihood estimator for the depth of a point source using the data from a vertical array of hydrophones. This was probably the first time that a sophisticated propagation model was used for source localization. Later, Bucker[2] proposed a scheme that has become known as "matched-field processing." Here, a search is made over forward solutions of the propagation model, where each solution assumes a particular source location. The estimation process consists of comparing these solutions, each of which constitutes a prediction of the field received on the measurement array, to the measured field. Bucker's method of making this comparison effectively consists of forming the inner product between the measured and predicted vector of array outputs. The squared magnitude of this estimator is then plotted on a range-depth map where the coordinates of the maximum constitute the estimate of the range and depth of the source. Various versions of this approach have been put forth.[3,4] More recently, direct inversion techniques have been investigated.[5–7] These approaches are based on normal-mode theory, which allows a set of linear equations in the modal amplitudes to be written that can be directly inverted. These amplitudes contain the range and depth information that must then be extracted as a second step. This second step can of course be carried out by a forward search just as in the matched-field method; the difference being that the direct inversion approach allows mode filtering, since the modal amplitudes are solved for directly.

In both the matched-field and direct inversion approaches, the scenario usually, but not necessarily, consists of a vertical array of hydrophones that samples the vertical structure of the field. The application of the model can then be viewed as being equivalent to the introduction of a matched beamformer, i.e., a beamformer that differs from a standard plane-wave beamformer in that it matches the configuration of the field peculiar to the particular acoustic propagation conditions.

It should be pointed out that although the matched-field techniques can involve prohibitively large computation times, since they must recompute the field for each assumed source location, they offer the luxury of allowing propagation models of any desired degree of complexity, such as range-dependent models, whereas the direct-inversion techniques are based on the linearity of the modal equations and thus are basically limited to the range-independent case. However, since the direct inversion techniques essentially constitute a modal filter, they allow solutions to be carried out using as many or as few modes as desired. This can be an important advantage since numerical studies have shown that errors in the assumed value of the sound velocity profile assumed for the model can impact the modes in diverse ways, depending on the nature of the error.[8,9]

In this article we wish to address the question of the estimator itself. In particular, we show that, given the normal-mode model of propagation, the maximum-likelihood estimator of the range and depth of the source follows directly. The inversion of the modal equations leads directly to a maximum likelihood estimator of the range and depth. This basically constitutes a generalization of the work of Hinich[1] in which the maximum-likelihood estimator of the depth only is treated. In Sec. I the maximum-likelihood estimator is derived. In Sec. II a numerical example is presented in which the performance of the maximum-likelihood estimator is evaluated, for a particular scenario, for several values of the signal-to-noise ratio (SNR). Also a comparison is made between the maximum-likelihood estimator and the

[a)] Presently at: Naval Underwater Systems Center, Code 01V, Newport, RI 02841.

 0001-4966/89/010214-06$00.80

matched-field estimator. In the last part of Sec. II the experimental results are presented. Section III contains a general discussion of the results along with comments on some of the theoretical aspects of the technique.

I. MAXIMUM-LIKELIHOOD ESTIMATOR

The scenario for the problem is as follows. An acoustic point source, located at depth z_0 below the surface, is radiating acoustic energy at circular frequency ω. The source is located at range r from a vertical array of N hydrophones. The acoustic propagation is assumed to be accurately described by the normal-mode model with M being the number of modes being supported by the channel. The acoustic pressure at the nth hydrophone is then given by

$$p(z,t) = b\, e^{i\omega t} \sum_{m=1}^{M} Z_m(z) Z_m(z_0) \frac{e^{ik_m r + i\theta}}{\sqrt{k_m r}}, \tag{1}$$

where k_m is the horizontal wavenumber associated with the mth mode. The modal functions are assumed normalized such that

$$\int_0^h \rho(z) Z_m^2(z) dz = 1. \tag{2}$$

Here, $\rho(z)$ is the density of the water at depth z (measured from the surface) and h is the depth of the water. The quantity b is a scaling parameter the square of which is proportional to the signal power and θ is an arbitrary phase term that must be removed by normalization.

The complex Fourier amplitude at frequency ω, using Eq. (1), is found to be

$$y(z) = b \sum_{m=1}^{M} Z_m(z) Z_m(z_0) \frac{e^{ik_m r + i\theta}}{\sqrt{k_m r}} + n(z), \tag{3}$$

where $n(z)$ is additive noise. It is assumed that the noise is expressible as a modal expansion with random coefficients.[10] That is,

$$n(z) = \sum_{m=1}^{M} e_m Z_m(z). \tag{4}$$

The e_m are assumed to be independent complex Gaussian random variables with zero mean and complex variance σ_m^2. This means that the real and imaginary parts of e_m are independent Gaussian variates with means of zero and variances equal to $\sigma_m^2/2$. This representation of the noise is based on the *crucial implication* of normal mode theory for propagating waves that any received signal is a linear combination of the normal modes. The e_m, of course, could obey different statistics, but this would depend upon whatever physical noise model was assumed. Lacking any particular noise model, these statistics are not unreasonable. In this regard it should be emphasized that the assumption of independence is not crucial to this development.

Now suppose that $y(z)$ is sampled at points z_n, $1 \leqslant n \leqslant N$. Equation (3) then can be written as

$$\mathbf{y} = \mathbf{Z}(\beta + \mathbf{e}), \tag{5}$$

where

$$\mathbf{y} = [y(z_1), y(z_2), \ldots, y(z_N)]^T, \tag{6}$$

$$\mathbf{e} = [e_1, e_2, \ldots, e_N]^T, \tag{7}$$

$$\beta = b\left(Z_1(z_0)\frac{e^{ik_1 r + i\theta}}{\sqrt{k_1 r}}, Z_2(z_0)\frac{e^{ik_2 r + i\theta}}{\sqrt{k_2 r}}, \ldots, Z_M(z_0)\frac{e^{ik_M r + i\theta}}{\sqrt{k_M r}}\right)^T, \tag{8}$$

and the m, n element of $\mathbf{Z}$ is given by $Z_m(z_n)$; $1 \leqslant n \leqslant N$, $1 \leqslant m \leqslant M$. We now define $\mathbf{x}$ via the pseudoinverse operation as

$$\mathbf{x} = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{y}, \tag{9}$$

where $\mathbf{x}$ is an M-dimensional vector given by

$$\mathbf{x} = \beta + \mathbf{e}. \tag{10}$$

At this point a few comments are in order. Equation (9) can be recognized as a form of the normal equation. Thus the vector $\mathbf{x}$ can be considered to be a least-squares estimate of the vector of modal amplitudes. For our purposes, however, it is to be considered simply as a linear transformation that effectively resamples the data in order to reduce their number to be equal to M, the number of degrees of freedom, i.e., the number of modes. Also, the transformation connects the covariance matrix in hydrophone space to that in modal space. Because of the statistics assumed for the noise model, the expected value of $\mathbf{e}\mathbf{e}^T$ is diagonal whereas the expected value of $\mathbf{Z}\mathbf{e}\mathbf{e}^T\mathbf{Z}^T$ is not. If the problem is not overdetermined, i.e., if $N = M$, the term $(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T$ simply reduces to $\mathbf{Z}^{-1}$. From a practical standpoint, Eq. (9) is a critical step in the solution of the problem since it contains a matrix inversion. Thus any sampling problems arising from improper hydrophone spacing or insufficient aperture can result in problems of ill-conditioning. Also, it is at this point that the noise representation given by Eq. (4) can be appreciated, since it leads directly to the form of $\mathbf{y}$ given by Eq. (5). In general, with other noise representations, $\mathbf{e}$ would not be a vector and thus the form of Eq. (9) would not be preserved. Hence, any algorithm based on Eq. (9) would break down.

Since the e_m are independent complex Gaussian variates with complex variance σ_m^2, the maximum-likelihood estimators of r and z_0 are solutions of the weighted least-squares problem given by

$$\sum_{m=1}^{M} \left| x_m - bZ_m(\hat{z}_0) \frac{e^{ik_m \hat{r}}}{\sqrt{k_m \hat{r}}} \right|^2 \Big/ \sigma_m^2 = \min. \tag{11}$$

Here, x_m is the mth element of $\mathbf{x}$. The variances could be measured in the absence of any sufficiently comprehensive knowledge of the noise field. Of course, Eq. (11) could be evaluated as an unweighted least-squares problem, at the cost of some degradation in performance, thereby eliminating the need for the noise variances.

In the evaluation of Eq. (11), x_m and $bZ_m(\hat{z}_0)e^{ik_m\hat{r}}/\sqrt{k_m\hat{r}}$ must be normalized both in magnitude and phase. This was done in practice by normalizing to the first-order mode. Normalizing in this manner, x_1 and $bZ_1(\hat{z}_0)e^{ik_1\hat{r}}/\sqrt{k_1\hat{r}}$ both equal unity.

Let $L(z_0, r)$ denote the logarithm of the likelihood func-

tion for the model. Denoting the expected value of x_m by $b\mu_m$, we have

$$\mu_m = Z_m(z_0)(e^{ik_m r}/\sqrt{k_m r}) \tag{12}$$

and

$$-2L(z_0,r) = M\log(2\pi) + 2\sum_{m=1}^{M}\sigma_m + \sum_{m=1}^{M}\frac{|(x_m - b\mu_m)|^2}{\sigma_m^2}, \tag{13}$$

which for our purposes can be written as

$$2L = \text{const} - \sum_{m=1}^{M}\frac{|(x_m - b\mu_m)|^2}{\sigma_m^2}. \tag{14}$$

The Fisher information matrix elements are given by

$$I_{11} = -\frac{\partial^2 L}{\partial z_0^2} = \sum_{m=1}^{M}\frac{b^2}{\sigma_m^2}\left(\frac{Z'_m(z_0)}{Z_m(z_0)}\right)^2|\mu_m|^2, \tag{15}$$

$$I_{22} = -\frac{\partial^2 L}{\partial r^2} = \sum_{m=1}^{M}\frac{b^2}{\sigma_m^2}k_m^2|\mu_m|^2[1+O(r^{-2})], \tag{16}$$

$$I_{12} = I_{21} = -\frac{\partial L}{\partial r\,\partial z_0} = -\frac{1}{2r}\sum_{m=1}^{M}\frac{b^2}{\sigma_m^2}\frac{Z'_m(z_0)}{Z_m(z_0)}|\mu_m|^2, \tag{17}$$

where $Z'_m(z_0) = \partial Z_m(z)/\partial z$ evaluated at $z = z_0$. The variance–covariance matrix for the parameter estimates $\hat{r}$ and $\hat{z}_0$ has a lower bound given by the relationship

$$\mathbf{v}^\dagger\Sigma^2\mathbf{v} \geq \mathbf{v}^\dagger\mathbf{I}^{-1}(z_0,r)\mathbf{v}. \tag{18}$$

In Eq. (18), r and z_0 are the true values of these parameters. The matrix Σ^2 is the variance–covariance matrix, $\mathbf{I}$ is the Fisher information matrix, and $\mathbf{v}$ is a nonzero weighting matrix. It can be shown that the estimators of z_0 and r are efficient and therefore the equality in Eq. (18) obtains in the limit as $N \to \infty$. When evaluating the Fisher matrix, an estimate of the value of the scaling constant b can be directly computed from the log-likelihood function L, which is linear in this parameter. The solution is given by

$$\hat{b} = \frac{\sum_{m=1}^{M}\text{Re}[x_m\mu_m^*]/\sigma_m^2}{\sum_{m=1}^{M}|x_m|^2/\sigma_m^2}. \tag{19}$$

The asterisk designates the complex conjugate.

As mentioned in the Introduction, the estimator usually used in studies of this type is the matched-field estimator that we designate as E_{mf} and is given by

$$E_{\text{mf}} = |\mathbf{y}_t^\dagger\mathbf{y}_d|^2. \tag{20}$$

Here, $\mathbf{y}_d$ is the vector of measured hydrophone data and $\mathbf{y}_t$ is the vector of hydrophone data as predicted by the model. The dagger designates complex conjugate transpose. In Eq. (20) it is assumed that $\mathbf{y}_t$ and $\mathbf{y}_d$ are normalized. Note that if Eq. (20) is averaged, it can be written as

$$E_{\text{mf}} = Y_d^\dagger\mathbf{R}Y_d, \tag{21}$$

where $\mathbf{R}$ is the covariance matrix of the data.

II. NUMERICAL AND EXPERIMENTAL RESULTS

A numerical and experimental study was carried out for a particular scenario, the configuration of which is depicted in Fig. 1. A vertical array of 32 hydrophones with 2-m spacing is located at a distance of 7.4 km from a 190-Hz acoustic source that is positioned at a depth of 50 m. The depth of the ocean is 103 m. The top hydrophone of the array is located 30 m below the surface of the ocean. The propagation conditions support nine modes. Thus $M = 9$ and $N = 32$.

The experimental data were taken in the Mediterranean Sea just north of the island of Elba, where the bottom is sufficiently flat and smooth to allow the propagation to be modeled as a parallel-plane waveguide and the bottom conditions are known. The sound velocity profile was pressure dominated (isothermal) since the experiment was carried out in late winter. The sound velocity profile and bottom parameters are tabulated in Table I.

The numerical study was based on synthetic data generated by SNAP, the SACLANT Centre's normal-mode model.[11] Noise was added to these data by drawing realizations of Gaussian random variables with a given variance from a random noise generator. The noise variance was defined by

$$\sigma_m^2 = |b\mu_m|^2/\text{SNR}, \tag{22}$$

where SNR is a signal-to-noise ratio to be specified. With this definition of variance the noise power per mode is proportional to the signal power per mode. Clearly, other forms for the noise variance could have been assumed.

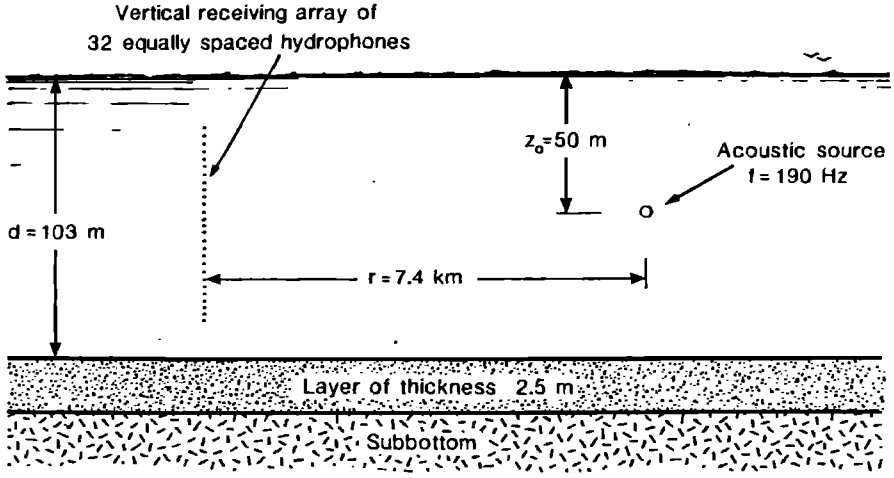

FIG. 1. Source–receiver configuration as used in the experiment and in the numerical example.

TABLE I. Acoustic channel parameters.

Sound velocity profile in water	
Depth (m)	Velocity (m/s)
0.0	1503.0
5.0	1503.1
10.0	1503.2
15.0	1503.3
20.0	1503.4
103.0	1504.7
Sound velocity profile in bottom layer	
Depth (m)	Velocity (m/s)
0.0	1598
0.1	1598
0.2	1523
1.6	1538
2.5	1553

Compressional velocity in layer = 1598 m/s
Density of layer = 1.84 g/m^3
Compressional attenuation = 0.15 dB/λ

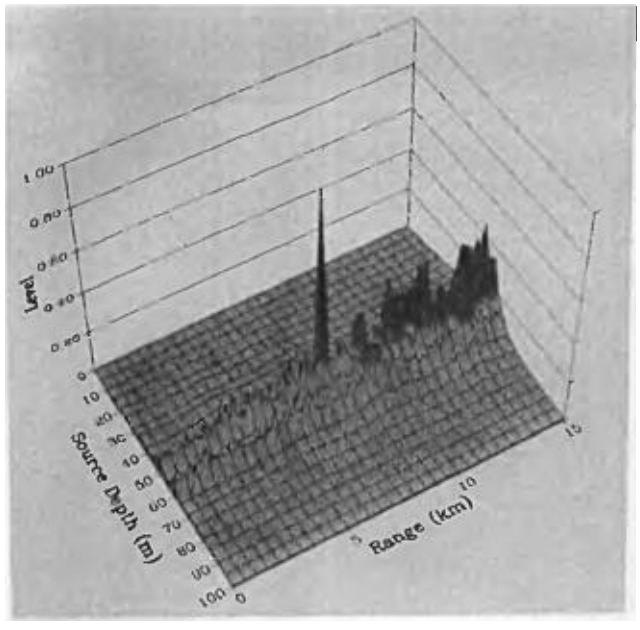

FIG. 3. Inverted log-likelihood surface for the case of SNR = 10 dB. The maximum indicates the minimum of the log-likelihood function.

The results of the numerical study are shown in Figs. 2–8. It should be pointed out that in the plots for the case of the maximum-likelihood estimator, the data have been inverted so as to show the minimum of the log-likelihood function as the maximum on the plot. This inversion was effected by dividing all the data into the minimum. Figures 2–4 show the results for SNR values of 0, 10, and 20 dB, respectively. Figure 5 is an example of the unweighted case, i.e., σ_m = constant, for the 0-dB case. Clearly, the weighting has a significant effect here. However, at higher SNR values, the weighting had no significant effect. The results for the matched-field estimator are shown in Figs. 6–8, which correspond to SNR values of 10 dB, 20 dB, and the no-noise case (SNR = ∞). These data are not inverted as in the case of the maximum-likelihood estimator since we seek the maximum of the matched-field estimator directly. It should be noted that the threshold in these plots has been elevated. This was necessary due to the poor sidelobe behavior of this estimator. With lower thresholds, the maximum was not easily discerned from the plot.

In Fig. 9 the experimental results for the maximum-likelihood estimator are shown. Here, as in the other maximum-likelihood plots, the data have been inverted. In this plot only the first seven modes have been included. There were no solutions using either eight or all nine modes. Similarly, with the plots of the matched-field results, the threshold has been elevated in order to clarify the plot. There were no successful solutions using the matched-field estimator. However, there were also no successful solutions using the

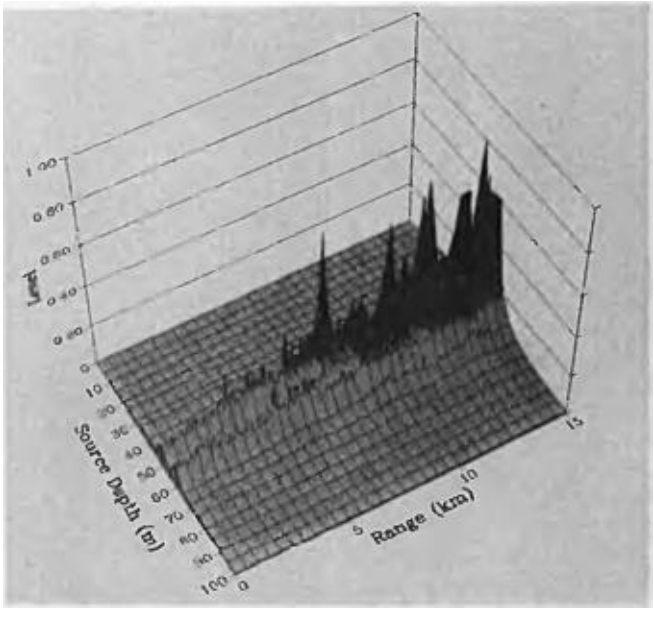

FIG. 2. Inverted log-likelihood surface for the case of SNR = 0 dB. The maximum indicates the minimum of the log-likelihood function.

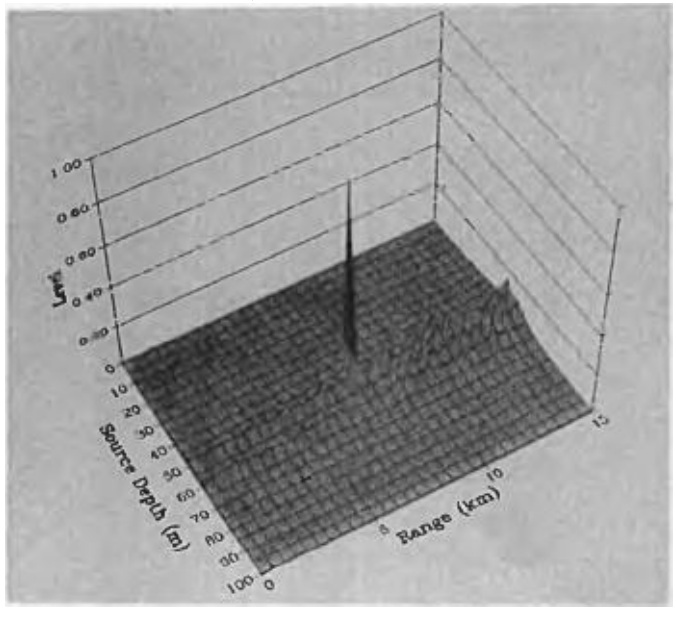

FIG. 4. Inverted log-likelihood surface for the case of SNR = 20 dB. The maximum indicates the minimum of the log-likelihood function.

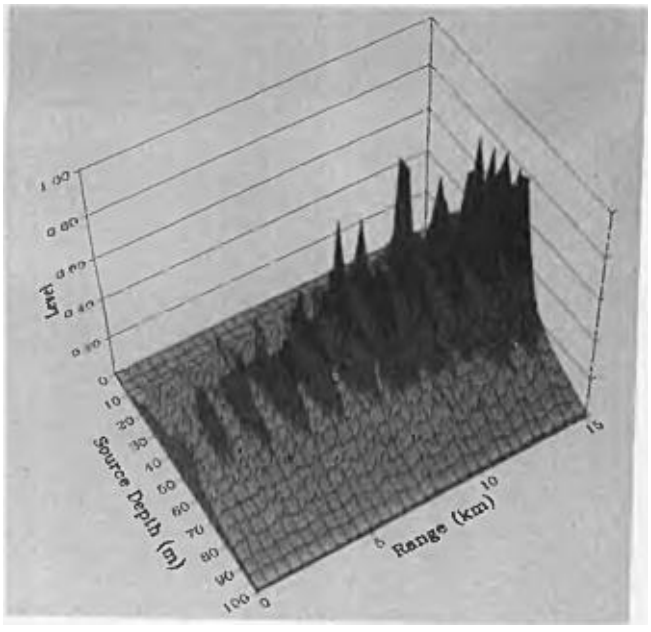

FIG. 5. Inverted log-likelihood surface for the case of SNR = 0 dB and no weighting, i.e., σ_m^2 = constant. The maximum indicates the minimum of the log-likelihood function.

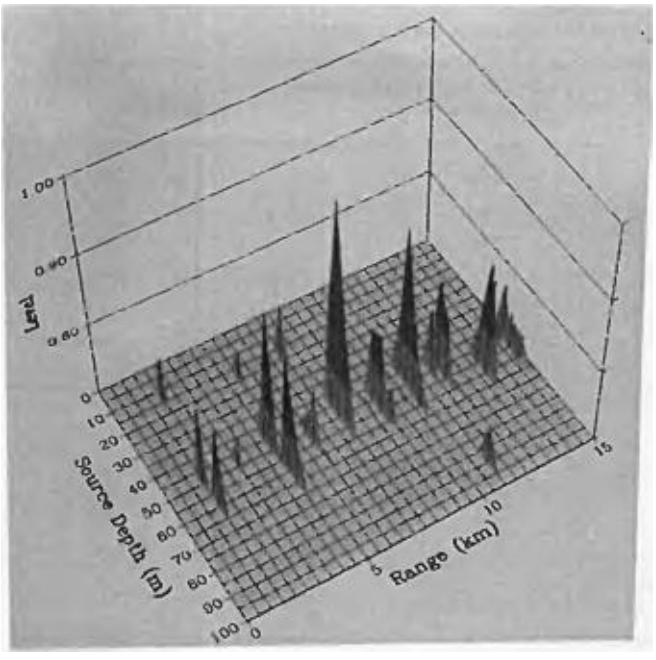

FIG. 7. Value of the matched-field estimator for the case of SNR = 20 dB. The threshold has been elevated for clarity.

maximum-likelihood estimator when the top two modes were not filtered out. Therefore, the failure of the matched-field estimator to produce an experimental solution probably was due to the deformation of the array, since it has been shown that array tilt can have disastrous results.[8]

III. DISCUSSION

It seems clear that, at least for this scenario, the maximum-likelihood estimator outperforms the matched-field estimator. Although this is true in regard to the behavior as a function of signal-to-noise ratio, the major difference between the two estimators is in their sidelobe behavior. Some investigators have shown that this sidelobe problem can be dealt with by preprocessing the data with a high-resolution beamforming process.[3,4] However, this requires an accurate knowledge of the noise covariance matrix, where the maximum-likelihood approach presented here gives good results even when the noise covariance matrix is unknown.

The effect of the weighting by the variances on the maximum-likelihood estimator for low values of SNR is clearly significant. This would have been expected on the basis of the values of the variances that vary over quite a large range. This can be seen from Table II in which the modal amplitudes are tabulated. Since the variances are proportional to these modal amplitudes, their values are equal to the variances for SNR = 1.

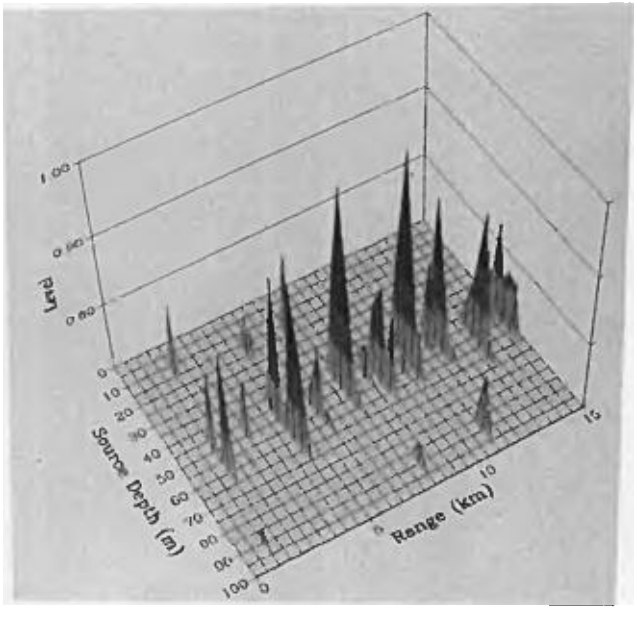

FIG. 6. Value of the matched-field estimator for the case of SNR = 10 dB. The threshold has been elevated for clarity.

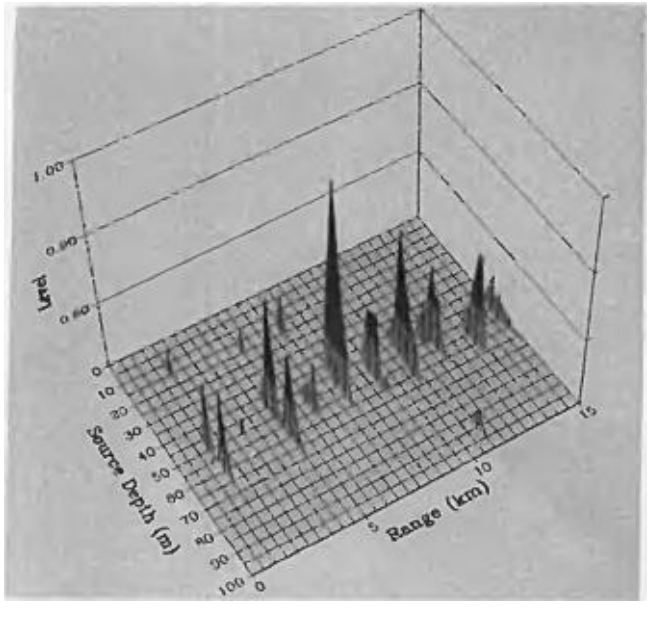

FIG. 8. Value of the matched-field estimator for the case of no noise (SNR = ∞). The threshold has been elevated for clarity.

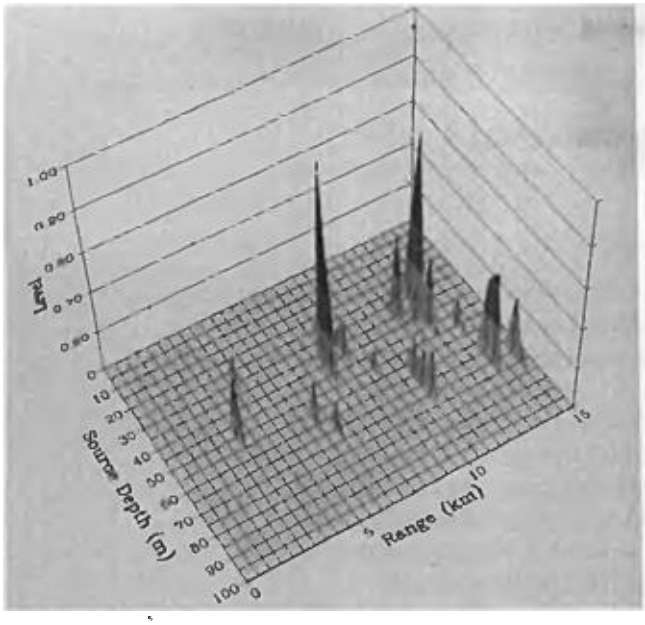

FIG. 9. Inverted log-likelihood surface for the case of experimental data. The maximum indicates the minimum of the log-likelihood function. In this case only the first seven modes are included in the calculation.

In order to gain some insight into the curvature of the log-likelihood surface, the Fisher information matrix was evaluated for the case of SNR = 0 dB. The results are given in Table III. Here, we see two interesting properties. First, the estimates are essentially uncorrelated, and second, the Cramer–Rao lower bounds on the estimates are quite small. As mentioned in Sec. I, these bounds are obtained as $N \to \infty$. From Eqs. (15)–(17), it can be seen that, not unexpectedly, these lower bounds vary directly in proportion to the variance of the noise. In the case of this work, one would not expect the variances on the estimates to be nearly this small since the data are not well averaged since N, the number of hydrophones, is 32.

In regard to the experimental results, it should be emphasized that the data were selected. There were great difficulties in maintaining the array vertical and straight due to currents. As a consequence, the array was constantly in motion and many of the time records produced poor solutions or no solution at all. Although several experimental solutions were obtained, that shown in Fig. 9 was selected for its low sidelobe behavior. In this solution a strong bias in depth is evident. This, most likely, is due to the distortion of the array. A further anomaly of the experimental results is that

TABLE II. Modal amplitudes.

Mode number	Magnitude of modal amplitudes
1	0.149
2	0.027
3	0.123
4	0.064
5	0.091
6	0.078
7	0.057
8	0.071
9	0.017

TABLE III. Elements of the Fisher information matrix for SNR = 0 dB.

$$I_{11} = 0.44$$
$$I_{22} = 3.4$$
$$I_{12} = I_{21} = -9 \times 10^{-6}$$

in no case could a solution be found without filtering out the two highest modes. It is speculated that this also is connected to the deformation of the array, since it would be expected that the phase information of the modes would be more disturbed the higher the mode order. Whatever the reason, this points out the importance of the mode filtering property of this estimator. Even in the case of a perfectly positioned array this filtering property can be of importance for other reasons. For example, in the case of an array of short aperture compared to the depth of the water column or an array with incorrect hydrophone spacings, one would expect that all modes would not be well sampled. In this case, the order of the system could be estimated from the data. Given this order estimate, the offending modes could then be filtered out.

The issue of the array distortion is an example of a more general problem usually referred to as the "mismatch problem," that is, the effect on the estimator of errors in the model parameters. This has been shown to be a major limitation to these model-based passive localization techniques.[8,9] Recently, it has been demonstrated that by using a horizontal (towed) array instead of a vertical array, only knowledge of the modal wavenumbers is necessary to determine range. That is to say, direct knowledge of the sound velocity profile, water depth, and bottom conditions is not required. Furthermore, with a sufficiently long physical array or by using synthetic aperture techniques, these wavenumbers can be directly estimated from the data, thereby avoiding the need for *a priori* knowledge of the model parameters.[12]

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Silvio Bongi of the SACLANT Undersea Research Centre for his valuable assistance with the computer programming and graphics.

[1]M. J. Hinich, J. Acoust. Soc. Am. 56, 499–503 (1973).
[2]H. P. Bucker, J. Acoust. Soc. Am. 59, 368–373 (1976).
[3]R. Klemm, Signal Process. 3, 333–344 (1981).
[4]R. G. Fizell, J. Acoust. Soc. Am. 82, 606–613 (1987).
[5]E. C. Shang, C. S. Clay, and Y. Y. Wang, J. Acoust. Soc. Am. 78, 172–175 (1985).
[6]E. J. Sullivan, SACLANTCEN Rep. SR-117, SACLANT Undersea Research Centre, La Spezia, Italy (1987).
[7]G. R. Wilson, R. A. Koch, and P. J. Vidmar, Applied Research Laboratories Rep. ARL-TP-87-1, The University of Texas at Austin Applied Research Laboratories, Austin, Texas (1987).
[8]E. J. Sullivan and K. Rameau, SACLANTCEN Rep. SR-118, SACLANT Undersea Research Centre, La Spezia, Italy (1987).
[9]E. C. Shang, K. M. Lawson, and D. R. Palmer, J. Acoust. Soc. Am. 79, 958–963 (1986).
[10]M. J. Hinich, J. Acoust. Soc. Am. 66, 480–483 (1979).
[11]F. B. Jensen and M. J. Ferla, SACLANTCEN Rep. SM-121, SACLANT Undersea Research Centre, La Spezia, Italy (1979).
[12]J. V. Candy and E. J. Sullivan, J. Acoust. Soc. Am. (to be published).