SAIC-00/3003

Multi-Sensor Data Fusion Project
Final Report

Ronan J. LeBras

28 February, 2000

Sponsored by:
Defense Threat Reduction Agency (DTRA)
Contract No. DASG60-96-C-0151

Name of Contractor: Science Applications International Corporation
Business Address: 10260 Campus Point Drive
San Diego, CA 92121

Effective Date of Contract: 26 July 1996
Contract Expiration Date: 28 February 2000 (Base Program)
Short-Title of Work: Final Report

SAIC-00/3003

Multi-Sensor Data Fusion Project
Final Report

Ronan J. LeBras

28 February, 2000

Sponsored by:
Defense Threat Reduction Agency (DTRA)
Contract No. DASG60-96-C-0151

Name of Contractor:	Science Applications International Corporation
Business Address:	10260 Campus Point Drive
	San Diego, CA 92121
Effective Date of Contract:	26 July 1996
Contract Expiration Date:	28 February 2000 (Base Program)
Short-Title of Work:	Final Report

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	28 Feb 2000	Final Tech. Report 7/26/96-2/28/00

4. TITLE AND SUBTITLE
Multi-Sensor Data Fusion Project Final Report

5. FUNDING NUMBERS

6. AUTHOR(S)
Ronan J. LeBras

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Science ApplicationsInternational Corporation (SAIC)
10260 Campus Point Drive
San Diego, CA 92121

8. PERFORMING ORGANIZATION REPORT NUMBER
SAIC-00/3003

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Space and Strategic Defense Command (USASSDC)
P. O. Box 1500
Huntsville, AL 35807-3801
Program Mgr. R. Gustafson (DoD/DTRA)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approval for public release pending.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

See attached Executive Summary

14. SUBJECT TERMS
CTBT, Data Fusion, Seismic, Infrasonic, Hydroacoustic, Radionuclide Processing

15. NUMBER OF PAGES
75

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unclassified

1. Executive summary

The Comprehensive Nuclear Test-Ban Treaty (CTBT) calls for the establishment of networks of seismic, hydroacoustic, infrasonic, and radionuclide sensors to monitor the earth for unannounced underground, underwater, and atmospheric nuclear tests.

Seismic processing at the prototype International Data Center (pIDC) has been operating continuously since the GSETT-3 experiment that started in late 1994. Seismic is the most mature of the four monitoring technologies in automatic processing and interactive analysis. The other three technologies are in an earlier stage of development and evaluation.

Since late 1996, efforts have been made under the Multi-Sensor Data Fusion project funded by the Defense Threat Reduction Agency (DTRA) to integrate hydroacoustic and infrasonic monitoring technologies into the existing pIDC automatic processing and interactive review framework. Signal processing of hydroacoustic and infrasonic waveform data was integrated into the pIDC monitoring system, and network processing was enhanced to include two hydroacoustic phase types, *H*-phase (in-water generated) and *T*-phase (underground generated) and one additional infrasound phase, *I*. A joint bulletin including the three waveform technologies is produced at the pIDC on a continuous basis. The bulletin includes arrivals from natural and man-made events recorded by all three sensor types. Enhancements have been made to the travel-time handling system to take into account temporal variations which may be significant for both acoustic technologies. When no specific travel time tables are available for a station, a default constant velocity model is used which makes the software ready for network expansion in the three waveform technologies.

Case studies, network simulations and results from the pIDC operational system were used to normalize and evaluate the capability of data fusion in the pIDC monitoring system.

Case studies were conducted involving marine and atmospheric ground-truth explosions. A set of low-yield marine explosions off the coast of Japan has increased confidence in the path-dependent velocity and attenuation models used in hydroacoustic processing. Atmospheric explosions from the White Sands Missile Range were recorded at two infrasound stations, and we conducted automatic and interactive processing on that data set to gain knowledge about infrasonic propagation and to test our system on this ground-truth data. Mining explosions have been analyzed in the context of joined seismic and infrasonic processing, showing the improved location accuracy expected when azimuth determined from the infrasonic data is added to the seismic data.

The International Monitoring System (IMS) sensor networks are not yet complete. Therefore, simulations were used to test the capabilities of multi-technology association for the planned IMS sensor networks. Simulations were also used to evaluate the capability to detect and locate clandestine nuclear tests under a variety of evasion scenarios.

Results from continuous processing and interpretation of seismic, hydroacoustic and infrasonic data at the pIDC were evaluated. Hydroacoustic processing has been performed at the pIDC since

May 1996. Infrasonic processing has been performed since December 1998. This has provided an extensive data set for operational evaluation.

Less work has been done on fusion of radionuclide data with data from the three waveform technologies. A prototype program to associate events from seismic, hydroacoustic and infrasonic data to radionuclide detections has been developed as a first attempt to fuse waveform technology events with radionuclide detections.

The data fusion model was designed to integrate the four sensor technologies to produce a composite bulletin to meet the current CTBT requirements and schedule. Options to improve the monitoring system to meet future requirements were also evaluated. In particular, improved methods for spatial data management, integration of new processing algorithms, and alternative fusion algorithms and concepts were evaluated.

Table of Contents

1. Executive summary

The Comprehensive Nuclear Test-Ban Treaty (CTBT) calls for the establishment of networks of seismic, hydroacoustic, infrasonic, and radionuclide sensors to monitor the earth for unannounced underground, underwater, and atmospheric nuclear tests.

Seismic processing at the prototype International Data Center (pIDC) has been operating continuously since the GSETT-3 experiment that started in late 1994. Seismic is the most mature of the four monitoring technologies in automatic processing and interactive analysis. The other three technologies are in an earlier stage of development and evaluation.

Since late 1996, efforts have been made under the Multi-Sensor Data Fusion project funded by the Defense Threat Reduction Agency (DTRA) to integrate hydroacoustic and infrasonic monitoring technologies into the existing pIDC automatic processing and interactive review framework. Signal processing of hydroacoustic and infrasonic waveform data was integrated into the pIDC monitoring system, and network processing was enhanced to include two hydroacoustic phase types, *H*-phase (in-water generated) and *T*-phase (underground generated) and one additional infrasound phase, *I*. A joint bulletin including the three waveform technologies is produced at the pIDC on a continuous basis. The bulletin includes arrivals from natural and man-made events recorded by all three sensor types. Enhancements have been made to the travel-time handling system to take into account temporal variations which may be significant for both acoustic technologies. When no specific travel time tables are available for a station, a default constant velocity model is used which makes the software ready for network expansion in the three waveform technologies.

Case studies, network simulations and results from the pIDC operational system were used to normalize and evaluate the capability of data fusion in the pIDC monitoring system.

Case studies were conducted involving marine and atmospheric ground-truth explosions. A set of low-yield marine explosions off the coast of Japan has increased confidence in the path-dependent velocity and attenuation models used in hydroacoustic processing. Atmospheric explosions from the White Sands Missile Range were recorded at two infrasound stations, and we conducted automatic and interactive processing on that data set to gain knowledge about infrasonic propagation and to test our system on this ground-truth data. Mining explosions have been analyzed in the context of joined seismic and infrasonic processing, showing the improved location accuracy expected when azimuth determined from the infrasonic data is added to the seismic data.

The International Monitoring System (IMS) sensor networks are not yet complete. Therefore, simulations were used to test the capabilities of multi-technology association for the planned IMS sensor networks. Simulations were also used to evaluate the capability to detect and locate clandestine nuclear tests under a variety of evasion scenarios.

Results from continuous processing and interpretation of seismic, hydroacoustic and infrasonic data at the pIDC were evaluated. Hydroacoustic processing has been performed at the pIDC since

May 1996. Infrasonic processing has been performed since December 1998. This has provided an extensive data set for operational evaluation.

Less work has been done on fusion of radionuclide data with data from the three waveform technologies. A prototype program to associate events from seismic, hydroacoustic and infrasonic data to radionuclide detections has been developed as a first attempt to fuse waveform technology events with radionuclide detections.

The data fusion model was designed to integrate the four sensor technologies to produce a composite bulletin to meet the current CTBT requirements and schedule. Options to improve the monitoring system to meet future requirements were also evaluated. In particular, improved methods for spatial data management, integration of new processing algorithms, and alternative fusion algorithms and concepts were evaluated.

2. Project objectives

The Multi-Sensor Data Fusion Project, sponsored by the Defense Threat Reduction Agency (DTRA) was initiated in 1996 in support of the prototype International Data Center's (pIDC) effort to integrate operational processing of the four monitoring technologies, seismic, hydroacoustic, infrasonic and radionuclide.

The main objectives of the project were:

- Develop the capability to integrate seismic, hydroacoustic, infrasonic and radionuclide data processing and interpretation in the pIDC monitoring systems to meet current CTBT requirements.
- Design the next-generation system to exploit emerging technologies in data management, artificial intelligence, and data fusion to meet future CTBT requirements.

An initial goal was to bring processing of the acoustic technologies into pIDC operations by integrating newly-developed single-technology processing modules *DFX* and *StaPro* for hydroacoustic and infrasound as well as upgrading the seismic network processing module, *GA*, into a multi-technology network processing module. The approach taken at the pIDC has been to jointly process the acoustic and seismic technologies as much as possible, and thus build upon the large experience and software available for seismic processing.

The project contributed a series of five software deliveries to the CMR, either independently or jointly with the IDC project funded by DTRA to develop the system that would eventually be contributed by the United States to the Vienna International Data Center. These deliveries are summarized in Table 1.

Table 1: Software releases and Fusion project contribution

pIDC Release name	*Contributed IDC Release*	*Date*	*Contribution of the Fusion project to the Release*
(SH)		Jan. 1997	Initial seismic-hydroacoustic fusion capability. DFX_hydro, default rules StaPro and GA with default velocity model.
PIDC5.0		Mar. 1997	Minor enhancements to hydroacoustic capability. Initial release of hydroqc.
PIDC6.0 (SHI)	R1	Dec. 1997	Initial infrasound implementation with constant velocity model. Upgrade of the hydroacoustic velocity model with 2D variation capability. Initial release of SynGen. Initial release of hydroacoustic recall processing.
PIDC6.1	R2	Jul. 1998	Upgrade of the hydroacoustic velocity model with seasonal variation capability. Initial release of GAgrid.
PIDC7.0	R3	Jan. 2000	Upgrade to infrasound velocity models to include station-dependent 2D seasonally varying tables. Upgrade of the system for hydroacoustic azimuthal capability at multi-site hydroacoustic stations. Initial release of new process HAE to perform automatic hydroacoustic azimuth determination. Initial release of HART module to perform review of hydroacoustic azimuth determination. Initial release of SHIRPA to associate radionuclide detections to events formed from waveform data.

3. Development under the Fusion project

3.1 Introduction

A program plan was developed at the initial phase of the project (*Sereno*, 1996) to define the operational objectives and schedule for this project. This plan was subsequently revised (*Sereno and LeBras*, 1998) and a new schedule defined. The CTBT-related activities were organized into four subtasks:

1. Hydroacoustic and Infrasound integration. This is for support and integration of third-party contractor software for single-sensor signal processing.
2. Automated fusion of SHI sensor data. This is to upgrade the Seismic-Hydroacoustic-Infrasound fusion software.
3. Automated fusion of radionuclide data. This is develop an initial radionuclide fusion capability.
4. Visualization. This is to develop software visualization tools needed for processing of the non-seismic technologies.

The following sections expand on the developments in each of these four areas.

3.2 Hydroacoustic and Infrasound integration

3.2.1 Hydroacoustic Integration

The principal mission of the hydroacoustic network is to detect underwater explosions and help in detecting low-altitude atmospheric explosions. The hydroacoustic network also complements the seismic network by detecting *T* phases generated by underground events (generally earthquakes) and associating these phases to seismic events. The planned IMS hydroacoustic network consists of a total of 11 stations, 6 of which are hydrophone stations and 5 are *T*-phase stations. Most of these stations are not yet in operation. The current network in use at the pIDC is much smaller than the proposed IMS network (Figure 1). It currently consists of 3 hydrophone stations (Wake Island, Ascension Island and Point Sur, California) and 1 *T*-phase station (Victoria Island, British Columbia). A fourth hydrophone station off New Zealand's coast was used on a temporary basis for more than a year to provide additional data for the purpose of testing hydroacoustic processing. The Wake Island station has two hydrophones (WK30 and WK31) separated by 240 km, and the Ascension Island station has three hydrophones (ASC23, ASC24 and ASC26) with station spacings ranging between 3 and 100 km. The Point Sur station has been used at the pIDC in lieu of other data and only has 1 hydrophone (PSUR). It will *not* be part of the IMS network. The New Zealand station consisted of two hydrophones (NZL01 and NZL06) spaced 0.6 km apart. The *T*-phase station on Victoria Island (VIB) consists of a single vertical component seismometer. The stations at Wake, Ascension and Victoria islands are proposed IMS stations.

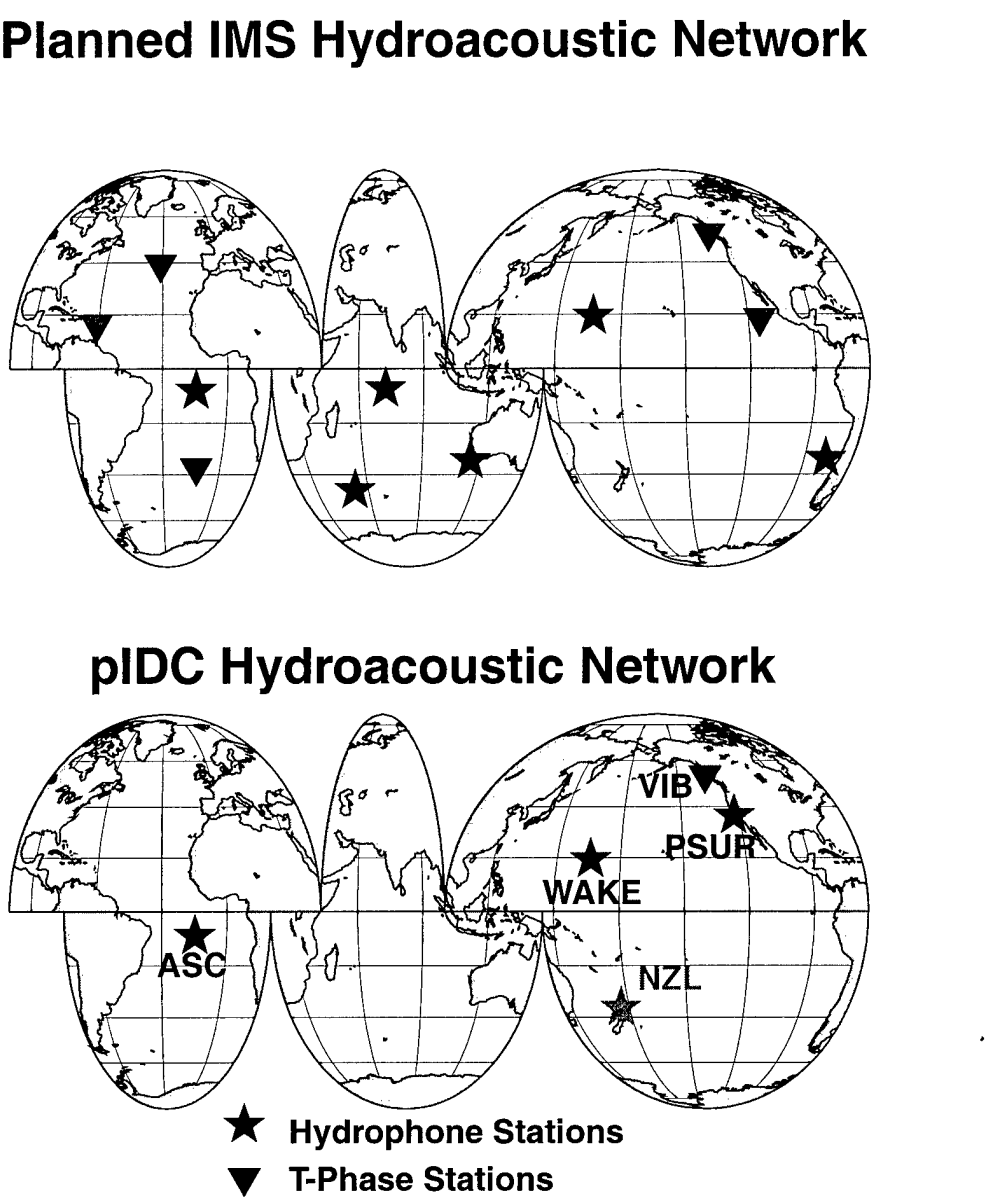

Figure 1. (from *Hanson et al.*, 2000). The planned IMS network has 6 hydrophone stations and 5 T-phase stations. The hydrophone stations, which are more sensitive than T-phase stations, are mainly located in the southern oceans to make up for the seismic network's lack of coverage. A hydrophone station will generally consist of two or more widely spaced instruments to reduce blockage and provide some azimuthal constraints. The current network used at the pIDC has 3 hydrophone stations and 1 T-phase station. A fourth hydrophone station near New Zealand was used on a temporary basis. The Wake and Ascension stations consist of two and three hydrophones respectively.

A large number of detections are recorded daily at the hydroacoustic stations when the current short-term average to long-term average (STA/LTA) signal to noise ratios are used in the detector. In the absence of discrimination between different types of arrivals, and of azimuthal information, since the current stations consist of single hydrophones, this large number of detections would lead the current network processing system to a large proportion of false events (*Le Bras and Sereno*, 1996).

Table 2: Simulations on synthetic hydroacoustic detections with varying numbers of random detections and precision on azimuth measurement.

No azimuth			
	Obtainable Events	Missed Events	False Alarms
Noise Free	10	0	0
12 hour Random	10	1	16
3 hour Random	10	3	219
3 hour Random with azimuth			
	Obtainable Events	Missed Events	False Alarms
Delaz = 3°	10	0	0
Delaz = 6°	10	0	3
Delaz = 12°	10	1	8

This is illustrated by Table 2 which summarizes the results of tests on 15-day synthetic data sets with three different levels of noise. The first is noise-free, the second has noise added at an average rate of 2 arrivals per day (called 12 hour random) and the third has noise added at a rate of 8 arrivals per day (called 3 hour random). There were 10 target events in the synthetic data set with 3 or more defining phases.

As expected, an increase in the number of noise arrivals results in an increase in the number missed events and a dramatic increase in the number of false alarms. The situation is much better when azimuths are available from the hydroacoustic stations as shown in the lower half of Table 2. Even with an azimuth uncertainty of 12° (Delaz), the number of false alarms dropped by almost a factor of 30 compared to the case without azimuth information (i.e., single hydrophone station).

The false alarm problem for hydroacoustic processing is mitigated by two features of pIDC processing.

- The IMS hydroacoustic stations will consist of several hydrophones and this project has developed the capability to use this multiplicity of stations to obtain an azimuth (or bearing) measurement for incoming hydroacoustic wavefronts. Currently, there are only two hydroacoustic stations (Ascension Island and Wake Island) with more than one hydrophone.

- Station processing cuts down the number of detections to be considered by network processing by differentiating between underwater explosions (*H*), underground sources, mostly earthquake-generated (*T*), and noise detections (*N*). The phases classified as *H* are the only ones used in forming events and in defining their location.

This approach is different from the seismic approach which classifies individual detections based on their path of propagation. Single-sensor processing of hydroacoustic signals includes the source characteristics into the classification, thus performing some level of source discrimination at single-station processing. A large number of features are measured on the hydroacoustic waveforms, for up to eight frequency bands. Classification into the three types of phases is performed either using a default set of rules on the frequency-dependent features or neural network weights adapted for individual stations. The default rules attempt to first identify positively the *T* phases mostly through their main characteristics of a long duration and low relative frequency content. Noise phases (*N*) are the next set of phases to be identified. Their frequency content characteristics overlap the *T* phase's frequency content, but they usually have a shorter duration. Finally, phases that do not fit into the *T* or *N* class are identified as in-water explosions (*H*). The main feature differentiators are a large ratio of high frequency to low frequency content and a short duration.

We received hydroacoustic products for integration into the pIDC processing suite from the SAIC hydroacoustic expert group in McLean:

- A DFX hydroacoustic signal processing and features extraction library has been added to the pIDC processing suite (*Laney*, 1997)

- A neural network - based station processing module processes the detection features and classifies detections into the three hydroacoustic phase types (*Dysart et al.*, 1997).

- A set of seasonally-and station-dependent 2D travel times was delivered for all existing stations. An example of these travel times is given in Figure 2 for station WK31.

One of the main differences between seismic and hydroacoustic travel times is that hydroacoustic times should be picked at the peak amplitude of the wavetrain as opposed to seismic travel times which are picked at the onset. The reason for this is that the hydroacoustic travel times are modeled for the SOFAR channel depth. Since the SOFAR channel tends to concentrate acoustic energy, the propagation time of the maximum energy arrivals will reflect travel times at the SOFAR channel depths.

In order to integrate the hydroacoustic travel time handling into the pIDC processing suite, we followed the same travel time representation model used for seismic travel times. This model assumes separation between measurement errors and modeling errors. To be able to estimate both a modeling error and a measurement error on the travel time for the maximum energy, we formulated the concept of probability-weighted time calculation (e.g., *Hanson et al.*, 2000). The probability-weighted time is an estimate of the peak amplitude time including the uncertainty on its value. The calculation takes into account the current background noise level for the measurement time and an average noise level for the modeled time. The path and season-dependent tables are generated using a ray-mode computation software, the Active Sonar Performance Model (ASPM), with DBDB5, a standard 5-minute bathymetric database provided by the US Navy (*Burger et al.*, 1994).

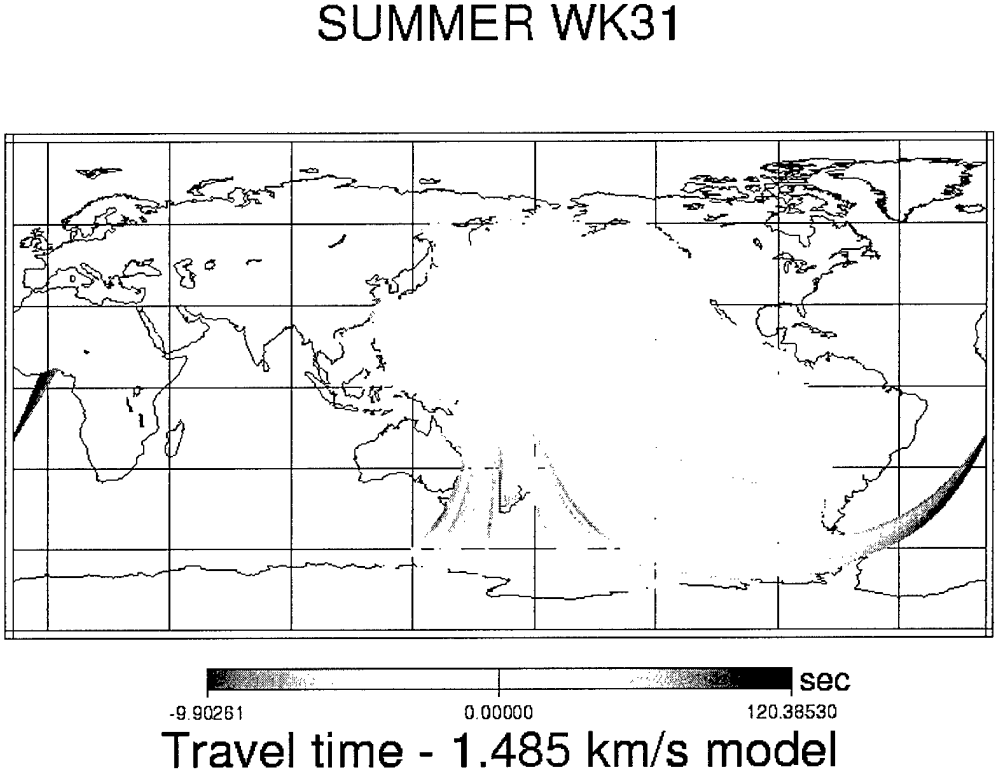

Figure 2. Map of travel time differences between the summer travel times for station WK31 and a constant velocity model (1485 m/s).

3.2.2 Infrasound Integration

The planned IMS infrasonic network is intended to detect unannounced atmospheric explosions anywhere on the globe. The design for IMS infrasound stations consists of an array of four microbarographs in a triangular pattern with a center element. Array processing for infrasound sensors is most similar to seismic array processing with the exception that the detections are based on a more sophisticated hybrid algorithm that takes into account a measure of waveform coherency in addition to the usual STA/LTA signal to noise measure. For a detection to be declared at an infrasound station, both measures must coincidently exceed a threshold specific to each measure (*Katz et al.*, 1998).

The atmosphere is traditionally divided into different layers with altitude: the troposphere, stratosphere, mesosphere and thermosphere, in order of increasing altitude (*McKisic*, 1997). Propagation of sound in the atmosphere is complex even in the absence of winds, due to the presence of two low velocity zones. For an atmosphere at rest, this sound velocity structure gives rise to three main groups of arrivals: the tropospheric, stratospheric and thermospheric arrivals. We are currently not attempting to differentiate between these different arrival types at the station processing level, although we are collecting a large set of feature measurements for each detection, which should help us in the future to identify these arrivals. In the current implementation of the initial phase identification module, a very simple phase velocity-based rule is used to differentiate detections between noise detections which are called *N* and signals of interest, which are all indiscriminately assigned the generic phase type *I*. Detections are

considered to a signal of interest (*I*) when their phase velocity is between two bounding values, currently set at 290 m/s and 660 m/s.

The version of network processing (PIDC_6.2) currently running at the pIDC has integrated the infrasound phase type by modeling the infrasound propagation as a single phase with a large time modeling error to account for the possibility that a given observed phase may be one of the three arrival types. The azimuth measurement is very precise and contributes the most to event location. The location algorithm will always locate an event at the surface if it includes an infrasound arrival (either purely infrasound or mixed event). There is currently an upper limit of 60 degrees on the range at which infrasound arrivals will be associated.

The travel time handling system has the capability to use station-specific and path-dependent travel times calculated using average seasonal wind patterns and we are in the process of integrating such tables for operational stations into the PIDC_7.0 release. These travel time tables were provided to us by D. Brown from the Center for Monitoring Research. An example of travel time map is shown in Figure 3.

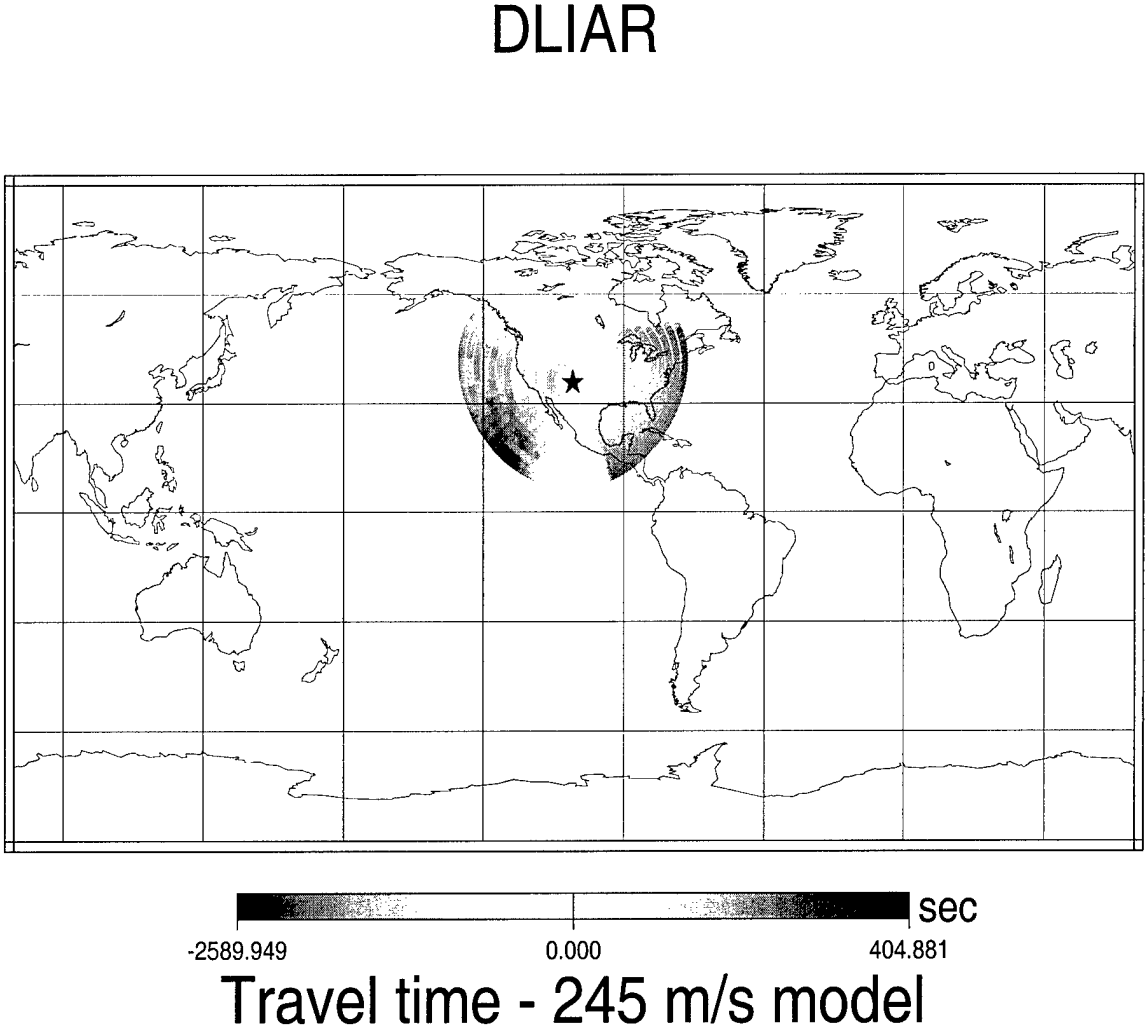

Figure 3. Travel time differences in seconds between a constant velocity model (245 m/s) and the path and season-dependent travel times (month of July) for station DLIAR up to a distance of 30 degrees. Note the east-west asymmetry of the travel times due to the effect of the dominant winds.

3.2.3 Automated fusion of SHI sensor data

Network processing (the GA subsystem) has been adapted to process hydroacoustic and infrasound phase types. The modifications made to the system in order to accommodate these additional data types were the following:

- The processing time windows have been adapted to the specific technologies in order to allow for sufficient time between arrivals at different stations to be able to form associations. There is one processing window per technology, and all windows are designed in a step-forward fashion, adapted to following origin times of events. The start time of the window is the same for all technologies.

- Conflict resolution has also been enhanced and specialized for acoustic technologies. For acoustic arrivals, a set of two parameters allows the resolution to favor either the events with the largest number of arrivals from the same technology or the events with the largest number of arrivals from another technology. The current setting at the pIDC is to resolve in favor of the events with the largest number of arrivals of the same technology.

- Another addition to network processing had been to set the appropriate weights to hydroacoustic, infrasonic or mixed technology events. Initial weights were developed based upon pure hydroacoustic and pure infrasonic events. The confirmation threshold per event is unique and independent of the technology type, including mixed events. The minimum pure hydroacoustic event is made up of 3 hydroacoustic arrivals with time only (no slowness information). The minimum pure infrasound event is made up of two arrivals with azimuth.

One of the main characteristics of hydroacoustic phases is that they propagate efficiently in the water column and that they can be blocked by significant continental masses after a few kilometers of underground propagation. To take this into account, blockage maps have been developed for individual hydroacoustic stations. Network processing determines whether an association is expected or not given the blockage constraints on the path. The maximum distance of propagation in a given azimuthal direction obtained from the blockage maps is compared against the distance between source and station in every hydroacoustic association. Associations that are found to be along blocked paths are disallowed.

Blockage is taken into account at several stages of the automatic association process (or network processing). This includes an initial grid-based exhaustive search where the phase list attached to each grid point takes it into account for every station. Blockage is also checked after location to verify the legitimacy of each hydroacoustic association. Finally, blockage maps are first checked at the prediction stage before adding a hydroacoustic phase to an association set. An interactive tool has also been developed and presents either an alphanumeric window indicating whether a path is blocked or clear or a map showing the great circle path color-coded according to its blockage status.

At the time of the initial implementation of hydroacoustic processing, travel time for hydroacoustic phases was handled using a simple constant velocity (1485 m/s) model. This has been upgraded starting with release PIDC_6.0 where path-dependency was introduced. The travel time handler was further upgraded with release PIDC_6.1 to allow for seasonal dependency in addition to path-dependency. In addition, this last improvement included the capability to specify path-specific modeling errors for acoustic phases. The season and path-specific travel times can be

generated for any hydroacoustic station with known location. Each station is user-configurable where the travel time uses by default the constant velocity model unless it is specified that the season and path-specific set of tables exist for the station and should be used. The path and season-dependent tables are generated using a ray-mode computation software, the Active Sonar Performance Model (ASPM), with DBDB5, a standard 5-minute bathymetry US Navy-provided oceanic velocity database. The theoretical travel time differences between a constant velocity model and these path-dependent travel times reaches a few tens of seconds and it is expected that their implementation will improve the fit with observed travel times. This has been observed and confirmed for a few specific paths (see section on case studies). The seasonal dependence on a given path is not as significant as the difference between the constant velocity travel time and the path-dependent travel time and so far we have not found any systematic variation with the seasons on pIDC hydroacoustic data.

The season and path-specific travel times are used for both *T* and *H* phases, however two differences are worth noting in the current configuration between these phase types. The first is that the travel time for *T* phases includes an underground path, usually short in length compared to the total distance between source and receiver, but undetermined. This is taken into account by adding a coupling term to the modeling errors for *T* phases. The second and most significant difference is that the *T*-phase travel time is not taken into account in the location as opposed to the *H*-phase travel time. The system is capable of locating a purely hydroacoustic event provided *H* arrivals have been detected and identified at a minimum of three stations. In the current configuration, *T* phases alone cannot form an event and they are always associated to pre-existing seismic events as non-defining phases and acquired by travel time prediction based on the origin time and location of the seismic event.

The vast majority of hydroacoustic phases associated to pIDC events are earthquake-generated *T* phases.

A statistical analysis of the hydroacoustic results at the pIDC shows that for a two-year period starting in May 1997, the Reviewed Event Bulletin (REB) contains 13,532 *T* phases that are associated to 8,437 events. The number of events possessing a hydroacoustic *T* phase is about 25% of the total number of REB events when taking into account hydroacoustic stations down time.

At the time of this report, no explosive marine events have been automatically detected and formed by the system, although there have been examples of automatic detections emanating from two series of small-yield marine explosions. Because of the low number of stations, the automatic detections were made at one or two stations and correctly identified as explosion-generated. The arrivals at the stations missing the detections had too low a signal to noise ratio to be detected. The fact that automatic processing correctly identified some of the arrivals as originating from in-water explosions incited the analyst into looking for additional detections and thus later completing the association set to a total of three stations. In both cases, the occurrence of the small yield explosions was independently confirmed.

A drawback of the initial hydroacoustic network is that, although two existing stations (Wake Island and Ascension) as well as planned IMS stations consist of multiple hydrophones, no

advantage is currently taken from that multiple hydrophone configuration to obtain directional information. We have developed this capability as a part of release PIDC_7.0.

Even though the software at the pIDC is ready to process infrasound sensor data all the way to automatic association in network processing, the infrasound waveform data has not been timely and the SEL3 bulletin does not reflect the expected performance of the current infrasound network. The automatic bulletin at the pIDC produced a low number of automatic infrasound associations (492) and automatic pure infrasound events (57) since implementation of release PIDC_6.1, between December 2, 1998 and August 2, 1999. The main characteristic of the distribution of pure infrasound events is its extreme irregularity with time. We observed a large number of automatic events in early December 1998, including one that was later confirmed and included in the REB. That distribution is very biased by the timeliness of the waveform data and to correct for this, we have reprocessed a year's worth of data through GA. The results of that reprocessing are reported in Section 6.2

3.3 Automated fusion of Radionuclide data

Twenty-two radionuclide stations currently exist and are operational, with an additional 58 stations planned to reach the total of 80 stations proposed in the CTBT. At each station, particles are collected continuously from the passing airmass for 24 hours. Each sample is then allowed to age for another 24 hours so that "background" radionuclides that may interfere with measurement can decay away. The radionuclide content of the sample is then determined through gamma ray spectroscopy. If anthropogenic radionuclides are detected, the sample is assigned a level 4 or 5 (respectively single or multiple anthropogenic radionuclides) alert status.

Meteorological measurements from the NOAA database and meteorological modeling software are used to determine the trajectories covered by the airmass. This locus of points in time and space from which the detected radionuclides might have been released is called the Field of Regard (FOR) of the sample.

The NOAA data set consists of wind velocity measurements gathered at 24-hour intervals at points on a regular grid covering the globe. OMEGA is a weather modeling prediction program used to calculate wind velocities at the grid points to a temporal resolution of 1 hour by modeling the airflow between the two endpoints that bound the 24-hour interval.

This information is used to simulate particle releases from each grid point at 1-hour intervals for a period of 3-14 days before a given sampling period. The locus of points from which a particle released a certain period of time before a sampling period has a significant probability of being detected in the sample is called the Field of Regard for that time period. It is used to calculate high resolution FORs and several other related data products. A radionuclide detection with a level 4 or 5 alert status triggers OMEGA to calculate 72, 48 and 24 hour FORs for the detecting station and sampling period. Automating radionuclide association to SHI events, then, consists simply of searching for events from the Standard Screened Events Bulletin (SSEB) whose error

ellipse has an intersection with a field of regard for a level 4-5 alert sample and whose origin time is compatible with the FOR's time period.

Particle trajectories could also be used to improve on the 24-hour granularity of the FORs, however the size of the FORs is usually of the order of at most a few error ellipses and the added complication of performing trajectory interpolation may not be worth the small gain in resolution.

Unlike the other three technologies in the Fused Event Bulletin, radionuclide detections represent direct physical evidence of a nuclear reaction at least, possibly of an explosion. Hence, greater care must be taken when publishing data relevant to this technology to avoid the appearance of making an accusation. In the context of the automated IDC system, an association drawn between a radionuclide detection and an SSEB event should not and cannot be construed as a claim that the SSEB event was indeed the source of the radionuclides. The association must be understood to indicate only that a detection of anthropogenic radionuclides was made at a station and that if such nuclides had been released from the location/time of the SSEB event, the probability of their being detected at that station in that sample is significant. Neither of these offers any basis on which to calculate the probability that such radionuclides *actually were* released from the location/time of the SSEB event.

We have developed a processing module, SHIRPA, to perform automatic association of FORs from radionuclide detections to SSEB events and publish the results in a Fusion Event Bulletin (FEB) (*Guern and Le Bras*, 1999).

3.4 Visualization

Four visualization tools were developed under this project, three of them as an aid to hydroacoustic processing and the third a utility tool to visualize the grid file used by the automatic associator, GA. The three hydroacoustic tools are the blockage visualization tool which allows the analyst to decide whether a particular path is blocked, the hydroacoustic recall tool to reset the onset and termination times for a hydroacoustic detection, and the hydroacoustic azimuth review tool (HART) which allows review of the automatically computed azimuth at multi-hydrophone hydroacoustic stations. A short description of each of these tools follows.

3.4.1 Blockage visualization tool, hydroqc

Unlike seismic propagation paths, hydroacoutic paths are mostly oceanic. The converted phases at the ocean-continent interface are rapidly attenuated. The propagation of hydroacoustic waves is stopped by a continent or a major island. We developed a new ARS tool to provide an analyst with the capability to determine which hydroacoustic paths are likely to be blocked based on the same blockage grid file that is used by GA. This tool, called hydroqc, includes a pop-up warning and an optional map display. It is invoked by a new button on the ARS toolbar.

3.4.2 Hydroacoustic recall tool

Detection times on hydroacoustic waveforms are not estimated on the onset times but rather on the peak amplitude of the waveform. We implemented a capability to review and edit the hydroacoustic onset and termination times automatically determined by DFX from ARS. The most probable cause of unreliable hydroacoustic feature measurements is an inaccurate estimate of the

signal onset and termination. If these can be adjusted by an analyst, then DFX can be initiated to recompute the other feature measurements either interactively (i.e., the analyst sends a message to DFX through the PMshell to recompute the features) or after analyst review as part of recall processing. The hydroacoustic recall tool allows interactive review of the onset and termination times to provide a consistent measure for the arrival time.

We enhanced ARS to review and edit the hydroacoustic onset and termination times. An example of the user-interface is shown in Figure 4. This display was generated after a hydroacoustic arrival was selected and "onset and termination time review" was selected from the ARS menu bar. The onset and termination times and filter parameters were extracted from the **hydro_features** table for the band whose peak time is equal to (± 0.1 s) the arrival time. The data were filtered by ARS in the appropriate band and the onset and termination times were displayed with arrows. The arrival time was displayed as a detection bar in the usual way. The analyst can now filter the data again and adjust the onset and termination times as required. When the event is saved, the new values will be saved in the ARS output account.

3.4.3 Hydroacoustic Azimuth Review Tool (HART)

We have developed a graphical tool invoked from ARS to perform the review of automatically determined azimuths at hydroacoustic stations with multiple hydrophones. The tool is analogous to XfkDisplay in its purpose, in that it is used to review and set the values of automatically computed azimuth picks. The difference with XfkDisplay is that it is invoked for a group of detections rather than a set of waveforms from an array of seismic or infrasound sensors an it has the added feature of displaying the waveform alignment appropriate for a given azimuth and slowness. This Hydroacoustic Azimuth Review Tool (HART) has been developed in the Java programming language. The tool is invoked by selecting an arrival from a group in ARS and selecting the appropriate button. In addition to *fstat* values as a function of azimuth and slowness, it displays segments of either the waveforms or their envelopes around the hydroacoustic detections. Selecting a value on the annular display will cause the waveform segments to align themselves according to the time delays derived from the chosen azimuth and slowness. Also shown are the values of azimuth, velocity and *fstat* at the current position of the cursor, the values of the automatically computed azimuth, the value of the manually selected azimuth, velocity, slowness and *fstat* as well as the station to event azimuth if available.

Figure 5 presents an example where HART is used to correct an erroneous azimuth value determined by the automatic system. In that case, the azimuth was determined to be 44 degrees by the automatic system because only two hydrophones were available out of the three. The interactive display allowed the azimuth to be corrected to the value of 180 degrees, very close to the station to event azimuth of 183 degrees.

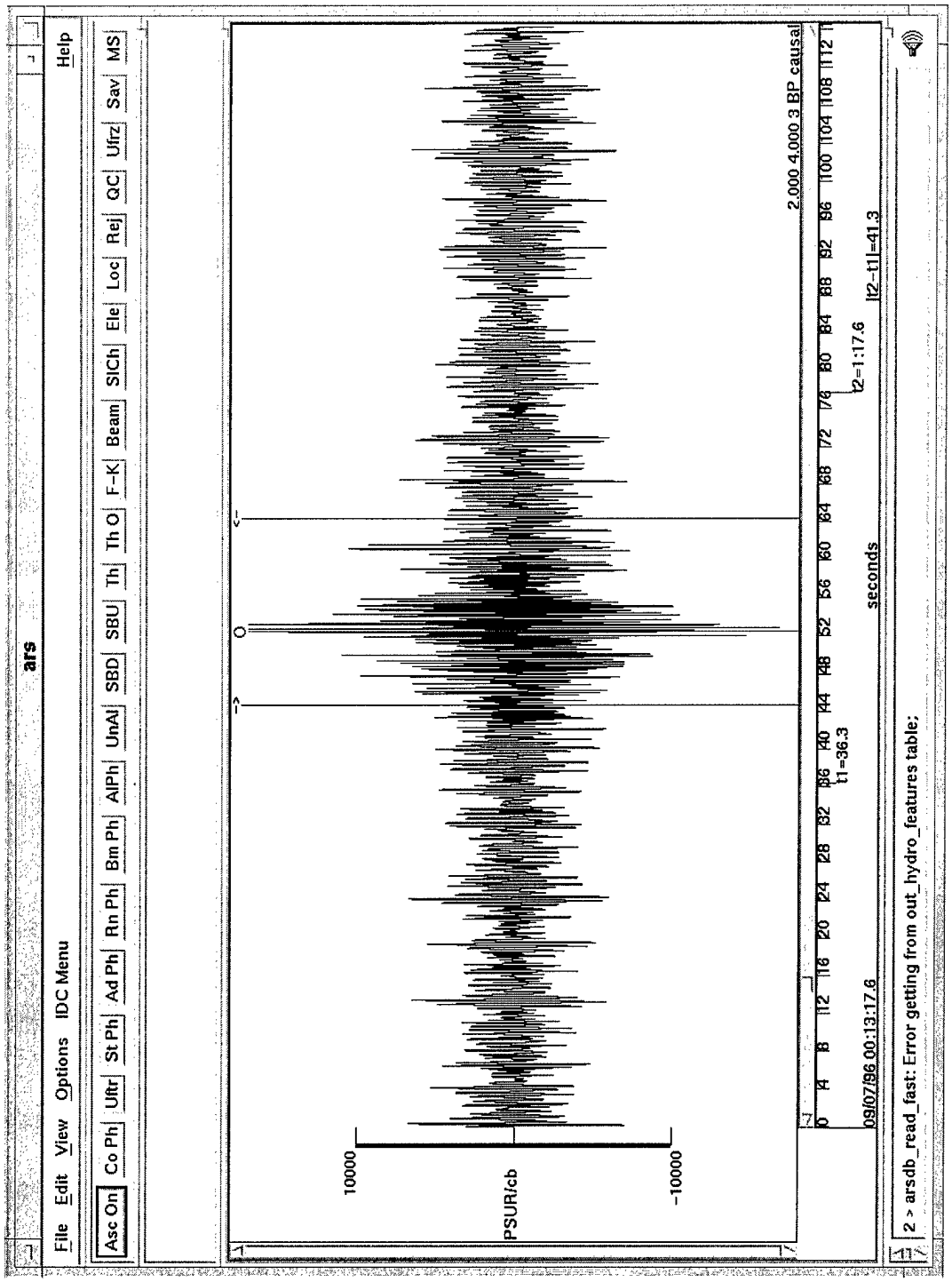

Figure 4. This is an example of the ARS user-interface for review of hydroacoustic onset and termination times.

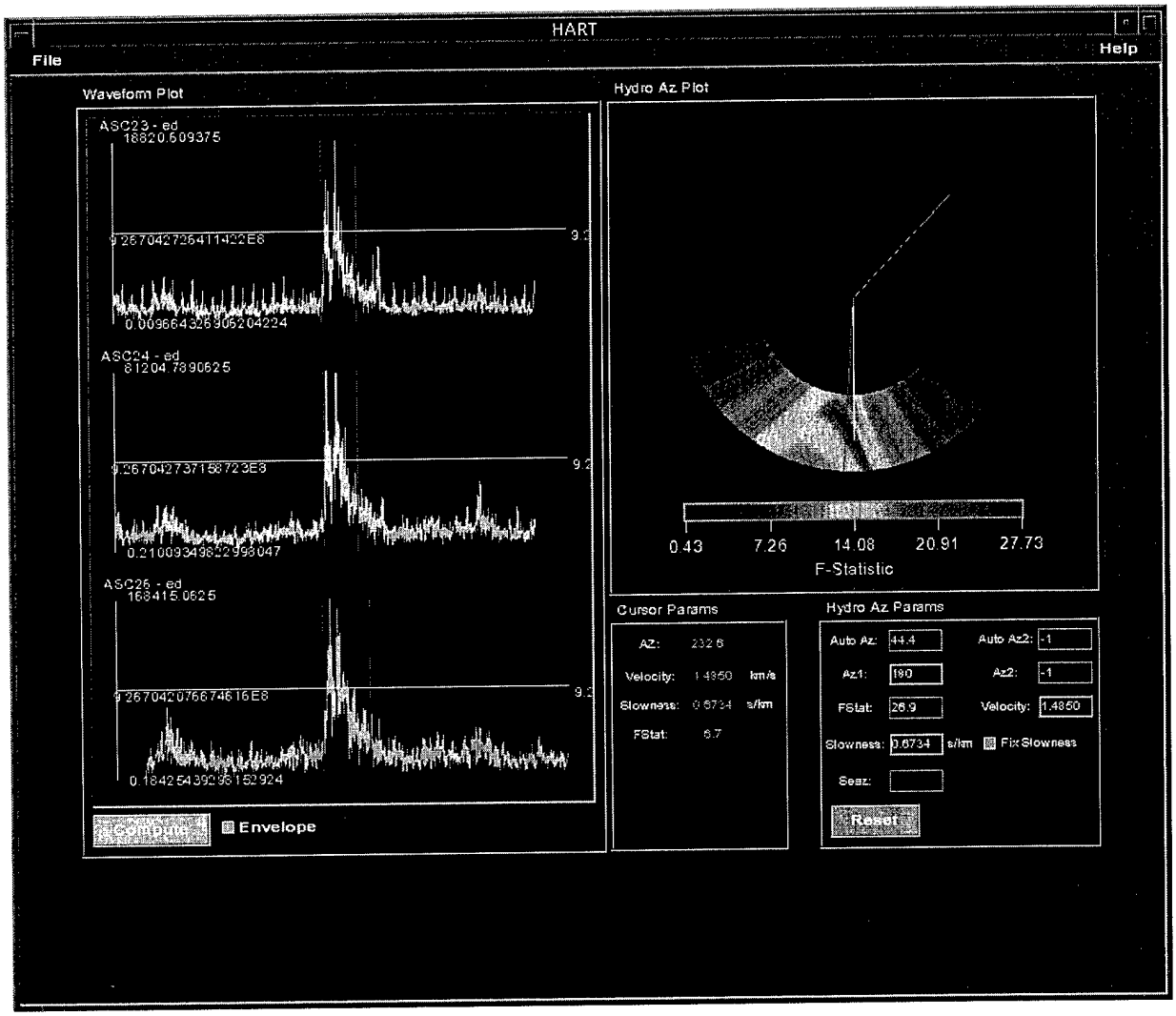

Figure 5. Example of a hydroacoustic phase seen at the three Ascension hydrophones. Use of the review tool is necessary in this case because the automatic pick was erroneous (azimuth of 44 degrees) and was corrected to 180 degrees, closer to the station to event azimuth value of 183 degrees.

3.4.4 Visualization tool for the automatic association knowledge base (GAgrid)

We have productized and enhanced the GAgrid utility and graphical user interface tool. The tool allows visualization of the grid geometry and contents stored in the GA grid file. The grid file content, including networks, stations, grid point locations and cell size, seismic and acoustic phases, can be visualized and displayed as alphanumeric tables.

The enhancements consisted in a display of all stations contained in a grid file and in color-coded display of the travel time for a selected station and phase. This last enhancement is illustrated in Figure 6.

GAgrid is intended as a medium term solution to the issue of grid content visualization and display. A long term solution to the issue of quality control and content visualization of the grid used in automatic association probably lies with the adoption of a commercial off-the-shelf (COTS) system for the visualization of spatially-oriented objects. Within the framework of this project, we have targeted this as one of the issues to be addressed in the next-generation design documents.

Figure 6. Examples of geographical display of travel time for a) the hydroacoustic H phase at one of the Ascension Island stations and b) the seismic Pn phase at array station ESDC. Note that the extent of definition of the H phase is limited by blockage. The Pn phase is limited within its range of definition.

4. Case studies

4.1 Introduction

Case studies have proved useful in evaluating the processing software and propagation models for the hydroacoustic and infrasound technologies where little operational experience with automatic systems is available compared with the seismic technology.

In the course of testing the different processing modules, we have gathered data from events of interest for all waveform technologies and performed a number of case studies. These case studies include sets of marine explosions from seismic refraction experiments in the Western Pacific (*Brumbaugh and Le Bras*, 1998), two atmospheric explosions with ground truth origin time and location, and mining events recorded at co-located seismic and infrasound arrays in Western Texas (TXIAR and TXAR). The marine explosions have provided useful data to test the hydroacoustic station processing module and in particular have allowed us to evaluate the phase identification module and test a cepstral-based bubble pulse delay extractor. Examination of the data from the atmospheric events in the White Sands Missile Range, for which ground-truth origin data is available, allowed evaluation of the detector and location modules. The mining events generated both seismic and infrasound arrivals and provided an opportunity to evaluate the usefulness of jointly processing these two data types.

4.2 Hydroacoustic case studies

4.2.1 Marine explosions offshore Japan

4.2.1.1 September 1996 refraction experiment

In the course of testing hydroacoustic signal processing algorithms to characterize explosion sources from earthquake-generated *T* phases, a series of over 100 distinct signals were identified at the two Pacific-based hydroacoustic stations: Point Sur, California (PSUR), and Wake Island (WAKE). The distance between the two stations is about 41°. The signals were grouped in three distinct 8- to 10-hour long subsets from September 6 to 10, 1996, and clearly emanated from in-water explosions because of the presence of bubble pulses identifiable in their spectrum, cepstrum, or autocorrelation. Furthermore, individual events could be correlated between the two stations and the time separation between them was quasi-constant (12-15 minutes). These characteristics pointed to human activity, and more specifically, to a seismic refraction survey as the source of these events. Since PSUR and WAKE were the only stations that recorded the series of explosions, no definite location was attempted. The locus of possible sources was a line crossing the Pacific ocean from the coast of Antarctica, passing between Australia and New-Zealand up to northern Japan.

After some investigation and intensive exchange of electronic mail with various institutions around the world, we established that the signals observed at these two stations emanated from a seismic refraction survey conducted by a Japanese scientific crew offshore northern Honshu, Japan (*Ryota Hino, Tohoku University, Sendai, Japan, personal communication*). The 141

underwater explosions ranged from 20 to 400 *kg* in size, and were detonated approximately 28° from WAKE and 71° from PSUR (Figure 7). The comparatively small size of these explosions, and the fact that they were recorded with good signal to noise ratios at cross-oceanic distances illustrate the extremely good propagation characteristics of the oceans. The signals generated by the refraction survey allowed us to evaluate the performance of automated processing programs and partially validate path-dependent hydroacoustic travel time tables. Marine refraction surveys such as this are now rare in the oceans and this ground truth data set is of great interest to validate propagation models and processing techniques.

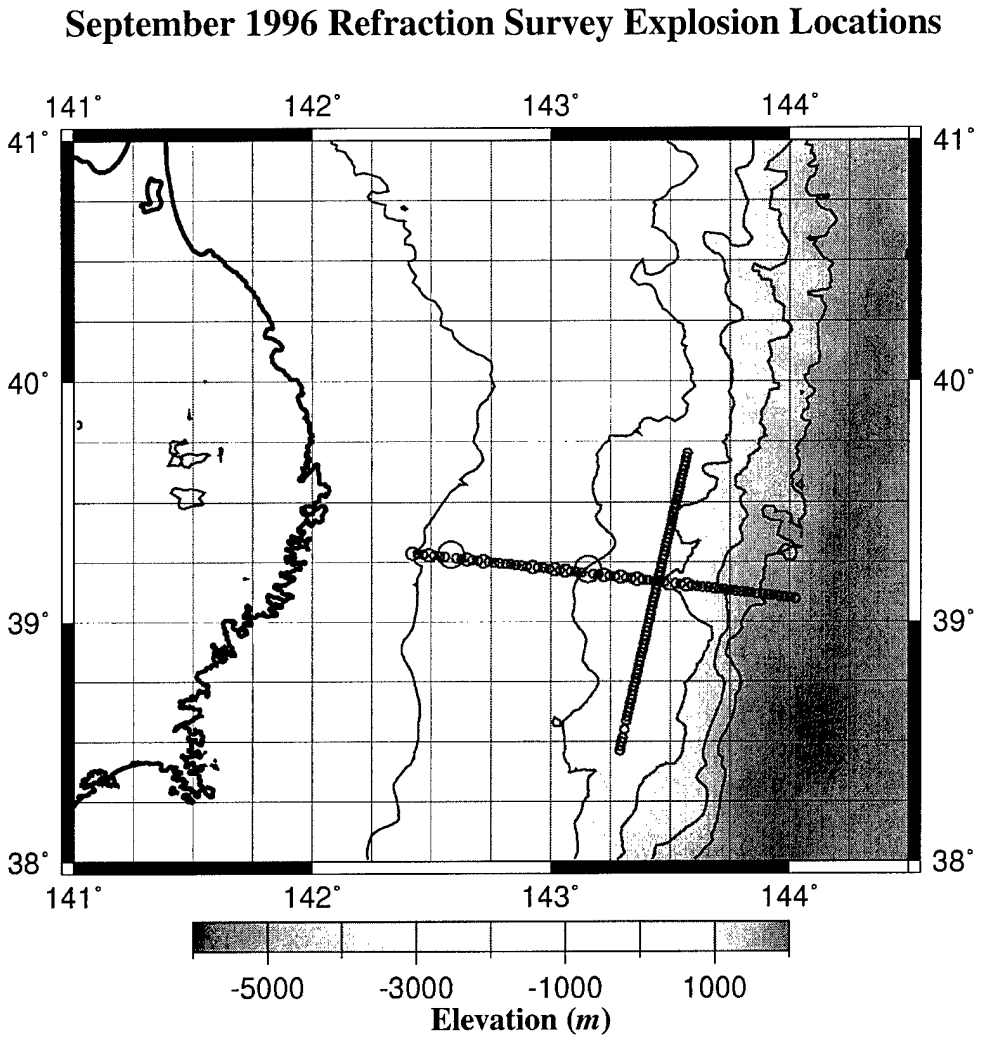

Figure 7. (from *Brumbaugh and Le Bras,* 1998). Location of the 141 individual explosions superimposed on bathymetric data and the coastline of northern Honshu, Japan. The smaller circles indicate explosions with a yield of 20 *kg*, the intermediate circles represent yields of 40 *kg* and the two large circles indicate 400 *kg* explosions.

Detection and phase identification statistics

The ground truth data set from the refraction survey allowed us to partially test and verify the automatic detection and phase identification system. Automatic association was not tested because only two stations detected the signals, short of the three needed to form the explosive events.

The system was able to detect and extract features for 334 arrivals at WAKE and 234 arrivals at PSUR during the time period of the refraction experiment. Of these, 138 arrivals at WAKE and 123 arrivals at PSUR were confirmed to have been generated from the 141 Japanese refraction experiment explosions. This corresponds to a detection rate of 98% at WAKE and 87% at PSUR for the explosive sources. The varying noise level and presence of additional T phase signals attributable to earthquake sources explains the few missed detections.

The phase identification stage of processing comes immediately after the detection and feature extraction stage and identifies the detections as noise (N), underground generated (T), or in-water generated (H). Two different methods may be used to perform initial phase identification. The first method relies on a set of default rules applied on measured features, mostly frequency content and duration. The second method uses neural net weights derived from a station-specific training set. The refraction experiment sources provided a means of measuring the improvement brought about by the application of the neural net weights. Of the 138 explosions detected at WAKE, 83.7% were correctly identified using the neural net weights developed for that station as opposed to 69.6% when using the default rules. At PSUR, the neural net recognized 72.5% of the detected explosions compared with 66.7% with the default rules. These results testify of the added value to the initial phase identification brought by the use of the neural nets in that process.

Bubble Pulse Delay

A feature of underwater explosions is the characteristic presence of one or several bubble pulses in the signal. The presence of bubble pulses may be detected through several means, including the autocorrelation and the cepstrum of the signal. Source depths may be computed from the measured cepstral delay and known yield through a standard relationship (*Urick,* 1983):

$$T_i = \frac{K_i w^{1/3}}{(z + 10.1)^{5/6}} \tag{1}$$

where T_i is the period of the i^{th} bubble pulse in seconds, w the yield of the explosion source in kilograms, z the source depth of the explosion in meters, and K_i proportionality constants (K_1 = 2.11, K_2 = 1.48, K_3 = 1.20). We performed cepstral delay statistics on arrivals in the 2-80 Hz band with very distinct cepstral peaks, defined as having amplitudes nine standard deviations above the mean. Unfortunately, this criteria eliminated all arrivals from the 400 kg shots, even though the 400 kg cepstrum did have recognizable peaks. Figures 8 illustrates the signature of a typical bubble pulse from a 20 kg explosion recorded at PSUR. The first bubble pulse peak, at about 0.15 s, clearly stands out. Also visible are the second bubble pulse peak and the delay between first and second bubble pulse.

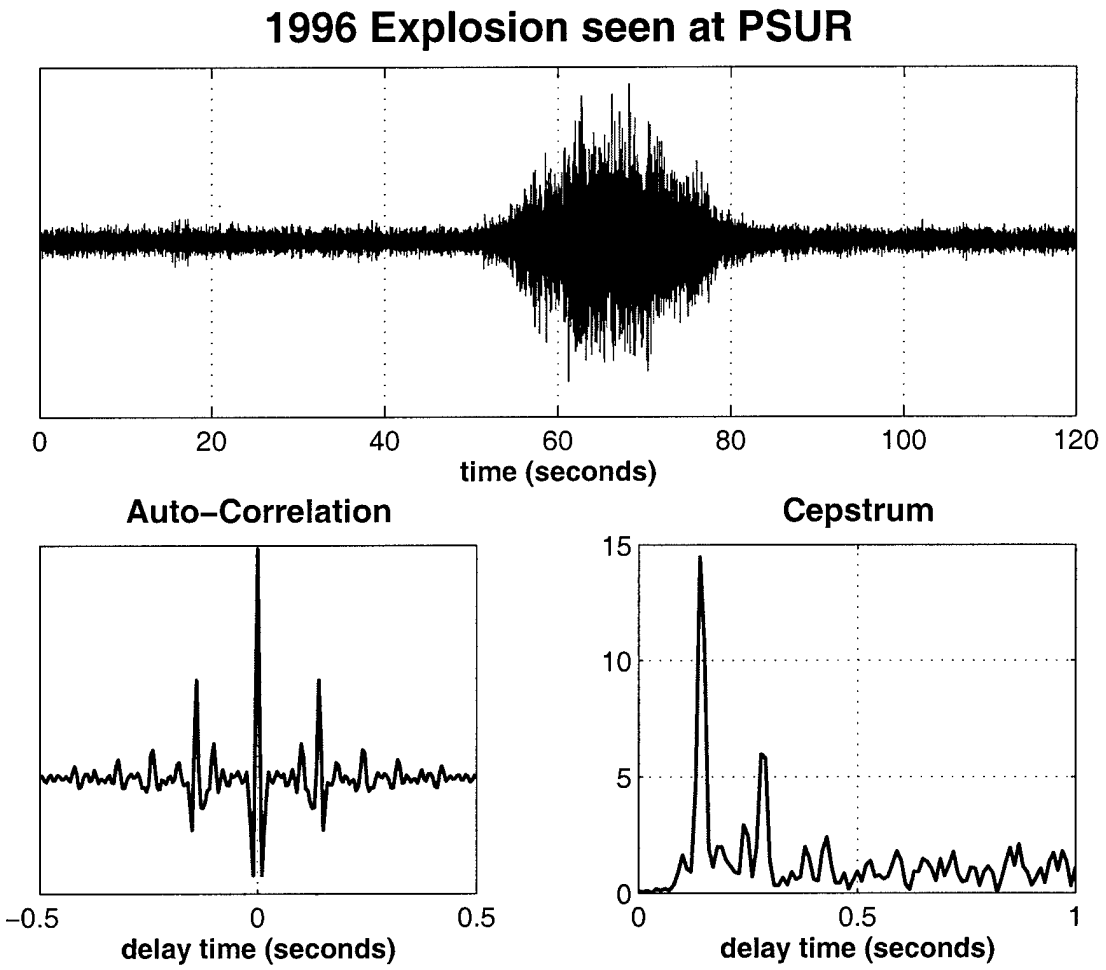

Figure 8. Waveform, autocorrelation and cepstrum of a 20 *kg* explosion recorded at PSUR. Note the presence of the large peak at about 0.15 *s* that we interpret as the first bubble pulse. The delay time of this bubble pulse is close to the center of the distribution of delay times for 20 *kg* shots. The peak at 0.26 *s* is interpreted as the delay between the initial explosion and the second bubble pulse. Note that the autocorrelation also suggests a smaller peak at 0.11 *s*, due to the delay between first and second bubble pulses. The sum of the two bubble pulse delay times (0.11 *s* + 0.15 *s*) is 0.26 *s*.

Table 3 summarizes the statistics of the first bubble pulse delay times at the two stations and the computed source depths for each yield group, assuming the energy released from the charges did not breach the surface. The depth estimation is obtained from equation (1) using the first bubble pulse delay. The number of explosion phases detected (n) and average bubble pulse delay times (period) are tabulated for each station. Note the nearly identical bubble pulse delay times measured at WAKE and PSUR for explosions of either yield. The bubble pulse delay times measured at the two stations agree to within 0.002 *s* for the same explosion source. However, the delay times measured may vary as much as 0.02 *s* from explosion to explosion of the same yield. A 0.02 *s* error corresponds to a depth error of about 12 *m* for a 20 *kg* explosion, while a 0.002 *s* error equates to a depth error of about 1 *m*. The impressive consistency of the bubble pulse delay measurements between the two stations for the same explosion lead us to believe that very good relative determination of the source depth could be achieved using this attribute. The information shown in Table 1 suggests that cepstral analysis techniques, when adequately calibrated, are useful in precisely determining the source depth of marine blasts.

Table 3: Comparison of Bubble Pulse Delay Time by Station and Yield

	20 *kg* Explosions		40 *kg* Explosions	
Station	**n**	**Period (*s*)**	**n**	**Period (*s*)**
WAKE	91	0.151 ± 0.009	8	0.182 ± 0.006
PSUR	43	0.151 ± 0.012	6	0.186 ± 0.006
Average	134	0.151 ± 0.010	14	0.183 ± 0.006

The average source depth computed from all bubble pulses passing the previously outlined criteria is 69 ± 6 *m*. These source depths are 38% deeper than the known source depth of 50 *m*, yet they are precisely grouped. Possible reasons for this discrepancy may be that the effective yields of the explosions were significantly less than the charge mass, or the K_1 constant may need to be corrected for the specific type of explosive used in this experiment.

Hydroacoustic Travel Times

We compared observed travel times to travel times predicted using two different velocity models. The first model is simply a 1D constant velocity model that assumes a hydroacoustic phase propagates at 1485 m/s. The second velocity model is a tabulated path-dependent model. The station-specific tables contain hydroacoustic travel time information for discrete source-receiver distances along discrete azimuths. All travel time measurements were calculated using modal analysis at 10 *Hz* at 0.5° intervals over both distance and azimuth. The travel time tables were computed using actual bathymetry and water velocities from a seasonally-varying database (dBdB5) established by the US Navy.

Histograms showing the travel time residuals (observed minus predicted travel times) at WAKE and PSUR for both models are shown in Figure 9. The travel time residuals at WAKE are shown in the top two panels, while the residuals at PSUR are shown in the bottom two panels. In this figure, the horizontal scaling is the same for all four histograms while the vertical scale differs from station to station.

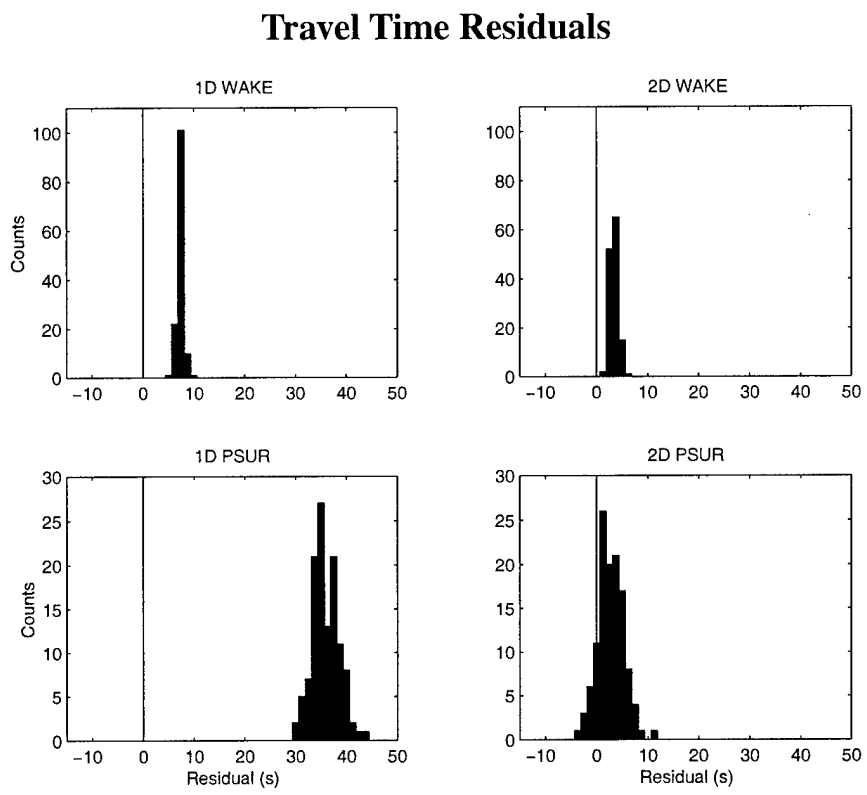

Figure 9. Histograms of travel time residuals (observed minus predicted travel times) computed at WAKE and PSUR using both the 1D constant velocity model and the 2D travel time tables. Note the vertical scale on the histograms differs between WAKE and PSUR. A significant improvement in travel time accuracy is obtained when the 2D travel time tables are employed, particularly for PSUR detections.

Most predicted travel times were shorter than observed travel times for both the 1D model and the 2D model. Using the 1D velocity model, the synthetic detections were 7 ± 1 *s* earlier than the observed arrivals at WAKE and 36 ± 3 *s* earlier at PSUR. The fit was improved when using the 2D travel time tables, particularly at PSUR. With the 2D model, the synthetic detections were 3 ± 1 *s* early at WAKE and 3 ± 3 *s* early at PSUR. The travel time accuracy noticeably improved with the use of the 2D tables, particularly for arrivals at PSUR. The gain in accuracy brought by these 2D tables for this particular experiment justifies using them to obtain more accurate locations in a monitoring environment.

Knowledge of the ground truth information for the set of marine explosions described in this article allowed us to establish the reliability of our detection and phase identification processes as well as of the cepstral delay time measurements. We established the extreme consistency of these delay times between the two stations that recorded the explosions. While variations from explosion to explosion were found to vary by about 0.02 seconds, likely due to differences in yield and depth, the difference between measurements at the two stations for the same explosion are within 0.002 second of each other. This makes the cepstral delay a suitable attribute to use as an association criterion.

We were able to evaluate and validate path-dependent hydroacoustic travel times derived from US Navy measurements. Synthetic travel times computed with the 2D station-specific travel time model were within five seconds of the observed travel times, whereas travel times generated assuming a constant hydroacoustic velocity were up to 40 seconds too early. Clearly, the predicted travel times benefit from using path-dependent travel time tables for each station, rather than a constant velocity model.

4.2.1.2 September 1998 refraction experiment

A second experiment was conducted in September 1998 by the same group that performed the previously described refraction experiment. This time, a third hydrophone was available which allowed us to perform locations and compare them to the locations announced beforehand. The previous survey was located at approximately 39°N, while this latest experiment was near 30°N. Unfortunately we have not yet been able to obtain exact ground truth for the 1998 experiment. We detected only a small fraction of the proposed explosions, and we have not yet received confirmation that the experiment went as planned.

The explosions in September 1998 were recorded at WK30, WK31 and PSUR. They were detected by the automatic system at WK30 and PSUR. The signals at WK31 had much lower *snr* compared with the other stations and were not detected by the automatic system. This is probably due to partial blockage by a local seamount. Analysts at the pIDC picked arrival times for the signal at WK31.

Because the variation in relative arrival times among the 24 explosions that we did observe is smaller than the expected measurement errors, we average the relative arrival times to obtain our best location estimate. We use these relative arrival times and our 2-D travel time tables to compute the travel time misfit over a grid surrounding the proposed shot locations. The RMS misfit contour of 2.5 seconds is shown in Figure 10. As expected there is little constraint in the E-W direction, but a fairly tight constraint in the N-S direction. The misfit shows the best locations to be slightly south (~15 km) of the proposed locations. It is possible that the experiment occurred slightly south of what was originally proposed, but more likely, our travel time model is slightly off. However the model would only require a modification of a few seconds to align the misfit contour with the proposed locations. This is consistent with *Brumbaugh and Le Bras* (1998) findings. The misfit contour's jagged edges are due to three factors: the size of the sampling grid, nonuniform changes in the 2-D model over the survey region, and the weak constraint provided by WK31.

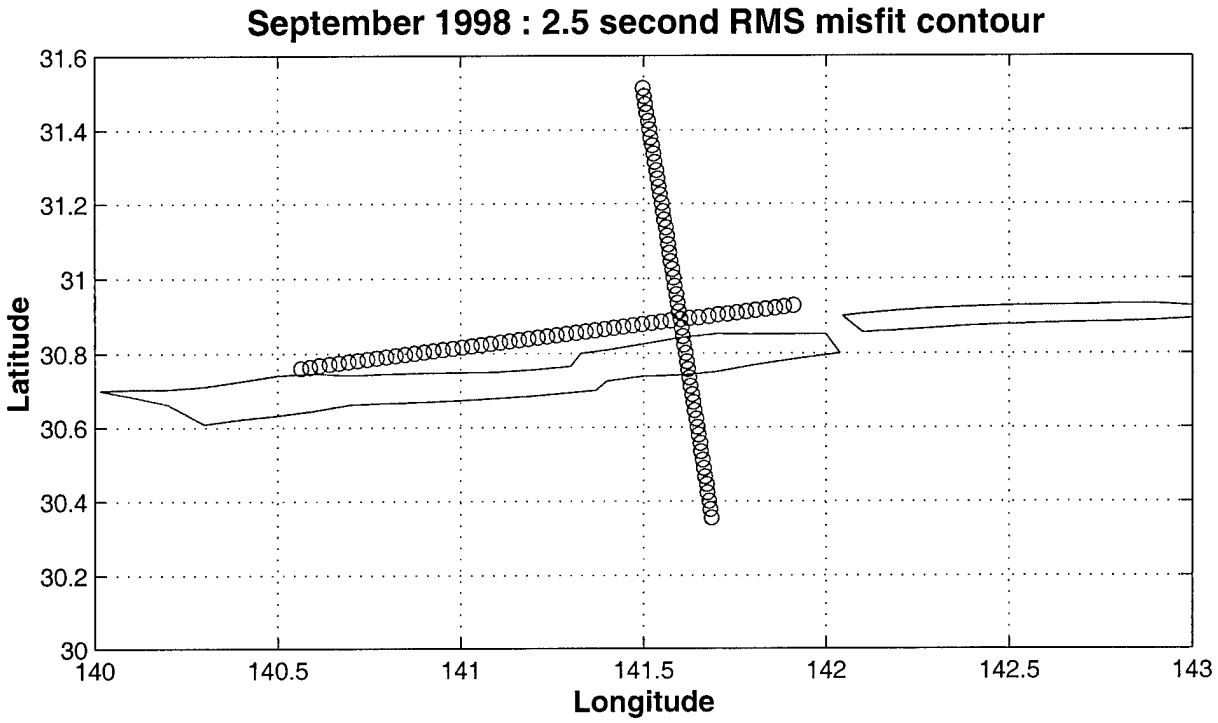

Figure 10. The RMS misfit for the average arrival times. The locations are well constrained in the north-south direction, but are virtually unconstrained in the east-west directions. The variation in the misfit in the east-west direction is due to the weak constraint added by the arrival time at WK31.

In conclusion, this data set provides further indication that our 2-D travel time models are highly accurate. It is also shown that hydrophone coverage is extremely important in constraining the location of an in-water event. Analysis of the bubble pulse for this experiment also confirms the consistency of bubble pulse delay measurements among several hydrophones (Figure 11). The use of two closely located hydrophones (WK30 and WK31) does not guarantee accurate locations, but depends on the event's location relative to the axis connecting the two stations.

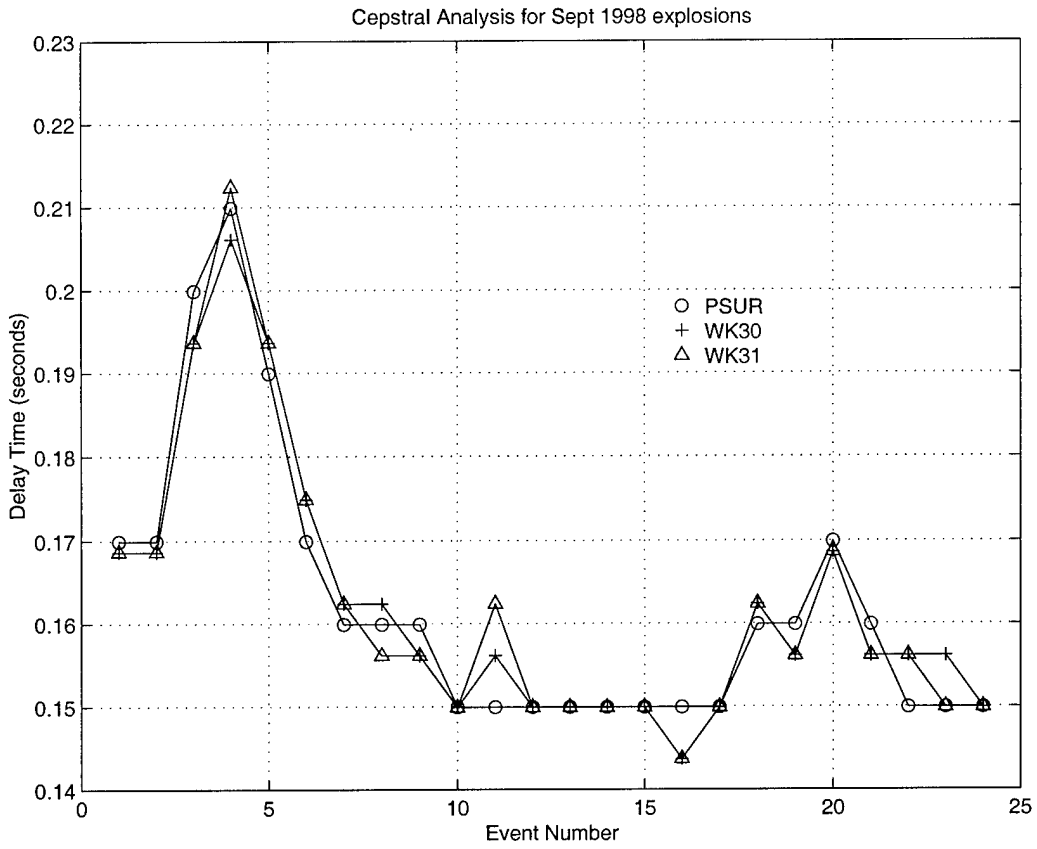

Figure 11. Bubble pulse delays computed from the cepstrum at each station for the September 1998 refraction experiment. The delay times vary from explosion to explosion but are extremely consistent between the three stations. The step-like appearance is due to the limited resolution (2 samples or 0.006 seconds) on the delay estimation.

4.2.2 T phase studies using the pIDC database

4.2.2.1 Blockage at Wake

Since the two sensors at Wake Island are widely separated, the effect of blockage by bathymetric highs may differ between them and this in turn may be of potential use in azimuth estimation by eliminating blocked azimuths from consideration at a given station. We have compared observed T-phase blockage patterns with seafloor topography. Rose diagrams for the two stations showing a quantity indicative of blockage is shown on Figure 12. The shadow of Wake Island at WK30 is very well defined (T-phases with back azimuths between 85 and 110 degrees are rarely seen).

However, WK31 does not display an area of complete blockage, and the expected shadow of the island and related bathymetric heights is not seen. This may be due to the closer proximity of the WK30 hydrophone to Wake island.

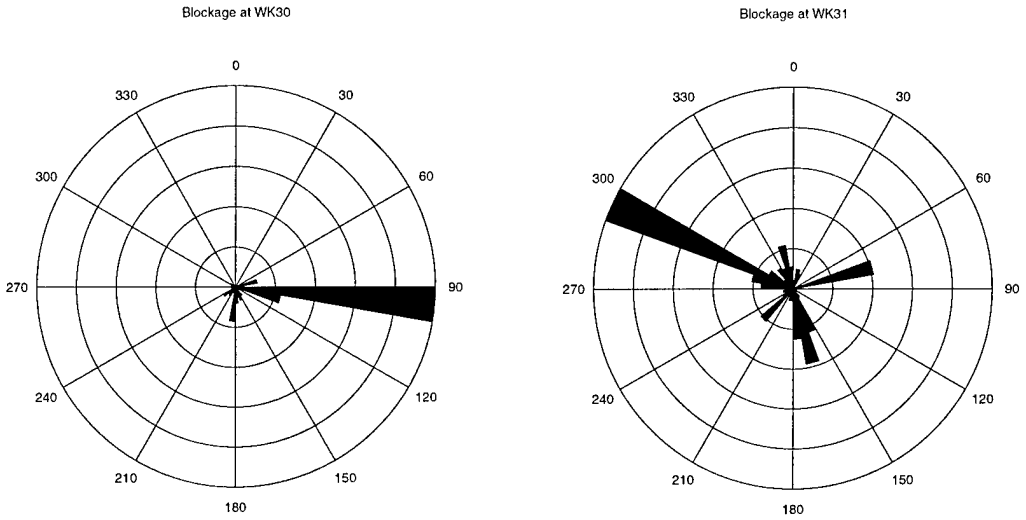

Figure 12. These rose diagrams show for each 10 degree bins a quantity computed from the number of REB events with an association only at one of WK30 (*wk30*) or WK31 (*wk31*) or at both stations (*both*). Bins with fewer than 15 events are not included in the display for lack of sufficient statistics. For station WK30, the quantity shown is $\frac{wk31}{wk30 + both + 1}$, and for station WK31 we show the converse. This rational function is designed to emphasize the bins where an event is seen only at the other station and presumably blocked at the station examined. Each plot is scaled according to its maximum value.

4.2.2.2 Travel time residuals in the Japanese subduction zone

During an analysis of residual travel times for T phases, it was discovered that there is a significant correlation between the T-phase arrival time residuals and the earthquake source depth. This means that hydroacoustic arrival time residual may be valuable as a depth discriminant. The time residuals of T phases for subduction zone earthquakes, which constitute a large portion of the T phase generators, systematically decrease with distance of the event to the coast as well as with event depth. Distance to the coast and depth are correlated at subduction zones and further study is needed to establish whether the T-phase residual can distinguish between the two. The observed decrease of the residual with distance from the coast and depth is expected since the path from the hypocenter to the place where coupling to the ocean basin occurs samples seismic velocities considerably higher than the water velocities. An illustration of this observation is shown for the Japanese subduction zone on Figure 13.

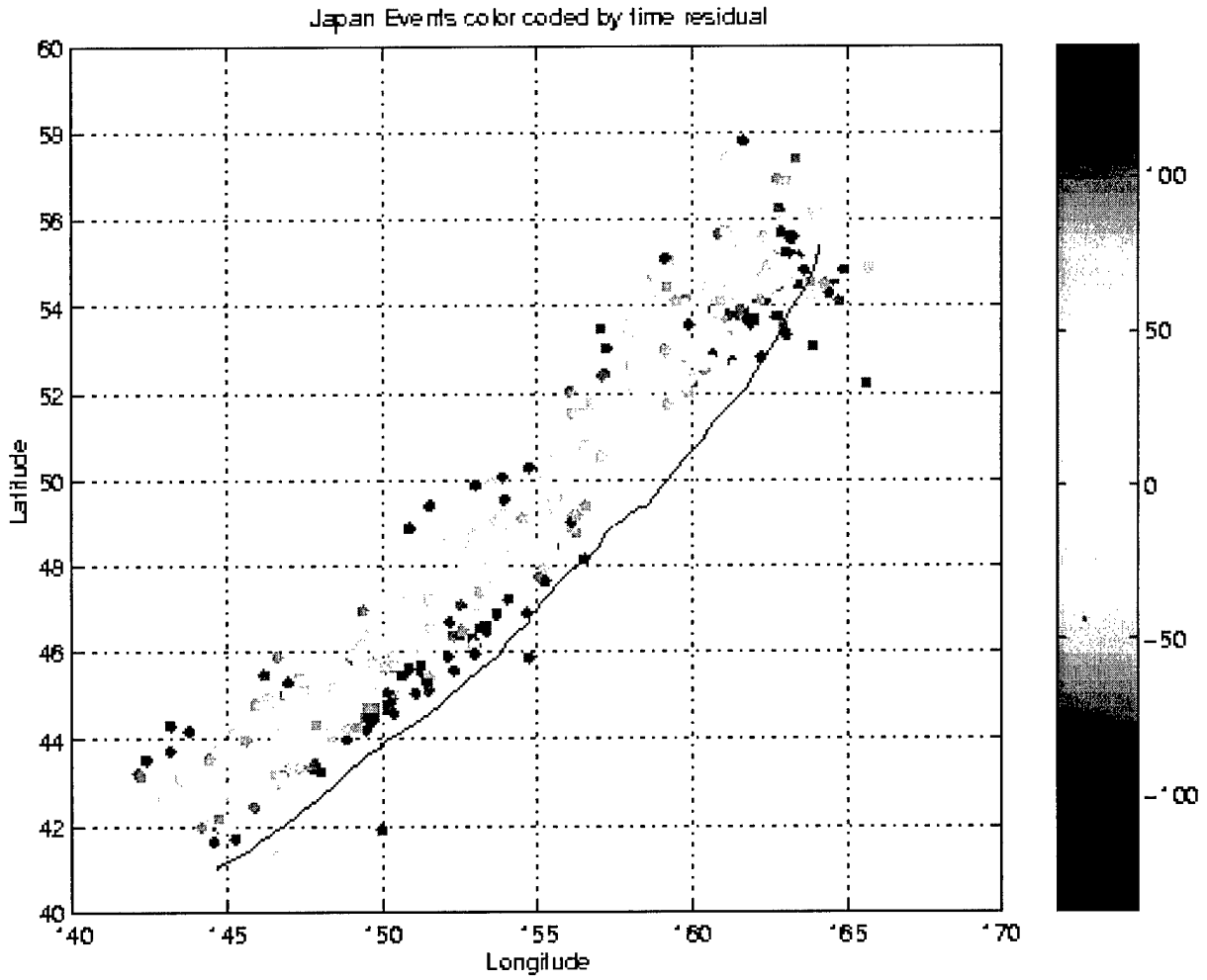

Figure 13. This map illustrates the systematic variation of the T phase residuals with distance from the axis of the trench and depth for the Japanese subduction zone. The residual is computed assuming a constant propagation velocity of 1485 m/s along the great circle path. The values of the residuals are color-coded according to the color bar shown beside the map.

4.3 Infrasound case studies

As attested by the state of development of the IMS infrasound network, the infrasound technology is not as developed as the seismic or even hydroacoustic technologies, where a few previously classified sensors have been made available to the R&D community. In order to acquire ground-truth data to help in understanding the waveform characteristics and propagation of infrasound waves, we have undertaken two cases studies involving infrasound detectors. The first case study is an examination of the signals from two low altitude atmospheric explosions in the White Sands Missile Range in New Mexico recorded at stations LSAR and TXIAR. The second case study involves two mining explosions recorded at TXIAR.

4.3.1 November 1997 Whites Sands Missile Range atmospheric explosions

A ground-truth data set for infrasound arrivals has been formed using two known explosions detonated at the White Sands Missile Range (WSMR) located in south-central New Mexico. The two explosions occurred in November 1997 and we show ground-truth information in Table 4.

Table 4: Ground-Truth of two WSMR Explosions.

Shot #	*Yield (pounds)*	*Date*	*Time*	*Latitude*	*Longitude*	*Elevation (meters)*
1	10,000	12 NOV 97	17:47:00.0248	33.6205	-106.4797	1473.3
2	20,000	19 NOV 97	18:00:00.0000	33.6209	-106.4797	1473.3

The Los Alamos, New Mexico (LSAR) and Lajitas, Texas (TXIAR) infrasound arrays are in the proximity of WSMR as shown in Figure 14. The expected distance and azimuth for the two stations relative to the two WSMR shots are shown in Table 5. Infrasound energy wavetrains directly attributable to the shots span a time interval of 4 to 5 minutes after the first arrival. The travel-time of the first arrival to LSAR and TXIAR was about 14 minutes and 27 minutes, respectively.

Table 5: Ground-Truth Station Information.

Shot #	*Station*	*Distance (km)*	*Azimuth (degrees)*
1,2	LSAR	249.5	183.1
1,2	TXIAR	545.1	331.4

At the LSAR array we identified four distinct arrivals that can be correlated between the two WSMR shots with similar timing and phase velocity within the sequence. The larger shot (November 19) has additional arrivals beyond these four. Four arrivals are also discernible at the TXIAR array for both shots, however the correlation of the timing between the shots is not as good as at LSAR. The quality of the arrivals at TXIAR is not as high as at LSAR due to the longer distance to TXIAR.

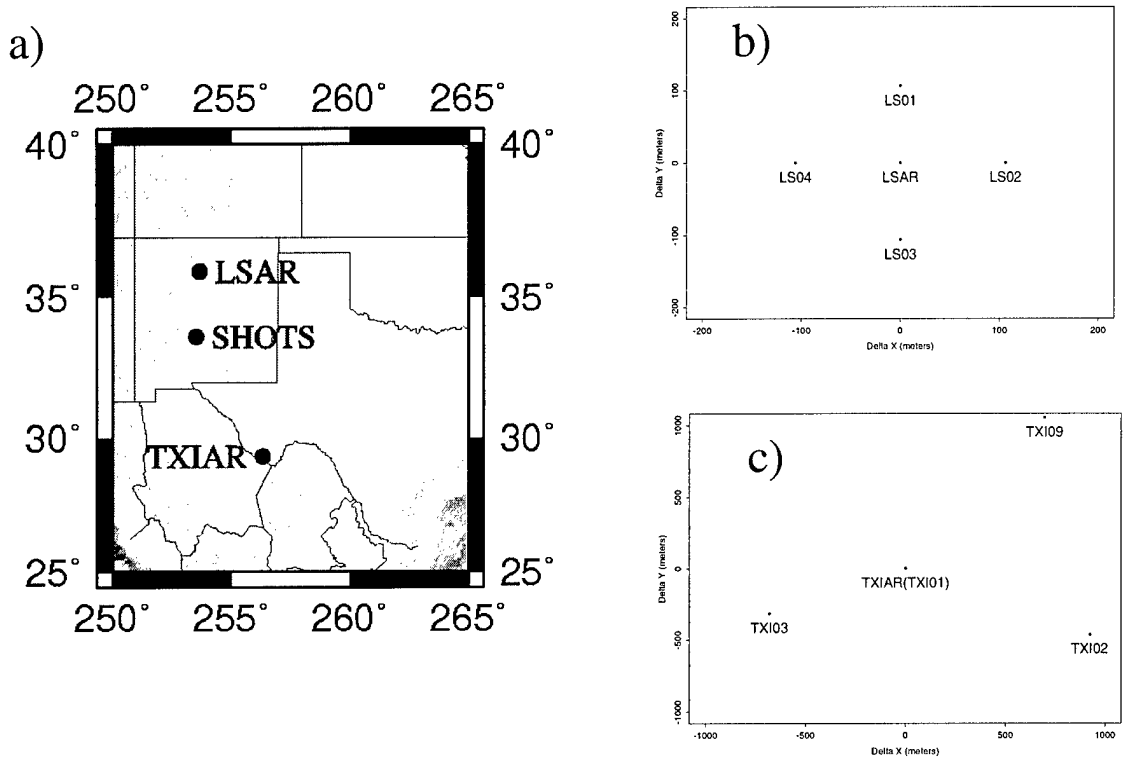

Figure 14. a) Location of WSMR shots and the LSAR and TXIAR infrasound arrays, b) Layout of LSAR channels, and c) Layout of TXIAR channels.

Approximate locations for both shots were estimated using first arrivals from the LSAR and TXIAR arrays. A constant velocity model of 320 m/s was used to derive these locations. Information for the two associated arrivals for the November 12 shot are listed in Table 6, while Table 7 lists the attributes for the two associated arrivals for the November 19 shot.

Table 6: Associations Used To Estimate The November 12, 1997 Shot.

Station	*Time*	*Distance (km)*	*Azimuth (degree)*	*Time Residual*	*Azimuth Residual*	*Slowness Residual*
LSAR	18:01:17	264	183.9	0.00007	0.0	-25.8
TXIAR	18:15:29	535	330.2	-0.0008	-16	7.89

Table 7: Associations Used To Estimate The November 19, 1997 Shot.

Station	*Time*	*Distance (km)*	*Azimuth (degree)*	*Time Residual*	*Azimuth Residual*	*Slowness Residual*
LSAR	18:14:02.8	276	181.9	1.75	0.1	-8.35
TXIAR	18:26:58.6	519	330.3	-1.75	0.3	-25.3

The event estimated location for the November 12 shot (first shot) was found to be 15 km south-southwest from the actual location and the origin time is 26 seconds late. For the November 19 shot (second shot), the estimated location was found to be about 27 km south-southeast from the actual location and the origin time was estimated to be 18 seconds early. The estimated events information is listed in Table 8 and the locations are shown on Figure 15. Note that the size of the error ellipses on that figure reflect a larger modeling error than is currently used in operations. A more detailed description of the data associated with these two shots can be found in *Jenkins et al.*(1997).

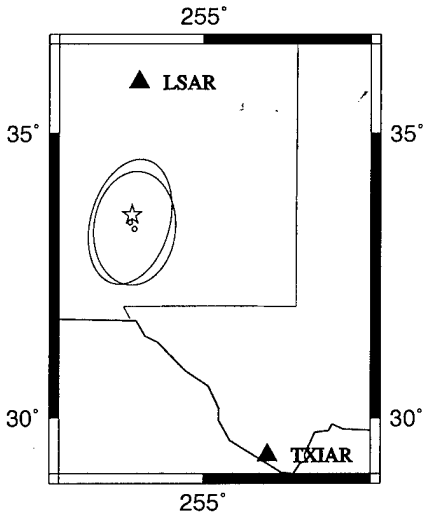

Figure 15. Estimated locations of November 12 and November 19 1997 explosions. First arrivals from the LSAR and TXIAR infrasonic arrays were used to form this location. Azimuth and time were defining at both stations for both shots. The 90% error ellipses are shown centered about the estimated locations. The actual locations, very close to each other, are represented by the star.

Table 8: Estimated Event Information for the two WSMR shots.

Origin Time	*Latitude*	*Longitude*	*Depth (km)*	*SMAJX (km)*
11/12/97 17:47:26.0	33.49	-106.53	0.0	132
11/19/97 17:59:42.0	33.38	-106.43	0.0	118

Interesting observations include that wavefronts with slower phase velocity arrived earlier than wavefronts with higher phase velocities at LSAR, indicating that the later rays reached higher altitudes and sampled higher velocities.

We computed theoretical travel times and phase velocities for arrivals from the troposphere (direct), stratosphere and thermosphere using a layered approximation to the standard

atmospheric model shown in Figure 16. These predictions are compared to the data observed at LSAR from the two White Sands shots in Figure 17 and at TXIAR in Figure 18.

The arrival time and phase velocity of the first main group of energy (three arrivals) at LSAR is consistent with predictions for arrivals from the stratosphere. This group of energy is more complicated for the first shot than for the second. The difference can probably be attributed to dynamic wind conditions since the paths are identical. The complication in this group may be the result of a triplication from a strong gradient within the stratosphere. The fourth observed arrival is more difficult to interpret. It arrives 40 s after the predicted arrival time for the first stratospheric multiple (i.e., two reflections from the stratosphere and one from the free surface). It is also about 60 s earlier than the predicted time of an arrival from the thermosphere. The observed phase velocity is closer to the predicted velocity for the stratospheric multiple. However, if the model is changed to match the predicted arrival of this multiple to the observed, then the predicted primary stratospheric arrival is shifted later. The interpretation as an arrival from the thermosphere is also problematic. It takes a significant change in the velocity model to move the predicted thermosphere arrival earlier by 60 s. One way to achieve this is to reduce the thickness of the low-velocity zone (LVZ) from 28 km to 20 km and to increase the velocity from 280 m/s to 290 m/s.

At TXIAR, the first three arrivals are consistent with respectively the direct, first stratospheric bounce and second stratospheric bounce. The direct arrival for the second shot has an especially good *snr* and *fstat* on the FK plot. The first arrival for the first shot is not as clear, with a very weak *snr* increase. No thermospheric arrival seem to be present in the data.

In conclusion, the two lower atmospheric explosions at the White Sands Missile Range in November 1997 provided us with ground-truth data against which to test our infrasound processing modules and propagation models. A detailed study (*Jenkins et al.,* 1998) of these two explosions, which occurred within a week of each other, pointed out variations in travel time (17 seconds difference at station LSAR for the first arrivals) and character of the waveforms between the two events. There were, however, similarities in the duration of the wavetrain within which individual arrivals can be correlated from one event to the other at both stations. A general observation is that the phase velocities of arrivals show an increasing trend with time within the wavetrain, consistent with the interpretation that they are sampling higher and higher layers of the atmosphere.

4.3.2 Mining explosions

A ground-truth data set for combined infrasonic and seismic arrivals at a single station has been formed using two known mining explosions: the first in the Minera de Carbonifera Rio Escondito (Micare) coal mining district in Coahuila, Mexico; the second in the Silver City copper mining region of New Mexico, USA. Both events were recorded at seismic station TXAR and infrasound station TXIAR and the data were provided to us by J. Bonner from Southern Methodist University (*Sorrels et al., 1997*). One of the events occurred on October 4, 1996 in the Micare coal-mining district in Northern Mexico and is referred to as the Micare event. The other event occurred on October 11, 1996 in a mining district in New Mexico and is referred to as the Tyrone event. A detailed study of the events is available in *Flanagan et al. (1999)*. Here we provide summarized descriptions and conclusions for each of the two events.

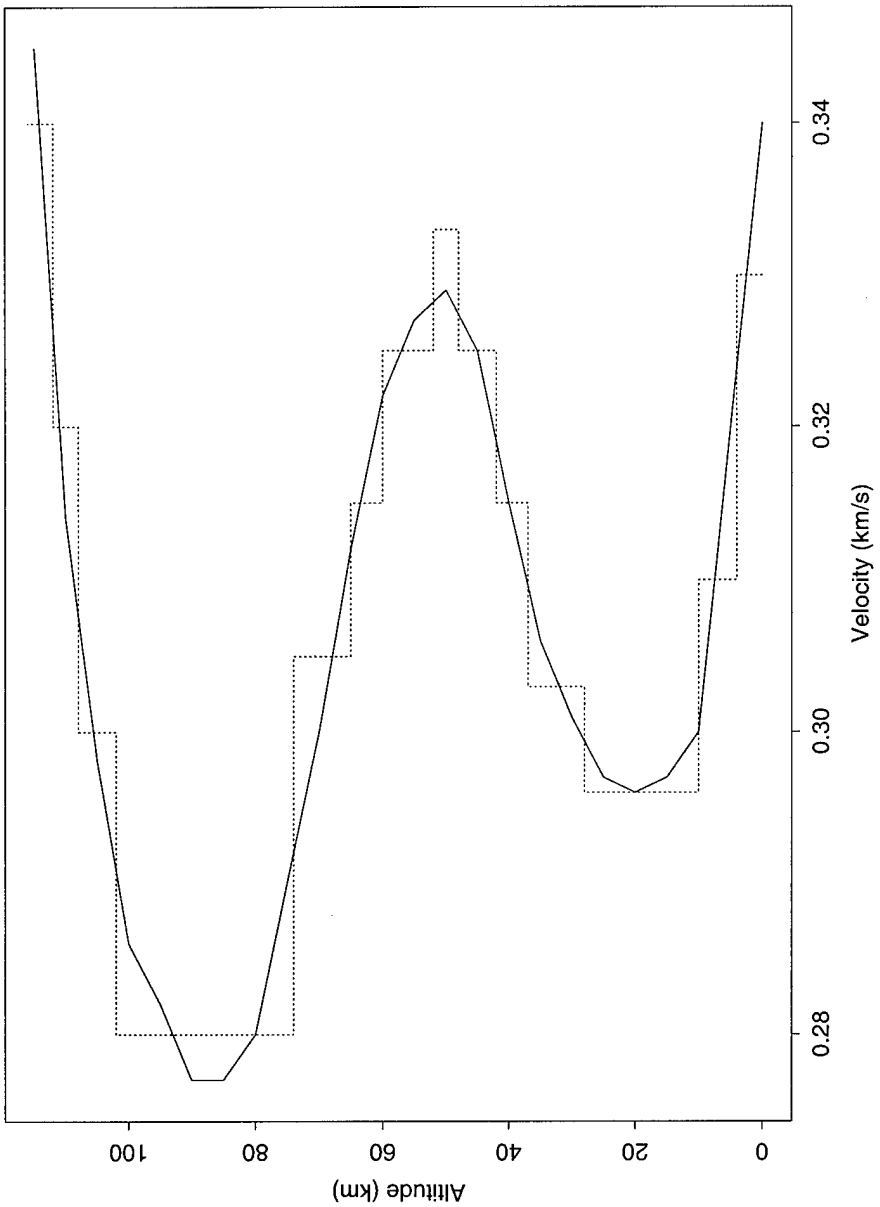

Figure 16. This shows an approximation to the sound speed as a function of altitude at 45°N in the absence of winds for spring-autumn from "The International Monitoring System for a CTBT," Preparatory Commission report CTBT/PC/IV/IMS/ITP/1 [adapted from *Donn and Rind*, 1972]. This was approximated by the layered model shown to compute theoretical arrival times and phase velocities for paths from White Sands to LSAR and TXIAR.

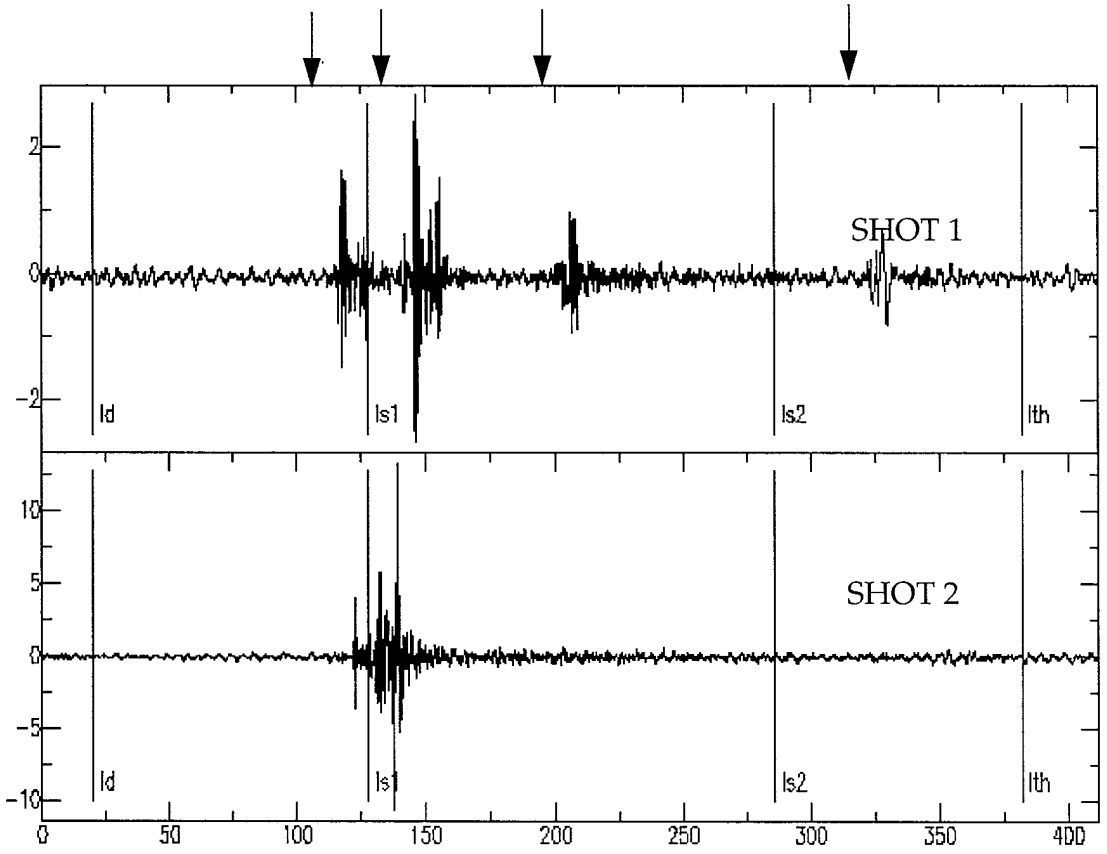

Figure 17. Waveforms at station LSAR on channel LS03 from two shots at White Sands about 250 km away. Four observed arrivals detected by DFX for the first shot are indicated by arrows on the top of the figure. The predicted arrival times of the direct (troposphere, Id), stratosphere (Is1), first stratosphere multiple (Is2), and thermosphere (Ith) paths are superimposed on the waveforms.

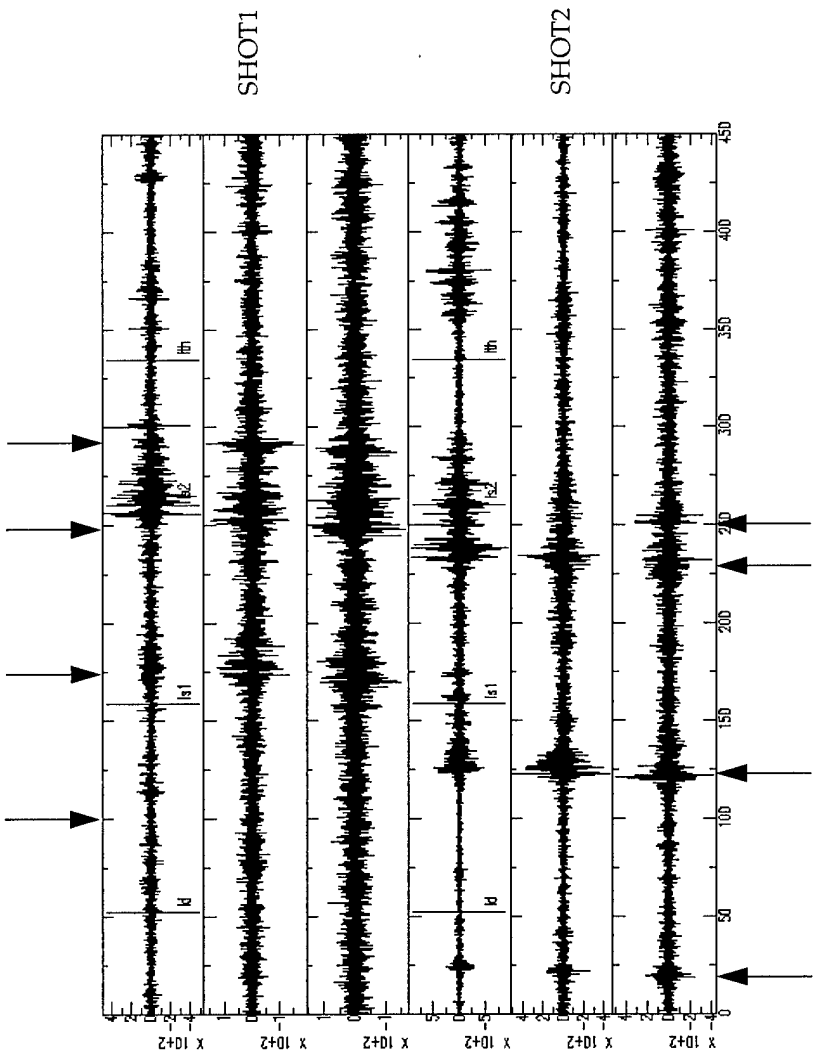

Figure 18. Waveforms at station TXIAR on channels TXI02, TXI03 and TXI09 for shots #1 and #2. Note that the *snr* is clearly lower, as expected, than at LSAR. The waveforms are aligned according to origin times. The start time on the waveforms is 26 m 40 s after origin time. Shown on channel TXI02 are the theoretical arrival times for the direct (Id), first bounce stratospheric (Is1), second bounce stratospheric (Is2) and thermospheric (Ith) arrivals. The vertical arrows on each trace show automatic picks. The earliest picks are used in the locations.

4.3.2.1 Micare Mining Explosion of October 4, 1996

One large infrasound arrival is recorded and two main seismic arrivals, *Pn* and *Lg*, are clearly observable and used in concert to form a location (Figures 19 and 20). The infrasound arrival shows high coherency across all four channels of the array. The measured travel-time for the *I* phase to reach TXIAR is approximately 20 minutes and 18 seconds. The picked arrival information used in the joint seismic-infrasound location is listed in Table 9 for both the seismic and infrasound phases.

Table 9: Arrival Information For Micare Mine Explosion.

Arid	*Delta Time (after Pn)*	*Time*	*SNR*	*Lead,Lag (sec)*	*Fstat*	*Phase (trace) velocity (m/s)*	*Average velocity (m/s)*	*Azimuth (deg)*
Pn	0.0s	17:03:49.7	good	1.05, 2.35	23.3	8561.	6530.	109.1
Lg	41.5s	17:04:31.2	good	1.05, 2.35	7.9	4660.	3448.	98.3
I	19m 31.8s	17:23:21.6	good	3.0, 5.0	10.6	376.	248.	104.7

The location obtained using the seismic arrivals alone is slightly to the north of the final location estimated using all three *Pn, Lg*, and *I* arrivals. The addition of the *I* arrival improves the azimuth, reduces the error, and moves the location farther to the south where it overlaps the mining area (Figure 21). The high phase velocity is compatible with the interpretation of the lone infrasound phase for this event as a thermospheric arrival. This is further confirmed by its travel time of 20:18.3 (or 1218.3 sec) at a distance of 303 km. The location is determined using a constant velocity of 320 m/s for the infrasound waves in the atmosphere, but this results in a positive time residual of 243.24 s.

We estimated another location where the infrasound constant velocity was set to a value more compatible with a thermospheric arrival (245 m/s). The location and error ellipse are only slightly different when using that slower constant velocity, although the time residual, as expected, is much smaller (-20 s vs. 243 s).

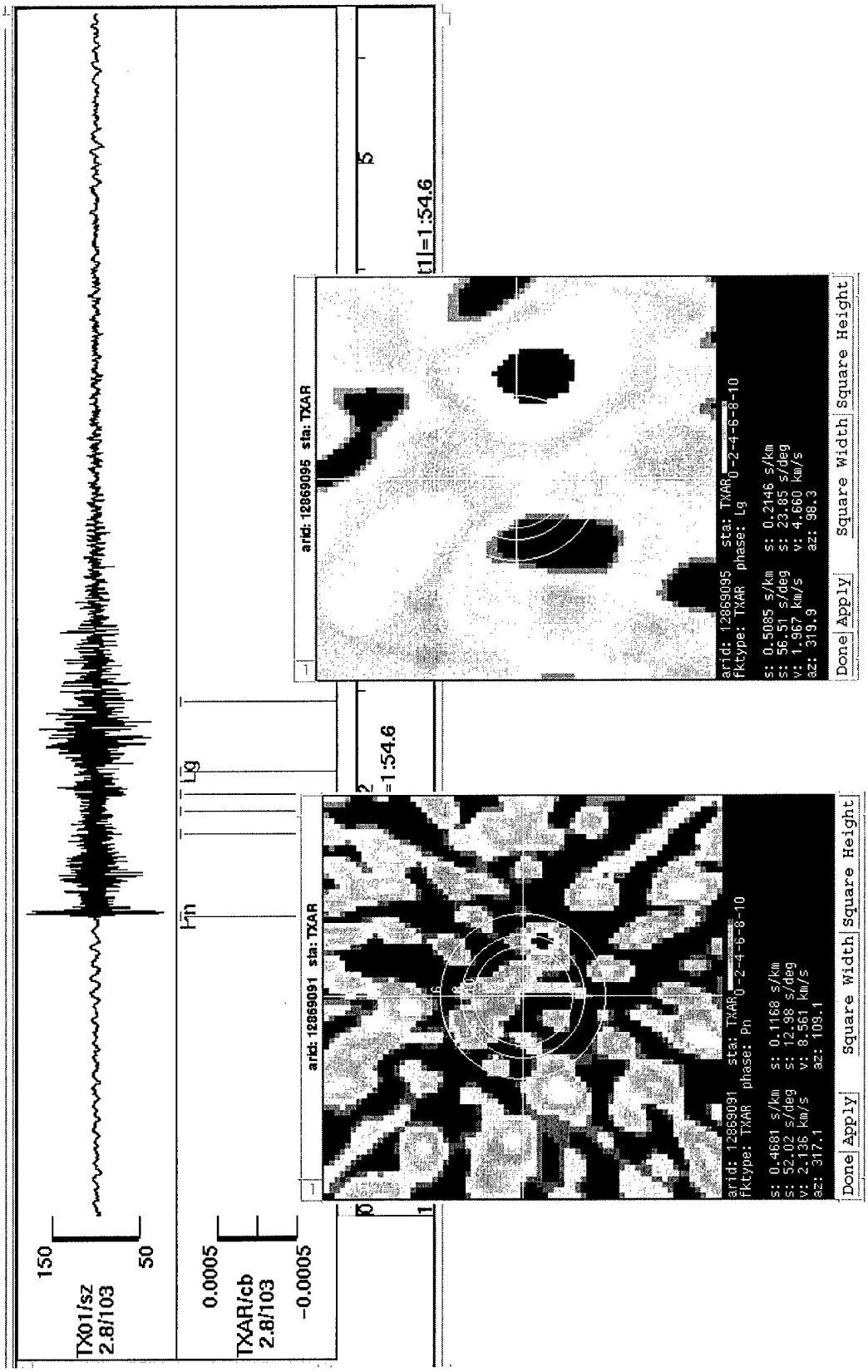

Figure 19. This is a display of one of the waveform on the seismic channels. Also shown on the figure are the FK displays for the Pn and Lg arrivals. The event to station azimuths are consistent with the direction of the MICARE mine from the TXAR station.

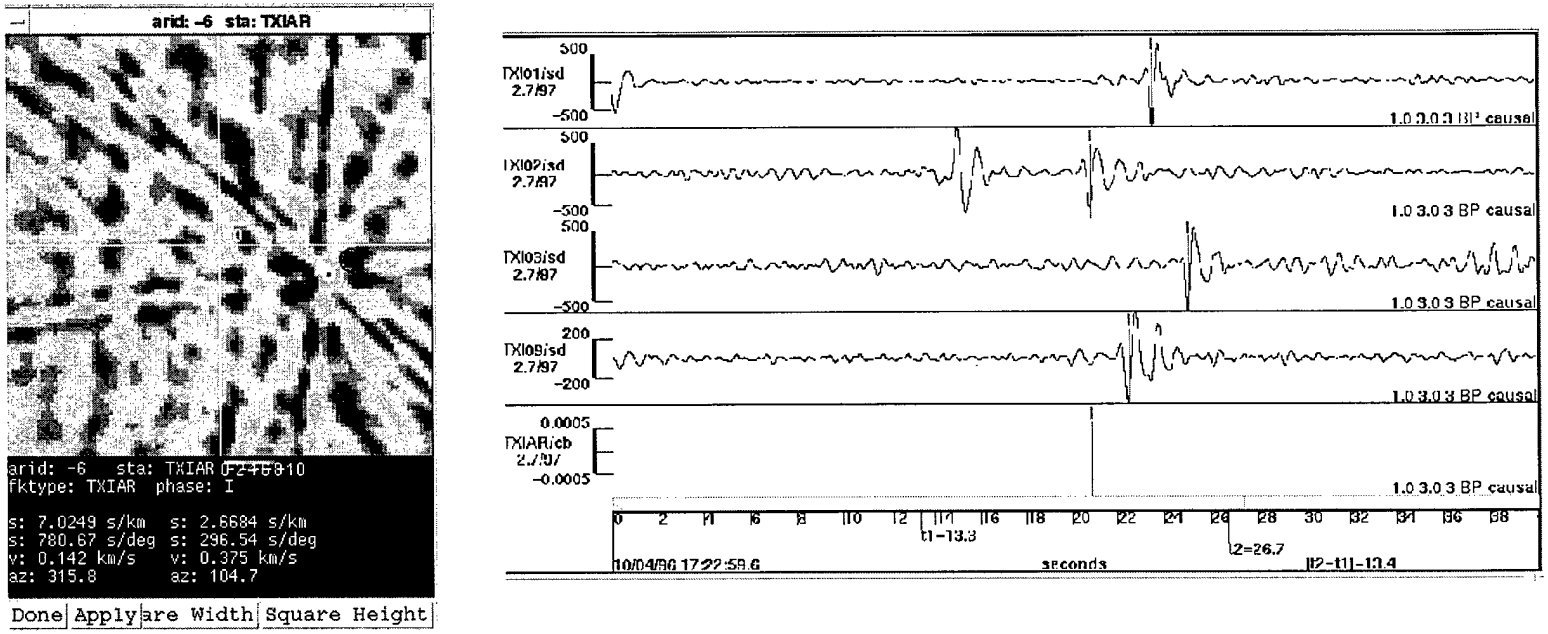

Figure 20. This is a display of the waveforms on the infrasound channels for the MICARE event. Shown next to the waveforms is the FK display for the high *snr* infrasonic arrival picked on the TXIAR/cb channel.

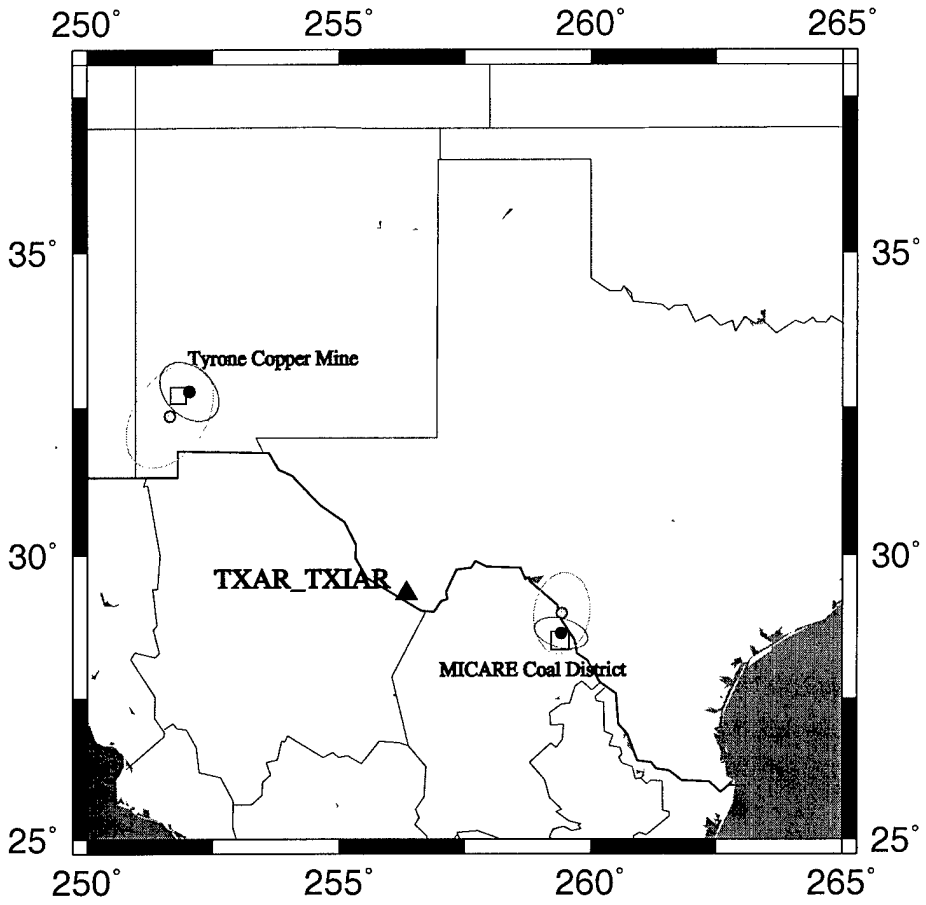

Figure 21. Estimated locations of October 4 and October 11, 1996 explosions. The circles on which the larger error ellipses (green) are centered indicate locations estimated using the Pn and Lg arrivals only while the centers of the smaller error ellipses (red) indicate the final locations estimated using the first infrasound arrival, I, together with Pn and Lg to form the location. The 90% error ellipses are computed in all cases. Note the improvement in the azimuth and reduced area error ellipse of the combined seismic-infrasound location. The actual mine locations are represented by the gray squares.

4.3.2.2 Tyrone- Silver City Explosion of October 11, 1996

Four infrasound channels are available for this event and there appears to be about 7 distinct packets of energy arriving between 27 and 33 min after the origin time, a situation quite different from the MICARE events situation where only one major arrival was present from the event. These individual waveforms are often seen on three out of the four channels, however, it was difficult to identify smaller arrivals on channel TXI02 as it was consistently noisier than the other 3 channels. Also, the multiple-arrival nature of the infrasound waves indicates a complicated path through the atmosphere possibly due to multiple reflections within atmospheric layers.

The ground-truth arrivals found for this station-event pair are listed in Table 10 Both seismic phases *Pn* and *Lg* are clearly observed and contain high frequency energy as seen in Figure 23 where the trace is filtered between 3.0 and 6.0 Hz.

Table 10: Arrival Information For Tyrone Explosion.

Arid	*Delta Time*	*Time*	*SNR*	*Lead,Lag (sec)*	*Fstat*	*Phase velocity (m/s)*	*Average velocity (m/s)*	*Azimuth (deg)*
Pn	0.0s	16:14:42.2	good	1.05, 2.35	4.7	8770.	7170.	302.0
Lg	1m 23.3s	16:16:05.5	good	1.05, 2.35	4.5	3640.	3460.	305.8
I1	0.0s	16:40:33.1	fair	5.0, 10.0	2.7	346.0	343.	295.5
I2	1m 25.5s	16:41.58.6	weak	-	-	-	326.	-
I3	2m 02.0s	16:42:35.1	fair	5.0, 10.0	2.0	322.0	319.	324.1
I4	3m 47.6s	16:44:20.7	fair	5.0, 10.0	3.2	309.0	301.	297.5
I5	4m 41.7s	16:45:14.8	weak	-	-	-	293.	-
I6	5m 38.1s	16:46:11.2	weak	-	-	-	284.	-
I7	6m 17.9s	16:46:51.0	fair	10.0, 10.0	2.2	335.0	279.	314.8

The emergent or packet-like appearance of the infrasound arrivals made them difficult to window for the F-K analysis and while we observed 7 individual detections (shown in Figure 23 and listed in Table 10) within the wavetrain only 4 of them produced acceptable, if rather low quality, F-K results (those with *snr*=fair). The optimal bandpass filter for the infrasound channels is found to be 0.8-3.0 Hz, and once again we experimented with the lead/lag window time and bandwidth parameters of the F-K analysis.

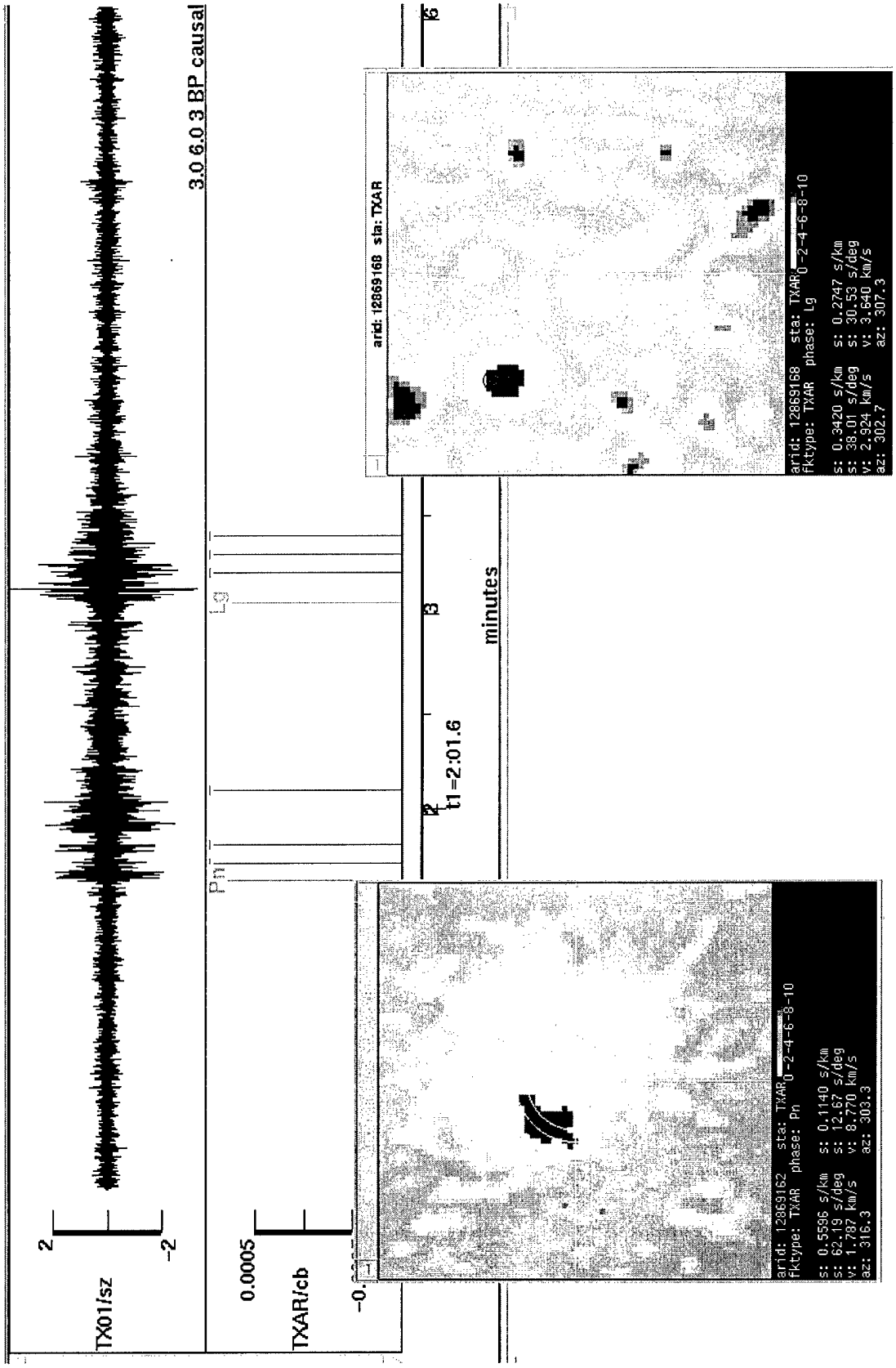

Figure 22. This is a dislay of waveforms on one of the seismic channels and FK-displays for both Pn and Lg arrivals for the Tyrone-Silver-City event of October 11, 1996.

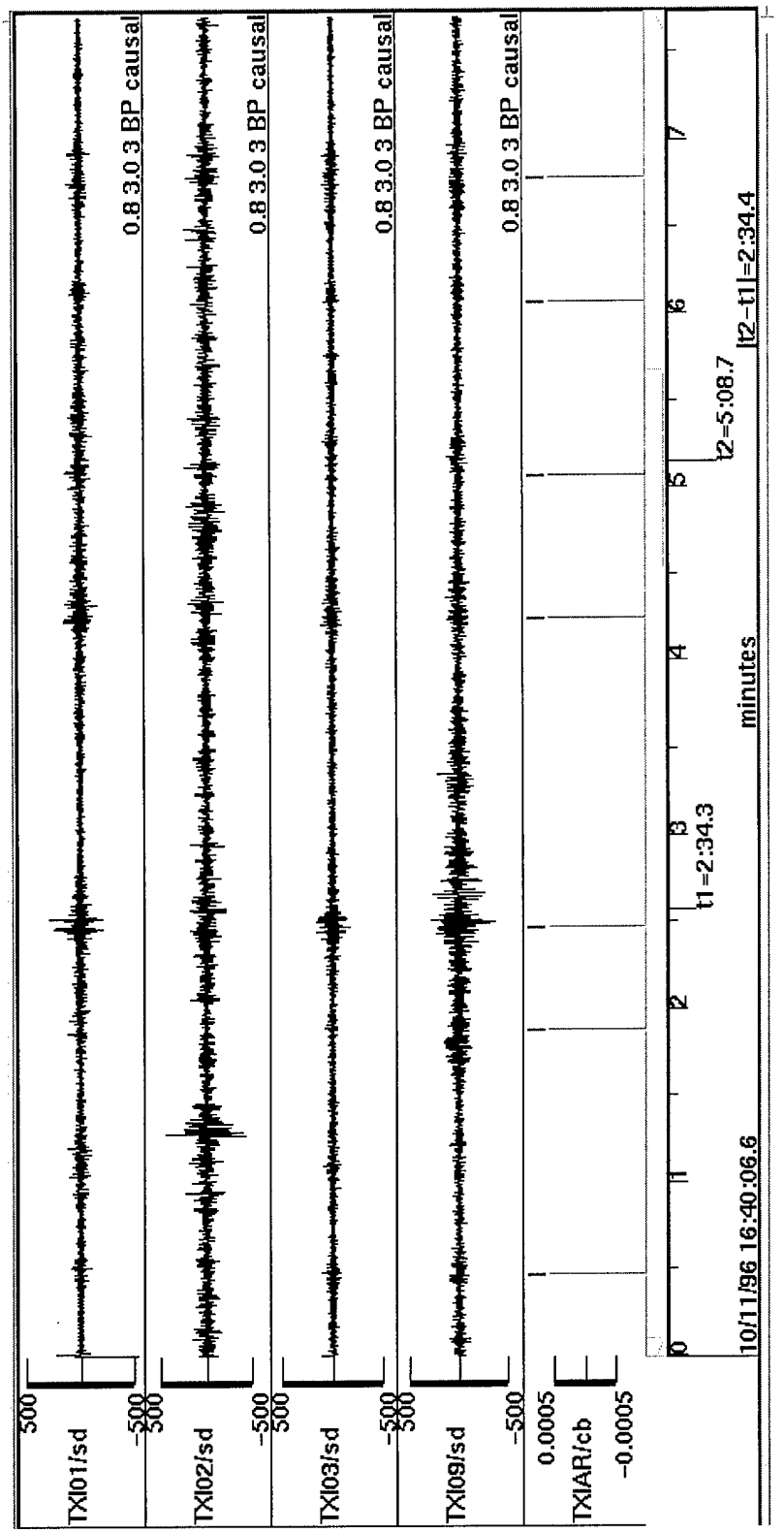

Figure 23. This is a display of the waveforms on all four infrasound channels for the Tyrone event. Note that there is a distinct coherent arrival with good *snr* (this is the arrival I3 in Table 10) around 16:42:35 with a duration of about 15 seconds. About 1.5 minutes later, a longer, lower snr wavetrain is visible. The snr of the signal for this event however is much lower than for the Micare event.

We find no reasonable azimuth or phase velocity for the *I2, I5* or *I6* arrivals. The F-K analysis of the remaining arrivals *I1, I3, I4*, and *I7* yields rather poor estimates of the azimuth. Phase velocities are consistent with predicted infrasound waves in the troposphere. While none of the arrivals produce a high quality *fstat,* we proceeded with computing a location by combining *I3* (the highest amplitude arrival as seen on all four channels) with the *Pn* and *Lg* arrival information shown in Figure 23. However, we do not use the azimuth determined by the F-K analysis to compute a location for this event; instead, we have employed an alternative technique (based on envelope functions described below) which appears to be more robust in extracting azimuth information from signals of poor quality.

Envelope function-derived azimuth and slowness determination

We present the results of an alternate means of estimating the azimuth of wavetrains impinging on the infrasound array, where the frequency content of the signal in the wavetrains is such that spatial aliasing or loss of phase coherency may occur. The method used is to take the envelope of the infrasound channels and produce an azimuth-slowness dependent display of an *fstat* value using the envelope functions low-pass filtered at 0.25 Hz. We are in the process of implementing this approach at the pIDC to estimate an azimuth for hydroacoustic signals using widely spaced sensors. Figure 24 shows the envelope-derived display for a 120 seconds segment around the highest *snr* arrival named *I3* in Table 10. The maximum value of *fstat* is obtained for an azimuth of 315 degrees as opposed to the F-K-determined azimuth of 324 degrees (Figure 24). When a longer time segment of 500 seconds is used to compute the envelope functions, the maximum of the *fstat* is obtained for the same azimuth value of 315 degrees.

The high frequency content of these infrasound arrivals makes coherent beam analysis very difficult on a long baseline array such as TXIAR. This example however shows that some benefit can be derived by using the envelope of the signals.

Location using seismo-acoustic data

The two main seismic arrivals, *Pn* and *Lg*, are used together with the infrasound arrival *I3* to form the location displayed in Figure 21. Although the *I3* arrival used in this location did not produce the highest *fstat* (see list below), it is clearly in the most prominent and largest amplitude wave packet as observed on the filtered infrasound records (Figure 23). The location estimated using the seismic arrivals alone is slightly to the southwest of the final location estimated using all three *Pn, Lg*, and *I3* arrivals. Similarly to the MICARE event, the inclusion of the infrasound arrival in the inversion improves the azimuth and reduces the error ellipse. The azimuth used to compute the location is the envelope function analysis-derived azimuth of 315 degrees, since this is by far the best estimate.

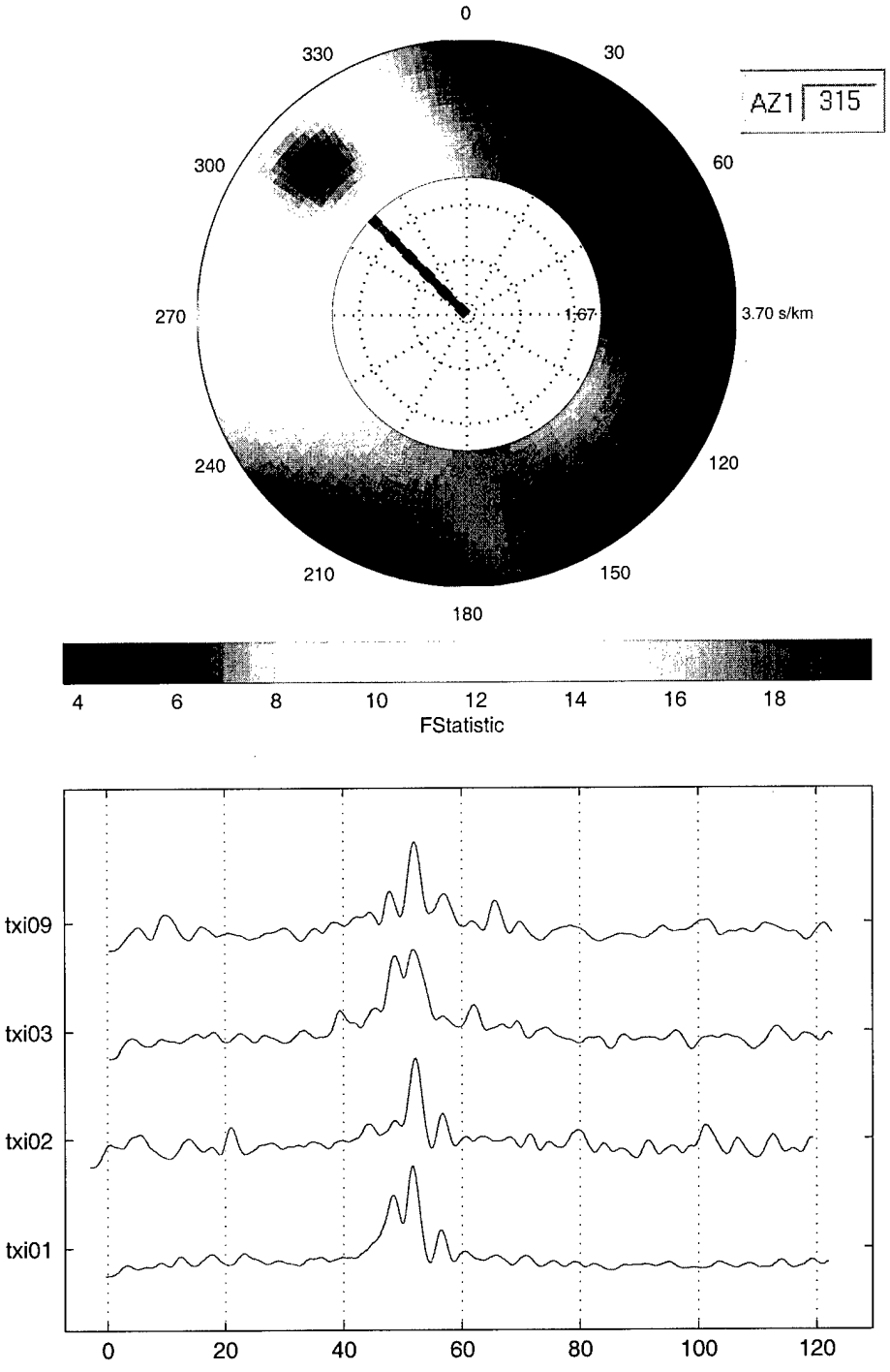

Figure 24. Envelope functions and slowness-frequency display of an *fstat* value for a short waveform segment around arrival *I3*. The data was filtered between 2.0 and 4.0 Hz before taking the envelope which is then low-pass filtered at 0.125 Hz. The maximum of *fstat* is obtained for an azimuth of 315 degrees.

5. Network processing simulations

5.1 Synthetic detection program SynGen

To achieve the objective of upgrading the network processing software in the absence of fully implemented networks and to satisfy the need for its thorough testing, a synthetic detections generator, SynGen (*Velasco and Brumbaugh,* 1998), has been developed to allow simulations of planned large networks. There are no data sets currently available that will test the required GA capabilities to detect events under various evasion scenarios with the full IMS network. Generating synthetic data is the only way to test these capabilities. SynGen may be configured to generate either natural seismicity or a prescribed set of events. It has the capability to generate seismic, hydroacoustic and infrasonic detections from these events, following a probability of detection approach. In addition to the detections generated from the natural seismicity or prescribed events, random detections may be added at each station to make the synthetic data set more realistic. The rate of random detections can be adjusted at each station, creating a data set matching the characteristics of a real data set. Each station of the network may be calibrated individually using noise background and threshold parameters. The simulated natural seismicity events follow a spatial and temporal distribution patterned after actual seismicity. This is achieved through the attribution of b values at a set of grid cells on the surface of the globe.

To calibrate the processing parameters and evaluate earth models, several case studies were undertaken involving hydroacoustic and infrasound data.

5.2 Simulation studies accomplished using SynGen

5.2.1 Stress testing of GA using the full IMS network

The current pIDC network falls short of the planned network specified by the CTBT, with 34 out of 49 of the primary seismic stations installed. We have conducted a test of the processing ability of the Network Processing software, GA, under stressful conditions, using the full IMS network. Specifically, we built a six-day hybrid data set. This hybrid data set consists of arrival data from all current pIDC stations that are also part of the planned IMS network coupled with synthetic data for the remaining proposed IMS stations. The synthetic data were computed using the REB event locations and a probability of detection algorithm at all stations with no real data available. These six days were chosen because they include an interval that caused the longest processing time observed over the more than two-year period that GA was in operations at the pIDC (August 28 through September 2, 1998). We used actual arrival data for the 34 operating IMS primary seismic stations, 3 hydroacoustic stations (ASC24, WK30, and VIB), and 1 infrasound station (WRAI). Realistic synthetic data for the remaining 15 IMS primary seismic stations, 8 hydroacoustic stations, and 58 infrasound stations were generated using SynGen and added to the real data to form the hybrid data set. Default noise characteristics were used at the stations that are not yet in operations.

One of the main objectives of this study is to assess the speed performance of GA for the full IMS network compared to the current network, on a particularly taxing interval. We collected CPU time statistics for the GAassoc portion of the processing for the six-day data set for three different configurations:

- We reproduced the initial run on the current pIDC network, with the version of GA that was in place at the end of August 1998. The results are the red curve in Figure 26, showing the dramatic peak for the problematic interval. Most of the processing time was spent in the large event extractor for that interval.

- An improved version of GA, delivered with PIDC_7.0 (the basis for R3), was used to re-process the six-day interval on the current pIDC network. The results are the blue curve in Figure 26, showing the dramatic improvement in efficiency for the problematic interval, while the other intervals remain approximately at the same level.

- The improved version of GA was ran on the hybrid data set for the full IMS network. The results are shown on the green curve in Figure 26. Note that the interval that was problematic with the current pIDC network did not cause as much of a problem for the larger IMS network. This can be attributed to the fact that the problem resided in the large event extractor and adding stations to the network can actually improve the situation in that module. GAassoc spent an average of 2 minutes more CPU time on a typical interval of the full IMS network hybrid data set than it did on a typical interval with the current pIDC network. This is expected since the planned IMS network is larger and there is a corresponding increase in the number of arrivals.

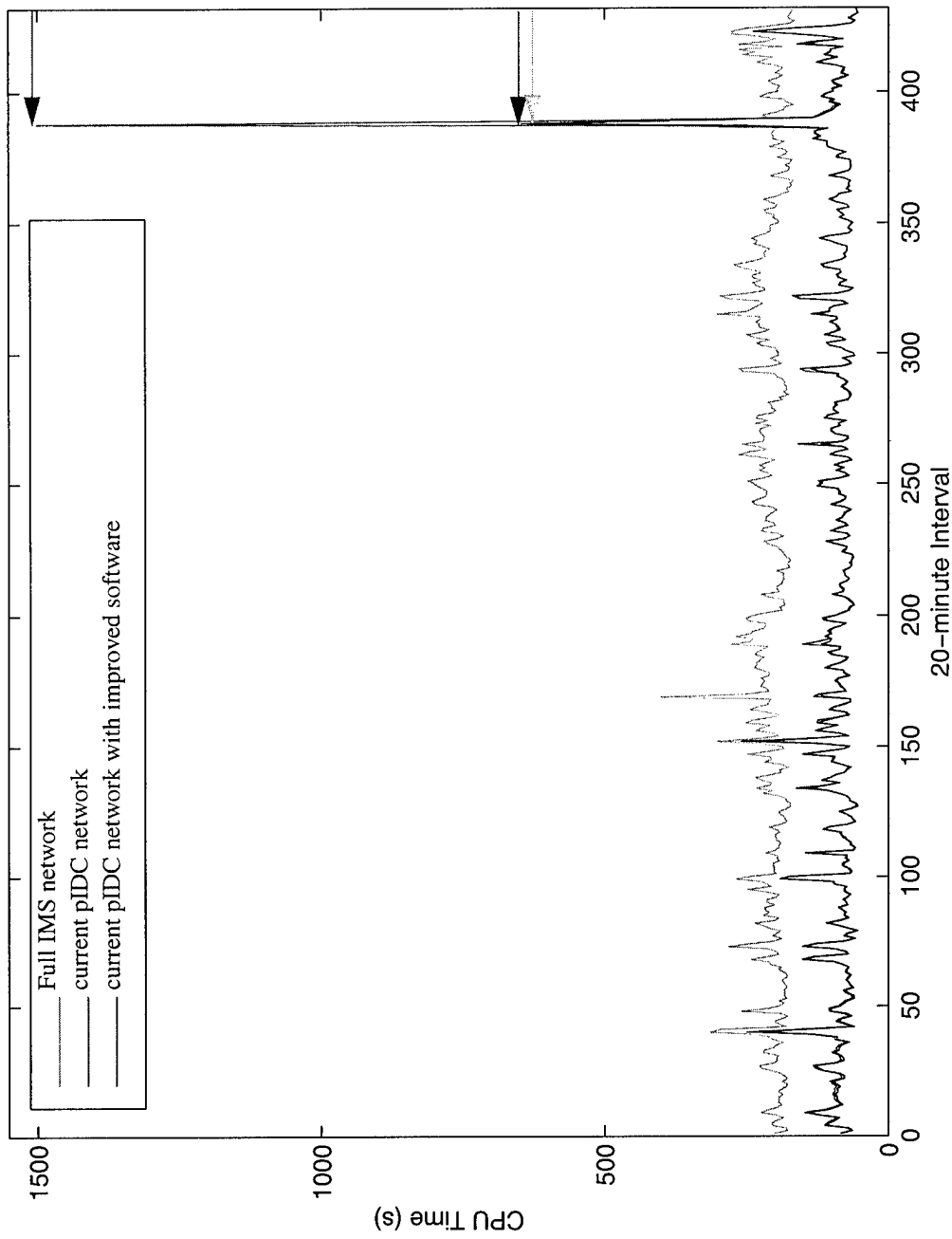

Figure 25. Comparison of the GAassoc CPU times for the full IMS network ($\mu = 211.8 \pm 34.8$), the current pIDC network ($\mu = 96.8 \pm 77.5$), and for GAassoc with the modified version of the large event extractor ($\mu = 91.8 \pm 40.3$). The colored arrows on the side indicate the processing times reached for the peak interval.

This study suggests that the average GAassoc processing time scales with the size of the network whereas the processing time associated with the non-linear, combinatoric nature of occasional heavy intervals is not necessarily higher when using a larger network. The GA software seems to be appropriately suited to process the expected data rate from the full IMS network.

5.2.2 Sensitivity of GA results to the presence of random noise and random variations in arrival time and slowness attributes

An important aspect of the characterization of a system is to evaluate its sensitivity to random variations in the input data stream. We have performed a study to do this evaluation for the GA automatic association system.

The SynGen simulation software has been calibrated at existing stations to generate realistic sets of arrivals from events representative of typical seismicity to which random arrivals are added. We used this software to quantify the sensitivity of GA bulletins to the presence of random arrivals and to the addition of perturbations in the values of the time, slowness and azimuth attributes for the arrivals expected from the actual events. The *deltim*, *delslo*, and *delaz* database attributes are used to set the amount of random perturbations added to the exact values of time, slowness, and azimuth.

A synthetic data set to be used as input to GA was generated by SynGen from a two-day set of 372 synthetic events (368 $m_b \geq 3.0$ earthquakes, two 1 kT marine explosions, and two 1 kT atmospheric explosions) for the 119 IMS stations. This data set will be referred to as the "baseline" for this study. The target bulletin of 268 events consisted of the 264 earthquakes with weighted count scores of 3.55 or greater, along with the four explosions.

In addition to the baseline data set, we generated a total of 41 synthetic data sets. The first set, referred to as the "detections and random noise" set, contains the baseline data to which a realistic set of additional random arrivals at all stations has been added. The other 40 sets were built to evaluate the effect of random perturbations in arrival time, slowness and azimuth, 4 different probability distributions of these parameters around their nominal values were tested. One distribution uses the data-calibrated values of *deltim*, *delslo* and *delaz*. The other 3 use respectively 2 times these values, 5 times these values and 1/10 of the value. The probability model used is a Gaussian distribution for all parameters. For each of the 4 distributions, 10 data sets were generated (10 realizations for each distribution).

We processed the 42 data sets with the PIDC_6.2.12 release of the GA software. The resulting bulletins were then compared with the target bulletin via BullComp. The number of linked (or detected) events, missed events, false alarms, and split events were determined using BullComp for each of the 42 simulations. The average distance, depth, and origin time residuals were then computed for all GA-formed events that were linked to events in the target bulletin.

Figure 26 is a graphical summary of the impact of random arrivals and Gaussian noise in the arrival data on the GA bulletins. Not surprisingly, the baseline data set, with no added Gaussian noise or random arrivals, yielded the fewest numbers of missed events, false alarms, and split events, and had the smallest residuals. When adding random arrivals and an increasing amount of perturbation to the detection features, the number of missed events, false alarms and split events increases as well as all location residuals.

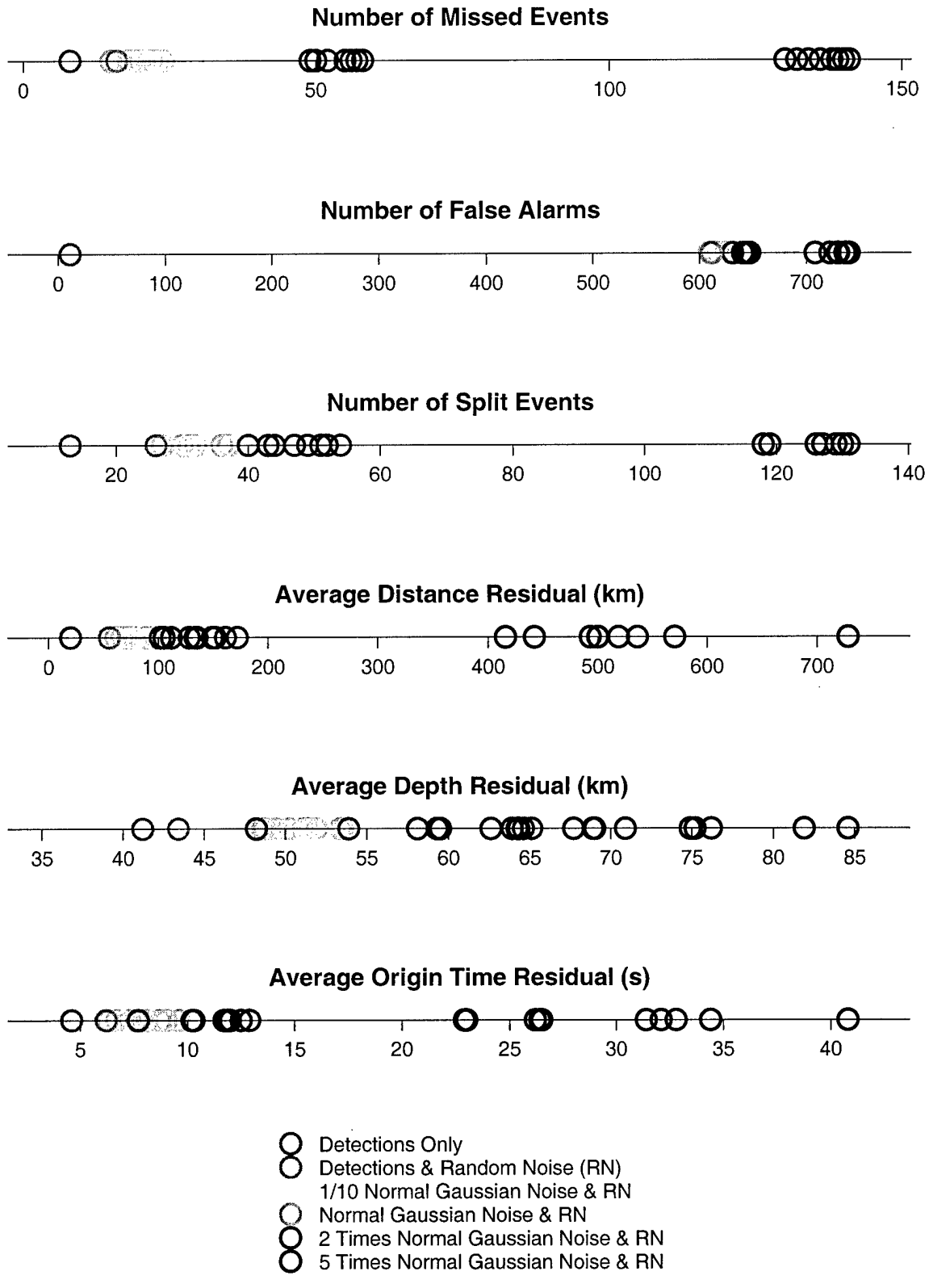

Figure 26. Comparisons of 42 GA bulletins with a target bulletin using six metrics. The 42 bulletins were derived from synthetic arrival data sets that contained variable numbers of detections and differing amounts of Gaussian noise.

Table 11 and Table 12 list the average values of the metrics for each set of simulation results shown in Figure 26. The data in this table provide a quantitative measure of the impact that random arrivals and Gaussian noise have on the bulletins produced by GA. Note that addition of random noise arrivals to the baseline data set (with no modifications to the *deltim*, *delslo*, or *delaz* parameters of the associated arrivals) resulted in a doubling in the number of missed events and, not surprisingly, the formation of a large number of false alarms (71%) as well as a 5% - 175% increase in the location and origin time residuals. Addition of "normal," or typically observed values of Gaussian noise to the probabilistic detections yielded a further 31% increase in the number of missed events, a very small 1% increase in the number of false alarms, and a further 19% - 51% increase in location and origin time residuals. Doubling the "normal" values of *deltim*, *delslo*, and *delaz* increased the amount of Gaussian noise added to the arrival time, slowness, and azimuth of the probabilistic detections, which resulted in a nearly 160% increase in the number of missed events, a 5% increase in the number of false alarms, and a 18% - 63% increase in the location and origin time residuals over the "normal" simulation results. Quintupling the "normal" values of *deltim*, *delslo*, and *delaz* resulted in a 550% increase in the number of missed events, an 18% increase in the number of false alarms, and a 43% - 520% increase in the location and origin time residuals over the "normal" simulation results.

Table 11: Missed events, false alarms and split events under various noise characteristics

	Missed events		False Alarms		Split Events	
	number	detected events (percentage)	number	percentage	number	percentage
Baseline	8	97.0%	11	4%	13	4.9%
Random Noise only	16	94.0%	611	71%	26	9.7%
Gaussian Noise and Random Noise	21±3	92.2%	615±5	71%	32±3	11.9%
1/10 Time Gaussian Noise and Random Noise	25±2	90.1%	598±2	71%	19±2	7.1%
2 Times Gaussian Noise and Random Noise	54±3	79.9%	643±5	75%	48±5	17.9%
5 Times Gaussian Noise and Random Noise	137±4	48.9%	728±9	85%	125±5	46.6%

Table 12: Location accuracy performance under various noise characteristics

	Distance Residual	Depth Residual	Origin Time Residual
Baseline	20±57	41±78	5±8
Detections and Random Noise	55±189	43±86	6±12
1/10 Gaussian Noise and Random Noise	22±56	41±75	5±12
1 Time Gaussian Noise and Random Noise	83±253	51±84	8±17
2 Times Gaussian Noise and Random Noise	135±457	60±93	11±28
5 Times Gaussian Noise and Random Noise	513±1260	73±98	30±68

GA accurately located both marine explosions in all 42 simulations. However, as the number of random arrivals and the amount of Gaussian noise increased, GA's ability to accurately locate the atmospheric explosions progressively decreased (Table 12).

A non-intuitive result of this sensitivity study showed that division of the "normal" values of *deltim*, *delslo*, and *delaz* by 10 increased the number of missed events by 19% over the "normal" simulation results, but the average values for the remaining metrics decreased. This is likely due to the fact that the same error values used in generating the synthetics were also used to normalize the time and slowness residuals.

This study of the sensitivity of GA to changes in the error attributes shows that larger values of these parameters lead to significantly larger numbers of missed events, false alarms, and location and origin time residuals in the final bulletins produced by GA. Any additional random detections further degrade the bulletin accuracy.

5.2.3 Minimum magnitude detectability study for underground and marine explosions using the combined Seismic and Hydroacoustic networks

The Comprehensive Test Ban Treaty (CTBT) specifies no lower limit on the yield of nuclear devices for which the ban holds. In principle, explosive events of any magnitude are of interest to the monitoring community. In the portion of the oceans where detection of an event is possible at three hydroacoustic stations, we expect the lowest detectable yield to be of the order of a few tens of pounds due to the extremely good acoustic conductivity of the oceans. Such extremely low yield explosive events, while of interest to test the detection, association and location system, are of relatively little concern, since they are unlikely to be of nuclear origin. In the portion of the oceans where an event cannot be detected using the hydroacoustic network alone because propagation paths are blocked, the primary seismic network will detect events above a magnitude threshold which is location dependent. This detection threshold will be lowered if the joined processing of hydroacoustic and seismic detections within the automatic association process allows for the formation of hybrid events, composed partly of seismic detections and partly of hydroacoustic detections. The objective here is to assess the usefulness of combining the seismic and hydroacoustic networks to determine the minimum magnitude of explosions that may be detected by three or more primary seismic and/or hydroacoustic IMS stations. The simulations were performed on a global scale to assess the benefit of forming hybrid seismic-hydroacoustic events. Specifically, these simulations help to characterize events that could be formed using a combination of seismic and hydroacoustic data but that could not be formed using either type of data alone.

Figure 27 quantifies the decrease in detection threshold that comes from combining seismic and hydroacoustic data as opposed to using just one of the technologies. The minimum threshold is decreased by 2 to 3 orders of magnitude in some very localized coastal regions, such as the North Atlantic off the coast of Spain, the Ross Sea, and portions of the Indian Ocean off the coasts of Africa and Indochina where seismic stations are located near the coastlines. We should note however that:

- The regions where the decrease in detection threshold is observed are limited in area and occur mostly near the coast where an infringement of the Treaty is less likely because of the heavy ship traffic close to most coastal areas.
- The improvement in detection level is of relatively little interest to the monitoring community because of the low magnitudes for which this is observed.
- Previous simulations have shown that the number of false alarms formed by GA when the event formation criterion allows combined seismic and hydroacoustic is too high a price to pay for the benefit that would ensue of detecting smaller yield explosion in a few areas of the oceans.

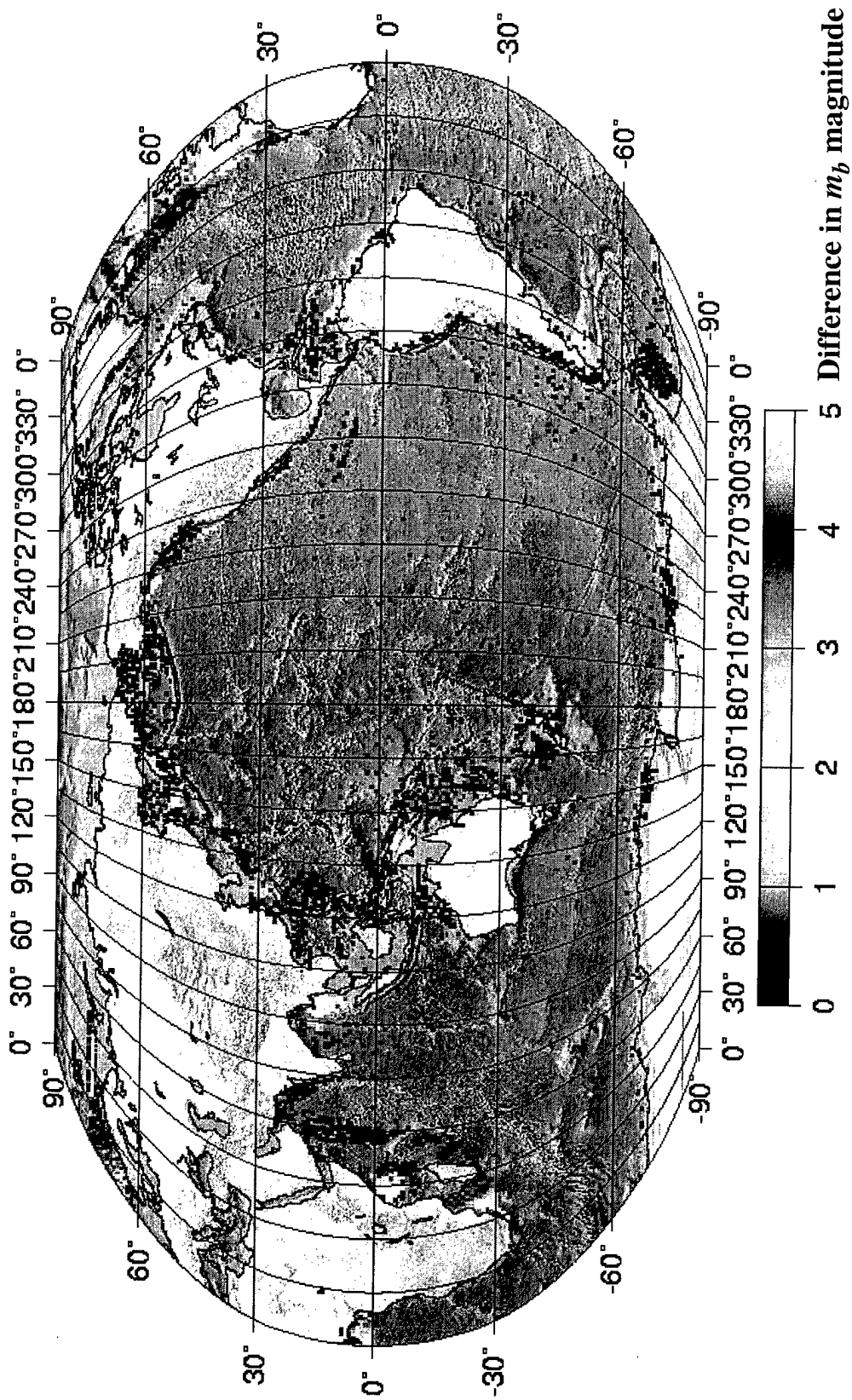

Figure 27. Difference in minimum magnitude detectable by at least three of the 60 combined IMS stations (locations of primary seismic and hydroacoustic stations indicated by triangles and starts, respectively) at 100% network detection criterion.

5.2.4 Simulation of Detections and Location of Minimum Detection Threshold Events

In the previous paragraph showing results also presented in *Gustafson et al.* (1999) and *Le Bras et al.* (1999), we have shown 99% global minimum detection maps derived using our SynGen program. These maps show the minimum magnitude for which detections will be present at a minimum of 3 stations. Such an event is detectable in theory, because of the presence of a sufficient number of detections. It does not mean, however, that the network processing step will always detect the event because of the presence of natural seismicity, random noise and possible mis-associations by any automatic association algorithm. We were able to quantify GA's ability to detect the minimum magnitude events using simulated data sets generated by SynGen.

Using SynGen, we generated 8 hours of synthetic seismicity. We then allowed SynGen to create probabilistic detections at the IMS primary and hydroacoustic stations for the 71 synthetic earthquakes, which had magnitudes in the range $3.00 \leq m_b \leq 5.27$. This set of 665 detections served as the background seismicity for this simulation. Note that no random arrivals at the IMS stations are present in this first set of simulations.

Next we created a set of global grid points at which we simulated synthetic explosions. We generated the 6307 grid points using GAgrid and a standard 3 degree spacing between adjacent grid centers.

We proceeded to generate a synthetic explosion at each of the 6307 grid points. Each explosion had an origin time corresponding to the midpoint of the 8 hour background seismicity interval and a source depth of 0.1 km. If the grid point lay in an oceanic area, a magnitude coupling coefficient of 0.4 magnitude unit was applied. The yield of each explosion was taken to be the minimum detection threshold determined at the nearest set of geographical coordinates as computed in the Global Event Detection Threshold study using the combined IMS primary seismic and hydroacoustic networks.

The probabilistic detections made at the IMS primary and hydroacoustic networks for each explosion were then added to the background seismic detections. We then processed a 20-minute interval containing this set of arrivals with the PIDC_6.2.32 version of GA. Note that the detections from each explosion were processed independently of the detections from other explosions; that is, we processed the arrival data using GA as described above, then removed all the explosion detections from the arrival table before generating the next explosion and its associated detections.

Finally, we used the PIDC_6.2.0 version of BullComp to determine which events formed by GA could be linked to the ground-truth explosions generated by SynGen. The criteria to establish that a GA event is linked to a SynGen event have been used previously to compare analyst bulletins with automatic bulletins:

- At least 2 defining phases must be shared by the GA event and SynGen event (note that the two defining phases may be H phases).
- The GA and SynGen event locations must be within 20 degrees of each other, or at least 50% of the arrivals linked/matched from the GA event to the SynGen event must be defining phases.

Simulations with no random arrivals

A set of 10 simulations has been completed for the synthetic data with no random arrivals added. GA was run on the synthetic data and BullComp was used to compare the bulletin produced by GA with the SynGen ground-truth to determine GA's effectiveness in locating the explosions. Averaging the results of the 10 simulations, we determined that BullComp linked 6049 ± 11, or 95.9%, of the SynGen ground-truth explosions to events formed by GA, where the uncertainty is one standard deviation.

Explosions that were not linked to GA events in at least one of the 10 runs were primarily located along the coastlines and in the oceans (Figure 28). These are also the areas where these minimum yield events are smaller in magnitude, mitigating the importance of detecting them. The failure to detect may be due to the parameter setting in the GA runs, where no mixed seismic-hydroacoustic event is allowed at the initial phase of event formation. No conclusion can be drawn at this point however without further simulations with a different setting of the parameters. Only a few of the underground explosions in continent centers were missed or mislocated by GA.

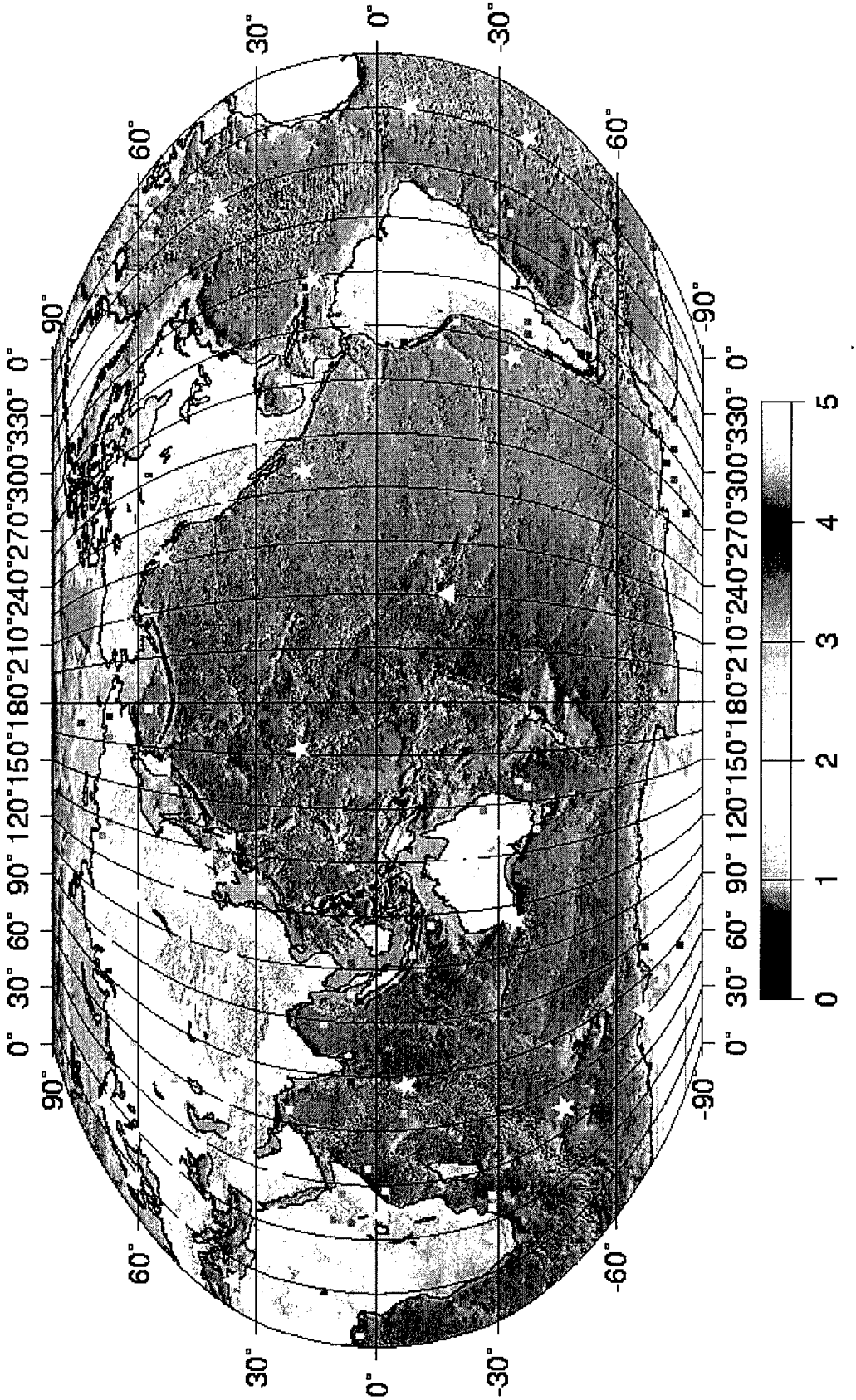

Figure 28. The locations of ground-truth explosions that were not linked by `BullComp` to events formed by GA. Two hundred and seventy-two (272) of the 6307 explosions were not linked to GA Events.

Simulations with random arrivals

Averaging the results of five simulations with random noise detections, we determined that BullComp linked 5765 ± 35, or 91.4%, of the ground-truth explosions to events in the GA bulletin, where the uncertainty is one standard deviation. This is a 4.7% decrease from the 6049 ± 11 events linked for the 10 simulations where only detections from global earthquakes were used as background noise. We expected to see a decrease in the number of linked events because the additional random detections increase the potential for misassociations.

Explosions that were not linked to GA events in at least one of the five runs were located primarily along the coastlines and in the oceans (Figure 29), a result duplicated from the 10 simulations without random detections.

Comparison of GA results to ground truth

We computed the average, standard deviation, and median distance, depth, and origin time residuals between the ground-truth explosions and the GA events linked to them for the two sets of simulations (Table 13). The average distance and origin time residuals for the five simulations with random detections are an order of magnitude larger than the same residuals computed for the 10 simulations without random detections. This observation is not surprising because of the additional number of random arrivals present in the five simulations which could be misassociated to events and thereby skew the final hypocenter and origin time. The depth residuals do not vary much between simulation sets - most of the GA events linked to ground-truth explosions were placed at the surface in both simulation sets (the ground-truth explosions were placed at depths of 0.2 km).

Table 13: Average Residuals Between "Ground-Truth" and GA Events

	Distance (km)	Depth (km)	Origin Time (s)
Average of 5 Simulations with Random Noise Detections	147 ± 621	0.9 ± 7.5	55 ± 243
Average of 10 Simulations without Random Noise Detections	11 ± 141	0.8 ± 4.6	4 ± 56
Median of 5 Simulations with Random Noise Detections	0.901	0.20	0.080
Median of 10 Simulations without Random Noise Detections	0.003	0.20	0.001

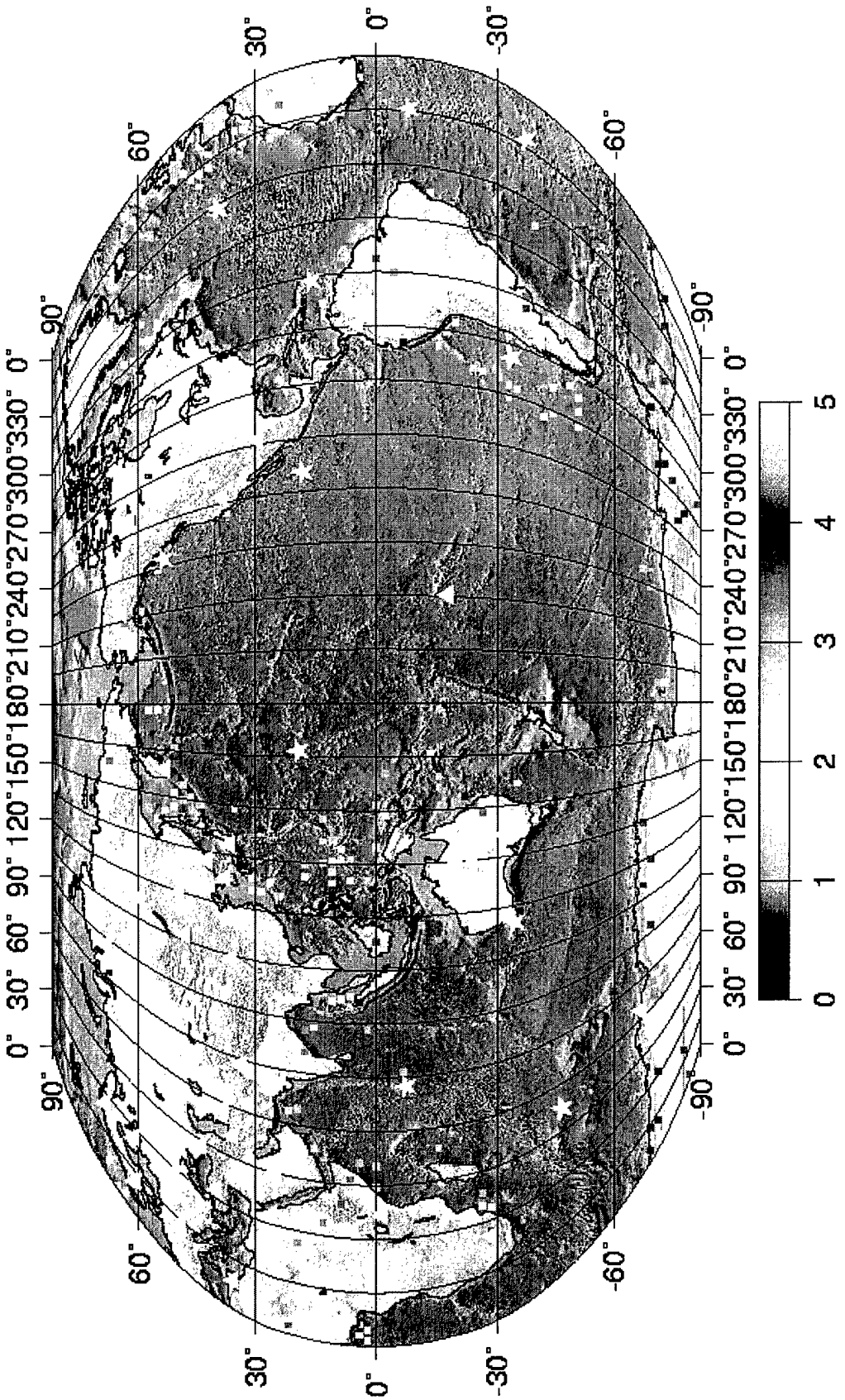

Figure 29. Locations and magnitudes of "ground-truth" explosions that were not linked to GA events in two or more of the five simulations with random noise detections. IMS Seismic and hydroacoustic station locations are indicated by triangles and stars, respectively.

6. Operational Results at the pIDC

6.1 Hydroacoustic processing results

The pIDC has performed routine analyst review and association of automatically detected hydroacoustic (mostly T) phases since May 1996. The automatic detector in place at that time was an adaptation of the seismic detector. A more elaborate detector adapted to the hydroacoustic sensor types has been put in place since May 1997. At the same time as this new detector was installed, preliminary identification and association of hydroacoustic phases has taken place. This section summarizes the statistics of hydroacoustic phase association at the pIDC.

6.1.1 The Reviewed Event Bulletin and the Standard Event List

The REB provides a large database of analyzed events from which to evaluate the automatic system (SEL). It is the largest global database available that associates *T* phases to seismic events (Figure 30), however we must keep in mind that *T* phases only get verified by analysts if they are associated to events that meet the REB criteria. The system has not been fully tested for its primary purpose of detecting large underwater explosions since none has taken place since the beginning of hydroacoustic processing. The SEL database tables contain an arrival for each detection made whether or not it is associated to an event, whereas the REB only contains arrivals associated to events. This means that many *T* phases in the SEL that are not in the REB are, in fact, real and probably generated by events which fall below the seismic detection threshold. In evaluating the automatic system relative to the REB we cannot determine the exact mis-classification rate of *T* phases in the SEL that do not make it into the REB, but we can examine how much of the REB is essentially produced by the automatic system. In addition, special studies have been conducted to measure the systems performance in classifying *H* phases which provides confidence that the system is able to distinguish between phase types.

There are some general observations about hydroacoustic signals in the bulletins that can be made. Over a two-year time span (May 1997 to May 1999) 13,582 *T* phases have been associated to 8,437 events in the REB (Figure 30). Taking station down time into account, this roughly translates into 25% of all REB events have an associated *T* phase. This percentage should increase by a few percent with the introduction of the new IMS hydrophone stations, particularly with the Indian Ocean stations. Even though the completed IMS hydroacoustic network will have much better coverage than the current network, the expected increase in REB associations is moderate because most near or below ocean events likely to produce *T* phases occur in the western Pacific which is well covered by the Wake Island station. A handful of events in Figure 30 clearly have *T* phases incorrectly associated to them, and are likely due to human error. For example, the event in the Persian Gulf is unlikely to have created a *T* phase which could travel to any of the current stations. However, these events represent a very small fraction of the total (< 0.1%). Below, we discuss each station individually and then compare results to classify performance based on station characteristics. Table 14 and Table 15 provide the overall statistics for each of the stations.

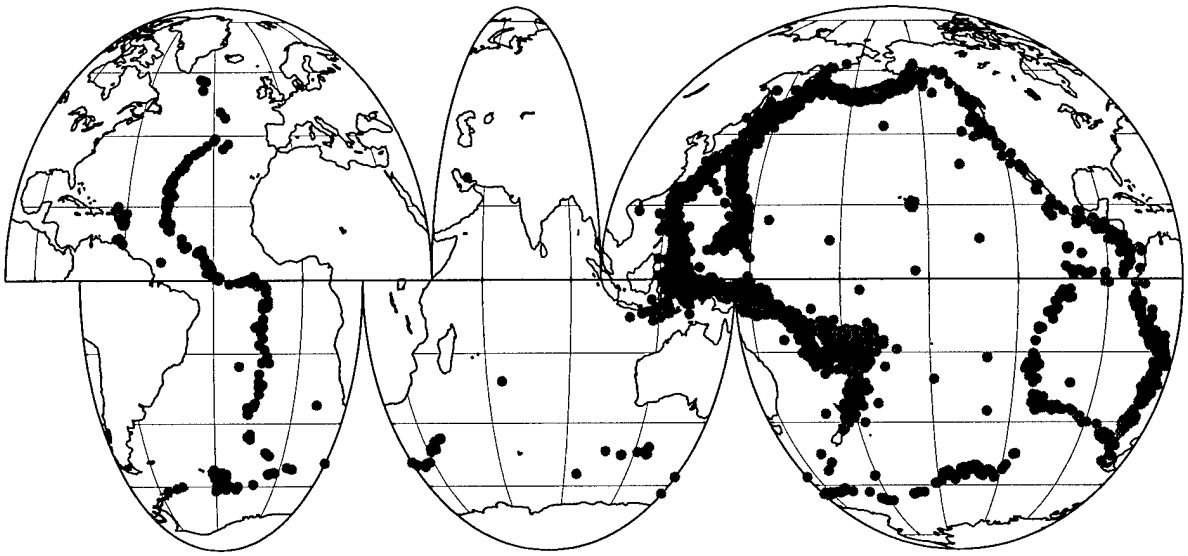

Figure 30. (from *Hanson et al.*, 2000). REB events with *T* phases between May 1997 and May 1999. There are 8,437 events in the two year period that have an observed *T* phase at a least one hydrophone station. Most of the events are along the western edge of the Pacific. There are a few events that appear questionable, but they compose a very small percentage of the total (< 0.1%).

6.1.2 Performance of Automatic Processing

The number of detections per day varies between stations (Table 14). The average number of detections - which have been corrected for station down time - range from 24 to 104 per day. These variations are due to factors such as ambient noise, local biological and man-made transient signals and the location of the station relative to seismically active oceanic regions.

Table 14: Average Number of Detections

	PSUR	WAKE	ASC	NZL	VIB
T phases detected in SEL per day	14	28	11	4	11
N phases detected in SEL per day	9	52	43	94	34
H phases detected in SEL per day	1.5	1.0	0	6.3	0
Total phases detected in SEL per day	24	81	54	104	45
T phases associated in SEL per day	3.9	10	1	0.6	1.5
H phases associated in SEL per day	0.9	0.2	0	3	0
T phases associated in REB per day	4.6	11	1.5	0.3	0.8
Total # of detections since start of processing	13,246	59,458	16,790	32,863	25,854

The performance of all the stations is remarkably similar even though the *T* phase association rate in the REB varies significantly (Table 15). The automatic processing of data from PSUR and Wake correctly classifies and associates *T* phases at nearly identical rates even though Wake

observes more than twice the number of *T* phases and 5 times the number of *N* phases. The T-phase station VIB misses more detections than the other stations, but the phase identification is quite comparable. Both NZL and VIB detect a lower percentage of the REB *T* phases associated to them than the deep water hydrophone stations. This is probably due to the high noise at these stations. However, VIB's identification rate is comparable to Ascension's. The phase identification performance at NZL is very poor because the low frequency energy in the signal, which is a major *T* phase discriminant, does not propagate through the shallow water. Once the automatic system detects and correctly identifies a *T* phase, there is a 65% to 75% chance it will get associated to the event that eventually makes it into the REB. This percentage does not seem to depend on station type which is expected since the association is only based on phase type and travel time residual. Because the association rate does not depend on the distance to events, we can conclude the majority of error in the arrival time is controlled by the event's location and/or coupling error and not the travel time models.

Table 15: Performance of Automatic Processing

	PSUR	WAKE	ASC	NZL	VIB
T phases associated in REB per day	4.6	11	1.5	0.3	0.8
% detected in SEL	77	85	81	70	61
% of these detections correctly identified as T	90	89	83	32	78
% of correctly identified *T* phases associated to same event in SEL and REB	71	73	65	75	67

6.2 Infrasound processing results

The pIDC received waveform data from infrasound stations with a very large delay and the majority of the arrival data did not make it on time to be included in the automatic bulletin.

In order to evaluate automatic processing of infrasound data, we re-processed a whole year's worth of data, between 1998335 and 1999334. The infrasound stations listed in Table 16 were operating during that time period.

Table 16: Average Number of Detections

	LSAR	DLIAR	SGAR	TXIAR	WRAI	PDIAR
Number of days with at least one detection	304	263	264	30	208	168
I phases automatically detected per day	26.0	4.1	12	0.3	2.1	3.2
N phases automatically detected per day	5.0	5.5	2.7	3.4	5.6	2.9
Phases automatically detected per day	31.0	9.6	14.7	3.7	7.7	6.1
I phases associated in SEL per day	6.0	0.6	5.4	0.1	0.1	2.7
Total number of associations during the year	1840	147	1429	3	21	459
Total number of detections during the year	9407	2527	3877	112	1613	1029

Table 16 shows the detection and association statistics for the whole year. The propagation model used in the association process is a constant velocity model (320 m/s). An automatic event may be formed with a minimum of two infrasound arrivals. We are presenting the results of the automatic processing in Figure 31 for events with two or more associated *I* arrivals and Figure 32 for events with three or more associated arrivals. Most of these events have not been reviewed by analysts or infrasound specialists and we have no definite criteria to label these events as real or false alarms, however a review of a similar set of events was performed at the CMR (*Gault and Brown*, 1999) where it was concluded that 20 to 30% of the events they looked at were real transient events, although no ground truth is available to confirm their location and time of origin. It is likely that some of these events originate from mining or quary blasts or ordinance disposal.

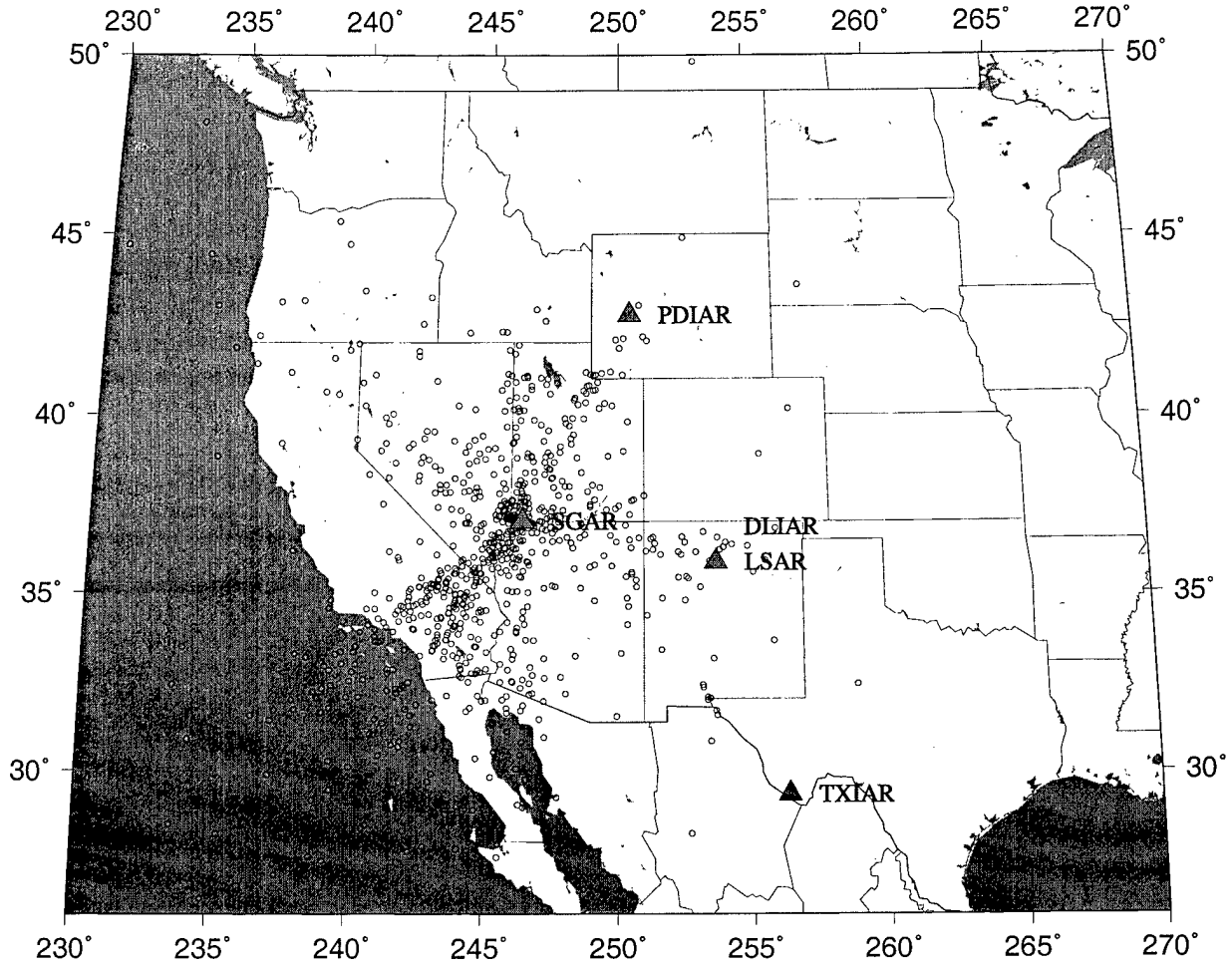

Figure 31. Location map showing the North America infrasound stations available at the pIDC (triangles) and the locations of the 946 events with two or more 'I' phases that occurred between 1998335 and 1999334.

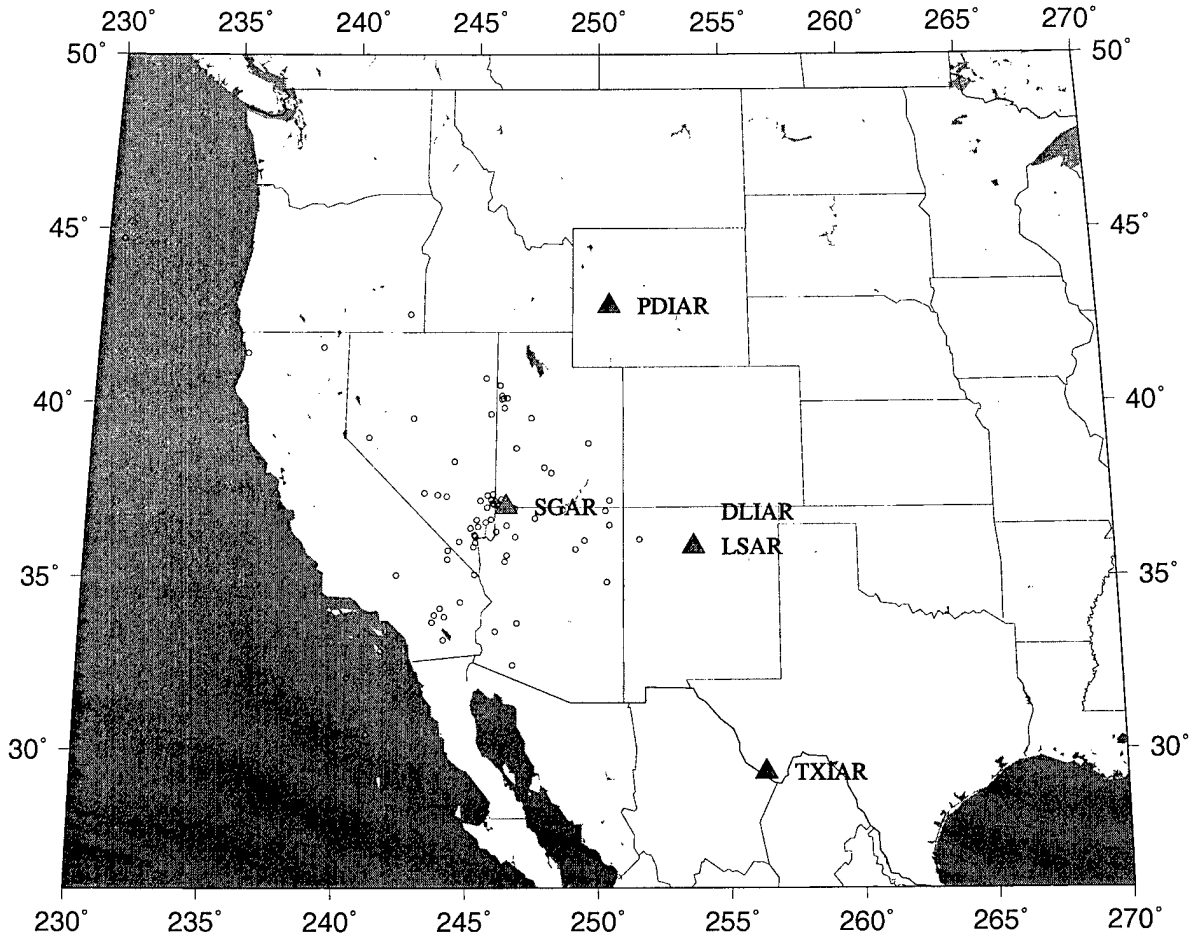

Figure 32. Location map showing the North America infrasound stations available at the pIDC (triangles) and the locations of the 94 events with three or more 'I' phases that occurred between 1998335 and 1999334.

7. Next-Generation Multi-Sensor Data Fusion

7.1 Introduction

To provide the best answer to design issues, it is important to keep abreast of developments in the literature and commercial world. In the next-generation part of this project, we have conducted several studies including a literature review, spatial representation product evaluation and a design for insertion of new sensor processing modules. In this section, we summarize the content of these three studies.

7.2 Literature review

We have performed a review of the multi-sensor data fusion literature to survey the state-of-the-art in that field and to provide input to the design of the next-generation multi-sensor data fusion systems (*Hanson et al.,* 2000). In this section, we summarize the main aspects of this literature review.

Most of the work in the area of data fusion as taken place in the military arena where the most common applications are target or missile tracking. An effort at classifying levels of abstractions in data fusion has been conducted by the Data Fusion Group of the Joint Directors of Laboratories (JDL), which recognizes five levels of abstraction: preprocessing, single object refinement, situation refinement, impact assessment and process refinement. The various components of the current CTBT data fusion system can be classified within these five abstraction levels, with for example DFX fitting in the context of level 0 and 1(preprocessing and single object refinement), GA in the context of level 1 and 2 (single object refinement and situation refinement) and Event Characterization in the context of level 3 (impact assessment).

The very essence of fusion is to combine data of different types, where the raw measurements usually consist of time series or spectra. We have attempted to classify the various level of data compatibility between different measurements or extracted features. Raw measurements at close-by sensors can be correlated and combined directly, as in seismic or infrasound beam forming. We call this type 1 data compatibility. Type 2 compatibility is when a transfer function has to be used to transform raw measurements to meta-data in order to combine these meta-data measurements at various sensors. Type 3 compatibility refers to extracted features measuring different physical values that can be related and combined together, as in using an azimuth from an infrasound sensor to corroborate time picks on three hydroacoustic stations. Finally type 4 is for weakly related data types.

Systems that deal with data fusion can be built using a number of different high-level architectures. We have recognized three different models of architecture. The first one is the centralized fusion with raw data, where the fusion into physical models and identity declaration is done without any intermediate step. The centralized fusion with feature vector data architecture performs feature extraction and association as intermediate steps. Finally, the autonomous fusion architecture has an extra step of classification or estimation after feature extraction so that the association step is performed on state vectors and identities rather than features. Figure 33 illustrates these three different architectures. Real world fusion systems tend to be hybrid models

of the three above architectures. The centralized fusion models often make up the subsystem portion of a more autonomous fusion architecture.

Positional fusion determines if a target exists and associates all sensor data to that target. A location is then derived from the features of the association set. Identity fusion is related to position fusion in the sense that the process involved is to identify the target rather than its position.

The computational techniques used to perform the decision making process in data fusion include classical inference, bayesian inference and the Dempster-Shafer's theory of evidence which is becoming of widespread use to combine various types of evidence. Other techniques used in data fusion include fuzzy logic and neural networks.

The applications of the above data fusion techniques to the CTBT problem lead to an analysis of some deficiencies of the current system and suggestions for improvements.

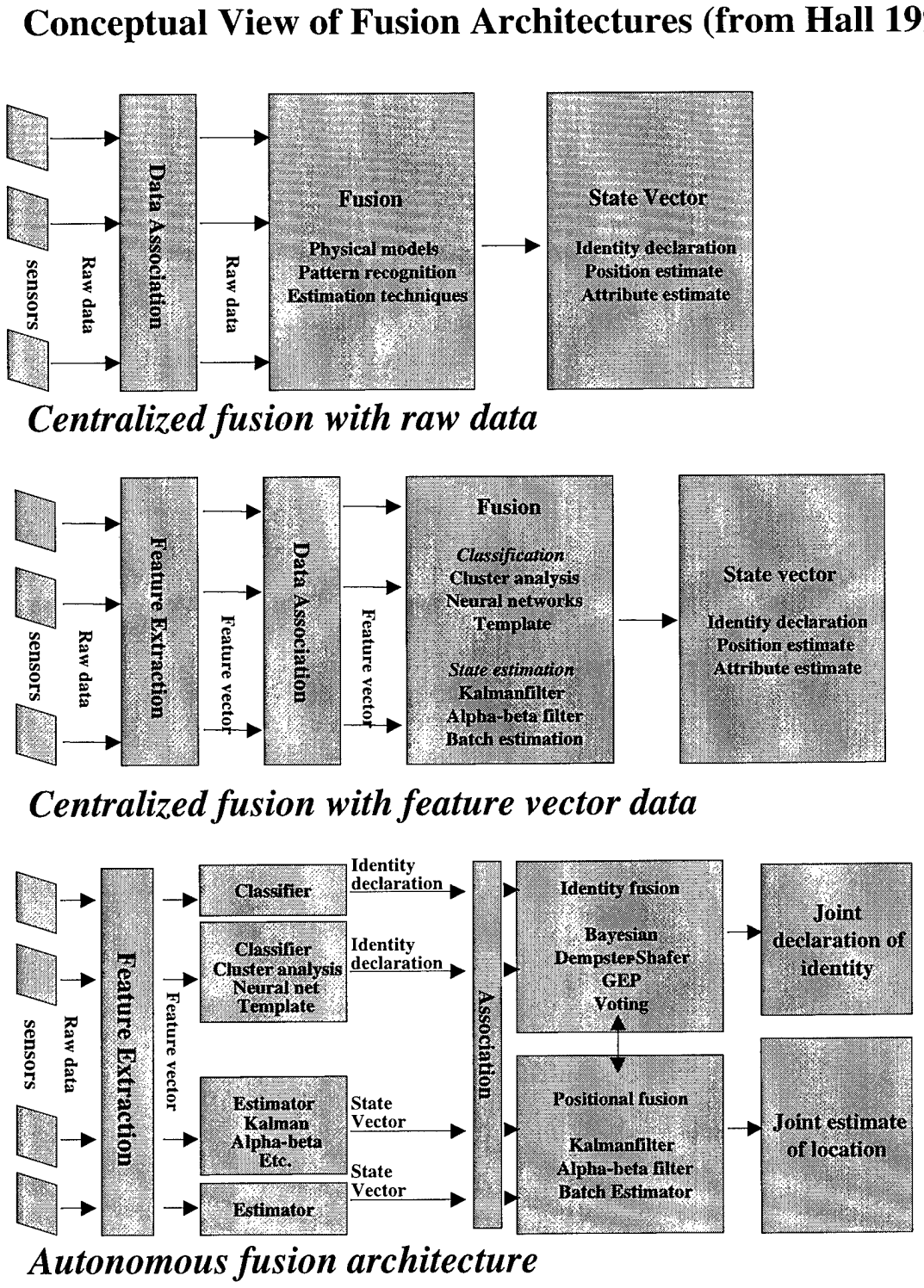

Figure 33. Three high-level architectures for multi-sensor data fusion systems (from *Hanson, et. al.*, 2000).

7.3 Knowledge representation of spatial information

One of the areas critical to the development of the next generation of monitoring systems is the area of spatial object representation and geographical information systems. We have outlined in our report the theory of spatial database and examined the commercial spatial database systems. We have also analyzed the trends in future research and development in this area, including spatio-temporal databases, and spatial data mining. The section on theory explains spatial concepts, spatial data models and structures, spatial data types and operations and access methods and performance. The section on commercial systems provides a market overview of existing spatial products and spatial databases (*Hanson et al.,* 2000).

7.4 Integration of single-sensor processing into the pIDC operations

Integration of a new processing unit or module into a large existing system involves the application of scientific and engineering knowledge. Any project has the prerequisite that the parties involved understand the existing system, including the operational model and the relationship of the new functionalities to the existing system. The engineering approach should be chosen and specified beforehand as well as acceptance criteria, deliverables and a model for maintenance.

The software infrastructure present at the pIDC includes the data management, distributed application control system, interprocess communication, global libraries and operational configuration. An understanding of each of these elements and how they fit with each other and how the new element will fit within the system is important.

Software standards to be used in the project should be specified. The pIDC offers support with several standard libraries such as a time handling library.

8. Conclusions

We implemented in pIDC operations a system that performs joint processing of data from the three waveform technologies; seismic, hydroacoustic and infrasonic. We have also developed an initial capability to associate radionuclide detections with events formed from the waveform technologies.

More work remains to be done to take full advantage of the acoustic data types. For instance, the hydroacoustic stations of the IMS are planned to consist of groups of multiple hydrophones. To take advantage of this configuration and enhance the usefulness of multi-site hydroacoustic stations, we have developed two new modules, one for automatic azimuth determination, the other for interactive review of the azimuth determination. We also enhanced the existing network processing module, GA, to use the hydroacoustic azimuth information. These will be in place with the PIDC_7.0 release and will need further evaluation in an operational context.

The hydroacoustic propagation model is relatively mature and complete. It takes into account blockage as well as path-dependent and seasonal variations in travel-time and attenuation. However, the infrasound model is still based upon a constant velocity in the current operational system. This results in large modeling errors leading to potentially erroneous associations. We have enhanced the infrasound propagation model with azimuthally and seasonally-varying travel times to allow for smaller travel-time modeling errors. This is planned for implementation in pIDC Release R3.

While a few cases of identifiable infrasound sources have been observed, an effort should be made to understand the events formed on the daily basis by the infrasound network and minimize the false alarm rate.

Case studies have proven quite useful in evaluating the pIDC processing software and more should be conducted to evaluate the planned enhancements. For example, we recommend further comparisons of ground truth infrasonic travel times with predictions from propagation models.

We developed, calibrated and used the simulation tool, `SynGen`, which has been used to simulate the planned IMS monitoring network, perform GA stress tests, and evaluate GA capability for a variety of evasion scenarios. We have also evaluated the operational performance of pIDC data fusion using seismic, hydroacoustic, and infrasonic data collected over several years.

The next generation of Multi Sensor Data Fusion system would benefit from a centralized system of spatial representation of model parameters, which would among other advantages simplify the handling of spatially dependent velocities for the hydroacoustic and infrasonic data types. It would also benefit from improvements to single-sensor processing algorithms. We have defined the prerequisites and issues associated with integrating new algorithms into the pIDC monitoring system. Finally, advances in artificial intelligence and data fusion methodologies could be beneficial for future enhancements to the pIDC monitoring systems. In particular, new techniques are likely to improve radionuclide fusion, event characterization and event screening.

9. Acknowledgments

This project was sponsored by DTRA under Contract No. DASG60-96-C-0151. We are grateful to Steve Bratt and Gus Gustafson for their guidance through the contract, and to the staff at CMR who assisted in designing, installing and evaluating the software developed under this project.

10. References

Brumbaugh, D. and R. J. Le Bras, Hydroacoustic Signals from a Ground-Truth Data Set of Marine Explosions, *SAIC Tech. Report, 98/3008*, 27pp., (1998).

Burger, M. D., C. E. Boucher, E. M. Daley, W. W. Renner, V. L. Pastor, L. C. Haines, and A. I. Eller. Acoustic System Performance Model ASPM 4.0A: User's Guide, *Tech. Rep. SAIC-94/1000*, Science Applications International Corporation, McLean, VA, (1994).

Dysart, P., H. Laney, R. Le Bras, and H. Freese. New Version of Hydroacoustic **StaPro**. Memorandum *CCB-PRO-97/26.* 1997.

Flanagan, M., R. Le Bras, J. Hanson and R. Jenkins, Analysis of Two Mining Explosions Recordings at the TXAR Seismic and Infrasound Arrays, *SAIC Tech. Report, 99/3002*, 25pp., (1999).

Gault, A. and D. Brown, Analysis of all infrasound phases recorded at the pIDC, *CMR 99/25*(1999).

Guern, J. and R.J. Le Bras, Requirements and Design for Seismic-Hydroacoustic-Infrasound and Radionuclide Potential Association (SHI-RPA), *SAIC Tech. Report, 99/3011, 15pp.,* (1999).

Gustafson, R., R. J. Le Bras, J. Hanson and Paul Dysart, Hydroacoustic Processing at the pIDC - Plans and Results, Informal Workshop on Hydroacoustic Technology for the CTBT, Papeete, Tahiti, French Polynesia, September 1999.

Hanson, J., R. Le Bras, P. Dysart, D. Brumbaugh, A. Gault and J. Guern, Operational Processing and Special Studies of Hydroacoustics at the Prototype International Data Center, P.A. Geophysics, in press (2000).

Hanson, J., Multi-Sensor Data Fusion and the CTBT, SAIC Tech. Report, 00/3022, (2000).

Jenkins, R., C. N. Katz, R. Le Bras, and T. Sereno. Ground-Truth Analysis of Explosions Recordings at LSAR and TXIAR Infrasound Arrays, *SAIC Tech. Report 98/3035*, 25pp., (1998).

Katz, C. N., D. J. Brown, A.K. Gault, R. J. LeBras, and J. Wang, PIDC 6.0: Implementation of infrasonic processing in PIDC operations, *PIDC CCB Memo*, April 03, 85 pp., (1998).

Laney, H., P. Dysart, H. Freese and R. Willemann, An automated system for detecting and classifying in-water explosions and T-phases, *J. Acoust. Soc. Am., 100, 2641,* (1996).

Laney, H., Hydroacoustic Feature Extractor (DFX-H Version 1). Technical Description. *CMR Tech. Rep. 97/06,* 1997.

Le Bras, R. and T. Sereno, Monitoring the Earth for Nuclear Testing in Underground and Underwater Environments, *Processings of the 1996 AGU Fall Meeting* (1996).

Le Bras, R. J., J. Guern, D. Brumbaugh, J, Hanson, and T. Sereno, Integration of Seismic, Hydroacoustic, Infrasound and Radionuclide Processing at the Prototype International Data Centre, in *Proceedings of the 21st Annual Seismic Research Symposium on Monitoring of a Comprehensive Nuclear-Test-Ban Treaty, (1999).*

McKisic, J. M., Infrasound and the infrasonic monitoring of atmospheric nuclear explosions: A Literature Review, *Tech. Rep. PL-TR-97-2123*, 320 pp., (1997).

Sereno, T. J., Multi-Sensor Data Fusion to Support Verification of a Comprehensive Nuclear-Test

Ban Treaty (CTBT) - Program Plan, Tech. Rep., SAIC-96/1176, 36 pp., (1996).

Sereno, T. J. and R. Le Bras, Multi-Sensor Data Fusion to Support Verification of a Comprehensive Nuclear-Test Ban Treaty (CTBT) - Revised Program Plan, Tech. Rep., SAIC-98/3020, 23 pp., (1998).

Smith, W. H. F. and D. T. Sandwell, Bathymetric prediction from dense altimetry and sparse shipboard bathymetry, *J. Geophys. Res.*, 99, 21803-21824, (1994).

Sorrells, G. G., E. T. Herrin, and J. L. Bonner, Construction of regional ground-truth databases using seismic and infrasound data, *Seism. Res. Lett., vol. 68, no. 5,* 743-752, (1997).

Urick, R. Principles of Underwater Sound. *McGraw-Hill Book Co., 3rd ed. (*1983).

Velasco, A. and D. Brumbaugh. **SynGen** - Synthetic Event and Detection Generator Design Document. *Tech. Rep. SAIC-98/3007.* Science Applications International Corporation, San Diego, CA, (1998).

Willeman, R., H. Laney, P. Dysart and H. Freese, A global monitoring system for a nuclear test ban treaty, *J. Acoust. Soc. Am., 100, 2642,* (1996).