LOCALIZATION OF SUB-SEA EARTHQUAKES USING HYDROACOUSTIC REFLECTIONS AND MATCHED-FIELD PROCESSING

Zachary M. Upton and Jay J. Pulli

BBN Technologies

Sponsored by National Nuclear Security Administration
Office of Nonproliferation Research and Engineering
Office of Defense Nuclear Nonproliferation

Contract No. DE-AC03-02SF22560

ABSTRACT

The purpose of this study is to investigate the localization of long-range hydroacoustic events, such as sub-sea earthquakes and underwater explosions, received at a single monitoring station. A matched-field technique is used to predict synthetic waveforms that would be received at a hydroacoustic station for a given source location. A grid of potential source locations is selected based on the areas of natural seismicity defined by the mid-ocean ridges in the Indian Ocean and the subduction zones to the east (Carlsberg Ridge, Mid-Indian Ridge, Southeast & Southwest Indian Ocean Ridges, and Java Trench). This matched-field technique makes use of a previously designed propagation model (Pulli, *et al*, 2000) along with a detailed knowledge of bathymetry to predict arrivals at the receiver including reflections from bathymetric features. An earthquake detected at the Diego Garcia hydroacoustic station in the Indian Ocean, for which bathymetric reflectors have already been identified (Harben and Boro, 2001), is used to validate the synthetic waveform algorithm. Earthquakes in the Indian Ocean are then localized using the matched-field algorithm and signals recorded at Diego Garcia. These localizations are used to study the sensitivity of the matched-field algorithm. This research shows that hydroacoustic events can be localized using data from single stations if reflections from bathymetric features are included.

OBJECTIVE

The objective of this research is to assess the feasibility of using matched-field techniques to localize sub-sea earthquakes and underwater explosions using a single monitoring station and hydroacoustic reflections. Previous research has shown that reflected arrivals from an explosion or earthquake can be predicted for a given source-receiver pair. This study seeks to extend that work by creating a field of modeled waveforms and comparing them to the measured waveform using matched-field techniques.

The value of single station localization using hydroacoustic reflections is that it has the potential to extend the detection and localization coverage of the International Monitoring System (IMS) hydroacoustic network. Acoustic energy from a strategically placed explosion might have a direct acoustic path to only one IMS station. Worse, the energy might have no direct acoustic path. However, if reflections can be used as a basis for detection and/or localization, areas that were once not covered by the IMS hydroacoustic network could be considered within its coverage.

RESEARCH ACCOMPLISHED

Validation of reflector model in the Indian Ocean

The first step in this study was to validate past research on reflections for use in the Indian Ocean. Pulli, *et. al*. (2000) designed a model that, for a given source-receiver pair, calculated potential bathymetric reflectors in an ocean basin. The model used these potential reflectors to form the impulse response of the ocean basin. By convolving this impulse response with a synthetic source function, the model was able to predict a received waveform envelope based on the direct arrival and bathymetric reflections. Pulli *et. al*. (2000) then used this model to predict waveforms received at stations in the Atlantic and Pacific Oceans.

The Indian Ocean is chosen as the ocean basin for this study because of the IMS station HA08 located near Diego Garcia. This station not only provides a consistent, reliable source of hydroacoustic data, but it is the first IMS hydroacoustic station to begin to transmit data to the International Data Centre (Vienna, Austria) and Prototype International Data Center (Virginia, USA). Earthquakes regularly occur in the Indian Ocean and are a good source of acoustic information. Analyses of the signals from some of these earthquakes show what appear to be reflected arrivals (Harben and Boro, (2001); Hansen (2001).

In a paper written in 2001, Harben and Boro study a magnitude 6.8 earthquake that occurred January 18, 2000, near Jawa, Indonesia. Figure 1 is a map depicting the location of this earthquake, determined using seismic techniques, with respect to the Diego Garcia IMS station. The waveform envelope of this earthquake, recorded on the north array at Diego Garcia, is shown in the top graph of Figure 2. To validate the reflections model for use in the Indian Ocean, the model was run using the origin of the earthquake as the source location and a nominal, centralized location for the northern array of the Diego Garcia IMS station as the receiver position. The bottom graph in Figure 2 shows the synthetic waveform result of the model calculations.

Note that the major reflections at approximately 55 minutes and 59 minutes are identified by the synthetic waveform. Harben and Boro identify a reflector at the Seychelles/Mauritius Plateau that corresponds to the reflections at approximately 55 minutes. There is a peak in the envelope that is not in the replica waveform (at ~42 minutes). This peak could be the result of some non-reflected energy, like a low-level aftershock. It could also simply be a deficiency in the model that caused the discrepancy between the real and synthesized waveforms. These results are comparable to those in the Atlantic and Pacific Oceans, and therefore validate the use of the reflections model, and its resulting synthetic waveforms, in the Indian Ocean Basin.

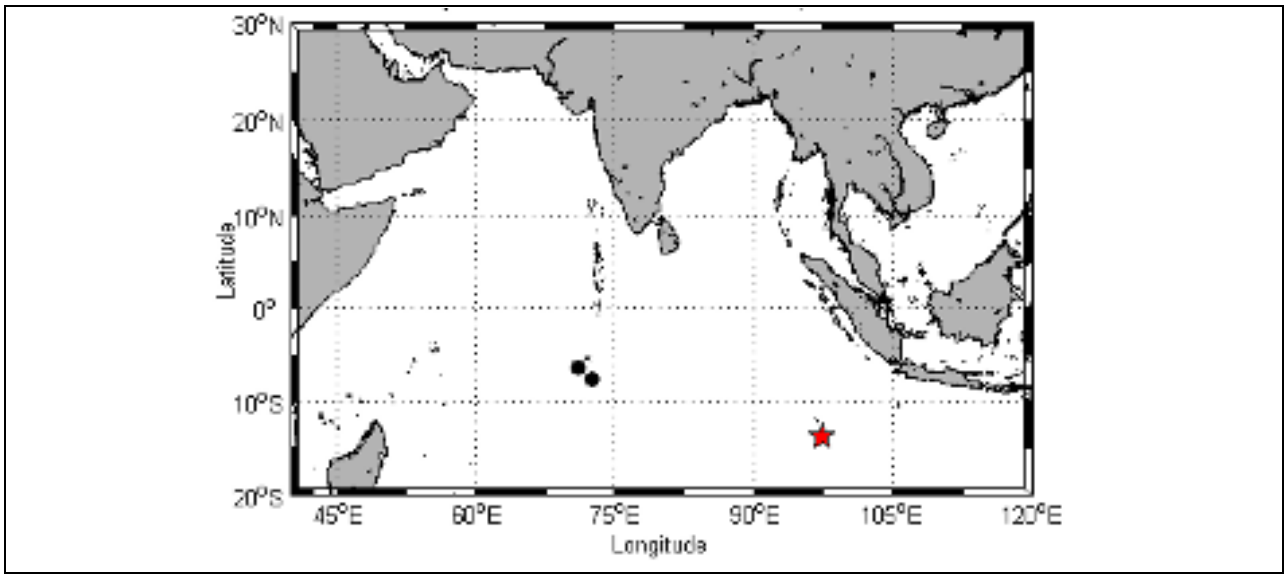

Figure 1. Location of the January 18, 2001, earthquake near Jawa, Indonesia (red star). The Diego Garcia IMS station is also shown in black circles.

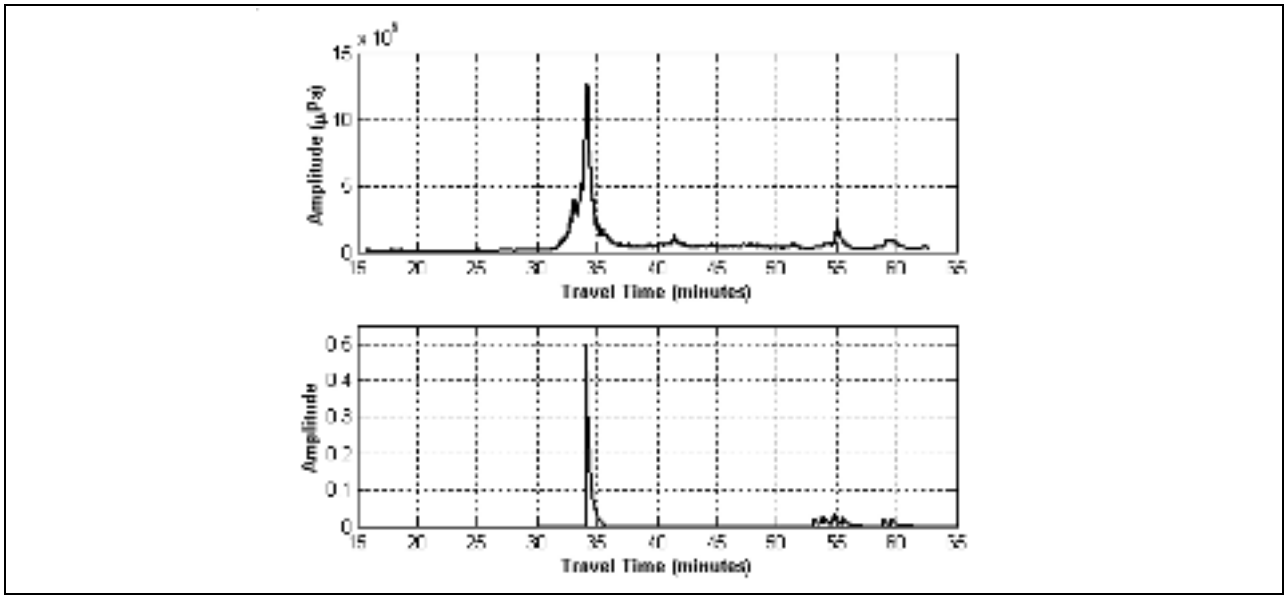

Figure 2. Real received waveform envelope from Diego Garcia North (top) and the output of the reflections model, the synthetic waveform (bottom).

Indian Ocean seismicity grid

The first step in the matched-field localization process was to generate a geographically disparate set of source locations. To attempt matched-field localization on an ocean-size scale, one might begin with a coarse grid of the entire ocean. However, time, measured data, and computer processing power were limiting factors in this research. There are no available recordings of explosions in the Indian Ocean. However, there are recordings of earthquakes. These earthquakes occur in a relatively limited number of geographic areas under the ocean (i.e. along mid-ocean ridges). Earthquakes, therefore, are an excellent source to test the methodology.

To minimize the time and processing power spent calculating synthetic waveforms, a series of source locations was selected that roughly matched the areas of peak seismic activity in the Indian Ocean. Figure 3 shows the map of

seismic activity in the Indian Ocean for 2000 and 2001 at left and the matched-field replica locations at right. The reflections model was used to create synthetic waveforms for each replica location.

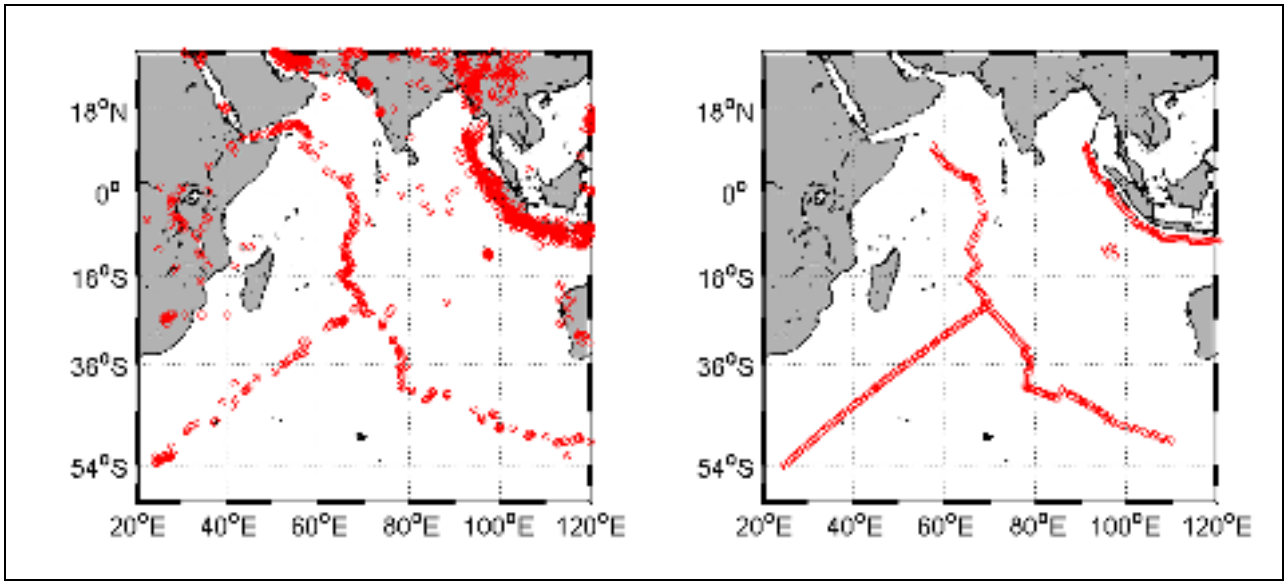

Figure 3. Map of seismic event locations for 2000-2001 (United States Geological Survey) at left, matched field replica locations at right.

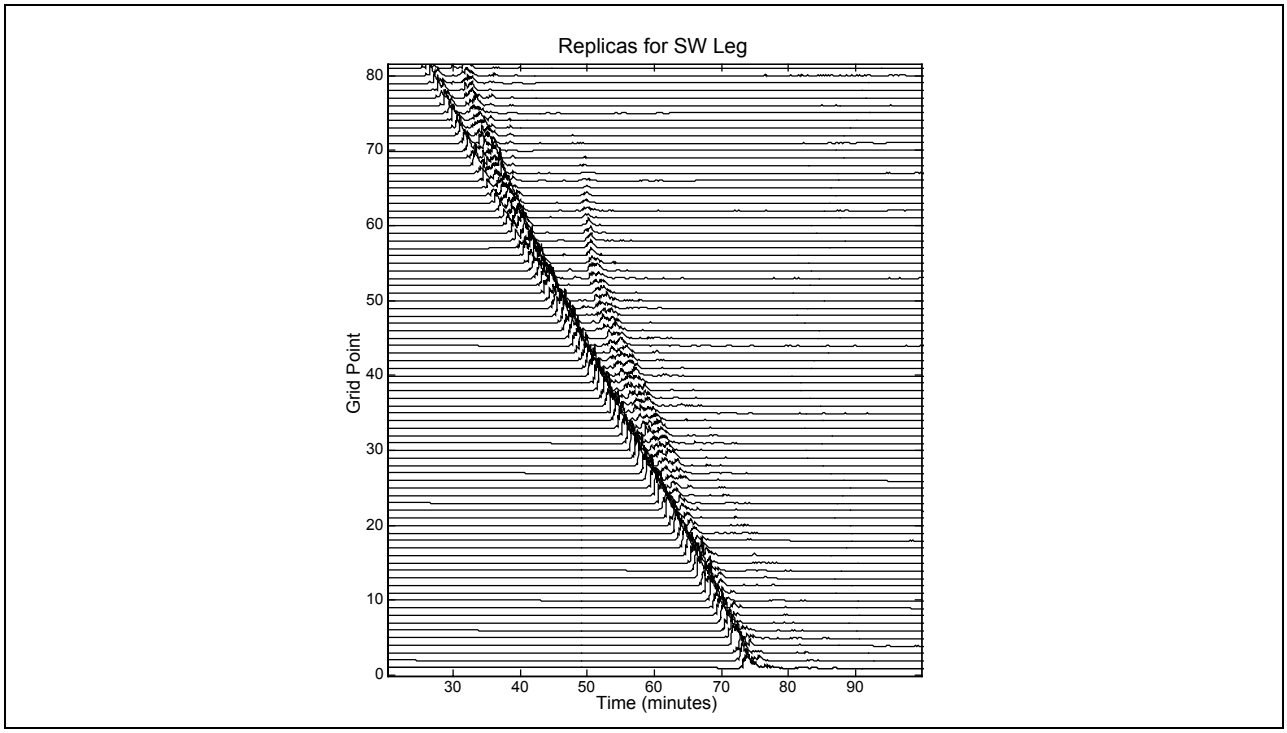

Figure 4. Synthetic waveforms for the Southwest leg of the source replica locations.

Localization Algorithm

A matched-field algorithm was developed to efficiently and effectively localize the earthquake signals and to display the results in a manner that was easily understood. Comparing 300+ replicas to over an hour of signal could be a very time consuming task. Since the only parts of the waveform that are of interest in the localization are the direct

and reflected arrivals, it makes sense to only correlate around them and not over the entire waveform. The real measured envelope is analyzed first. A noise estimate is calculated in 5-sec windows across the envelope. Since the noise estimate is only to be used as a basis for thresholding, it need not be terribly precise. Therefore, in each 5-sec window after the direct arrival time, the slope of the envelope is calculated using the beginning and end samples. Windows that have a slope greater than a prescribed maximum slope are considered signal. Windows with a slope less than a prescribed maximum slope are considered noise. An average value of noise is calculated within each of the noise windows. To detect the direct and reflected arrivals, a threshold is set above the calculated noise estimate. The points at which the waveform is above the threshold are considered direct or reflected arrivals and the rest of the waveform is set to zero. Figure 5 shows an example of the result of the noise estimate and detection. The top graph shows the original waveform. The bottom graph shows the detections in the thresholded waveform. These arrivals are found using a threshold of 6 dB above the noise estimate.

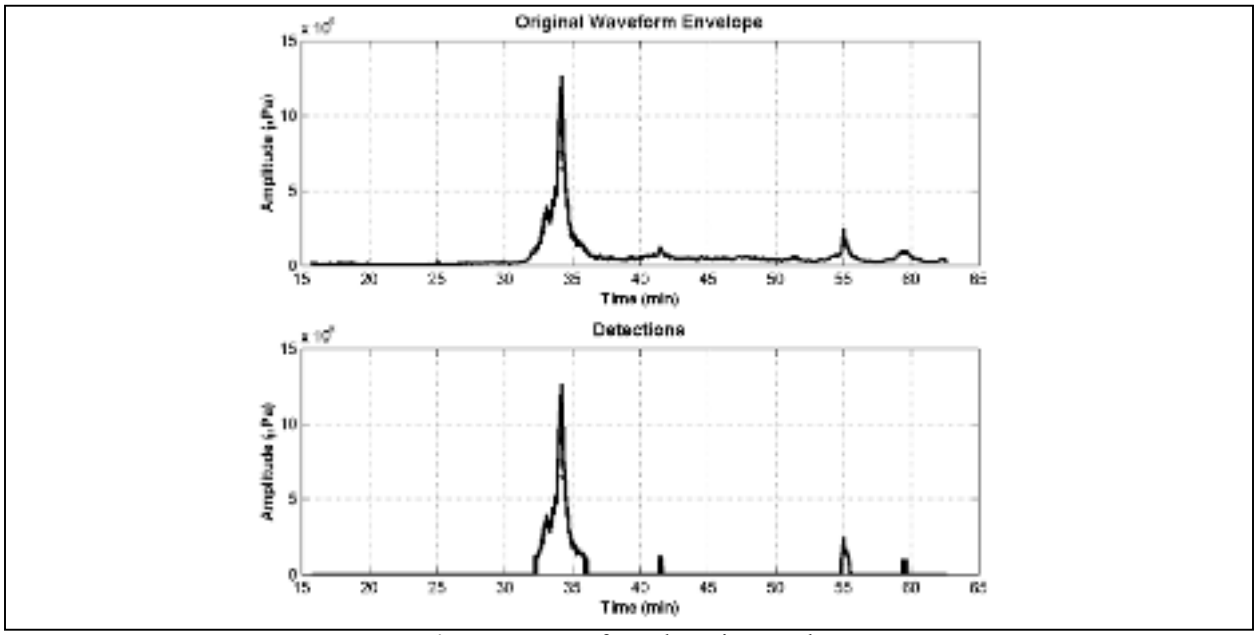

Figure 5. Waveform detection results.

To complete the matched-field localization, the replica waveforms are correlated with this thresholded waveform. However, in this instance and many others, the direct arrival is much stronger than the reflected ones. Each detection area (in time) is normalized by dividing by its maximum to consider the direct and reflected arrivals equally. This operation is performed on both the measured and the replica waveforms.

For the last piece of pre-processing before the actual localization, it is necessary to make sure that the direct arrival in the thresholded waveform would be correlated with the first arrival in the replica waveforms. This eliminates any large correlation between the direct arrival of the thresholded waveform and a reflected arrival in a replica. To achieve this goal, the replica is tested. If the time of the first arrival in the replica does not match the first arrival in the real waveform, the replica is set to zero around the time of the real direct arrival, thus eliminating the chance of any undesired high correlation. Reflected arrivals are correlated simply by their occurrence in time.

To correlate multiple replicas with the thresholded waveform, a loop is set up that cross-correlates the replica waveforms with the raw waveform during the detection time periods. Cross-correlations are performed in 5-sec windows. A replica waveform that is a strong match to the raw waveform in a particular time window has a high cross-correlation at the zero lag time. To determine the relative localization strength of a particular replica waveform, the zero-lag correlation is summed over the correlated waveform segments. Therefore, the replica with the highest summed zero-lag correlation over the direct and reflected arrivals is the strongest localization.

Display of results

Typically, matched-field localization results are displayed as ambiguity surfaces that allow the viewer to pick the most likely localization (for example, in range and azimuth). However, in this case, the potential source positions are located on an irregular grid over the areas of frequent seismicity in the Indian Ocean. The results of the localization are displayed on a map similar to that of Figure 3. However, in this case, the source locations are color-coded by their respective summed zero-lag correlation, with the blue end of the color spectrum being low correlation level and the red end being high correlation level. In this way an ambiguity 'line' is formed analogous to the traditional ambiguity surface. Jensen, *et al* (1994) provide more information on ambiguity surfaces.

Localization Case 1 – The January 18, 2000, Jawa, Indonesia, earthquake revisited

The localization algorithm was tested first using the Jawa, Indonesia, earthquake described in Harben and Boro, 2001. A map of the seismic (United States Geological Survey -- USGS) location of the earthquake was shown in Figure 1. The waveform envelope of this earthquake was shown in Figure 2. The detections have already been shown in Figure 2. The output of the localization algorithm is shown in Figure 6. The results displayed in Figure 6 show that the strongest localization strength, or the highest overall correlation occurred at the only red dot on the map. Table 1 is a comparison between the ground truth of the earthquake and the calculated location.

Table 1. Ground Truth Location vs. Calculated Location – Jawa, Indonesia, Earthquake

Location Type	Latitude	Longitude
Ground Truth Source Location (USGS)	13.802S	97.453E
Calculated Location	13.80S	97.45E

The results of this localization demonstrate that, when there is a grid point in close proximity to the event location, this algorithm can localize the event accurately. The distance between the ground truth location and the calculated location is 0.4 km. What is interesting in this result is that there is only one point in the area whose relative correlation level comes near that of the calculated location. Its correlation level, however, is only about two-thirds of that of the localization point. Therefore, based on only the initial test, the algorithm is fairly sensitive to location.

Figure 7 compares the replica waveform at the calculated location with the replica at the point that is nearest to it. These waveforms illustrate the difference in response that occurs with changes in location. The two points are 150 km apart.

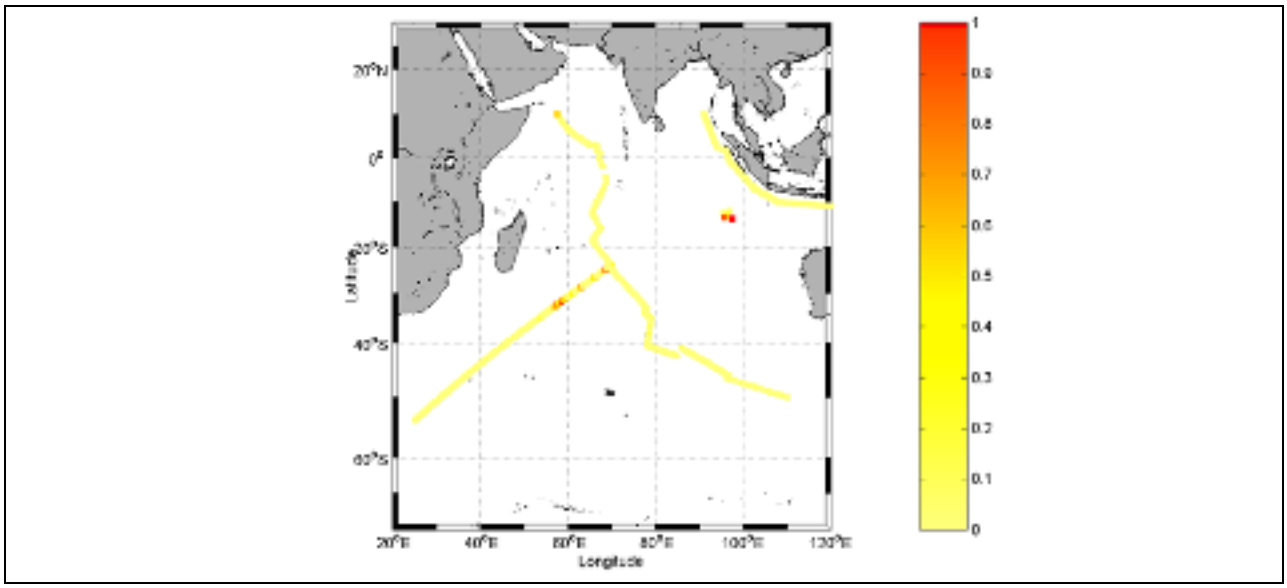

Figure 6. Visualization of the localization of the January 18, 2000, earthquake near Jawa, Indonesia.

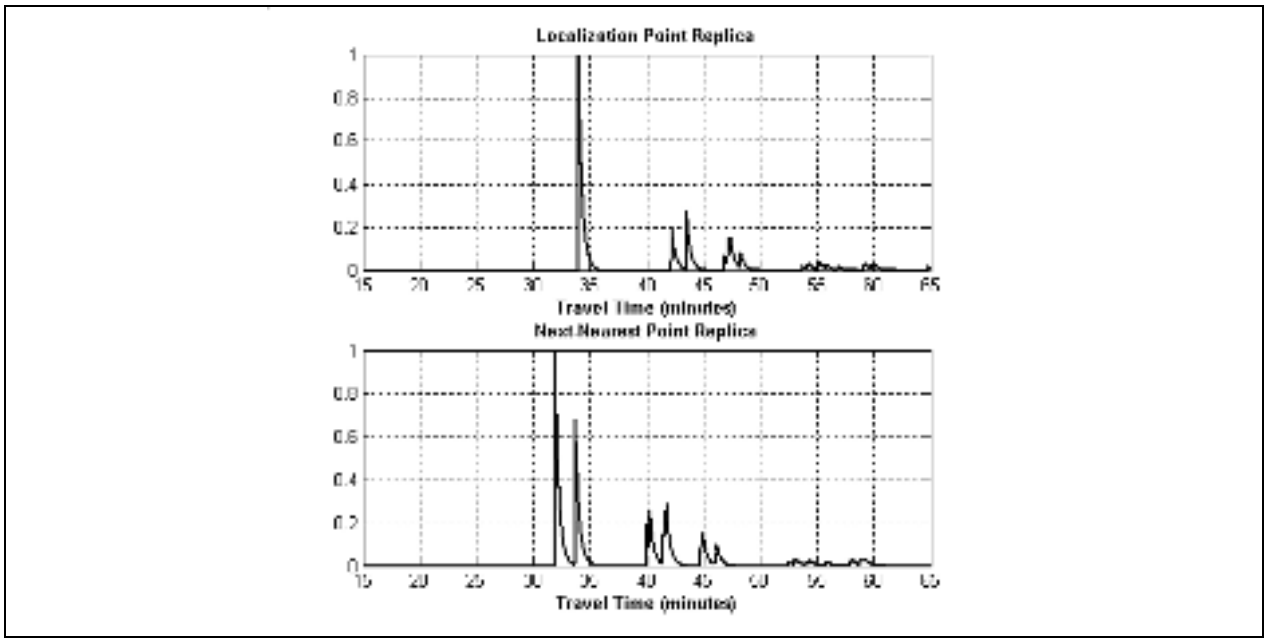

Figure 7. Comparison of the localization point replica with the next-nearest (geographically) replica.

Localization Case 2 – The April 19, 2001, Mid-Indian Ridge Earthquake

This 4.5-m_b earthquake is another event that exhibits direct and reflected arrivals suitable for study and attempted localization. The earthquake's ground-truth location and its relationship to Diego Garcia are shown in Figure 8. In contrast to the first case, there is not a replica source point within 50 km of the actual event location. The mid-ocean ridges are wide in comparison to the discrete grid of replica source locations, and this earthquake occurred away from the grid.

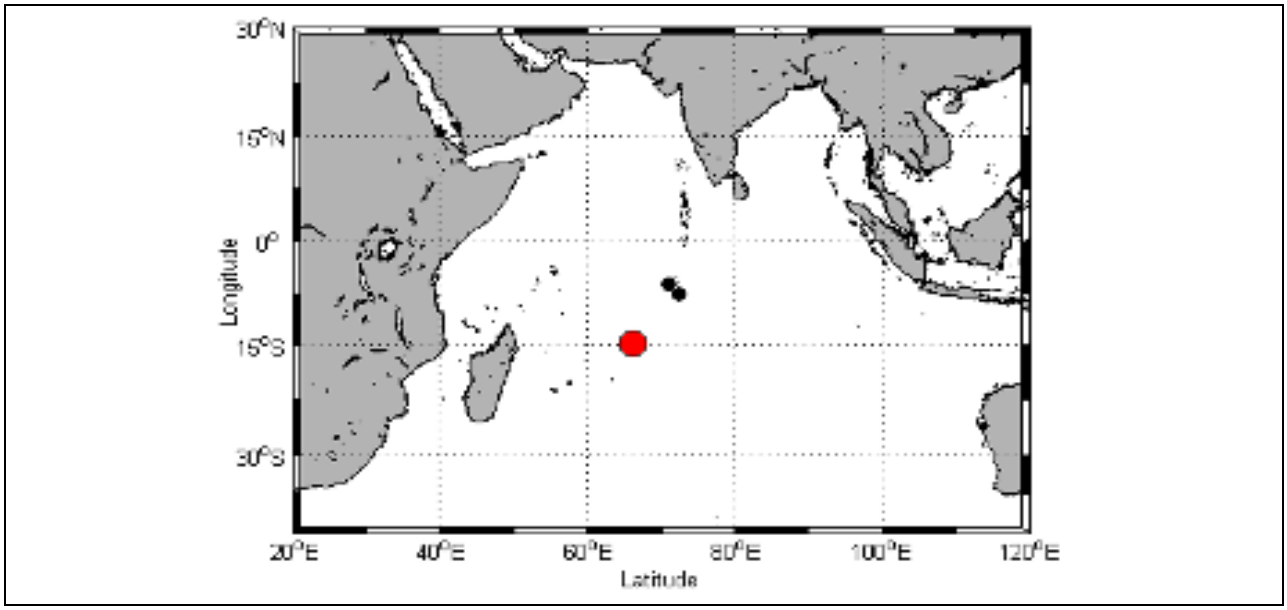

Figure 8. Location of the April 19, 2001, Mid-Indian Ridge Earthquake (red dot) and the Diego Garcia IMS Stations (black dots).

The earthquake's signal envelope, as recorded at the northern Diego Garcia array, is shown in Figure 9. The direct arrival (at ~12 minutes) and the reflected arrivals at 27 and 30 minutes, respectively, are immediately apparent. The localization algorithm output is shown in Figure 10. The comparison between the ground truth location and computed location is shown in Table 2.

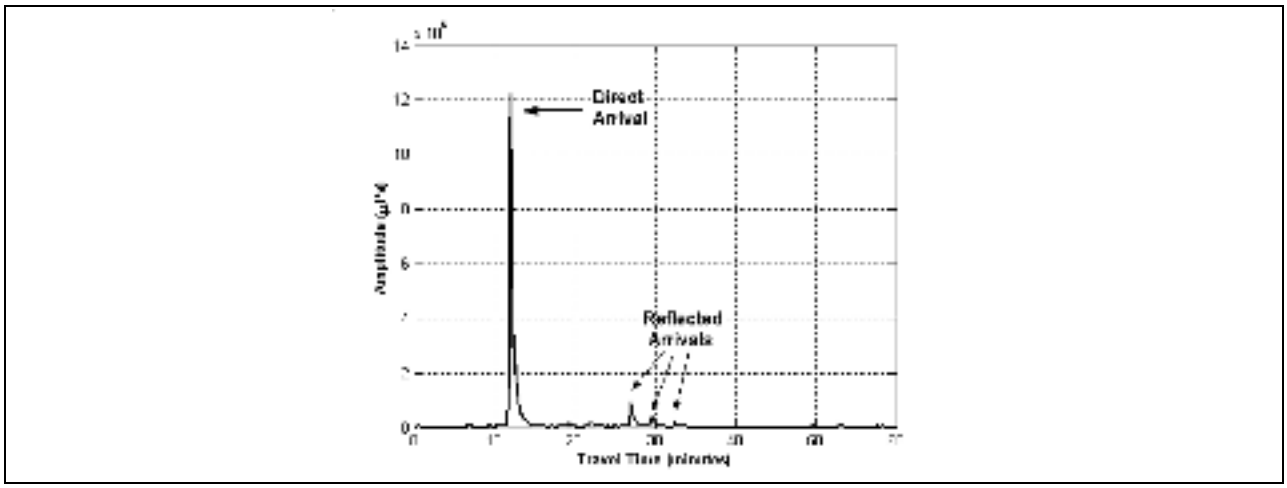

Figure 9. Waveform envelope of the April 19, 2001, Mid-Indian Ridge Earthquake as recorded at DGN.

Table 2. Ground truth location vs. localization location – Mid-Indian Ridge Earthquake

Location Type	Latitude	Longitude
Ground Truth Source Location (USGS)	14.63S	66.21E
Localization Location (Northernmost)	15.13S	67E
Localization Location (Southernmost)	23.67S	69.5E

The analysis of this localization is somewhat more difficult than the first event. The localization output shows two red boxes of nearly equal relative localization strength. The northernmost red box is near the earthquake location (101 km). The southern localization is much further from the earthquake location (1058 km).

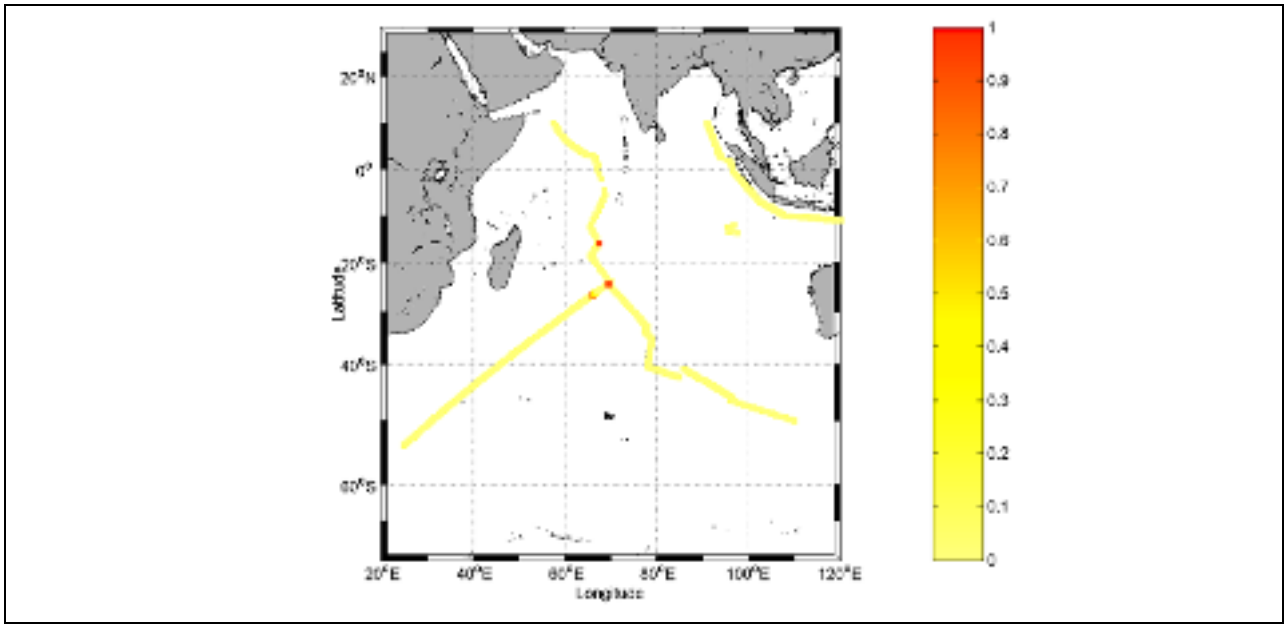

Figure 10 . Localization Algorithm Output.

Many factors could contribute to this result. First, there is not a grid point in close proximity to the earthquake. In the first example, it was shown that the algorithm is very sensitive to location. For an earthquake localization in which the event location does have a corresponding grid point, the surrounding points could be much weaker in correlation level. Therefore, the fact that the localization algorithm produced two localizations in this case is not a surprise. Second, the localization algorithm, as designed, eliminates the amplitude discrimination capability of the cross-correlation function. By normalizing all the arrivals (direct and reflected) by their respective maximum value, the direct and reflected arrivals were weighted equally in terms of the correlation. However, in doing so, the relative strength of the arrivals is masked, since each arrival will have a peak value of one.

CONCLUSIONS AND RECCOMENDATIONS

In this study, it has been shown that long-range hydroacoustic reflections can be used as a basis for the creation of replica waveforms, at a specific receiver location, for a geographically disparate set of source locations. It has also been shown that these replica waveforms can be used in a matched-field localization algorithm to determine the source location of an acoustic signal.

The localization algorithm exhibited sensitivity to the location of the replica positions. If there was a grid point near the event location, the event could be accurately localized. Also, nearby grid points did not exhibit spurious high localization strengths that might damage the accuracy of localization. However, in the case of an event that did not have a nearby replica grid point, the algorithm produced a dual localization. One of these localizations was a true localization, in that the grid point selected was near the event location. However, the other localization point was over 1000 km away from the event location. The two results, though very different, yield some similar conclusions. This algorithm is capable of localizing sources on an ocean-basin-sized scale. However, for the algorithm to make accurate localizations anywhere in the ocean, the grid of potential source locations would have to be fairly dense. The density of the grid, in fact, would probably have to be correlated to the desired localization accuracy.

Though the results of this study are very encouraging, there are many areas of research that could improve the capability of localizing events using long-range hydroacoustic reflections. The most obvious is to extend the calculations here to a grid of an entire ocean. A whole-ocean case would validate this method for use in the IMS. The localization algorithm and reflections models exhibit a few weaknesses that are worthy of further study. For instance, the localization algorithm deliberately eliminates the amplitude discrimination power of cross-correlation. A limited amount of change to the pre-correlation processing could allow the localization algorithm to take full advantage of this characteristic.

A brief list of other possible future study areas follows:

- A study of reflector strength calculations. In the current synthetic waveform model, the amplitudes of the arrivals do not match well to the data. Including a better calculation of reflector area or a better impulse response weighting function might improve the match between synthetic and reflected data.
- A study on the use of higher resolution bathymetry to develop a better measure of bathymetric slope in the reflections model. In what direction should slope be calculated? How can the directional gradient be determined? This study could be performed using higher resolution topographic databases to obtain an accurate, yet unclassified result.
- An investigation of the use of different source functions for convolution with the ocean basin impulse response model in the reflections model. Is there a better source function that is both more representative of the event in question and provides better synthetic replicas?
- An investigation of the use of more than zero-lag correlation. Is there a method that will yield better localizations?
- An investigation of the use of back-azimuth calculations in conjunction with travel time to improve the reflections model.

Finally, HydroCAM, the modeling software used to perform the travel-time calculations for this study, assumes a source in the water column when it makes its travel-time calculations. Earthquakes occur many kilometers below the ocean floor. For an earthquake that takes place 30 km under the ocean floor, seismic waves must travel in earth for about 30 km before their energy is converted to sound. Assuming a seismic velocity of approximately 6 km/sec, these waves will take about 5 sec to reach the earth/water interface. The best modeling of earthquakes that could be accomplished in this research was to set up the source on the bottom of the ocean. In the scope of calculations that span hours, the 5-sec bias does not seem to hamper localization, but a final topic of future study would be to analyze the benefit of adding a few seconds to the travel-time measurements to account for this bias.

ACKNOWLEDGEMENTS

The initial research and development of the reflections model for this study was performed under Defense Threat Reduction Agency (DTRA) contract number DSWA 01-97-C-0164. Additional support was provided through internal funding at BBN Technologies. We thank Phil Harben of Lawrence Livermore National Laboratory for providing data and insight during the course of this work. We also thank Dr. Anthony Lyons of Penn State University for useful discussions and suggestions.

REFERENCES

Angell, J., T. Farrell, and J. Pulli (1998), Characterization of reflected hydroacoustic signals, Proc. 20th Annual Seismic Research Symposium, 650-659.

Bendat, Julius S. and Allen G. Piersol. *Engineering Applications of Correlation and Spectral Analysis*. Second Edition, John Wiley and Sons, Inc. 1993.

Hanson, J. (2001), Initial Analysis of Data from the New Diego Garcia Hydroacoustic Station, 23rd Seismic Research Review, Jackson Hole, WY, LA-UR-01-4454.

Harben, P. and C. Boro. (2001), Implosion Source Development and Diego Garcia Reflections, Proc. 23rd Seismic Research Review, Jackson Hole, WY, LA-UR-01-4454.

Jensen, Finn B, W.A. Kuperman, M.B. Porter, and H. Schmidt. *Computational Ocean Acoustics*. American Institute of Physics. 1994

Pulli, J.J., Z. Upton, R. Gibson, T. Farrell (2000), Modeling Long-Range Hydroacoustic Reflections in the Atlantic and Pacific Oceans, Proceedings of the 22nd Seismic Research Symposium, New Orleans, LA, Defense Threat Reduction Agency, OMB No. 0704-0188.

Pulli, J.J., T. Farrell, and R. Gibson (1999), Characterization and utilization of hydroacoustic signals reflected from continents and bathymetric features, 21st Seismic Research Symposium, Las Vegas, NV, LA-UR-99-4700.