ABSTRACT

Historically, event characterization has focused on the task of separating earthquakes and explosions. However, with the increasing availability of high-quality regional seismic data, including seismic array data, there is a need to characterize smaller events (mb 3.5 and below). The present challenge of seismic event identification therefore includes the task of identifying mining explosions, which fall within this lower range of magnitudes. The increasing availability of infrasound data, and in particular the advent of the International Monitoring System (IMS) infrasound array, affords new opportunities in the combined use of seismic and acoustic data for the discrimination of mining explosions. As detailed in last year’s report, we have assembled a comprehensive database of earthquakes and mining explosions in three mining regions in the US and in the Altai-Sayan mining region in Russia. By establishing a good working relationship with a large mine operator in Wyoming, we have obtained a detailed ground-truth dataset of explosions in the Powder River basin. By focusing on this region initially, we have developed and tested three types of discrimination algorithms based on high-frequency phase amplitude measurements, spectral content and infrasound.

An algorithm for identifying delay-fired mining explosions by exploiting the time-independent spectral modulations produced by such explosions has been developed. This algorithm has been enhanced to make it applicable to seismic arrays. By applying the algorithm to the dataset of earthquakes and explosions at the Pinedale seismic array (PDAR) in Wyoming, it is shown that the use of array data significantly improves the discriminant. In a blind test, the method successfully identifies 97% of the events out of a group of 76 earthquakes and large delay-fired mining explosions (cast blasts).

The potential of infrasound for contributing to the event identification scheme is assessed using data from the PDIAR infrasound array in Wyoming, which is co-located with the PDAR array. Acoustic detections at the PDIAR array, generated using an array-based signal detection algorithm, are compared with predicted arrival times of cast blast acoustic signals. Predicted arrival times are based on ground-truth information from the mine, which have been adjusted by seismic observations. For one year of data, over fifty percent of cast blasts are associated with detections during July and August, with significantly fewer detections at other times of year. This result suggests that the use of infrasound has good potential for contributing to our identification scheme, although the potential is possibly seasonally dependent.

Amplitude ratios at PDAR have been considered for separating delay-fired mining blasts from nearby earthquakes (within 1,000 km of the Wyoming coal mine). Historically, in other regions, high-frequency P/S ratios have been successful at discriminating these types of events; however, in this region that discriminant fails. We believe this is due to a combination of both path effects (as described in Goforth and Zhou, 2006; Zhou and Stump, 2004) and variability in the source. In order to discover more about the role of source variability in successfully discriminating mining events, current work is designed to minimize the role of path effects by developing regional phase attenuation tomography models of the western US.

In order to assess the portability of the various discriminants tested in Wyoming to different regions, the above techniques are being applied to the dataset acquired for the Altai-Sayan region in Russia. Initial results on this effort are presented.
OBJECTIVES

Enhancements of Comprehensive Ground Truth Database

In order to relate discriminant performance to physical source processes, we have obtained more detailed ground truth information on a number of shots recorded at the PDAR. We have also obtained ground truth information on additional shots for testing the use of infrasound as a potential discriminant.

Further Development of Discriminants and Application to the Subset of Data

We have continued the development of the three discriminants discussed in Arrowsmith et al. (2005): time-varying spectral estimation, infrasound and amplitude ratios. A time-varying spectral estimation method is developed for application to single-component or three-component seismographs, and for either single stations or arrays. We have begun the development of an infrasound discriminant by evaluating the probability of observing infrasound signals from ground truth mining explosions, and how this varies with time-of-year. We also give progress on the amplitude ratio discriminant, in which we are working to refine the path corrections applied to the amplitude measurements via developing regional phase attenuation models of the western US. As outlined in Arrowsmith et al. (2005), we have developed a high-quality training dataset for application and testing of discriminants. We present the results of applying each of the three discriminants to the training dataset.

RESEARCH ACCOMPLISHED

Enhancements of Comprehensive Ground Truth Database

Arrowsmith et al. (2005) outline the development of a comprehensive ground truth database using earthquakes and mining events in three mining regions of the US (the Powder River Basin, southeast Arizona and northwestern Minnesota) and in the Altai-Sayan region in Russia. Due to the high-quality ground truth and close ties with local mining operators, the Powder River Basin has been chosen as a training dataset for the application and testing of regional discriminants. We have improved the training dataset outlined in Arrowsmith et al. (2005) by obtaining more detailed information on select mining explosions and by obtaining ground truth on additional mine shots.

The new ground truth information includes detailed shot-time information for a number of select cast blasts (i.e., number of shots in blast sequence, inter-shot delay times, inter-row delay times and depths of charges). This new information allows us to model regional waveforms from these complicated blast sequences and will be useful in further improving and testing the regional discriminants. The detailed ground truth information should allow us to relate discriminant performance to the source physics.

The development and testing of a regional discriminant based on infrasound data requires a ground truth dataset of mining explosions and earthquakes that covers every season, in order to properly assess the effect of seasonal changes of the atmosphere. The training dataset described in Arrowsmith et al. (2005) has been extended to include additional mining explosions and earthquakes in order to ensure that each season is adequately sampled.

Further Development of Discriminants and Application to the Subset of Data

Time-Varying Spectral Estimation Discriminant

We have developed a methodology for identifying delay-fired mining explosions that utilizes time-independent spectral modulations. The methodology is described in detail in Arrowsmith et al. (2006a) and the extension of the methodology for seismic arrays is described in Arrowsmith et al. (2006b). The methodology is applied to the training dataset described above (which consists of 98 mining explosions and 43 earthquakes), and shown in Figure 1. Briefly, rather than discriminating between two or more possible event classes, the technique can identify simply whether or not an event is a delay-fired mining explosion. The technique can provide seven separate discriminants based on the binary spectrograms of seismic events recorded at a three-component seismograph. Examples of binary spectrograms for a typical earthquake and cast blast in the training dataset are shown in Figure 2.
The seven discriminants include the cepstral mean, the three values of cross-correlation of the binary spectrograms (evaluated on all three components) and the three values of autocorrelation of the binary spectrograms (on each component separately). To calculate the cepstral mean we first compute the two-dimensional Fourier transform of the binary spectrogram matrices, providing a 2D cepstral matrix. It is then straightforward to isolate energy periodic in frequency and independent of time, yielding 1D cepstra. The separate 1D cepstra from each of the individual components are then stacked and averaged over the first few cepstral coefficients. We have found the cepstral mean to be a more effective discriminant than simply taking the value of the maximum cepstral peak.

The seven discriminants exploit similar properties of the binary spectrograms: the regular pattern of spectral scalloping, its time independence and the correlation on all three components. Therefore, the discriminants are not completely independent of each other, providing some redundant information about the nature of the source. We perform a simple feature selection procedure in order to ensure that the addition of each discriminant is significant, and therefore that each discriminant contributes to the overall separation between delay-fired mining explosions and the remaining event population. The feature selection procedure also identifies the best combinations of d discriminants to use (where 1 ≤ d ≤ 7). We start by identifying the best single discriminant (using the Mahalanobis distance as a measure of discriminant quality). Next, we identify the best combination of d discriminants, where d = 2, 3, …, 7. At each step, we calculate an F statistic as a guide to determine whether the addition of the new discriminant is significant. The F-statistic (Hand, 1981; Taylor and Hartse, 1997) is given by the following:

\[
F = \frac{(n-d-1)n_1n_2(D_{d'}^2 - D_d^2)}{(d-d')(n(n-2)+n_1n_2D_{d'}^2)},
\]

(1)
where \( n \) is the total number of events, \( n_1 \) is the number of earthquakes, \( n_2 \) is the number of delay-fired mine blasts, \( d \) is the total number of discriminants and \( d' \) is the number of discriminants for a particular iteration. \( D_d^2 - D_{d'}^2 \) is the difference in the Mahalanobis distance on the subset of \( d \) and \( d' \) discriminants. \( F \) is then compared with the tabulated \( F \) distribution for \((d-d')\) and \((n-d-1)\) degrees of freedom to examine whether the addition of the new discriminant is significant. The top panel in Figure 3 indicates which discriminant or discriminants are selected after each step in which a new discriminant is added. In the discussion that follows, and in Figure 3, we use \( Z, N \) and \( E \) to refer to the vertical, north and east component recordings respectively. In the figure, “auto” denotes the autocorrelation of a binary spectrogram on an individual component, “cc” denotes the cross-correlation between binary spectrograms recorded on two components, and “Mean cep” denotes the mean cepstrum. The corresponding Mahalanobis distance and \( F \)-statistic are also shown for each step in order to illustrate whether the addition of a new discriminant is significant. The overall Mahalanobis distance between the earthquakes and cast blasts increases as each new discriminant is added, however the final two discriminants \([cc(Z&N) \text{ and } cc(Z&E)]\) do not result in a significant increase in the Mahalanobis distance. This observation fits with the \( F \)-statistics that show that five of the seven discriminants improve the discrimination significantly but that the poorest two discriminants do not. Therefore, \( cc(Z&N) \) and \( cc(Z&E) \) could have been removed from the discrimination procedure without degrading the overall performance. However, the inclusion of the two discriminants does not degrade the Mahalanobis distance between the two groups and we have included them in our results.

![Figure 2](image-url)

**Figure 2.** Input seismic waveforms with corresponding spectrograms and binary spectrograms for two example events. For each event, the portion of the waveform used in the computation of the spectrograms is shown. The first event, an example mine blast, exhibits clear spectral modulations that are time independent. The second event, which is a typical earthquake, shows no evidence of spectral modulations.

For a single-component seismogram, only two separate discriminants are computed (cepstral mean and autocorrelation on the vertical component). For any given station, the algorithm can effectively be tuned with a
reference set of events, in order to maximize the success rate in identifying delay-fired and non-delay-fired events. This is accomplished by searching for the optimum values of the free parameters that maximize the separation between the two classes of reference events using the mean Mahalanobis distance. The free parameters in the methodology are the spectrogram duration \((w)\), the averaging windows used in converting the spectrograms into binary form \((sp1 \text{ and } sp2)\), and the window over which the mean cepstral value is evaluated \((\text{cep})\). The results of the optimization procedure applied to the training dataset show that the performance of this technique is highly dependent on the values of the input parameters (Figure 4). An inherent assumption in this approach is that the nature of delay-fired blasting in a particular area does not vary significantly, and that delay times between individual shots and rows in blasting arrays are similar. Otherwise it could be possible that a single optimum set of input parameters would not exist, as the optimum parameters would be different for each source.

Figure 3. Results of the seven-step feature selection procedure. Each step represents the optimum combination of \(d\) discriminants \((1 \leq d \leq 7)\). Top panel: Dots indicate the optimum combination of discriminants for each search step. Middle panel: Mahalanobis distance between the earthquake and cast blast groups calculated for each search step (using the appropriate number of discriminants). Bottom panel: Calculated F-statistic from Equation 4 (solid line) and corresponding tabulated value of the F distribution for \((d-d')\) and \((n-d-1)\) degrees of freedom (dashed line).

We have developed a procedure to correct for the effect of pre-event noise, which can exhibit low amplitude spectral modulations. First, we compute a spectrogram of the pre-event noise and evaluate the mean scaled logarithm of the pre-event spectrogram in each frequency band \((\bar{x}_{f,\text{pre}})\). Next, we compute the following parameter for each pixel in the spectrogram of the signal:

\[
\alpha = \left| x_{tf} - \bar{x}_{f,\text{pre}} \right|
\]  

where \(x_{tf}\) is the scaled logarithm of the signal spectrogram at time \(t\) and frequency \(f\). We then randomize the pixels in the binary spectrogram where \(\alpha\) is less than a threshold value.
For array data, we evaluate a spectrogram for each element of the array separately, and stack the spectrograms from all the array elements. Since each spectrogram is computed for a time window that begins at the picked first arrival at each array element (with duration $w$ set as a free parameter), they can be stacked without the need for aligning based on the speed and direction of the incoming wavefront. The results of applying the methodology to 43 earthquakes and 98 delay-fired mining explosions (Figure 1) are shown in Figure 5. The cast overburden shots separate well from the earthquakes, although the smaller shot types do not (for description of shot types see Table 1). The separation improves if (1) the full PDAR array is used (rather than an individual instrument) and (2) the noise correction is applied to remove pre-event noise. In a drop-one event test, this method can identify 97.4% of the earthquakes and cast blasts (i.e., 74 out of the 76 events are identified correctly). The two events that are classified incorrectly (Figure 1) appear to be failures of this methodology. We hope that by combining this discrimination procedure with the other regional discriminants that these misidentified events would be successfully identified.

Figure 4. Cross-sections of the objective function (Mahalanobis distance) that show the optimal input parameters. Left panel: Mahalanobis distance as a function of $cep$ and $w$ with $sp1$ and $sp2$ held at their optimal values. Right panel: Mahalanobis distance as a function of $sp1$ and $sp2$ with $cep$ and $w$ held at their optimal values. Note the requirement that $sp2 > sp1$. 
Figure 5. Values of the two discriminants (cepstral mean—top, autocorrelation—bottom) obtained for each event using a single station only (left), the full array stack (middle) and the full array stack with the noise correction applied (right). Black circles represent earthquakes, plus signs represent parting shots, crosses represent coal extraction shots in the upper coal seam, open squares represent coal extraction shots in the main coal seam, open diamonds represent TS overburden blasts and open stars represent cast overburden shots (refer to Table 1 for more information on shot types). The solid lines show the extreme upper and lower bounds of discriminants evaluated from 50 samples of stacked pre-event noise spectrograms (using all 15 array elements). The dashed lines show the equivalent upper and lower bounds for a single array element (i.e., without stacking).

Table 1. Blast types for a large coal mine in the Powder River Basin, Wyoming

<table>
<thead>
<tr>
<th>Blast Type</th>
<th>Description</th>
<th>Min Yield (lb)</th>
<th>Max Yield (lb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cast Overburden</td>
<td>Overburden is cast and removed into adjacent empty pit</td>
<td>300,000</td>
<td>2,500,000</td>
</tr>
<tr>
<td>TS Overburden</td>
<td>Blasts in overburden material to be excavated by shovels loading into trucks</td>
<td>100,000</td>
<td>600,000</td>
</tr>
<tr>
<td>Coal - Main</td>
<td>Blasts in the main coal seam, which is 60–70 ft in thickness</td>
<td>20,000</td>
<td>200,000</td>
</tr>
<tr>
<td>Coal - Upper</td>
<td>Blasts in the upper coal seam, which is 10 ft in thickness</td>
<td>2,000</td>
<td>10,000</td>
</tr>
<tr>
<td>Parting</td>
<td>Blasts of waste material layer between the upper and lower coal seam, ranging from 0 to 40 ft in thickness</td>
<td>200</td>
<td>500</td>
</tr>
</tbody>
</table>
Infrasound Discriminant

We are currently developing a discriminant based on the presence (or absence) of a regional infrasound signal. The premise of the discriminant is surface sources generate more infrasound than equivalent sources located at depth. Therefore, infrasound has the potential to be used as an effective depth-discriminant, and the presence of an infrasound signal associated with an event would support the hypothesis that the event was a mining explosion rather than an earthquake. The first step in developing this discriminant is to assess the relative probabilities that ground-truth mining explosions or earthquakes are associated with infrasound signals. Since atmospheric temperatures and winds (which vary seasonally) affect infrasound propagation, it is important that these relative probabilities are assessed as a function of time of year. To this end, we have assembled a dataset of ground truth mining explosions that cover every month of the year.

We have developed a simple automatic procedure to search for an associated infrasound signal from each mining explosion based on signal arrival time and back azimuth. We have obtained waveform data from the PDIAR infrasound array in Wyoming (Figure 1). Signal detections are obtained using the Progressive Multichannel Cross-Correlation procedure (PMCC) (Cansi, 1995), which is an array-based algorithm that uses cross-correlation of the waveforms from the separate array elements to determine signal detections (with associated back azimuths, speeds, etc.). For each ground-truth mining explosion, the predicted arrival time (using a celerity of 300 m/s) and true great-circle back azimuth are calculated. We then search for detections from the PMCC algorithm with an arrival time and back azimuth that match the predictions (within an appropriate time-window and back azimuth range that reflect deviations from the assumed celerity of 300 m/s and true back azimuth). Figure 6 shows the relationship between the actual numbers of cast blasts (for which we have been able to obtain waveform data at PDIAR) and the corresponding number of associated detections (using an allowed time-offset of +10 minutes and azimuth-offset of –5 degrees). In July and August, greater than 50% of explosions are associated with detections, whereas during other months there are no detections (except November, when there are unusually many ground truth explosions). This fits with our knowledge of stratospheric wind directions, which assist stratospheric returns from east to west during the summer, but inhibit returns during winter months. This indicates that the use of infrasound as a discriminant will likely be seasonally dependent.

The next steps in the development of this discriminant will be (1) to further study detections that are associated with mining explosions in order to improve confidence in the tie-in and (2) to repeat this analysis for the earthquakes in the training dataset. If the results obtained indicate that there is potential for infrasound as a discriminant, we will work on developing an automatic infrasound discriminant and evaluating procedures for integrating the discriminant with the results from other regional discriminants.

Figure 6. Number of explosions with corresponding waveform data (blue bars) and number of associated detections with ground truth mining explosions (red bars) as a function of month of the year.
Amplitude Ratio Discriminant

Previously, we tested traditionally well-performing discriminants, such as high-frequency $P_g/L_g$, on regional phase amplitudes from the events shown in Figure 1 recorded at the PDAR. The phase amplitudes were corrected for path and source effects using the MDAC methodology (Walter and Taylor, 2002). We chose a simple 1D path correction based on an average Q model along the path between the mine and the Pinedale array. The results (Figure 7) indicate that this is a poorly performing discriminant, evidenced by variability in $P_g$ and $L_g$ amplitudes for similar event types at high frequencies.

In order to investigate why the discriminant performs as it does, we need to first determine if the 1D path correction is adequate for this region so that we can rule out underlying path effects as the cause for the poor discriminant performance. Previous studies in this geographic region have shown that path effects greatly influence regional waveforms (Zhou and Stump, 2004; Goforth and Zhou, 2006), indicating that simple 1D path corrections may in fact be inadequate. Therefore, we are developing 2D path corrections via construction of attenuation tomography maps of the western United States for the regional phases. Such maps can be imported into the MDAC processing scheme and allow for path corrections based on variable Q along a given path.

We are in the process of our first round of data collection, which includes gathering earthquake information for earthquakes with magnitudes $> 4.0$ in the region [120 W 104W 30N 50N] recorded at broadband stations in the same region. This effort has amounted in 488 earthquakes recorded about 35 stations. We have also incorporated earthquake data from the western US database developed by Walter et al. (2003). Figure 8 summarizes the data we have collected thus far. As our data coverage is limited in the Intermountain and Great Plains regions, our next round of data collection efforts will focus on smaller magnitude events recorded at stations in the northern and middle Rocky Mountains and Great Plains. The processing scheme on this data includes making regional phase picks and amplitude measurements, using coda magnitude methodology to obtain reliable moments, and finally, to apply inversion techniques similar to those utilized by Lay et al. (2006).

Figure 7. (Left) MDAC corrected $P_g/L_g$ (6-8 Hz) amplitude ratio for mining events (colored stars) and earthquakes (yellow circles) as a function of $M_w$ and distance. (Right) Selected waveforms for two cast blasts (red) and three earthquakes (yellow) with regional phase window onset times marked.
Figure 8. Summary of data gathered in the first round of collection efforts for the attenuation tomography study of the western US. The top left figure shows stations, where red stars represent new stations utilized and blue stars represent those incorporated from the Walter et al. (2003) dataset. Red stars outlined in blue are overlapping stations. The top right figure shows events used, with the same color scheme as the stations plot. The bottom left figure illustrates path coverage for the new data added for this study. The bottom right figure shows path coverage for the Walter et al. (2003) dataset. In all figures, the black circle indicates the location of the Wyoming coal mine referred to in this paper.

CONCLUSIONS AND RECOMMENDATIONS

This report builds on the work discussed in Arrowsmith et al. (2005) in two primary ways: (1) enhancement of ground truth testing dataset and (2) development of discriminants and application to testing dataset. Under the first point we have obtained the following: (a) improved ground truth on select cast blasts for the purpose of relating discriminant performance to source physics, and (b) ground truth on additional shots for testing the infrasound discriminant. Under the second point, we have outlined significant progress in developing three regional discriminants: time-varying spectral estimation, infrasound and amplitude ratios. A new time-varying spectral estimation is discussed, which is applicable to seismic arrays and successfully identifies 97.4% of events from a testing dataset of 43 earthquakes and 33 cast blasts. We outline the initial work on developing an infrasound discriminant, whereby we have begun evaluating the probability of detecting infrasound signals from cast blasts and found a strong seasonal dependence, with over 50% of events being associated with detections in July and August. We also give progress on the amplitude ratio discriminant; because significant variability was seen in regional waveforms, we are working to refine the path corrections applied to the amplitude measurements via developing regional phase attenuation models of the western US. Future work will focus on three main issues: (1) continued development of the three discriminants outlined in this report, (2) development of additional discriminants (e.g., correlation, mb/Ms), and (3) development of a scheme for integrating all the regional discriminants into a comprehensive discrimination package.
ACKNOWLEDGEMENTS

We thank Steve Taylor of Rocky Mountain Geophysics for his help with the feature selection procedure; Steve Beil of ArchCoal for providing ground truth shot time data; and Michele Kelley at AFTAC for her help in providing us with PDIAR data.

REFERENCES


