DETECTION AND LOCATION CAPABILITIES OF MULTIPLE INFRASOUND ARRAYS

Robert H Shumway

University of California, Davis

Sponsored by Defense Threat Reduction Agency

Contract No. DTRA01-00-C-0082

ABSTRACT

We have developed an integrated approach to locating an infrasound source that fuses local-array wave-number parameters and uncertainties into an overall location procedure. For local estimation of the velocity and azimuth, a small-array theory, based on maximum likelihood, has been given in earlier work that characterizes the large-sample uncertainty of the estimates and evaluates the theoretical missed-signal and false alarm probabilities. We have verified the theoretical uncertainties by computing empirical estimates using the frequency domain bootstrap on a gas-pipe explosion, a Titan IV B missile launch, and a Hawaii meteorite. Detection probabilities and wave-number uncertainties are then integrated into a Bayesian nonlinear regression procedure for evaluating the location capabilities of the particular global infrasound array that is proposed for the International Monitoring System (IMS).

We show contour maps for the average expected areas of the 90% confidence ellipses produced by the overall fusion procedure. The results indicate that for a single-array false alarm probability of 10^{-4}, the standard proposed IMS 4-element array will detect over 90% of the signals at signal-to-noise ratios as low as .6, with sufficient bandwidth. Location accuracies will require higher signal-to-noise ratios on the order of 2-4 and high single-array detection probabilities (>.90) to guarantee reasonable coverage (1000 km^2) for the 90% posterior probability ellipses. Eastern Hemisphere and Western Hemisphere 90% contour plots show almost complete coverage by expected uncertainty areas of 1000 km^2 or less.

OBJECTIVES

The objectives of this project have been to (1) develop the detection and estimation capabilities of small infrasound arrays and (2) to integrate these single-array directional estimation statistics into a procedure for assessing the predicted global performance of the infrasound component of the proposed International Monitoring System (IMS).

In support of (1), our sub-objectives were to develop local-array performance capabilities for estimating velocities and azimuths of propagation and to characterize single-array signal detection probabilities at low false alarm rates. In support of (2), our sub-objectives were to develop fusion posterior-probability ellipses for location and to incorporate single-array detection probabilities into a procedure for developing a global coverage map giving expected areas of 90% uncertainty regions.

RESEARCH ACCOMPLISHED

We have investigated a number of proposed procedures based on plane wave models for detecting infrasound signals at small arrays and for estimating velocities and azimuths, along with their predicted uncertainties. Velocities and azimuths are functionally related to the coordinates in the wave-number plot, say $\boldsymbol{\theta}=(\theta_1 ,\theta_2)'$ and their predicted uncertainties. The estimated wave numbers and their covariance matrices for the detecting sub-arrays are fused into an overall location and its posterior probability ellipse. A large scale simulation using predicted detection probabilities and locations then is used to develop contour plots of the areas of the 90% error ellipses for the Eastern and Western Hemispheres.

Detection of Infrasound Signals

We have investigated three wave-number detectors from the literature, the Capon (1969) *high-resolution estimator*, the F-detector suggested by Shumway (1971), and the MUSIC estimator suggested by Schmidt (see Stoica, 1989).

The high-resolution estimator of Capon is the inverse of a Hermitian form in the probe vector, $\mathbf{x}(\boldsymbol{\theta})$, involving the inverse of the covariance matrix (see Shumway, 2001). Difficulties are in estimating the covariance matrix and in using the statistical distribution, which depends on the unknown theoretical covariance matrix. The multiple signal characteristic (MUSIC) estimator keeps the same form but replaces the inverse spectral matrix by the spectral matrix of the noise, approximated by an inner product of the residual eigen vectors (see Shumway, 2001). The difficulties with this detector relate to the intractability of its distribution under the *noise-alone* and the *signal-plus-noise* hypotheses.

The usual F-detector is defined as the ratio of the scaled beam power to the scaled error power, $F(\theta)$, where θ is the two-dimensional wave-number vector corresponding to a given velocity and azimuth (Shumway, 2001). In contrast to the high-resolution and MUSIC detectors, the performance is determined by observing that $F(\theta)$ is distributed as $(1+rN)F_{2L,2L(N-1)}$, i.e., as an F-statistic with 2L and 2L(N-1) degrees of freedom, where L=BT is half the bandwidth and N is the number of elements in the array. The parameter r is the signal to noise ratio on a single channel. Since the distribution under the noise-alone and signal-plus –noise hypotheses both involve the F, with r=0 under the noise-alone hypothesis, the preceding theory allows us to predict the detection probability as a function of any given false alarm probability. The result also allows unbiased estimation of the signal-to-noise ratio r by equating the value of the F-statistic to the expectation of $(1+rN)F_{2L,2L(N-1)}$.

Figure 1 shows the predicted detection probability for two hypothetical combinations of bandwidth (2L/n) and sub-array size, where L is the number of frequencies smoothed to obtain the test statistic and n is the total number of time points. Note that even with the smaller size array characteristic of those used in this paper, the signal detection is very high for relatively low signal-to-noise ratios and a false alarm probability of 10^{-4}. Extremely low false alarm probabilities are of interest in order to maintain an overall false alarm rate that is sufficiently low when there are many wave-numbers to test. For example, if 100 wave numbers are potentially of interest in this case, Bonferonni's inequality guarantees that the overall false alarm probability will be less than .01.

For illustration purposes, we show a contour plot in Figure 2 exhibiting the behavior of the three statistics given above and the beam power. Note that all statistics give comparable results for the Hawaiian bolide observed on April 23, 2001 from the Pinion Flat Array (shown in Table 1). This array was chosen because, with N=6 elements, it

represents a likely result from processing an event using an array that is close to the IMS configuration We simply note here that all processors detect the event and that F=172 is highly significant. It should also be noted that the signal to noise ratio was extremely high here and the plots are unusually unambiguous. The apparently superior resolution of the Capon and MUSIC estimators does not lead to lower variances or to better resolution of multiple signals (Shumway, 2001) Comparable plots from the Los Alamos, Lac du Bonnet and Mina , Nevada showed multiple maxima that sometimes corresponded to velocities in the neighborhood of .3 km/sec and sometimes at unreasonable velocities. We note also that these arrays, for various reasons, only recorded on N=3 channels. To resolve some of these ambiguities, a general nonlinear optimization was employed that started in the neighborhood of a velocity and azimuth corresponding to the Hawaii location.

Estimation of Velocity and Azimuth Parameters

Figure 2 suggests that we simply read the wave-number coordinates corresponding to the maximizing value of the appropriate statistic and specify the velocity and azimuth corresponding to these wave-numbers as our estimated values. For the F-detector, the maximizers are equivalent to the maximum likelihood estimators, as was shown in Shumway *et al* (1998). Using the Cramer-Rao lower bound, they obtained variance-covariance matrices for the estimated wave-number parameters in Figure 2 and for the derived velocities and azimuths.

It was recognized early in the investigation that analytical computations for the variance-covariance matrices of the Capon and MUSIC estimators would be difficult and a version of the frequency domain bootstrap (see Shumway and Stoffer, 200, p244) was employed for these cases. This involves drawing a random sample from the frequencies determining the maximum likelihood estimator repeatedly and computing the mean and variance over a large number of bootstrap samples. This was not only done for the Capon and MUSIC estimators but was used to check the large-sample covariance matrices computed for the maximum likelihood estimators.

Table 1. Estimated azimuths and uncertainties (bootstrap std. dev.) for sample events. LANL denotes value given by Los Alamos National Laboratories

Event	F	S/N	Capon	Music	LANL
Gas Pipe Explosion	256(1.3)	3.15	258(1.8)	257(1.3)	
Titan IV Missile	267(1.0)	.75	263(.4)	263(.4)	
4/23 Hawaiian Bolide					
Los Alamos	263(.3)	1.91	263(.4)	263(.4)	259
Lac du Bonnet	240(.1)	2.86	240(.2)	240(.1)	244
Pinion Flat	256(.8)	28.7	256(.8)	256(.8)	247
Hawaii	61(.6)	.83	59(1)	61(.6)	61
Mina, NV	237(.9)	3.79	238(1.3)	237(.9)	236
St. George, UT	278(2.3)	23.1	269(2.4)	276(2.3)	252

Preliminary results are shown in Table 1 for a 1998 gas pipe explosion and a Titan IV missile launch, both were recorded on four elements at the Los Alamos array. Additionally, we show results from a suite of arrays that recorded the April 23, 2001 Hawaii bolide. All events reside in the CMR R&D Test Bed Infrasound Waveform Library. As mentioned earlier, in order to handle severe aliasing that resulted in multiple maxima for the Los Alamos, Lac du Bonnet, and Mina arrays, non-linear optimization was used, starting in the vicinity of the assumed location. Because the asymptotic likelihood theory will not apply with multiple maxima, we used the frequency domain bootstrap to estimate the variances and covariances. The results for the bolide show estimated azimuths that are comparable with those obtained by Los Alamos National Laboratories with the exception of St. George Utah, where the angle should have been in the neighborhood of 246 degrees, LANL was close but we were off by 17-25 degrees

For the case where more than one array records a single event such as given by Table 1, one would be interested in the estimated location made by combining or fusing the results of the single-array wave-numbers and their covariance matrices into and overall location. The location theory is summarized in the next subsection.

Location Using Fused Array Wave-Number Parameters

Location results, given previously in Shumway (2000), require estimated wave-number coordinates $\theta_k(\mathbf{x})$ for k=1,2,…,n sub-arrays and their respective estimated covariance matrices, $\sum_k$, computed from one of the procedures in the previous section. The estimated wave-number coordinates are assumed to be a function of the location $\mathbf{x}=(x_1, x_2)'$. Linearizing $\theta_k(\mathbf{x})$ about an initial location $\mathbf{x}_0$, and performing the exact Bayesian analysis with a uniform diffuse prior on $\mathbf{x}$ and an inverted chi-square distribution for the location variance σ^2 in the location covariance matrix $\sigma^2 \sum$, we obtain the posterior probability distribution of the location vector $\mathbf{x}$ as a bivariate t-distribution. Then, using the fact that the posterior probability distribution of the quadratic form in $\mathbf{x}$ will have an F-distribution with 2 and 2(n-1)+m degrees of freedom, we obtain a posterior probability ellipse with a given posterior coverage area. Note that m degrees of freedom are associated with the chi-squared random variable defining the uncertainty in the scaling variance. The solution for independent errors ($\sum=\sigma^2 I$) was given by Jordan and Sverdrup (1981).

Global Array Performance of Proposed IMS Array

In order to develop a reasonable measure of global performance we considered incorporating the infrasound recording arrays proposed for the infrasound part of the IMS. For these plots, we used the asymptotic covariance matrix implied from maximum likelihood, as given by Shumway *et al* (1999) and in a previous SRS report (Shumway, 2000). For this simulation, we need the assumed signal-to-noise ratio (r=4) at a single station, the smoothing constant (L=51 frequencies), and the number of channels in a sub-array, assumed to be N=7, composed of 1km outer triangles with 1km sides and an inner inverted triangle with 2km sides. For location capability, we look at various simulated random configurations of detecting stations. For this initial simulation, we were lacking precise information on detection probabilities for single arrays although from Figure 1, we can infer that they will generally be high, with malfunctions excepted. Overall, we might optimistically expect an average of 9 out of 10 stations to detect and this was assumed for the simulation. We summarize the simulation procedure as:

1. Fix a hypothetical event location
2. Set input parameters as N=7 elements, signal to noise ratio r=4, and a bandwidth of .1 Hz.
3. Simulate a configuration of detecting stations (3<n<28, 3<n<32) for Western and Eastern Hemispheres respectively and P=.9, the probability for single sub-array detection. Assign the chi-squared distribution for the scaling variance as chi-squared with m=10 and expectation 1.
4. Compute the Bayesian posterior probability ellipse and area.
5. Repeat 3.and 4. 500 times, averaging the areas obtained.
6. Record average on grid and increment the hypothetical event location in 1. by 5 degrees.
7. Contour the results for the 90% posterior probability areas separately for the Eastern and Western Hemispheres.

Figure 3 shows the resulting contours in thousands of km^2 separately for the Western and Eastern Hemispheres. We note that doing the computation this way leaves out the possible detections between hemispheres and will distort the performance at the edges of both plots. We note that the 1000 km^2contours include the majority of regions of interest (-80 degrees to 50 degrees for the Western Hemisphere and –80 degrees to 80 degrees for the Eastern Hemisphere). One could obtain a better index of performance for the western Pacific by folding in some of the Eastern Hemisphere sub-arrays and re-doing the computation. We tried relaxing the signal-to-noise ratio to r=2 or decreasing the detection probabilities to P=.5. Either strategy increases the 1000 km^2 to 3000 km^2.

CONCLUSIONS AND RECOMMENDATIONS

Theoretical computations have shown that the wave-number F-detector at a single array can be expected to perform very well for low signal-to-noise ratios and false alarm probabilities as low as 10^{-4}. We have not been able to collect enough historical detections from IMS stations to know that the empirical false alarm and signal detection probabilities will be equally impressive. Historical precedents set by seismic detections would imply that the empirical false alarm rates will be higher than the theoretical ones, implying that thresholds will need to be set high

to avoid significant numbers of false alarms. Every attempt should be made to develop empirical rates based on real events.

For the fused location estimators, theory suggests that we need higher signal-to-noise levels for the purposes of locating accurately with the seven element sub-arrays. The theoretical covariances for the conditions that guaranteed the good worldwide coverage for the 90% 1000 km^2 area contours were on the order of 10^{-4} km^2 whereas the bootstrap simulations using the Hawaiian bolide suggested that the covariance matrix would be on the order of 10^{-3} km^2. An indication of the overall scale variance that multiplies the covariance matrix can possibly be inferred by computing a location and a 90% posterior probability ellipse for the Hawaiian bolide. We intend to try this before the end of the contract.

REFERENCES

Capon, J. (1969). High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE*, **57**, 1408-1418.

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.(1971). On detecting a signal in N stationarily correlated noise series. *Technometrics*, **13**, 499-519.

Shumway, R.H. (2000). Detection and location capabilities of multiple infrasound arrays. *Proc. 22nd Seismic Research Symposium, 12-15 September*, New Orleans, Louisiana, Chapter 7, 7-12.

Shumway, R.H. (2001). Detection and location capabilities of multiple infrasound arrays. *Proc. 23rd Seismic Research Symposium, 2-5 October, Jackson Hole, Wyoming*, Chapter 7, 7-10.

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..

Stoica, P. and A Nehorai (1989). Music, maximum likelihood, and Cramer-Rao lower Bound. *IEEE Trans. Acoustics, Speech and Signal Processing*, **37**, 720-741.

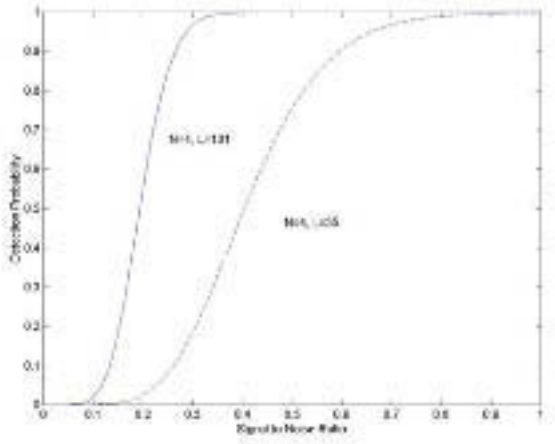

Figure 1. Detection probabilities for different signal-to-noise and smoothing options.

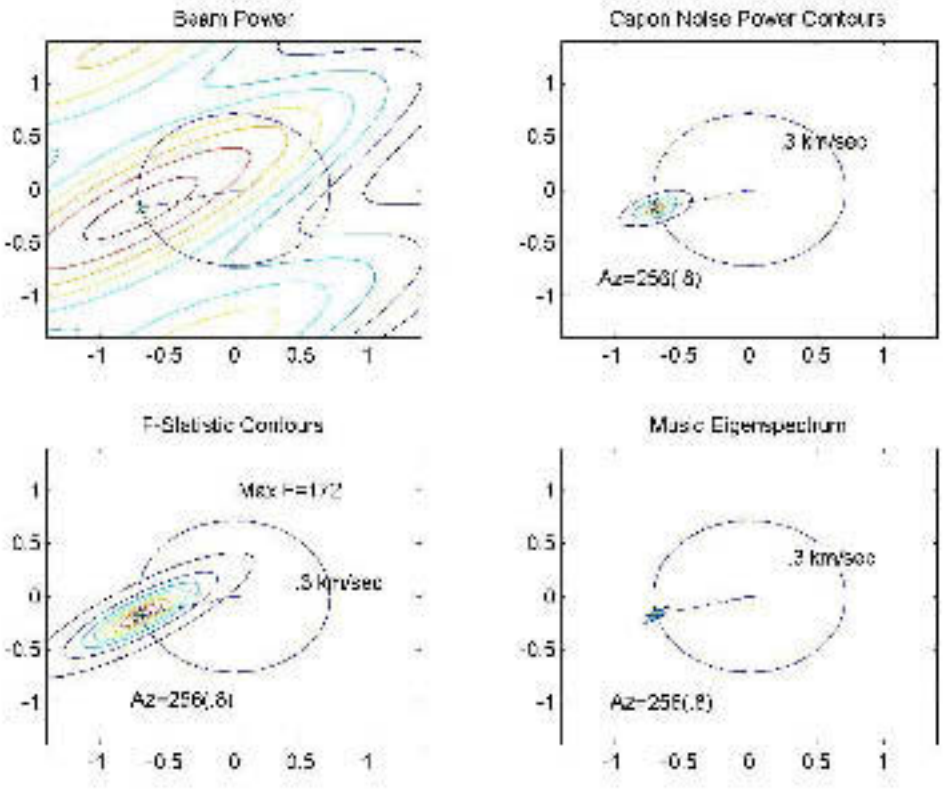

Figure 2. Wave-number analysis for the Hawaii bolide observed at Pinion Flat.

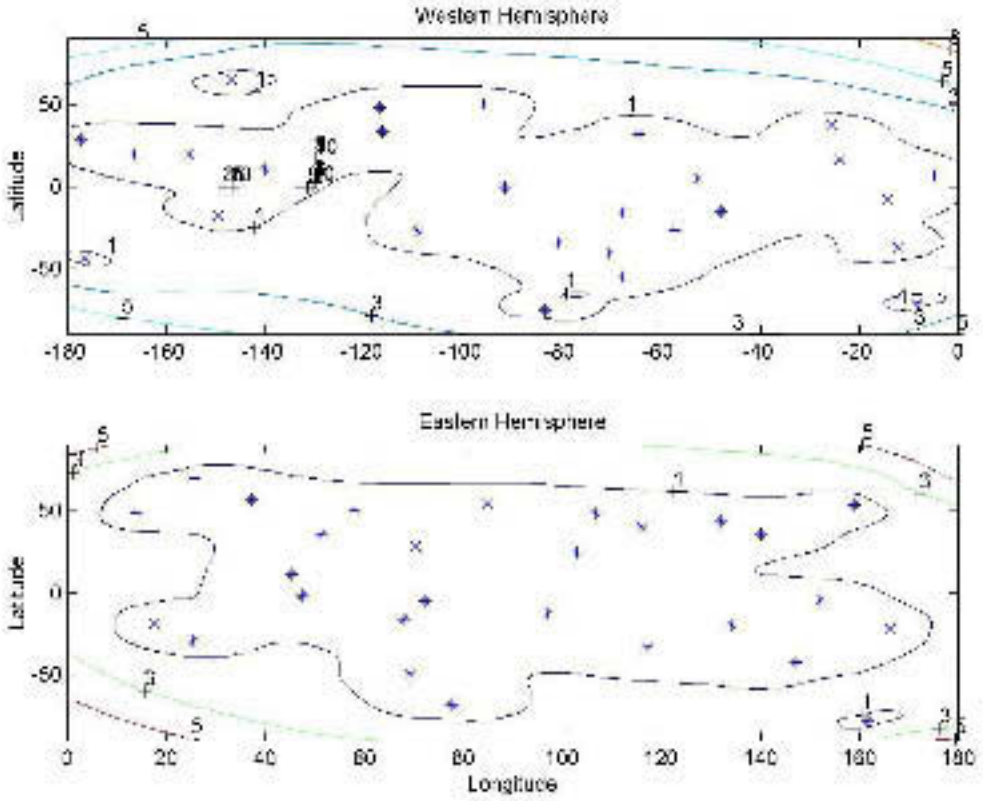

Figure 3. Contoured 90% posterior probability areas for a signal-to-noise ratio of 4 and a single array detection probability .9. Note that the 1000 square km contour includes most of both hemispheres. Array locations are shown as *.