AFRL-VS-HA-TR-2006-1091

Advances in Mixed Signal Processing for Regional and Teleseismic Arrays

Robert H. Shumway

University of California
Department of Statistics
One Shields Avenue
Davis, CA 95616

Final Report

15 August 2006

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
29 Randolph Road
AIR FORCE MATERIEL COMMAND
Hanscom AFB, MA 01731-3010

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the Electronic Systems Center Public Affairs Office and is available to the general public, including foreign nationals. Qualified requestors may obtain additional copies from the Defense Technical Information Center (DTIC) (http://www.dtic.mil). All others should apply to the National Technical Information Service.

AFRL-VS-HA-TR-2006-1091 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//Signature//	//Signature//
ROBERT RAISTRICK Contract Manager	ROBERT BELAND, Chief Battlespace Surveillance Innovation Center

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
August 2006	Final Report	30 MAY 2004 - 30 JUNE 2006

4. TITLE AND SUBTITLE

Advances in Mixed Signal Processing for Regional and Teleseismic Arrays

5a. CONTRACT NUMBER
FA8718-04-C-0013

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62601F

6. AUTHOR(S)

Robert H. Shumway

5d. PROJECT NUMBER
1010

5e. TASK NUMBER
SM

5f. WORK UNIT NUMBER
A1

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Statistics
University of California
One Shields Avenue
Davis, CA 95616

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/VSBYE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-VS-HA-TR-2006-1091

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release: Distribution Unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project considers possible approaches to resolving mixtures of propagating signals observed on arrays. In particular, conventional approaches such as beam-forming, MUSIC and single-signal F-statistics have flaws that will not adapt to certain mixtures. In order to solve this problem, we derive the partial F-statistic for testing for an added signal in a multiple-signal model. In this case a combination of sequential non-linear partial F-statistics in combination with Akaike's corrected model selection criterion AICC leads to determining the correct configuration of signals and their velocities and azimuths.

The conventional estimators and the new estimators are applied to known and unknown configurations of regional signals from China and to a teleseismic mixture involving two known earthquakes and noise caused by an ocean storm. We also analyze a regional event with propagating noise and show that a deconvolution based on the two velocities and azimuths gives an enhanced view of the depth phase. Software as well as internal and external documentation is provided in the form of MATLAB subroutines that can be incorporated into government research tools such as MATSEIS. We provide software for (1) time-frequency analysis, (2) F, MUSIC and Capon detectors, (3) multiple signal analysis, (4) bootstrap confidence intervals and (5) deconvolution.

15. SUBJECT TERMS
Seismic arrays, Multiple signal models, Least squares, Model selection, Deconvolution.

16. SECURITY CLASSIFICATION OF: UNCLASSIFIED			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Robert J. Raistrick
UNCLAS	UNCLAS	UNCLAS	SAR	48	19b. TELEPHONE NUMBER (include area code)

Table of Contents

Figures

Tables

1. Introduction

Several signal detection and estimation problems arise in monitoring teleseismic and regional events using data from long period and short period arrays. Both on-line and off-line detection and estimation software must be able to operate successfully in the presence of either relatively high noise or interfering events. Besides lowering detection thresholds, the interfering events will produce distortions in velocity and azimuth estimators resulting in poor locations for teleseismic events and in an inability to sort out velocities for the different phases for near and far regional arrivals on the short period arrays. It should also be noted that amplitude or power measurements made for discrimination purposes will also be distorted.

Fitting data from an array of distributed sensors to a multiple-signal model for the purpose of extracting possible mixed signals is a method used routinely in seismic data processing, e.g., Smart(1972). In general, however, such procedures do not constitute detectors. They have served, among other purposes, to disentangle mixed (simultaneously arriving) long-period surface-wave signals when one of them was sought as a critical discriminant to distinguish an explosion from a possible earthquake. Such processors have had no independent detection statistics, and, consequently, for verification and validation of waveforms extracted by their use, alternative means are employed. Those means include the search, in the same time interval, for body-waves of possible associated seismic events in the records of a complementary network of short-period seismic stations. Any tentatively identified surface waveform, recovered by multiple-signal modeling, may be verified by association with one of the seismic events detected and located by that search. If a surface waveform so extracted meets the criteria of arrival time, dispersion, and magnitude predicted from the estimated location, magnitude and time of occurrence of a seismic event discovered in the short-period data, the surface-wave is associated with that event and treated as 'genuine'.

To illustrate, consider Figure 1, which contains a verifiable mixture of signals from two earthquakes, one from the south of Africa and the other from the Philippines, observed at a long period array in Korea. As we will note later, conventional methods applied to this data indicate a single event from somewhere between the two observed events. Rather than discovering such incorrect results using an alternate network or short-period records, it would be strongly preferable to have a detection statistic that sorted out multiple arrivals from the long-period records in Figure 1. We show that the sequential detection procedure given in this paper identifies the two signals, as well as a complementary unidentified component that may be another event or noise.

Another powerful advantage of the multiple-signal detector is its immunity to interference from persistent coherent noise, natural or anthropogenic, from such phenomena as, say, the microseisms of storms at sea, or from power plant emanations. Conventionally persistent, nuisance coherences are dealt with by notch filters to prevent the detector from responding to the coherent noise, and from missing smaller signals in adjoining frequencies. These notch filters generally operate right in the midst of the band of frequencies of interest, and that critical portion of the band is sacrificed to accommodate the vulnerabilities of the detector

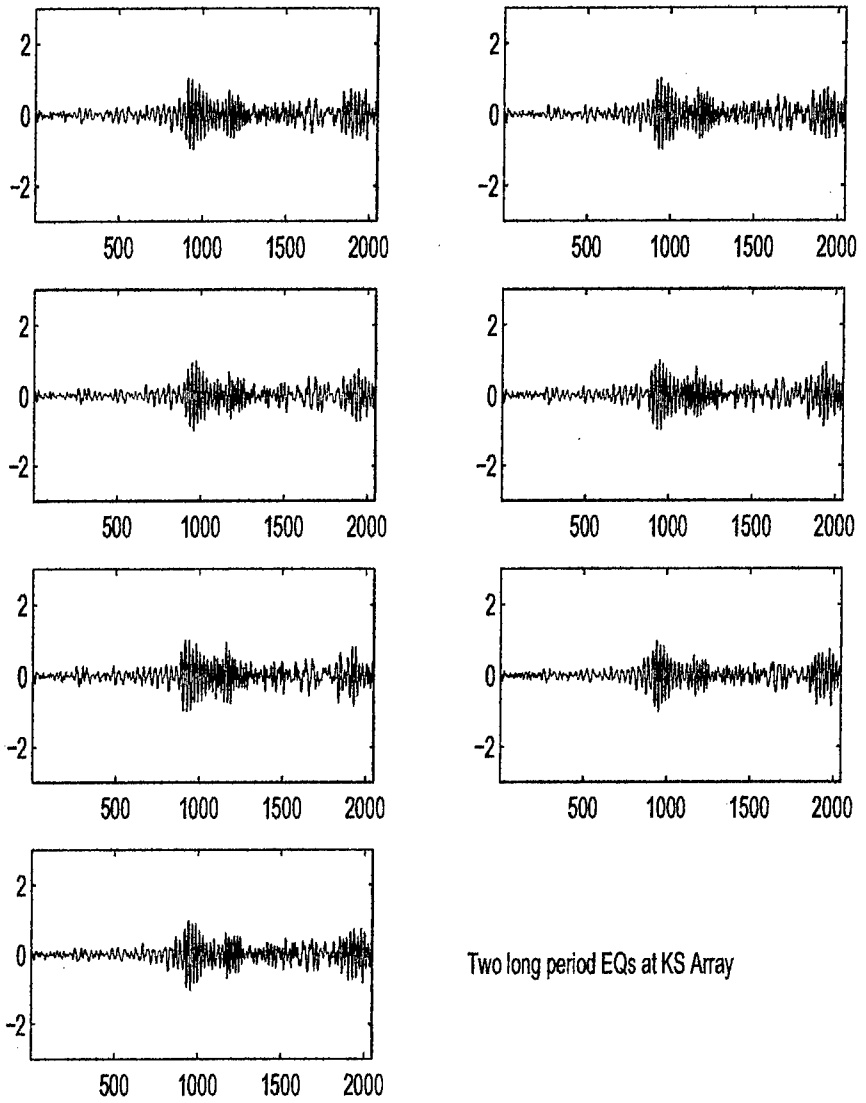

Figure 1: Mixture of signals from two earthquakes from south of Africa and the Philippines observed at USAEDS long-period seismic array in Korea. Correct geodesic back-azimuths are 226 degrees and 198 degrees.

(Clauter, 2004). Finally, we note another benefit of the sequential F-detector proposed in this paper is its use of the noise computed during the signal window. Conventional detection using the ratio of the short-term average mean-square error (STA) to the corresponding long-term average (LTA) requires that the LTA be refreshed after the onset of a signal to remain relevant, causing a blind window where the detector will miss valid signals.

Approaches to detecting signals on arrays all focus on the basic model that expresses the observed channel as sums of delayed signals and a unique noise process. The delays are functionally dependent on velocity and azimuth if the signals are propagating plane waves, and this is the assumption that is usually made. Methods that are commonly in use for analyzing such data, when a single signal is assumed to be present, can be roughly categorized as (a) beam-forming and plotting the power as a function of slowness, which can be converted to estimators of velocity and azimuth,(b) Capon's estimator (see Capon, 1969, Capon and Goodman, 1970), (c) beam-forming converted to an F-statistic by dividing by an estimator of the noise power (see Melton and Bailey, 1957, Blandford, 1970, Shumway, 1970, 1971, 1983, 2000), (d) Multiple Signal Characteristic (MUSIC)(Schmidt, 1979, Stoica and Nehorai, 1989), and (e) cross correlation (Tribuleac and Herrin, 1997). Several algorithms such as the sequential F-detector proposed here and the multiple signal characteristic (MUSIC) algorithm are available that offer promise for handling array data with low signal-to-noise ratios and contamination from interfering signals. In this proposal, we exhibit the performance of currently available algorithms on teleseismic and regional data containing mixed signals and demonstrate the superior performance of the sequential F-statistic. A sequential analysis of power using the F-statistic is proposed that estimates the correct number of signals and their velocities and azimuths. This is contrasted with results using conventional f-k estimators that do not handle the mixed signal case.

2. The Multiple Signal Model

The usual model expresses the received data $y_j(t), j = 1, \ldots, N, t = 1 \ldots, n$ at N sensors

$$y_j(t) = \sum_{k=1}^{M} s_k(t + \boldsymbol{r}_j'\boldsymbol{\theta}_k) + v_j(t) \tag{1}$$

as the sum of M propagating plane waves with time delays $T_j(\boldsymbol{\theta}_k) = \boldsymbol{r}_j'\boldsymbol{\theta}_k$ for the k^{th} signal at the j^{th} sensor where $\boldsymbol{r}_j$ is the two dimensional coordinate of sensor j in km and $\boldsymbol{\theta}_k$ is the two-dimensional slowness vector of signal k in sec/km. The two-dimensional slowness vector $\boldsymbol{\theta}_k = (\theta_{1k}, \theta_{2k})'$ is related to the velocity $V_k = ||\boldsymbol{\theta}_k||^{-1}$ and azimuth $\alpha_k = \tan^{-1}(\theta_{2k}/\theta_{1k})$.

Because the plane wave model formulates more easily in the frequency domain, we may consider a model for the discrete Fourier transforms, say

$$Y_j(\omega_\ell) = \sum_{k=1}^{M} \exp\{2\pi i \omega_\ell \, \boldsymbol{r}_j'\boldsymbol{\theta}_k\} S_k(\omega_\ell) + V_j(\omega_\ell) \tag{2}$$

evaluated at $\ell = 1 \ldots L$ frequencies of the form $\omega_\ell = l/n$ in the neighborhood of some target frequency ω, measured in cycle per point. A convenient representation of the model is obtained by stacking the sensors in an $N \times 1$ vector $\boldsymbol{Y}_\ell = (Y_1(\omega_\ell) \ldots, \boldsymbol{Y}_N(\omega_\ell)'$. Then, the above equation becomes

$$\boldsymbol{Y}_\ell = Z_\ell(\Theta)\boldsymbol{S}_\ell + \boldsymbol{V}_\ell \tag{3}$$

where $\boldsymbol{S}_\ell = (S_1(\omega_\ell), \ldots, S_M(\omega_\ell))'$ denotes the unknown vector of signals at frequency ω_ℓ and

$$Z_\ell(\Theta) = \left\{ \exp\{2\pi i \omega_\ell \, \boldsymbol{r}_j'\boldsymbol{\theta}_k\}, j = 1, \ldots, N, k = 1 \ldots, M \right\} \tag{4}$$

is an $N \times M$ matrix that defines the way the plane waves map into the observed sensor elements. Suggested detectors are based on the statistical structure assumed for the elements of the model (2) under the fixed and stochastic signal assumptions.

Suppose that the signal is deterministic and the noise has $N \times N$ spectral matrix $f_v(\omega)$. Then, it is clear that asymptotically, the inferences can be based on a model that assumes that $\boldsymbol{Y}_\ell, \ell = 1, \ldots, L$ are independently distributed with mean $Z(\Theta)\boldsymbol{S}_\ell$ and covariance matrix f_v. Various specializing assumptions lead to the classic beam and to the various forms of the F-detector. We remark that it is clear that the deterministic model has the nonlinearity in the mean function. The primary interest will be in the parameter matrix $\Theta = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_M)$ which enters non-linearly in the mean function,

$$E\boldsymbol{Y}_\ell) = Z_\ell(\Theta)\boldsymbol{S}_\ell.$$

If the signal is random with $q \times q$ spectral matrix $f_s(\omega)$, $\boldsymbol{Y}_\ell, \ell = 1, \ldots, L$ will be independently distributed with mean zero and covariance matrix

$$\Sigma_y(\omega) = Z_\ell(\Theta) f_s(\omega) Z_\ell^*(\Theta) + f_v(\omega) \tag{5}$$

and the nonlinearity in the slowness parameters is concentrated in the covariance structure. Such a model can be used to argue in the single-signal case for Capon's (1969) high-resolution estimator and for the MUSIC estimator of Schmidt (1979). Furthermore, a maximum likelihood estimator can be derived by maximizing the quasi-Gaussian log likelihood implied under the complex Gaussian assumption (see Shumway et al, 1999).

There are a number of conventional approaches to estimating the slowness parameters and hence, the derived velocities and azimuths. Most common are the beam power and F-statistics described in Section 3.1 below and variations on the correlation method, described in (3.4). Note that the F-statistic is a monotone function of the statistic, semblence. We also adapt the Multiple SIgnal Characteristic (MUSIC) algorithm of Schmidt (1979) to seismic arrays. Finally, we look at the Capon (1969) "high-resolution" estimator as another variation based on the eigen vectors and eigen values of the spectral matrix.

The following Section 4 introduces the multiple signal approach using the partial F-statistic which is shown to improve on the conventional estimators in Section 5.

3. Conventional Approaches

We summarize first some common approaches to determining the velocity and azimuth of a possible mixture of propagating signals. The model is expressed in terms of slowness in seconds per km $\boldsymbol{\theta}_1 = (\theta_1, \theta_2)'$, where θ_1 and θ_2 are the horizontal and vertical components. Velocity $c = 1/||\boldsymbol{\theta}||$ and back azimuth $\alpha = \tan^{-1}(\theta_2/\theta_1)$ are the propogating plane wave velocities in km/sec and back azimuths from due north. We will see later that the techniques described below do not work well on some common mixtures.

3.1 Beam Power and F-statistics

Consider the single signal version of (1), namely, take $M = 1$ so that the model becomes

$$y_j(t) = s_1(t + r_j'\boldsymbol{\theta}_1) + v_j(t) \tag{6}$$

for the j^{th} sensor. Transforming this model to the frequency domain and stacking the observed series yields

$$\boldsymbol{Y}_\ell = z_{\ell 1}(\boldsymbol{\theta}_1)S_{\ell 1} + \boldsymbol{V}_\ell \tag{7}$$

where

$$\boldsymbol{z}_{\ell 1}(\boldsymbol{\theta}_1) = (\exp\{2\pi i\omega_\ell\, \boldsymbol{r}_1'\boldsymbol{\theta}_1\}, \ldots, \exp\{2\pi i\omega_\ell\, \boldsymbol{r}_N'\boldsymbol{\theta}_1\})' \tag{8}$$

is the frequency domain version of the beam-steering vector when it is concentrated at the true slowness $\boldsymbol{\theta}_1$. The beam probe

$$B_\ell(\boldsymbol{\theta}) = N^{-1}\boldsymbol{z}_{\ell 1}^*(\boldsymbol{\theta})\boldsymbol{Y}_\ell \tag{9}$$

will be large at $\boldsymbol{\theta} = \boldsymbol{\theta}_1$ and will be small off the beam. If the spectral matrix $f_v(\omega) = s_v(\omega)I_N$ is diagonal with equal spectra on all channels, then the beam power

$$P_B(\boldsymbol{\theta}) = \sum_{\ell=1}^{L} |B_\ell(\boldsymbol{\theta})|^2 \tag{10}$$

is distributed proportionally to a chi-squared random variable with $2L$ degrees of freedom when $\boldsymbol{\theta} = \boldsymbol{\theta}_1$ the true slowness value. Unfortunately the proportionality constant is a function of the unknown noise spectrum $P_v(\omega)$. The noise spectrum can be estimated from a sample of data prior to the signal, but the problems with that practice have already been mentioned. Note that $L = BT$ where B is the bandwidth in Hz, and T is the sample length in seconds.

It is also the case that the value of $\boldsymbol{\theta}$ that maximizes the beam-power (6) also minimizes the mean squared error

$$SSE(\boldsymbol{\theta}) = \sum_{\ell=1}^{L} ||\boldsymbol{Y}_\ell - \boldsymbol{z}_{\ell 1}(\boldsymbol{\theta})B_\ell(\boldsymbol{\theta})||^2, \tag{11}$$

and it can be shown that the likelihood ratio test of no signal in the model (5) leads to the F-Statistic

$$F(\hat{\boldsymbol{\theta}}) = c\,\frac{N\sum_{\ell=1}^{L}|B_\ell(\hat{\boldsymbol{\theta}})|^2}{SSE(\hat{\boldsymbol{\theta}})}, \tag{12}$$

which converges to a non-central F-distribution with $df_1 = 2(L+1)$ and $df_2 = 2[L(N-1)-1]$ degrees of freedom as $\hat{\boldsymbol{\theta}} \rightarrow \boldsymbol{\theta}_1$. The constant is $c = df_2/df_1$ The advantage of the F-statistic is that its distribution does not depend on nuisance parameters so that thresholds can be set directly from the tabulated values. Furthermore, an estimator of the noise power is available from the signal window as $\hat{P}_v(\omega) = SSE(\hat{\boldsymbol{\theta}})/N$.

3.2 Capon's High Resolution Method

Estimators available in the engineering literature depend primarily on the sample spectral matrix and on the average wave-number value over a frequency band. For example Capon's(1971) estimator uses the sample spectral matrix

$$\hat{\Sigma}_y = \sum_{\ell=1}^{L} \boldsymbol{Y}_\ell \boldsymbol{Y}_\ell^* \tag{13}$$

as input to the proposed estimator obtained by maximizing the function

$$C(\boldsymbol{\theta}) = [\boldsymbol{z}_{\bar{\ell}}^*(\boldsymbol{\theta}) \hat{\Sigma}_y^{-1} \boldsymbol{z}_{\bar{\ell}}(\boldsymbol{\theta})]^{-1}, \tag{14}$$

which has a distribution proportional to the chi-squared distribution with $2(L - N + 1)$ degrees of freedom (see Capon and Goodman, 1970). The test statistic is evaluated at at an integer near the mean frequency, say $\bar{\ell}$. The distribution here also depends on the unknown signal and noise spectra P_s and P_v, or on the noise spectrum only when there is no signal, i.e., $P_s = 0$

3.3 The MUSIC Algorithm

The MUSIC approach suggested by Schmidt (1979) has been extended to arrays by Shumway (2002) for possible use in detecting infrasound signals. The approach is based on orthogonality properties of the eigen vectors of the spectral matrix under a simple multivariate random signal model. In general, the test statistic (11) is expected to have large values under a model that assumes an appropriate number of input signals.

An estimator based on the eigen vectors of Σ_y forms the main component of the Multiple SIgnal Characteristic (MUSIC) detector proposed by Schmidt (1979). A good summary of the statistical properties of this estimator is in Stoica and Nehorai (1989). In particular, letting $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N$ be the eigen vectors of Σ_y, take the MUSIC detector as the maximizer of

$$M(\boldsymbol{\theta}) = \left[\boldsymbol{z}_{\bar{\ell}}^*(\boldsymbol{\theta}) \Big(\sum_{j=M+1}^{N} \boldsymbol{e}_j \boldsymbol{e}_j^* \Big) \boldsymbol{z}_{\bar{\ell}}(\boldsymbol{\theta}) \right]^{-1}. \tag{15}$$

where we assume that there are M possible signals.

The approach to determining the number of signals and their velocities and azimuths reduces to trying various values of M in the above equation, say $M = 1, 2, 3, 4$, looking for peaks in the plotted function $M(\boldsymbol{\theta})$.

It is interesting to note that the Capon detectors can also be written in terms of the eigen values and eigen vectors of the spectral matrix. We note that (14) becomes

$$C(\boldsymbol{\theta}) = \left[\boldsymbol{z}_{\bar{\ell}}^*(\boldsymbol{\theta}) \Big(\sum_{j=1}^{N} \frac{1}{\lambda_j} \boldsymbol{e}_j \boldsymbol{e}_j^* \Big) \boldsymbol{z}_{\bar{\ell}}(\boldsymbol{\theta}) \right]^{-1}.$$

The equation exhibits the estimator in terms of a matching of the probe vector with the eigen vectors.

3.4 Correlation Methods

There are several simple variations that use the cross correlation pairs measured over the array to estimate the velocity and azimuth. Note first that the single signal model (5) implies that the sample cross correlation function between $y_j(t)$ and $y_k(t)$ will tend to be maximized at lag $(\boldsymbol{r}_j - \boldsymbol{r}_k)'\boldsymbol{\theta}_1$.

One simple method might be to simply average the cross correlation functions which turns out to be nearly the same as broad band beam-forming. A somewhat better method is to think of the estimated lag T_{jk} that maximizes the cross correlation between $y_j(t)$ and $y_k(t)$ to be represented by the model

$$T_{jk} = (\boldsymbol{r}_j - \boldsymbol{r}_k)'\boldsymbol{\theta}_1 + e_{jk} \tag{16}$$

for $j \neq k = 1 \ldots, N$ and minimize some objective function of the residuals e_{jk}. Tribuleac and Herrin (1997) use the sum of the absolute errors, robustified by dropping out some the largest ones.

The computations done later here simply minimized the sum of squared errors with no robustification. For example, just chose $\boldsymbol{\theta}_1$ to minimize

$$sse = \sum_{j<k} [T_{jk} - (\boldsymbol{r}_j - \boldsymbol{r}_k)'\boldsymbol{\theta}_1]^2. \tag{17}$$

One problem with these approaches is that standard errors of the estimator for $\boldsymbol{\theta}_1$ are not easy to come by. The distribution of the maximizing values T_{jk} is not available, particularly when it is noted that the errors should be highly correlated because of common sensors used in calculating the cross correlations, for example, $\hat{\rho}_{12}, \hat{\rho}_{13}, \ldots, \hat{\rho}_{1N}$ all involve the first sensor.

3.5 Remarks

Given the erratic performance of conventional estimators to be demonstrated later in the case of mixed signals, developing a means for automatically detecting and isolating multiple interfering signals should have high priority. First of all, one would like to have a reliable means of detecting a single signal over noise, and all methods appear to do an acceptable job of indicating local maximums that agree. The question of statistical significance of the single-signal detection rests on the probability of obtaining a larger value for any one of detectors (a)-(d) or the cross correlation method. A satisfactory solution to the problem of statistical significance is available only for the F-statistic, where a rigorous P-value is available based on asymptotic approximations (see Shumway, 1970, 1971, 1983, 2000). The cross correlation is based on the lag corresponding to the local maximum, a function of a highly variable measurement that ignores the fact that one must use cross correlations involving common sensors. The Capon estimator has a distribution (see Capon and Goodman, 1979) that depends on the noise spectrum, a nuisance parameter that would have to be estimated from noise preceding a potential signal. The MUSIC estimator depends on deriving distribution

theory for a statistic depending on the eigen vectors of the spectral matrix and will be unnecessarily complex (see Stoica and Nehorai, 1989), given the erratic performance that the method exhibits when applied to determining the exact number of signals. Nevertheless, the array version of this statistic has not been applied in a seismic context, and we retain it as an option.

4. Multiple Signal Detection and Estimation

It is obvious from the preceding section that the conventional methods are unreliable for detecting the correct number of signals and for estimating their velocities and azimuths. In this section, we develop a sequential extension of the F-statistic for multiple signals that provides effective methodology for identifying the correct number of signals and their estimated velocities and azimuths. We also provide estimators for the uncertainty using a multiple signal bootstrap procedure. Finally, we show how to develop deconvolved waveforms for each of the component signals.

Solutions to the above problems would enable online detection procedures that would identify single signals and would also identify multiple signals or contaminating propagating noises. Such mixtures contribute to errors in arrival time estimation that is essential for location purposes. Uncertainty estimators for velocity and azimuth would help to identify travel time uncertainties in un-calibrated areas. Procedures for estimating velocities of P and S components could be helpful in constructing travel time predictions for separate phases for regional data. Deconvolutions could be used to determine better amplitudes and spectral ratios for discrimination purposes. It should be noted that work on multiple signal models has been common in early seismic research (see Shumway and Dean, 1968, Shumway, 1970, 1971, Blandford et al 1973, Shumway, 1983).

4.1 Maximum Likelihood Estimation

The proposed solutions to the above problems rest squarely on formulating the multiple signal model as a nonlinear regression problem in the frequency domain. We formulate a model for a time series observed at sensor $j = 1, \ldots N$ at time $t = 1, \ldots, n$ as the delayed sum of M signals and noise, of the form (1). In particular, we wish to have an unambiguous sequential procedure for estimating the number of components M and for estimating the slowness matrix Θ for a given value of M. In this analysis, the signals are assumed to be deterministic unknown functions, and the noises are assumed to be stationarily correlated over time but independent from channel to channel. As usual, we work in the frequency domain by vectorizing $y_j(t)$ and taking Discrete Fourier Transforms to obtain the model (2). In general, we evaluate all statistics over $L = BT$ frequencies, where B denotes the bandwidth of interest in Hz and T denotes the sample length in seconds. The $N \times 1$ vector $\boldsymbol{Y}_\ell$ is expressed in terms of the unknown signal transform $\boldsymbol{S}_\ell$ and the $N \times M$ slowness matrix $Z_\ell(\Theta)$ as in (2) and (3). The unknown slowness parameters are collected in the $2 \times M$ matrix $\Theta = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \ldots, \boldsymbol{\theta}_M)$. Regression theory (see Shumway and Stoffer, 2000, Section 5.4) applied

for a given slowness matrix Θ implies that we may estimate the unknown signal by

$$\hat{\boldsymbol{S}}_\ell(\Theta) = \Big[Z_\ell^*(\Theta)Z_\ell(\Theta)\Big]^{-1} Z_\ell^*(\theta)\boldsymbol{Y}_\ell. \tag{18}$$

Generally, $\hat{\boldsymbol{S}}_\ell(\Theta)$ is concentrated out of the likelihood, and we determine $\hat{\Theta}$ as the minimizer of

$$SSE(\Theta) = \sum_{\ell=1}^{L} ||(Y_\ell - Z_\ell(\Theta)\hat{\boldsymbol{S}}_\ell(\Theta))||^2. \tag{19}$$

Searching for a minimum in the squared error (19) over Θ or, equivalently, maximizing the F-statistic for testing $\boldsymbol{S}_\ell(\Theta) = 0$ leads to

$$F(\hat{\Theta}) = c\,\frac{P_B(\hat{\Theta})}{SSE(\hat{\Theta})}, \tag{20}$$

where

$$P_B(\hat{\Theta}) = \sum_{\ell=1}^{L} \boldsymbol{Y}_\ell^* Z_\ell(\hat{\Theta})\Big[Z_\ell^*(\hat{\Theta})Z_\ell(\hat{\Theta})\Big]^{-1} Z_\ell^*(\hat{\Theta})\boldsymbol{Y}_\ell \tag{21}$$

is the generalized beam-power and $c = df_2/df_1$ is the constant required to convert (20) to an F-statistic with $df_1 = 2(L+M)$ and $df_2 = 2[L(N-M)-1]$ degrees of freedom. We call (21) the generalized beam power because it reduces to the ordinary beam power when $M = 1$.

4.2 Sequential Analysis Using Partial F-statistics

In order to develop a sequential F-test for adding in the signal components $\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_L$, we appeal to the likelihood ratio test of $\boldsymbol{S}_\ell = (\boldsymbol{S}_{\ell 1}, \boldsymbol{0})$ against the alternative $\boldsymbol{S}_\ell = (\boldsymbol{S}_{\ell 1}, \boldsymbol{S}_{\ell 2})$ where the signal vector is partitioned into M_1 and M_2 components ($M = M_1 + M_2$) respectively. The likelihood ratio criterion yields

$$\tilde{F} = \tilde{c}\,\frac{SSE(\hat{\hat{\Theta}}_1, \boldsymbol{0}) - SSE(\hat{\Theta}_1, \hat{\Theta}_2)}{SSE(\hat{\Theta}_1, \hat{\Theta}_2)} \tag{22}$$

where $\tilde{c} = df_2/df_1$ and $\hat{\hat{\Theta}}_1$ denotes the estimator for the slowness matrix Θ_1 under the reduced model. The test statistic compares the mean squared error under the hypothesis that the added signal is absent under the hypothesis that the added signal is present (numerator) with the denominator which measures the mean squared error under the full model. Under the null hypothesis, $\tilde{F}$ has an F distribution with with $df_1 = 2M_2(L+1)$ and $df_2 = 2[L(N-M)-1)]$ degrees of freedom. To apply the result, we add the potential signals in sequentially, each time taking M_1 as the number currently in the model and $M_2 = 1$. It is intuitively pleasing that the numerator is essentially the difference between the generalized beam power under the reduced model and the full model. A sequential search proceeds by performing the partial F-test (22) with M taking values $1, 2, 3, \ldots$ and stopping when F is no longer statistically significant.

Note that a different sequential procedure, based on residuals from lower order signal models, was proposed in Smart (1972) who searched for a second signal by first stripping the first signal away, i.e., by computing

$$\hat{\boldsymbol{V}}_{\ell 1} = \boldsymbol{Y}_\ell - Z_{\ell 1}(\hat{\Theta}_1 \hat{\boldsymbol{S}}_{\ell 1}$$

and then selecting $\hat{\Theta}_2$ as the value minimizing

$$\sum_{\ell=1}^{L} ||\hat{\boldsymbol{V}}_{l1} - Z_\ell(\Theta_2)\hat{\boldsymbol{S}}_{\ell 2}||^2.$$

where $\hat{\boldsymbol{S}}_{\ell 1}$ denotes the solution of (18) under the reduced model and $\hat{\boldsymbol{S}}_{\ell 2}(\Theta_2)$ denotes the solution of (18) with the residuals $\hat{\boldsymbol{V}}_{\ell 1}$ as observations.

An additional statistic of interest in applications is the proportion of power accounted for by any given multiple signal model, given as

$$R^2(\hat{\Theta}) = \frac{P_B(\hat{\Theta})}{\sum_{\ell=1}^{L} ||\boldsymbol{Y}_\ell||^2}, \tag{23}$$

which is immediately recognized as the ratio of the generalized beam power to the total power. Smart has suggested using the proportion of power accounted for as a measure of how well the model fits.

4.3 Model Selection

An alternate approach can be taken in terms of a model selection criterion in the AIC family. We take here a modification to the corrected AIC, developed especially for regression models by Hurvich and Tsai (1989). In the case of the signal model (2), we develop a real regression model isomorphic to (2) and then expand in a Taylor's series about the $2M \times 1$ vector $\boldsymbol{\theta}$. Taking the Hurvich-Tsai approach essentially doubles the numbers appearing in the usual form. We obtained

$$AIC_C(M) = \log \frac{SSE(\hat{\Theta})}{NL} + \frac{2NL + 2M(L+1)}{2NL - 2M(L+1) - 2}. \tag{24}$$

The value is computed for signal models with respectively $M = 1, 2, 3, \ldots$ and the value M that minimizes $AIC(M)$ determines the number of signals chosen for the model.

As a confirming check, we also compute the values of the model selection criterion (17) as a function of m and look for a confirming minimum. We illustrate the approach for the long and short period events given in Figures 1 and 3 and show how it leads to resolving the mixtures of signals for these two data sets.

4.4 Confidence Intervals

First, note that we have ended with a multiple signal model and need the variances and covariances of the slowness matrix $\hat{\Theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_M)$. For the vector case, it would be

possible to express the non-linear regression model (2) as a first-order Taylor's series expansion in Θ and $\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_L$. Then, the Gauss-Newton covariance matrix derived from the nonlinear regression analysis prevails. The variance covariance matrix of the estimated velocity and azimuth parameters could then be computed by the delta method. This procedure seems rather involved, and we adopt a simpler approach that depends on a frequency domain version of the bootstrap (Paparoditis and Politis, 1999, Shumway and Stoffer, 2000, p. 245). Note that the residuals

$$\hat{\boldsymbol{V}}_\ell = \boldsymbol{Y}_\ell - Z_\ell(\hat{\Theta})\widehat{\boldsymbol{S}_\ell(\hat{\Theta})} \tag{25}$$

should be approximately independent with mean zero and variance $f_v(\omega)$ under the model. In order to obtain one bootstrap sample, take a random sample of the residuals above, with replacement. Suppose this random sample is denoted by $\boldsymbol{V}_1^{(1)}, \boldsymbol{V}_2^{(1)} \ldots, \boldsymbol{V}_L^{(1)}$. For this initial bootstrap sample, construct a pseudo-sample of observations

$$\boldsymbol{Y}_\ell^{(1)} = Z_\ell(\hat{\Theta})\widehat{\boldsymbol{S}_\ell(\hat{\Theta})} + \boldsymbol{V}_\ell^{(1)} \tag{26}$$

and apply the estimation procedure, obtaining an estimated $\hat{\Theta}^{(1)}$ for this first bootstrap sample. Then, repeat the above sampling procedure, obtaining a large number, say 500, estimated values $\hat{\Theta}^{(1)}, \hat{\Theta}^{(2)}, \ldots, \hat{\Theta}^{(500)}$. Finally, convert these values to velocity and azimuth and use the two sampling distributions to obtain variances and lower and upper confidence intervals for velocity and azimuth.

4.5 Deconvolution of Multiple Signals

It should be noted that the waveforms of the signals recovered from the final model may show features that are not available in the simple beam. The signals may be recovered using the inverse finite Fourier transform of the frequency domain version (14) which will be available over a limited band spanned by the vector signal. In order to estimate the vector of time functions $\boldsymbol{s}(t) = (s_1(t), \ldots, s_M(t))'$, we expand the the frequency range to $(-1/2, 1/2)$ or $(0, 1)$ by taking $\boldsymbol{S}(\omega_\ell) = S_\ell$ over the band and $\boldsymbol{S}(\omega_\ell) = 0$ for $\omega_\ell = \ell/n, \ell = 0, 1, \ldots, n/2$ and by completing to frequencies between $1/2$ and 1 using $\boldsymbol{S}(\omega_{\ell+n/2}) = \text{conj } \boldsymbol{S}(\omega_{\ell-n/2}),$. Taking the inverse Fourier transform

$$\hat{\boldsymbol{s}}(t) = n^{-1/2} \sum_{\ell=0}^{n-1} \hat{\boldsymbol{S}}(\omega_\ell) \exp\{2\pi i \omega_\ell t\} \tag{27}$$

and shifting the coefficients gives the deconvolved series. Note that increasing the original n to $2n$ will eliminate the end effects of the discrete Fourier transform.

5. Multiple Signal Analysis of Recorded Events

The value of multiple signal techniques developed in Section 4.1-4.3 over and above conventional multiple signal techniques such as MUSIC or single signal techniques like the F-statistic (semblence) or the Capon estimator ultimately rests on their success in analyzing real data.

Motivation for the multiple signal analysis is provided initially by analyzing a teleseismic recording known to contain two earthquakes where it is shown that conventional techniques fail and that a combination of the partial F-statistic and model selection separate the recording into two earthquakes and a noise source, known to be generated by a Pacific storm. Verification of the methodology is provided by using a contrived mixture of signals with known velocities and azimuths. A third example involves analyzing a noisy China event where the elimination of propagating noise provides a visual verification of an obscured depth phase.

While the examples make it clear that conventional methods work poorly, that is not the only motivation for using the techniques derived in Section 4. We mention the following specific advances for the new methodology:

A. Nonlinear optimization in the sequential procedure provides computationally more accurate velocities and azimuths than are obtained by grid search methods typically used by conventional methods in Section 3. All velocities and azimuths are estimated by maximum likelihood and hence, will be efficient.

B. The sequential procedure using partial F-statistics and AIC_C shows the results for all possible multiple signal configurations and settles on the one that minimizes AIC_C

C. Uncertainties for the best model in 2. are computed, giving standard deviations and confidence intervals for velocity and azimuths.

D. The separate waveforms of the component signals are estimated using least squares deconvolution in the frequency domain.

We should make it clear from the onset that the same systematic procedure for identifying the components of the possible mixture and estimating the resulting velocities and azimuths of the components will be followed in each of the examples in the following sections. We summarize the approach for each example as:

1. An initial ***time-frequency spectral analysis*** is used to determine an appropriate bandwidth for the conventional and sequential analysis. This analysis can also be used to focus on particular arrival phases and to isolate the appropriate time interval and scale for the sequential analysis.

2. Conventional slowness analysis plots for ***(a) beam power, (b) Capon, (c) F-statistic*** and ***(d) MUSIC*** estimators are used to suggest regions in slowness space where there might be signals. This will often help in setting initial slowness values for the multiple signal optimization sequence.

3. An initial maximum for the number of signals (usually no more than four) and their initial slowness coordinates is used to start a sequential analysis, with the final number of signals chosen by minimizing the fit ***Corrected Information Theory Criterion*** AIC_C. The ***proportion of variability*** accounted for and the ***partial F-statistic***

for adding a signal are shown at each step. The final configuration is determined by a subjective weighting of these criteria.

4. The ***frequency domain bootstrap*** is used to develop 95% confidence intervals for the velocities and azimuths in the optimal configuration.

5. ***Deconvolution*** is used to produce estimated waveforms for the component signals.

5.1 A Teleseismic Recording Containing Two Earthquakes and a Noise Source

We begin by considering the data that precipitated interest in multiple signal models. This data, shown in Figure 1, contains a verifiable mixture of two earthquakes, one from the south of Africa and another from the Philippines observed at the Korean Seismic Array (KSAR). Figure 2 repeats the last channel of this mixture and shows its time-varying spectrum.

The time-frequency spectrum suggests a band (.02-.08 Hz) for analysis and we adopted this range for the slowness statistics shown in Figure 3. The time-frequency spectrum shows two main arrivals with 2-3 possible secondary arrivals. Figure 3 shows the slowness plots for methods (a)-(d) and we note that the beam power and F-statistic show only a single arrival at 203 degrees. All methods show initial arrivals in this range, with the Capon and MUSIC (3 signals assumed) estimators showing bulges in the 150-160 degree range. The correct azimuths are 226 degrees and 198 degrees which are both different than the estimated primary azimuth obtained by all methods. The results are summarized in Table 1 below.

Table 1: Conventional Estimates for Long Period Event

Estimates	F-stat	Correlation	MUSIC	Capon
Azimuth 1	203	204	204	207
Velocity-1	3.5	3.7	3.7	4.1
Azimuth-2			158	149
Velocity-2			5.3	4.5

We complete Step 3. by setting an initial guess for the maximum signals at 4, with initial guesses for the nonlinear optimization focusing first on the third quadrant (-.1 -.1), with a possible secondary in (.1, -.1). The other two start slowness values were (.1, .1) and (-.1, .1). Table 2 below shows the results of the model selection procedure and the results indicate three signals by AIC_C and possibly four signals by the partial F-test. Values around 2 are marginal for this particular F and we settle on three signals.

Table 2: Model Selection for Long Period Event

Model	R^2	AIC_C	Partial F-stat
1-Signal	.86	-1.25	35.1
2-Signal	.95	-1.87	9.8
3-Signal	.98	-2.21	7.4
4-Signal	.99	-1.57	2.2

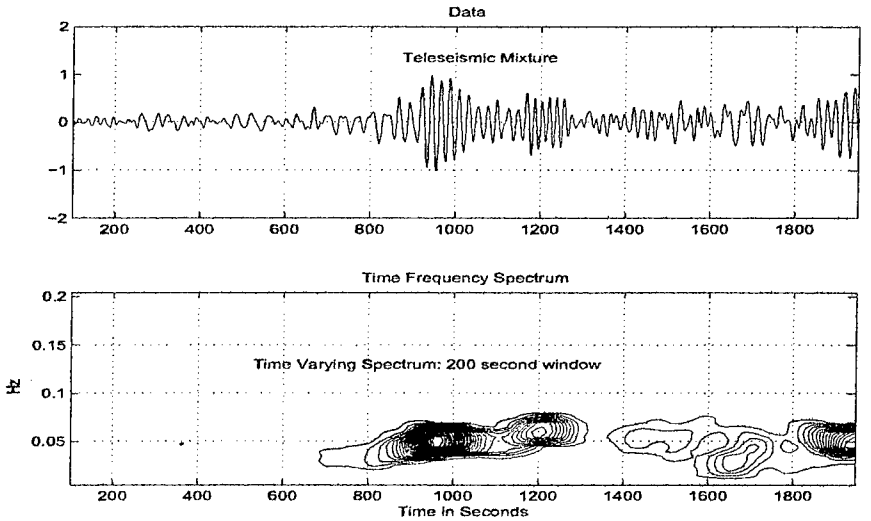

Figure 2: Teleseismic signal of Figure 1 at Korean Seismic Array with time-varying spectrum.

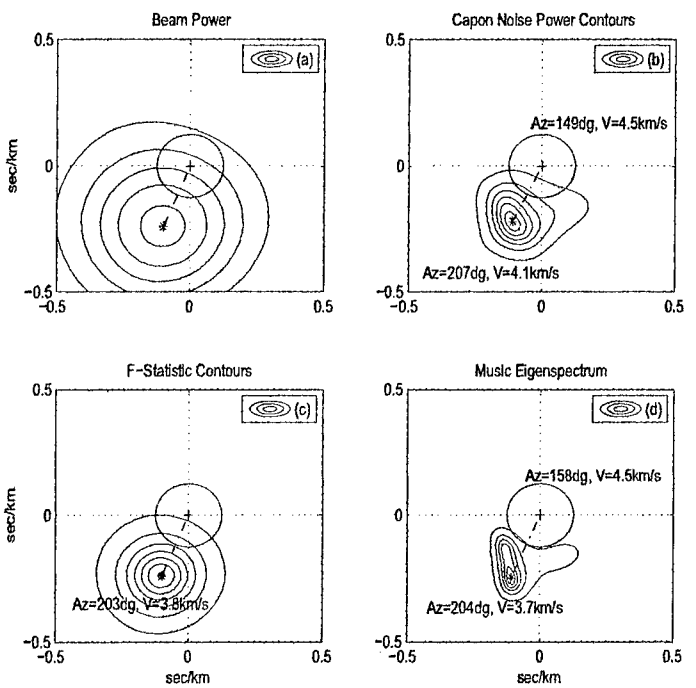

Figure 3: Slowness plots for the teleseismic signal in Figure 1. The circle denotes a velocity of 3km/sec.

The final results, including the uncertainty estimates using the bootstrap as in Step 4 of the suggested procedure leads, are summarized in Table 3 below. We note that the estimators are still several degrees off the known azimuths but that the confidence intervals include the true values.

Table 3: Summary Estimators and Uncertainties for Long period Event

Parameter	Known Value	Estimate(sd)	95% Confidence
Azimuth 1	198	200(1.7)	197-203
Azimuth 2	-	130(2.2)	125-134
Azimuth 3	226	223(2.9)	216-229

The final step involves deconvolving the signals and the separate components are shown in Figure 4. The signal at 223 degrees looks reasonable for the primary signal, as does the second signal at 200 degrees. The noise at 130 degrees seems large and there must be some cancelation for the three to add to the observed series. For this reason, it seems essential to construct a mixture where the component signals are known and we do this in the next section.

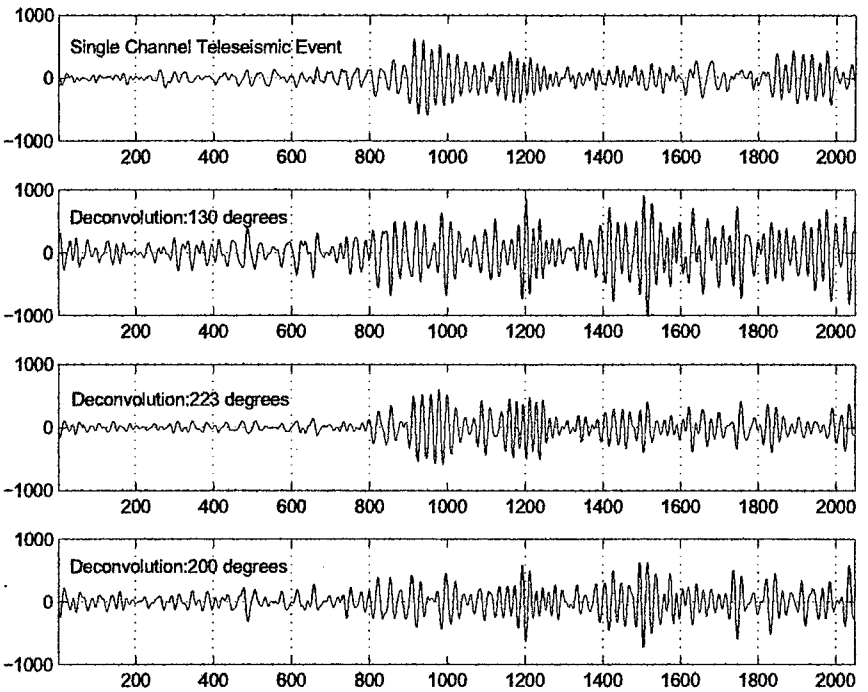

Figure 4: Deconvolution of two identified signals at known azimuths of 226 and 198 degrees and a third unidentified signal or noise source at 135 degrees at a long period array.

5.2 Two Regional Events and a Contrived Mixture

In order to verify that the procedures described in Sections 4.1-4.5 perform as claimed, we took regional signals recorded at the Korean seismic array from 38 degrees and 196 degrees respectively and first analyzed them separately using Steps 1-4. Figure 5 shows the time-frequency spectrum of the regional signal and suggests a frequency range of .3-3.3 Hz for analysis. The first 10 seconds seems to contain a mixture of three different frequencies, centered at about .75,1.5 and 2.8 Hz respectively. The signal is quite impulsive. Figure 6 shows a strong arrival at 38 degrees (F-statistic and cross correlation) or 39 degrees (Capon and MUSIC). The results are summamrized in Table 4 below.

Table 4: Conventional Estimates for Regional Signal from 38 Degrees

Estimates	F-stat	Correlation	MUSIC	Capon
Azimuth	38	38	39	39
Velocity	11.9	11.9	15.6	15.6

Step 3 can be performed as a check against the possibility of additional signals. Starting with the assumption that there are a maximum of two signals present, we chose (.1, .1) for the first slowness and (-.1, -.1) for the start point in a search for a potential additional signal, we obtained the results as summarized in Table 5 below. We note that AIC_C is minimized for the single signal model and that the partial F-statistic of 1.7 is not large enough to add the second signal. Note also that the percentage of variation accounted for (57%) is much smaller for this noisier regional data.

Table 5: Model Selection for 38 Degree Regional Signal

Model	R^2	AIC_C	Partial F-stat
1-Signal	.57	-2.96	22.8
2-Signal	.61	-2.93	1.7

The single signal model then gives estimated azimuths of 38.3(.3) degrees with the 95% confidence interval (37.9-39.0) via the bootstrap. The estimated velocity was 11.9(.1) km/sec with 95% confidence interval (11.8-12.1) km/sec. Using the slowness vector (.0519, .0657) corresponding to the these velocities and azimuths leads to the deconvolution estimator shown in Figure 9.

As a second regional event, consider another regional event, shown in Figure 7, recorded at the same array. This particular event generated a longer lower frequency signal, as can be seen from the single channel and time-frequency spectrum shown in Figure 7. Figure 8 shows the slowness plots for this second event and it can be seen that all methods indicate a single peak for velocity and azimuth, with the summary results as given in Table 6.

We note here the beginnings of a persistent tendency of the Capon and MUSIC estimators to produce higher than normal velocities. As can be seen in (14) and (15), both of these

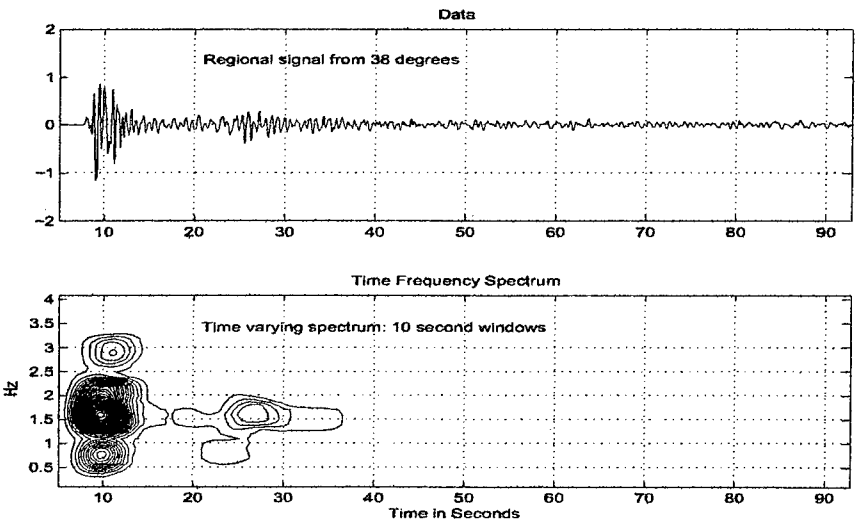

Figure 5: Regional signal (38 degrees) at Korean Seismic Array with time-varying spectrum.

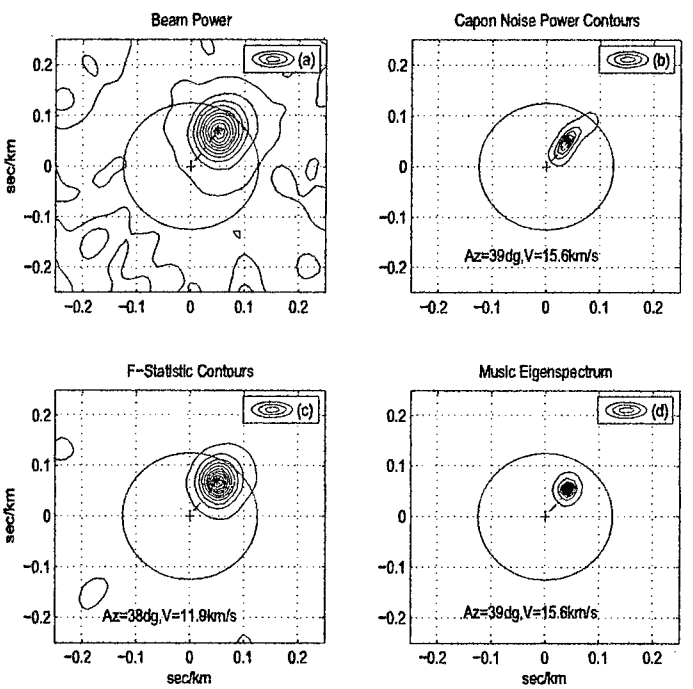

Figure 6: Slowness plots for the regional signal from 38 degrees at Korean Seismic Array. The circle denotes a velocity of 8km/sec.

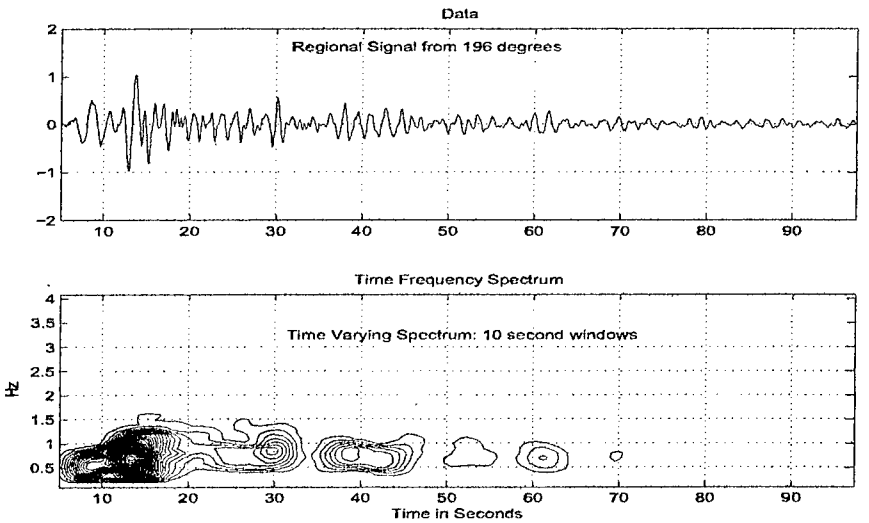

Figure 7: Regional signal (196 degrees) at Korean Seismic Array with time-varying spectrum.

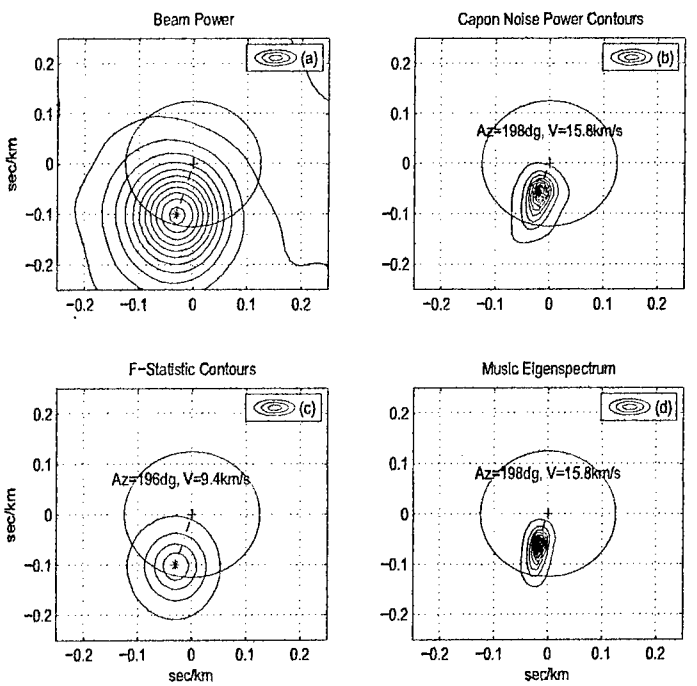

Figure 8: Slowness plots for the regional signal from 196 degrees. The circle denotes a velocity of 8km/sec.

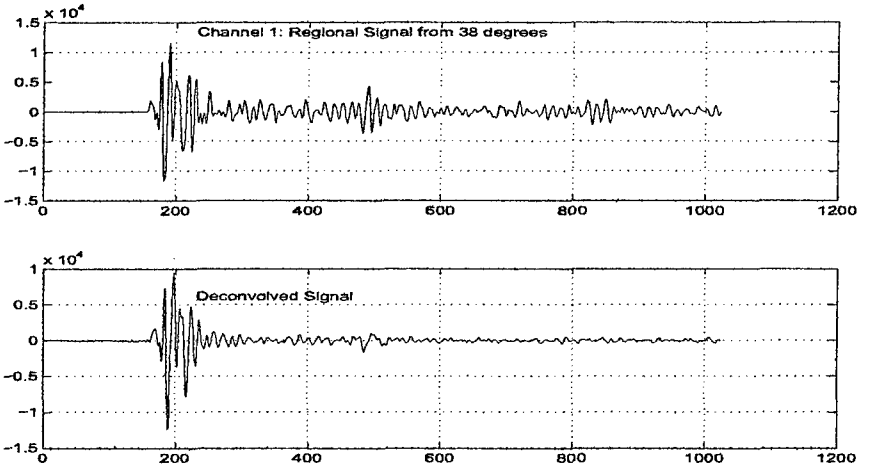

Figure 9: Deconvolution of regional signal from 38 degrees.

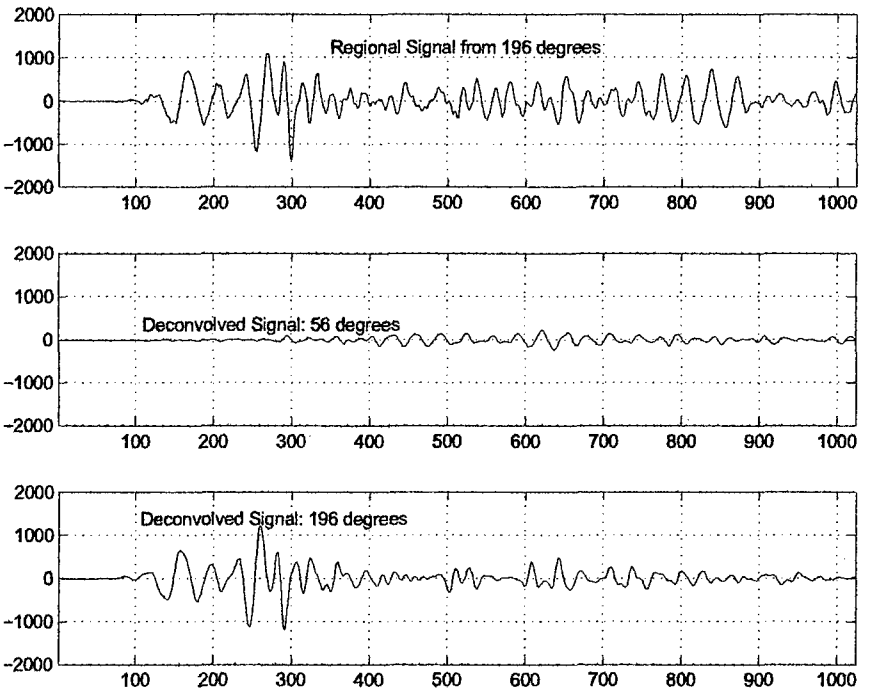

Figure 10: Deconvolution of regional signal from 196 degrees.

estimators involve the spectral matrix (13) which must be estimated from a band of of $L > N$ frequencies where N denotes the number of sensors. The matching vector $z_{\bar{\ell}}(\boldsymbol{\theta})$ is chosen at the center $\bar{\ell}$ corresponding to the average frequency over the interval. It might be better to modify both statistics by averaging over the single frequency matches but we have not tried this approach. The azimuths for all methods match quite nicely.

Table 6: Conventional Estimates for Regional Signal from 196 degrees

Estimates	F-stat	Correlation	MUSIC	Capon
Azimuth	196	196	198	198
Velocity	9.4	9.2	15.8	15.8

Table 7: Model Selection for 196 Degree Regional Signal

Model	R^2	AIC_C	Partial F-stat
1-Signal	.59	-1.61	24.3
2-Signal	.65	-1.62	7.5
3-Signal	.68	-1.56	1.4

The model selection results for this particular regional signal indicate that a second contaminating process may be present. The model selection statistic AIC_C is minimized for the two-signal model and the F-statistic for this signal may be too large to ignore. Our tentative conclusion for this event is that there is a primary signal at 196(.6) degrees with 95% confidence interval (195-197) degrees and a secondary signal at 56(1.1) degrees with 95% confidence interval (55-59) degrees. The velocity of this second signal is slower at 3.1(.05) km/sec. However, the joint deconvolution of the two components using the slowness vector resulting from the two-signal model is shown in Figure 10. Clearly, the results show that the contaminating signal has low amplitude and we note parenthetically that it probably will not exert a significant effect on the contrived mixture.

As a test of the estimation procedure, the two previous arrays were mixed by adding the two events together on common channels with equal amplitudes. Figure 11 shows the contrived mixture along with its time-frequency spectrum. Note first that the time-frequency spectrum suggest the frequency range .2-3.3 Hz and this is what was used as input. The slowness plots all show both signals as being present at the correct azimuths. Note that the Capon and MUSIC again over-estimate the velocities if the average frequency of 1.75 Hz is chosen. Choosing .9 Hz on the basis of the time-frequency spectrum leads to the more reasonable estimates shown in the plot and in Table 8.

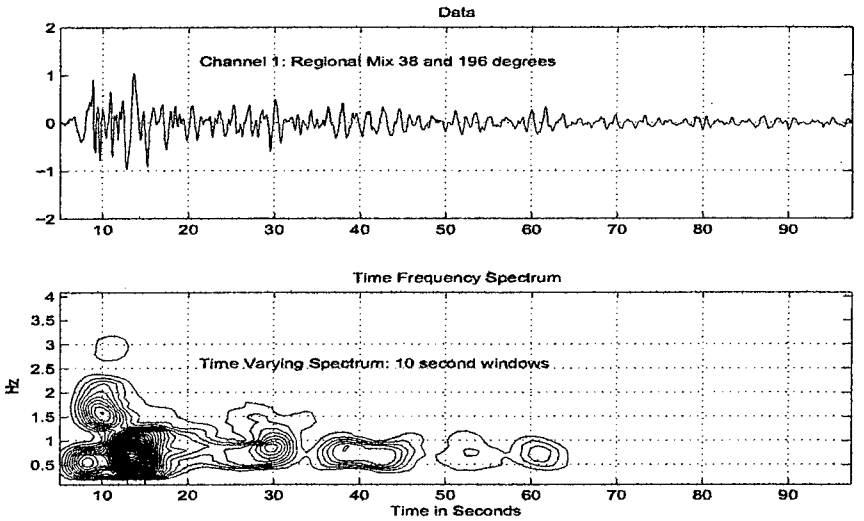

Figure 11: Contrived mixture of regional signals (38 degrees, 196 degrees) at Korean Seismic Array with time-varying spectrum.

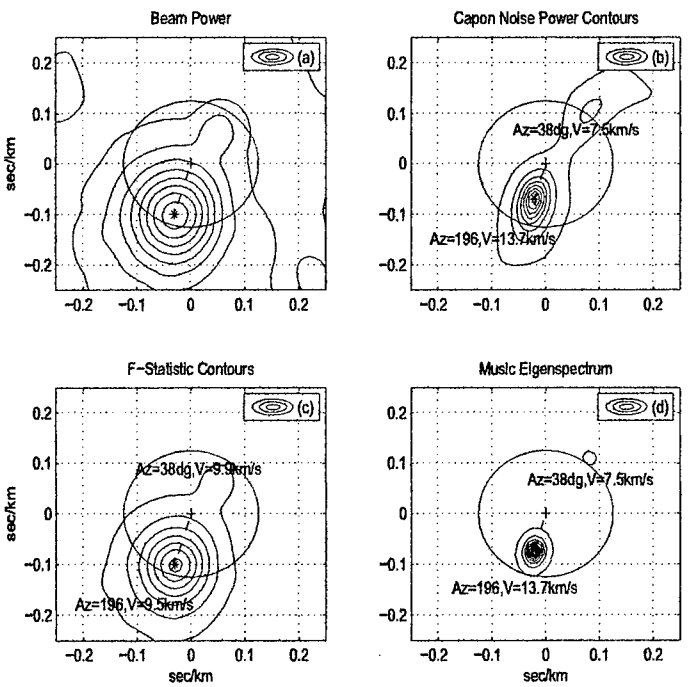

Figure 12: Slowness plots for the mixture of regional signals. The circle represents a constant velocity of 8km/sec.

Table 8: Conventional Estimates for Contrived Regional Mixture

Estimates	F-stat	Correlation	MUSIC	Capon
Azimuth-1	196	196	196	196
Velocity-1	9.5	9.4	13.7	13.7
Azimuth-2	38		38	38
Velocity-2	9.9		7.5	7.5

It was noted in Figure 12 that the slowness plots for the Capon and MUSIC (with two signals assumed) detectors showed the second weaker signal to be present at fairly clearly defined azimuths. While this simple case is one where a weaker but well separated signal is detected by these two methods, tests of significance and confidence intervals are not provided. The single signal version of the F-statistic shows only a weak indication of a second signal so that it is clear that more must be done.

Hence, we apply the partial F-statistic and AIC_C, fitting a sequence of models with $M = 1, 2, 3$ signals. The results, summarized in Table 9 below, show the correct number of signals by either the model selection or partial F-statistic criteria. Therefore, the two-signal model can be used to develop the final estimators and confidence intervals for velocity and azimuth which will produce the deconvolution of the two signals.

Table 9: Model Selection for Contrived Regional Mixture

Model	R^2	AIC_C	Partial F-stat
1-Signal	.46	-1.90	14.5
2-Signal	.60	-2.05	5.3
3-Signal	.64	-2.00	1.6

Table 10 summarizes the final estimators and their uncertainties for the optimum two-signal model. Note that both azimuth estimators are close to those obtained for the separate event analyses so this simple example verifies that the methodology used for the long period event is reasonable.

Table 10: Summary Estimators and Uncertainties for Contrived Mixture

Parameter	Known Value	Estimate(sd)	95% Confidence
Azimuth 1	196	195.5(.62)	194.1-196.5
Velocity 1		9.5(.11)	9.2-9.7
Azimuth 2	38	38.4(.88)	37.8-41.2
Velocity 2		12.0(.14)	11.8-12.3

Finally, Figure 13 shows the result of the least squares deconvolution of the two signals. Comparing these deconvolved signals with the single-event deconvolutions in Figures 9 and 10 shows excellent agreement.

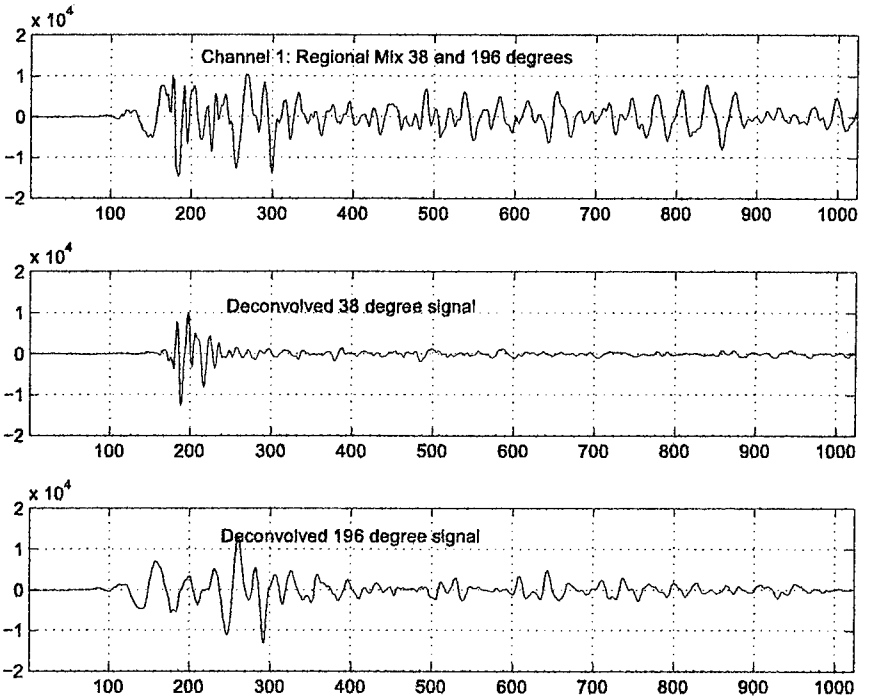

Figure 13: Deconvolution of mixture (38, 196 degrees). degrees.

5.3 A Regional Event from China With Depth Phase Obscured by Noise

In this example, we look further at the event analyzed by Stroujkova and Rieter (2006) in their preliminary report. The event, from NE China, had magnitude of 4.7, with published depths between 0 and 15 kilometers. Because of the high noise level, as can be noted from Figure 14, there will be difficulty in establishing whether or not there is a depth phase. We show in this example how a multiple signal approach eliminates enough of the noise so that the deconvolved signal shows a clear depth phase.

The time-frequency analysis in Figure 14 covers a relatively long time interval. The broader time interval is primarily to emphasize the multiple arrivals and relatively complex structure of the single channel time-frequency spectra.

Figure 15 shows the slowness analysis for this event and we note that the conventional estimators show a primary arrival in the neighborhood of 284 degree except for the correlation analysis which is off by about 40 degrees. Furthermore, the Capon and MUSIC estimators show a secondary peak in the 150-160 degree range.

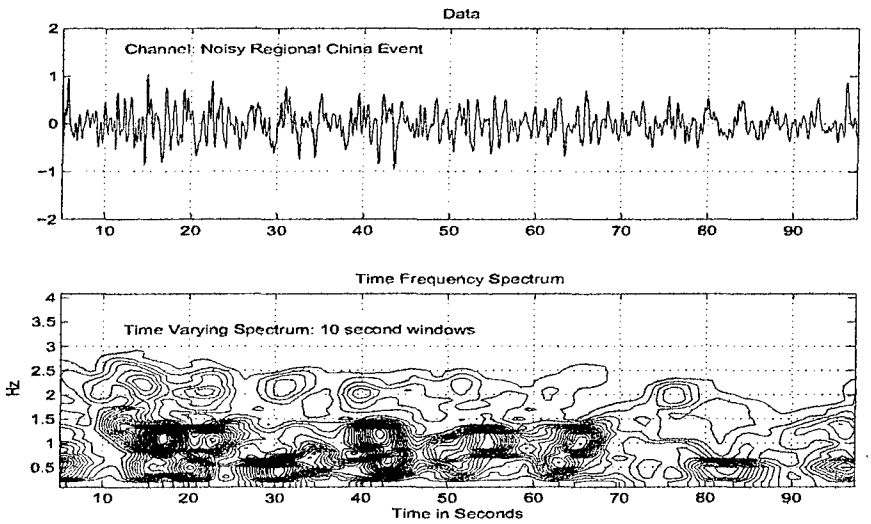

Figure 14: The 1991 China event and its time varying spectrum.

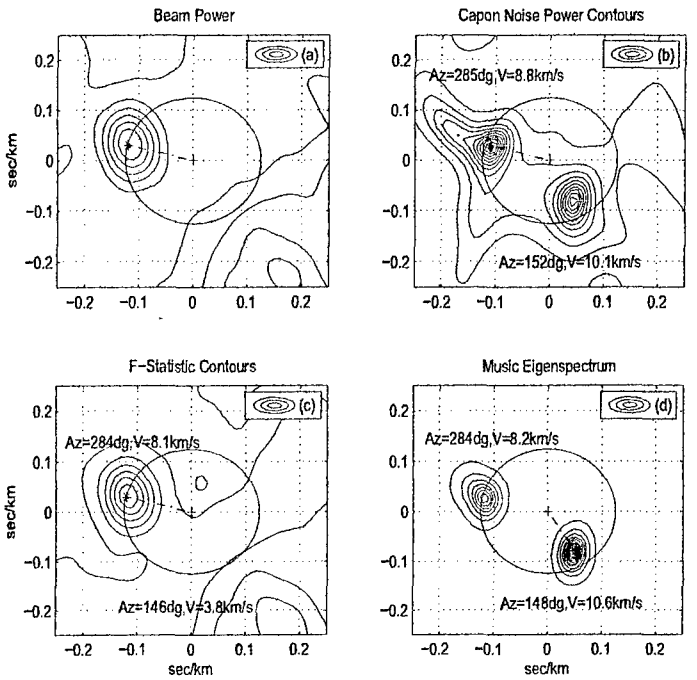

Figure 15: Slowness plots for the 1991 China event. The circle denotes a velocity of 8km/sec.

Table 11: Conventional Estimates for Regional Event From China

Estimates	F-stat	Correlation	MUSIC	Capon
Azimuth-1	287	247	284	286
Velocity-1	8.1	24.9	12.1	13.7
Azimuth-2			153	159
Velocity-2			13.5	14.1

The results of a combined multiple signal analysis using the partial F-statistics and the model selection criterion AIC_C are shown in Table 12. Note that two signals each give statistically significant added power, as measured by the partial F-statistics and that the third one does not add substantially to the power. Noise is still quite high, with only 50% of the power accounted for by the two signals. accounted for by the two main signals. The estimated velocities and azimuths of the two signals are 287 degrees and 8.2 km/sec for the first and 151 degrees and 3.4 km/sec for the second signal. The estimated velocity of the second signal differs from that estimated by MUSIC but some runs at different frequency centers yielded estimators in that range for the other estimators as well. The estimators are summarized in Table 13.

Table 12: Model Selection for China Event

Model	R^2	AIC_C	Partial F-stat
1-Signal	.34	-.40	8.1
2-Signal	.46	-.43	3.4
3-Signal	.52	-.41	1.6

Table 13: Summary Estimators and Uncertainties for Noisy China Event

Parameter	Estimate(sd)	95% Confidence
Azimuth 1	150.8(1.3)	148.2-153.8
Velocity 1	3.4(.08)	3.3-3.6
Azimuth 2	286.8(1.1)	284.5-288.4
Velocity 2	8.2(.11)	8.0-8.5

The deconvolution corresponding to the two components noted in the previous section, shown in Figure 16, gives the estimated signal and noise components in the regional data. The estimated noise component from 148 degrees shows the character of the low frequency noise before the signal enters. The estimated signal from 286 degrees shows the depth phase pP much more clearly than does the single channel mixture. In this case, one obtains from the plot a delay of about 4.6 seconds which is comparable to that obtained by Stroujkova and Reiter (2006). We also note the noise reduction capabilities of the two-signal beam.

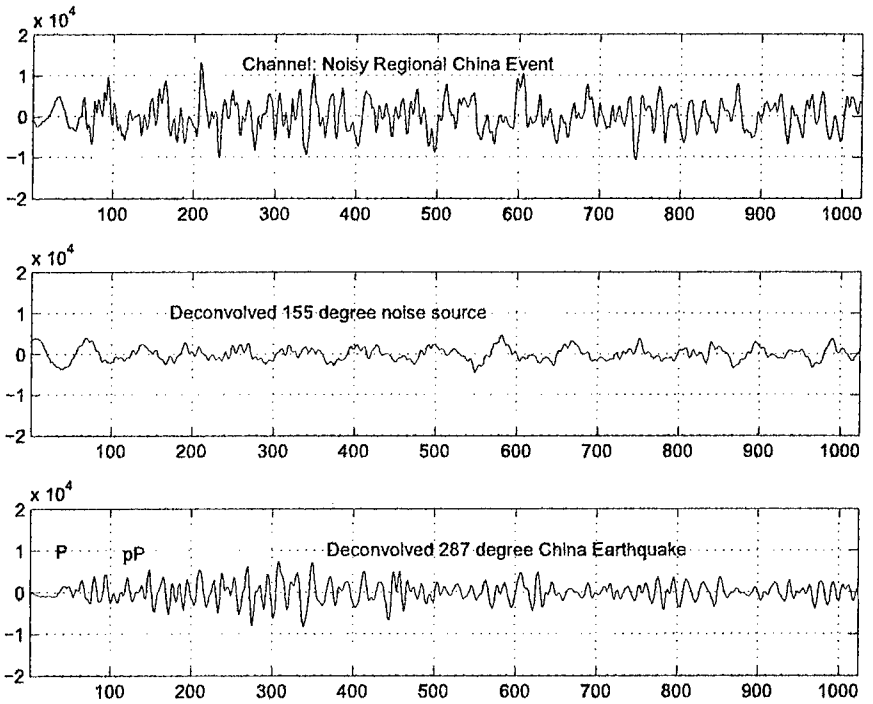

Figure 16: Deconvolution of Chinese event into a noise source and a 287 degree signal.

6. Software Documentation and Data Files

6.1 General Structure

The software delivered with this final report is structured as a main driving program with function subroutines for computing and plotting all quantities given in Sections 3, 4 and 5 of this report. The one exception is the computation and plotting for the original series and its time-frequency spectrum which we assume will be available to any user. The main program, as indicated below calls the five functions that are needed to perform the multiple signal analysis.

The function *beam.m* plots four conventional estimators (beam, Capon, F-statistic and MUSIC) in slowness space and gives the estimated velocities and azimuths. A separate function *corr_sl.m* gives the broad-band correlation estimators. For the contrived mixture of signals shown in the text, the two functions produce Figure 12 and the entries necessary for Table 8.

The function *mul_sig.m* develops the multiple signal analysis described in Section 4. This gives the sequence of partial F-statistics and values of AIC_C necessary to settle on on optimum number of signals. The function identifies the optimum number of signals and gives the corresponding slownesses for input to the deconvolution. The model selection criteria

are computed for inclusion in Table 9. The estimators for velocities and azimuths corresponding to the optimum model are given and the option to compute standard errors and 95% confidence intervals is included in the function *conf_int.m*. This would complete the entries for Table 10. Finally, the option for deconvolving the component signals is given by *decon.m*. We summarize the structure below where each main function is identified with the appropriate section of the report:

Main Program: *M_signal*

1. function *beam.m*(3.1-3.3)
   - function *fprob.m*
   - function *v_az.m*
2. function *corr_sl.m*(3.4)
3. function *mul_sig.m*(4.2-4.3)
   - function *fprob.m*
   - function *v_az.m*
   - function *SSEL.m*
4. function *decon.m*(4.5)
5. function *conf_int.m*(4.4)
   - function *v_az.m*
   - function *SSEL.m*

Note that there are three additional functions that are used by certain some of the five main functions. The function *fprob.m* computes the P-value for the F-statistics obtained at various stages. The function *v_az.m* computes the velocity and azimuth corresponding to a given slowness. The squared error objective function (19) in computed in the subroutine *SSEL.m*.

6.2 Data

We summarize the data files provided in the report in Table 14 along with the sections where an analysis is provided. More details for the inputs are given in Section 3.3

Table 14: Data Provided with Report

Description	Section	Data File	Array Coordinates
1. Teleseismic	5.1	k_array	r_k_array
2. Region(38dg)	5.2	ks1994202	r_ksar
3. Region(196dg)	5.2	ks199517	r_ksar
4. Mixture	5.2	ks9495mix	r_ksar
5. China Regional	5.3	ks1991101	r_ksar_1991_101

6.3 Main Program and Function Subroutines

The code for the main program is given below with the calls to the function subroutines. The inputs are the data files and sensor locations, along with the sample rate, frequency and slowness ranges and an initial slowness vector for the largest model to consider.

In general, starting the vector with four bivariate slowness coordinates, i.e. sl_full=$(sl_{11}, sl_{12}, sl_{21}, sl_{22}, \ldots sl$ has been fairly effective. The single-signal model, $M = 1$ uses the first two coordinates as start values, the $M = 2$ signal model uses the first two pairs and the $M = 4$ model uses all four pairs as start value. Best results were obtained by starting with vectors in the middle of the quadrant, e.g. $(s_{11}, s_{21}) = (.1, .1)$. The recommended pattern will be clear for the example in the next section.

```
 % M_Signal.m
% Multiple signal analysis
% Reads Data and Controls the Separate Analysis Features
    % Conventional (Beam, Capon, F-statistic, MUSIC, Correlation)
        % Sections 3.1-3.4 of Shumway (2006)
    % Multiple Signal Analysis (Partial F, AIC, % Variation)
        % Sections 4.1-4.3 of Shumway (2006)
    % Bootstrap Confidence Intervals
        % Section 4.4 of Shumway (2006)
    % Deconvolution
        % Section 4.5 of Shumway (2006)
% Input Information
    % Data File (data): n(points) X N(sensors) rectangular array
    % Sensor Locations (r): N(sensors) X 2 Coordinates (E, N)
    % Sample Rate (sr): points/s
    % Frequency Band: (LO, HI) in Hz
    % Slowness Values: Range for s=(s_1,s_2)
    % Initial Slownesses for Multiple Signal Analysis(sl_full)
    % rd=1/V_c, V_c=constant velocity circle in slowness plot

id=input('File Number')

M=input('Assumed number of signals for MUSIC estimator')
% Set the maximum lag for the cross correlation estimator
mxlag=input('Maximum lag in pts for correlation estimator (25)')

% Load Data
if id==1
    load k_array.dat, data=k_array;
    % Sampling rate, frequency range
    sr=1,LO=.02,HI=.08
```

```
    % Slowness range
    s=-.5:.01:.5;,rd=1/8;
    % Array Coordinates (km) East-North
    sl_full=[-.1 -.1 .1 -.1 .1 .1 -.1 .1]
    load r_k_array, r=r_k_array;
end

if id==2
    load ks1994202, data=ks1994202;
    % Sampling rate, center frequency, bandwidth
    sr=20,LO=.3, HI=3.3
    % Slowness Range
    s=-.25:.01:.25;, rd=1/8;
    % Array Coordinates (km) East-North
    load r_ksar, r=r_ksar;
    sl_full=[.1 .1 -.1 -.1]
end

    if id==3
    load ks199517, data=ks199517;
    % Sampling rate, center frequency, bandwidth
    sr=20,LO=.2,HI=2.0
    % Slowness Range
    s=-.25:.01:.25;, rd=1/8;
    % Array Coordinates (km) East-North
    load r_ksar, r=r_ksar;
    sl_full=[-.1 -.1 .1 .1 .1 -.1 ]
end

if id==4
    load ks9495mix, data=ks9495mix;
    % Sampling rate, frequency range
    sr=20,fo=1, LO=.2, HI=3.3
    % Slowness Range
    s=-.25:.01:.25;, rd=1/8;
    % Array Coordinates (km) East-North
    load r_ksar, r=r_ksar;
    sl_full=[-.1 -.1 .1 .1 .1 -.1]
end

if id==5
   load ks1991101
```

```
   sr=20
% LO=.3;, HI=3.0
LO=.2;, HI=1.5
    data=ks1991101(3501:4524,:);
  % data=ks1991101(2501:3524,:);
rd=1/8;
 s=[-.25:.01:.25];
%  s=[-.4:.01:.4];
sl_full=[.1 -.1 -.1 .1 -.1 -.1 ]
 load r_ksar_1991_101, r=r_ksar_1991_101;
 end

 % Sections 3.1-3.3 of Shumway(2006)
beam(data,r,sr,LO,HI,s,rd,M);

% Section 3.4 of Shumway (2006)
[az_corr, vel_corr]=corr_sl(data,r,sr,mxlag)

% Section 4.2-4.3 of Shumway (2006)
[M_min,az_min,vel_min,sl_min]=mul_sig(data,r,sr,LO,HI,sl_full)

% Section 4.5 of Shumway (2006)
idecon=input('1 for deconvolution, 0 otherwise') if idecon==1
    jdecon=input('0 for optimum input, 1 for user input')
    if jdecon==1
        sl_min=input('Slowness vector')
    end

    decon(data,r,sr,LO,HI,sl_min)
end

% Section 4.4 of Shumway (2006)
id_boot=input('1 if bootstrap CI 0 otherwise Bootstrap takes time')
if id_boot==1
   n_boot=input('# of bootstrap reps, 500 is suggested')
end
 if id_boot==1
  conf_int(data,r,sr,LO,HI,sl_min,n_boot)
 end
```

6.4 A Worked Example

We illustrate a run of the program for the mixture of signals considered in Section 5.2. For this case, we restricted the start slowness vector to a maximum of three signals to cut down on the printout. For sl_full=[-.1 -.1 .1 .1 .1 -.1], we looked at a potential for $M = 3$ signals with the primary searches in quadrants containing suspected signals. Different start configurations can produce different looking end results so the above start value should be checked against other permutations of the two dimensional slowness vectors.

Two plots are produced which are not shown. These are exactly Figures 12 and 13 with no labeling on the plots. Generally, labeling is data dependent and can be done using the gtext('label') instruction. The axes can also be changed after the fact by statements like axis([1 1024 -20000 20000]).

```
>> M_signal
File Number4 id =     4
 Assumed number of signals for MUSIC
estimator2
 M =     2
Starting with file number 4 and the 2-signal assumptions for MUSIC

Maximum lag in pts for correlation estimator (25)25
 The starting value should cover the maximum time lag (positive or
 negative.

mxlag =    25

sr =    20 fo =     1 LO =    0.2000  HI =    3.3000
 Sampling rate and frequency ranges are typical for regional event

sl_full =

   -0.1000   -0.1000    0.1000    0.1000    0.1000   -0.1000

avg_frequency =    1.7500
 The average of high and low cutoffs may not be the best for the
 probe vector used in the Capon and MUSIC estimators

Input user center frequency.75

fo =    0.7500

R =
```

```
    8.3774    0.0231
    0.0231    5.5422

F_stat =   14.5768
 Single-signal F-statistic and its P-Value
P_val =  7.6690e-019

az_F =  196.6992

vel_F =    9.5783

az_C =  198.4349

vel_C =   10.5409

az_M =  198.4349

vel_M =   10.5409

az_corr =  195.5305

vel_corr =    9.3569


M =     3

 Single Signal Model Results
sl_est =   -0.0287   -0.1018

No_Signals =     1

F_Azimuths =  195.7325

F_Velocities =    9.4540

 Double Signal Model Results
sl_est =   -0.0281   -0.1015    0.0516    0.0651

No_Signals =     1     2

F_Azimuths =  195.4661   38.4138
```

```
F_Velocities =    9.4983   12.0364

 Triple Signal Model Results

sl_est =   -0.0279   -0.1020    0.0514    0.0650   -0.0506   -0.3025

No_Signals =     1     2     3

F_Azimuths =  195.2898   38.3312  189.4961

F_Velocities =    9.4615   12.0728    3.2607

R_squared =    0.4620    0.5961    0.6356

Corrected_AIC =   -1.8968   -2.0507   -2.0029

Error_SS =  281.6107  211.3944  190.7318

Total_SS =  523.3957

Partial F-statistics

Number_signals =     1
F_value =   14.5475

Number_signals =     2
F_value =   14.5475    5.2958

Number_signals =     3
F_value =   14.5475    5.2958    1.6189

 Best Model
M_min =     2
 az_min =  195.4661   38.4138
 vel_min =9.498312.0364
 sl_min =   -0.0281   -0.1015    0.0516    0.0651

1 for deconvolution, 0 otherwise1
 idecon =     1

0 for optimum input, 1 for user input0
 jdecon =     0
```

```
M =     2

nplt =     3

1 if bootstrap CI 0 otherwise Bootstrap takes time1

id_boot =     1

# of bootstrap reps, 500 is suggested200

n_boot =
   200

M =     2
 Standard deviations for Azimuths and Velocitie
std_Az =    0.6476    0.9115

std_Vel =    0.1076    0.1475 95%
 Confidence Intervals
 Az_025 =  193.9662   37.7199

Az_975 =  196.4673   41.3768

Vel_025 =    9.2418   11.7695

Vel_975 =    9.6738   12.3464
```

7. Discussion

This work was primarily motivated by the problem of detecting mixtures of signals on teleseismic and regional arrays. Such undetected mixtures are shown to give incorrect velocities and azimuths when treated by traditional single-signal detection methods based on cross-correlation or frequency wave-number methods. Furthermore, the separation of interfering phases and coherent noise sources should improve detection statistics and lead to improvements in location and magnitude estimates.

Using current improved computing platforms such as MATLAB make the nonlinear estimation problems implicit in multiple signal modeling tractable and easy to implement. Using early work involving multiple signal estimation (Shumway, 1970) and its extensions to wave-number methods (Smart, 1972, 1976), we were able to formulate the problem in a regression framework that led to a sequential detection approach for reliably determining the number of signals and coherent noises present along with their estimated velocities and azimuths. We

have provided methods for comparing detection performance using the F-statistic, the AIC_C model selection statistic, and confidence intervals for the velocities and azimuths using the bootstrap. Finally, we show that the regression model provides a method for deconvolving the component signals and exhibit results for both long and short period seismic data.

8. Acknowledgements

Dr. Gene Smart and Dean Clauter of the Air Force Technical Applications Center (AFTAC) have provided much of the data used in this analysis as well as valuable insights. In particular, the previous seminal work of Smart (1972, 1976) is acknowledged. I am also grateful to Anastasia Stroujkova and Delaine Reiter of Weston Geophysical, Inc. for providing the data from the regional Chinese event and for making available their preprint.

References

Blandford, R.R. (1970). An automatic event detector at TFO. Seismic Data Laboratory Report No. 263, Teledyne Geotech, Alexandria, Virginia.

Blandford, R.R. (2002). A plan of development for detection systems for seismic and infrasound arrays. *AFTAC-TR-02-005*, Air Force Technical Applications Center, Patrick AFB, FL 32925-002.

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

Capon, J. and N.R. Goodman (1970). Probability distribution for estimators of the frequency wavenumber spectrum. *Proc. IEEE* (letters), 58, 1785-1786.

Clauter, D.A., Personal Communication, 2004.

Hurvich, C.M. and C-L Tsai (1989). Regression and time series model selection in small samples. *J. Time Series Anal.*, 19, 19-46.

Melton, B.S. and L.F. Bailey (1957). Multiple signal correlators. *Geophysics* ,3, 565-588.

Paparoditis, E. and D.N. Politis (1999). The local bootstrap for periodogram statistics. *J. Time Series Analysis*, 20, 193-222.

Schmidt, R.O.(1979). Multiple emitter location and signal parameter estimation. *Proc. RADC Spectral Estimation Workshop*, 243-258. Rome, Italy.

Shumway, R.H. and Dean, W.C.(1968). Best linear unbiased estimation for multivariate stationary processes. *Technometrics*, 10, 523-534.

Shumway, R.H.(1970). Applied regression and analysis of variance for stationary time series. *J. Amer. Statist. Assoc.*, 65, 1527-1546.

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

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. (1999). Signal detection and estimation of directional parameters for multiple arrays. Technical Report DTRA-TR-99-50, Defense Threat Reduction Agency, 8725

John. J. Kingman Road, MS-62301, Fort Belvoir, VA 22060-6201

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

Shumway, R.H. (2002). Detection and location capabilities of multiple infrasound arrays. Final Scientific Report. DTRA01-00-C0082, Preprint.

Smart, E. (1972). FKCOMB, A Fast General-Purpose Array Processor, Report No. 9, Seismic Array Analysis Center, Teledyne Geotech, Alexandria, Virginia.

Smart, E. (1976). Linear high-resolution frequency-wavenumber analysis. PhD Dissertation, Southern Methodist University.

Smart, E. (2003). Develop a mixed signal processor. Personal communication, October 2, 2003.

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

Strouujkova, lA. and D. Reiter (2006). Improving focal depth estimates: Studies of depth-phase detection at regional distances. Preprint for the 28th Seismic Research Review.

Tribuleac, I.M. and E.T. Herrin (1997). Calibration studies at TXAR. *Seis. Res. Lttrs*, 68, 353-365.