1. INTRODUCTION

A major concern in the Marginal Ice Zones Experiments (MIZEX) during the summers of 1983 and 1984 was the ice-ocean eddies. These eddies have diameters of 20 to 50 km, periods and lifetimes of at least 20 days, vertical structures penetrating to depths of 500 to 1000 m (with the most significant variation over the first 100 m), and advection rates of 5 to 15 km per day (O. M. Johannessen et al., unpublished manuscript, 1986). The formation of eddies due to various mechanisms [Johannessen et al., this issue] is common; they are perhaps the dominant mesoscale process in the MIZ. During the summer, the abundant eddies along the ice edge advect warm Atlantic water under the ice which enhances the melting of the edge by 1 to 2 km/d [Johannessen et al., this issue]. Josberger [1984] has reported that without the presence of the eddies, the retreat of the ice edge could be reduced by a factor of 10. The eddies also exist during winter, as has been verified by satellite observation, but less is known about their effects during this season.

In order to increase our understanding of the generation, propagation, and dissipation of MIZ eddies and their influence on the dynamics of the ice edge and the polar climate, we must measure eddy structures adequately in space and time. However, their relatively short lifetimes, length scales, and time scales (as compared with mid-ocean eddies) impose some difficulties in sampling. In order to overcome these obstacles, a hybrid observational system was employed during MIZEX in 1983 and 1984: remote-sensing satellites and aircraft were used to first locate the eddies, and this information was then used to guide ships in real time to conduct hydrographic surveys. Without the guidance of the remote sensors, the ships could easily have missed the eddies, and without the hydrography, the internal structures of the eddies could not have been observed. However, unfavorable weather conditions can be a major obstacle to this type of approach. Clouds can prevent satellites and aircraft with sensors outside the microwave range from observing the external eddy signatures directly below them. Furthermore, shipborne surveys are slow and easily affected by unfavorable weather.

In view of the limitations of aerial remote sensing and shipborne surveys, acoustic tomography comes to mind as an alternate and supplementary observational system for MIZ eddy fields. The advantages of acoustic tomography are that (1) the system can be implanted in the ocean on a semipermanent basis for continuous observations, (2) it is not affected greatly by weather conditions, (3) it has high temporal resolution, (4) it can cover an extensive volume of the ocean interior and probe the different parts simultaneously, and (5) only a few moorings are needed, thus minimizing the effort in deployment and maintenance.

In brief, an ocean acoustic tomography system consists of a sparse horizontal array of moored sources and receivers (or transceivers for reciprocal, two-way transmissions to monitor currents) surrounding the area of interest. Each of the sources (or transceivers) repeatedly sends out phase-modulated pulse
codes [Spindel, 1985]. The transmitted sound energy propagates to the receivers (or transceivers) along many distinct routes (i.e., the acoustic multipaths) that cycle through different sections of the water column. The detected signals are then demodulated and decoded to give the impulse response of the ocean. Time series of multipath arrival times contain information on the cumulative changes of the ocean along the multipaths. Perturbative inverse techniques are then applied to the travel time anomalies and the travel time differences of reciprocal transmission paths to optimally reconstruct the four-dimensional temperature and flow fields, respectively.

From a systems viewpoint, the adequacy of an oceanographic observational method is judged by the precision of the estimates of the fields generated. The mean square error of the estimate is one of the most important measures of system performance, since it accounts for the total error. The mean square error consists of the sum of two independent components. The first component is the variance, which quantifies the effect of random experimental noise on the estimate. The other component is the square of the bias. This systematic error is a consequence of the finite number of experimental data and thus limits resolution. Minimization of the mean square error is an intuitively meaningful and widely accepted criterion for designing optimal estimators. This, together with the criterion that the estimates be linear combinations of the data [Liebelt, 1967] is the foundation on which our analysis of the performance of acoustic tomography in the MIZ is built.

Resolution is another important, though not independent, measure of system performance. Its consideration can lead to the determination of the size and period of the smallest and most oscillatory ocean feature that can be detected by the system. In a linear, minimum mean square error reconstruction of the perturbed field, resolution and mean square error are linearly related. As resolution increases, the error decreases. Although the variance of the estimate vanishes in a noise-free experiment, the other error component, the square of bias, remains nonzero owing to the finite resolving power exerted by a data set of finite size. In the presence of experimental noise, a trade-off between variance and resolution exists: as resolution increases, so does variance. Thus resolution may be adjusted to reduce the deviation of an estimate from its mean. However, any of such adjustments of resolution from that determined by the minimum mean square error estimator can only result in an increase of bias, which in turn can only lead to an increase of the total error.

The purpose of our work here is to give a preliminary evaluation of the system performance of acoustic tomography in the MIZ. The main focus is on the spatial resolution of the eddy fields in a time interval much shorter than the eddies' characteristic time and, on the other hand, much longer than the travel times of acoustic pulse. A suitable choice of the time interval is about 1 day, since during this time the eddy fields change little, whereas the acoustic travel time along each ray path can be sampled several times to yield an improved signal-to-noise ratio through averaging. Temporal resolution is of no concern, since the transmission rate is completely controllable and can always be adjusted to fulfill a particular Nyquist sampling criterion. An assumption used in our analysis is that there are no correlations between eddy fields inferred at different times. From a systems viewpoint, this assumption simply expresses an ignorance of eddy dynamics and eliminates the input of as yet uncertain dynamical information. Our study of resolution, or similarly of the statistical errors in the estimates, can determine the potential benefits and limitations of the acoustic system as well as give us hints on how to counteract system weaknesses.

Using the well-known Backus-Gilbert geophysical inverse theory [Backus and Gilbert, 1967, 1968, 1970; Eisler et al. [1982] have studied the horizontal planar resolution of ocean acoustic tomography, particularly of the array used in the Ocean Tomography Experiment of 1981 [Ocean Tomography Group, 1982]. (The Backus-Gilbert theory is closely analogous to the minimum mean square error method we will employ.) To simplify their calculations, Eisler et al. made the assumption that the acoustic ray paths are straight lines. Implicit in this is the assumption of a uniform medium. Therefore the resolution they calculated is determined completely by array configuration in the x-y plane and can only be interpreted as the resolution of the depth-averaged field. In a sparsely explored environment such as the MIZ, however, a study of the resolution of the depth-averaged field is perhaps not the most appropriate study for the evaluation of system performance. In reality, the vertical sound-speed structure determines the trajectories of the acoustic multipaths. The horizontal and vertical resolution in the vertical planes containing the multipaths depend strongly on the differences of the paths' cycle distances and depths of turning points, respectively. To examine ocean effects on system performance, both the horizontal and vertical resolutions in these vertical planes must be considered using realistic sound-speed profiles.

The organization of our paper is as follows. In section 2 we present the relations between the acoustic observations, i.e., travel time anomalies, and the desired unknown variables, the sound-speed perturbation $\delta c$ and current $v$. (Note that $\delta c$ is approximately proportional to the temperature perturbation.) Attention is paid to the resolvability of ray arrivals, upon which the success of tomographic travel time inversions critically depends. The procedures for constructing linear estimates and the mathematical expressions for variance, bias, resolution, etc., are well known at this time. Thus without too much elaboration, we write out any needed inverse theory expressions in discrete form in section 3, using the matrix notation of Wiggins [1972]. Local measures of resolution, analogous to those used by Eisler et al. [1982], are also defined. Some of the more intricate details of the constructions of the inverse solutions, error estimates and resolution kernels in this study may be found in the appendix, if the reader is interested. In section 4, the results of our study on resolution and accuracy, as well as simulated inversions, are presented and discussed. In that section we also examine how the inclusion of satellite and in situ point observations results in improved system performance. Conclusions are discussed in section 5.

2. ACoustic TRAVel TIMES

Normal modes and rays are two commonly employed representations of the acoustic field due to a point source in the water column. While the former constitutes an exact, frequency-dependent solution to the governing wave equation, the latter is a frequency-independent, asymptotic approximation of geometric optics. For most underwater acoustic applications, the variation of the refractive properties of the medium is small over an acoustic wavelength, and so geometric optics is adequate. But the ray approximation cannot account for sound diffraction and thus needs corrections at
The basic input to ray acoustics calculations is the sound speed (or alternately, index of refraction) of the ocean in the area of interest, which is in general a function of all three spatial coordinates and time. It is the operational philosophy of acoustic tomography that in general the sound speed can be represented in most cases by the sum of an overall background profile \( c(z) \), which governs the basic arrival structure of acoustic energy and is a function of depth only, and a small perturbation field \( \delta c(x, y, z, t) \), which is the desired result of the inverse theory calculations. The background sound speed profile we will use in this study is typical of the MIZ in sen et al., 1985]. The speed of sound in water depends strongly on both pressure and temperature. Below a depth of 700 m, the pressure dominates, so that sound speed increases almost linearly with depth. However, the warm Atlantic water confined in the upper layer increases the sound speed there. As a result, a weak sound speed minimum is formed.

It should be noted that in performing our analysis, we are purposely limiting ourselves to the case of a MIZ summer profile which is disturbed by the presence of ice-ocean eddies. Thus our study does not include the effects of (1) seasonal changes in the background profile, (2) the often sharp discontinity in the upper (≈ 50 m) layer of the sound speed profile caused by the ice edge (if it is in the transmission path), and (3) the extremely large perturbation in the background profile (over the top 500 m or so of the water column) caused by the front between the warm West Spitsbergen Current and the Arctic-influenced waters typical of the East Greenland Current (again, if it intersects the transmission path). Our avoidance of examining seasonal profiles other than summer is purely due to a lack of data, hopefully to be remedied in part by the upcoming MIZEX in the winter of 1989.

The question of how to deal with fronts, such as those due to the ice edge or current boundaries, is a thornier one and worthy of a full study in its own right. However, we can at least comment on how one might approach these frontal transitions, even if we do not pursue this matter much further here. The front due to the ice edges in the MIZ can to a large extent be ignored except for shallow turning rays, since the ice edge front itself is shallow. To incorporate it for rays which it does affect appreciably, it is obvious that one must break up the background into (at least) two sections described by two background depth profiles, one for each side of the front. This means one must get as background information the sound speed profiles of both sides of the front as well as the position of the front. The background sound speed profiles are obtained from the hydrographic survey made as a standard part of ocean acoustic tomography experiments—it is the position of the front that is harder to obtain. In the case of the ice edge, it seems reasonable that satellite remote sensing, particularly in the microwave range, could be employed to supplement the acoustic system by giving the position of the ice edge (R. A. Shuchman et al., unpublished manuscript, 1985). In the case of the West Spitsbergen Current boundary, the general approach of breaking up the background into two regions is the same as in the ice edge case; indeed, hydrographic surveys show the characteristics of this transition very clearly (F. R. DiNapoli, private communication, 1986). Again, satellite studies (in particular, passive radiometry) should be able to locate the frontal boundary in open water. However, for regions where the front runs under ice, the problem of location could be considerably harder.

To evaluate the importance of incorporating bottom-reflected arrivals in the inversions, a bottom boundary is included in the ocean model. For simplicity and without loss of generality, a constant depth of 2.5 km, corresponding to the nominal depth in this region, is assumed. Bottom-reflected tomography signal arrivals in the MIZ have been shown by Lynch et al. [this issue] to be detectable and usable in a tomographic analysis. The issues of ray instability, ray identification difficulity, and nonlinearity in the tomographic inversion due to errors or unknown variations in the bottom bathymetry [Palmer et al., 1983] are avoided in our analysis by treating the bottom as perfectly flat, but will have to be accounted for in real world experiments. Our reason for omitting this error term is that we presently do not know the magnitude of the bathymetry error or even, if given the error, how to calculate precisely its effect on the inversion.

Figure 2 is a ray diagram, generated using geometric optics, showing the trajectories of the non-bottom-interacting multi-
Fig. 2. Ray diagram, showing the trajectories of the non-bottom-interacting multipaths that connect two transceivers both located at a depth of 350 m and separated by a range of 150 km.

Fig. 3. Arrival time structure of the rays for an ideal source with infinite bandwidth.
For an ocean that has a mesoscale sound speed perturbation field \( \delta c(x, z) \) and a mesoscale horizontal current field \( v(x, z) \) during a time period much longer than the transmission intervals in the vertical slice where the multipaths lie, and neglecting smaller-scale internal wave fluctuations, the time required for an acoustic pulse to travel from transceiver to transceiver along a (perturbed) ray path is given by

\[
t^f = \int \frac{ds}{c + \delta c + v(dx/ds)}
\]

where \( x \) and \( z \) are the horizontal and vertical coordinates, respectively, and \( s \) is the arclength along that path. The pulse travel time for the same path but in the reverse direction is

\[
t^b = \int \frac{ds}{c + \delta c - v(dx/ds)}
\]

Under the mild condition that \( c \gg |\delta c| \gg |v| \), and for shallow-angle rays with \( dx/ds \approx 1 \), the deviation \( \delta t^+ \) of the sum \( t^b + t^f \) and the difference \( \delta t^- = t^b - t^f \) of the two reciprocal travel times, as caused by \( \delta c \) and \( v \), can be linearized to

\[
\begin{align*}
\delta t^+ &= -2 \int \frac{\delta c}{c^2} ds \\
\delta t^- &= -2 \int \frac{v}{c^2} ds
\end{align*}
\]

respectively. Typically, \( c \approx 1500 \text{ m/s} \), \( \delta c \) is of the order of meters per second, and \( v \) is of the order of centimeters per second. Thus the condition is almost always satisfied. Furthermore, for stable rays, that is, any rays that exist in the background and do not disappear or alter their trajectory drastically in the perturbed state, the integrals in (3) and (4) can often simply be evaluated along the unperturbed trajectory rather than the perturbed one with little loss of accuracy [Hamilton et al., 1980]. And even if the perturbation is large enough to make the problem somewhat nonlinear [Mercer and Booker, 1983], one can iterate the inverse problem and use updated ray paths in the estimates. The time averages over this time period of interest of the observed \( \delta t^+ \) and \( \delta t^- \) from each of the multipaths constitute the database for the inverse problem. The data sets are generally contaminated by experimental noise. Because the two components are statistically independent, \( \delta t \) can be expressed as the sum of the two corresponding error covariance matrices, that is,

\[
\delta \hat{p} = C_e A^T C_e^{-1} y
\]

where \( \hat{p} \) is a \( n \)-dimensional vector parameterizing an unknown field, either \( \delta c(x, z) \) or \( v(x, z) \), in the vertical slice. In other words, the components of \( p \) are either the sound speed perturbations or the currents in each of the small boxes. Moreover, \( y \) is a \( m \) dimensional data vector that (depending on what \( p \) parameterizes) is composed by the (time averaged) anomalies of the sums or differences of the reciprocal travel time observed from \( m \) resolvable paths; \( e \) is a \( m \) dimensional vector denoting the random experimental noise in the data, and all vectors are column vectors. We assume uncorrelated noise and that \( e \) has zero mean and known covariances. Since the two sets of kernels that operate linearly on \( \delta c \) and \( v \) are identical, the corresponding discrete linear operators are also identical, and is represented by the \( m \times n \) matrix \( A \) in (5). The product \( A \hat{p} \) in (5) thus represents the signal produced by the perturbations parameterized by \( p \).

In general, (5) constitutes a highly underdetermined system with \( n \gg m \). Without additional constraints on the solution for \( p \), the system admits infinitely many plausible solutions. The objective in an inversion of data is to obtain the best possible solution that satisfies some well-defined "optimal" criteria in addition to fitting the data well. Following Cornuelle et al. [1985], we define the best estimate of \( p \) to be the one that is linear with the data \( y \) and has the least mean square error. With \( p \) and \( e \) uncorrelated, the Gauss-Markoff theorem asserts that such estimate is unique and can be written as

\[
\hat{p} = C_e A^T C_e^{-1} y
\]

where

\[
C_e = C_p - (C_p A^T(A C_p A^T + C_o)^{-1} (C_p A^T))^T
\]

is the covariance matrix of the total error \( e = \hat{p} - p \) in the estimate, and \( C_p \) and \( C_o \) are the covariance matrices of noise \( e \) and unknown parameters \( p \), respectively [Liebelt, 1967]. Through the specification of \( C_p = "a \ priori" \) knowledge concerning the perturbed field can be parameterized and input to the estimator. The diagonal elements of \( C_e \) are the mean square errors of the estimates in each of the boxes, which are the smallest among all linear estimates.

The total error \( e \) in the estimate can be broken into two components, namely the bias \( b = \langle \hat{p} \rangle - p \) and the random error \( \delta \hat{p} = \hat{p} - \langle \hat{p} \rangle \), where \( \langle \hat{p} \rangle \) denotes the expected value of \( \hat{p} \). While the former corresponds to a systematic error, attributed solely to the finiteness of the sample size, the latter arises solely from the randomness in the data, as caused by experimental noise. Because the two components are statistically independent, \( C_e \) can be expressed as the sum of the two corresponding error covariance matrices, that is,

\[
C_e = \text{bb}^T + C_{\delta \hat{p}}
\]

Using (6), it is easily verified that the covariance matrix of \( \delta \hat{p} \) can be expressed as

\[
C_{\delta \hat{p}} = C_e A^T C_e^{-1} A C_e
\]

We can define

\[
R = VA(I + \Lambda^2)^{-1} AV^T
\]

as the resolution matrix, where \( \Lambda \) and \( V \) are matrices containing the eigenvalues and eigenvectors, respectively, obtained from a matrix factorization of \( A \) as discussed in the appendix. Each column (or row) of this symmetric matrix corresponds to a resolution kernel. The \( i \)-th resolution kernel, that is, the \( i \)-th
column of \( R_i \) represents the best estimate of a field that is perturbed only at the \( i \)th box, in the case that the perturbations at different boxes are uncorrelated. When correlations exist, this perturbation field corresponds to a broader pulse centered at the \( i \)th box, and the widths of this pulse in the horizontal and vertical directions are approximately the horizontal and vertical correlation lengths of the field, respectively. In general, each of the resolution kernels peaks at the correct location, but energy leaks into the other boxes. This results in the broadening of the main peak, the reduction of the amplitude of the peak, and the generation of side lobes. Although some energy can generally be found at faraway distances, the significant portion of the energy is always confined within the main peak.

A simple measure of local resolution at each box is the amplitude of the main peak in the corresponding resolution kernel. The closer the amplitude to unity, the less the energy leaks and hence the better the resolution. The relation between mean square error and resolution is

\[
C_t = C_{pt}/2(I - R)C_{pt}/2 (11)
\]

(This expression can be obtained by substituting (10) in (A9) in the appendix.) It is seen that the relation between resolution and mean square error is direct and linear. As resolution increases, the error decreases. In the limit of ideal resolution, that is, when all the resolution kernels resemble delta functions, \( R = 1 \) and \( C_t = 0 \).

In this study, two other appropriate measures of local resolution are used for physical reasons, namely the minimum horizontal and vertical resolution lengths, \( \lambda_h \) and \( \lambda_v \). At the location of the \( i \)th box, \( \lambda_h \) and \( \lambda_v \) are defined respectively as the horizontal and vertical distances from this box, within which half of the total energy of the \( i \)th resolution kernel associated with a noise-free experiment is confined. Physically, they define the size of the smallest eddy that can be adequately detected by the acoustic system. For perturbations with horizontal and vertical length scales shorter than \( \lambda_h \) and \( \lambda_v \), respectively, the acoustic system can not resolve them unambiguously.

4. Numerical Results

Since the structure of MIZ eddies normally extends to a depth of about 500 m, we have taken \( \delta c = 0 \) and \( \delta v = 0 \) below a depth of 750 m in our simulation study. Further, since the diameters of MIZ eddies are typically 20 to 50 km, it is appropriate to assume the fields \( \delta c(x, z) \) and \( \delta v(x, z) \) to be highly correlated within these horizontal distances. Vertically, such eddies trap cold polar water, producing large negative sound-speed anomalies within a thin surface layer, typically the first 50 to 100 m. Therefore it is also appropriate to assume the perturbation fields in the surface layer to be highly correlated within these vertical distances. The vertical correlation scale also generally increases with depth. To exhibit these scales, we show in Figure 4 the sound speeds across a vertical section at 78.55°N from 2.64°W to 2.58°E, obtained by Johannessen et al. [1985] during MIXEX '84 using standard hydrographic techniques. The survey reveals the existence of two typical MIZ eddies trapping cold water of polar origin in a layer close to the surface.

In order to produce conservative estimates of resolution and accuracy, we must avoid the possibility of over-constraining the eddy fields. Consequently, we have taken the estimates of the horizontal and vertical decorrelation lengths, \( L_h \) and \( L_v \), to be at the lower limits, i.e., 20 km and 50 m, respectively. However, we must keep in mind that the actual eddy fields may be more correlated. We further assume that the eddy fields are statistically homogenous and that the covariance function \( C(\Delta x, \Delta z) \) is Gaussian in shape, such that

\[
C(\Delta x, \Delta z) = \sigma^2 \exp - [(\Delta x/L_h)^2 + (\Delta z/L_v)^2] (12)
\]

where \( \Delta x \) denotes the horizontal distance, \( \Delta z \) denotes the vertical distance between two points in the vertical slice, and \( \sigma^2 \) is the energy or variance of the perturbations. Using (12), the covariance matrix \( C_p \) is specified. For \( \delta c \), \( \sigma \) is of the order of meters per second, while for \( \delta v \), \( \alpha \) is of the order of centimeters per second. The actual value of \( \sigma \) is not important for the determinations of the minimum horizontal and vertical resolution lengths, \( \lambda_h \) and \( \lambda_v \), since these measures are derived in the
Fig. 5. (a) Minimum horizontal resolution lengths (in kilometers) at different locations associated with the 12 non-bottom-interacting ray paths. (b) Minimum vertical resolution lengths (in meters) at different locations associated with the 12 non-bottom-interacting ray paths.
C(Δx, Δz) produces only higher-order effects on error and resolution estimates and also on the inverse solution, as long as the changed function is still monotonically and slowly decreasing and gives high correlations for Δx < L, and little correlation otherwise.

For the time being, let us assume that all the rays that exist in the reference (background) state also exist in the perturbed state. Using the twelve non-bottom-interacting paths, we have computed $F(x, z)$ and $Y(x, z)$, the results are displayed in Figures 5a and 5b, respectively. At each point in the vertical
The average horizontal resolution changes only very slightly. As more and more bottom-reflected paths are incorporated in the observing system, better resolution is generally attained. But the improvement becomes insignificant after a certain number of bottom-reflected paths have been included. The solid lines in Figures 7a and 7b relate the depth- and range-averaged minimum horizontal and vertical resolution lengths to the number of bottom-reflected paths incorporated, showing clearly that the resolution increase saturates beyond a number of about 14. In fact, the local measures \( \mathcal{H}(x, z) \) and \( \mathcal{Y}(x, z) \) change only very slightly beyond the addition of the 14 paths at the different locations. The implication is that the incorporation of bottom-reflected paths beyond this number produces redundant information about the ocean changes.

The dashed lines shown in Figures 7a and 7b are associated with the "missing-path syndrome." By retracing those reference (background) rays through the eddy field shown in Figure 4, we found that while the bottom-reflected paths are very stable, the non-bottom-interacting paths are not. Some of the latter paths become less regular, and some disappear due to the perturbation; stability generally tends to decrease for paths with smaller launch angles. The dashed lines in Figures 7a and 7b, when compared with the solid lines, reveal general increase of the average minimum resolution length resulting from the disappearance of five non-bottom-interacting paths. The average horizontal resolution changes only very slightly. In fact, the local horizontal resolution at different locations, as measured by \( \mathcal{H}(x, z) \), is similar to that associated with the ideal situation of no missing paths. On the other hand, there is a uniform increase of about 15 m in the average minimum vertical resolution length as a function of the number of bottom-reflected paths included in the missing-path scenario. Although a 15-m increase seems small, we must be careful in our interpretation, since the 15 m is an average measure. The fact is, the distribution of the degradation of vertical resolution is uneven in the vertical slice. Below a depth of 200 m, \( \mathcal{Y}(x, z) \) does not increase significantly, but above this depth, an increase of 50 m occurs almost uniformly. This increase is approximately one vertical scale of the perturbations due to the eddies. Therefore drastically degraded vertical resolution can result from the instability of the non-bottom-interacting paths.
The square of the bias in the estimate associated with the use of 25 paths, which includes 13 bottom-interacting paths, is displayed in Figure 8 as a ratio to the variance of the perturbations. Similar to the minimum resolution lengths, the systematic errors in the estimates of both $\delta c$ and $v$ are the same percentagewise. These errors, as is indicated in the figure, are significant. They are everywhere larger than 50% and increase toward greater depths. However, we must keep in mind that

Fig. 8. Ratio of the square of the bias in the inverse solution obtained using 25 ray arrivals, including 13 from the bottom-reflected paths, to the variance of the actual perturbation field.

Fig. 9. Ratio of the variance of the inverse solution to the variance of the actual field for an experimental noise of 2 ms rms using 25 ray arrivals, including 13 from the bottom-reflected paths.
SOUND SPEED PERTURBATION (M/S)

**Fig. 10.** A measured sound speed perturbation field used for our simulation inversions.

they correspond to conservative estimates, determined basically by the decorrelation lengths we specified. In our estimates, we have taken the smallest possible decorrelation lengths. If in fact the eddy field is correlated over longer distance, the bias will be smaller.

The random error in the estimate, when compared with the bias, is much smaller. To show this, let us first consider the noise variance of the data obtained in a reciprocal transmission experiment using standard tomography transceivers, which send bihourly phase-modulated pulse codes that have an equivalent power bandwidth of 100 Hz [Worcester et al., 1985], and assume that the optimal field estimates are derived from daily mean travel times, as is commonly done. Each daily mean travel time is an average of 12 samples with independent noise and therefore possesses a noise variance that is 12 times smaller than that of each individual sample. By accounting for all the error including internal wave-induced fluctuations and measurement error due (mainly) to finite signal bandwidth, Cornuelle et al. [1985] has estimated the noise variance of the daily mean travel times to be 2 ms² for a transmission range of 300 km. With a range half the distance, i.e. 150 km, we consider here, the noise variance should be smaller than 2 ms². Nevertheless, this number represents an upper limit in our case and thus will be used in the following analysis. The daily mean sums of reciprocal travel times, therefore, have a noise variance of $2 + 2 = 4$ ms². As for the differences of reciprocal travel times, Howe [1986] has found that their noise variance is much smaller, only about 10% of that of the sums. The reason is that a lot of the internal wave noise is eliminated by the subtraction. By taking the noise variances to be 4 ms² and 0.4 ms² for the daily mean sums and differences of reciprocal travel times, and the rms values of $\delta c$ and $v$ to be 3 m/s and 10 cm/s, respectively, the signal-to-noise ratios for both data sets become approximately the same. As a result, the ratios of the variances of the estimates of $\delta c$ and $v$ to the variances of the corresponding perturbations are also approximately the same, and are displayed in Figure 9. It is seen that these ratios are small and are less than 0.05 everywhere in the vertical slice.

To further demonstrate the adequacy of the horizontal resolution as a result of the incorporation of the bottom-reflected paths and the overall inadequacy of vertical resolution, we have performed some simulated inversions for $\delta c$. In these inversions, we took the first 50 km of the measured eddy field shown in Figure 4 to be the field in the first 50 km of our simulated ocean. We further took the perturbations to be zero beyond 50-km range and assumed noise-free experiments. The corresponding $\delta c$ in the first 50 km with respect to the reference profile is displayed in Figure 10. Travel time anomalies produced by this 50-km-long eddy field are of the order of 10 ms. A tomographic inversion using only seven non-bottom-interacting paths give the estimate shown in Figure 11. (We could not use all 12 non-bottom-interacting paths in the inversion because five of them had disappeared, as a result of the perturbations.) The figure shows that the regions of colder and warmer water, corresponding to negative and positive perturbations respectively, are only roughly determined in the inverse solution, and the estimated amplitudes are much smaller than the actual perturbed field. A large amount of energy leaks vertically in the estimate. In particular, the energy of the large negative perturbation near the surface leaks downward, thus destroying or weakening the positive perturbations right below. However, it is also fair to say that the estimate is not bad qualitatively, considering that only a minimal number of paths were used.

By adding seven bottom-reflected paths in the inversion, horizontal resolution improves significantly. This can be seen in Figure 12, where the corresponding inverse solution is displayed. The width of the region the colder water occupies, the location of the peak of the positive perturbations, and the amplitudes of the positive perturbations are now more accurately estimated. However, energy still leaks vertically, as is
SOUND SPEED PERTURBATION ESTIMATE (M/S)

seen through the estimated negative surface layer perturbations falsely penetrating to greater depths than were actually reached.

Thus far we have only examined the resolution of acoustic tomography in the absence of support by other types of observation. When other observations are available, whether from remote satellites or moored point sensors, it is only logical that they be combined with the acoustic observations to pro-

Fig. 11. The sound speed perturbations estimated using seven non-bottom-interacting arrivals.

SOUND SPEED PERTURBATION ESTIMATE (M/S)

Fig. 12. The sound speed perturbations estimated using both non-bottom-interacting and bottom-reflected arrivals, seven of each type.
duce the best possible field estimates. The improvement in system resolution due to the incorporation of these other types of observation is the subject we will investigate next.

We first examine the addition of conventional point sensors, such as temperature recorders and current meters, which are usually mounted on the same moorings as the transceivers. These point measurements improve resolution, both horizontally and vertically, in the vicinities of where they are taken. By distributing four temperature sensors evenly over the first 400 m on the acoustic mooring at zero range and combining them with the 14 acoustic paths in the inversion, the solution in the vicinity of this mooring is much improved, as is shown in Figure 13. The thickness of the thin colder water layer and the amplitudes of the perturbations near the point measurements are now quite accurately estimated. Significant improvement of resolution has occurred within a horizontal distance and a vertical distance approximately equal to the horizontal and vertical correlation lengths of the eddy field, respectively, from the location of each point sensor.

We next examine the cooperative performance of acoustic tomography and satellites in space. Polar-orbiting satellites can provide data for sea surface temperature (SST) as well as other dynamical fields at the sea surface as often as every 30 min with an excellent spatial resolution of a few kilometers [Robinson, 1985]. Figures 14a and 14b show the minimum horizontal and vertical resolution lengths, $\lambda$ and $\nu$, respectively, resulting from the combination of the 25 acoustic ray paths used previously and satellite observations of the surface boundary temperature field. The influence of the satellite observation penetrates to a depth of 100 m, generating a horizontally uniform, exceptional high resolution in this layer. In this first 100 m, $\lambda$ and $\nu$ are now 10 to 20 km and 50 to 100 m, respectively, highly compatible with the eddies’ length scales there. This improvement is particularly desirable in monitoring MIZ eddies, since the finer-scale eddy fluctuations are confined to the first 100 m. Below this depth, the influence of the satellite observation diminishes, and acoustic tomography takes over to adequately monitor the internal structures. Figure 15 shows the solution obtained from an inversion of the same 14 ray travel times used before together with satellite measurements of the SST. The thicknesses and widths of the colder and warmer water sections, as well as the magnitudes, are now accurately estimated. The results of a joint inversion of the composite data set consisting of the 14 ray travel times, 4 internal temperature measurements from moored point sensors, and the SST measured by satellite are shown in Figure 16; the result is virtually indistinguishable from the true solution.

5. Conclusions

The applicability of acoustic tomography using point receivers depends critically on the resolvability of ray and/or mode arrivals. In the MIZ for horizontal ranges of the order of 150 km, modes are unresolvable, since the time-bandwidth products are smaller than unity for practical ranges of acoustic frequencies. Ray arrivals, however, are resolvable with a transmission bandwidth larger than 25 Hz, since the multipath arrivals are separated by approximately 40 ms and longer. Thus a bandwidth of 100 Hz as has been proposed for an experiment in the Greenland Sea and MIZ in 1989 should be adequate.

The system performance of acoustic tomography can be...
evaluated using two measures of resolution, namely, the minimum horizontal and vertical resolution lengths $o$ and $\gamma$. In a minimum mean square error estimate, resolution and accuracy are directly related. The incorporation of bottom-reflected arrivals improves horizontal resolution significantly, bringing $o$ well within the horizontal scale of the eddies. Without using such arrivals, $o$ is only marginal. Resolution does not increase without bound, as the number of bottom-

Fig. 14. (a) Same as Figure 6a, except for the inclusion of satellite observation. (b) Same as Figure 6b, except for the inclusion of satellite observation.
reflected arrivals being incorporated increases but saturates at a certain number beyond which little independent information about the ocean is gathered.

The vertical resolution of unaided acoustic tomography in the MIZ is generally not adequate. Unlike horizontal resolution, vertical resolution cannot be improved by the incorporation of bottom-reflected arrivals. The inadequacy is due to the fact that MIZ has an upward refracting sound channel.
that makes most of the rays turn at the same depth, i.e. the surface. Most of the rays thus sample the entire upper column and cannot provide sufficient discrimination of the perturbations in different layers. Moreover, vertical resolution can be degraded significantly by the loss of unstable paths, whereas horizontal resolution is hardly affected.

Significant improvement of system resolution is obtained with the inclusion of moored point measurements and satellite observations. In particular, the combination of acoustic tomography and satellite observation gives a resolution which is quite adequate for monitoring MIZ eddies. Point sensors, which are usually mounted on the same moorings as the acoustic transceivers, and polar-orbiting satellites, which give high resolution data at the sea surface, are used to be highly desirable for MIZ eddy tomography.

Finally, if a full tomographic array is used as opposed to the single slice considered here, resolution improves considerably near the locations where the ray paths cross horizontally. Also, the incorporation of dynamical constraints such as the relation between $d$ and $v$ and space-time correlations of the eddy fields can only further improve resolution and accuracy, providing that the constraints are accurate. Thus the results presented in this preliminary study are associated only with a “worst case”, even so, they are already quite encouraging.

APPENDIX: CONSTRUCTION OF INVERSE SOLUTIONS AND ERROR ESTIMATES USING SINGULAR VALUE DECOMPOSITION

The singular value decomposition of matrices is a powerful tool for determining parameter resolution and information distribution among the observations [Wiggins, 1972]. Its application to linear, minimum mean square error estimates involves the factorization of the transformed (scaled) matrix operator

$$A' = C_p^{-1/2} A C_p^{1/2}$$

into a diagonal matrix $\Lambda$ of eigenvalues and two matrices of eigenvectors, $U$ and $V$, such that

$$A' = U \Lambda V^T$$

The corresponding eigenvalue problem can be cast as

$$AV = U \Lambda$$

and

$$A^T U = V \Lambda$$

where $\Lambda$ contains all the eigenvalues except those with vanishing values. Thus the $i$th diagonal element $\lambda_i$ in $\Lambda$ is the $i$th nonzero eigenvalue and the $i$th columns of $U$ and $V$, denoted respectively by $u_i$ and $v_i$, are the two eigenvectors associated with $\lambda_i$. $\lambda_i^2$ is an eigenvalue of the nonnegative-definite symmetric matrix $(A^T A)$, or equivalently of $A(A^T)^T$. $\lambda_i$ is always positive. The number of nonzero eigenvalues, $k$, which is generally referred to as the rank of $A$, is the number of pieces of independent information supplied by the data and hence is always smaller than or equal to the number of data $m$ itself; i.e., $k \leq m$. The dimensions of $U$, $A$, and $V$ are $m \times k$, $k \times k$, and $n \times k$, respectively.

The eigenvectors $u_i$ and $v_i$ that compose $U$ and $V$ constitute two conjugate sets of orthogonal basis vectors in the transformed data $y'$ and parameter $p'$ spaces, respectively. The transformations are

$$y' = C_p^{-1/2} y$$

$$p' = C_p^{1/2} p$$

Both sets of basis vectors are incomplete; however, they define the regions in the two spaces where information is available. By convention, these basis vectors are normalized to have unit energy, so that

$$U^T U = I \quad V^T V = I$$

where $I$ denotes identity matrices.

Formally, the data vector $y'$ can be expanded as a linear combination of the eigenvectors $u_i$. The weight on $u_i$ in the expansion is given by the dot product of $u_i$ and $y'$, i.e., $u_i^T y'$. The $\lambda_i^2$ values are analogous to the signal-to-noise ratios associated with each of the components in the expansion. For a $\lambda_i^2$ large compared with unity, signal dominates noise in that particular component. But for $\lambda_i^2 \ll 1$, the reverse is true. Similarly, the estimate $\hat{p}$ of $p'$ can also be expanded as a linear combination of the eigenvectors $v_i$. Moreover, the weight of $v_i$ in the estimate can only be determined by the information distributed in $u_i$ in the data space. Using the matrix factorization given in (A2), (6) can be recast as

$$C_p^{-1/2} \hat{p} = V[(I + \Lambda^2)^{-1} A]U^T (C_p^{-1/2} y)$$

or equivalently as

$$\hat{p} = \sum_{i=1}^{k} \frac{\lambda_i}{\lambda_i^2 + 1} (u_i^T y') v_i$$

It is clear in (A8) that the linear, minimum mean square error estimator minimizes the noise effect of the data on the estimate by down-weighting those components that are associated with a high noise level, that is, with $\lambda_i \ll 1$. In general, a small $\lambda_i$ is associated with a $v_i$ that parameterizes a more oscillatory function. Therefore a smoothed version of the exact solution is generally estimated.

Using the matrix decomposition given in (A2), we can further rewrite (7) and (9) as

$$C_x = C_p^{-1/2} [(I - V A (I + \Lambda^2)^{-1} A V^T) C_p^{1/2}$$

$$C_y = C_p^{-1/2} [(I + \Lambda^2)^{-1} A V^T] C_x$$

In the absence of experimental noise, all values of $\lambda_i$ approach infinity, so that $C_y \rightarrow 0$ in (A10), revealing that the estimate has zero variance. In this case, the mean square error is just the square of the bias, $\mathbf{b} \mathbf{b}^T$, which can be evaluated from (A9) by letting all values of $\lambda_i$ go to infinity. The result is

$$bb^T = C_p^{-1/2} (I - V V^T) C_p^{1/2}$$

In the other extreme, that is, in the absence of signal, all values of $\lambda_i$ approach zero, but once again $C_y \rightarrow 0$. This shows the estimator is optimal in that it always tries to minimize the effect of random noise. Naturally, $\hat{p} = 0$ and $C_x = C_y$ in this extreme, implying that no new information about the field is gained from this experiment.

Acknowledgments. The authors would like to thank Robert C. Spindel for his useful comments and discussions. Our thanks also go to the two referees of this article for their careful and thoughtful reviews. Our appreciation to Chris Dunn for generating some of the figures. This work was supported by the Office of Naval Research under contract N00014-86-C-0127. Woods Hole Oceanographic Institution contribution 6253.

REFERENCES


---

C.-S. Chiu, Building 1100, Institute of Naval Oceanography, National Space Technology Laboratories, MS 39529.

O. M. Johannessen, Nansen Remote Sensing Center, Geophysical Institute, University of Bergen, Bergen, Norway.


(Received June 30, 1986; accepted February 3, 1987.)