DETECTION AND LOCATION CAPABILITIES OF MULTIPLE INFRASOUND ARRAYS

Robert H. Shumway

Sponsored by The Defense Threat Reduction Agency
Arms Control Technology Division
Nuclear Treaties Branch

Contract No. DTRA01-00-C0082

ABSTRACT

We develop an integrated approach to estimating wave-number parameters form a collection of local arrays and fusing these estimators to obtain an accurate location, along with its uncertainty ellipse. A small-array theory has been given in previous work that characterizes the detection probabilities and large sample variances of the local optimal detectors and we show how to incorporate the results of the local-array performance into a global network assessment. The wave-number estimators and their uncertainties are used as input to a Bayesian nonlinear regression that produces fusion ellipses for event locations using probable configurations of detecting stations in the global infrasound array proposed by the Prototype International Data Center (PIDC). The network capability is characterized as a function of separate local-array characteristics, including signal-to-noise ratios, bandwidth, array geometry, local correlation and coherent interfering signals.

Key Words: *Maximum likelihood, Bayes, signal estimation and detection, nonlinear estimation, regression, F-Statistics.*

OBJECTIVE

We are beginning a study to characterize the detection and location capabilities of the world-wide infrasound arrays configured for the Prototype International Data Center (PIDC) network. The study will utilize results of our previous work (see, for example, Blandford, 1997, Shumway et al, 1999) to characterize the detection and wave-number estimation capabilities of local arrays. In that study (Shumway and Kim, 1999), we have also developed *fusion posterior probability ellipses* for location that incorporate the two dimensional covariance matrices of the wave-number vector, e.g., velocity and azimuth, for each of the local arrays. The theory allows prediction of of a fusion ellipse for arbitrary configurations of recording arrays, under assumed local signal-to-noise ratios, array geometry and signal decorrelations. Extending these results to a full analysis of the detection and location capabilities of the PIDC network requires a two-stage approach, consisting first of characterizing local-array detection and estimation uncertainty and then incorporating these results into a detailed study of the detection and estimation capabilities of the full network.

Note that the theory is in place for following the above procedure through for perfectly correlated signals and uncorrelated noise. A generalization being developed in this study is the extension of the local-array performance capabilities to the case where there are coherent interfering signals or noises generated by such local features as wind or microbaroms or by multiple arrivals occurring in the same time window. If the local features are not readily available, this requires a *multiple stochastic signal* approach, where the interfering signals may have either known or unknown coherence properties. We are covering both the case where the interfering signals propagate according to some unknown velocity and azimuth and the case where the coherence structure must be estimated from the currently observed data. Extension of conventional stochastic signal models to the multiple signal case has been considered by Shumway (1983, 1988) and Shumway and Stoffer (2000). The above study should complete the characterization of *local-array performance* for infrasound.

The second stage will incorporate the results of the local-array performance into a summary *global-network performance* in terms of the uncertainty regions predicted by fusion ellipses. Usually, only a few arrays will record any given event so that local detection probabilities must be combined to develop an overall average fusion ellipse. In this study, we are expanding conventional network detection, as set out in Wirth (1976), to develop an *expected location contour map* of the world that could be expected for the proposed PIDC network. The network capability will be a function of separate inputs for each array, involving signal-to-noise ratios, bandwidth parameters, array geometry, local correlation and coherent interfering signals.

RESEARCH ACCOMPLISHED

Local-Array Estimation and Detection

In the case of a local array measuring a propagating plane wave from an infrasonic signal, we are generally dealing with a small number of sensors and a model of the form

$$y_j(t) = s_j[t - T_j(\boldsymbol{\theta})] + n_j(t), \tag{1}$$

corresponding to expressing the received process $y_j(t)$ as a signal $s_j(t)$ imbedded in in a noise series $n_j(t)$ on each of $j = 1, 2, \ldots, N$ sensors. The signal incurs a time delay $T_j(\boldsymbol{\theta})$ at the j^{th} sensor where $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ is a vector of wave-number parameters that are nonlinearly related to velocity and azimuth and to the source location of the signal, say $\boldsymbol{x} = (x_1, x_2)'$. It is the vector $\boldsymbol{x}$ that we are primarily interested in estimating, where x_1 and x_2 will denote latitude and longitude in km, based on some arbitrary origin. In Shumway et al (1999), we have assumed that the signals had some correlation structure that could either be estimated from data or could conform to the distance, frequency correlation structure advanced by Mack and Flinn (1971). If the signals in (1) were perfectly correlated, replacing $s_j(t)$ by the common signal $s(t)$ simplifies results considerably, as shown in Shumway et al (1999).

As we will see later, what is needed is a theory that will lead to an optimum detector and a signal detection probability for each local array. We also need an estimator for the wave-number vector $\boldsymbol{\theta}$ (velocity and azimuth) and a variance-covariance matrix that will depend on bandwidth, signal-to-noise ratio and the geometry of the local array. This has been done in Shumway et al (1999) under the two signal-correlation structures mentioned above. Then, noting that the true wave-number vector is a known nonlinear function of location, say $\boldsymbol{\theta}_k(\boldsymbol{x})$ for the k^{th} array, $k = 1, 2, \ldots, n$ gives a set of values from which we may infer location and a confidence or posterior probability ellipse for that location. This is covered in Shumway and Kim (1999) for the case of a perfectly correlated signal.

In the new work, we are extending the above theory to the case where one or more interfering signals are present. They may be generated by consistent wind effects (see Hedlin, 1998) or by ground, atmospheric, stratospheric or thermispheric reflections (see Armstrong, 1998). Many times, separate phases can be analyzed in isolation but it is also possible that more information will accrue from a broad time window that increases the time-bandwidth product. For a simple double-signal example, suppose that the received process is of the form

$$y_j(t) = s[(t - T_j(\boldsymbol{\theta})] + u[t - D_j(\boldsymbol{\gamma})] + n_j(t), \tag{2}$$

where $u(t)$ represents a second interfering signal with an unknown set of time delays, depending on a second set of wave-numbers $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)'$ corresponding to an interfering signal, possibly generated by wind (see Hedlin et al, 1998). While there there are models and computer code available (see, for example, Dighe et al, 1998) for generating wind profiles at a given site as a function of weather conditions, the development of a theory for a general interfering signal with unknown velocity and azimuth characteristics would still be a useful exercise. In this proposal, we discuss isolating either a propagating signal or a signal with completely unknown coherence structure. It should be noted that the restriction in (2) to just two unknown signals is not necessary.

Signal Detection

In order to handle the optimum signal detection problem, we review the results for a single signal with unknown wave-number parameters that obtains when $s_j(t) = s(t)$ in (1), namely, the case of a perfectly correlated signal. The requisite theory follows from writing the model in the frequency domain and applying the Whittle likelihood to obtain the optimum detector for testing the absence of a signal (see Shumway et al, 1999) for details. For the likelihood ratio detector, we obtain a ratio of beam-power to noise power in the neighborhood of a center frequency that is expressed as an F-Statistic of the form

$$F(\hat{\boldsymbol{\theta}}) = \frac{N^{-1} B(\hat{\boldsymbol{\theta}})}{SSE(\hat{\boldsymbol{\theta}})}, \tag{3}$$

introduced by Shumway (1983), where the numerator has the well-known beam-power and the denominator is an estimated noise power, obtained by summing the squared deviations of each channel from the overall beam. The parameter $\hat{\boldsymbol{\theta}}$ is the maximizer of the beam-power and the log likelihood. It is commonly computed by searching over the two-dimensional space spanned by the vector or over some suitable restricted space. The distribution of the detector (3) is proportional to an F-statistic with $2BT$ and $2BT(N-1)$ degrees of freedom, where BT is the time-bandwidth product. That is

$$F(\hat{\boldsymbol{\theta}}) \sim (1 + rN) F_{2BT, 2BT(N-1)}, \tag{4}$$

where $\sim$ denotes *is distributed as* and

$$r = \frac{P_s}{P_n} \tag{5}$$

denotes the signal-to noise ratio, computed as the ratio of the signal power spectrum P_s to the noise power spectrum P_n.

In the sequel, we propose extending the above theory to the model (2), possibly containing the interfering signal. Portions of the requisite theory are available in Shumway and Stoffer (2000), where tests of hypotheses involving multiple signals are developed from classical regression theory. It is clear that a stepwise approach will be involved, with searches over $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ needed to find maximum likelihood estimators for those two parameter vectors. An approach proposed for the fixed multiple-signal case is given in Shumway (1983) and it is proposed that the approach be extended to the random signal case.

In the case that the simpler model (2) does not hold, we are also investigating the possibility of estimating the propagation structure of the interfering interfering signals, using models of the form

$$y_j(t) = s[(t - T_j(\boldsymbol{\theta})] + \sum_{k=1}^{q} \lambda_{jk}(t-\tau)u_k(\tau) + n_j(t), \tag{6}$$

where the unknown filters $\lambda_{jk}(t)$ can be estimated using a generalization of factor analysis in the frequency domain. For example, taking discrete Fourier transforms on both sides of (6) and vectorizing leads to a frequency domain model of the form

$$\boldsymbol{Y} = \boldsymbol{A}S + \Lambda\boldsymbol{U} + N, \tag{7}$$

which can be recognized as factor analysis in the frequency domain and the estimated Λ may be a reasonable approximation to some generalized modal structure in the frequency domain. The vector

$$\boldsymbol{A} = (e^{2\pi i\nu T_1(\boldsymbol{\theta})}, \ldots, e^{2\pi i\nu T_N(\boldsymbol{\theta})})'$$

still contains the propagating plane wave. Of course, the bandwidth requirements of such a procedure are fairly stringent and a limited number of *modal* components U_j can be fitted for a small array.

Optimal Estimation of Wave-number Parameters

Although the signal detection solution discussed in the previous section required maximum likelihood estimators for the wave-number parameters, obtained by maximizing the beam-power in the uncorrelated signal case, it is convenient to give a separate discussion of the problem of estimating the wave-number (velocity and azimuth) parameters of a given local array. The main reason for needing this additional discussion is to specify the array-dependent uncertainty of the estimated wave-number parameters which serve as critical inputs for the location procedure. For any given source, it is clear that different recording arrays will have wave-number estimators with different variances and covariances. Providing estimators of the differential that can be attributed to differing values for signal-to-noise ratios, bandwidth and geometry of the array will be important for determining the uncertainty region of the location procedure. It should be noted here that additional variability due to geophysical causes other than those given above will be incorporated later, in part, through a variance scaling factor introduced in the location model.

For the uncorrelated signal case, Shumway et al (1999) have obtained the large-sample covariance matrix (Cramer-Rao Lower Bound) of the two-dimensional estimated wave-number vector $\hat{\boldsymbol{\theta}}$ as

$$\Sigma \approx \frac{1}{(2\pi)^2}\frac{1}{2BT}\frac{1}{rN}\left(1 + \frac{1}{rN}\right)R^{-1}, \tag{8}$$

where, in addition to depending on the signal-to-noise ratio r, the covariance matrix depends on the geometry of the array, as determined by the sample covariance matrix of the sensor coordinates $\boldsymbol{r}_j = (r_{j1}, r_{j2})', j - 1, \ldots, N$, say

$$R = \frac{1}{N}\sum_{j=1}^{N}(\boldsymbol{r}_j - \bar{\boldsymbol{r}})(\boldsymbol{r}_j - \bar{\boldsymbol{r}})'. \tag{9}$$

Equations (8) and (9) exhibit several advantages that accrue when formulating location in terms of the wave-number parameters θ_1 and θ_2 rather than in terms of velocity and azimuth, which are more complicated nonlinear functions of the location. Furthermore, for typical arrays, such as the ones illustrated later, coordinates can be chosen so that the off-diagonal elements of R are zero, leading to uncorrelated estimators for the two wave-number parameters. In the case of a correlated signal, this is not the case and Shumway et al (1999) show what (8) becomes when one specifies the coherence matrix of the signals in (1) but still uses the simple beam-forming estimator for $\boldsymbol{\theta}$. Detailed estimators for the variances of the azimuth estimators as a function of array-baseline size have been computed in Shumway et al (1999). Those computations show that the increase in variance attributable to signal decorrelation is not very substantial over the frequency and array baseline distances contemplated for PIDC (see, also Armstrong, 1998).

In order to cover wave-number estimation and uncertainty for multiple signal models of the form (2), we will need to extend the approach in Shumway et al (1999) to that more complicated model. The benefits from such an extension will be estimators for the wave-number vector $\boldsymbol{\theta}$ that are uncontaminated by the interfering signal $u(t)$. We will also obtain estimators for the variance-covariance matrix of $\hat{\boldsymbol{\theta}}$. As a by-product of the estimation, there will be separate estimators of the wave-number parameters of the interfering signal and its waveform. it should be noted that the estimator for the primary signal will no longer be the straight beam and that the test statistic will no longer be a monotone function of the the beam power. Some examples of multiple signal estimation in the deterministic signal case are given in Shumway (1983) and extensions to the stochastic signal case are discussed in Shumway and Stoffer (2000).

RESEARCH ACCOMPLISHED

Global Location Capabilities

Integrating or fusing data from the single-array sources into a best overall location, with an uncertainty region provided, will be an important aspect of evaluating the predicted performance of the PIDC network. Hence, we consider a methodology for using the information developed in the detection and estimation portions of this paper for estimating the location vector $\boldsymbol{x} = (x_1, x_2 2)'$. In general, we propose a model of the form

$$\hat{\boldsymbol{\theta}}_k = \boldsymbol{\theta}_k(\boldsymbol{x}) + \boldsymbol{e}_k, \tag{10}$$

where $\hat{\boldsymbol{\theta}}_k = (\hat{\theta}_{1k}, \hat{\theta}_{2k})', k = 1, \ldots, n$ is the estimated wave-number vector, as computed from maximizing the beam-power, or F-statistic (3) at the k^{th} array and

$$\boldsymbol{\theta}_k(\boldsymbol{x}) = \frac{\nu}{V} \frac{\boldsymbol{x} - \boldsymbol{c}_k}{||\boldsymbol{x} - \boldsymbol{c}_k||} \tag{11}$$

gives the theoretical connection between the wave-number parameters and location. In (11), ν is the center frequency, V is velocity and $\boldsymbol{c}_k = (c_{1k}, c_{2k})'$ denote the coordinates of the k^{th} array ($||\boldsymbol{a}|| = \sqrt{a_1^2 + a_2^2}$ is the norm of the vector $\boldsymbol{a}$). It can generally be assumed that velocity is known or can be inferred from the wave-number plot. It should be noted that there are often separate phases at the same array that may have their own estimated $\hat{\boldsymbol{\theta}}_k$ and these are included under those that possibly contribute to the model (11). Under certain conditions, there may also be estimators of travel times $\hat{t}_k$, where $t_k(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{c}_k||/V$, perhaps from cross-correlation or other means, that could be added to the observations on location $\boldsymbol{x}$.

A possible assumption for the bivariate error terms in (10) is that they are independent and identically distributed with mean zero and 2×2 covariance matrix

$$\text{cov } \boldsymbol{e}_k = \sigma^2 \Sigma_k, \tag{12}$$

where the matrix Σ_k comes from (8), specialized to the k^{th} array. Note that the components of (8) will vary according to the array size and geometry, signal-to-noise ratio and the time-bandwidth product. The scaling variance σ^2 is to account for additional variability from geophysical sources or from the observed error in a particular event location. If there are consistent biases associated with particular regions or subsets of arrays, constant correction terms can be added to the defining Equation (10). If a number of events are available, the correction terms may even be estimated by least squares using consistent source-receiver pairs.

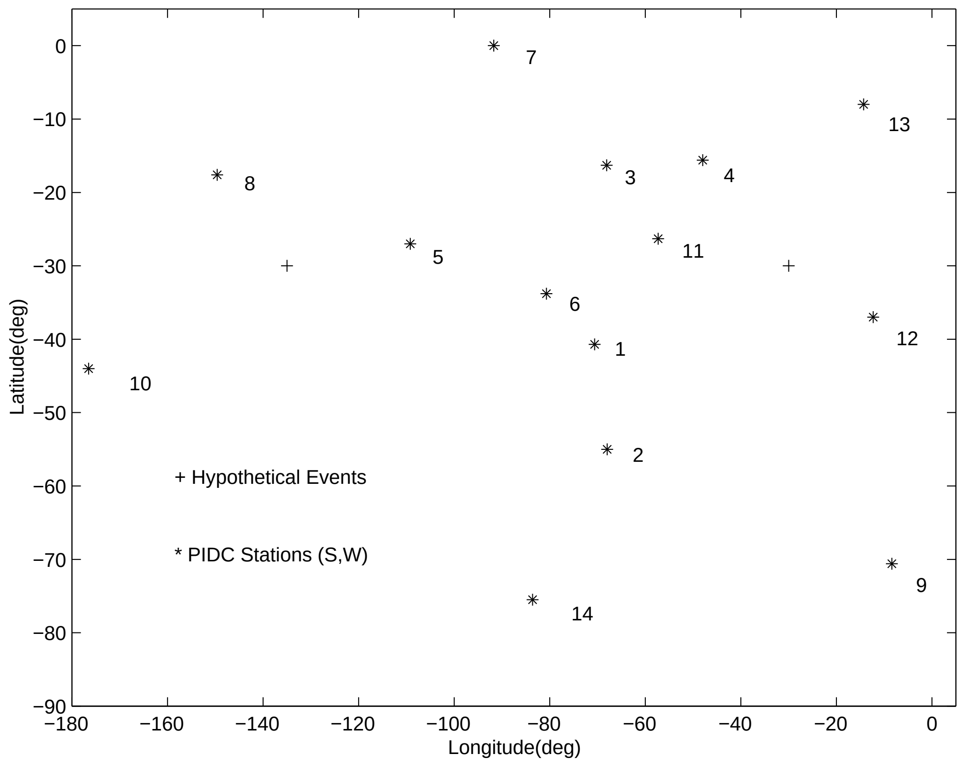

Figure 1: Proposed PIDC infrasound arrays in the southwest quadrant and two hypothetical event locations

As an example of a small demonstration set, we consider Figure 1, which shows the 14 PIDC array centers $\boldsymbol{c}_k$ proposed for the southwest quadrant of the world-wide network. For purposes of illustration, we also show two hypothetical events that might generate observed wave-number parameters $\hat{\boldsymbol{\theta}}_k, k = 1, \ldots, n$ for incorporation into the location model defined in (11) and (12). Later, we discuss an approach to combining or fusing the wave-number parameters into an overall location, based on a given fixed set of recording arrays and their characteristics, as defined by the signal-to-noise ratio r, the time-bandwidth product BT and the array geometry, as it is expressed in terms of R and N in (8).

The solution for a fixed configuration of recording arrays is not the only factor of interest since there will be multiple possible recording configurations possible for any given event. That is, the capability of the network will be an expected value, accumulated over possible recording configurations, weighted by the probability of each particular configuration. The resulting network capability can be contoured by setting an event at each location on the world-wide grid and then computing the probability that each station detects the event, given by (4). We may then compute a weighted average of some parameter reflecting the location capability. For this discussion, we assume that the area of the 90% ellipse is of interest, where the ellipse may be a confidence ellipse under classical assumptions or a posterior probability ellipse under the Bayesian paradigm. We are estimating this kind of network capability using a method that essentially parallel the *Networth* calculations made by Wirth et al (1976) in the case of seismic arrays.

We extend the classical methods first to the case where we observe wave-number parameters and their covariance matrix from n arrays and wish to combine or *fuse* the information into an overall location. The nonlinear model (10) and (11) can be treated in the usual way. That is, expand $\boldsymbol{\theta}_k(\boldsymbol{x})$ around some initial value, say $\boldsymbol{x} = \boldsymbol{x}_0$ and write a linearization as

$$\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k(\boldsymbol{x}_0) = A_k(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{e}_k, \tag{13}$$

where

$$A_k(\boldsymbol{x}) = \frac{\partial \boldsymbol{\theta}_k(\boldsymbol{x})}{\partial \boldsymbol{x}} \tag{14}$$

is the usual 2×2 matrix of partial derivatives of $\boldsymbol{\theta}_k(\boldsymbol{x})$. Then, stacking the n, 2×1 wave-number vectors and minimizing the weighted sum of squared errors can be done by successively estimating $\boldsymbol{\beta} = \boldsymbol{x} - \boldsymbol{x}_0$. This leads to

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}_0 + C^{-1}(\boldsymbol{x}_0) \sum_{k=1}^{n} A_k(\boldsymbol{x}_0)' \; \Sigma_k^{-1} [\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k(\boldsymbol{x}_0)], \tag{15}$$

where

$$C(\boldsymbol{x}_0) = \sum_{k=1}^{n} A_k(\boldsymbol{x}_0)' \; \Sigma_k^{-1} A_k(\boldsymbol{x}_0). \tag{16}$$

It follows that the estimated covariance matrix of the final estimator is

$$\widehat{\text{cov}}\; \hat{\boldsymbol{x}} = \sigma^2 C^{-1}(\hat{\boldsymbol{x}}). \tag{17}$$

Equations (10) and (11) exhibit the *fusion estimators* at each stage as pooled estimators over the n arrays as long as the variances are known. We may also develop a confidence ellipse for the fusion estimators under assumptions (A), (B) and (C), as given below.

(A) **Variance Known:** We may assume that the variance σ^2 is known, either from the statistical variances of the computed wave-number estimators or from a combination of factors including the statistical wave-number variances. In this case, the generalization of the usual chi-squared ellipse considered by Evernden (1969) can be computed from the fact that

$$(\boldsymbol{x} - \hat{\boldsymbol{x}})' C(\hat{\boldsymbol{x}}) (\boldsymbol{x} - \hat{\boldsymbol{x}}) \sim \sigma^2 \chi_2^2, \tag{18}$$

where $\sim$ denotes *is distributed as* and χ_2^2 denotes a chi-squared distribution with 2 degrees of freedom. Note that the statistical uncertainty of the wave-number estimators is already in the matrix Σ_k so that a plausible estimator for σ^2 in the absence of other factors might by unity.

(B) **Variance Unknown:** If variances are known only up to the constant σ^2, this scaling variance may be estimated from the set of arrays that record the event. For the Gaussian case, the maximum likelihood estimator is proportional to the unbiased estimator

$$s^2 = \frac{1}{2(n-1)} \sum_{k=1}^{n} \big(\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k(\hat{\boldsymbol{x}})\big)' \; \Sigma_k^{-1} \big(\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k(\hat{\boldsymbol{x}})\big). \tag{19}$$

This case, originally considered in Flinn(1965), leads to a confidence interval based on the F-distribution, namely

$$(\boldsymbol{x} - \hat{\boldsymbol{x}})' C(\hat{\boldsymbol{x}})(\boldsymbol{x} - \hat{\boldsymbol{x}}) \sim 2s^2 F_{2,2(n-1)}, \tag{20}$$

where $F_{2,2(n-1)}$ denotes the F-distribution with 2 and $2(n-1)$ degrees of freedom.

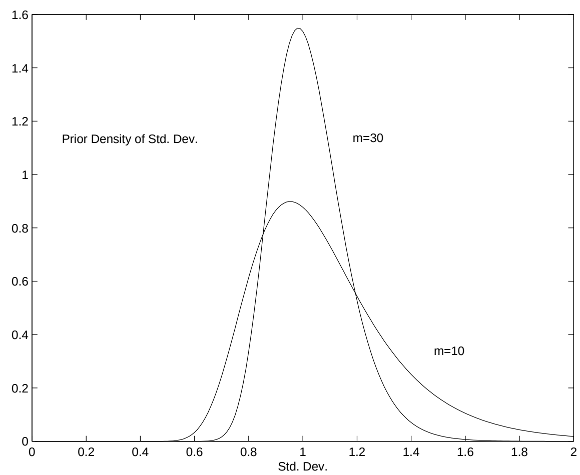

Figure 2: Possible prior distributions for standard deviations of measured wave-number estimates

(C) **Variance Subject to Prior Distribution:** It is often the case that it is unrealistic to assume that the variance is known exactly because (18) becomes too small. For a small number of arrays, the ellipse based on the F-statistic (20) is often much too large. A useful compromise, introduced by Jordan and Sverdrup (1981)and continued by Bratt and Bache (1988), is to quantify the initial uncertainty about σ^2 by assigning it a prior distribution with density function $\pi(\sigma^2)$. It is convenient to use the inverted chi-squared distribution with parameters m, representing the equivalent sample size embodied in the prior information and σ_0^2, representing a prior centering value for the variance. Figure 2 plots the density function for the standard deviation σ for $\sigma_0 = 1$ and $m = 10, 30$. We note that the two values put the standard deviation between .4 and 2 for $m = 10$ and between .6 and 1.6 for $m = 30$ For a fully Bayesian approach, we assume a non-informative prior on $(-\infty < x_1, x_2 < \infty)$ for the location $\boldsymbol{x}$ and compute the posterior distribution, given the wave-number observations, as a bivariate t-distribution with 2 and $2(n-1)+m$ degrees of freedom. The posterior estimator for the variance is

$$\sigma'^2 = \frac{2(n-1)s^2 + m\sigma_0^2}{2(n-1)+m} \tag{21}$$

implying that the best approach is simply to pool the initial variance σ_0^2 and the sample variance s^2, weighted by their degrees of freedom. The quadratic form involving the location vector $\boldsymbol{x}$ in the multivariate t has an F-distribution, making the 95% posterior probability ellipse for the location expressible as

$$\big(\boldsymbol{x} - \hat{\boldsymbol{x}}\big)'C(\hat{\boldsymbol{x}})\big(\boldsymbol{x} - \hat{\boldsymbol{x}}\big) \sim 2\sigma'^2 F_{2,2(n-1)+m} \tag{22}$$

It is interesting that the form of the posterior probability ellipse (18) is similar to (20) but will be tighter because of the additional degrees of freedom for the F-statistic. Hence, the Bayesian solution represents a compromise between (18) and (20), the methods of (A) and (B).

CONCLUSIONS AND RECOMMENDATIONS

Development of Global Network Performance Measures

We are proposing to combine the theoretical information available from the previous sections into an overall assessment of expected global network performance. Since there will likely be relatively few stations recording each event, we will make computations based on the Bayesian posterior probability interval given by (22).

The natural measure to use for the overall analysis will be the area of the ellipse, taken here to be the area of the 90% posterior probability ellipse.

First, consider the area for a fixed event location and a specified set of detecting arrays. Crucial inputs are as follows:

Signal-to-Noise Ratio

The signal-to-noise ratio r should be determined for each array recording a signal from that particular location. One option is to just assume a value, based on experience, or to sample randomly from an assumed distribution of values. Given an observed event and a recording station, estimation of the signal and noise spectra by maximum likelihood is solved in Shumway et al (1999) for both perfectly correlated and decorrelated signals. These estimators can be used directly if events and recordings are available. In the example given here, we arbitrarily take signal-to-noise ratios of $r = 2, 3$.

Array Geometry

The geometry of each array affects the detection probability for that array and the accuracy of the input wave-number $\hat{\boldsymbol{\theta}}_k(\boldsymbol{x})$, large through N, the number of sensors and R, the sample covariance matrix of the array. Again, Shumway et al (1999) show expressions for the estimated variance under the two different assumptions for correlation. In the example that comes next we used two array geometries; the first was a triangle with baseline 1km and a center element; the second was the first triangular array with an additional inverted inner triangle with baseline distance .2km (see Blandford, 1997).

Time-Bandwidth Product, Center frequency, velocity

The time-bandwidth product BT will depend on the assumed frequency band containing the signal and the length of the signal window. We assumed here a center frequency of 1 Hz and a time-bandwidth product of $BT = 17$, which represented a reasonable compromise from our earlier work (Shumway et all, 1999) with a Pacific Islands event. Velocity was assumed to be .3 km/sec in the example.

Note that the values of r and N in the preceding discussion are sufficient for computing the detection probability at a fixed false alarm rate from the F-distribution given by (4). The input covariance matrix Σ_k can be computed from (8), using an array assumption and a configuration of recording stations. For the posterior probability ellipse (22) we require an assumed location $\boldsymbol{x}$ which we took as the two events marked with + in Figure 1. This leads to a value for the matrix $C(\boldsymbol{x})$ in (16). As assumptions for the prior scaling variance and its uncertainty, we took $\sigma_0^2 = 1$ and $m = 10$, assuming the sample variance $s^2 = 1$.

Figure 3 shows the predicted posterior 90 and 95% ellipses for the locations of the two events, assuming various recording array configurations (see also Table 1). We note that ellipses produce reasonably satisfactory regions for a small number of arrays detecting but that the areas are still not up to the often-mentioned standard of areas less than 1000 km^2 at 90% confidence. Of course, this could be achieved by increasing the assumed time-bandwidth product or the signal-to-noise ratio or by changing the nature of the recording array. Table 1 gives a short introduction to what might happen under various scenarios. Increasing the number of elements in the array decreases the area by approximately 10%; increasing the signal-to-noise ratio 2 to 3 leads to reductions of about 40%. Assuming that there are more detecting arrays also has a fairly substantial effect.

Finally, we plan to develop an expression for the average area expected over all possible configurations of detecting arrays. That is, define a detection indicator D_k that is one if the array detects and is zero otherwise, for the full set of possible detecting arrays, say, for $k = 1, 2, \ldots, K$ of them. If we define $p_k = Pr\{D_k = 1\}$,

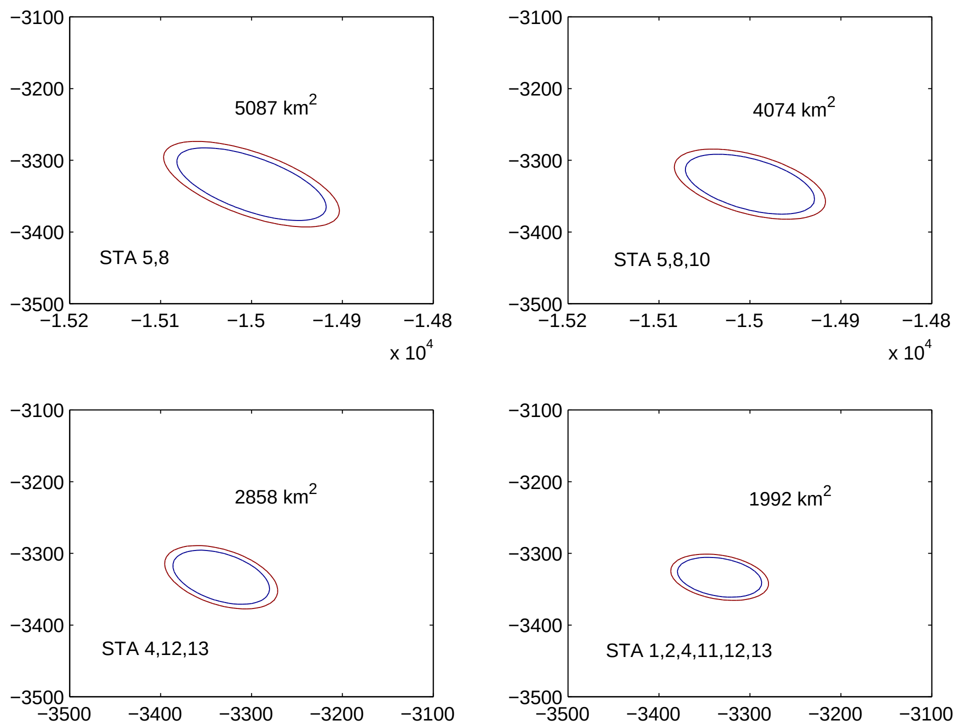

Figure 3: Posterior probability (90 and 95%) ellipses for various array detection configurations assuming 7-element arrays, perfect signal correlation and a single channel $S/N = 2$.

the joint density of the random variables $D_1, D_2, \ldots, D_K$ would be

$$P(D_1, \ldots, D_k) = \prod_{k=1}^{K} p_k^{D_k} (1 - p_k)^{1-D_k}. \tag{23}$$

The probabilities can be arbitrary for each station, based on recorded data, but our preferred approach is to use the estimated probabilities computed from the result (4). Now, for any given event, we can observe 2^K possible configurations of detecting stations and each one of them will give a predicted area. We can multiply each configuration probability by its area and add them up to get the average areas of location for that particular source. Alternately, and perhaps, easier would be to simulate values of $D_1, \ldots, D_K$ repeatedly and simply average the resulting areas. In either case, it is clear that a relatively small computing effort will yield an average predicted 90% uncertainty area for each location. Plotting contours on the map will give us an index of global network performance.

Table 1: Areas (km^2) of 90% Posterior Probability Ellipses for Simple and Extended Triangular Arrays

	Triangular	Array	Extended	Array
Stations/Event	$S_N = 2$	$S_N = 3$	$S_N = 2$	$S_N = 3$
5,8/1	5554	3566	5087	3316
5,8,10/1	4448	2856	4074	2656
4,12,13/2	3129	2002	2858	1863
1,2,4,11,12,13/2	2176	1397	1992	1299

REFERENCES

Armstrong, W.T. (1998). Comparison of infrasound correlation over differing array baseline. In *Proceedings of 21st Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty (CTBT).* 543-554

Blandford, R.R. (1997). Design of infrasonic arrays. *AFTAC Report AF-TR-97-013*, Headquarters, Air Force Technical Applications Center, Patrick AFB FL 32925-3002.

Bratt, S.R. and T.C. Bache (1988). Locating events with a sparse network of regional arrays. *Bull. Seismolog. Soc. Amer*, **78**, 780-798.

Dighe, K.A., R.W. Whitaker and W.T. Armstrong (1998). Modeling study of infrasonic detection of 1 kt atmospheric blast. In *Proceedings of 21st Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty (CTBT).* 571-585.

Evernden, J.F. (1969). Precision of epicenters obtained by small numbers of world-wide stations. *Bull Seismolog. Soc. Amer.*, **59**, 1365-1398.

Flinn, E.A. (1965). Confidence regions and error determinations for seismic event location. *Rev. Geophys.*, **3**, 157-185.

Hedlin, M.H., F.L. Vernon and J. Berger (1998). Surveying infrasonic noise on oceanic islands. In *Proceedings of 20th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty (CTBT)*, 586-595.

Jordan, T.H. and K.A. Sverdrup (1981). Teleseismic location techniques and their application to earthquake clusters in the South-Central Pacific, *Bull Seismolog. Soc. Amer*, **71**, 1105-1130.

Shumway, R.H.(1983). Replicated time series regression: An approach to signal estimation and detection. *Handbook of Statistics Vol. 3*, Chapt. 18, 383-408, Time Series in the Frequency Domain. D.R. Brillinger and P.R. Krishnaiah ed., North Holland.

Shumway, R.H., S.E. Kim and R.R. Blandford (1999). Nonlinear estimation for time series observed on arrays. Chapter 7, Ghosh ed. *Asymptotics, Nonparametrics and Time Series*, 227-258. New York: Marcel Dekker.

Shumway, R.H. and D.S. Stoffer (2000). *Time Series Analysis and Its Applications.* New York: Springer-Verlag. Forthcoming.

Shumway, R.H. and S.E. Kim (1998). Detection and estimation on small infrasonic arrays. *Proceedings of 20th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty (CTBT)*, 628-637.

Shumway, R.H. and S.E. Kim (1999). Estimation and location using multiple infrasound arrays. *Proceedings of 21st Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty (CTBT).*

Wirth, M.H. R.R. Blandford and R.H. Shumway(1976). Automatic seismic array and network detection. *Bull. Seismolog. Soc. Amer.*, **66**, 1375-1380.