Matched-Field Processing for Broad-Band Source Localization
Zoi-Heleni Michalopoulou, Member, IEEE and Michael B. Porter, Associate Member, IEEE

Abstract—In the Hudson Canyon experiment, a sound source moved at a constant depth in 73 m of water while transmitting four tonals. The signal was received on a vertical array of hydrophones that spanned the water column. The data set from this experiment has become a standard test case for studying source tracking using matched-field processing. As part of that process it was important to first determine a suitable environment model and demonstrate the feasibility of matched-field processing. In this paper, we provide the background on the original data processing that was done to accomplish this. Several interesting results emerged from that study. Frequency averaging was demonstrated to be extremely beneficial when used with the Bartlett processor. However, the popular Minimum Variance processor performed poorly. Finally we discuss a very simple approach to combining the energy coherently that provided significantly improved results.

I. INTRODUCTION
In 1993, a matched-field processing [4], [5] workshop was held at the Naval Research Laboratory to provide a blind test of methods for source-tracking [6]. Four of the test cases were synthetic in that the acoustic field was generated by a model. However, data from the Hudson Canyon experiment [7], [8] were also made available. The site of the experiment is in an area off the New Jersey coast where the ocean bottom is nearly flat. (The experiment tracks are near but not actually in the Hudson Canyon where the bottom is much more complicated.)

For a detailed discussion on matched-field processing algorithms we refer the reader to the monograph by Tolstoy [9]. In this work, we will only review matched-field processing very briefly. Fig. 1 shows a simulation of the acoustic intensity due to a 50-Hz source in an environment similar to that of the Hudson Canyon site. The actual array in the experiment sees a characteristic intensity pattern depending on the array’s range from the source. This pattern provides an acoustic fingerprint that uniquely identifies the range to the source. Using an acoustic model one can then simulate the field that would be seen on a replica array placed at various ranges from the source. The location that gives the best match in the intensity pattern corresponds to the true position. In practice, the array is normally fixed and the source moves; however, the idea is the same in both cases.

The simplified view of matched-field processing described above conceals two key challenges. First, an accurate channel model is required to simulate the acoustic intensity pattern. Fortunately, there are many robust and well-documented acoustic models readily available today [10] and for our application the accuracy of the models is primarily limited by environmental knowledge. Second, a variety of signal processing methods can be used to identify the best match between the acoustic fields seen on the real and replica arrays. The simplest method measures the similarity of the data and replica vectors by an inner product.

The structure of the paper is as follows. In Section II, we provide an overview of the Hudson Canyon experiment and demonstrate both narrow-band and broad-band tracking using the Bartlett estimator. We then discuss the sensitivity of those results to environmental parameters in Section III. In Section IV, we study the performance of a popular high-resolution estimator (the minimum variance method). In Section V, we discuss an approach to exploiting broad-band energy coherently. The relative performance of the coherent and incoherent processors versus mismatch is then discussed in Section VI. Finally we end with a summary and conclusions in Section VII.

II. SOURCE TRACKING USING CONVENTIONAL PROCESSING

The scenario for the Hudson Canyon experiment is shown in Fig. 2. The water depth is about 73 m and varied by less
Fig. 2. The Hudson Canyon experiment.

than a meter along the track. The acoustic field was received at an array consisting of 24 vertically separated hydrophones spanning the water column.

Data were collected for two source tracks. The source initially moved away from the receiving array from approximately 500 m to a distance of 4.5 km and then returned towards the receiver. The transmitted signal consisted of tones at 50, 175, 375, and 425 Hz on the outgoing leg and 75, 275, 525, and 600 Hz on the incoming leg. Data were provided for ten different source ranges (frames) for each leg and for each range ten observations of the acoustic field were collected. The source depth was always maintained at approximately 36 m.

As mentioned above, the source localization is done by comparing the actual data vector \( \mathbf{d} \) to a replica vector \( \mathbf{e}(r, z) \). The replica vector represents the field that would be seen on the array if the source were at a hypothetical position \((r, z)\). Both vectors contain 24 complex values representing the amplitude and phase of the acoustic field on each phone. Scanning across a domain of such coordinates we trace out a surface (usually called an ambiguity surface) which measures the similarity of the two fields. Thus, if the modeling is accurate the ambiguity surface will have a peak at the correct source position.

Using the Bartlett matched-field processing scheme, the ambiguity surface is the dot-product of the normalized data and replica vectors. However, to mitigate against noise in the data, such dot products are formed for each of the 10 observations obtained at each range and the results are then averaged. More precisely, one forms

\[
P_{\text{Bart}}(r, z) = \mathbf{e}^*(r, z) \mathbf{C} e(r, z),
\]

where \( P_{\text{Bart}}(r, z) \) is the Bartlett ambiguity surface, \( \mathbf{C} \) is the covariance matrix of the received data \( \mathbf{C} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{d}_i \mathbf{d}_i^* \), \( i = 1, \ldots, N \) where \( N = 10 \) is the number of data observations, and the asterisk denotes the conjugate transpose. In this work the replica vectors are computed with the program KRAKEN, which uses the method of normal modes [11], [12].

The resulting ambiguity surfaces are shown in Fig. 3 for a case where the source was at a range of 1.8 km. Each of the four tonals yields its own ambiguity surface as shown in the upper four plots. The location of the peak on each surface is shown by the large square. When there is a correct localization, the larger square surrounds the small square which indicates the actual source position. Thus, the processing is seen to be successful for the two higher frequencies but fails on the lower ones. Here we define a correct localization as one that is within 5 m in depth and 300 m in range of the true position.

The ambiguity surfaces with incorrect localizations still contain useful information in that they tend to have a high sidelobe at the correct position. Thus, researchers have sought improved performance by averaging the ambiguity surfaces across frequency. This process was studied in simulation in [13]. In deep water the ambiguity surfaces tend to have a regular structure across frequency so that the incoherent averaging is expected to be less beneficial. However, in shallow water the summed surface shown on the bottom of Fig. 3 is improved in terms of peak-to-sidelobe level. Here we have performed a linear summation:

\[
P_{\text{inc, Bart}}(r, z) = \sum_f P_{\text{Bart}, f}(r, z)
\]

Fig. 3. Bartlett ambiguity surfaces for the four tonals and their incoherent sum.
problem (such as array tilt), that it is not difficult to obtain a peak at any desired position. (This motivated the use of blind test cases in the MFP workshop.) Thus, a persuasive demonstration requires successful localizations at many ranges keeping variables that should be static, unchanged. This is shown in Fig. 4, which presents the ambiguity surfaces corresponding to the 10 frames of the outgoing leg. The source is tracked continuously as it moves from the array out to about 5 km.

We have shown the full sequence in Fig. 4 to illustrate clearly the processing that is done and provide a sense of the sidelobe structure that arises at different ranges. These results can be further distilled by plotting the range-depth coordinate of the peak vs. the actual range number as shown in Fig. 5. Solid lines show the true coordinates of the source position; asterisks indicate the estimated source location. The incoherent Bartlett processor identifies the correct position of the source in 18 out of 20 cases.

Fig. 5 shows the robustness of the processing as the source moves in range. Similarly, one may ask whether the frequency averaging provides a consistent improvement. A sense of this may be derived from Table I which shows the number of correct localizations when Bartlett processing is applied to each of the four tonals. The percentage of correct localizations with individual tones varies from 0% to 70%. Thus the incoherent average is significantly better than the results obtained with any single tonal.

A final point of interest in this plot is the performance across frequency. At low frequencies the ambiguity surfaces tend to have low peak-to-sidelobe level even in the absence of environmental mismatch. Thus there are many candidate source positions that lead to a similar acoustic fingerprint. At high frequencies the ambiguity surface shows good peak-to-sidelobe level but the peak is much more sensitive to environmental perturbations. Considering these factors, one intuitively expects that performance will be optimal at an intermediate or mid-frequency band. The results in Table I suggest that the optimum frequency for localization in the Hudson Canyon scenario is around 300 Hz.

### Table I

<table>
<thead>
<tr>
<th>Outgoing Leg</th>
<th>Incoming Leg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td># correct localizations</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>175</td>
<td>5</td>
</tr>
<tr>
<td>375</td>
<td>7</td>
</tr>
<tr>
<td>425</td>
<td>7</td>
</tr>
</tbody>
</table>

III. Sensitivity to Environmental Mismatch

One of the well-known difficulties of matched-field processing is the sensitivity to environmental information. In real data experiments mismatch between the true and assumed environmental parameters is inevitable. In this section, we try to understand the sensitivity of the estimation process to the environment.

Initially, a different model of the ocean subbottom structure than that shown in Fig. 2 was considered. This environment, to which we refer as the original environment, is shown in Fig. 6. The original subbottom was obtained through a sophisticated acoustical inversion process [8]. However, that inversion was done along a track that was near but not coincident with the track used in the experiment. The resulting matched-
field processing estimates are shown in Fig. 7. While these results are not as good as the ones obtained with the new subbottom, they were extremely encouraging. Interestingly, the new subbottom (Fig. 2) was obtained by a nonacoustic method [14], [15]. However, with such limited data it is not useful to try to draw conclusions about the relative merits of the two inversion schemes; the performance difference may not be statistically significant.

Another variable in these results was the ocean sound speed profile. Initially four profiles were provided as shown in Fig. 8. These profiles represent snapshots of the ocean at times a few hours apart and, therefore, show the ocean variability during the experiment. The results presented above used a single representative profile (dotted line) that was obtained by drawing a curve that appeared to summarize qualitatively all four profiles. We refer to this as the empirical profile.

In an attempt to improve our results we calculated the mean profile precisely. As may be seen in Fig. 9 empirical and mean profiles differ by a few meters per second. The localization results obtained using the mean profile are shown in Fig. 10 and are actually somewhat degraded compared to those of Fig. 5. One may tentatively conclude from these results that a practical system will require regular measurements of the ocean sound speed profile.

IV. MINIMUM VARIANCE PROCESSING

The MV processor [9], [16]–[18] has been widely considered as an alternative to the Bartlett estimator. Judging its performance is a subtle process. When it provides a correct localization it tends to produce a surface with higher peak-to-sidelobe level than the Bartlett processor. One tends to intuitively (but incorrectly) have more confidence in the results. However, the high peak-to-sidelobe level is also seen with incorrect results. Critics also point to the interesting but often irrelevant property that its peak level varies significantly with small changes in the environment. The main issue for this application is whether the MV processor provides a correct localization more consistently. The data from this experiment provide a small but useful sample that contrasts with our past experience in which both processors offered similar performance.
The MV ambiguity surface $P_{MV}(r,z)$ is calculated in a manner quite similar to the Bartlett surface:

$$P_{MV}(r,z) = \frac{1}{e^*(r,z)C^{-1}e(r,z)}.$$  \hspace{1cm} (3)

Usually the covariance matrix $C$ is formed from enough data vectors that it is invertible. However, here only 10 observations are available, so, following a standard procedure we loaded the matrix diagonally by adding a small multiple of the identity matrix to it. As before, we also sum the MV ambiguity surfaces across frequency.

The results with the MV processor, shown in Fig. 11, are quite poor showing only about 10% correct localizations. To conjecture about the possible reasons for the poor MV performance, it is useful to look at Fig. 12 which shows the MV ambiguity surfaces for 50, 175, 375, and 425 Hz along with their sum for one particular case. Note that the power level of the top surface (50 Hz) is much higher than that of the other three. The sum (bottom ambiguity surface) is, therefore, almost a replica of the 50-Hz MV ambiguity surface so that there is no real gain from the incoherent average.

Having identified this problem, there are several obvious fixes one might try. In particular, we implemented a new incoherent MV processor by adding the narrowband surfaces after first normalizing them to a maximum of 1. This raises the interesting argument about how one should weigh ambiguity surfaces across frequency. The weighting is especially important for the MV processor since its peak level can vary significantly across frequency. However, even after the normalization scheme, the MV localization results shown in Fig. 13 are only mildly improved yielding a success rate of approximately 30%.

Another factor degrading the performance of the incoherent MV processor is that the peaks on the MV surface are much sharper than those of the Bartlett surface. In many cases, there is a frequency-dependent bias in the peak position. Because of the narrowness of the peaks along with the presence of a bias, stacking across frequencies does not lead to reinforcement.

Finally, it should be noted that we experimented with different levels of diagonal loading of the covariance matrix. The source location estimates varied with the loading level.
but the statistics of successful and poor localization remained essentially unchanged.

V. COHERENT BROAD-BAND SOURCE LOCALIZATION

The processors discussed in Sections II and IV have an intrinsic weakness in that they are discarding useful information. They are incoherent space-only processors in the sense that they exploit spatial coherence within a single frequency and ignore coherence across frequencies. The implementation of coherent processing has been proposed in [9], [19], [20].

Here, we suggest a space-frequency processor that exploits correlations among the acoustic fields at different frequencies. The new processor is based on the formulation of “super-vectors.” To be precise, we denote the data vector for the \( m \)th frequency by \( \mathbf{x}_M \). The super-vector is then:

\[
\mathbf{y} = \begin{bmatrix}
\mathbf{x}_1 \\
\mathbf{x}_2 \\
\vdots \\
\mathbf{x}_M
\end{bmatrix}
\]  (4)

If there are \( L \) hydrophones and \( M \) frequencies, the super-vector is a column vector of \( ML \) elements. We may then construct an extended covariance matrix from the super-vectors and form Bartlett and MV surfaces in the usual way. However, there is one important complication. The source spectrum is not generally known, but the acoustic models predicting the replica vectors assume it is given. If the replica vectors are not adjusted for the source spectrum, then they will not match the received field and the processor performance will be severely degraded. (The Bartlett and MV surfaces are insensitive to the phase and amplitude of a single frequency but are sensitive to such variations across frequencies in the super-vectors.)

To compensate for this problem, we have here scaled the data and replica vectors at each frequency so that they have zero phase on the first phone and unit length. (This process can be extended in several ways. For instance, in a low signal-to-noise Ratio (SNR) case it would be better to use the phase of the phone with the highest SNR for phase normalization.) The results of the coherent Bartlett and MV processors are shown in Figs. 14 and 15. Both figures indicate that the new Bartlett and MV processors are successful 90% of the time. The Bartlett performance is the same as that of the incoherent Bartlett processor (Fig. 5). Fig. 15, however, shows a great improvement for the MV processor, now yielding nearly perfect tracking of the source in range and depth.

Ambiguity surfaces for both coherent processors computed for the data used in Figs. 3 and 12 are shown in Fig. 16. Comparing the coherent MV ambiguity surface of Fig. 16
to the incoherent MV surface of Fig. 12, one can notice the substantial improvement when frequency coherence is accounted for in the estimation process. The source is localized correctly in the coherent framework (estimated range was 1.79 m and estimated depth was 34 m), whereas the source location estimates were wrong when the incoherent MV processor was used (estimated range was 3.25 km and estimated depth was 62 m).

Extending the vectors across frequency may be compared to increasing the spatial aperture of an array. Increasing spatial aperture tends to improve robustness against mismatch. Adding multiple frequencies coherently provides effectively an increase in ‘aperture’ and a similar improvement in robustness.

VI. PERFORMANCE EVALUATION OF INCOHERENT AND COHERENT PROCESSORS

In the above sections we have seen that incoherent averaging provided a great improvement for the Bartlett surface but not for the MV surface. We have also seen that a simple coherent version of both processors provides excellent results in both cases. One would like to understand whether this is generally true. In the absence of additional data, we look now at simulations to see if we can gain confidence in the conclusions by duplicating them in simulation.

We conjecture that the performance differences can be understood in terms of differences in the sensitivity of each of the processors to environmental mismatch. In particular, we have chosen to vary the ocean sound speed profile.

When the environment is known exactly and the SNR is high, all four processors localize the source correctly. The SNR for the simulations is set to 10 dB. In our simulations we used the empirical profile, $c_{emp}$, shown in Fig. 8 to generate the replica fields. We then constructed an empirical orthogonal function, $\Psi(z)$ where $z$ is depth, that represents the shape of the dominant variation in the ensemble of sound speed profiles [21].

A mismatched sound speed profile was then generated by $c(z) = c_{emp}(z) + \mu \Psi(z)$ where $\mu$ is a mismatch coefficient. Fig. 17 shows the original profile and the perturbed profiles for $\mu = 10$ and $\mu = 50$ (smallest and largest values of $\mu$ used here). Data were generated for frequencies of 50, 175, 375, and 425 Hz for different values of $\mu$. Ten observations of data vectors were created for every examined case adhering to the parameters of the real data from the experiment.

Now, we will compare the four main processors discussed above, that is the incoherent and coherent versions of the Bartlett and MV estimators. (The incoherent MV processor implemented here was the first one implemented in Section IV, that is, the narrow-band ambiguity surfaces were added without any normalization.) The statistic we study is the probability of miss (wrong localizations) as the mismatch coefficient $\mu$ is increased. The probability of miss was calculated
after taking into account localization results for a number of runs corresponding to different source locations and noise realizations.

The results shown in Fig. 18 show many interesting features. First, the incoherent MV processor appears to have a much poorer performance than the other three processors, having a high probability of miss even for a small mismatch in the sound speed profile. The incoherent Bartlett and the coherent Bartlett and MV processors are very similar in their performance. However, there is a crossover point near $\mu = 25$. For $\mu > 25$ the two coherent processors show the best performance but all three processors seem to converge for a large degree of mismatch ($\mu = 50$). Note also that the coherent Bartlett processor actually does better as the mismatch is increased from $\mu = 20$ to $\mu = 30$. This is believed to be a consequence of the limited number of trials that could be included. We also note that simulations were done with the number of observations increased from 10 to 100. However, no substantial alteration of the comparative performances of the processors was noticed.

In summary, it seems very difficult to conclude which processor should be recommended. When mismatch with respect to other parameters is considered, the performance patterns of the different processors will likely be still more complicated.

VII. SUMMARY AND CONCLUSION

Matched-field processing in shallow water has been generally viewed as a challenging proposition because of the difficulty of accurately predicting the acoustic field in such environments. In contrast to deep-water problems, there is typically a strong effect of the ocean bottom which in turn is typically known very inaccurately. Despite these difficulties, the results presented here clearly demonstrate the feasibility of continuous source tracking using matched-field processing. However, the exploitation of broad-band information was seen to be important in its success.

The comparison of different estimators revealed some interesting properties. The MV processor performed much worse than the simpler Bartlett estimator and an explanation has been suggested. A simple scheme for exploiting coherence between frequencies has also been demonstrated and shows a significant pay-off in performance. In practice such coherences may be strong or weak depending on the environment and the coherence being considered. In simulation, we have also seen that the relative performance of the processors depends in a complicated way on the environmental mismatch.

ACKNOWLEDGMENT

Several scientists at NUWC, the Woods Hole Oceanographic Institution, and Science Applications International Corporation were instrumental in both processing the data and providing scientific information about the experiment (E. Hug, J. Doutt, R. Evans). The Hudson Canyon experiment was supported by ONR Ocean Acoustics and the Naval Undersea Warfare Center (Chief Scientist, W. Carey).

REFERENCES


Zoi-Heleni Michalopoulos (M’93) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from Duke University, Durham, NC, in 1990 and 1993, respectively.

Between January 1994 and August 1995, she was a research assistant professor at the Department of Mathematics, New Jersey Institute of Technology, Newark, NJ. Since September 1995, she has been an Assistant Professor at the Department of Mathematics and the Department of Electrical and Computer Engineering at the New Jersey Institute of Technology. Her research interests include signal processing, detection and estimation theory, and underwater communications. She is a member of the Acoustical Society of America and the Technical Chamber of Greece.
Michael B. Porter (A'95) was born in Quebec, Canada, on September 19, 1958. He received the B.S. degree in Applied Mathematics from the California Institute of Technology (Pasadena, CA) in 1979. In 1984 he received a Ph.D. in Engineering Science and Applied Mathematics from Northwestern University (Evanston, IL).

From 1983 to 1985 Dr. Porter was a scientist at the Naval Ocean Systems Center, San Diego, CA, where he performed research in numerical modeling of antennas, optical fibers, sonar transducers, and ocean acoustics problems. From 1985 to 1987 he was employed as a research physicist at the Naval Research Laboratory in Washington, DC. At NRL his interest included matched-field processing and computational acoustics.

From 1987–1991 he continued research in these areas as a senior scientist at the SACLANT Undersea Research Centre (NATO) in La Spezia, Italy. He is presently a Professor of Mathematics at the New Jersey Institute of Technology.

Dr. Porter is a member of the Acoustical Society of America and the Society for Industrial and Applied Mathematics.