REGIONAL SEISMIC DISCRIMINATION OPTIMIZATION WITH AND WITHOUT NUCLEAR TEST DATA: WESTERN U.S. EXAMPLES

William R. Walter, Kevin M. Mayeda, Rengin Gok, Arthur J. Rodgers, Alan Sicherman, Tracy Hickling, Doug Dodge, Eric Matzel, Mike Ganzberger, and Veronica Parker

Lawrence Livermore National Laboratory

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

Contract No. W-7405-ENG-48

ABSTRACT

The western United States (U.S.) has abundant natural seismicity, historic nuclear explosion data, and widespread mine blasts, making it a good testing ground to study the performance of regional source-type discrimination techniques. We have assembled and measured a large set of these events to systematically explore how to best optimize discrimination performance. Nuclear explosions can be discriminated from a background of earthquakes using regional phase (Pn, Pg, Sn, Lg) amplitude measures such as high frequency P/S ratios. The discrimination performance is improved if the amplitudes can be corrected for source size and path length effects. We show good results are achieved using earthquakes alone to calibrate for these effects with the magnitude and distance amplitude correction (MDAC) technique (Walter and Taylor, 2002). We show significant further improvement is then possible by combining multiple MDAC amplitude ratios using an optimized weighting technique such as linear discriminant analysis (LDA). However, this requires data or models for both earthquakes and explosions. In many areas of the world regional distance nuclear explosion data are lacking, but mine blast data are available. Mine explosions are often designed to fracture and/or move rock, giving them different frequency and amplitude behavior than contained chemical shots, which seismically look like nuclear tests. Here we explore discrimination performance differences between explosion types, the possible disparity in the optimization parameters that would be chosen if only chemical explosions were available and the corresponding effect of that disparity on nuclear explosion discrimination.

There are a variety of additional techniques in the literature also having the potential to improve regional high frequency P/S discrimination. We explore two of these here: three-component averaging and maximum phase amplitude measures. Typical discrimination studies use only the vertical component measures and for some historic regional nuclear records these are all that are available. However, S-waves are often better recorded on the horizontal components and some studies have shown that using a three-component average or a vertical-P/horizontal-S or other three-component measure can improve discrimination over using the vertical alone (e.g., Kim, et al., 1997; Bowers, et al., 2001). Here we compare the performance of vertical and three-component measures on the western U.S. test set.

A complication in regional discrimination is the variation in P- and S-wave propagation with region. The dominantly observed regional high frequency S-wave can vary with path between Sn and Lg in a spatially complex way. Since the relative lack of high frequency S-waves is the signature of an explosion, failing to account for this could lead to misidentifying an earthquake as an explosion. The regional P phases Pn and Pg vary similarly with path and also with distance, with Pg sometimes being a strong phase at near regional distances but not far regional. One way to try and handle these issues is to correct for all four regional phases but choose the phase with the maximum amplitude. A variation on this strategy is to always use Pn but choose the maximum S phase (e.g., Bottone et al., 2002). Here we compare the discrimination performance of several different (max P)/(max S) measures to vertical, three-component and multivariate measures. Our preliminary results show that multivariate measures perform much better than single ratios, though transportability of the LDA weights between regions is an issue. Also in our preliminary results, we do not find large discrimination performance improvements with three-component averages and maximum phase amplitude measures compared to using the vertical component alone.

OBJECTIVES

Monitoring the world for potential nuclear explosions requires characterizing seismic events and discriminating between natural and man-made seismic events, such as earthquakes and mining activities, and nuclear weapons testing. We continue developing, testing, and refining size-, distance-, and location-based regional seismic amplitude corrections to facilitate the comparison of all events that are recorded at a particular seismic station. These corrections, calibrated for each station, reduce amplitude measurement scatter and improve discrimination performance. We test the methods on well-known (ground truth) data sets in the U.S. and then apply them to the uncalibrated stations in Eurasia, Africa, and other regions of interest to improve underground nuclear test monitoring capability.

RESEARCH ACCOMPLISHED

As part of the overall National Nuclear Security Administration's Ground-Based Nuclear Explosion Monitoring Research and Engineering program, we continue to pursue a comprehensive research effort to improve our capabilities to seismically characterize and discriminate underground nuclear tests from other natural and man-made sources of seismicity. To reduce the monitoring magnitude threshold, we make use of regional body and surface wave data to calibrate each seismic station. Our goals are to reduce the variance and improve the separation between earthquakes and explosion populations by accounting for the effects of propagation and differential source size.

Western U.S Data Corrected for Magnitude and Distance Effects

We have been re-examining the large database of the western U.S. underground nuclear tests and earthquakes we assembled under a prior year broad agency announcement award (Walter et al., 2003). This western U.S. nuclear explosion data covers a wide range of depths and material properties and has excellent ground truth information (Springer et al., 2002). This is unlike the situation in most of the world where regional recordings of nuclear tests are scarce and discrimination optimization needs to be done in their absence. In addition we have chemical explosions recorded at the same stations from the Arizona Source Phenomenology Experiment (AZSPE). The AZSPE carried out dedicated single shot chemical explosions under a variety of depth and confinement conditions in two mining regions, a soft rock coal mine and a hard rock copper mine (see Bonner et al., 2005, this Proceedings). These mining regions also routinely detonate ripple-fired production blasts that can be observed at regional distances. The availability of both nuclear and chemical explosions lets us examine the differences in optimization and performance of the two source types relative to the earthquakes. The locations of the data and stations discussed in this paper are shown in Figure 1.

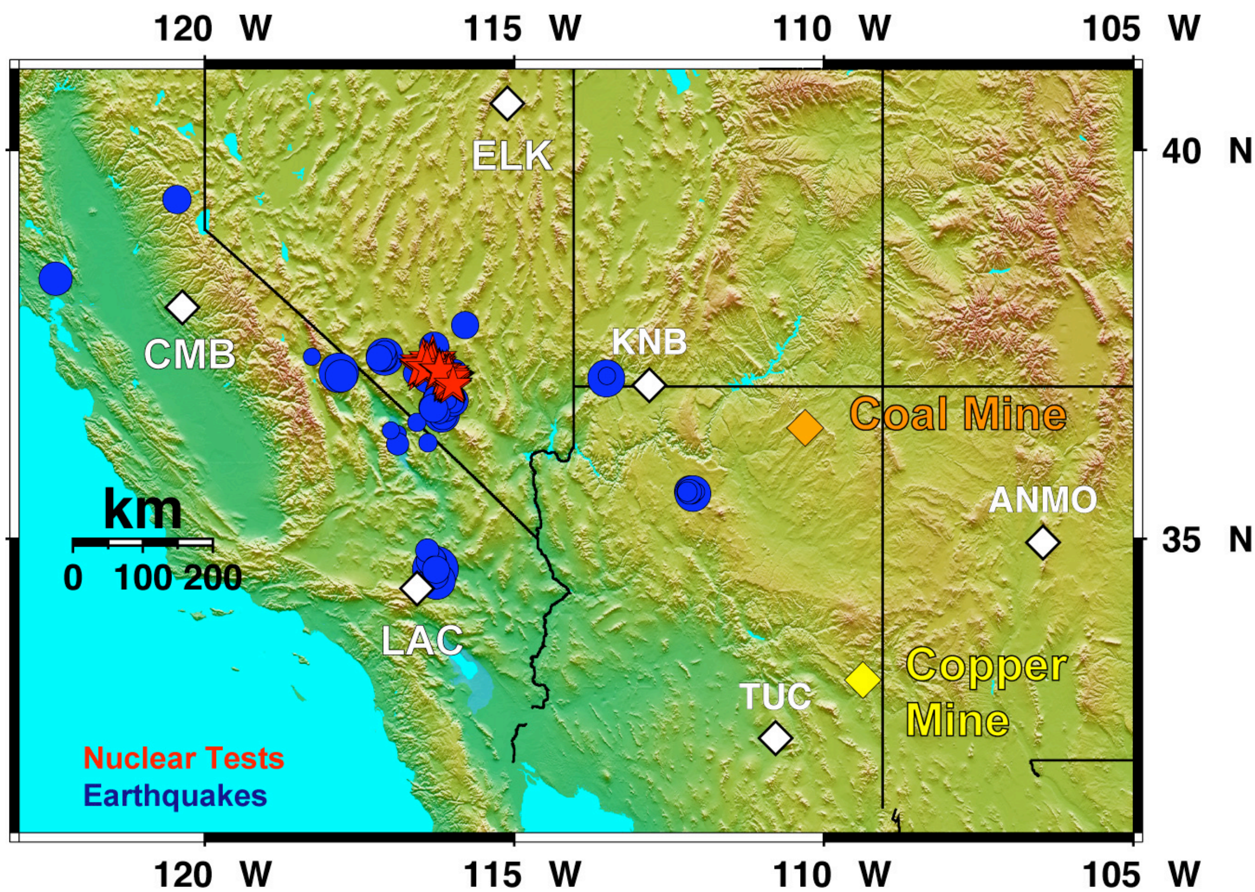

Figure 1. Map showing the location of earthquakes, historic nuclear explosions, mining explosions and stations discussed in this paper.

Effective earthquake-explosion discrimination has been demonstrated in a broad variety of studies using ratios of regional amplitudes in high-frequency (primarily 1- to 20-Hz) bands (e.g., Walter, et al., 1995; Taylor, 1996; Hartse et al., 1997; Rodgers and Walter, 2002; Taylor et al., 2002; Battone et al., 2002; and many others). When similar-sized earthquakes and explosions are nearly co-located, we can understand the observed seismic contrasts, such as the relative P- to S-wave excitation, in terms of depth, material property, focal mechanism and source time function differences. However, it is well known that path propagation effects (e.g. attenuation, blockage) and source scaling effects (e.g., corner frequency scaling with magnitude) can make earthquakes look like explosions and vice versa. We have developed a technique called MDAC (Walter and Taylor, 2002) that can account for these effects with proper calibration. We use the earthquakes alone to determine the MDAC parameters such as geometrical spreading, frequency dependent Q and the average apparent stress. After calibration the MDAC formulation provides expected spectral amplitudes as a function of phase, magnitude and distance. These can then be subtracted from the actual observations. For earthquakes the corrected data should exhibit a close zero mean, and a magnitude and distance detrended population. Explosions should have significant non-zero mean residuals, leading to improved discrimination. We show the results of a low to high frequency Lg spectral ratio before and after MDAC correction in Figure 2.

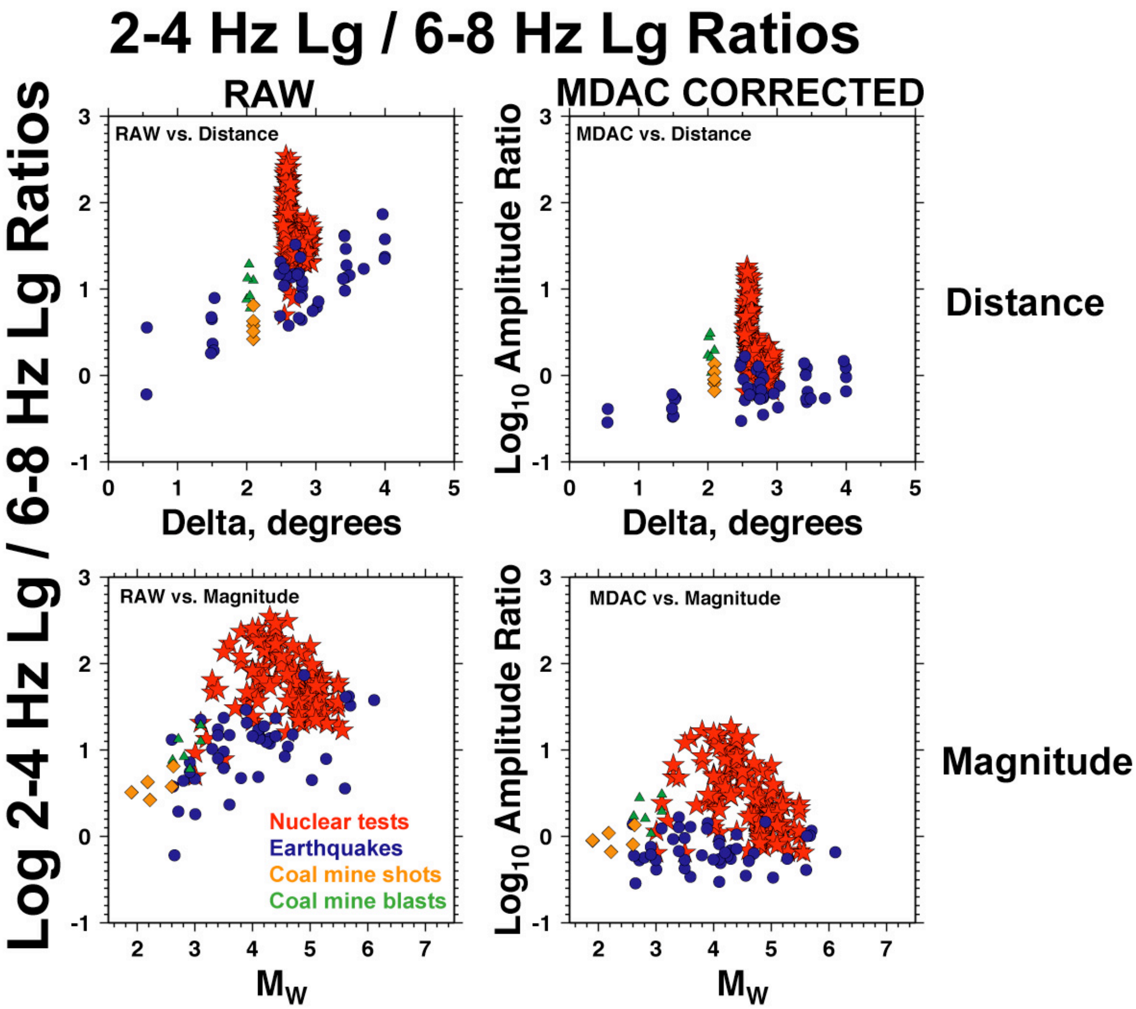

Figure 2. Western U.S earthquakes (blue circles), nuclear explosions (red stars), northern Arizona coal mine dedicated shots (orange diamonds) and regular production mine blasts (green triangles) for the discriminant ratio of (2–4 Hz Lg)/(6–8 Hz Lg) at station KNB. The left-hand side shows raw data as a function of distance (top) and magnitude (bottom). The right-hand side shows MDAC corrected data. Note that strong distance and magnitude trends apparent in the raw data are removed by MDAC, improving discrimination.

After the MDAC correction we can explore optimal combinations of particular regional discriminants (e.g., Taylor, 1996). We use LDA to find the optimal coefficients to combine the measurements. As an example of this we show in Figure 3 a combination at station KNB of three different regional phase and spectral ratios. The metric of performance we use is the equiprobable point, which provides a measure of the overlap of the earthquake and explosion populations. It is the point on a receiver operating characteristic (ROC) tradeoff curve where the error rates are equal. For example an equiprobable point of 0.1 implies that 10% of the earthquakes are misclassified as explosions and 10% of explosions are misclassified as earthquakes. In practice, one might choose a decision line with unequal error rates, such as by picking a low probability of misclassifying an explosion. The equiprobable point provides a single numerical measure of performance that is much more intuitive than other measures such as Mahalanobis distance, though it can be related to that measure.

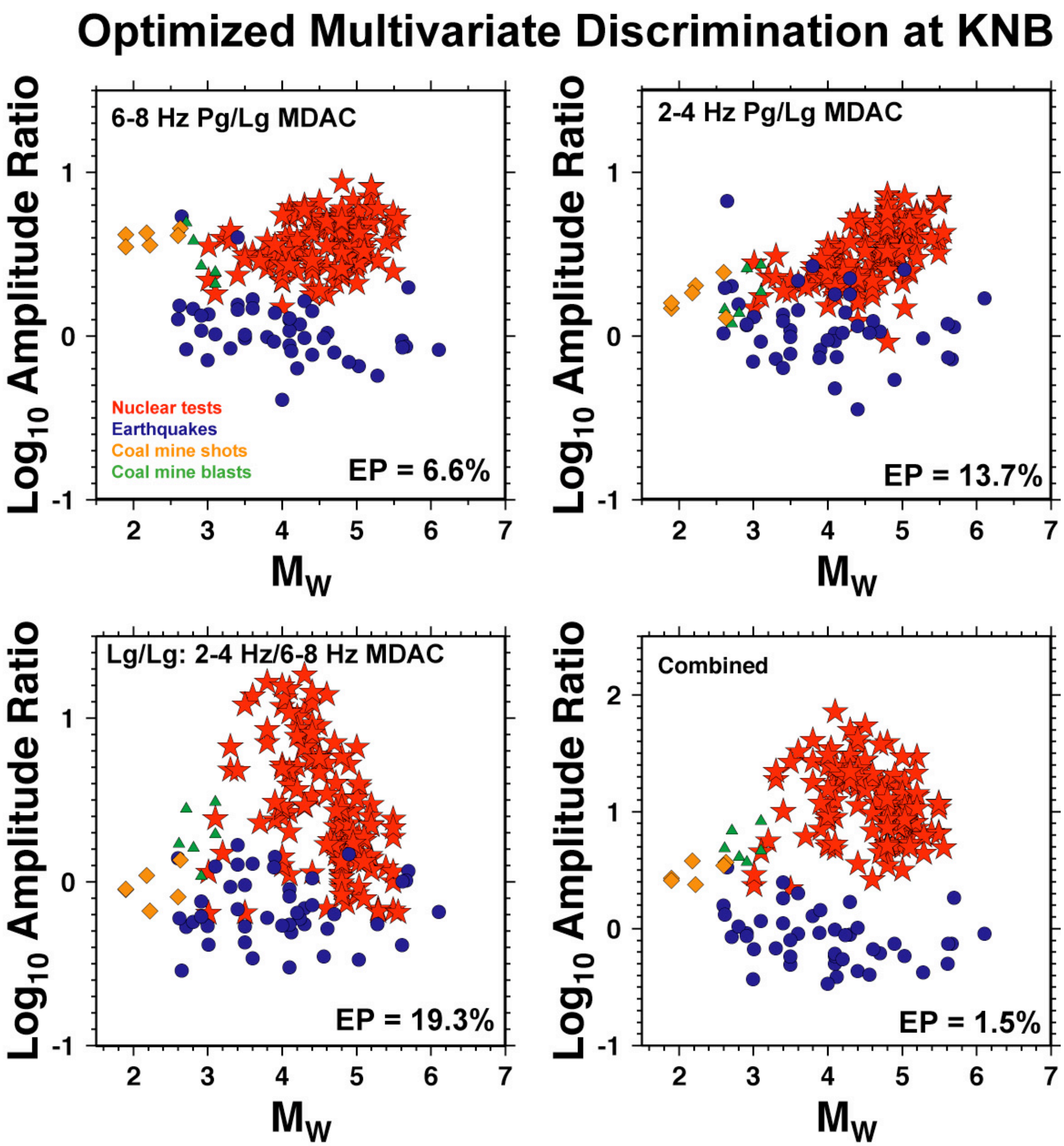

Figure 3. We show nuclear explosion discrimination from earthquake performance at station KNB for three different regional phase ratios after MDAC corrections were applied. In the lower right we combine these three ratios using an optimal set of weights determined using LDA to get a dramatic increase in performance. The combination is 0.71 (6–8 Hz Pg/Lg) + 0.88 (2–4 Hz Pg/Lg) + 0.57 (2–4/6–8 Hz Lg/Lg). This shows how optimally combining even mediocre discriminants can improve the performance of very good discriminants because new information always helps. Note that the mine shots track the nuclear tests for the P/S ratios but not for the low to high Lg ratio. Using the mine shots to obtain LDA weights would degrade the nuclear explosion discrimination performance.

A very interesting result demonstrated in Figure 3 is that by adding together several different mediocre discriminant measures using LDA coefficients we can greatly improve performance. In fact, using LDA we can always improve performance by adding another discriminant measure because it provides new information. In practice we have found that after combining about three to five different regional amplitude ratios using LDA, further improvement by adding additional measures is limited as the new measures do not provide much new information. We have found significant improvements by using LDA to combine measures for all the stations where we have done such analysis, covering a wide variety of regions. The challenge is that the best few discriminant measures and their optimal LDA coefficients vary from region to region in ways we do not yet fully understand, complicating transportability from region to region.

The two different types of chemical explosions (single contained shots versus ripple fired production blasts) show some interesting similarities and differences to the nuclear explosions. They all have similar high frequency P/S ratios as shown in Figure 3. However, in looking at regional seismic coda derived spectra (e.g., Mayeda and Walter, 1996; Mayeda et al., 2003) in Figure 4 we find the production shots have steeper spectral decay between 1–8 Hz and this accounts for the differences we see in the low to high frequency ratios. For this reason it is clear that doing an LDA analysis on the production chemical explosions and the earthquakes would produce different coefficients and discrimination performance. This is an area of research we are actively exploring.

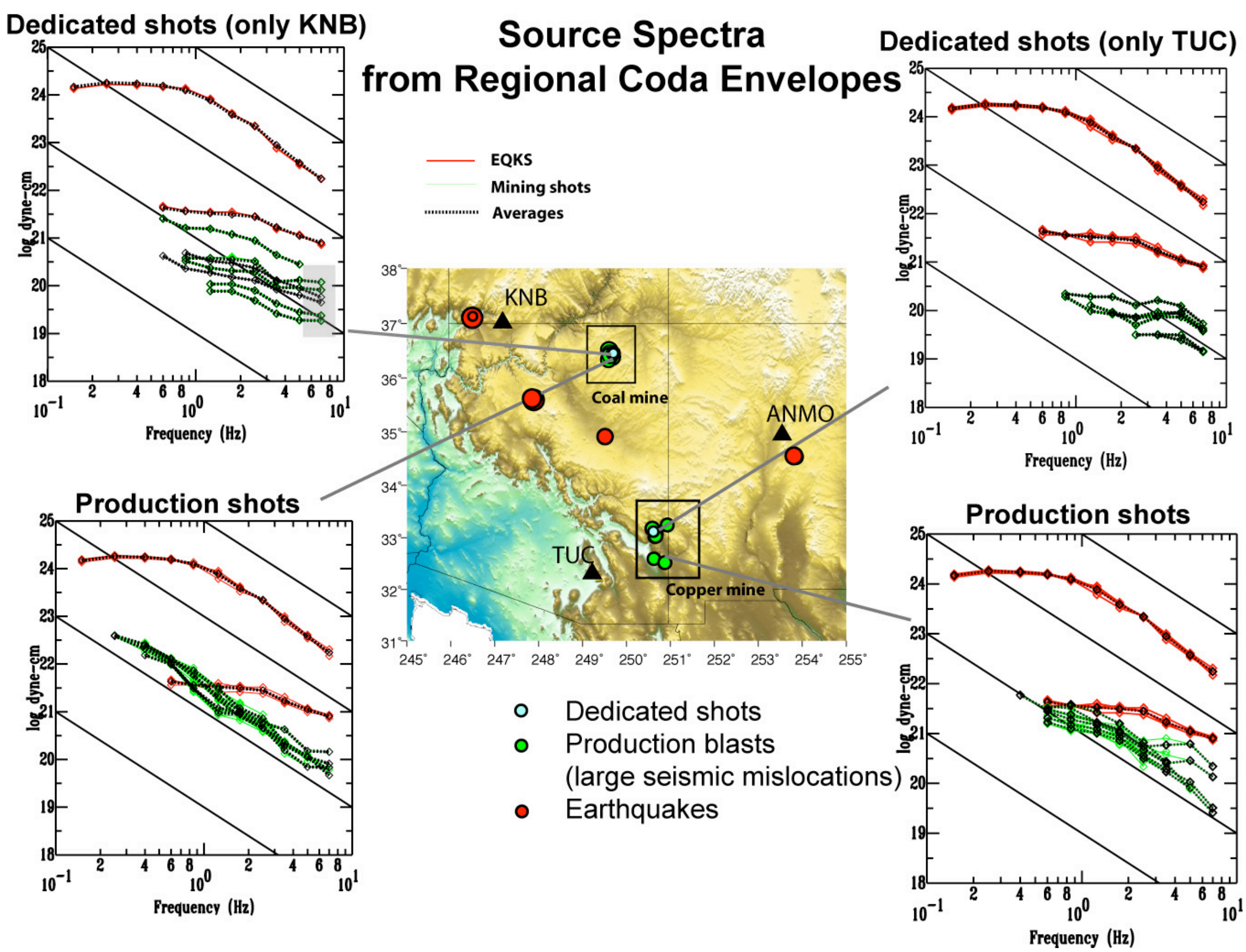

Figure 4. Regional coda envelope derived S-wave spectra of earthquakes (red) and dedicated single shot chemical explosions (light blue) and normal mine production explosions (green). The coda calibrations were done using the Colorado Plateau earthquakes shown. Left-hand side plots show the coal mine region single shots at top and ripple-fired production shots below. Similar plots for the copper mine region are shown on the right-hand side. Note that most of the ripple-fired shots have much steeper spectral falloff than the single shots.

Vertical Component Versus Three-Component

When three-component instruments are available the S-waves are often more clearly observed with larger amplitude on the horizontal components. Similarly, the P-waves often are more clearly observed with larger amplitude on the vertical component. For this reason a variety of studies have suggested that P/S ratio discrimination can be improved if all three components are used (e.g., Kim et al., 1997; Bowers et al., 2001). There are a variety of ways to make the measures, such as vertical P and horizontal S and rotation of the horizontals to radial and transverse and then using vertical and radial P and transverse S; but the simplest way is to average all three components together. In the prior study of Kim et al. (1997) the simple average did not perform much worse than the more sophisticated ways of separating out the P-SV and SH waveform components. Here we compare the performance of vertical and three-component measures on the western U.S. test set. We compare the vertical alone to the three-component average for 6–8 Hz Pg/Lg at station ELK in Figure 5 for a western U.S. set of earthquakes and nuclear explosions. The improvement in discrimination performance is modest. We are repeating these tests at other stations and using several different ways of doing the measures.

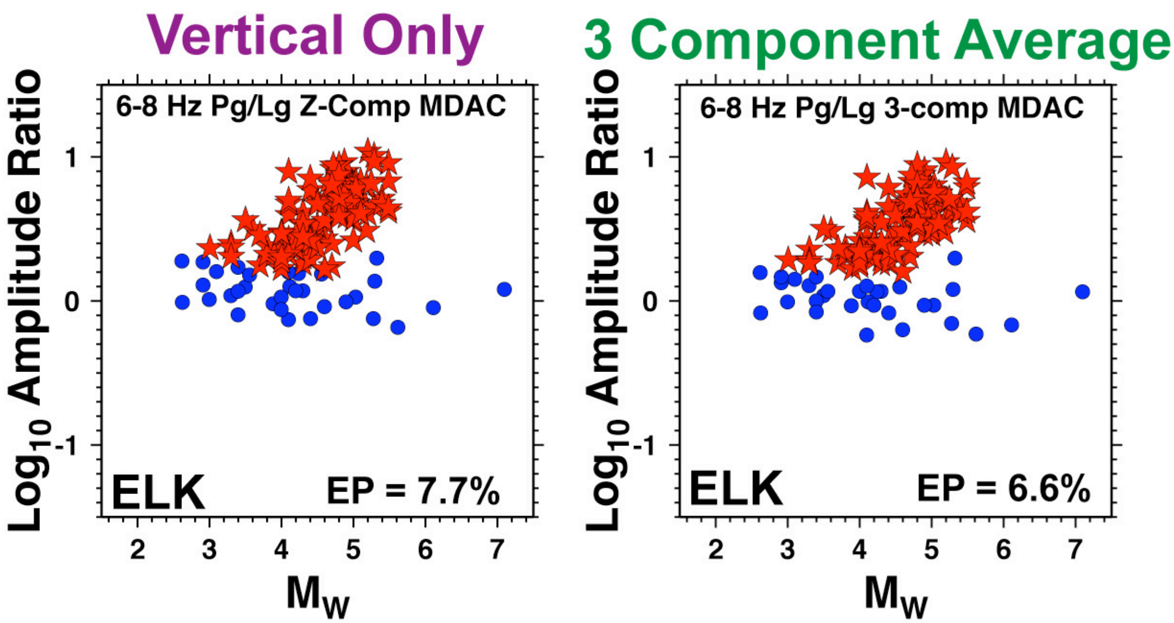

Figure 5. Here we use the same set of earthquakes and explosions at station ELK to test the improvement provided by measuring the 6–8 Hz Pg/Lg discriminant by averaging all three components (right-hand side) versus just using the vertical component (left-hand side). The improvement both in the scatter plot and as measured quantitatively by the equiprobable value is modest.

Maximum P and Maximum S

A complication in regional discrimination is the variation in P- and S-wave propagation with geophysical province. The dominantly observed regional high frequency S-wave can vary with path between Sn and Lg in a spatially complex way. Since the relative lack of high frequency S-waves is the signature of an explosion, failing to account for this could lead to misidentifying an earthquake as an explosion. The regional P phases Pn and Pg vary similarly with path and also with distance, with Pg sometimes being a strong phase at near regional distances but not far regional. One way to try and handle these issues is to correct for all four regional phases but choose the phase with the maximum amplitude. A variation on this strategy is to always use Pn but choose the maximum S phase (e.g., Bottone et al., 2002).

An important point in using maximum amplitude methods or any regional discrimination technique is that separate source, path and site effects still need to be determined for each of the four phases that might be used. The mantle phases Pn and Sn have geometrical spreading and attenuation parameters which tend to be quite different from crustal Pg and Lg phases, so one still needs to know which of the phases is giving the maximum amplitude to make the appropriate correction. For this reason the maximum amplitude techniques require the same amount of calibration effort as the traditional fixed phase ratio methods. The main operational advantage of maximum amplitude methods is simplicity in plotting all the results together rather than trying to form an optimal combination or discarding measurements when a phase is not present due to blockage or attenuation below the noise level.

The question of whether using a maximum amplitude measure will help or hurt discrimination performance depends on the style of variation of the regional phase amplitudes in the area. We show examples in Figure 6 of 6–8 Hz P/S ratios where the maximum (Pn, Pg) to maximum (Sn, Lg) measure would improve or worsen discrimination performance relative to a straight Pg/Lg measure. If the number of events where performance is hurt are small relative to the number where performance is helped, than the maximum amplitude measure will do better than the single ratio.

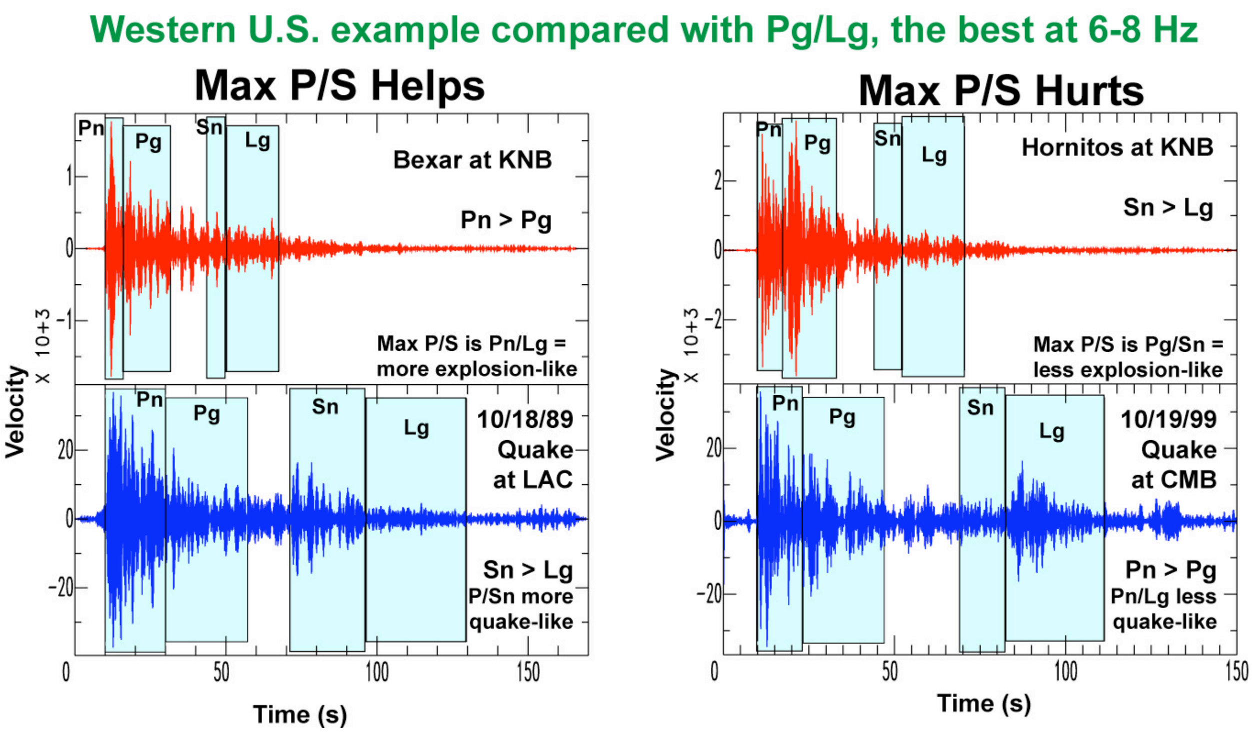

Figure 6. This figure demonstrates how using a maximum P from Pn or Pg and/or a maximum S from Sn and Lg can help or hurt the discrimination of event relative to just using a single ratio such as Pg/Lg. Here we show an example nuclear explosion and an example earthquake seismogram for each of the possible cases. The comparison is to 6-8 Hz Pg/Lg, which is the best 6-8 Hz P/S ratio. At the upper left we show an example where using the maximum P, in this case Pn, makes the explosion event more explosion-like, helping the discrimination process. In contrast, on the upper right we show a case where using the maximum S, in this case Sn, makes the explosion more earthquake-like. We show similar behavior for earthquakes in the bottom half of the figure.

Here we compare the discrimination performance of several different (max P)/(max S) measures to vertical, three-component and multivariate measures. Figure 7 shows the results in terms of the equiprobable value for several different 6–8 Hz P/S ratios combining the results from four stations in the western U.S. First it is clear that the MDAC corrections improve the discrimination performance of all the different P/S ratios. Second we note that the simple Pg/Lg ratio has better overall performance than the Pn/Maximum (Sn, Lg) technique. Given the uncertain results of which phases will have the maximum amplitude for a given event-station path, and the remarkable improvements available using LDA combinations demonstrated in Figure 3, we believe that in practice it makes the most sense to measure standard ratios of the major observed regional phases and then form LDA combinations. For example in the western U.S. we would expect that 6–8 Hz Pg/Lg, as the best single measure in combinations with other measures involving major observed phases such as 6–8 Hz Pn/Lg, will provide the best overall performance, as was demonstrated in Figure 3.

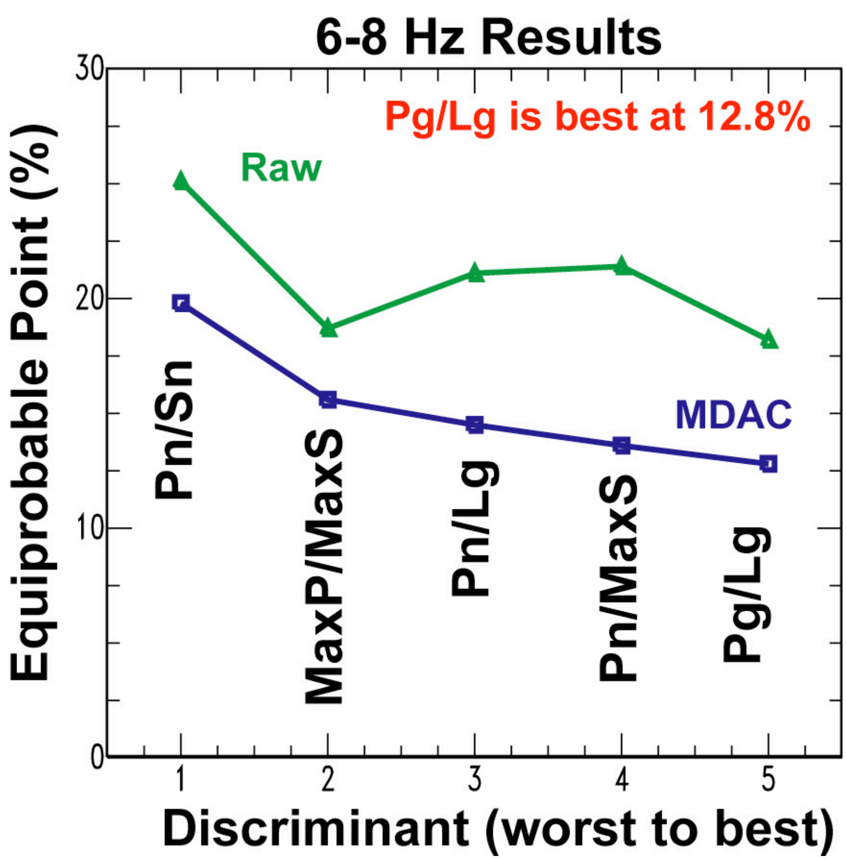

Figure 7. We combined data from four stations (CMB, ELK, KNB, LAC) to evaluate discriminant performance using equiprobable value as a metric. Equiprobable value is the level when the earthquake and explosion misidentification rates are equal, so the lower the number the better. We tried a variety of 6–8 Hz measures including taking the max (Pn, Pg) and max (Sn, Lg) and found Pg/Lg does the best. Note that the MDAC corrections significantly improve performance for all discriminants.

CONCLUSIONS AND RECOMMENDATIONS

Regional discrimination algorithms require calibration at each seismic station to be used for nuclear explosion monitoring. We have developed a revised MDAC procedure to remove source size and path effects from regional body-wave phases. This allows the comparison of any new regional events recorded at a calibrated station with all available reference data and models. This also facilitates the combination of individual measures to form multivariate discriminants that can have significantly better performance.

ACKNOWLEDGEMENTS

We thank Stan Ruppert, Terri Hauk and Jennifer O'Boyle for loading and maintaining the Lawrence Livermore National Laboratory (LLNL) database.

REFERENCES

Battone, S., M. D. Fisk and G. D. McCarter (2002), Regional Seismic-Event Characterization Using a Bayesian Formulation of Simple Kriging, *Bull. Seism. Soc. Am.* 92: 2277–2296.

Bowers, D., P. D. Marshall, and A. Douglas (2001), The Level of Deterrence Provided by Data from the SPITS Seismometer Array to Possible Violations of the Comprehensive Test Ban in the Novaya Zemlya Region, *Geophys, J. Int.* 146: 425–438.

Hartse, H., S. R. Taylor, W. S. Phillips, and G. E. Randall (1997), A Preliminary Study of Regional Seismic Discrimination in Central Asia with an Emphasis on Western China*, Bull. Seism. Soc. Am*. 87: 551–568.

Kim, W.-Y., V Aharonian, A. L. Lerner-Lam, and P. G. Richards (1997), Discrimination of Earthquakes and Explosions in Southern Russia Using Regional High Frequency Three-Component Data from the IRIS/JSP Caucasus Network, *Bull. Seism. Soc. Am.* 87: 569–88.

Mayeda, K. M. and W. R. Walter, (1996), Moment, Energy, Stress Drop and Source Spectra of Western U.S. Earthquakes from Regional Coda Envelopes, *J. Geophys. Res*. 101: 11,195–11,208.

Mayeda, K., A. Hofstetter, J. L. O'Boyle, and W. R. Walter (2003). Stable and Transportable Regional Coda Magnitudes Based on Coda-Derived Moment Rate Spectra, *Bull. Seism. Soc. Am.* 93: 224–239.

Rodgers, A. J. and W. R. Walter (2002), Seismic Discrimination of the May 11, 1998, Indian Nuclear Test with Short-Period Regional Data from Station NIL (Nilore, Pakistan), *Pure Appl. Geophys.* 159: 679–700.

Springer, D. L., G. A. Pawlowski, J. L. Ricca, R. F. Rohrer, and D. K. Smith (2002), Seismic Source Summary for All U.S. Below-Surface Nuclear Explosions, *Bull. Seism. Soc. Am.* 92: 1806.

Taylor, S. (1996), Analysis of High-Frequency Pg/Lg Ratios from NTS Explosions and Western U.S. Earthquakes, *Bull. Seism. Soc. Am.* 86: 1042–1053.

Taylor, S., A. Velasco, H. Hartse, W. S. Philips, W. R. Walter, and A. Rodgers (2002), Amplitude Corrections for Regional Discrimination, *Pure. App. Geophys*. 159: 623–650.

Walter, W. R., K. Mayeda, and H. J. Patton (1995), Phase and Spectral Ratio Discrimination between NTS Earthquakes and Explosions Part 1: Empirical Observations, *Bull. Seism. Soc. Am.* 85: 1050–1067.

Walter, W. R. and S. R. Taylor (2002), A Revised Magnitude and Distance Amplitude Correction (MDAC2) Procedure for Regional Seismic Discriminants, Lawrence Livermore National Laboratory report UCRL-ID-146882.

Walter, W. R., K. D. Smith, J. L. O'Boyle, T. F. Hauk, F. Ryall, S. D. Ruppert, S. C. Myers, M. Anderson, and D. A. Dodge (2003), Improving the Fundamental Understanding of Regional Seismic Signals with a Unique Western

United States Dataset, in *Proceedings of the 25th Seismic Research Review—Nuclear Explosion Monitoring: Building the Knowledge Base*, LA-UR-03-6029, 1: pp. 486–494.