

SAND97-1692C
SAND-97-1692C

Testing the Waveform Correlation Event Detection System: Teleseismic, Regional, and Local Distances

CONF-970967--5

Christopher J. Young, Judy I. Beiriger, J. Mark Harris, Susan G. Moore, Julian R. Trujillo

Sandia National Laboratories

Mitchell M. Withers, Richard C. Aster

New Mexico Institute of Mining and Technology

Sponsored by U.S. Department of Energy

Comprehensive Test Ban Treaty Research and Development Program **OSTI**

RECEIVED

JUL 14 1997

ABSTRACT

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED *xj*

Waveform Correlation Event Detection System (WCEDS) prototypes have now been developed for both global and regional networks and we have extensively tested them to assess the potential usefulness of this technology for CTBT (Comprehensive Test Ban Treaty) monitoring. In this paper we present the results of tests on data sets from the IDC (International Data Center) Primary Network and the New Mexico Tech Seismic Network. The data sets span a variety of event types and noise conditions. The results are encouraging at both scales but show particular promise for regional networks.

The global system was developed at Sandia Labs and has been tested on data from the IDC Primary Network. We have found that for this network the system does not perform at acceptable levels for either detection or location unless directional information (azimuth and slowness) is used. By incorporating directional information, however, both areas can be improved substantially suggesting that WCEDS may be able to offer a global detection capability which could complement that provided by the GA (Global Association) system in use at the IDC and USNDC (United States National Data Center).

The local version of WCEDS (LWCEDS) has been developed and tested at New Mexico Tech using data from the New Mexico Tech Seismic Network (NMTSN). Results indicate that the WCEDS technology works well at this scale, despite the fact that the present implementation of LWCEDS does not use directional information. The NMTSN data set is a good test bed for the development of LWCEDS because of a typically large number of observed local phases and near network-wide recording of most local and regional events. Detection levels approach those of trained analysts, and locations are within 3 km of manually determined locations for local events.

Keywords: event detection, event location, waveform correlation, CTBT monitoring

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

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, makes 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.

MASTER

OBJECTIVES

The goal of the Waveform Correlation Event Detection System (WCEDS) Project at Sandia has been to assess the potential usefulness for Comprehensive Test Ban Treaty (CTBT) monitoring of a seismic event detection system based on the correlation of waveforms. Shearer (1994) showed that this type of system may have an ability to detect events that traditional event detection systems cannot, but the potential relevance to CTBT verification was not clear due to the type of data used (long-period) and the less-stringent event detection requirements of that study.

We believe that pursuing alternative event detection technologies for CTBT monitoring is warranted for at least two reasons. First, although the current system (Global Association or GA) is highly sophisticated and is a vital part of the current monitoring pipelines at both the IDC (International Data Center) and the USNDC (United States National Data Center), the quality of the bulletins produced still shows significant room for improvement, particularly in reducing the number of false events (i.e. those that are not real and which must be disassociated by an analyst). Second, the USNDC is seeking to develop new methods to "spotlight" certain regions of interest to substantially lower the global threshold achieve in those regions. Many approaches to improving event detection could be investigated, of course, but to create the most sensitive detector possible a full waveform approach is especially attractive. Such an approach has not been practical in the past due to the processing and disk storage requirements, but the ever-increasing rate of technological innovation in computer hardware continues to reduce the severity of these problems.

INTRODUCTION

The WCEDS Project was initiated in the Spring of 1995. We started with the algorithm developed by Shearer to process long-period data, and by May of 1995 a rudimentary prototype was available. Details of the initial prototype development and testing were given in a presentation at the 17th Seismic Research Symposium (Young et al., 1995). The prototype was then refined based on the results of repeated testing with a set of continuous data from the Incorporated Research Institutions for Seismology (IRIS) broadband three-component global network, and the results of that study were presented at the 18th Seismic Research Symposium (Young et al., 1996a). At the same time, a local/regional version of the system was being developed and tested at New Mexico Tech using data from the New Mexico Tech Seismic Network (Withers, 1997).

In the past year, both systems have substantially matured and have undergone extensive testing. The global version has been modified to process array data, allowing us to conduct tests with data from the IDC Primary Network. The local/regional version has been fully-automated and is now routinely used by New Mexico Tech to detect and locate events in New Mexico. In this paper we will briefly review the basic features of the WCEDS system, and then discuss in detail recent modifications and testing results for the global and local/regional versions. The global and local/regional systems are described in detail in Young et al. (1996b) and Withers (1997).

REVIEW

WCEDS is a grid-based automatic system which examines processed continuous seismic data streams for matches with some portion of a master set of waveforms which we refer to as the *Master Image* (MI). The grid covers the area which the user is interested in monitoring, and should have a spacing appropriate for the desired accuracy of the locations and the frequency content of the data. To understand how the detector works, it is useful to think of it as creating a waveform profile for each grid point at a given time and comparing this observed profile with the expected

profile if an event had indeed happened at that point and time. A "good" match would indicate that an event had occurred at the location at that given time. The comparison of observations and the MI is accomplished via simple dot products of the waveforms.

Prior to forming the profiles, the waveforms are processed (e.g. bandpass filtering followed by STA/LTA processing) to enhance phases relative to noise and to generalize the waveforms for correlation with the generalized MI. We refer to this step as *preprocessing* because it occurs prior to execution of the detector itself. The detector is then run on the preprocessed data at some specified time discretization interval.

Detection and location are accomplished by computing the correlation value for each grid point for each time point over some specified period (e.g. an hour), and then finding the grid and time point combination with the maximum value; if this value exceeds a threshold value, an event is declared. Other events in the processed period can be detected by masking the detected event and repeating the search.

To improve the efficiency of the algorithm, at each potential origin time all possible epicentral distance correlations for each processed waveform can be pre-calculated and stored in an array which we call the C matrix:

$$C_{ij} = \sum_{k=1}^{N_t} M_{ik} D_{kj} \quad (\text{EQ 1})$$

Where D is the data matrix whose columns are the observed waveforms at each station, M is the MI matrix whose rows are the expected waveforms at each discrete distance, and N_t is the number of time points. i and j refer to the distance (discrete) and the station, respectively. Note that C has one column for each station and a number of rows equal to the number of discrete distances in the MI.

The detector can be run much more efficiently using C because no further dot products are needed; C contains all of the information necessary to calculate the correlation sum for any grid point. The detector output at each grid point is determined by summing one element from each column of C:

$$O = \sum_{j=1}^{N_s} C_{ij} \quad (\text{EQ 2})$$

Where N_s is the number of stations. Which row index is used for each column is dependent on the epicentral distance between the station and the grid point for which the summation is calculated. This indexing information, which we refer to as the *grid map*, is static and need only be calculated once, prior to running WCEDS, and stored as a look-up table.

RESEARCH ACCOMPLISHED

WCEDS (Global System)

Data Set

We have had difficulty in getting continuous Primary Network data from either the IDC or the USNDC and so have had to limit our testing to the few short intervals which we were able to obtain. For the purposes of this paper we will discuss 2 hours of data from June 19, 1996 (00:00 - 02:00). Despite the obvious limitations of such a small data set, we feel that this interval offers a

number of excellent tests of the WCEDS systems. The events reported by the IDC for the interval are listed in Table 1.:

TABLE 1. IDC REB events for 6/19/96 00:00 to 02:00

origin time	latitude	longitude	depth	mb	region	in AEB?
00:01:31.5	7.36 S	123.22 E	619.7	3.2	Banda Sea	no
00:18:08.9	36.11 N	35.87 E	51.0	4.0	Turkey	yes
00:55:18.9	13.5 S	174.59 W	0.0	4.2	Samoa	yes
01:16:17.7	5.69 N	126.18 E	87.9	4.3	Mindanao	yes
01:38:57.8	6.58 S	132.92 E	0.0	3.4	Tanimbar	yes
01:47:34.9	41.92 N	142.47 E	71.0	3.6	Hokkaido	yes

This is a challenging set of events for any automatic detection system. None of the events is particularly large, most of them are in the southern Pacific region where Primary Network coverage is poor, and one was small enough to have been missed in the AEB (Alpha Event Bulletin). According to the REB (Reviewed Event Bulletin), only Turkey and Mindanao had substantial signal above background at more than a few Primary Network stations.

New Preprocessing

Because the Primary Network includes arrays, to process this data set we had to develop a strategy to pre-process array data. The simplest approach is to treat each channel of an array as a separate station and process in the standard manner, but to do so with such a coarse global grid (we use 2 degrees) would give none of the coherent signal enhancement which arrays are designed to provide. In fact, because our Master Image is smoothed and therefore the data to be correlated must also be smoothed (i.e. run through an STA/LTA filter), it is essential that array noise cancellation be accomplished in the pre-processing prior to the smoothing.

One obvious method to achieve the noise cancellation is to form a beam to each grid point, which is essentially the strategy currently used for array signal detection at the IDC and USNDC (a set of beams is formed which spans azimuths and slownesses and then each is run through a signal detection algorithm). We did not pursue this approach, however, because it requires the calculation and storage of a huge number of beams, which would significantly degrade our system performance.

In fact, given that our full matrix correlation approach uses an azimuth-independent Master Image, we would prefer to use one "best beam" for each array. This could be accomplished by forming some sort of a composite maximum value beam from a beam set, but we found that the spatial coherency analysis technique of Wagner and Owens (1996) provides a much more efficient means to do this. To calculate a spatial coherency stream for an array, we transform all data channels to the frequency domain, form the covariance matrix, and then solve for the principal eigenvalue. This process is repeated for a moving window to generate an output time series. Thus the technique, which can be applied to a station with any number of data channels (we have used for the Primary Network 3-component stations as well), yields one directionally independent stream for each array, as we require. Further, this stream should always have a high response to any spatially coherent signal, regardless of the azimuth, slowness, or even phase delay between channels. In our testing, we found that the technique can yield results quite comparable to, if not superior to, traditional beam-forming but that it is very sensitive to parameter choice. We selected optimal parameters (10 second window, 3 frequency bands: 0.5-1.5 Hz, 1.5-2.5 Hz, 2.5-3.5 Hz) based on

extensive analysis of the smaller, more poorly detected events from segments of our Primary Network test data sets.

Test Results

Despite the excellent signal enhancement provided by the spatial correlation technique, when we ran the original WCEDS prototype on our Primary Network test sets using relative timing only, we found that we had difficulty properly building many of the smaller events due to the sparse coverage provided by the Primary Network. Particularly troublesome are small events (less than mb 4.0) in the South Pacific, where network coverage is especially poor (the closest station can be more than 40 degrees away). These events are difficult for WCEDS to build properly because typically the only observations are P phases at the Australian arrays WRA and ASAR, which are unfortunately quite close together and hence function much like a single array.

Figure 1 compares the REB events with the set of WCEDS events for the 2 hour test set. We used a total of 35 stations: 12 arrays and 23 three-component single site stations.

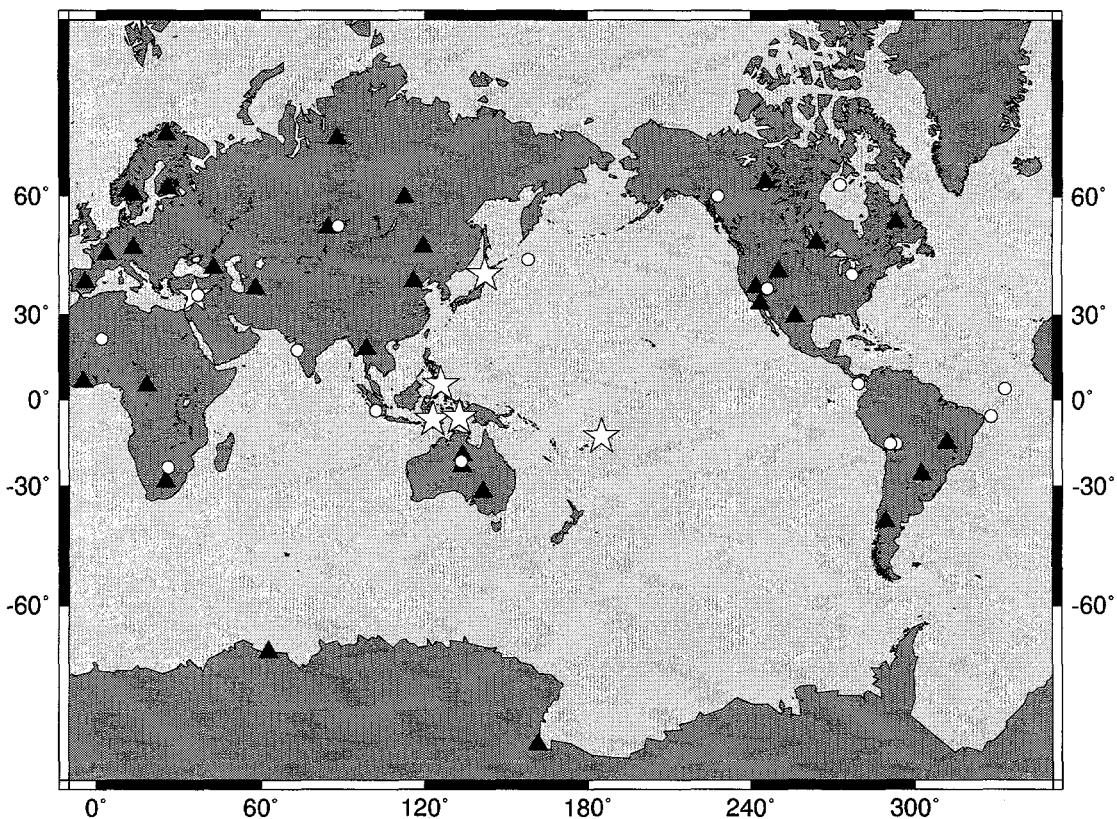


FIGURE 1. WCEDS vs. REB event locations - A comparison of REB events (white stars) and WCEDS events (white circles). The stations in the network used by WCEDS are indicated with black triangles.

The large number of events built by WCEDS is due to the fact that we let the system continue to try to build events in the time interval well down into the noise level to give the system every opportunity to build all of the REB events. In spite of this, very few of those events seem to have been built. There is one excellent agreement (the event in Turkey), but in general we found very little correspondence between the WCEDS events and the REB events. This is because very few of the REB events are large enough to be seen by more than a few stations, so relative timing does

not constrain them well. In our tests with the IRIS network we found that if an event is at all widely recorded, as was the event in Turkey, WCEDS will do an excellent job of locating it because the relative timing information across the network provides a unique constraint on location and origin time. Conversely, when very few stations see an event, the relative timing provides a very non-unique constraint and the event may be very poorly built, perhaps so poorly that it will show almost no resemblance to the true event. In fact, we have found that for several of the REB events the relative timing information could be fit equally well by so many different grid and time point combinations that the wrong combination was often picked in order to falsely correlate an unrelated noise feature at a non-observing station with a predicted phase in the Master Image.

Fortunately, with array data (and perhaps three-component data, as well), we can significantly improve this result by taking advantage of the available directional information. By investigating the events in our sample data set, we found that a simple way to use directional information is to verify directional compatibility with a grid point location. Almost all of the legitimate signals from the REB events seen at the arrays had observed azimuths which were within +/- 10 degrees of the predicted azimuths, suggesting that this simple test could be used to prevent non-compatible high-amplitude features from contributing incorrectly to event detection and location. Thus we modified the software to require that the observation fall within these limits, or the correlation for that particular station/phase will not be added to the overall grid point value. This prevents the random inclusion of unrelated signals with the true signals to get a higher correlation (but at the wrong grid and time point).

To implement this test, for each array we require a continuous stream of azimuth as a function of time. We obtain the additional streams by measuring the azimuth of the peak of an FK transform of a moving window of the same length as the spatial coherence calculation. In many cases, there is no obvious peak and the azimuth is indeterminate which could lead to erroneous results, but we avoid this problem by only referring to directional information when a sufficiently strong signal is seen on the spatial coherency stream.

The results of processing the same two hour interval using the arrays only and incorporating the azimuthal verification step are shown in Figure 2. The improvement is dramatic. In this case, the first 6 events that WCEDS built were the REB events, including the Banda Sea event which did not appear in the AEB. The sensitivity of the system is excellent, even though we are using only 12 arrays. Some of the mislocations are large, but very understandable when one examines the contributing stations and considers how the technique works. For example, the Tonga event was seen only by ASAR and WRA, hence the azimuth to those stations is well-constrained but the distance is not. Similarly, the Turkey event was seen only by the northern European arrays and so is azimuthally constrained to be somewhere along the arc connecting those arrays and the true location.

An obvious way to improve these mislocations is to incorporate slowness information as a verification as we did with azimuth, and we have begun to investigate this. We also generate a continuous slowness stream from the moving FK window, so the data is readily available. So far, the results have been mixed. In some cases slowness can dramatically improve location while in others it can prevent events from being properly built. It may be that in order to use this information properly, at some stations we will need to account for local heterogeneity which causes observed slownesses to deviate from simple radial Earth model predictions.

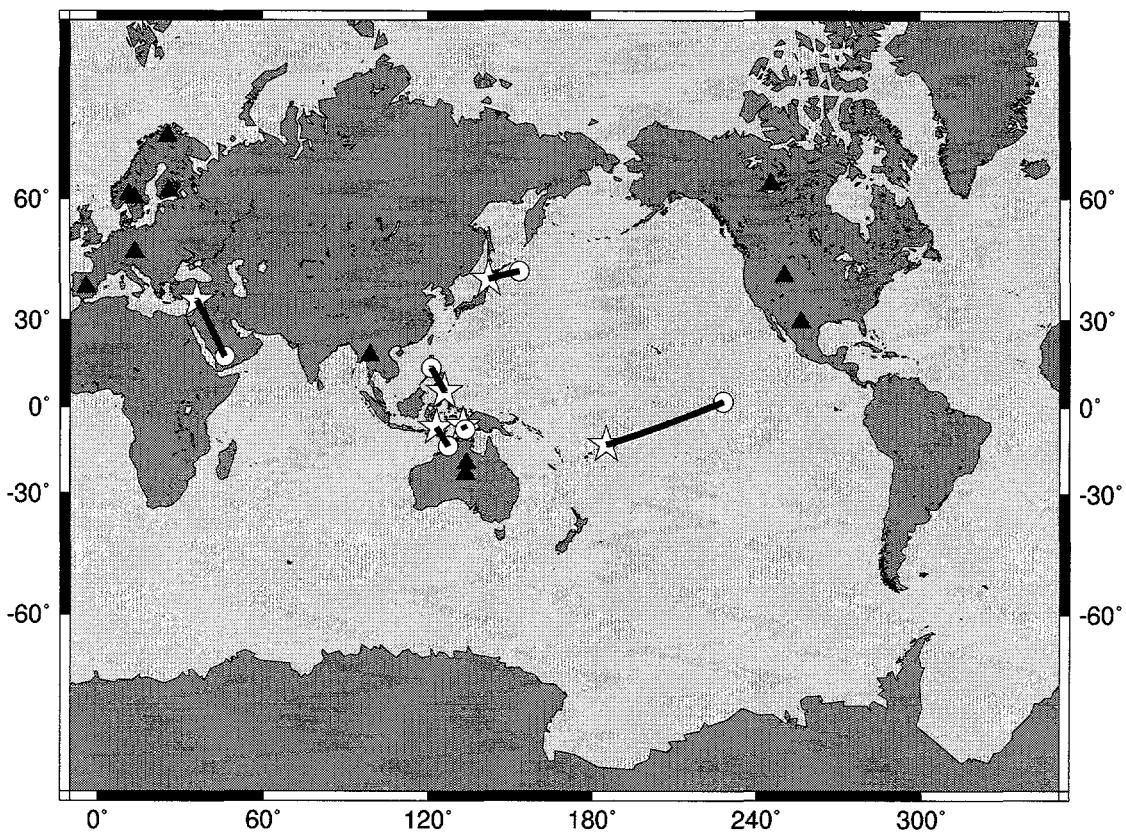


FIGURE 2. WCEDS vs. REB event locations: directional information used - A comparison of REB events (white stars) and the first 6 WCEDS events (white circles). Corresponding events are linked with black lines. The stations (all are arrays) in the network used by WCEDS are indicated with black triangles.

LWCEDS (Local/Regional System)

Data Set

New Mexico Tech operates a short-period digital seismic network consisting of 17 vertical-component seismometers, nine of which are employed in an approximate 60 km aperture centered on the central Rio Grande rift near Socorro, New Mexico. We will present results from testing using data from those nine stations. The test set used for this paper consisted of 10 local events selected to span a reasonable range of location and event quality. The events are listed in Table 2. The events were located by New Mexico Tech personnel who hand-picked the arrivals and then used the program *Seismos* (Hartse, 1991) to invert for location. The magnitudes are duration magnitudes (Ake, 1983).

TABLE 2. New Mexico Tech Seismic Network test events

origin time	latitude (N)	longitude (W)	depth	Md	RMS
09/20/95, 09:10:36.16	34.1158	106.8282	5.89	0.84	0.218
09/22/95, 10:37:57.84	34.3043	106.9240	6.35	0.84	0.362
09/22/95, 22:45:56.25	34.3550	106.7097	8.71	1.35	0.265
09/24/95, 16:04:51.57	34.0333	107.0288	7.56	1.37	0.369
09/24/95, 16:06:44.75	34.0345	107.0328	9.22	0.16	0.242
09/24/95, 16:11:25.43	34.0323	107.0318	8.89	0.33	0.234
09/24/95, 17:58:40.95	34.0297	107.0282	7.81	1.57	0.348
09/24/95, 18:10:55.02	34.0285	107.0283	8.80	0.59	0.202
10/14/95, 04:24:03.87	34.4477	107.0180	4.93	1.43	0.191
09/27/96, 08:35:33.71	34.4275	107.0678	2.08	1.32	0.392

PreProcessing

Unlike the teleseismic problem where often only a P arrival can be clearly identified, at regional or local distances several phases are often readily apparent within observed waveforms even for small events. Thus a full waveform technique should be particularly effective at these distances. Unfortunately, however, local and regional phases span a wide range of frequencies, so it is not possible to choose a single set of STA and LTA windows which will do a good job of emphasizing all phases. For this reason, we developed an adaptive STA/LTA algorithm (Withers et al., 1997) which maintains the same proportion between the window lengths but changes their absolute lengths to adapt to the dominant frequency of the signal. Our testing indicates that this algorithm has an excellent ability to emphasize each phase in a local or regional waveform, from high-frequency Pn to low-frequency Lg.

Test Results

The LWCEDS system is a much simpler system than the global system because it does not contain any of the masking or continuous processing capabilities. Masking is unnecessary because overlapping events are rarely recorded by the network due to the relatively low level of seismic activity in New Mexico and the size of the network (LWCEDS does not try to locate teleseismic or far-regional events, which it identifies by frequency content). Also, true continuous operation is not necessary because the algorithm is launched by a trigger in the acquisition system which is set to "go off" when a significant signal is measured at a user defined number of stations (currently 4) in the network during some specified period of time. For the test set, the trigger was assumed to have occurred (it did occur originally -- hence the presence of these events in the archive) and LWCEDS was launched manually on the archived interval for each of the 10 events.

A comparison of the LWCEDS and Seismos events is shown in Figure 3. The results are very encouraging. The largest mislocation is 3.53 km, and none of the others is more than 2 kms. Despite the fact that we use no directional information, these results are at least as impressive as those of the global system, suggesting that a low threshold ($>\sim 0.1$) and high-quality locations can be achieved using relative timing information only if the event is recorded by several stations in

the network, as we discussed above. Tests with an expanded test set including regional events from New Mexico and eastern Arizona show a similar detection sensitivity to those events, though the mislocations increase dramatically (10 to 100 km), depending on the size of the azimuthal gap.

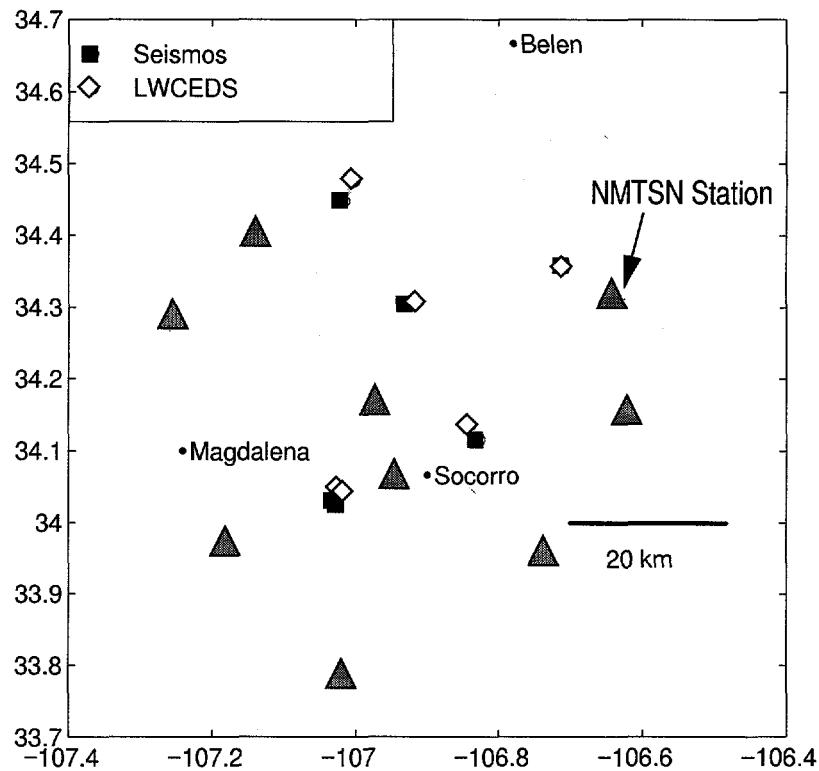


FIGURE 3. LWCEDS vs. Seismos locations - A comparison of LWCEDS locations and analyst locations for the 10 event test set. The shaded triangles show the locations of the elements of the New Mexico Tech Seismic Network (NMTSN).

CONCLUSIONS AND RECOMMENDATIONS

Mature, second-generation WCEDS software has been developed for both global and local/regional networks and tested with data sets from the IDC Primary Network and the New Mexico Tech Seismic Network, respectively. Our results indicate that the technique can work well at both scales but shows particular promise for local/regional networks where multiple phases for a given event are often recorded by many stations. In fact, the local/regional version of WCEDS (LWCEDS) has proven so successful that it is now routinely used at New Mexico Tech to automatically locate events which trigger their acquisition system. In regions of the world where dense single station coverage is available, this technology could be used to develop highly sensitive, highly accurate monitoring systems.

Our results from the global system have been much less clear cut. Testing on IDC Primary Network data with a version of the system which does not use directional information (ala LWCEDS) yielded unsatisfactory results for many events in the IDC REB. However, if directional information is included results improve dramatically. By including a simple azimuthal consistency check we found that WCEDS can achieve a detection threshold comparable to if not superior to GA with a low false alarm rate, though the locations can be off by as much as 30 degrees

for events with large azimuthal gaps. We have begun to test schemes using slowness information as well, and have found the results to be mixed: mislocations are generally considerably reduced and the number of misassociated detections decreases dramatically, but in some cases obvious phases are missed because observed slownesses do not match predictions.

Our future work will focus on the use of the WCEDS technology for "spotlighting". We believe that both the global and the local/regional results suggest that WCEDS shows promise for this problem. Which version of the algorithm is most appropriate will depend on the proximity of the monitoring stations to the region of interest. Our testing of both versions suggests that a very low detection threshold could be achieved.

REFERENCES

Ake, J. (1983). A magnitude scale for central New Mexico based on signal duration, *Geophysics Open-file Report*, 45, 8 pp.

Hartse, H. (1991). Simultaneous hypocenter and velocity model estimation using direct and reflected phases from microearthquakes within the central Rio Grande Rift, New Mexico, *Ph.D. Dissertation, New Mexico Tech*, 251 pp.

Shearer, P. M. (1994). Global seismic event detection using a matched filter on long-period seismograms, *J. Geophys. Res.*, 99, 13713-13725.

Withers, M. M. (1997). An automated local/regional seismic event detection and locations system using waveform correlation, *Ph.D. Dissertation, New Mexico Tech*, 195 pp.

Withers, M. M., R. C. Aster, C. J. Young, J. I. Beiriger, J. M. Harris, S. G. Moore, J. R. Trujillo (1997). A comparison of select trigger algorithms for automated global seismic phase and event detections, *submitted to Bull. Seismol. Soc. Amer.*

Young, C. J., J. I. Beiriger, J. M. Harris, S. G. Moore, J. R. Trujillo, M. M. Withers, and R. C. Aster (1996a). The waveform correlation event detection system project: issues in system refinement, tuning, and operation, *Proceedings of the 18th Annual Seismic Research Symposium on Monitoring a CTBT*, p. 789-798.

Young, C. J., M. Harris, J. I. Beiriger, S. G. Moore, J. M. Trujillo, M. M. Withers, and R. C. Aster (1996b). The waveform correlation event detection system project, phase I: issues in prototype development and testing, *SAND Report #SAND96-1916*.

Young, C. J., J. I. Beiriger, J. M. Trujillo, M. M. Withers, R. C. Aster, L. Astiz, and P. M. Shearer (1995). WCEDS: a waveform correlation event detection system, *Proceedings of the 17th Seismic Research Symposium on Monitoring a CTBT*.

DISCLAIMER

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