

USING THE DOE KNOWLEDGE BASE FOR SPECIAL EVENT ANALYSIS

Christopher J. Young, J. Mark Harris, Hillary M. Armstrong

Sandia National Laboratories

Aaron A. Velasco, Julio Aguilar-Chang

Los Alamos National Laboratory

Sponsored by U.S. Department of Energy
Office of Nonproliferation and National Security
Office of Research and Development
Contract No. DE-AC04-94AL85000

ABSTRACT:

The DOE Knowledge Base is a library of detailed information whose purpose is to support the United States National Data Center (USNDC) in its mission to monitor compliance with the Comprehensive Test Ban Treaty (CTBT). One of the important tasks which the USNDC must accomplish is to periodically perform detailed analysis of events of high interest, so-called "Special Events", to provide the national authority with information needed to make policy decisions. In this paper we investigate some possible uses of the Knowledge Base for Special Event Analysis (SEA), and make recommendations for improving Knowledge Base support for SEA.

To analyze an event in detail, there are two basic types of data which must be used: sensor-derived data (waveforms, arrivals, events, etc.) and regionalized contextual data (known sources, geological characteristics, etc.). Currently there is no single package which can provide full access to both types of data, so for our study we use a separate package for each: MatSeis, the Sandia Labs-developed MATLAB-based seismic analysis package, for waveform data analysis, and ArcView, an ESRI product, for contextual data analysis. Both packages are well-suited to prototyping because they provide a rich set of currently available functionality and yet are also flexible and easily extensible.

Using these tools and Phase I Knowledge Base data sets, we show how the Knowledge Base can improve both the speed and the quality of SEA. Empirically-derived interpolated correction information can be accessed to improve both location estimates and associated error estimates. This information can in turn be used to identify any known nearby sources (e.g. mines, volcanos), which may then trigger specialized processing of the sensor data. Based on the location estimate, preferred magnitude formulas and discriminants can be retrieved, and any known blockages can be identified to prevent miscalculations. Relevant historic events can be identified either by spatial proximity searches or through waveform correlation processing. The locations and waveforms of these events can then be made available for side-by-side comparison and processing. If synthetic modeling is thought to be warranted, a wide variety of relevant contextual information (e.g. crustal thickness and layering, seismic velocities, attenuation factors) can be retrieved and sent to the appropriate applications. Once formed, the synthetics can then be brought in for side-by-side comparison and further processing.

Based on our study, we make two general recommendations. First, proper inter-process communication between sensor data analysis software and contextual data analysis software should be developed. Second, some of the Knowledge Base data sets should be prioritized or winnowed to streamline comparison with observed quantities.

Key Words: knowledge base, special event analysis, arcview, matseis

Sandia is a multiprogram laboratory
operated by Sandia Corporation, a
Lockheed Martin Company, for the
United States Department of Energy
under contract DE-AC04-94AL85000.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible electronic image products. Images are produced from the best available original document.

OBJECTIVES

In this paper we investigate the use of the Knowledge Base for Special Event Analysis (SEA) for seismic signals. SEA involves the use and analysis of previously collected data. The analysis can be varied since it will depend on the objectives needed for a special event. The data can include earthquake catalog information, reference events (i.e., identified events from mines, nuclear explosions, etc.), digital waveform data, and the use of contextual geographical information system (GIS) data. For this investigation, we use the MATLAB-based product MatSeis to demonstrate accessibility to the Knowledge Base for catalog, reference event, and waveform data, while we use the ESRI product ArcView for accessing GIS information. Both are well-suited to this study because they are flexible, extensible, and provide a tremendous variety of potentially relevant tools. Our objectives are to test the utility of the first generation of the Knowledge Base for SEA, and to develop a suite of SEA tools which can be used by the USNDC for SEA.

INTRODUCTION

It is anticipated that when operational, the USNDC will routinely process on the order of two hundred events per day in order to fulfill its mission of monitoring the CTBT. The vast majority of these events will be quickly determined to be irrelevant to that mission and dismissed with little if any additional processing. Occasionally, however, an event which could potentially indicate a treaty violation will be found and every effort must then be made to characterize that event as completely and quickly as possible. This process of intense scrutiny of a particular event is referred to as SEA.

SEA is a much different process than the event analysis that goes on as part of the routine monitoring at the USNDC. Routine event analysis has a clearly defined set of procedures which will be followed with very little variation regardless of the particulars of the event being analyzed. Automatic results are checked first and if they are found to be sufficiently accurate, the location and magnitude of the event can be used to determine whether or not additional processing is warranted. If obvious errors are found, they are corrected (e.g. arrivals are re-timed) and the location and magnitude are re-calculated, then the determination of a need for additional processing is made.

For SEA, the processing can start out similarly but quickly diverge. Again, the analyst will evaluate the results of the previous processing, correcting any obvious errors, but then procedures become much less predictable. For example, if the event was small, had poor station coverage, and was near a known test site, obtaining a precise location might become the primary concern. To do this, great pains might be taken to re-pick arrivals and add any secondary arrivals which might be available. Regional or local velocity models might be used to relocate the event. Additional data from other stations not initially available might be brought in and processed. If well-characterized previous events are available for the source area, these events might be used for master-event re-location or for waveform correlation. This procedure would differ if, on the other hand, the event was well-located (i.e., an announced nuclear test), but the size was poorly determined. The critical concern would be the yield of the event instead of event location. The SEA would involve researching the appropriate magnitude-yield formula for the region, which might involve transporting a well-calibrated formula or developing a new theoretical formula. In either case, the analyst must use whatever means are necessary to arrive at an acceptable characterization of an event. One common aspect of SEA, however, is the need for detailed regional information to characterize the event. This is the type of information that will be provided by the DOE Knowledge Base.

DATA SET

In this paper we use a portion of the China data set developed by Los Alamos National Laboratory and delivered to Sandia National Laboratories in June of 1998 (Velasco, 1998). This data set consists of several catalogs of events, a set of reference events with waveform segments, and a large set of contextual geographic information (geologic units, depth to Moho, etc.).

RESEARCH ACCOMPLISHED

Event Location with Improved Path Correction Information

Knowledge Base information can be used in several ways to improve event location. In this study we examine the use of travel-time correction information to improve the performance of the USNDC location program, EvLoc. Like many event location codes, EvLoc locates an event by minimizing the misfit between observed and predicted travel times (e.g., Bratt and Bache, 1988). The predicted travel times come from a one-dimensional model (e.g. IASP91) plus a series of corrections to account for the differences between the model and the true Earth: an ellipticity correction, an elevation correction, a static station correction, and a path correction. All of these are currently available in Evloc. Path corrections, however, have suffered from overly-smoothed representations and poorly-characterized error information. The Knowledge Base offers more precise path correction information and provides robust error estimates. A detailed discussion of how travel time correction information is processed for, stored in, and retrieved from the Knowledge Base is given in the presentation by Young et al. (1998).

To facilitate the use of Knowledge Base travel time correction information, we developed a MatSeis locator interface (Figure 1). This interface provides access to all of the current locator parameters as well as new parameters

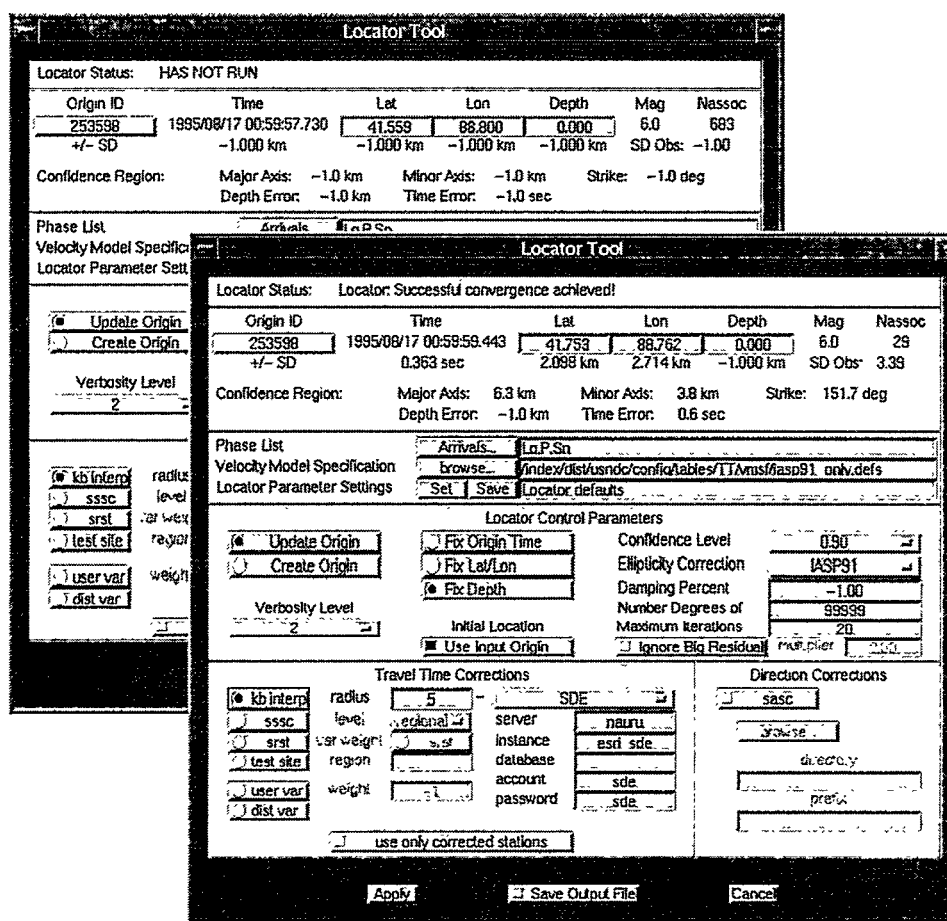


FIGURE 1. MatSeis locator interface.

needed for using the Knowledge Base. Figure 2 shows an event from China located with and without Knowledge Base correction information. In this case we know the true location, so it is apparent that the use of the Knowledge

Base has decreased the mislocation. Note also that the size of the error ellipse has decreased reflecting the decrease in modeling error due to the use of travel-time corrections within the Knowledge Base.

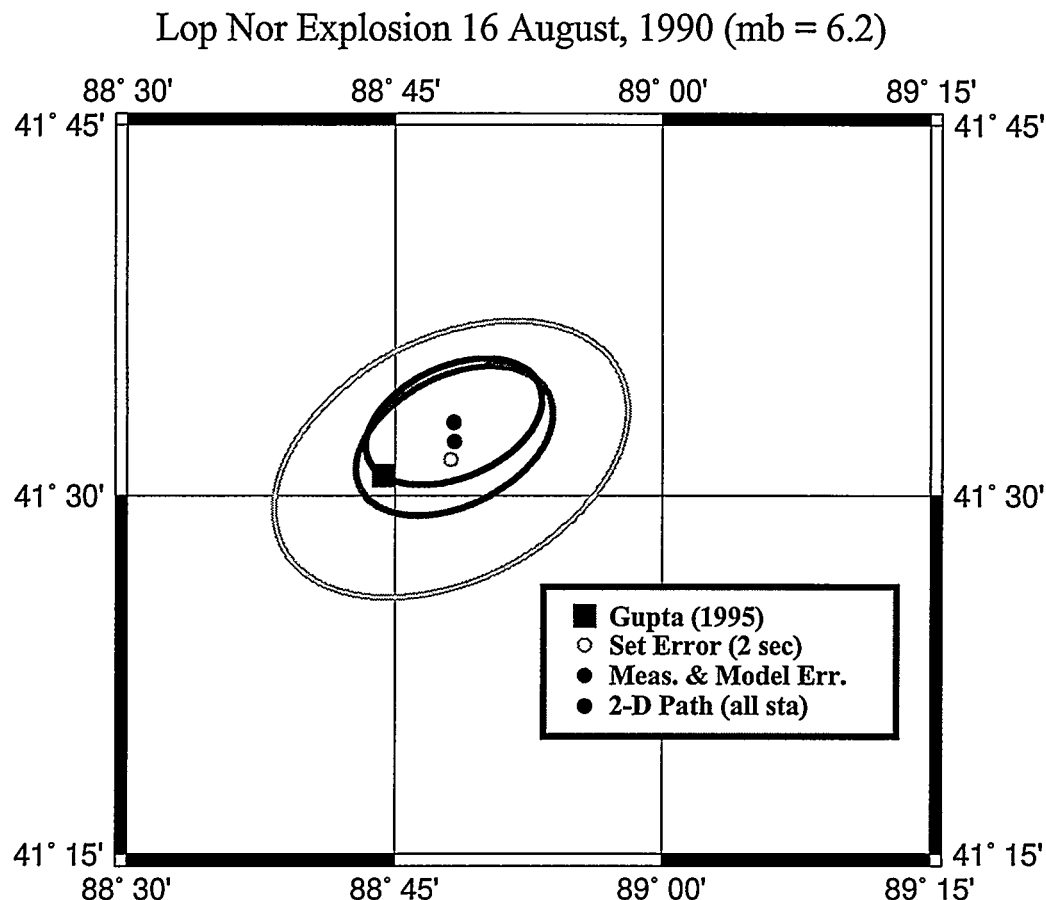


FIGURE 2. Event location using Knowledge Base correction information for a Lop Nor Nuclear explosion. The largest error ellipse is for the location using a fixed measurement error of 2 seconds. The next smaller ellipse uses the picked measurement error and modeling error defined by IASP91. The smallest error ellipse is for the location that incorporates the Knowledge Base 2-D travel time corrections.

Discrimination at a Seismic Station

Unlike the case for teleseismic data where the same formulas for magnitude and discriminants can be used regardless of source region, for regional data preferred discriminants for a station will be assigned to each source region. For example, station AAK has been studied extensively by Hartse et al. (1997). Their work has shown that several discriminant ratios are effective for discriminating between nuclear explosions at Lop Nor and earthquakes in the region. The preferred discriminants, distance correction formula, and a ranking can be queried from the Knowledge Base.

Specialized Waveform Processing

Since the Knowledge Base has catalog and waveform information, specialized waveform processing can be incorporated into SEA. For example, event location and characterization can be improved through waveform correla-

tion. When a suspicious event occurs in an area for which we have a catalog of known events, the event can be evaluated for similarity by correlating waveforms at one or more stations. Using MatSeis, we developed a suite of tools to do this. We use basic cross-correlation normalized by the auto-correlations of the individual waveforms. More complex approaches are possible and may be added in the future, but we found our simple approach to work quite well. As an example, we choose as a "suspicious" event an explosion from the Lop Nor test site (event 271418) and correlate it against the waveforms from other known explosions at the site, as well as earthquakes from the surrounding region. In Figure 3, we show the results for correlation of a 5 second window of the P phase for the station AAK. Our tool shows the results in three displays. First, the correlations are shown in a table, sorted by correlation value. Next the correlated waveform segments are shown in another window, again sorted by correlation and with the appropriate lags. Finally, the correlation values are shown on a map with the size of the symbol proportional to the degree of correlation. The tool has correctly identified event 271418 as being most similar to the nearest explosion, event 217186.

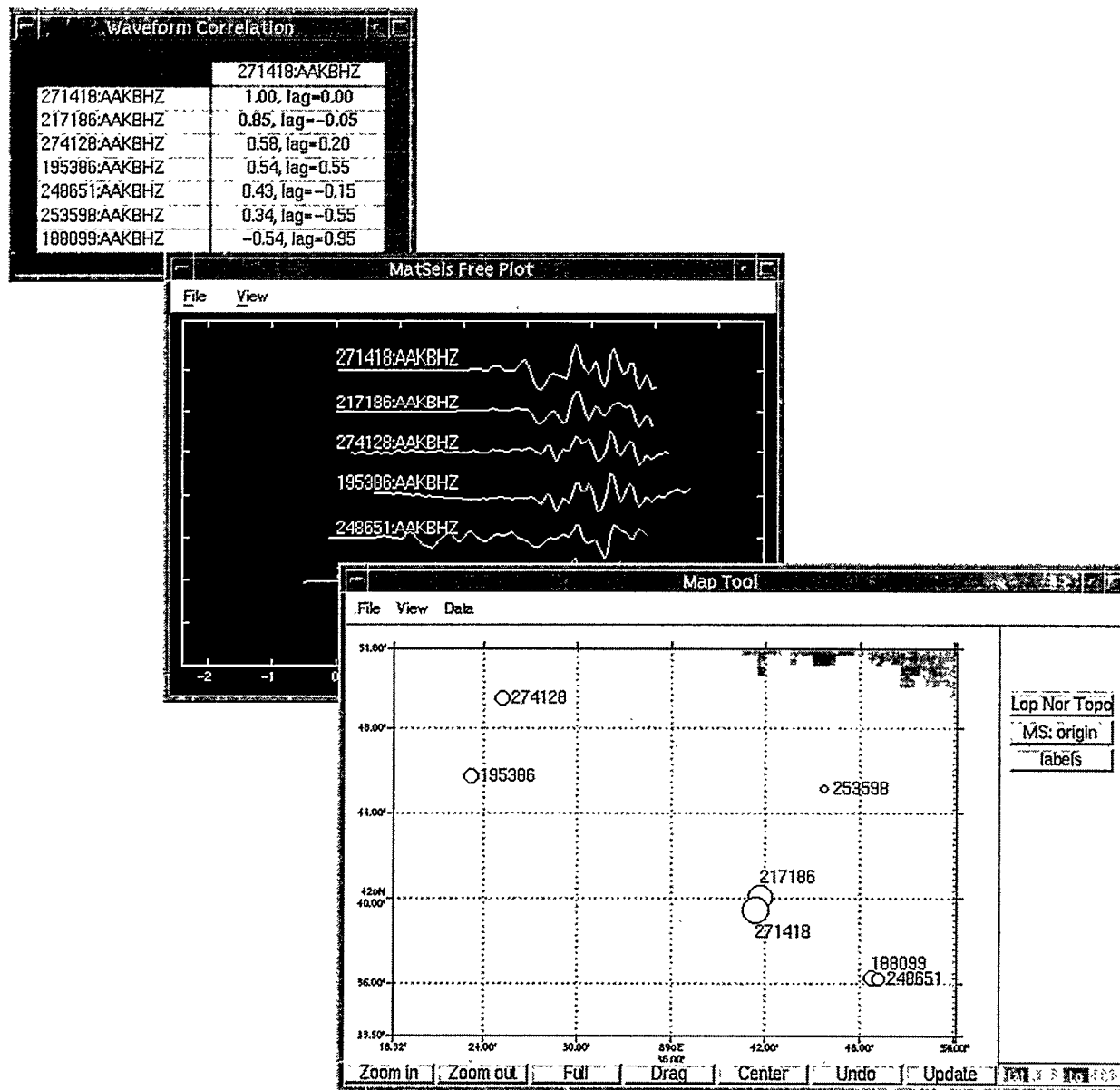


FIGURE 3. Waveform correlation for AAKBHZ between suspect event 271418 and events in the Lop Nor data set. Event 217186 has been identified as most similar, with a correlation coefficient of 0.85.

Identification and Use of Relevant Contextual Information

Once an event has been located, the location and the error ellipse can then be used to search Knowledge Base geographic data sets for relevant information. The China data set provides a variety of data sets which might be useful for identifying the event: e.g. locations of mines, earthquakes, and faults. Extracting potentially relevant information can be accomplished either by visual inspection or by using the built-in functionality provided by ArcView. Figure 4 shows an epicenter and 95% confidence ellipse displayed in ArcView. We have used ArcView to highlight all of the data sets within 50 km of the ellipse.

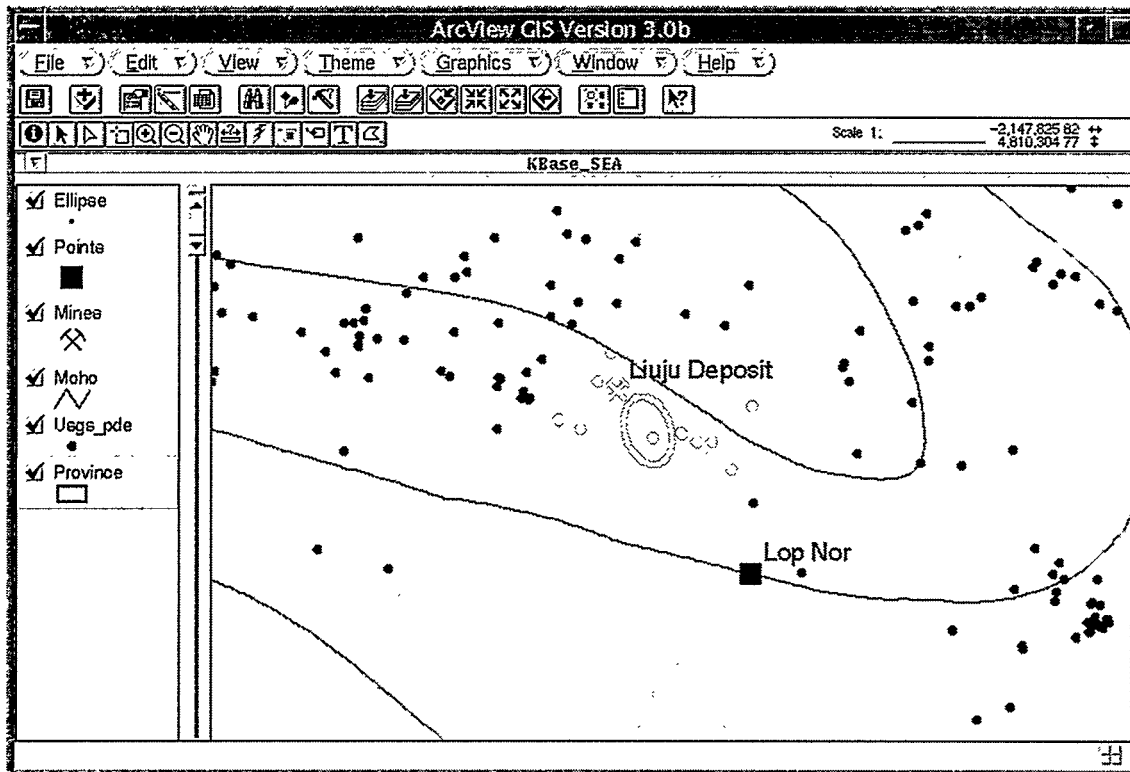


FIGURE 4. Identification of contextual information related to an event. Mine locations and seismicity within 50 km of the ellipse are highlighted.

Once relevant contextual information has been identified, it can be used to trigger specialized waveform processing. Continuing with our case from above, we note that there is a mine close to the calculated location. We might then want to perform cepstral or binary spectrogram analysis to search for ripple fire characteristics. Alternatively, we might choose to further refine or corroborate the location by performing FK and three-component analysis to extract directional information from the data. All of this functionality is available in MatSeis.

The contextual information in the Knowledge Base provides excellent opportunities for waveform modeling. The China data includes topographic elevation, depth to Moho, depth to base of lithosphere, and surface geology. As an example, we extract a profile from an event at the Lop Nor test site to the station WMQ. The top of Figure 5 shows a map view with the event, the station, topography, and contours for the depth to Moho. The bottom of Figure 5 shows a profile along the line shown in the map view, with topography and depth to Moho. Information from this profile may be used as input parameters for a synthetic seismogram code.

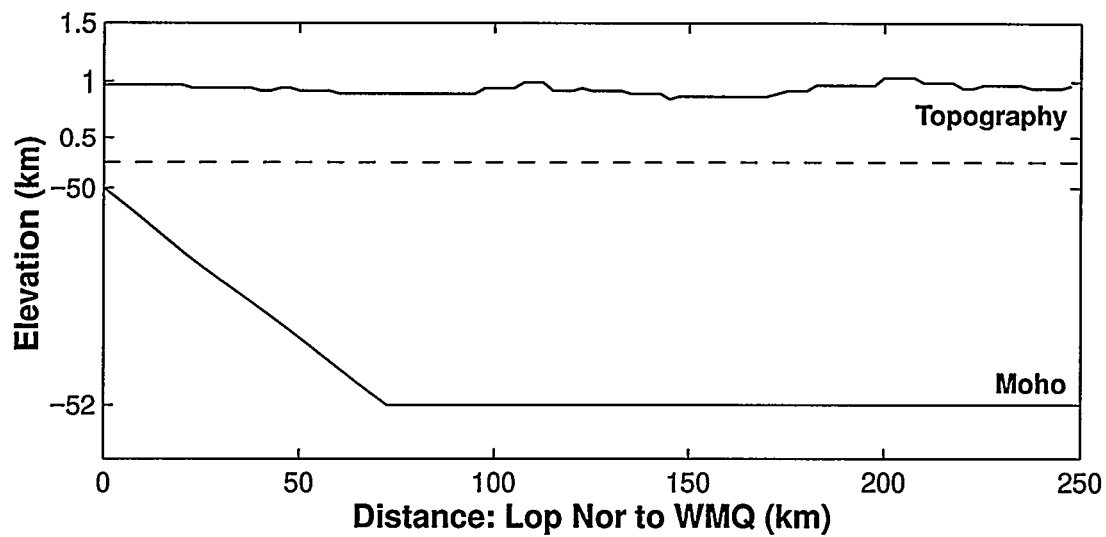
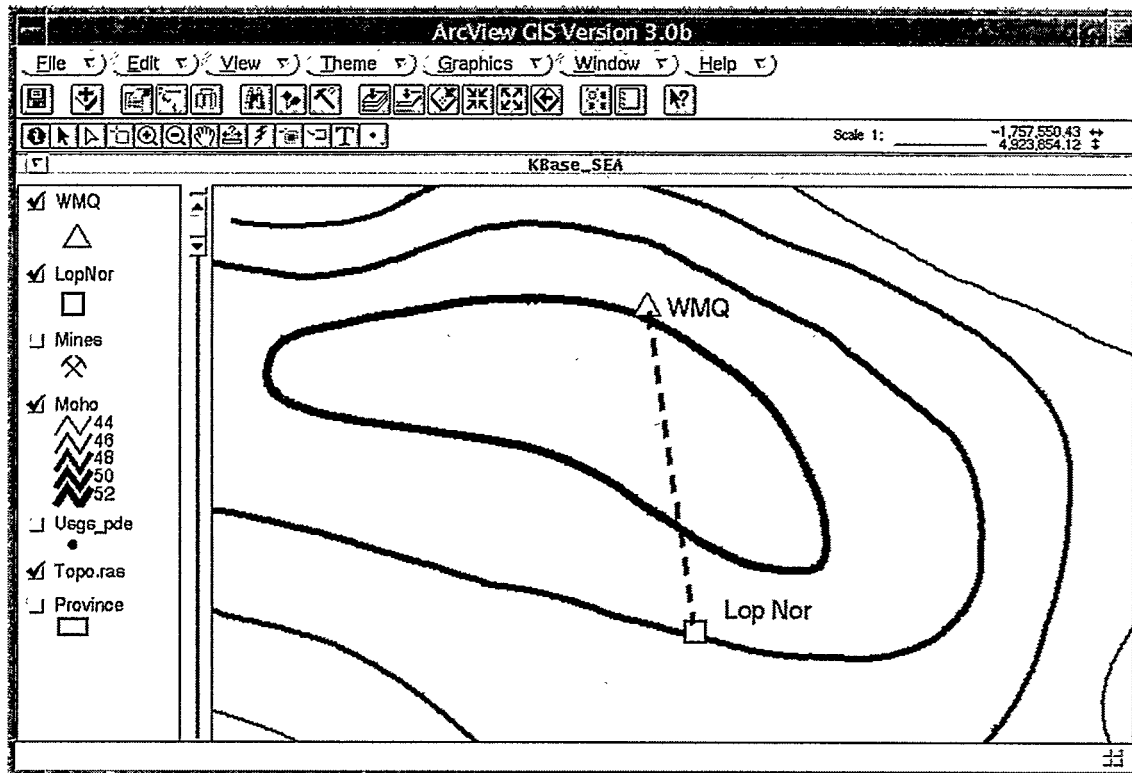


FIGURE 5. Waveform modeling profile extracted from Knowledge Base contextual information.

CONCLUSIONS AND RECOMENDATIONS

Even with a relatively small amount of data available and prototype tools, we have shown that the Knowledge Base can greatly aid Special Event Analysis. We chose to use MatSeis for waveform analysis and ArcView for GIS analysis and found both to be excellent choices due to their power and flexibility. Our highest priority for the immediate future will be to improve the communication between the packages, because we found that there is an almost constant need to communicate back and forth between them. Results of waveform analysis often lead to GIS analysis and vice versa. Additional tools will no doubt prove useful, but already both packages provide a wealth of tools which we have yet to fully explore.

We also recommend that some level of prioritization be provided with the Knowledge Base data sets where the number of data is large or there are seemingly redundant data sets. For instance, the reference event set we used for China contained approximately 300 events, many of which seem to cluster in certain areas. Based on our analysis, we believe that unless these events are substantially different from each other, it would be far more useful to either winnow the original data set down to a smaller number of reference events, or provide a prioritized index whereby the analyst could readily find the best event to use for a given region. This also applies to the GIS data sets. The China data set provides four sets of tectonic boundaries. For a SEA analyst under pressure to produce a result quickly, this wealth of information will likely prove a nuisance. While it is possible that more than one set might prove useful for some analysis, we believe that the analyst will be considerably aided by providing some sort of prioritization which will quickly guide the analyst to the preferred boundary set.

REFERENCES

- Armstrong, H. M., and R. G. Keyser (1997), The Browser Prototype for the CTBT Knowledge Base, Proceedings of the 19th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty.
- Bratt, S.R., and T.C. Bache (1988), Locating events with a sparse network of regional arrays, *Bull. Seismol. Soc. Am.*, 78, 780-798.
- Gupta, V. (1995), Locating nuclear explosions at the Chinese test site near Lop Nor, *Science and Global Security*, 5, 205-244.
- Harris, M., and C. Young (1996), MatSeis: A Seismic GUI and Toolbox for MATLAB, *Seismol. Res. Lett.*, 67(2), 267-269.
- Hartse, H.E., S.R. Taylor, W.S. Phillips, and G.E. Randall (1997), Regional event discrimination in central Asia with emphasis in western China, *Bull. Seismol. Soc. Am.*, 87, 551-568.
- Velasco, A. A. (1998), Lop Nor special event data set: Knowledge Base product, LA-UR-98-1148, 26pp.
- Young, C.J., J. R. Hipp, E. Shepherd, and S. G. Moore (1998), The DOE model for improving seismic event locations using travel time corrections: description and demonstration, submitted to the 20th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty.