Vol. 135, No. 1 The American Naturalist January 1990

PASSIVE LOCALIZATION OF CALLING ANIMALS AND SENSING OF THEIR ACOUSTIC ENVIRONMENT USING ACOUSTIC TOMOGRAPHY

JOHN L. SPIESBERGER AND KURT M. FRISTRUP

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Submitted November 21, 1988; Accepted January 5, 1989

Understanding temporal and spatial patterns in the distribution of animals is a fundamental goal of both aquatic and terrestrial ecologists. Marine biologists have used the acoustic signals from calling organisms to estimate the positions of the animals (Walker 1963; Watkins and Schevill 1972; Clark 1980). With one exception (Magyar et al. 1978), terrestrial biologists have not used similar techniques, and the use of acoustic tracking in either environment is rare. Pioneering attempts to use this technique have achieved some success, but the methods are time-consuming and tedious. Even so, these studies have yielded results that are impractical to obtain by other means. This paper discusses optimal techniques that can be used to map automatically the positions of animals and to map the wind and sound-speed fields in the region from the acoustic calls themselves.

The simultaneous mapping of the location of animals and the surrounding fields of sound speed (temperature) and winds would enhance many investigations important to naturalists.

1. Studies of the social behavior of animals would be enhanced. Investigations of vocalizations and communication, mating behavior, and home ranges would be accomplished more easily. Since many organisms could be studied simultaneously, a large sample of a population could be investigated. A map of the locations of most calling animals in regions 1 km square or larger are probably feasible in a forest. Such maps could be updated often enough to permit studies of the temporal evolution of the locations.
2. The relation between the winds, sound speeds (temperature), other acoustic effects, and the behavior of animals may be investigated.
3. Conservation efforts may be enhanced. We provide two examples. First, by obtaining precise information on the location of animals, the investigator may travel to a specific location to observe the natural food sources. Second, acoustic tomography can provide rapid measurements of population density, home-range overlap, and minimum critical size for the design of nature preserves.

It is possible to localize calling animals by measuring the arrival-time differences of the acoustic call at different pairs of acoustic receivers. The arrival-time

Am. Nat. 1990. Vol. 135, pp. 107–153.

differences constrain the position of the animal to some region of space. A different pair of receivers constrains the position to another spatial region. The intersection of the possible regions from many different pairs of receivers determines the location of the animal. It is important to estimate the arrival-time differences as accurately as possible and to account for fluctuations in the acoustic-propagation speed caused by fluctuations in the winds (currents) and the speed of sound. It is also important to account for the error of the receiver positions. The techniques used to map the locations of the animals and the acoustic environment are taken from the fields of acoustics, signal processing, and geophysical inverse theory. The automatic algorithms require the use of computers, which are now inexpensive and fast enough for the task.

In the first section, we discuss the relevant topics associated with acoustic propagation in the atmosphere and in the ocean. The following section concerns the signal-processing requirements and the acoustic characteristics of calling animals. The manner in which the localization can proceed if the fluctuations in the winds (currents), the sound speeds, and the errors of the receiver positions are ignored is discussed in the third section. Next, we discuss geophysical inverse theory, which is used to make optimal localizations of the animals and to map the wind and sound-speed fluctuations using tomographic (inverse) techniques. Examples of tomographic and non-tomographic localization are presented.

SOUND TRANSMISSION IN THE ATMOSPHERE AND THE OCEAN

This is a review of some processes in the atmosphere and the ocean that influence the transmission of sound. The discussion is confined to the principles of sound transmission and to some equations for quantitatively estimating acoustic propagation.

The Speed of Sound

The speed of sound is related to the compressibility of the medium by the wave equation (Rayleigh 1945; Lighthill 1978, pp. 1–7). Because speed increases as compressibility decreases, sound travels faster in water than in air. The relation between speed, c, and compressibility, $d\rho/dp$, is

$$c^2 = (d\rho/dp)^{-1}, \tag{1}$$

where the differential change in the density is $d\rho$ and the differential change in the pressure is dp. Since compressibility does not depend on acoustic frequency for the cases of interest here, the speed of sound is the same for different frequencies.

The speed in air.—The atmosphere can be taken as a perfect gas. The speed of sound in the atmosphere can be derived from the perfect-gas law and from considerations of the change in its energy when work by compression is done by the pressure waves associated with sound propagation. In the atmosphere,

$$c \simeq (\gamma R')^{1/2}\Theta^{1/2}(1 + 0.25 w_r)\ \text{(cm/s)} = 20.05\Theta^{1/2}(1 + 0.25 w_r)\ \text{(m/s)}. \tag{2}$$

The expression for the speed of sound in dry air was given by Rayleigh (1945) and is modified to include the considerations discussed by Rogers (1976, pp. 1–

25), which deal with moist air. γ, which is about 1.40, is a dimensionless number that involves the heat capacity of dry air. R' is the individual gas constant for dry air:

$$R' = 2.87 \times 10^6 \quad (\text{ergs g}^{-1}\,^\circ\text{K}^{-1}).$$

The temperature, Θ, is expressed in degrees Kelvin and is equal to the centigrade temperature plus 273°C. The ratio of the mass of water vapor to the mass of dry air per unit of volume is denoted by w_r. It is related to the relative humidity, f_{hum}, which is expressed in percent by

$$w_r = \frac{0.622}{p}\left(\frac{f_{hum}}{100}\right)e_{s0}\exp\left[\frac{L}{R_v}\left(\frac{1}{\Theta_0} - \frac{1}{\Theta}\right)\right], \tag{3}$$

where $p \simeq 760$ mbar (1 atmosphere pressure), $e_{s0} = 6.11$ mbar (saturation vapor pressure at Θ_0), $L \simeq 2500$ joules/g (latent heat of vaporization), $R_v = 0.461$ joules $^\circ\text{K}^{-1}\text{mole}^{-1}$ (individual gas constant for water vapor), and $\Theta_0 = 273^\circ$K.

At 20°C, the speed of sound in dry air is about 343.2 m/s at sea level (Mach 1). Changes in the sound speed, δc, are related to changes in the temperature and the moisture content by

$$\delta c(\Theta, w_r) \simeq \frac{10.03}{\Theta_{ref}^{1/2}}\delta\Theta + 5.012\,\Theta_{ref}^{1/2}\delta w_r, \tag{4}$$

where the temperature of the air before any fluctuations is Θ_{ref} (in °K). At 20°C, the speed increases by about 0.58 m/s for a temperature increase of 1°C, and it increases by about 0.9 m/s from dry air ($f_{hum} = 0$) to completely saturated air ($f_{hum} = 100$). For many cases of interest, the effect of the humidity variation is small compared with the effect of the temperature variation; however, humidity variations in jungles are often important. The speed of sound also increases with increasing pressure because of the attendant increase in temperature. If the pressure on a parcel of air is increased adiabatically (the entropy of the parcel is constant), then its temperature increases by about 9.8°C per kilometer of decreased elevation. This is the adiabatic lapse rate (Rogers 1976).

Vertical profiles of the temperature and the sound speed near the ground typically exhibit a significant variation between day and night (fig. 1). During the day, the ground is heated by the sun and the air above it is virtually transparent. The ground thus warms the air immediately above, and the temperature and the sound speed decrease with height. At night, the ground emits thermal radiation into the air more quickly than the cooler air can radiate. Since the ground quickly becomes colder than the air, the temperature and the sound speed increase with height. The vertical profiles of temperature have different characteristics beneath the canopies in forests. Typically, the temperature is a few degrees warmer in the canopy than near the ground in the daytime (Evans 1939; Eyring 1946). At night, the temperature profile below the canopy becomes more uniform.

The speed in water.—Sound travels about five times as fast in water as in air (equivalent to about Mach 5 in air or about 1500 m/s). It is a complicated function of the temperature, pressure, and the salinity (salt concentration). The most

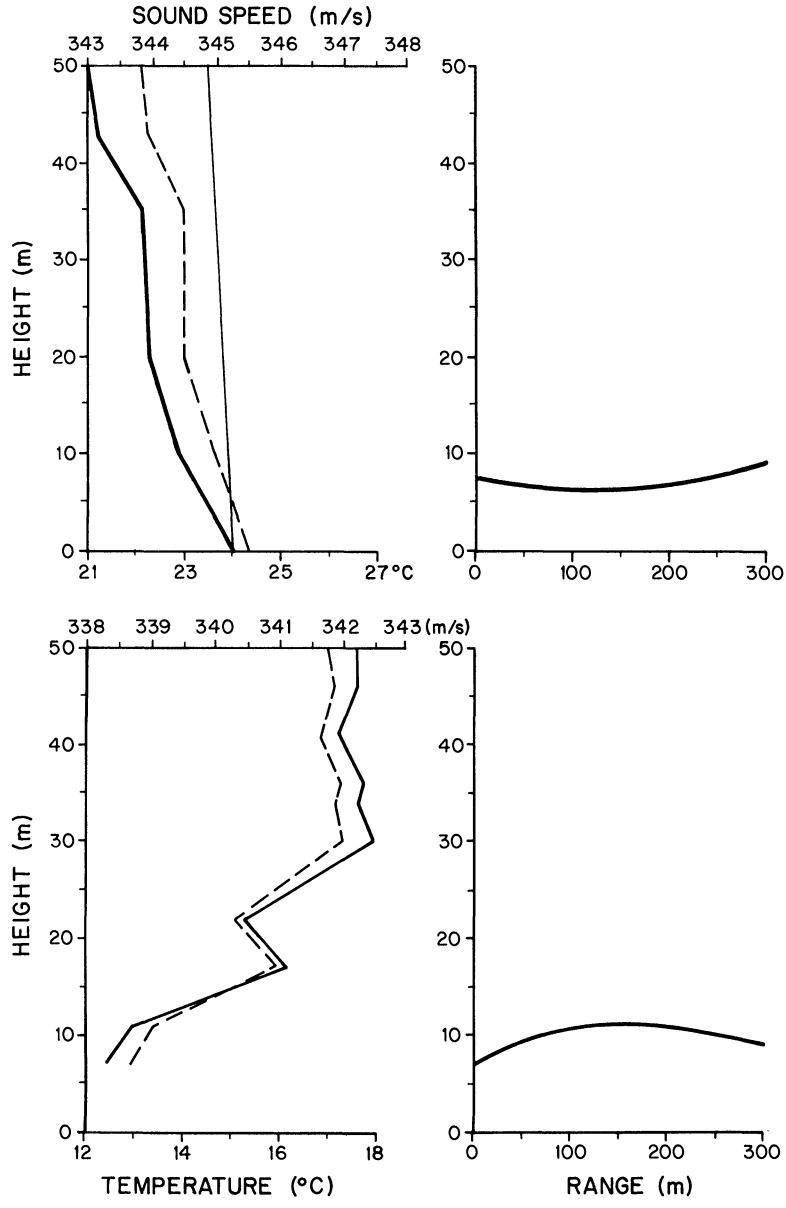

FIG. 1.—Profiles of the temperature (*heavy solid line*) and the sound speed (*dashed line*) during the day (*top left*) and the night (*bottom left*). The daytime profile is taken at about 10:00 A.M. over a prairie grassland (Hall et al. 1975). *Light solid line*, The slope of the temperature if the conditions were adiabatic. The nighttime profile is taken at about 5:20 A.M. (Emmanuel 1972). The right-hand panels show the ray paths that sound traverses if the source and the receiver are at elevations of 7 m and 9 m, respectively, and their horizontal separation is 300 m.

TABLE 1

SIMPLE ALGORITHMS FOR THE SPEED OF SOUND IN WATER (m/s)

Algorithm	Validity	Source
$c = 1449.2 + 4.6\Theta - 0.055\Theta^2 + 0.00029\Theta^3 + (1.34 - 0.010\Theta)(S - 35) + 0.016\mathscr{Z}$	$0 \le \Theta \le 35$ $0 \le S \le 45$ $0 \le \mathscr{Z} \le 1000$	Medwin 1975
$c = 1448.96 + 4.591\Theta - 5.304 \times 10^{-2}\Theta^2 + 2.374 \times 10^{-4}\Theta^3 + 1.340(S - 35) + 1.630 \times 10^{-2}\mathscr{Z} + 1.675 \times 10^{-7}\mathscr{Z}^2 - 1.025 \times 10^{-2}\Theta(S - 35) - 7.139 \times 10^{-13}\Theta\mathscr{Z}^3$	$-2 \le \Theta \le 30$ $30 \le S \le 40$ $0 \le \mathscr{Z} \le 8000$	Mackenzie 1981

NOTE.—Θ, Temperature in degrees centigrade; S, salinity in parts per thousand; $\mathscr{Z}$, depth in meters. Medwin's algorithm differs from Del Grosso's (1974) by amounts ranging between about 0.1 and 0.6 m/s. Mackenzie's algorithm has a standard error of about 0.07 m/s.

accurate algorithms for computing the sound speed in terms of these variables are complicated (Del Grosso 1974; Chen and Millero 1977), but they should be used for accurate estimates. Medwin (1975) and Mackenzie (1981) formulated simpler algorithms that are less accurate (table 1).

As table 1 indicates, sound speed is an increasing function of temperature, pressure, and salinity. Change in sound speed is approximately related to changes in temperature, pressure, and salinity:

$$\delta c(\Theta, p, S)/c \simeq 3.19 \times 10^{-3}\delta\Theta + 1.12 \times 10^{-7}\delta p + 0.96 \times 10^{-3}\delta S, \quad (5)$$

where the temperature is expressed in degrees Kelvin, the pressure in millibars, and the salinity (S) in grams of salt per kilogram of water (parts per thousand; Munk 1974). The speed increases by about +4.78 m/s per degree of increased temperature, by +0.000168 m/s per mbar of increased pressure (1 cm of water is about 1 mbar), and by +1.44 m/s per part-per-thousand increase in the salinity. In most cases, the effects of the pressure and salinity variations are small compared with the effect of the temperature variation.

A typical profile of sound speed at mid-latitudes exhibits a minimum near a depth of about 1 km (fig. 2). The minimum in the speed of sound near a depth of 1 km is typical in the tropics and the subtropics. The speed of sound increases near the surface because the water is warmer there, and the speed increases in the deep ocean because the pressure increases and the temperature stays about the same.

Ray Theory and Travel Time

A solution of the acoustic-wave equation provides a complete solution to the propagation of sound in the air or the water (Rayleigh 1945; Lighthill 1978). Its solution depends on the sound-speed and wind (current) fields as well as a specification of the conditions at the boundaries (e.g., the ground, trees, leaves, rocks, the sea surface or the bottom, etc.). In general, this equation is not a feasible method of solution because of enormous computational problems. There are several approximations to this equation whose solutions can be computed for most cases of interest. The simplest one is called the ray equation or the Eikonal equation (Pierce 1981). In this approximation, the sound travels along ray paths

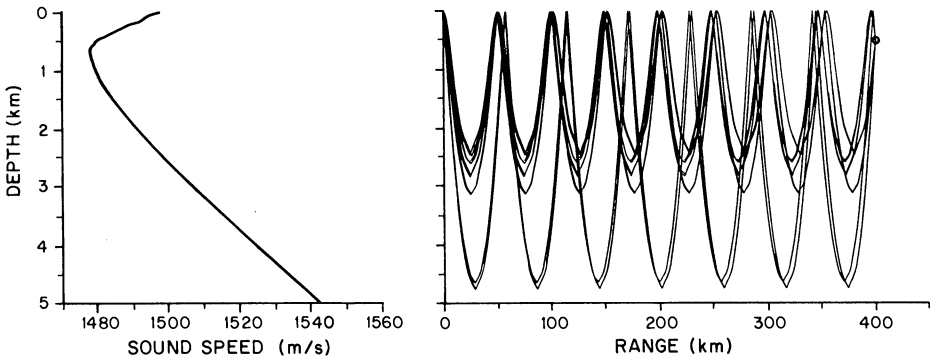

FIG. 2.—*Left*, A climatological profile of sound speed in the Northeast Pacific near 31° N and 157° W. *Right*, The ray paths between a finback whale at 35 m depth at a range of 0 km and an acoustic receiver at a depth of 500 m and a range of 400 km. The paths all reflect from the surface. The travel times of the first and last ray arrivals are 269.132 s and 270.211 s, respectively. The ray paths do not portray the locations in the ocean in which the sound travels with much accuracy because ray theory is not a good approximation in the ocean for frequencies below about 100 Hz. However, the travel times computed from ray theory are probably a good approximation of the actual travel times that would be observed.

that are determined by solving Snell's law (Pierce 1981), which governs how the sound refracts through the medium. For example, the paths that different colors of light take as they are refracted through a prism is predicted from Snell's law. The ray solutions are valid when the acoustic wavelength is small compared with the scales of fluctuation in the medium. Therefore, the ray equation is often referred to as the high-frequency solution to the wave equation. Ray theory can be adapted for the presence of the wind or currents as well as for the sound-speed field. The theory of ray tracing has been discussed by Lighthill (1978) and by Pierce (1981).

When the calls from animals are measured at a microphone, the experimenter does not know the actual speed of sound or the actual wind or current field. However, we can make a good guess for the sound speed or the wind, which we call the *reference field*. The rays are traced through the reference field and a reference travel time is obtained. For simplicity, the reference fields will not include the wind or the current. The reference travel time is

$$T_0 = \int_{\Gamma_0} ds/c_0(\Gamma_0)\,, \tag{6}$$

where the integral is evaluated along the reference ray path, Γ_0, which connects the animal to the receiver, the differential length of the path is ds, and the reference speed along the path is $c_0(\Gamma_0)$. The actual travel time between the animal and the receiver is

$$T_1 = \int_{\Gamma_1} ds/[c_1(\Gamma_1) + \mathbf{u}(\Gamma_1) \circ \hat{\mathbf{s}}]\,, \tag{7}$$

where the actual ray path is Γ_1, the wind vector is $\mathbf{u}$, the unit vector along the ray path is $\hat{\mathbf{s}}$, the circle in the denominator is the dot product, and the actual speed along the path is $c_1(\Gamma_1)$. The actual speed is related to the reference speed at

location **x** by

$$c_1(\mathbf{x}) = c_0(\mathbf{x}) + \delta c(\mathbf{x}) . \tag{8}$$

The ray path is a function of the sound speed and the wind field. It is convenient if the reference and actual ray paths are nearly identical because the rays do not have to be retraced if more information about the sound or wind field is available. In fact, this is often the case. In this paper, acoustic travel times are used to locate the animal as well as to make maps of the sound-speed and wind fields.

The difference between the actual and the reference travel times is about

$$\delta T_1 = T_1 - T_0 = -\int_{\Gamma_0} ds[\delta c(\Gamma_0) + \mathbf{u}(\Gamma_0) \circ \hat{\mathbf{s}}]/c_0^2(\Gamma_0) + \text{nonlinear terms} \tag{9}$$

(Munk and Wunsch 1979). The nonlinear terms vanish when the reference and actual paths are identical (Spiesberger 1985*a*), and in many cases of interest they can be neglected. When this is the case, the change in the travel time is

$$\delta T_2 = -\int_{\Gamma_0} ds[\delta c(\Gamma_0) + \mathbf{u}(\Gamma_0) \circ \hat{\mathbf{s}}]/c_0^2(\Gamma_0) . \tag{10}$$

This approximation has the mathematically convenient property that the medium fluctuations are linearly related to the travel-time change.

A straight line is a good approximation to the ray path between the source and the receiver when localizations and maps of the sound-speed, wind, and current fields are not significantly distorted by this approximation. We derive an approximate expression for the travel-time error that arises from assuming that the path is straight rather than curved. Since we cannot find a simple and approximate expression in the literature in terms of physically meaningful variables such as the range or the sound-speed gradient, a derivation is presented here. Consider a source at (x_1, z_1) and a receiver at (x_2, z_2), where the horizontal and vertical coordinates are x and z, respectively. For simplicity, let the receiver be higher than the source. The distance between the source and the receiver is denoted by R, and the angle of inclination of the line from the source to the receiver is denoted by ψ_0 and is measured positive up from the horizontal. Let the speed of sound vary linearly with height from a value of c_1 at the source to c_2 at the receiver. The actual ray path between the source and the receiver is an arc of a circle whose radius of curvature is

$$R_{\text{rad}} = c_1/c_z \cos\psi_1 , \tag{11}$$

where the vertical derivative of the sound speed is c_z and the ray angle with respect to the horizontal at the source is ψ_1 (Appendix A). The ray angle at the source is given by

$$\psi_1 = \psi_0 \pm \tan^{-1}\left(\frac{\epsilon \cot\psi_0}{2 + \epsilon}\right), \tag{12}$$

where the plus sign is used when the speed is greater at the receiver than at the source and the minus sign is used otherwise. The parameter

$$\epsilon = c_2/c_1 - 1 \tag{13}$$

TABLE 2

ESTIMATES OF THE ACOUSTIC-TRAVEL-TIME CHANGES CAUSED BY WIND (OR CURRENT) AND TEMPERATURE FLUCTUATIONS ENCOUNTERED IN AIR AND WATER

Variable	Air	Water
c_0	343 m/s	1500 m/s
R	300 m	300 m
Wind/current		
u	5 m/s	1 m/s
δT	−0.013 s	−0.00013 s
Temperature		
$\delta\Theta$	10°C	1°C
δT	−0.0150 s	−0.00065 s

NOTE.—Equation (15) is used to compute the estimates, and the travel-time fluctuation in air is computed for a starting temperature of 20°C. The wind represents a steady-state condition or a gust with a horizontal scale exceeding 300 m. The atmospheric temperature fluctuation is typical of the difference between night and day. c_0, Reference speed of sound; R, the distance between the source and the receiver; u, the magnitude of the wind or current; $\delta\Theta$, the temperature change; and δT, the travel-time change.

is introduced to express the small change in the speed between the two locations. The difference between the actual travel time and the travel time along the straight line is

$$\delta T_1 \simeq -c_z^2 R^3/24\, c_1^3 \tag{14}$$

(Appendix A). For the daytime conditions measured by Hall and for the acoustic-ray path considered for the daytime condition (fig. 1), $R = 300$ m, $c_1 = 345.08$ m/s at height of 7 m, $c_2 = 344.96$ m/s at height of 9 m, and the error given by expression (14) is $\delta T_1 \simeq -0.0001$ s.

Even at a large distance, the error of the straight-line approximation is small compared with the actual travel time of about 300 m/(345 m/s) = 0.87 s.

If the acoustic travel times are used to construct maps of the sound-speed or the wind fields, then the error of the travel-time measurement should not be much larger than the travel-time perturbation from the media fluctuations. It is of interest to estimate the travel-time perturbations caused by the media. If the temperature or wind fluctuation, u, has the same value along the straight ray path, then the travel-time perturbation given by equation (10) simplifies to

$$\delta T_2 \simeq -Ru/c_0^2 \quad \text{or} \quad -R\delta c/c_0^2 . \tag{15}$$

The travel-time perturbations are typically larger in the air than in the water because the Mach number, u/c, or $\delta c/c$ is larger in the air than in the water (table 2). Because the travel-time changes are smaller in water, we expect to have more difficulty measuring the currents and sound speed in water than the wind and sound speed in air.

When the temperature decreases with height, the ray paths are refracted up-

ward; thus, if a distant receiver is not high enough, the sound may not be heard. Regions in which the sound is refracted away from the receiver are called *shadow zones*. The sound intensity in these regions fades considerably, and its effects can be computed using extensions to ray theory or more-complete solutions of the wave equation (Pierce 1981). Shadow zones occur in the atmosphere and in the ocean. Shadow zones in air are discussed by Wiley and Richards (1982, p. 157). If the temperature decreases with height in a linear manner, and if the heights of two receivers above the ground are z_1 and z_2, then the distance to the shadow zone is

$$R_s = 2(\Theta_1/|\Theta_z|)^{1/2}\,(z_1^{1/2} + z_2^{1/2})\,,$$

where Θ_1 is the temperature at receiver one and the temperature gradient is Θ_z. The relation is derived by finding the ray that grazes the ground, where the ray's radius of curvature is given by equation (11) and the speed of sound is related to the temperature by equation (2). For the daytime profile of temperature measured by Hall (fig. 1) and for $z_1 = 2$ m, $z_2 = 2$ m, and $\Theta_1 = 23.78$°C, the distance to the shadow zone is $R_s = 83$ m.

In practice, turbulent fluctuations in the daytime atmosphere can scatter significant amounts of sound energy into the shadow zone (D. Thomson, pers. comm.). Shadow zones are more common at night when the turbulent fluctuations decrease. At night, the ray paths typically curve toward the surface, and shadow zones are more commonly observed for receivers mounted at high-enough positions (Thomson, pers. comm.).

For long-range propagation in the ocean, the ray paths are refracted by the sound channel. The typical cycle distance of the paths is about 50 km (fig. 2). Although these paths appear to deviate significantly from a straight line, they are actually rather flat, having an inclination with respect to the horizontal typically 12° or less.

Absorption in Air and Water

Sound energy is converted into heat by the inelastic response of the molecules to the pressure waves associated with the sound propagation. In air, the absorption is a function of temperature, humidity, and the acoustic frequency. A tabulation of the absorption for these variables is given elsewhere (American National Standards Institute 1978). In general, the absorption increases as the frequency increases. However, given a frequency, the absorption reaches a maximum value for some value of the humidity.

Attenuation coefficients are often expressed in units of decibels (dB) per unit of distance. A decibel is a measure of the relative sizes of two things. If I_1 and I_2 are two intensities (amplitude squared), then I_1 is $10\log_{10}(I_1/I_2)$ dB from I_2. A summary of some attenuation coefficients in air is shown in table 3.

Water is almost transparent to low-frequency sound. Urick (1983, pp. 102–110) provided some attenuation coefficients as a function of frequency in the ocean (table 4). Thorp (1965) provided a value of the absorption coefficient in seawater given by

$$\text{dB attenuation} \simeq 1.87 \times 10^{-7} f^{1.5} \quad \text{(dB/100 m)}\,, \tag{16}$$

TABLE 3

THE ABSORPTION COEFFICIENTS FOR AIR (dB/100 m) AS A FUNCTION OF THE ACOUSTIC FREQUENCY AND THE RELATIVE HUMIDITY FOR AMBIENT TEMPERATURES OF 5°C AND 20°C (FROM HARRIS 1967)

	RELATIVE HUMIDITY, %				
f (Hz)	10	30	50	70	90
		AMBIENT TEMPERATURE, 5°C			
250	0.225	0.096	0.079	0.070	0.063
500	0.719	0.235	0.189	0.166	0.151
1000	2.06	0.734	0.466	0.404	0.369
2000	3.67	2.52	1.47	1.06	0.912
4000	4.71	8.00	4.93	3.47	2.70
8000	6.16	17.0	15.9	11.6	9.06
12,500	9.09	23.1	28.1	24.5	19.3
		AMBIENT TEMPERATURE, 20°C			
250	0.133	0.088	0.073	0.064	0.059
500	0.354	0.211	0.175	0.155	0.141
1000	1.16	0.513	0.422	0.376	0.343
2000	3.90	1.29	1.04	0.924	0.843
4000	11.0	4.12	2.65	2.31	2.14
8000	20.2	13.8	8.32	6.22	5.47
12,500	25.4	29.4	17.9	13.1	10.8

TABLE 4

APPROXIMATE VALUES OF THE ABSORPTION COEFFICIENT IN SEAWATER

f (Hz)	dB/100 m	f (Hz)	dB/100 m
125	0.00026	2000	0.013
250	0.00074	4000	0.021
500	0.0021	8000	0.044
1000	0.0059		

NOTE.—The values from 125 to 1000 Hz are evaluated from equation (16). For higher frequencies, the values are taken from some curves given by Urick 1983.

for frequencies between 112 and 1780 Hz where the frequency, f, is evaluated in Hertz. The attenuation in distilled water is less than in seawater by a substantial amount (Urick 1983).

Scattering Due to Turbulence in Air and Water

Both the atmosphere and the ocean contain inhomogeneities that are impractical to predict in a deterministic manner and that have a significant effect on the transmission of sound. A useful method of characterizing the resulting acoustic variability can be formulated by pretending that the variability in the medium is random, in that it can be described using some probability density function or that it has certain scales. Scattering theory relates the random fluctuations in the medium to the subsequent random fluctuations of the sound after it passes through the medium. It is important to understand something about the scattering process

because it imposes limits on the performance of a localization system. The theories that follow should only be used to indicate the character and approximate values of the acoustic fluctuations from turbulence.

When a steady wind blows above the ground, a region containing turbulent fluctuations or eddies lies between the ground and the steady wind in the essentially friction-free atmosphere. This region is referred to as the planetary boundary layer (PBL). If the steady-wind speed is slow enough, then the wind speed in the boundary layer increases linearly with height to the steady value. However, for all practical cases, the wind speed is sufficiently fast that this region becomes turbulent. Thus, a statistical description of its properties is appropriate. The height of the PBL depends on whether the air is dynamically stable, neutral, or unstable (Panofsky and Dutton 1984, p. 108). If a parcel of air is displaced vertically, and it is then released (with zero speed), oscillating about its original position, then the conditions are stable. If it remains at the displaced position, then the conditions are neutral. If it continues to move away from its original position after it is displaced, then the conditions are unstable. Neutral conditions are usual. The conditions of stability can be understood with the notion that the density of air decreases as its temperature increases. This follows from the ideal-gas law, $p/\rho = R'\Theta$, where the density of air is ρ. In typical daytime conditions, the ground is warmed by the sun, the temperature is warmer near the surface than above, and less-dense air is formed below the denser air above. This is an unstable situation, and thermal plumes rise from the ground. In typical nighttime conditions, the ground cools quickly, and the temperature near the surface is cooler than above. The air near the surface is denser than that above, and stable conditions prevail. We are primarily concerned with about the lower 10% of the PBL, which is called the surface layer. During still nights, the surface layer may only be about 2 m high and the PBL only about 20 m thick. On windy days, the surface layer may be 100–200 m high and the PBL about 2000 m thick (Stull 1988).

For neutral conditions, the thickness of the PBL is approximately given by

$$H_{\mathrm{PBL}} \approx 0.2u^*/f_{\mathrm{cor}}, \tag{17}$$

where f_{cor} is the Coriolis frequency, which is

$$f_{\mathrm{cor}} = 2\Omega \sin\theta_{\mathrm{lat}} \quad (\text{radians s}^{-1}), \tag{18}$$

where Ω is the earth's rate of rotation (7.29×10^{-5} s^{-1}), the latitude is θ_{lat}, and the friction velocity, u^*, is obtained through

$$U(z) = u^* \ln(z/z_0)/\kappa, \tag{19}$$

and κ is Von Karman's constant and equals about 0.4 (Haugen 1973). The wind field is decomposed into an average or mean portion, $U(z)$, which depends only on height, and a random portion. The mean wind speed in the surface layer increases logarithmically with height above the roughness length, z_0. The roughness length depends on whether the ground is grassy, smooth, or covered with snow, etc. It is about 10^{-4} m for water or smooth ice, 10^{-2} m for mown grass, and about 0.2 m for pastures. (These values for z_0 and those for other surfaces are found in Panofsky and Dutton 1984, tables 6.1, 6.2.) Weather reports usually give the wind speed at a

height of 10 m. Thus, if $U(z = 10\text{ m})$ is known, equation (19) can be solved for u^*, and the PBL thickness is given by

$$H_{\text{PBL}} \approx 0.2\kappa U(10)/f_{\text{cor}} \ln(10/z_0)\,, \tag{20}$$

where z_0 is expressed in meters. The thickness of the PBL for other conditions near the surface can be estimated from other considerations (Panofsky and Dutton 1984).

In the surface layer, the largest scales of turbulence (or, equivalently, the largest eddy) is approximately given by

$$L_o \approx 0.4z\,, \qquad z \leq 0.2H_{\text{PBL}}\,, \tag{21}$$

where L_o is the *outer scale*. The smallest scales of turbulence are imposed by dissipation processes that act on scales on the order of 1 mm or less. This is the *inner scale*, or *Kolmogorov microscale*. The scales between the outer and inner scale are called the *inertial subrange*.

In the surface layer, most of the energy is in the mean wind rather than in the turbulent field. The standard deviation of the turbulent wind in the direction of the mean wind for neutral or unstable conditions is about

$$\sigma_u \simeq 2.4u^*\,. \tag{22}$$

For stable conditions, this value can be much larger because of the presence of waves (Panofsky and Dutton 1984). Equation (19) is solved for u^*, and along with expression (22), we get

$$\sigma_u \simeq U(10)/\ln(10/z_0)\,, \tag{23}$$

where the mean wind at 10 m is used to evaluate u^*.

If the turbulence is weak enough, then the travel time along the ray path fluctuates by a small amount and the amplitude of the ray remains about constant. In geometrical optics, this is called *weak scattering*. The requirement for weak scattering is approximately given by

$$\frac{0.53\ln(10/z_0)c^{19/12}z^{1/3}}{U(10)f^{7/12}R^{11/12}} > 1\,, \tag{24}$$

where the acoustic frequency is f (Hz), and the distance is R (Appendix B). The standard deviation of the acoustic travel time at the receiver is about

$$\sigma_t \simeq \frac{0.63\,U(10)(Rz)^{1/2}}{\ln(10/z_0)c^2} \tag{25}$$

(Appendix B).

If the turbulence is great enough, the single ray path is broken up into randomly fluctuating neighboring ray paths, which destructively and constructively interfere with each other at the receiver. The amplitude of the signal undergoes deep fades and the phase fluctuates from 0° to 360°. This is called full saturation, and the conditions under which it is achieved for straight ray paths are

$$0.66\left(\frac{R}{z}\right)^{3/2}\frac{U(10)}{c\ln(10/z_0)} > 1 \tag{26}$$

and

$$\frac{4.0f(Rz)^{1/2}U(10)}{c^2\ln(10/z_0)} > 1 \tag{27}$$

(Appendix B). When full saturation is achieved, the standard deviation of the acoustic travel time increases and the precision of the measurements of the acoustic travel time is decreased. The statistical description of the acoustic signal in this regime is complicated, and the reader is referred elsewhere for details (Flatte 1979). Consider $z_0 = 0.20$ for pastureland, $U(10) = 1$ m/s, $c = 340$ m/s, $z = 2$ m, $f = 1000$ Hz, and $R = 30$ m. For this case, the left-hand side of inequality (24) equals 36; thus, the weak-scattering approximation is valid, and equation (25) yields a value of about 11 μs. If the mean wind speed at 10 m increases to 10 m/s and if the range increases to about 100 m, then the weak-scattering approximations are questionable.

Below the forest canopy, the logarithmic law for the mean wind as a function of height is not valid because the flow depends on eddies that form as the wind propagates around the trees and the other obstacles. The above considerations for the scattering are therefore invalid.

At long ranges in the ocean (on the order of 50 km) and at frequencies above about 100 Hz, the internal wave field is responsible for most of the short-term acoustic fluctuations between a stationary source and a stationary receiver (Flatte 1979). At short ranges ($\simeq$ 1 km) and at high frequencies ($\simeq$ 1–100 kHz), the acoustic fluctuations between stationary sources and receivers are principally due to turbulent scales on the order of 1 m, and the scattering is weak (Duda et al. 1988).

Scattering and Absorption from Obstacles in Air

Scattering and absorption from tallgrasses, trees, shrubs, leaves, other obstacles, and the ground lead to a significant loss of energy at a distant receiver. The total loss caused by these processes and by the absorption in air alone has been measured in tropical forests by Eyring (1946), Marten and Marler (1977), Marten et al. (1977), and Waser and Waser (1977). Morton (1970) made similar measurements in Central America. Wiley and Richards (1982, pp. 138–140) reviewed the losses from different types of terrain, and ground absorption has been studied by Attenborough (1985). The frequency-dependent loss from scattering and absorption from obstacles can be estimated by subtracting the appropriate values of the loss from absorption in air (table 3; American National Standards Institute 1978).

The absorption from obstacles usually increases with frequency, and its effect often dominates that resulting from absorption in air. For example, in the Kibale Forest in Uganda, the absorption from obstacles alone at 1000 Hz is about 3 dB per 100 m and at 4000 Hz is about 7 dB per 100 m (table 3; Waser and Waser 1977).

Multipaths

In the absence of boundaries, it is possible to emit a single pulse yet receive many pulses because of the differential refraction of sound for paths that leave the source at different angles (multipaths). In the absence of turbulence, the paths can

often be determined by ray tracing. The presence of multipaths depends on the details of the sound-speed, wind, and current fields. If the speed of propagation is a linear function of height, there are no multipaths because the travel time monotonically increases for steeper ray angles at the source (Spiesberger 1985*a*).

Near the ground, the surface or bottom of the sea, or other obstacles, multipaths may result from reflections at interfaces or from partial transmission of the sound through the boundary, with a subsequent bending or reflecting of the sound back into the air or water.

MEASUREMENT OF ARRIVAL-TIME DIFFERENCES

In general, we do not know the instant at which an animal starts its call. Therefore, we cannot directly measure acoustic travel times and perform tomographic inversions as is done in the ocean (Munk and Wunsch 1979). However, a measurement of the arrival-time difference at two receivers is independent of the time of the origin of the call. We therefore discuss an optimal method for measuring arrival-time differences. A visual inspection of the received records may be adequate if the calls are sharply peaked and if they have large signal-to-noise ratios. However, visual inspection is a tedious process, and the technique does not work when the signals are below the background-noise level. There are signal-processing techniques that estimate the arrival-time difference in an optimal manner and that can be implemented automatically. The technique discussed here is widely used and is referred to as *cross-correlation* (Helstrom 1975). The technique is effective when the animal call at two different receivers is the same except for a travel-time delay and for a difference in the amplitude.

Description of the Cross-Correlator

The properties of the cross-correlator are as follows. Let the digitally sampled records at receivers 1 and 2 be

$$\phi_1[\mathrm{k}] = a_1 \nu[\mathrm{k}] + e_1[\mathrm{k}], \tag{28}$$

$$\phi_2[\mathrm{k}] = a_2 \nu[\mathrm{k} + \mathrm{p}] + e_2[\mathrm{k}], \tag{29}$$

where [k] denotes the digital sample number k and the total number of samples is K. The signal at receiver 2 arrives p samples before it arrives at receiver 1. The amplitudes of the animal call at each microphone are a_1 and a_2, and the animal call is $\nu[\mathrm{k}]$. The noises at the microphones are e_1 and e_2. The average signal-to-noise ratios at the two microphones are defined by

$$d_i^2 = a_i^2 \frac{1}{\mathrm{K}} \sum_{\mathrm{k}=1}^{\mathrm{K}} \nu^2[\mathrm{k}] \Big/ \sigma_i^2, \tag{30}$$

where $i = 1, 2$. The noises at the microphones have zero mean and have variances given by

$$\mathrm{var}\{e_i\} = \sigma_i^2, \tag{31}$$

where $i = 1, 2$. For simplicity, we assume that the noises at the two receivers are uncorrelated and that the noises have a white spectrum (Helstrom 1975). The

cross-correlator has an output given by

$$\zeta[j] = \sum_{k=1}^{K} \phi_1[k]\phi_2[k - j], \tag{32}$$

where $j = 0, 1, 2, \ldots, K - 1$. There are always end effects associated with the evaluation of the cross-correlation. The usual solution for evaluating equation (32) is to set unmeasured values of the signals to zero. The signal portion of equation (32) is

$$\zeta_s[j] = a_1 a_2 \sum_{k=1}^{K} \nu[k]\nu[k + p - j], \tag{33}$$

which takes on a maximum value at

$$\zeta_s[p] = a_1 a_2 \sum_{k=1}^{K} \nu^2[k]. \tag{34}$$

When both the signal and the noise are present and the noise samples are uncorrelated, the standard deviation of the noise from the cross-correlator output is

$$\zeta_n = \left(K\sigma_1^2\sigma_2^2 + a_1^2\sigma_2^2 \sum_{k=1}^{K} \nu^2[k] + a_2^2\sigma_1^2 \sum_{k=1}^{K} \nu^2[k]\right)^{1/2}. \tag{35}$$

The expected signal-to-noise ratio of the cross-correlator at the peak value is defined as

$$d^2 = \zeta_s^2[p]/\zeta_n^2, \tag{36}$$

which can be written as

$$d^2 = Kd_1^2d_2^2/(1 + d_1^2 + d_2^2). \tag{37}$$

The expected ratios of signal to noise of the cross-correlator output to the signal-to-noise ratios at either of the microphones are

$$d^2/d_1^2 = Kd_2^2/(1 + d_1^2 + d_2^2) \tag{38}$$

and

$$d^2/d_2^2 = Kd_1^2/(1 + d_1^2 + d_2^2). \tag{39}$$

For most cases of interest, these ratios are much larger than one. The cross-correlation gain cannot be increased arbitrarily by sampling the signal at such a large frequency that K is huge. The cross-correlation gain is determined by the duration of the signal, by the bandwidth of the signal, and by the covariance scale of the noise. Equations (38) and (39) are true only when the noise samples are uncorrelated, as in

$$\overline{e_i[k]e_i[l]} = \begin{cases} \sigma_i^2 & \text{if } k = l \\ 0 & \text{otherwise,} \end{cases}$$

where the overbar denotes the expected value. If the noise samples are correlated, the formula for the cross-correlation gain is more complicated (Appendix C).

The constant K has an important interpretation. Suppose each record is sampled for a period of T seconds throughout which the animal's call is arriving.

Suppose that the acoustic frequencies in the call fall in a frequency range given by W Hz, which is centered about 0 Hz after complex demodulation. Then, the highest frequency is $W/2$ Hz, and it is possible to show that the signal can be adequately reconstructed by taking digital samples at twice the highest frequency, which is $2(W/2) = W$ Hz. These are not obvious statements, and the reader is referred to texts concerning complex demodulation (Helstrom 1975) and aliasing (Oppenheim and Schafer 1975, p. 28) for explanations. Then, the number of digital samples is

$$K = T/(1/W) = TW. \tag{40}$$

The expression in equation (40) is a useful quantity known as the time-bandwidth product. The cross-correlator usually has a larger signal-to-noise ratio than does either receiver because the time-bandwidth product is large (fig. 3).

It can be shown that the expected time at which the cross-correlator reaches its largest signal-to-noise ratio is the arrival-time-difference interval, $k = p$. Of all linear estimators of the travel-time difference, the cross-correlator is the maximum-likelihood estimator for the difference when the noise fields are white-Gaussian processes (Helstrom 1975).

It can also be shown that the width of the cross-correlator peak is about $1/W$ s regardless of the duration of the signal (Helstrom 1975, p. 20). When the bandwidth of the call is large, the cross-correlation peak is narrow (fig. 3).

When the peak of the envelope of the cross-correlator output is used to measure the arrival-time difference, the standard deviation of the estimate is given by

$$\sigma_t \simeq 1/2\pi W_{rms} d \tag{41}$$

(Helstrom 1975, p. 276). Here, W_{rms} is the root mean square (rms) of the bandwidth of the call in cycles per unit of time (Helstrom 1975, p. 18), and d^2 is given by equation (37). However, the arrival difference can be measured with much greater precision when the arrival phase can be determined. Then, the precision of the arrival time is given by

$$\sigma_t \simeq 1/2\pi f_c d, \tag{42}$$

where f_c is the center frequency (Hz) of the call (Helstrom 1975, p. 176). Phase measurements of the travel time are often difficult because it is necessary to know which cycle of the phase is being received.

Practical Implementation of the Cross-Correlator

In practical applications it is useful to compute the circular cross-correlation (Oppenheim and Schafer 1975, p. 105). Suppose the same signal is present in two records, and suppose the records contain no noise. Then, if the signal is K_s digital samples long, the circular cross-correlator output is identical to the cross-correlator output when $K \geq 2K_s$. If two receivers are separated by a distance equal to R, then the maximum travel-time difference between a source and the receiver is R/c seconds, where c is the speed of sound. If the samples are measured at the Nyquist frequency, f_{ny} (Oppenheim and Schafer 1975), then there are only $J_1 = 2f_{ny}(R/c)$ lags that must be computed (instead of K lags). Equation

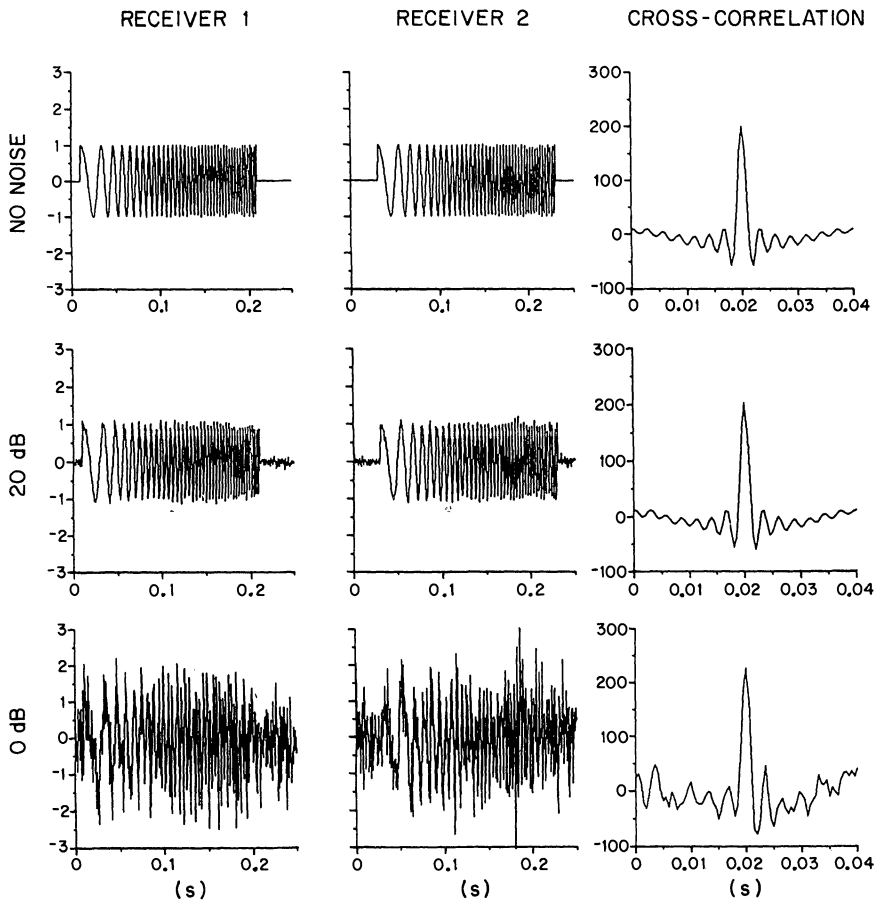

FIG. 3.—Examples of the cross-correlation of a call that lasts 0.20 s and changes frequency from 20 to 400 Hz as a linear function of time. The call would sound like a short upward sweep in pitch. The two left columns show the signal at receivers one and two, respectively, for signal-to-noise ratios of infinity (no noise), 20 dB, and 0 dB. Notice that the signal arrives 0.02 s later at receiver 2 than at receiver 1. The right column displays a portion of the cross-correlation of the two receivers on an expanded scale so that the width of the peak can be examined. In this example, the signal and the noise are digitized at a frequency of 2000 Hz and the noise samples are independent. The time-bandwidth product of the signal is about 0.2 s × 200 Hz = 40, and the root mean square of the bandwidth of the signal is about 200 Hz. The theoretical gain of the cross-correlator is given by equation (38) and is about 11 dB for the bottom row. The arrival-time difference at the two receivers can be measured even when the signal is invisible at the receivers.

(32) requires about $N_{\text{direct}} = 2J_1K$ additions plus multiplications to evaluate directly.

When J_1 and K are large, the circular correlation can be evaluated much more efficiently by applying the convolution theorem (Oppenheim and Schafer 1975, p. 110). The construction of the circular correlation uses fast Fourier transforms (FFT). This technique is most efficient when K equals 2^n, where n is a positive

integer, but it also offers considerable savings when K is composed of many prime factors (Brigham 1974). A wide variety of FFT algorithms exists. The choice of a particular algorithm depends on the value of K and the kind of hardware used for the computation (Burrus and Parks 1985). If all K lags are required, the FFT technique requires about $N_{\mathrm{FFT}} \simeq 15\mathrm{K}(\log_2 \mathrm{K} - 1)$ arithmetic operations. When the number of correlation lags desired is much less than K, the number of arithmetic operations required is less than N_{FFT} (Rader 1970). However, consider an example that compares the direct and FFT methods. For $R = 300$ m, $c = 331$ m/s, $f_{\mathrm{ny}} = 20{,}000$ Hz, and $\mathrm{K} = 5f_{\mathrm{ny}}$ for a 5-s record, the number of arithmetic operations required for the cross-correlator for the direct and FFT methods is

$$N_{\mathrm{direct}} \simeq 7.2 \times 10^9,$$

$$N_{\mathrm{FFT}} \simeq 1.6 \times 10^7.$$

In this example, the direct technique evaluated only $J_1 \simeq 36{,}254$ lags and the FFT evaluated all 100,000 lags. Minicomputers can currently do about 2–5 million arithmetic operations per second. For 2 million operations per second, the direct and FFT approaches require about 1 h and 8 s, respectively. In the next year, minicomputers will achieve speeds of 10–50 million arithmetic operations per second. Thus, a combination of efficient cross-correlation algorithms and fast computers will allow the investigator to work in real time.

Signal-to-Noise Ratio of the Cross-Correlator

The peak signal-to-noise ratio after cross-correlation is approximately given by

$$\mathscr{D} \simeq \mathscr{D}_{\mathrm{sl}} - \mathscr{D}_{\mathrm{gl}} - \alpha R + 10\log(d^2/d_i^2) - \mathscr{D}_{\mathrm{noise}} - \mathscr{D}_{\mathrm{refl}} \quad \text{(dB)}, \tag{43}$$

where $\mathscr{D}_{\mathrm{sl}}$ is the source level of the animal (its loudness); $\mathscr{D}_{\mathrm{gl}}$, the geometric spreading loss of the sound; αR, the absorption and scattering losses; $10\log(d^2/d_i^2)$, the gain from the cross-correlator; $\mathscr{D}_{\mathrm{noise}}$, the background noise; and $\mathscr{D}_{\mathrm{refl}}$, the reflection losses from the ground or water boundaries. The range, R, is the distance to receiver i. The source level is given by

$$\mathscr{D}_{\mathrm{sl}} = 136.0 + 10\log P \quad \text{(dB) in air} \tag{44}$$

$$= 171.5 + 10\log P \quad \text{(dB) in water}, \tag{45}$$

where P is the acoustic power from the animal, in watts, at a distance of one yard. The acoustic power is related to the rms value of the acoustic pressure, p_1, at a distance of one yard from the animal by

$$P = 2.525 \times 10^{-14} p_1^2 \quad \text{(watt) in air} \tag{46}$$

$$= 70.0 \times 10^{-19} p_1^2 \quad \text{(watt) in water}, \tag{47}$$

where the pressure is measured in micro-Pascals (μPa; Urick 1983). The geometric spreading loss can often be estimated from ray theory (Pierce 1981). In the ocean, this loss usually falls somewhere between spherical and cylindrical spreading:

$$\mathscr{D}_{\mathrm{gl}} = 10\log R^2 \quad \text{(spherical)}$$

$$= 10\log R \quad \text{(cylindrical)}$$

TABLE 5

SOURCE LEVELS ($\mathcal{D}_{sl}$), FREQUENCY RANGES, CALL DURATIONS (T), AND TIME-BANDWIDTH PRODUCTS (K) OF SOME TERRESTRIAL AND MARINE ORGANISMS

Animal	Call	$\mathcal{D}_{sl}$ (dB)	Range (Hz)	T (s)	K	Source*
		TERRESTRIAL ORGANISMS				
Cerocebus albigena (grey-cheeked mangabey monkey)	whoop-gobble	115–117	100–400	2	600	1
Colobus guereza (black-and-white colobus monkey)	roaring bout	111–121	100–1000	1	900	1
Ceropithecus mitis (blue monkey)	pyow	119–123	100–2500	0.2	480	1
Ceropithecus ascanius (redtail monkey)	hack	107–115	100–1000	0.2	180	1
Acheta firmus (male field cricket)	calling male	80–110	3600–4400	0.1	80	2
Myotis lucifugus (bat)		58	39,000–78,000	0.0023	90	3
Grallaria perspicillata (speckled antpitta bird)		125–131	1600–1900	4	1200	4
Vireo flavoviridus (yellow-green vireo)		106–112	2200–5000	1	2800	4
Hyla cinerea (green treefrog)	4 0.1-s bursts	126–128	500–7000	0.4	2600	5, 6
Bufo terrestris (southern toad)		116–117.5	2200–2500	1.7	510	5, 6
		MARINE ORGANISMS				
Balaenoptera physalus (finback whale)		160–186	18–23	1	5	7
Leptonychotes weddelli (Weddell seal)	type T	153–193	100–12,800	20	254,000	8
Micropogon undulatus (croaker fish)	10 0.01-s bursts	136	300–600	0.1	30	9, 10

NOTE.—The values of the frequency, duration, and time-bandwidth product are visually estimated from the spectrographs shown in the references or from tables provided by the authors; they are probably accurate to a factor of two.

* 1, Waser and Waser 1977; 2, Alexander 1960; 3, Griffin 1958; 4, Morton 1970; 5, Gerhardt 1975; 6, Bogert 1960; 7, Watkins et al. 1987; 8, Thomas and Kuechle 1982; 9, Moulton 1963; 10, Fish and Mowbray 1970.

(Spiesberger, pers. obs.). For short distances, away from boundaries and before refractive effects become important, the spherical law is valid. The geometric spreading loss for an acoustic source near the ground or near an ocean boundary is more complicated to compute because the impedance of the boundary requires consideration. The absorption term is due to attenuation in the media or due to excess loss as a result of scattering and of absorption by obstacles along the acoustic path. The gain from cross-correlation is given by equations (38) and (39). The noise level per Hz is given by

$$\mathcal{N} = 10\log(p^2_{\text{noise}}/1\ \mu\text{Pa}) \qquad (\text{dB per Hz}), \tag{48}$$

where the rms pressure level from the noise per Hz is p_{noise}. The reference pressure level used throughout this paper, including table 5, is the ocean standard

equal to 1 μPa. (The standard level used for atmospheric propagation is 20 μPa.) The total background noise ($\mathcal{D}_{\text{noise}}$) admitted by the receiving filter of bandwidth W is about $10 \log W$ decibels greater than $\mathcal{N}$ if the level of the noise is approximately constant in the frequency band of interest.

There are few measurements of the source levels of the vocalizations of animals in air. Some levels for primates have been measured by Waser and Waser (1977; see also table 5). Dumortier (1963) reviewed the source levels for some arthropods. The time-bandwidth products for frogs (Bogert 1960) and birds (Lanyon 1960; Marler 1960; Morton 1970) are often about 200 or greater. Morton (1970) has measured the source levels of some birds in the Canal Zone in Central America. Gerhardt (1975) has measured the source levels of many North American frogs and toads, and the source levels for many marine organisms have been measured (table 5; Fish and Turl 1976; Watkins and Schevill 1979; Thomas and Kuechle 1982; Watkins et al. 1987). The recordings at the Cornell Library of Natural Sounds (Cornell University), the Borror Laboratory of Bioacoustics (Ohio State University), and the Bioacoustics Laboratory and Archives (Florida Museum of Natural History, University of Florida) and the recordings tabulated by Watkins and Wartzok (1985) could also be used to estimate the time-bandwidth products for many other organisms.

The noise levels in tropical forests have been measured by Eyring (1946), Saby and Thorpe (1946), Morton (1970), and Waser and Waser (1977). The levels varied by about 10 to 20 dB with time of day, largely depending on the calling times of other animals. For example, Waser and Waser (1977) measured noise levels in the Uganda forest in the band from 100 to 1500 Hz from 54 to 66 dB relative to 1 μPa at 1 yard. The dawn chorus of birds raises the level by about 10 dB near sunrise. Waser and Waser (1977) also measured noise levels between 67 and 76 dB in the band from 50 to 7500 Hz, and from 81 to 101 dB from 50 to 20,000 Hz. The increased noise level in the last frequency band is due to Orthoptera.

Noise levels in the ocean were summarized by Urick (1983). The levels in deep water vary with shipping activity, wind speed, and frequency. Near 20 Hz, shipping noise is dominant, and typical levels vary from 68 to 85 dB relative to 1 μPa at 1-Hz bands. Near 1000 Hz, wind noise and sea state are dominant, and typical levels vary from 44 to 70 dB in 1-Hz bands. The noise levels in shallow water also exhibit great variation with wind speed, rain, and location. At 1000 Hz, noise levels vary from 50 to 80 dB in 1-Hz bands.

The range at which the call from an animal can be detected increases when cross-correlation is used to estimate the arrival-time differences at the receivers. Let R_1 be the maximum range at which a signal can be detected without a cross-correlator. The signal-to-noise ratio at the receiver is

$$\mathcal{D}_1 \simeq \mathcal{D}_{\text{sl}} - \mathcal{D}_{\text{gl1}} - \alpha R_1 - \mathcal{D}_{\text{noise}} \quad (\text{dB}). \tag{49}$$

If cross-correlation is used, then the same signal-to-noise ratio, $\mathcal{D}_1$, is achieved for a range R_2 given by

$$\mathcal{D}_1 \simeq \mathcal{D}_{\text{sl}} - \mathcal{D}_{\text{gl2}} - \alpha R_2 + 10 \log (d^2/d_i^2) - \mathcal{D}_{\text{noise}} \quad (\text{dB}). \tag{50}$$

When the geometric spreading loss is spherical, the two ranges are related by

$$20 \log R_2 + \alpha R_2 = 20 \log R_1 + \alpha R_1 + 10 \log (d^2/d_i^2). \tag{51}$$

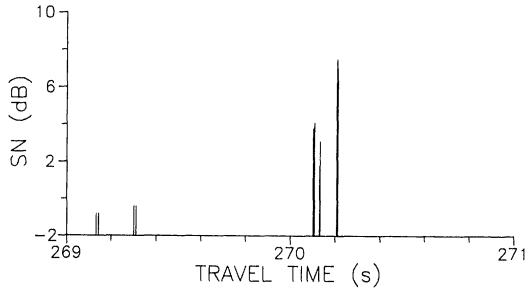

FIG. 4.—The travel times of the acoustic-ray paths versus their predicted signal-to-noise ratios for a finback whale. The geometry is identical to that in figure 2 (400-km range). The source level is 186 dB (Watkins et al. 1987); the noise level in a 5-Hz band near 20 Hz is taken to be 77 dB (Urick 1983); the absorption is 0.1 dB (see the text); and the geometric spreading loss is calculated from ray theory.

For many cases of interest, the detection range of an animal is increased significantly. An example is illuminating.

Consider a call that is received at a range of R_1 with a signal-to-noise ratio, $\mathcal{D}_1$, of 10 dB before cross-correlation. Then, for the black-and-white colobus (table 5), $\mathcal{D}_{sl} = 116$ dB, $\alpha \simeq 0.03$ dB/m near 500 Hz (Waser and Waser 1977), $\mathcal{D}_{noise} \simeq 60$ dB in the band from 100 to 1500 Hz (Waser and Waser 1977); and for a spherical spreading loss the signal is received with a 10-dB signal-to-noise ratio at a range of about 120 m. The cross-correlation gain for this animal depends on the time-bandwidth product of the call and the properties of the background noise (Appendix C). If we assume that the noise samples are independent, as in equation (38), and if the signal-to-noise ratios at the two receivers are about 10 dB, then the gain of the cross-correlator is approximately

$$d^2/d_i^2 = (900)10/(1 + 10 + 10) = 430,$$

or $10\log 430 = 26$ dB. After cross-correlation, the call would be received at 10 dB with receivers at a distance of about 550 m instead of 120 m!

The cross-correlator will not enhance the signal-to-noise ratio very much for the finback whale because the time-bandwidth product of its call is small (table 5). However, because the absorption of sound near 20 Hz in the ocean is only about 0.00003 dB per 100 m (Urick 1983), it is possible to detect its call at a range of at least 400 km (fig. 4). Since the call is about 1 s long, the actual energy would start to arrive at the start of each ray path's arrival time and last for about 1 s for each ray. The total duration of the received energy would be about 1 s because the first ray arrivals are weak. The noise level can drop by another 10–15 dB for periods of a day (Spiesberger, pers. obs.). In that case, the finbacks' call would be detected at a signal-to-noise ratio of about 20 dB. If the receiver is moved from the 400- to 1000-km range and the depths of the whale and the receiver are unchanged, the travel-time difference between the first and last ray arrival is about 2 s. If one of these quiet times occurs, then the signal-to-noise ratio would be about 7–10 dB.

If necessary, the signal-to-noise ratio can be increased with the use of steerable acoustic arrays by about

$$\mathcal{D}_{array} \simeq 10\log N_e \quad \text{(dB)}, \tag{52}$$

where the number of elements in each array is N_e (Urick 1983). The gain from an array comes from the property that it can steer its listening direction toward the source and also attenuate the noise originating from other directions. The price to pay for the array gain is the additional computational burden required when the steering direction must be determined.

Another complication can arise when the animal moves with respect to the receivers. The received acoustic frequencies are changed by an amount that is approximately given by

$$\Delta f \simeq u_{\text{animal}} f/c\,, \tag{53}$$

where the speed of the animal along the line between it and the receiver is u_{animal}, the original acoustic frequency is f, and the speed of sound is c. This is the Doppler effect (Jackson 1975). Different receivers receive different frequencies because the oncoming speed of the animal to each receiver is different. In this case, the recorded signals at each receiver must be corrected before cross-correlation. This is a significant computational problem. Fortunately, many animals do not move when calling.

ANIMAL LOCALIZATION: IGNORING ENVIRONMENTAL FLUCTUATIONS AND RECEIVER-POSITION ERRORS

The positions of calling animals can be estimated when the fluctuations of the wind and sound-speed fields are ignored and when the errors of the receiver positions are small. Watkins and Schevill (1971) have used four hydrophones to locate the position of whales during their calls. They measured the three travel-time *differences* (hereafter, TTD's) between hydrophones 2, 3, and 4 relative to hydrophone 1. The location of the whale has three unknowns: x, y, and z. The position can be determined by solving the three equations that relate the measured TTD's to the unknown position. These equations assume that the speed of sound is known, that effects from the currents are negligible, and that the errors in the receiver positions are insignificant. Magyar et al. (1978) applied the same technique to localize quail. They also assumed that the speed of sound was known and that wind effects were unimportant in their application.

A primary conclusion of this paper is that ignoring the effects of sound speed and wind and ignoring the errors in the receiver positions lead to significantly less accurate localizations than do localizations that account for these effects. However, a starting estimate for the source position (called the reference position), which can be made under ideal conditions, is presented next. The models that incorporate actual fields of sound speed, wind, current, and receiver-position error are discussed in later sections.

Suppose the speed of sound is constant, that there is no wind or current, and that the exact receiver positions are known. If the acoustic TTD between two receivers is known, then the source must lie on a hyperbola (for two-dimensional problems) or on a hyperboloid (for three-dimensional problems) (fig. 5). Every TTD measurement constrains the source to a position on a hyperboloid, and the intersection of the hyperboloids determines the source position. With four receiv-

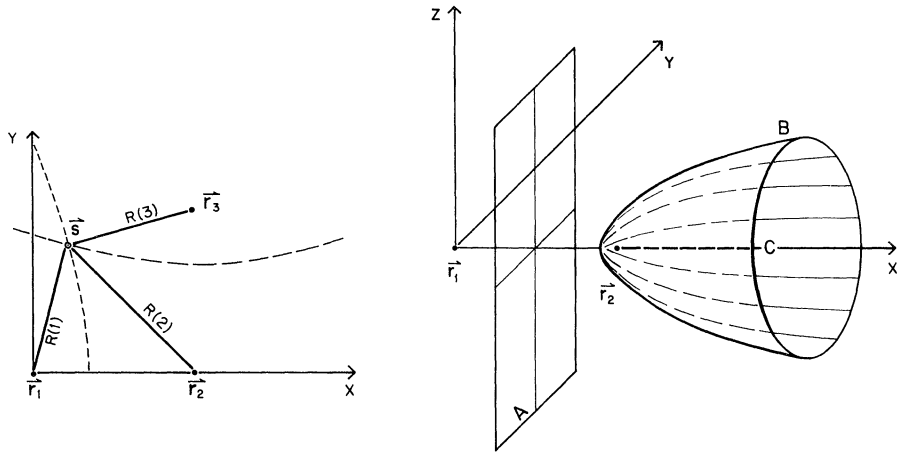

FIG. 5.—*Left*, Localization in two dimensions. The three acoustic receivers are at $\mathbf{r}_i$, where $i = 1, 2, 3$. The acoustic source is at $\mathbf{s}$. The ranges between the source and the receivers are $R(i)$, where $i = 1, 2, 3$. The curves are hyperbolas that define the possible positions of the source based on the measured travel-time differences for receivers 1 and 2 (*short-dashed line*) and for receivers 2 and 3 (*long-dashed line*). *Right*, In three dimensions, the difference in the arrival times (receiver 2 minus receiver 1) constrains the source location to the surface of a hyperboloid. If the travel-time difference is zero, the hyperboloid reduces to a plane (A). If the travel-time difference is positive (the signal arrives at receiver 2 first), the hyperboloid looks like surface B. If the travel-time difference equals the time required for an acoustic signal to travel from receiver 1 to 2, the hyperboloid reduces to line segment C, which starts at $\mathbf{r}_2$ and extends in the positive x direction.

ers, the source is constrained to lie on the intersection of three hyperboloids that may intersect at two points. With five or more receivers, the intersections constrain the position to a single point. The mathematics for obtaining the source position based on five or more receivers follows.

A reference model is taken to have a sound speed c_0. The wind field is neglected. Since there is no wind and since the sound speed is a constant, the acoustic paths are straight.

The reference source position is denoted by

$$\mathbf{s}_0 = (s_{0x}, s_{0y}, s_{0z}) \tag{54}$$

and the reference receiver position by

$$\mathbf{r}_0(i) = \left(r_{0x}(i), r_{0y}(i), r_{0z}(i)\right), \tag{55}$$

where $i = 1, 2, \ldots, N_r$ and boldface characters denote vectors. The number of receivers is N_r, and there are N_r equations that describe the relation between the TTD's and the locations of the source and the receivers:

$$\|\mathbf{s}_0 - \mathbf{r}_0(i)\|^2 = c_0^2[T(1) + \tau_1(i)]^2, \tag{56}$$

where $i = 1, 2, \ldots, N_r$ and the measured TTD's are

$$\tau_1(i) = T(i) - T(1) + e_i, \tag{57}$$

where $i = 1, 2, 3, \ldots, N_r$, $T(i)$ is the travel time to receiver i, and e_i is the error in the measurement. In equations (56), the modulus squared of a vector is denoted by $\|\mathbf{x}\|^2$, which is the square of the length of the vector. As suggested by Watkins and Schevill (1971), equation (56) is subtracted from the remaining $N_r - 1$ equations, and receiver one is placed at the origin of the coordinate system:

$$\mathbf{r}_0(1) = (0, 0, 0).$$

The resulting equations are

$$2r_{0x}(i)s_{0x} + 2r_{0y}(i)s_{0y} + 2r_{0z}(i)s_{0z} + 2c_0^2\tau_1(i)T(1) = -c_0^2\tau_1^2(i) + \|\mathbf{r}_0(i)\|^2, \quad i = 2, 3, 4, \ldots, N_r. \tag{58}$$

Equation (58) is linear in the four unknowns:

$$\mathbf{m} = \begin{pmatrix} s_{0x} \\ s_{0y} \\ s_{0z} \\ T(1) \end{pmatrix}. \tag{59}$$

Therefore, we can rewrite equation (58) in matrix notation:

$$\mathbf{Am} = \mathbf{b}, \tag{60}$$

where the ith row of the matrix $\mathbf{A}$ is

$$[2r_{0x}(i + 1) \; 2r_{0y}(i + 1) \; 2r_{0z}(i + 1) \; 2c_0^2\tau_1(i + 1)],$$

where $i = 1, 2, \ldots, N_r - 1$ and $\mathbf{A}$ has $N_r - 1$ rows. The ith row of $\mathbf{b}$ is

$$b_i = -c_0^2\tau_1^2(i + 1) + \|\mathbf{r}_0(i + 1)\|^2, \tag{61}$$

where $i = 1, 2, \ldots, N_r - 1$. Note that with five or more receivers, a reference position for the source can be determined from equation (60) by solving linear equations. The solution to equation (60) that minimizes the sum of the squares of the residuals is

$$\mathbf{m}_1 = \mathbf{VS}^*\mathbf{U}^T\mathbf{b}, \tag{62}$$

where

$$\mathbf{A} = \mathbf{USV}^T \tag{63}$$

is the singular-value decomposition of the matrix $\mathbf{A}$ and

$$\mathbf{S}_{4\times(N_r-1)} = \begin{pmatrix} \tilde{\mathbf{S}} \\ \mathbf{O} \end{pmatrix}. \tag{64}$$

$\tilde{\mathbf{S}}$ is a square, nonsingular, diagonal submatrix of dimensions $k \times k$, where the rank of $\mathbf{A}$ is k and the matrix $\mathbf{S}^*$ is related to $\tilde{\mathbf{S}}$ through

$$\mathbf{S}^*_{4\times(N_r-1)} = (\tilde{\mathbf{S}}^{-1} \; \mathbf{O}) \tag{65}$$

(Lawson and Hanson 1974, chaps. 4, 7). The superscript T denotes the transpose operation. The uncertainty in the model can be estimated using standard linear-

error analysis. To the first order, the noise component of b_i given in equation (61) is

$$\delta b_i \simeq -2c_0^2\tau_1(i + 1)e_i, \tag{66}$$

where $i = 1, 2, \ldots, N_r - 1$, $\tau_1(i)$ is considered a known number (the measured data), and e_i is considered a random variable. Define the deviation from the reference model, $\mathbf{m}_1$, as

$$\delta\mathbf{m}_1 = \begin{pmatrix} \delta s_{0x} \\ \delta s_{0y} \\ \delta s_{0z} \\ \delta T(1) \end{pmatrix}. \tag{67}$$

The error-covariance matrix for $\delta\mathbf{m}_1$ because of errors e_i is

$$\mathbf{C}_0 \simeq \overline{\delta\mathbf{m}_1\delta\mathbf{m}_1^T} = \mathbf{V}\mathbf{S}^*\mathbf{U}^T\overline{\delta\mathbf{b}\delta\mathbf{b}^T}\mathbf{U}(\mathbf{S}^*)^T\mathbf{V}^T. \tag{68}$$

The expected value of a random variable e_i is $\bar{e}_i$. The error-covariance matrix defines the uncertainty of the model estimate given by equation (62). The i,jth element of $\overline{\delta\mathbf{b}\delta\mathbf{b}^T}$ is

$$(\overline{\delta\mathbf{b}\delta\mathbf{b}^T})_{ij} = 4c_0^4\tau_1(i + 1)\tau_1(j + 1)\,\overline{e_ie_j}. \tag{69}$$

ANIMAL LOCALIZATION: ACOUSTIC TOMOGRAPHY

In the preceding section, a reference model was defined. If the effect of winds (currents), sound-speed fluctuations, and errors in the receiver positions are significant, acoustic tomography provides a method by which these quantities can be estimated from the data. Tomography is the study of sections of a medium. Tomographic reconstruction is the synthesis of the three-dimensional structure from the two-dimensional sections. In our case, the sections are the acoustic paths between the calling animal and the receivers. The three-dimensional fields of sound speed and wind or current are synthesized. We also obtain improved estimates for the positions of the animal and the receivers.

If the effects of winds, sound-speed fluctuations, etc., are significant, the errors (e_i) used to determine the error-covariance matrix in equation (69) for the reference position must include contributions from the neglect of the wind, etc. For example, if the error of the travel-time difference (TTD) is a millisecond and the effect of the wind on a TTD is on the order of 10 ms, one should set the standard deviation of e_i in equation (69) equal to

$$(0.001^2 + 0.010^2)^{1/2} \simeq 0.01 \text{ s}.$$

In this manner, the error-covariance matrix for the reference model will contain errors large enough to accommodate the explicit neglect of wind effects in model equation (68). In actual applications, it is important to estimate the error-covariance matrix in order to obtain accurate estimates of the errors of the source position. The errors in the TTD are significantly correlated because the spatial scales of the wind and the sound-speed perturbation are usually large enough to

influence the travel times along many of the acoustic paths. A method of estimating the diagonal and off-diagonal terms of the error-covariance matrix is given later.

The first part of the acoustic-tomography problem discusses how the environmental fluctuations and the errors of the receiver positions affect the acoustic travel-time differences. In this discussion, the sound-speed and wind fields and the errors in the receiver positions are assumed to be known and the travel-time data are calculated. The second part of the acoustic-tomography problem deals with the more difficult problem of estimating the sound-speed and wind fields and the errors of the receiver positions, assuming that the data are known. These problems are known as the forward problem and the inverse problem, respectively.

Forward Problem

The forward problem is formulated as follows. The reference TTD's are calculated as in

$$\tau_0(i) = n_0\|\mathbf{s}_0 - \mathbf{r}_0(i)\| - n_0\|\mathbf{s}_0 - \mathbf{r}_0(1)\|, \tag{70}$$

where $i = 2, 3, 4, \ldots, N_r$ and the reference sound slowness is defined by

$$n_0 = 1/c_0. \tag{71}$$

The sound slowness is used instead of sound speed because the subsequent mathematics are simplified. We model the measured TTD by

$$\begin{aligned}\tau_1(i) = &\int_{\mathbf{s}_1}^{\mathbf{r}_1(i)} [n_0 + \delta n(\mathbf{l}_i)]dl_i - \frac{1}{c_0^2}\int_{\mathbf{s}_1}^{\mathbf{r}_1(i)} \mathbf{u}(\mathbf{l}_i)\circ\mathbf{dl}_i \\ &- \int_{\mathbf{s}_1}^{\mathbf{r}_1(1)} [n_0 + \delta n(\mathbf{l}_1)]dl_1 + \frac{1}{c_0^2}\int_{\mathbf{s}_1}^{\mathbf{r}_1(1)} \mathbf{u}(\mathbf{l}_1)\circ\mathbf{dl}_1 + e_i,\end{aligned} \tag{72}$$

where $i = 2, 3, 4, \ldots, N_r$.

The actual positions of the source and receiver are

$$\mathbf{s}_1 = \mathbf{s}_0 + \delta\mathbf{s} \tag{73}$$

and

$$\mathbf{r}_1(i) = \mathbf{r}_0(i) + \delta\mathbf{r}(i), \tag{74}$$

where $i = 1, 2, 3, \ldots, N_r$, respectively. The increment of acoustic path length (direction) between the actual source position and receiver i is denoted by dl_i ($\mathbf{dl}_i$). In general, the path is curved because of acoustic refraction. The actual sound-slowness field along path $\mathbf{l}_i$ is

$$n_1(\mathbf{l}_i) = n_0 + \delta n(\mathbf{l}_i). \tag{75}$$

The wind field along path i is $\mathbf{u}(\mathbf{l}_i)$. In equation (72), the unknown models are (1) $\delta\mathbf{s}$, the correction to the reference position of the source; (2) $\delta\mathbf{r}(i)$, the correction to the reference position of receiver i; (3) $\delta n(\mathbf{l})$, the correction to the reference sound slowness; and (4) $\mathbf{u}(\mathbf{l})$, the wind field.

Because we do not care about the absolute geographical location or the particular rotation of the coordinate frame in which the animal calls are mapped, some of the corrections of the reference positions of the receiver can be set to zero. The origin of the reference frame is located at receiver number 1 by definition; therefore, the variances of the x, y, and z components of its position are zero. The x-axis is defined such that receiver 1 is at x equal zero and receiver 2 has a y-coordinate equal to zero. The variance of the y component of the correction to receiver 2 is zero. This leaves one more parameter to determine: the rotation about the x-axis. This is defined such that the z-coordinate of receiver 3 is at a given height from the x-y plane. Then, by definition, the correction of the z position of receiver 3 is zero.

The differences between the measured and reference TTD's are

$$\delta\tau(i) = \tau_1(i) - \tau_0(i)\,, \tag{76}$$

where $i = 2, 3, 4, \ldots, N_r$, and these are called the differential travel-time differences (hereafter, DTTD). The DTTD's are the data used in the inversion. The DTTD's are decomposed into

$$\delta\tau(i) = \delta\tau_2(i) + \delta\tau_3(i)\,, \tag{77}$$

where $i = 2, 3, 4, \ldots, N_r$, $\delta\tau_2$ is a linear function of the models, and $\delta\tau_3$ is a nonlinear function of the models. The tomographic problem is linearized by neglecting the nonlinear constituent in the forward problem. The validity of this approximation can be checked separately. We believe that the linear approximation is adequate for many cases.

The linear constituent is given (Appendix D) by

$$\begin{aligned}\delta\tau_2(i) = {} & n_0[F(i) - F(1)] + e_i \\ & - \frac{1}{c_0^2}\int_{\mathbf{s}_0}^{\mathbf{r}_0(i)} \mathbf{u}(\boldsymbol{\xi}_i)\circ d\boldsymbol{\xi}_i + \frac{1}{c_0^2}\int_{\mathbf{s}_0}^{\mathbf{r}_0(1)} \mathbf{u}(\boldsymbol{\xi}_1)\circ d\boldsymbol{\xi}_1 \\ & + \int_{\mathbf{s}_0}^{\mathbf{r}_0(i)} \delta n(\boldsymbol{\xi}_i)d\boldsymbol{\xi}_i - \int_{\mathbf{s}_0}^{\mathbf{r}_0(1)} \delta n(\boldsymbol{\xi}_1)d\boldsymbol{\xi}_1\,,\end{aligned} \tag{78}$$

where $i = 2, 3, 4, \ldots, N_r$, and where

$$\begin{aligned}F(i) = {} & \frac{[r_{0x}(i) - s_{0x}][\delta r_x(i) - \delta s_x] + [r_{0y}(i) - s_{0y}][\delta r_y(i) - \delta s_y]}{\|\mathbf{r}_0(i) - \mathbf{s}_0\|} \\ & + \frac{[r_{0z}(i) - s_{0z}][\delta r_z(i) - \delta s_z]}{\|\mathbf{r}_0(i) - \mathbf{s}_0\|}\,,\end{aligned} \tag{79}$$

where $i = 2, 3, 4, \ldots, N_r$. In equation (78), the incremental length (direction) along the *straight* line from the source to receiver i is $d\xi_i$ ($d\boldsymbol{\xi}_i$). The vector that points from the origin to points along the path from a source to receiver i is ξ_i. In equation (79), the Cartesian coordinates between the exact and reference positions of the source and receiver are $\delta\mathbf{s}(i) = (\delta s_x(i), \delta s_y(i), \delta s_z(i))$ and $\delta\mathbf{r}(i) = (\delta r_x(i), \delta r_y(i), \delta r_z(i))$, respectively. The linearized forward problem is given by equation (78), where the DTTD's are used in place of the $\delta\tau_2(i)$.

Inverse Problem

The principal concern of inverse theory is to describe properties of the unknown fields of interest from the measurements. In this case, the unknowns are the source position, the sound-speed and wind fields, and possibly the corrections to the positions of the receivers. In this paper, the data are the TTD's. It is customary to call the unknown quantities the *model*. The *forward problem* is the calculation of the data given some model. The *inverse problem* is the estimation of the model given some data.

Much has been written concerning inverse theory in the geophysical literature (Backus and Gilbert 1968; Parker 1977). A summary of some concepts is described here.

Only a finite amount of data can be measured. For all but a few pathological inverse problems, an infinite number of models fit the data, even if the data contain no errors. Additionally, there is a fundamentally lower limit on the scales that can be resolved from the data (Backus and Gilbert 1968). Although usually an infinite number of models fit the data, the bounds in these models are often small enough that useful information is obtained. For example, it may be possible to infer that the sound speed is between 330 and 331 m/s in a certain region in the atmosphere.

There are two kinds of inverse problems, linear and nonlinear. If the forward problem contains a linear relation between the model and the data, the inverse problem is linear. Otherwise, the inverse problem is nonlinear. It is possible to estimate the precision and resolution of all the models that fit the data for linear inverse problems. For nonlinear inverse problems, it is usually difficult to estimate the bounds on all the models that fit the data. However, if the forward problem can be linearized and if the problem is nearly linear, useful bounds on the acceptable models can be estimated (Parker 1977).

For nonlinear inverse problems, it is helpful to construct the best guess possible for the model. The starting model is called the reference model. It is then necessary to solve for the deviation from the starting model. We adopt this procedure because the relations between the animal position, the sound-speed field, and the wind field are nonlinearly related to the data (TTD's).

The models that fit the data can be constructed using many techniques. The theory described here was invented by Wiener and is called Wiener filtering. This technique is also known as *objective analysis* (in oceanography) and *Kriging* (in economics). Two of the many other inverse techniques that would provide useful results are parameter-estimation schemes, which usually rely on some least-squares solutions (Lawson and Hanson 1974), and Kalman filters (Ghil and Cohn 1983), which allow the data to be assimilated in numerical models of the atmosphere or the ocean.

The Wiener filter is a linear estimator of the model based on the data. Suppose there are N data and that the data are placed into a vector, $\mathbf{g}$. This is referred to as the data vector. In our case, there are $N_s \times (N_r - 1) = N$ data, where the number of sources is N_s and the number of receivers is N_r. The data vector could also include measurements of the winds (currents) at certain locations or measure-

ments of the sound speed at certain locations. All the data are placed in the data vector, and the data need not be of the same type or even have the same units.

Refer to the quantities that we wish to estimate from the data as the models. Some choices for the models could be the wind (current) at a certain location, the sound speed at a certain location, or the correction to the reference locations of one of the sources or one of the receivers. Each of the models is written as a linear combination of the data:

$$\hat{m} = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{g}\,, \tag{80}$$

where the N unknown weights that multiply the data are contained in the weight vector $\boldsymbol{\alpha}$. The rest of the analysis involves choosing the weights in order to minimize the error of the estimate of the model. For simplicity, the mean value of the data and the model is taken to be zero. This assumption need not be made (Bretherton et al. 1976), but in practice many useful problems may be solved with this assumption by properly formulating the forward problem. The weights are chosen such that the expected mean-square error between the estimated model, $\hat{m}$, and the true model, m, is a minimum. In other words, the weights are chosen such that

$$Q = \overline{(m - \hat{m})^2} \tag{81}$$

is a minimum. The quantity Q is just a number, and the choice of $\boldsymbol{\alpha}$ that minimizes Q is obtained by taking the partial derivative of Q with respect to each of the elements of $\boldsymbol{\alpha}$ and setting each equation to zero. This results in a set of N equations (one for each partial derivative) in N unknowns (the N alpha weights). The solution for $\boldsymbol{\alpha}$ that minimizes Q (minimum-mean-square-error solution) is given by

$$\boldsymbol{\alpha} = \boldsymbol{\Gamma}^{-1}\boldsymbol{\beta}\,. \tag{82}$$

The solution involves a symmetrical square matrix, $\boldsymbol{\Gamma}$, having N rows and N columns. The i,jth element of this matrix is

$$\Gamma_{ij} = \overline{g_i g_j}\,, \tag{83}$$

where the ith datum is g_i. This matrix is referred to as the data-data covariance matrix. Its construction relies on being able to estimate the covariances between the different data. In practice, this can often be done, and we will come back to this construction later for this application.

The solution also involves a vector, $\boldsymbol{\beta}$, which has N rows. Its ith row is

$$\beta_i = \overline{m g_i}\,. \tag{84}$$

This vector is referred to as the model-data vector. Its construction relies on being able to estimate the covariance between the model and each of the data. Again, this can be done, and the construction is discussed later.

Now that a criterion has been used to determine the weights that multiply the data to give the model, it is of interest to estimate the error of the model estimate. The variance of the model estimate is

$$Q_{\min} = \overline{m^2} - \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\boldsymbol{\beta}\,. \tag{85}$$

The first term on the right-hand side is the variance of the true model (hereafter, the a priori variance). For example, if the model is the correction to the x-coordinate of source 1, then its variance is $\overline{m^2}$. The second term on the right-hand side is the amount by which the a priori variance is reduced because of the information provided by the data. If the data are not correlated with the model, then $\boldsymbol{\beta}$ is zero and the second term on the right-hand side is zero. In this case, the variance of the model is just the a priori variance because the data have told us nothing about the model. For example, if the model is the sound speed at a location 10^6 km from a small array of receivers, then we expect that the acoustic travel-time data will tell us nothing about the value of that model, and all we know is the a priori variance of that model. Moreover, if the model is chosen to be the sound speed in the center of the array, then the variance of the estimated model is less than the a priori variance. (For more-complete discussions of the Wiener filter, see Bretherton et al. 1976; Cornuelle 1982; Spiesberger 1985*b*.)

Construction of the Wiener Filter

With the data given by equation (78), the i,jth element of the data-data covariance matrix is

$$\begin{aligned}
\overline{\delta\tau_2(i)\delta\tau_2(j)} = \Gamma_{ij} = n_0^2\overline{[F(i) - F(1)][F(j) - F'(1)]} \\
+ \frac{1}{c_0^4}\int_{s_0}^{r_0(i)}\int_{s_0'}^{r_0(j)} \overline{\mathbf{u}(\boldsymbol{\xi}_i)\circ d\boldsymbol{\xi}_i \mathbf{u}(\boldsymbol{\xi}_j)\circ d\boldsymbol{\xi}_j} - \frac{1}{c_0^4}\int_{s_0}^{r_0(i)}\int_{s_0'}^{r_0(1)} \overline{\mathbf{u}(\boldsymbol{\xi}_i)\circ d\boldsymbol{\xi}_i \mathbf{u}(\boldsymbol{\xi}_1')\circ d\boldsymbol{\xi}_1'} \\
- \frac{1}{c_0^4}\int_{s_0}^{r_0(1)}\int_{s_0'}^{r_0(j)} \overline{\mathbf{u}(\boldsymbol{\xi}_1)\circ d\boldsymbol{\xi}_1 \mathbf{u}(\boldsymbol{\xi}_j)\circ d\boldsymbol{\xi}_j} + \frac{1}{c_0^4}\int_{s_0}^{r_0(1)}\int_{s_0'}^{r_0(1)} \overline{\mathbf{u}(\boldsymbol{\xi}_1)\circ d\boldsymbol{\xi}_1 \mathbf{u}(\boldsymbol{\xi}_1)\circ d\boldsymbol{\xi}_1'} \\
+ \int_{s_0}^{r_0(i)}\int_{s_0'}^{r_0(j)} \overline{\delta n(\boldsymbol{\xi}_i)\delta n(\boldsymbol{\xi}_j)}\, d\xi_i d\xi_j - \int_{s_0}^{r_0(i)}\int_{s_0'}^{r_0(1)} \overline{\delta n(\boldsymbol{\xi}_i)\delta n(\boldsymbol{\xi}_1')} d\xi_i d\xi_1' \\
- \int_{s_0}^{r_0(1)}\int_{s_0'}^{r_0(j)} \overline{\delta n(\boldsymbol{\xi}_1)\delta n(\boldsymbol{\xi}_j)} d\xi_1 d\xi_j + \int_{s_0}^{r_0(1)}\int_{s_0'}^{r_0(1)} \overline{\delta n(\boldsymbol{\xi}_1)\delta n(\boldsymbol{\xi}_1')} d\xi_1 d\xi_1' + \overline{e_i^2}\delta_{ij},
\end{aligned} \tag{86}$$

where $F(i)$ is given by equation (79). The prime indicates that the source for datum j may not be the same as the source for datum i. The Kroenecker delta function, δ_{ij}, is one if i equals j and is zero otherwise. This means that the noise components of data i and j are uncorrelated. For simplicity, the covariances of the following quantities are also set to zero:

$$\begin{aligned}
\overline{\delta n \mathbf{u}} &= \mathbf{0}, && \text{sound slowness is not correlated with the wind;} \\
\overline{\delta n e_i} &= 0, && \text{sound slowness is not correlated with the noise;} \\
\overline{\mathbf{u} e_i} &= \mathbf{0}, && \text{wind or current is not correlated with the noise.}
\end{aligned} \tag{87}$$

Even if these quantities are correlated, the Wiener filter still produces models that are consistent with reality except that its estimates of the models' variance is too large. In other words, the assumption that these covariances are zero places fewer constraints on the models and places more burden on the data to produce models of sufficient accuracy.

In equation (86), the term that deals with the corrections to the positions of the sources and receivers is

$$n_0^2\overline{[F(i) - F(1)][F(j) - F'(1)]},$$

which contains the covariances (with $\delta_{ss'} = 0$ for different sources and $\delta_{ss'} = 1$ for the same source)

$$\begin{aligned}
\overline{\delta r_x(i)\,\delta r_x(j)} &= \sigma^2_{\delta r_x(i)}\delta_{ij}\,,\\
\overline{\delta r_y(i)\,\delta r_y(j)} &= \sigma^2_{\delta r_y(i)}\delta_{ij}\,,\\
\overline{\delta r_z(i)\,\delta r_z(j)} &= \sigma^2_{\delta r_z(i)}\delta_{ij}\,,\\
\overline{\delta r_x(i)\,\delta r_y(j)} &= \overline{\delta r_x(i)\,\delta r_z(j)} = \overline{\delta r_y(i)\,\delta r_z(j)} = 0\,,\\
\overline{\delta s_x\,\delta s'_x} &= \sigma^2_{\delta s_x}\delta_{ss'}\,,\\
\overline{\delta s_y\,\delta s'_y} &= \sigma^2_{\delta s_y}\delta_{ss'}\,,\\
\overline{\delta s_z\,\delta s'_z} &= \sigma^2_{\delta s_z}\delta_{ss'}\,,\\
\overline{\delta s_x\,\delta s_y} &= \overline{\delta s_x\,\delta s_z} = \overline{\delta s_y\,\delta s_z} = 0\,.
\end{aligned} \tag{88}$$

The simplifying assumptions set the covariances of different models to zero. For example, the correction to the x component of receiver i's position is taken to be uncorrelated with the y or z component of receiver i's position. The covariances for the source corrections are obtained from the diagonal elements of the error-covariance matrix of the reference position of the source (eq. 68). The covariances of the receiver corrections are obtained from surveys that the investigator measures. In other words, an investigator places some receivers in certain spots and estimates their (x, y, z) positions and the errors of those positions. As before, the simplifying step that sets certain covariances to zero increases the errors of the models estimated from the data. If the cross-covariances are estimated and the Wiener filter is written to include the extra cross terms, then the errors of the model estimates decrease.

The terms in the data-data covariance matrix that involve double integrals of the form

$$\frac{1}{c_0^4}\int_{\mathbf{s}_0}^{\mathbf{r}_0(i)}\int_{\mathbf{s}'_0}^{\mathbf{r}_0(j)}\overline{\mathbf{u}(\boldsymbol{\xi}_i)\circ d\boldsymbol{\xi}_i\,\mathbf{u}(\boldsymbol{\xi}_j)\circ d\boldsymbol{\xi}_j}$$

involve covariances of the wind (current) fields. The Cartesian coordinates of the wind (current) are $\mathbf{u} = (u, v, w)$. In this paper, the covariances between u and v, u and w, and v and w are set to zero for simplicity. Setting these covariances to zero has the effect of placing fewer constraints on the inversion, resulting in error bars larger than necessary. For more-complicated implementations of this inverse, these terms can be calculated and incorporated into the data-data covariance matrix. Similarly, the terms in the data-data covariance matrix that involve double integrals containing integrands like

$$\overline{\delta n(\boldsymbol{\xi}_i)\,\delta n(\boldsymbol{\xi}_j)}$$

require that the covariance of the sound-slowness field be specified.

A common way to specify the covariance function for the winds (currents) or sound-slowness field is to choose some covariance lengths, L_x, L_y, and L_z, which have the same scales as the shortest scales in fields containing a significant amount of energy. The covariance function can be obtained from some physically possible covariance function such as the Gaussian function:

$$\begin{aligned}&\overline{\delta n(x_i, y_i, z_i)\delta n(x_j, y_j, z_j)}\\ &= \sigma^2_{\delta n} \exp\left[-(x_i - x_j)^2/L_x^2 - (y_i - y_j)^2/L_y^2 - (z_i - z_j)^2/L_z^2\right].\end{aligned} \tag{89}$$

Similar relations can be used for each component of the wind. The variance of this function, $\sigma^2_{\delta n}$, is chosen to be the variance of the sound-slowness variations expected in the region. Similar expressions can be used for the winds. On windy days, the variance of the u, v, or w components would have larger values than on calm days.

Recall that the reference locations and the errors of the reference locations of the animals are obtained from equations (62) and (68). The data-data covariance elements given in equation (86) can be used to estimate the error-covariance elements, $\overline{e_i e_j}$ in equation (69). We simply set the variance of the source uncertainty in equation (86) to zero to allow estimation of the errors in the travel-time differences resulting from the neglect of the receiver-position uncertainties, sound-speed uncertainties, and winds.

The Wiener filter also requires evaluation of the model-data covariance vector $\boldsymbol{\beta}$. We choose four models. (1) The corrections to the reference positions of the sources. If there are N_s sources, then there are $3N_s$ unknowns. (2) The corrections to the reference positions of the receivers. For N_r receivers there are $3(N_r - 1) - 2$ unknowns. The reference receiver is at the origin by definition, and the positions of the other receivers are therefore relative to it. Two other Cartesian coordinates of receivers 2 and 3 are defined quantities, since they determine the orientation of the coordinate frame. (3) The wind (current) field at $N_{\mathbf{u}}$ locations. Since each location requires the wind in all three directions, there are $3N_{\mathbf{u}}$ unknowns. (4) The sound-slowness field at N_{SS} locations. The total number of models is

$$N_{\text{models}} = 3N_s + 3(N_r - 1) + 3N_{\mathbf{u}} + N_{SS} - 2. \tag{90}$$

For each model, a model-data covariance vector needs to be computed. Some of these vectors are computed here. For the model denoted by $\delta r_x(i)$, the model-data vector is zero except on the rows where the travel-time datum from a source involves receiver i. For source 1, the nonzero element is

$$n_0 \frac{r_{0x} - s_{0x}}{\|\mathbf{r}_0(i) - \mathbf{s}_0\|} \sigma^2_{\delta r_x(i)}.$$

This is obtained by taking the expectation of the model with the data as expressed by equation (84). The model-data covariance elements for the remaining corrections for the sources and the receivers are similar. For a sound-slowness model at

a position $\mathbf{x}$, the ith row of the model-data covariance matrix is

$$\int_{\mathbf{s}_0}^{\mathbf{r}_0(i)} \overline{\delta n(\mathbf{x})\,\delta n(\boldsymbol{\xi}_i)}\,d\xi_i - \int_{\mathbf{s}_0}^{\mathbf{r}_0(1)} \overline{\delta n(\mathbf{x})\,\delta n(\boldsymbol{\xi}_1)}\,d\xi_1 .$$

The model-data vectors for the wind models are similar.

Terrestrial Examples

The following examples of tomographic localization presume mild atmospheric conditions. Nonetheless, localizations are significantly more precise when the atmospheric fluctuations are accounted for.

We examine the accuracy of locating a calling animal with and without the use of tomography. Five receivers are placed in a region of about 30 m on a side (fig. 6). The Gaussian-length scales in equation (89) for the horizontal variations of the wind and the sound speed are taken to be representative of gusts or a temporary "mean wind" with $L_x = L_y = 300$ m. The vertical scale for the wind is set to $L_z = 20$ m, and the vertical scale for the sound speed is set to $L_z = 5$ m. The standard deviation of each horizontal component of the wind is 1 m/s. The standard deviation of the vertical component of the wind is 0 m/s. The standard deviation of the sound speed is 1.2 m/s, which is the same as a change of 1°C at a reference temperature of 20°C. The evaluation of the standard deviation of the travel-time differences is estimated from equations (14), (25), and (42), which give the contributions from the straight-ray approximation, the turbulence, and the acoustic noise. For $c_1 = c(0) = 338.1$ m/s, $c(10\text{ m}) = 339$ m/s (nighttime conditions); $c_z = 0.090\ \text{s}^{-1}$; $R = 30$ m; $\delta T_1 \simeq 0.94 \times 10^{-6}$ s (for straight-ray approximations); $z_0 = 0.05$ m (roughness length for long grass); $U(10) = 1$ m/s; $z = 3$ m; $\sigma_t \simeq 9.6 \times 10^{-6}$ s (from turbulence); $W_{\text{rms}} = 1000$ Hz; $d = 19.95$ (equivalent to a signal-to-noise ratio after correlation of 26 dB); and $\sigma_t \simeq 8.1 \times 10^{-6}$ s from acoustic noise.

The standard deviation of the TTD from these three effects is approximately

$$16 \times 10^{-6} \simeq [2(0.94)^2 + 2(9.6)^2 + 8.2^2]^{1/2} \times 10^{-6}\ \text{s},$$

where the factors of 2 come from the independent errors at two receivers for the ray-bending and turbulence contributions. The relative positions of the receiver are assumed to be known with an accuracy of 1 cm horizontally and 5 cm vertically. Error maps of the source position are computed without tomography from equations (68), where the error-covariance matrix of the TTD's is estimated from the data-data covariance matrix in equations (86) without including the errors from the uncertainty of the source position. The errors range from a few centimeters near the center of the study area to a few meters (fig. 6). The errors of locating a single source are smaller by a factor of 2 to 20 when tomography is used (fig. 6). One of the reasons for using tomography is that the localization errors are significantly smaller because the effects of the environment and the receiver-position errors are taken into account. Without tomography, the only place to put the error is in the source position, but this is wrong and too harsh. When the error is distributed into the environment and the uncertainty of the receiver positions, the localization of the calling animal is improved.

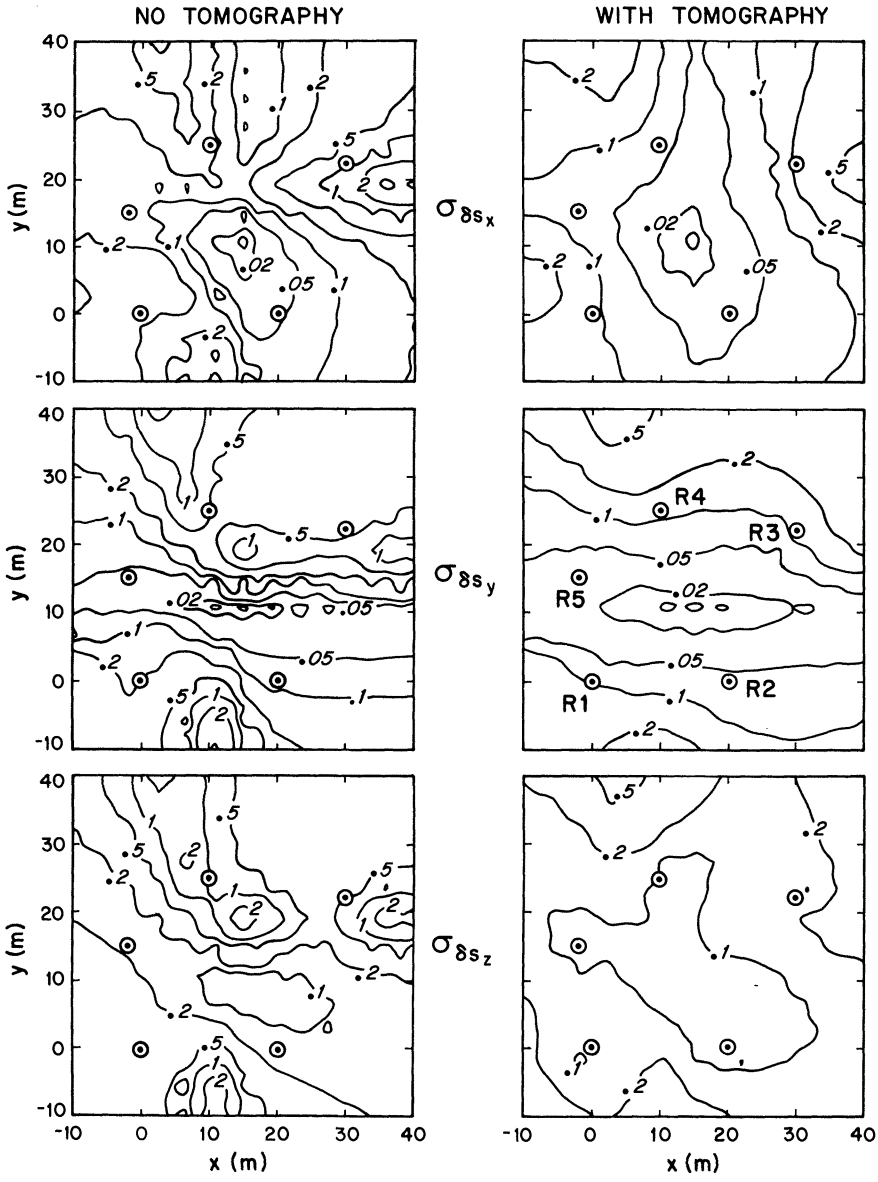

FIG. 6.—Error maps in the x, y, and z directions (*top*, *middle*, and *bottom*, respectively) for the location of an acoustic source at various places in the horizontal (x-y) plane at an elevation of 2 m without the use of tomography (left column) and with the use of tomography (right). Five receivers are placed at heights of 0, 5, 0, 5, and 0 m for R1–R5, respectively, and their horizontal positions are indicated in each panel and labeled in the middle panel on the right. The standard deviation of the horizontal (vertical) components of the receiver positions is assumed to equal 0.01 (0.05) m. See the text for a description of the scales assumed for the wind and the sound-speed perturbation. The source has a root mean square of the bandwidth of 1000 Hz, and the signal-to-noise ratio of the cross-correlated output is assumed to be 20 dB. The contour levels are logarithmic with levels at 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, and 50 m.

Figure 7 is the same as figure 6 except that the relative positions of the receivers are assumed to be known to within 50 cm for each horizontal component and 50 cm for the vertical component. Without tomography, the errors of localization range from 0.1 to 20 m. With tomography, the errors range from about 0.1 to 2 m. For accurate localizations, it is important to obtain relative positions of the receivers to about 1 cm and to use tomography. Acoustic tomography can also be used to obtain the relative positions of the receivers to about 1 cm (Spiesberger, MS).

An array of 12 receivers was distributed over a region of about 300 m in diameter to determine how accurately a black-and-white colobus (table 5) can be located when it emits a roaring bout (fig. 8). Assume that the Gaussian horizontal scales for the wind and the sound-speed perturbations are 300 m. The vertical scale is 20 m for the wind and 5 m for the sound-speed perturbation. The standard deviation of the horizontal components of the wind is 2 m/s and that of the vertical component is 0 m/s. The standard deviation of the sound-speed perturbation is 2 m/s. The standard deviation of the TTD's from the straight-ray approximation and the acoustic noise are given by $c_1 = c(5) = 345.0$ m/s and $c(10) = 344.6$ m/s (daytime conditions); $c_z = -0.080\ \mathrm{s}^{-1}$; $R = 300$ m; $\delta T_1 \simeq 0.00019$ s (straight-ray approximation); $W_{\mathrm{rms}} = 500$ Hz (table 5; this is a conservative estimate); $d = 3.2$ (equivalent to a signal-to-noise ratio after correlation of 10 dB); and $\sigma_t \simeq 0.0001$ s (from acoustic noise). The effects of turbulence are negligible in this case. The standard deviation of the TTD from these two effects is approximately $0.00029 \simeq [2(0.00019)^2 + 0.0001^2]^{1/2}$ s. The relative positions of the receivers are known to 2 cm horizontally and 6 cm vertically. The localizations errors in the horizontal plane are about a factor of 2 better with tomography than without. The vertical errors are reduced by a factor of 5 to 10 with tomography.

We examined the accuracy of tomographic maps of the sound speed and the wind field using an array of 10 receivers and four simultaneously calling animals (fig. 9). There are four sources times (10 − 1) receiver pairs, giving 36 different TTD's available for the tomographic inversion, even though there are only four animals. The horizontal scales for the sound speed and for each component of the wind are the same as for figures 6 and 7 (300 m). The vertical scale for the wind and the sound speed is set to $L_z = 5$ m. At a height of 1 m, the standard deviation of u and v is 0.81 m/s, and the standard deviation of the sound speed is set to a typical diurnal variation of 5.9 m/s, which is equivalent to 10°C. The reference temperature is the same as for figures 6 and 7. As before, the standard deviation of the TTD is set to 16×10^{-6} s. When the relative positions of the receivers are known accurately (fig. 9, left column), more than 90% of the a priori sound-speed variance is mapped, and about 86% of the wind variance is mapped. The equivalent precision of the temperature measurement is given by equation (4) and is about 0.68°C in the region between the receivers. The localization of the four sources is accurate to within about 1 cm for this case. When the relative positions of the receivers are less well known (fig. 9, right column), more than 90% of the sound-speed variance is still mapped, but little of the a priori variance of the wind is mapped. The localization of the four sources is accurate to within about 10 cm for this case. Since the relative positions of the receivers are less well measured, the tomographic maps of the wind degrade more than the tomographic maps of the

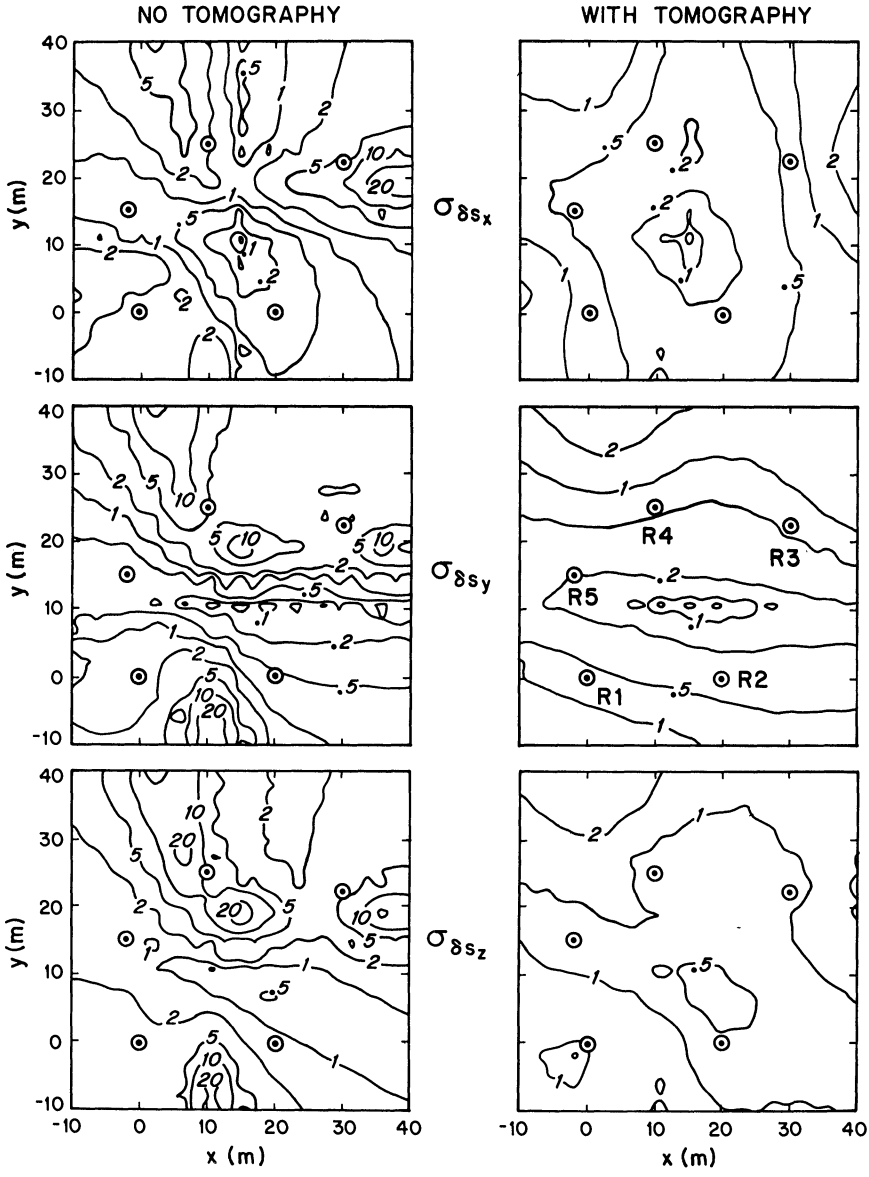

FIG. 7.—Same as figure 6 except the standard deviation of the horizontal and the vertical components of the receiver positions is assumed to equal 0.5 m.

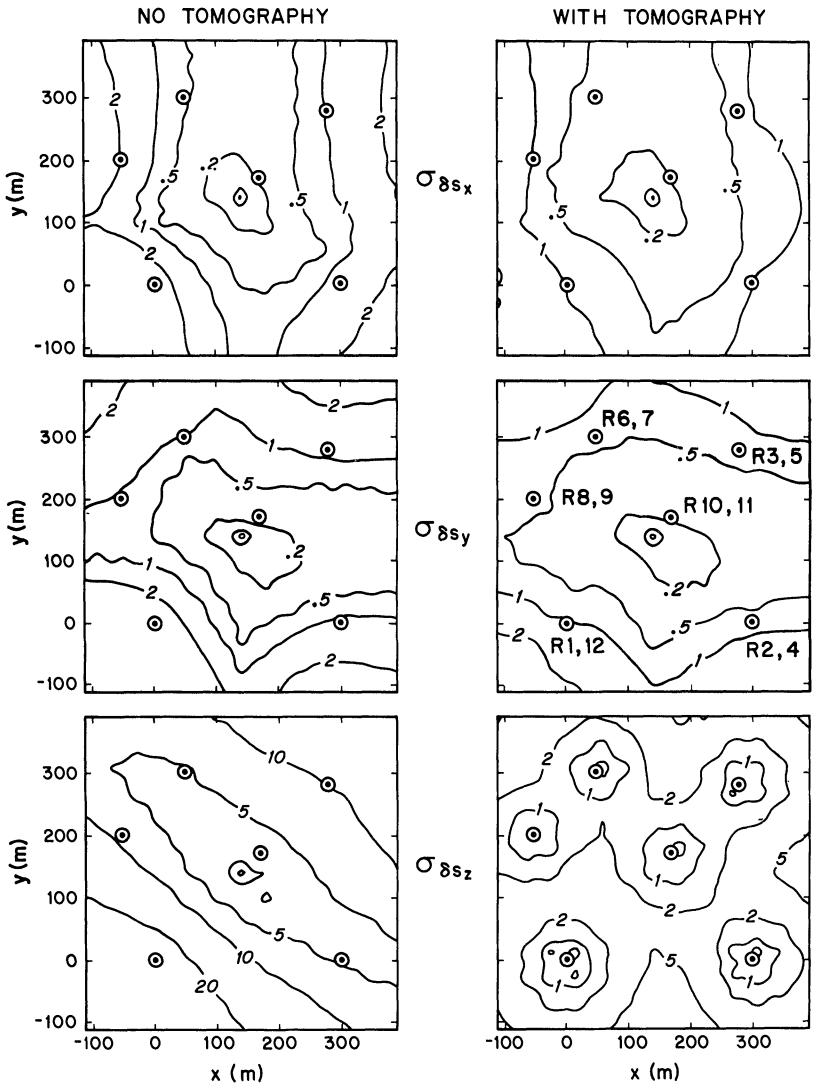

FIG. 8.—Error maps in the x, y, and z directions (*top*, *middle*, and *bottom*, respectively) for the location of one calling black-and-white colobus monkey at various places in the horizontal (x-y) plane at an elevation of 5 m without the use of tomography (left column) and with the use of tomography (right). Receivers are positioned at heights of 0 and 10 m at each position labeled in the horizontal plane. The standard deviation of the horizontal (vertical) components of the receiver positions is assumed to equal 0.02 (0.06) m. See the text for a further discussion. The contour levels are logarithmic with levels the same as for figure 6.

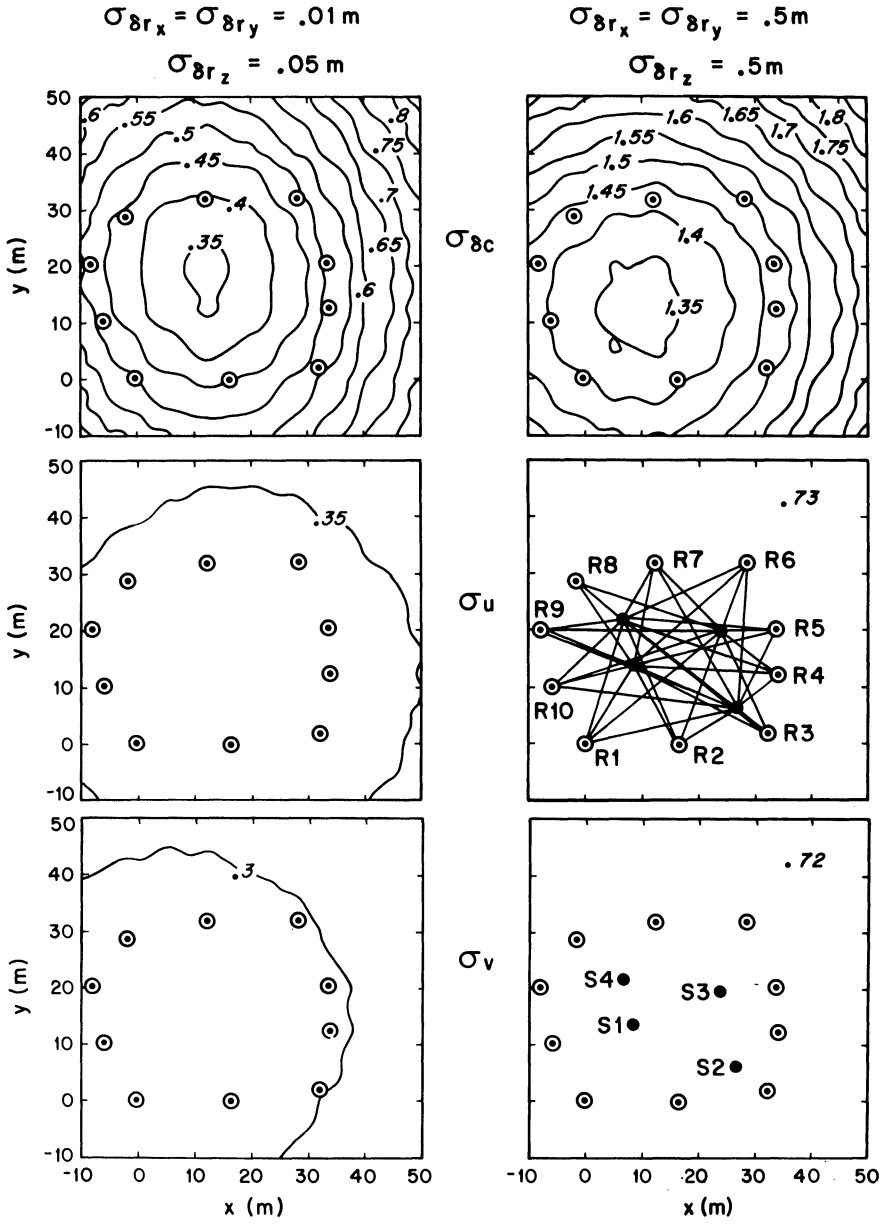

FIG. 9.—Ten receivers (R1–R10) are placed at elevations of 0, 7, 0, 7, 0, 7, 0, 7, 0, and 7 m, respectively, at the positions shown in the horizontal plane. Four animals (S1–S4) call simultaneously at elevations of 0, 0.5, 0.1, and 1 m, respectively, at the positions shown. The acoustic paths between all the sources and receivers are indicated. Tomographic error maps (one standard deviation) of sound speed and the two horizontal components of the wind (u and v) are shown at an elevation of 1 m for receivers whose relative positions are known precisely (left column, top labels) and for receivers whose relative positions are known less precisely (right column, top labels). The contours are drawn at intervals of 0.05 m/s. The a priori variability of the sound speed is taken to equal 5.9 m/s (about 10°C). The a priori variability of each component of the wind is taken to equal 0.81 m/s. The errors for the u and v components of the wind for the right-hand column are almost uniform and have the labeled values. See the text for the other environmental and acoustic data used to construct the maps.

sound speed. This occurs because the effect of the a priori wind variance on the TTD's is less than the effect of the a priori sound-speed variance for this example. In other words, it is more difficult to measure the relatively small effect of the winds in this case.

Marine Example

An array of five receivers was used to estimate how well the fish *Micropogon undulatus* (table 5, croaker) can be located when the relative positions of each coordinate of the receiver are known to an accuracy of 0.01 m (fig. 10). The ocean is assumed to be well mixed in the top 60 m, making the temperature nearly constant. The sound-speed fluctuations are set to zero. Each component of the horizontal current is assumed to have a scale of 300 m and has a standard deviation of 1 m/s. The vertical current is set to zero. The horizontal currents are uniform with depth. The acoustic noise level in shallow water is taken to be about 65 dB relative to 1 μPa at 1 m/Hz for a wind speed of about 2.5 m/s at the surface (Urick 1983). In the 300-Hz bandwidth of the croaker's call, $\mathcal{D}_{\text{noise}} \simeq 65 + 10 \cdot \log(300/1) = 90$ dB. For a source level of 136 dB and a range of about 100 m, the signal-to-noise ratio at a receiver is about 6 dB before correlation. Assuming the noise samples are uncorrelated, the signal-to-noise ratio after correlation is about 17 dB after cross-correlation. We conservatively assume that the signal-to-noise ratio is about 14 dB after correlation. Equation (42) is used to obtain the standard deviation of the TTD from noise, which is about 0.00011 s. The travel-time errors from ray bending and turbulence are negligible. Without tomography, the localization errors are 10–100 times those with tomography. The tomographic maps of the currents are poor because the travel-time fluctuations caused by the currents are small, as expected.

DISCUSSION

The equipment required to implement the techniques described above is readily available. For example, these algorithms have been implemented on a portable personal computer. The costs of multiple-channel recordings are modest.

Tomographic localizations are not always effective, but all instruments fail in particular conditions. Experiments are required to decide the utility of the theories presented here, particularly in the atmosphere where studies of pulsed propagation are rare. Some possible difficulties are considered.

1. When multipaths arrive closer together than the width of the peak of the cross-correlator output ($\approx 1/W$ s), then the precision of the travel-time differences may decrease. However, modern advancements in signal-processing theory exhibit the capability of resolving many multipaths that arrive within $1/W$ s (Feder and Weinstein 1988). A prudent choice of receiver location may minimize coincident multipaths.

2. When many animals call at about the same time, the output of the cross-correlator shows many peaks, but only one peak is due to any of the calling animals. The correct identification of cross-correlation peaks must be made for each pair of receivers. There are several ways to deal with this ambiguity. We believe that the details of the acoustic phases emitted from different members of

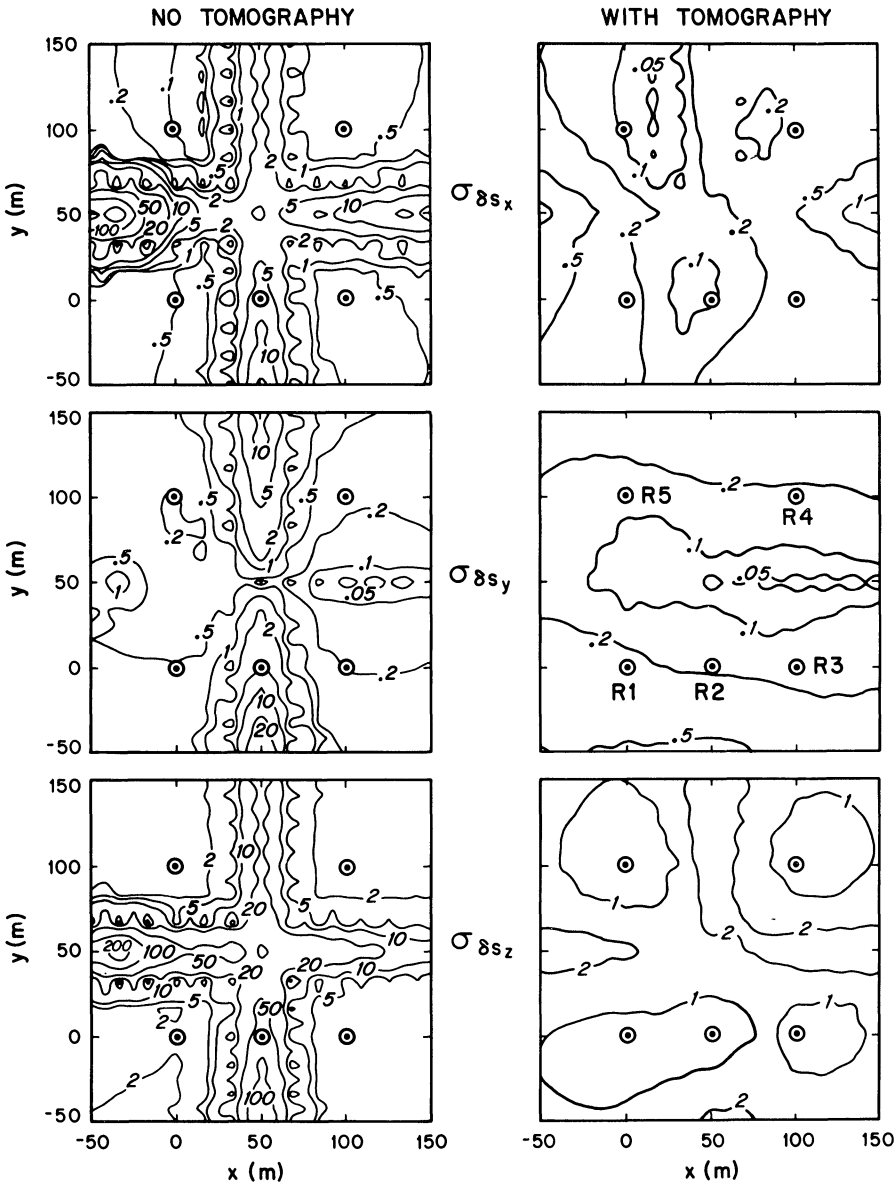

FIG. 10.—Error maps in the x, y, and z directions (*top*, *middle*, and *bottom*, respectively) for the location of a croaker fish (table 5) at various places in the horizontal (x-y) plane at a depth of 55 m without the use of tomography (left column) and with the use of tomography (right). Five receivers are placed at depths of 60, 60, 50, 50, and 60 m for R1–R5, respectively. The standard deviation of the horizontal and vertical components of the receiver positions is assumed to equal 0.01 m. See the text for a description of the scales assumed for the current and the sound-speed perturbation and for a description of the acoustic characteristics of the croaker. The contour levels are logarithmic with levels at 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200, and 500 m.

the same species often differ significantly from one another; thus, the cross-correlation of the calls from different animals is relatively small. The ambiguity can also be reduced if different animals call in nonoverlapping frequency bands. If the wrong set of peaks is used in the tomographic inversion, the mathematical result will be inconsistent with the data because the choice of peaks is not physical. Formally, this can always be checked by seeing if the model fits the data.

3. The acoustic-ray paths may not be stable in the presence of large atmospheric fluctuations. Large turbulent fluctuations may prevent reliable reception of the sounds. The stability of ray paths at long range in the ocean has been demonstrated elsewhere (Spiesberger et al. 1980).

Localizations of calling animals are significantly improved when the fluctuations of the acoustic environment and the receiver-position uncertainties are modeled rather than ignored in both the air and the water. Tomographic maps of the acoustic environment can map much of the acoustic variability in the atmosphere, but not in the ocean, when the sounds of calling animals are used as data. When tomographic maps of the terrestrial environment are made, it is necessary to reduce the errors of the relative positions of the receivers to about 1 cm.

SUMMARY

When monitored by several acoustic receivers, the sounds emitted by terrestrial and marine animals can be used to localize their positions. The localization can proceed quickly and automatically with the use of computers when cross-correlation of the acoustic records is used to optimally estimate the travel-time difference between signals at different receivers. The time-bandwidth product (call duration $\times$ acoustic bandwidth) of the call is an important parameter. When the time-bandwidth product of the acoustic call equals or exceeds 10 (which is the case for many animals), localization is possible even when the signal-to-noise ratios of the call are below the noise level, as long as cross-correlation is used to estimate the travel-time difference. Use of cross-correlation significantly increases the range at which a calling animal may be detected over that possible by visual inspection of acoustic records. The error maps for the localization can be improved by factors from 2 to 100 if tomographic techniques are used to account for the fluctuation in winds (currents) and sound speed and the errors in the receiver positions. In addition, useful tomographic maps of the atmospheric wind and the sound-speed fields can be obtained from the sounds emitted by the animals. Tomographic maps of the currents and the sound speed in water are much less precise because the fractional change in the acoustic propagation speed from currents and sound-speed fluctuations is much smaller in water than in the air. Studies in ecology, behavior, and conservation could be significantly enhanced by using acoustic tomography.

ACKNOWLEDGMENTS

We thank W. Watkins (Woods Hole Oceanographic Institution) for many useful comments. D. W. Thomson (Pennsylvania State University) suggested many

improvements in the description of the boundary layer and acoustic propagation in the atmosphere. We also extend our thanks to E. Terray (Woods Hole Oceanographic Institution), who provided guidance concerning the theory of acoustic scattering caused by turbulence. This work was partially supported by the education program at the Woods Hole Oceanographic Institution. Woods Hole contribution 6968.

APPENDIX A

Straight-Ray Approximation

Define z to be positive upward. The exact travel time between the source and the receiver is

$$T_1 = \int_{\Gamma_1} ds/c(\Gamma_1), \tag{A1}$$

where the ray path follows the contour given by Γ_1, the increment of path length on the contour is ds, and the speed of sound along the path is $c(\Gamma_1)$. The ray path is a function of the sound-speed field. For fields that vary only with the vertical coordinate, the ray path can be determined from Snell's law,

$$c_1/\cos\psi_1 = c/\cos\psi, \tag{A2}$$

where the angle of the ray with respect to the horizontal is ψ, and ψ_1 denotes the angle at the source.

The travel time along a *straight* path, Γ_0, between the source and the receiver is given by

$$T_0 = \int_{\Gamma_0} dx/c(\Gamma_0). \tag{A3}$$

The travel-time error resulting from the straight-ray-path approximation is

$$\delta T_1 = T_1 - T_0. \tag{A4}$$

We evaluate this error for a simple case.

Let the speed of sound vary in the vertical coordinate in a linear fashion, as in

$$c(z) = c_1 + c_z(z - z_1), \tag{A5}$$

where the speed at (x_1, z_1) is c_1 and $c_z = (c_2 - c_1)/(z_2 - z_1)$. For linear profiles, the acoustic paths are arcs of circles with radii given by equation (11) (Spiesberger 1985*a*). The solution for T_0 is

$$T_0 = \frac{1}{\sin\psi_0}\int_{z_1}^{z_2} \frac{dz}{c_1 + c_z(z - z_1)} = (c_z \sin\psi_0)^{-1} \ln(c_2/c_1). \tag{A6}$$

The solution for T_1 is

$$\begin{aligned} T_1 &= \int_{\psi_2}^{\psi_1} \frac{d\psi}{c(\psi)} = \frac{\cos\psi_1}{c_1}\int_{\psi_2}^{\psi_1} \frac{d\psi}{\cos\psi} \\ &= c_z^{-1} \ln\left[\left(\frac{1 - \tan\psi_0\tan\delta}{1 + \tan\psi_0\tan\delta}\right)\left(\frac{1 - \sin\psi_0\cos\delta + \cos\psi_0\sin\delta}{1 - \sin\psi_0\cos\delta - \cos\psi_0\sin\delta}\right)\right], \end{aligned} \tag{A7}$$

where δ is

$$\delta = \tan^{-1}[\epsilon\cot\psi_0/(2 + \epsilon)] \tag{A8}$$

and $\psi_1 = \psi_0 \pm \delta$, as in equation (12). A Taylor-series expansion for δT_1 can be obtained by expanding ϵ to the third order. The result is given by equation (14).

APPENDIX B

Acoustic Scattering in a Random Medium

The amount of wind variance per unit of change in the wavelength in the inertial subrange is approximately described by a power spectral density function given by

$$\mathcal{P}(\lambda) = a(1/\lambda)^{-5/3}, \tag{B1}$$

where inner scale $\leq \lambda \leq$ outer scale, the wavelength of the wind is λ, and a is a constant (Panofsky and Dutton 1984, p. 179). This spectrum describes the amount of energy at different-length scales in the wind, and it forms the basis for the theory that relates the random turbulence in the wind to the random behavior of acoustic signals that propagate through it.

The requirement for weak scattering in geometric optics for the spectrum of wind speed given by equation (B1) is derived from the requirement that

$$\Phi\Lambda^{5/12} < 1, \tag{B2}$$

where Φ and Λ are the strength and diffraction parameters (Flatte 1979, p. 93). For straight ray paths traveling through turbulence given by equation (B1),

$$\Phi^2 = q^2 R L_o \sigma_u^2/c^2, \tag{B3}$$

$$\Lambda = R/6L_o^2 q, \tag{B4}$$

where the acoustic wavenumber q is related to the acoustic frequency, f, by $q = 2\pi f/c$ (Flatte 1979, pp. 91, 92). In our notation, inequality (B2) reduces to

$$\sigma_u/c < 6^{5/12} q^{-7/12} L_o^{1/3} R^{-11/12}, \tag{B5}$$

which yields inequality (24). In this regime, the standard deviation of the phase at the receiver is given by Φ in units of radians per second. Equation (25) is derived from equation (B3).

The requirement for full saturation or strong scattering is

$$\Lambda\Phi > 1 \quad \text{and} \quad \Phi > 1, \tag{B6}$$

(Flatte 1979, pp. 128–129). Substituting equations (B3) and (B4) into inequalities (B6) yields inequalities (26) and (27).

APPENDIX C

Correlated Noise and the Cross-Correlator

When some of the different noise samples at a receiver are correlated with each other, the expression for the peak signal-to-noise ratio of the cross-correlator is given by equation (36), where

$$\zeta_n^2[j] = \sum_k^K \sum_l^K \left(\begin{array}{l} a_1^2 \overline{e_2[k-j]e_2[l-j]}\nu[k]\nu[l] + a_2^2 \overline{e_1[k]e_1[l]}\nu[k+p-j]\nu[l+p-j] \\ + a_1 a_2 \overline{e_2[k-j]e_1[l]}\nu[k]\nu[l+p-j] + a_1 a_2 \overline{e_1[k]e_2[l]}\nu[k+p-j]\nu[l] \\ + \overline{e_1[k]e_1[l]e_2[k-j]e_2[l]} + a_1 \overline{e_1[l]e_2[k-j]e_2[l-j]}\nu[k] \\ + a_2 \overline{e_1[k]e_1[l]e_2[l-j]}\nu[l] + a_1 \overline{e_1[k]e_2[k-j]e_2[l-j]}\nu[1] \\ + a_2 \overline{e_1[k]e_1[l]e_2[k-j]}\nu[l+p-j] \end{array} \right) \tag{C1}$$

(digital samples are in brackets). The terms that involve triple noise correlations in equation (C1) are negligible except when the source of the noise originates near the plane that is equidistant from the two receivers. Suppose these terms are negligible, and consider the noise near the peak of the cross-correlator ($j \simeq p$). Then, if

$$\overline{e_1[k]e_2[l - p]} \ll \overline{e_1[k]e_2[l]}$$

for $k \simeq l$ and if

$$\overline{e_i[k]e_i[l]} \ll \nu[k]\nu[l]$$

for $k \simeq l$ (large signal-to-noise ratio), then

$$\zeta_n^2 \simeq \sum_k^K \sum_l^K (a_1^2\overline{e_2[k]e_2[l]}\nu[k]\nu[l] + a_2^2\overline{e_1[k]e_1[l]}\nu[k]\nu[l]). \tag{C2}$$

In equation (C2), the statistics of the noise are assumed to be independent of the sample number. Equation (C2) is used to evaluate the cross-correlation gain for a single recording of a call where we set $a_1 = a_2$ and $e_1[k] = e_2[k]$, for simplicity. The gain of the cross-correlator is no longer proportional to K because ζ_n^2 increases with K also.

APPENDIX D

Linear and Nonlinear Decomposition of the Differential Travel-Time Differences

In equation (72) we have

$$\int_{s_1}^{r_1(i)} [n_0 + \delta n(\mathbf{l}_i)]\, dl_i = \mathcal{T}_1 + \mathcal{T}_2, \tag{D1}$$

where

$$\mathcal{T}_1(i) = n_0 \int_{s_0 + \delta s}^{r_0(i) + \delta r(i)} dl_i, \tag{D2}$$

$$\mathcal{T}_2(i) = \int_{s_0 + \delta s}^{r_0(i) + \delta r(i)} \delta n(\mathbf{l}_i) dl_i, \tag{D3}$$

and dl_i is an increment of the acoustic path between the actual position of the source and the ith receiver. Positions along the ray path are given by $\mathbf{l}_i$. Usually, this path is curved because of refraction. Using the calculus of variations, it is possible to show that changes in the acoustic-ray path relative to the reference-ray path (a straight line) relate the sound-slowness and wind (current) fields in a nonlinear manner to the travel-time differences (Spiesberger 1985*a*). Then,

$$\mathcal{T}_1(i) = n_0 \|\mathbf{r}_1(i) - \mathbf{s}_1\| + t_{nl}, \tag{D4}$$

where t_{nl} is the constituent of $\mathcal{T}_1$ involving changes in the acoustic path from the straight line from $\mathbf{s}_1$ to $\mathbf{r}_1$. Now,

$$\|\mathbf{r}_1(i) - \mathbf{s}_1\| = \|[\mathbf{r}_0(i) + \delta\mathbf{r}(i)] - (\mathbf{s}_0 + \delta\mathbf{s})\|, \tag{D5}$$

and to the first order in $\delta\mathbf{r}(i)$ and $\delta\mathbf{s}$,

$$\|\mathbf{r}_1(i) - \mathbf{s}_1\| \simeq \|\mathbf{r}_0(i) - \mathbf{s}_0\| + F(i), \tag{D6}$$

where

$$F(i) = \frac{[r_{0x}(i) - s_{0x}][\delta r_x(i) - \delta s_x] + [r_{0y}(i) - s_{0y}][\delta r_y(i) - \delta s_y]}{\|\mathbf{r}_0(i) - \mathbf{s}_0\|} + \frac{[r_{0z}(i) - s_{0z}][\delta r_z(i) - \delta s_z]}{\|\mathbf{r}_0(i) - \mathbf{s}_0\|}, \tag{D7}$$

$i = 2, 3, 4, \ldots, N_r$, and $\|\mathbf{r}_0(i) - \mathbf{s}_0\| \neq 0$. When $\|\mathbf{r}_0(i) - \mathbf{s}_0\| = 0$, there are no first-order terms in $\delta\mathbf{r}$ and $\delta\mathbf{s}$. The expression for $\mathcal{T}_1$ is now

$$\mathcal{T}_1(i) = n_0\|\mathbf{r}_0(i) - \mathbf{s}_0\| + n_0F(i) + t_{nl}, \tag{D8}$$

where t_{nl} absorbs all the higher-order terms in $\delta\mathbf{r}$ and $\delta\mathbf{s}$ and absorbs the nonlinearity caused by the difference between the actual and reference acoustic paths.

The expression for $\mathcal{T}_2$ can be written as

$$\mathcal{T}_2 = \int_{\mathbf{s}_0}^{\mathbf{r}_0(i)} \delta n(\xi_i)\, d\xi_i + t_{nl}, \tag{D9}$$

where the first term on the right-hand side of equation (D9) is linearly related to the data and the second term absorbs all the nonlinearities. The increment of distance between the straight line joining $\mathbf{s}_0$ to $\mathbf{r}_0(i)$ is denoted by $d\xi_i$.

Similarly, equation (72) contains the term

$$\int_{\mathbf{s}_1}^{\mathbf{r}_1(i)} \mathbf{u}(\mathbf{l}_i) \circ \mathbf{dl}_i = \int_{\mathbf{s}_0}^{\mathbf{r}_0(i)} \mathbf{u}(\boldsymbol{\xi}_i) \circ \mathbf{d}\boldsymbol{\xi}_i + t_{nl}, \tag{D10}$$

where the second term on the right-hand side of equation (D10) absorbs the nonlinear terms. The incremental vector, $\mathbf{d}\boldsymbol{\xi}_i$, points along the straight line from the reference position of the source to the reference position of the receiver.

By subtracting $\tau_0(i)$ in equation (70) from $\tau_1(i)$ in equation (72) and retaining up to first-order terms in the models, we get the linearization of the forward problem as written in equation (78).

LITERATURE CITED

Alexander, R. 1960. Sound communication in Orthoptera and Cicadidae. Pages 38–92 *in* Lanyon and Tavolga 1960.

American National Standards Institute. 1978. Method for the calculation of the absorption of sound by the atmosphere. Acoustical Society of America, ANSI, New York.

Attenborough, K. 1985. Acoustical impedance models for outdoor ground surfaces. J. Sound Vib. 99:521–544.

Backus, G. E., and F. Gilbert. 1968. The resolving power of gross earth data. Geophys. J. R. Astron. Soc. 16:169–205.

Bogert, C. M. 1960. The influence of sound on the behavior of amphibians and reptiles. Pages 137–320 *in* Lanyon and Tavolga 1960.

Bretherton, F. P., R. W. Davis, and C. B. Fandry. 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. Deep-Sea Res. 23:559–582.

Brigham, E. O. 1974. The fast Fourier transform. Prentice-Hall, Englewood Cliffs, N.J.

Burrus, C. S., and J. F. Parks III. 1985. DFT/FFT and convolution algorithms. Wiley, New York.

Chen, C. T., and F. J. J. Millero. 1977. Speed of sound in seawater at high pressures. J. Acoust. Soc. Am. 62:1129–1135.

Clark, C. W. 1980. A real-time direction finding device for determining the bearing to the underwater sounds of southern right whales *Eubalaena australis*. J. Acoust. Soc. Am. 68:508–511.

Cornuelle, B. 1982. Acoustic tomography. IEEE (Inst. Electr. Electron. Eng.) Trans. Geosci. Remote Sensing GE-20(3):326–332.

Del Grosso, V. A. 1974. New equation for the speed of sound in natural waters (with comparison to other equations). J. Acoust. Soc. Am. 56:1084–1091.

Duda, T. F., S. M. Flatte, and D. B. Creamer. 1988. Modeling meter-scale acoustic intensity fluctuations from oceanic fine structure and microstructure. J. Geophys. Res. 93(C5):5130–5142.

Dumortier, B. 1963. Ethological and physiological study of sound emissions in Arthropoda. Pages 583–654 *in* R. G. Busnel, ed. Acoustic behaviour of animals. Elsevier, New York.

Emmanuel, C. B. 1972. Observations of Helmholtz waves in the lower atmosphere with an acoustic sounder. Ph.D. diss. Colorado State University, Fort Collins.

Evans, G. C. 1939. Ecological studies on the rain forest of southern Nigeria. J. Ecol. 27:436–482.
Eyring, C. F. 1946. Jungle acoustics. J. Acoust. Soc. Am. 18:257–270.
Feder, M., and E. Weinstein. 1988. Parameter estimation of superimposed signals using the EM algorithm. IEEE (Inst. Electr. Electron. Eng.) Trans. Acoust. Speech Signal Process. 36(4): 477–489.
Fish, J., and W. H. Mowbray. 1970. Sounds of western North Atlantic fishes. Johns Hopkins Press, Baltimore, Md.
Fish, J., and C. Turl. 1976. Acoustic source levels of four species of small whales. Nav. Undersea Center (San Diego, Calif.) Tech. Pap. 547:1–14.
Flatte, S. M., ed. 1979. Sound transmission through a fluctuating ocean. Cambridge University Press, Cambridge.
Garrett, C., and W. Munk. 1972. Space-time scales of internal waves. Geophys. Fluid Dyn. 2:225–264.
Gerhardt, C. H. 1975. Sound pressure levels and radiation patterns of the vocalizations of some North American frogs and toads. J. Comp. Physiol. A, Sens. Neural Behav. Physiol. 102:1–12.
Ghil, M., and S. E. Cohn. 1983. Applications of sequential estimation to data assimilation. Pages 341–356 *in* Large-scale oceanographic experiments in the World Climate Research Programme. WCRP Publ. Ser. 1, Vol. II. World Meteorological Organization/International Council of Scientific Unions, Joint Science Committee (WMO/ICSU). Geneva, Switzerland.
Griffin, D. R. 1958. Listening in the dark. Yale University Press, New Haven, Conn.
Hall, F. F., Jr., J. G. Edinger, and W. D. Neff. 1975. Convective plumes in the planetary boundary layer, investigated with an acoustic echo sounder. J. Appl. Meteorol. 14:513–523.
Harris, C. M. 1966. Absorption of sound in air versus humidity and temperature. J. Acoust. Soc. Am. 40:148–159.
———. 1967. Absorption of sound in air versus humidity and temperature. National Aeronautics and Space Administration, Contractor Rep. CR-647. Washington, D.C.
Haugen, D. A., ed. 1973. Workshop on micrometeorology. American Meteorological Society, Boston.
Helstrom, C. W. 1975. Statistical theory of signal detection. Pergamon, New York.
Jackson, J. D. 1975. Classical electrodynamics. 2d ed. Wiley, New York.
Lanyon, W. E. 1960. The ontogeny of vocalizations in birds. Pages 321–347 *in* Lanyon and Tavolga 1960.
Lanyon, W. E., and W. N. Tavolga, eds. 1960. Animal sounds and communication. Symposium on animal sounds and communication, Indiana University, Bloomington, 1958. AIBS Publ. 7. American Institute of Biological Sciences, Washington, D.C.
Lawson, C. L., and R. J. Hanson. 1974. Solving least squares problems. Prentice-Hall, Englewood Cliffs, N.J.
Lighthill, J. 1978. Waves in fluids. Cambridge University Press, New York.
Mackenzie, K. V. 1981. Nine-term equation for sound speed in the oceans. J. Acoust. Soc. Am. 70:807–812.
Magyar, I., W. M. Schleidt, and D. Miller. 1978. Localization of sound producing animals using the arrival time differences of their signals at an array of microphones. Experientia 34:676–677.
Marler, P. 1960. Bird songs and mate selection. Pages 348–367 *in* Lanyon and Tavolga 1960.
Marten, K., and P. Marler. 1977. Sound transmission and its significance for animal vocalization. I. Temperate habitats. Behav. Ecol. Sociobiol. 2:271–290.
Marten, K., D. Quine, and P. Marler. 1977. Sound transmission and its significance for animal vocalization. II. Tropical forest habitats. Behav. Ecol. Sociobiol. 2:291–302.
Medwin, H. 1975. Speed of sound in water: a simple equation for realistic parameters. J. Acoust. Soc. Am. 58:1318–1319.
Morton, E. S. 1970. Ecological sources of selection on avian sounds. Ph.D. diss. Yale University, New Haven, Conn.
Moulton, J. M. 1963. Acoustic behavior of fishes. Pages 655–693 *in* R. G. Busnel, ed. Acoustic behaviour of animals. Elsevier, New York.
Munk, W. 1974. Sound channel in an exponentially stratified ocean, with application to SOFAR. J. Acoust. Soc. Am. 55:220–226.
Munk, W., and C. Wunsch. 1979. Ocean acoustic tomography: a scheme for large scale monitoring. Deep-Sea Res. 26:123–161.

Oppenheim, A. V., and R. W. Schafer. 1975. Digital signal processing. Prentice-Hall, Englewood Cliffs, N.J.

Panofsky, H. A., and J. A. Dutton. 1984. Atmospheric turbulence. Wiley, New York.

Parker, R. L. 1977. Understanding inverse theory. Annu. Rev. Earth Planet. Sci. 5:35–64.

Pierce, A. D. 1981. Acoustics: an introduction of its physical principles and applications. McGraw-Hill, New York.

Rader, C. M. 1970. An improved algorithm for high speed autocorrelation with application to spectral estimation. IEEE (Inst. Electr. Electron. Eng.) Trans. Audio Electroacoust. AU-18:439–441.

Rayleigh, J. W. S. 1945. The theory of sound. Vols. 1, 2. 2d ed. Dover, New York.

Rogers, R. R. 1976. A short course in cloud physics. Pergamon, New York.

Saby, J. S., and H. A. Thorpe. 1946. Ultrasonic ambient noise in tropical jungles. J. Acoust. Soc. Am. 18:271–273.

Spiesberger, J. L. 1985*a*. Ocean acoustic tomography: travel time biases. J. Acoust. Soc. Am. 77:83–100.

———. 1985*b*. Gyre-scale acoustic tomography: biases, iterated inversions, and numerical methods. J. Geophys. Res. C, Oceans 90:11869–11876.

Spiesberger, J. L., R. C. Spindel, and K. Metzger. 1980. Stability and identification of ocean acoustic multipaths. J. Acoust. Soc. Am. 67:2011–2017.

Stull, R. B. 1988. An introduction to boundary layer meteorology. Kluwer Academic, Boston.

Thomas, J. A., and V. B. Kuechle. 1982. Quantitative analysis of Weddell seal (*Leptonychotes weddelli*) underwater vocalizations at McMurdo Sound, Antarctica. J. Acoust. Soc. Am. 72:1730–1738.

Thorp, W. H. 1965. Deep-ocean sound attenuation. J. Acoust. Soc. Am. 38:648–654.

Urick, R. J. 1983. Principles of underwater sound. 3d ed. McGraw-Hill, New York.

Walker, R. A. 1963. Some intense, low-frequency, underwater sounds of wide geographic distribution, apparently of biological origin. J. Acoust. Soc. Am. 35:1816–1824.

Waser, P. M., and M. S. Waser. 1977. Experimental studies of primate vocalization: specializations for long-distance propagation. Z. Tierpsychol. 43:239–263.

Watkins, W. A., and W. E. Schevill. 1971. Four hydrophone array for acoustic three-dimensional location. Woods Hole Oceanogr. Inst. Tech. Rep. 71–60. Bar Harbor, Me.

———. 1972. Sound source location by arrival times on a non-rigid three-dimensional hydrophone array. Deep-Sea Res. 19:691–706.

———. 1979. Distinctive characteristics of underwater calls of the harp seal, *Phoca groenlandica*, during the breeding season. J. Acoust. Soc. Am. 66:983–988.

Watkins, W. A., and D. Wartzok. 1985. Sensory biophysics of marine mammals. Mar. Mamm. Sci. 1(3):219–260.

Watkins, W. A., P. Tyack, and K. E. Moore. 1987. The 20-Hz signals of finback whales (*Balaenoptera physalus*). J. Acoust. Soc. Am. 82:1901–1912.

Wiley, R. H., and D. G. Richards. 1982. Adaptations for acoustic communication in birds: sound transmission and signal detection. Pages 132–181 *in* D. E. Kroodsma and E. H. Miller, eds. Acoustic communication in birds. Vol. 1. Academic Press, New York.