Abstract. High-frequency regional records from small earthquakes (1.3 < magnitude < 4), and comparable magnitude explosions, are analyzed to find a reliable seismic discriminant in the eastern U. S. Over 500 digital, vertical-component seismograms recorded by the New York State Seismic Network in the distance ranges 10 to 610 km are used. Mean P/Lg spectral ratios in the band 1 - 25 Hz are about 0.5 and 1.25 for earthquakes and explosions, respectively, in the eastern United States. We find that the high-frequency P/Lg spectral amplitude ratio in the frequency band 5 - 25 Hz is a reliable and robust discriminant for classifying these events. A linear discriminant function analysis indicates that the P/Lg spectral amplitude ratio method provides discrimination power with a total misclassification probability of about 1%. Single-hole instantaneous explosions and ripple-fired quarry blasts have somewhat different P/Lg spectral ratios, but as a group are distinctly different from earthquakes.

Introduction

The discrimination of small earthquakes from large chemical explosions (from mines and quarries) based on seismic signals recorded at regional distances (10 - 1000 km) is an important issue facing numerous regional seismic networks. The seismic discrimination problem becomes especially severe in an area with low seismicity such as the eastern United States. For instance, in New York and adjacent states covered by the New York State Seismic Network (NYSSN; Figure 1), a few small earthquakes (magnitude < 4) occur per month, but about 20 chemical explosions per day may be recorded during weekdays.

Much previous work on seismic discrimination has focused on separating underground nuclear explosions from earthquakes in the context of a future comprehensive test ban [e.g., Evernden et al., 1986; Taylor et al., 1989]. In most earlier studies, data available for discrimination analyses were limited to frequencies below 10 Hz, and previous work on regional signals from earthquakes and explosions in the western United States indicated that P and S waves from explosions have higher frequency content than signals from earthquakes [e.g., Bennett & Murphy, 1986; Taylor et al., 1989; Chael, 1988]. These western U. S. results are contrary to our observations and to the claim by Evernden et al. [1986] that regional signals from explosions should show higher frequency content than those from earthquakes. Regional discriminants must be evaluated on a regional basis, however we presume the method reported here would work to discriminate small nuclear explosions from small earthquakes, in regions that support high-frequency signals.

Our typical observations of high-frequency (1-35 Hz) regional signals from earthquakes and explosions in the eastern U. S. may be summarized as: 1) P waves from single-hole explosions have much higher frequency content than S waves; 2) S waves from earthquakes have higher frequency content than P waves, as much as by a factor of about two; 3) P and S waves from ripple-fired quarry blasts have similar frequency content and show frequency banding due to spectral modulation; and 4) records from explosions often show a strong Rg phase out to about 100 km. [We use P and Pg interchangeably to indicate P waves with a group velocity around 6 km/s; and likewise S and Lg are used to indicate shear waves with group velocity about 3.5 km/s, without further classification.] Typical vertical-component digital seismograms from an earthquake and a single-hole explosion are shown in Figure 2. Unfiltered seismograms (top traces) support the differences 1) and 2) stated above. When the seismograms are filtered to show the conventional low-frequency band (1-10 Hz, middle traces), the spectral characteristics appear similar to the unfiltered signals. But bandpass filtered, 10-25 Hz, high-frequency seismograms (bottom traces) accentuate the differences between the earthquake and the explosion. Such high-frequency seismograms suggest that the PgLg spectral amplitude ratio can be made the basis of a useful discriminant. In the following sections, we report our measurements of the PgLgLg spectral amplitude ratio, using high-frequency, vertical-component digital seismograms from earthquakes and chemical explosions recorded by the NYSSN (Figure 1); and we evaluate its discrimination capability.

Data

We first present the analysis of digital seismograms from 30 explosions (16 quarry blasts and 14 single-hole explosions) and 30 earthquakes to obtain a specific discriminant, which later is applied to other events. Most of the data are from felt events in the magnitude range 1.3 to 3.5 sampling a wide range of propagation paths within the eastern U. S. (Figure 1). All quarry blasts are from eight known quarry sites. Only blasts for which quarry personnel have provided us with pertinent information (delay times, number of shot holes, maximum charge/delay period etc.) are included. The maximum charge-weight per delay period (within 8 milliseconds) of the ripple-fired blasts ranged from 0.1 to 27 tons. All single-hole explosions are from a seismic refraction experiment and had 1 - 2 ton charge weights. Distance ranges of the data are 5 to 610 km with means of 149 (±135) km and 129 (±107) km for earthquakes and explosions, respectively. PgLgLg signals are windowed with a Gaussian weighting function centered at group velocities around 5.9

1Now at Incorporated Research Institutions for Seismology.

Copyright 1993 by the American Geophysical Union.

Paper number 93GL00367
0094-8554/93/93GL-00367S03.00
Fig. 1. Locations of earthquakes (shaded circles), single-hole explosions (pluses), quarry blasts (open triangles), NYSSN stations (diamonds) and earthquakes used for testing the discriminant function (shaded squares). Size of the circles are proportional to the magnitude of the earthquakes.

Fig. 2. Typical vertical-component records from an earthquake and an explosion. Unfiltered (top), low-frequency bandpass filtered (middle) and high-frequency bandpass filtered (bottom) traces are plotted. Example Gaussian time windows used for \( P_g \) and \( L_g \) spectral amplitude measurements are shown on the unfiltered earthquake trace.

Fig. 3. Network averaged \( \log_{10}(P_g/L_g) \) spectral amplitude ratios at seven discrete frequency points used in discrimination analysis are plotted for earthquakes (circles), single-hole explosions (pluses) and quarry blasts (triangles).

Digital seismograms from the NYSSN are sampled at 100 samples/sec and instrument responses of the stations are nearly flat to ground velocity in the frequency band 1 - 25 Hz. The \( P_g \) and \( L_g \) signals, weighted by the Gaussian functions, are fast Fourier transformed. The resulting amplitude spectra are smoothed with another Gaussian function having \( \sigma = 2.5 \) Hz and are re-sampled at every 5 Hz interval from 5 to 35 Hz. Spectral amplitudes for the frequency band 1 to 10 Hz, with 1 Hz interval, are obtained using \( \sigma = 0.5 \) Hz. The S/N ratio becomes less than 2 at about 25 Hz for records from distant events (\( \Delta > 300 \) km). Network averaged \( \log_{10}(P_g/L_g) \) ratios are obtained for each event by averaging the discrete frequency values of the ratio at each station (Figure 3). Earthquake and explosion populations are well separated in the frequency band 10 to 25 Hz. At 5 Hz and at frequencies higher than 30 Hz, there is some overlap. Thus, the frequency band 5 - 25 Hz is the most reliable and will be used in the following discrimination analysis.

The Linear Discriminant Function

To test the discriminant power of the high-frequency \( P_g/L_g \) spectral ratio, we performed multivariate discriminant analysis on \( \log_{10}(P_g/L_g) \) spectral ratios for the data set of earthquakes and explosions. Each training group of 30 events (i.e., 30 explosions, 30 earthquakes) is described by a matrix of 5 rows \([\log_{10}(P_g/L_g) \text{ at } 5 \text{ Hz frequency intervals from } 5 \text{ to } 25 \text{ Hz}]\) and 30 columns. Preliminary analysis suggested somewhat higher \( P_g/L_g \) ratios from single-hole shots than for quarry blasts, but we merged both data sets into one explosion group because the differences between them was much less than their differences from earthquakes.

We introduce \( f_{Ex}(r) \) and \( f_{Eq}(r) \) as the probability densities of the two types of events, with \( r \) as a column vector representing the \( \log \) spectral ratios sampled at 5 values. And we take \( \pi_{Ex} \) and \( \pi_{Eq} \) as the a priori probability of the two types of events, so \( \pi_{Ex} + \pi_{Eq} = 1 \). We follow standard practice by assigning an event to the earthquake class if

\[
\frac{f_{Eq}(r)}{f_{Ex}(r)} > \frac{\pi_{Ex}}{\pi_{Eq}} \quad (1)
\]

and to the explosion class otherwise [see e.g., Seber, 1984]. This rule is optimum, in that it minimizes the total probability of misclassification. Our knowledge of \( f_{Ex}(r) \) and \( f_{Eq}(r) \) comes from the training groups, and we introduce and evaluate a linear discriminant function \( D(r) \) under the assumptions that:

a) the sample distributions are normal;

b) the dispersion...
matrices of the two groups are the same; and c) the training observations are correctly classified. The linear function is

\[ D(r) = \lambda^T \left[ r - (\mu_{Eq} + \mu_{Ex}) \right] / 2 \]

(2)

where \( \mu_{Eq} \) is the mean of the earthquake training values \( r_{Eq}^i \) \( (i = 1, \ldots, 30) \), and similarly for \( \mu_{Ex} \); \( \lambda = S^{-1} (\mu_{Eq} - \mu_{Ex}) \), and \( S \) is the average of the dispersion matrices \( S_{Eq} \) and \( S_{Ex} \), with \( (\text{for example}) S_{Eq} = \frac{1}{29} \sum_{i=1}^{30} (r_{Eq}^i - \mu_{Eq})^T (r_{Eq}^i - \mu_{Eq}) \); and \( ^T \) denotes a transpose. In terms of these easily calculated quantities, the discrimination analysis is very simple. The rule (1) becomes: assign an event \( r \) to the earthquake population if \( D(r) > \ln (\pi_{Eq} / \pi_{Ex}) \). If \( \pi_{Eq} = \pi_{Ex} = \frac{1}{2} \) the rule is even simpler: the event is labelled an earthquake or an explosion according as \( D(r) > 0 \) or \( D(r) < 0 \). In this case, the probability of misclassification is

\[ P(\text{misclassification}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Delta^2} \exp(-x^2/2) dx \]

(3)

where \( \Delta^2 = \lambda^T (\mu_{Eq} - \mu_{Ex}) \). \( \Delta^2 \) is the Mahalanobis D-squared measure of distance between the two means.

This type of multivariate analysis was developed and used by Fisher [1936], and has been used in seismological applications [e.g., Sandvin & Tjostheim, 1978]. If the dispersion matrices of the populations are not same, then a quadratic discriminant function is appropriate [Seber, 1984].

Four Applications of Discriminant Analysis

High-frequency network averaged \( P_g/L_g \) ratio: The sample data sets consisting of 30 earthquakes and 30 explosions were analyzed using the linear discriminant function given in (2). For each event, network averaged \( \log_{10} (P_g/L_g) \) ratios at frequencies of 5, 10, 15, 20 and 25 Hz correspond to the variables \( r_1, r_2, r_3, r_4, \) and \( r_5 \). The linear discriminant function obtained is

\[ D(r) = -1.3 + 15.2 r_1 - 43.9 r_2 + 17.5 r_3 - 0.5 r_4 - 34.7 r_5 \]

(4)

and the Mahalanobis D-squared measure is \( \Delta^2 = 20.768 \). Assuming equal prior probabilities for the two groups, we assign event \( r \) to the earthquake class if \( D(r) > 0 \). Applying this rule to the earthquake and explosion data, we find that all events are classified correctly and the misclassification probability is 0.0113. Values of \( D(r) \) may be called the discriminant score and are plotted in Figure 4 with respect to the mean \( \log_{10} (P_g/L_g) \) spectral amplitude ratio of each event. Vertical lines in the figure denoted as Eq and Ex are the projection of the multivariate mean of the earthquake and explosion populations, respectively. The vertical line, \( D_o \), is the line \( D(r) = 0 \), which serves to classify the events when the \( \text{a priori} \) probability of the two populations is the same. The distance between \( \text{Eq} \) and \( \text{Ex} \) is the Mahalanobis D-squared measure of distance between two populations, since, from (2),

\[ D(\mu_{Eq}) - D(\mu_{Ex}) = \Delta^2 \]

It is shown in Figure 4 that all the earthquake records from various paths in the eastern U. S. have a mean \( P_g/L_g \) spectral ratio of about 0.5, while the explosion records show a mean of about 1.25.

The equality of the dispersion matrices for the two groups is tested using a \( \chi^2 \) test [Seber, 1984]. The \( \chi^2 \) statistic, on 15 degrees of freedom in this case, is 10.89, i.e., well below the value 25 that would indicate rejection of the hypothesis that the dispersion matrices of the parent populations are different. The null hypothesis of equality of means of two groups is tested using an F statistic obtained from the Mahalanobis D-squared distance measure [Seber, 1984]. The F statistic for this analysis is 58.0 with (5, 54) degrees of freedom, which exceeds the critical value (\( F = 2.38 \) at the 5 % level of significance), so we reject the null hypothesis and conclude that our samples (\( P_g/L_g \) spectral ratios) indicate a difference in the means of the two populations.

Single record \( P_g/L_g \) ratio: When single-record \( \log_{10} (P_g/L_g) \) ratios of 230 explosions records and 255 earthquakes records are analyzed as the previous case, we find that about 14.9 % of earthquakes records and about 14.8 % of explosions records are classified incorrectly. The Mahalanobis D-squared measure of distance between two samples is 4.10 and the misclassification probability is 0.156 according to (3). This substantially higher misclassification probability of the single-record \( P_g/L_g \) ratio over the network-averaged \( P_g/L_g \) ratio indicates that the frequency content of regional \( P_g \) and \( L_g \) are strongly influenced by propagation paths as well as by local receiver site responses.

Low-frequency network \( P_g/L_g \) ratio: To compare the high frequency results with the more conventional lower frequency observations, we repeated the analysis using network averaged \( \log_{10} (P_g/L_g) \) ratios in the frequency band 1 - 10 Hz with 1 Hz intervals, so \( r \) had 10 elements. Discriminant function analyses for this low-frequency band suggest that most of the discriminant power is carried by frequencies higher than 5 Hz. The Mahalanobis D-squared measure is \( \Delta^2 = 10.426 \). The discriminant function analysis indicates that one event in each group in the training data is misclassified and the misclassification probability is 0.0532, suggesting that the low-frequency band performs more poorly than the high-frequency band. Using the low-frequency band, many ripple-fired explosions as well as several distant earthquakes clustered close to the classification line.

Discrimination test for another set of known events: We tested the performance of the above high-frequency discriminant function, eq. (4), by applying it to another set of five known earthquakes and 18 known explosions. We found
that all events except one were correctly classified. The one exception, an explosion misclassified as an earthquake, is a single-hole shot (#20 in Figure 1), which was fired in a water-filled quarry site [water depth about 200 m; see Kim et al., 1991]. The shot excited very strong $P_g$ and $L_g$ waves in the range 1–10 Hz due to the effects of reverberations in the water column. This is an abnormal explosion, to be recognized as not an earthquake by the dominance of $P$ over $S$, at distances up to 200 km. This single misclassification by our $P_g/L_g$ spectral ratio method is instructive but does not undercut the general utility of the method.

**Discussion and Conclusions**

We find that $P_g$ waves from explosions have stronger high frequency content than $L_g$ waves over the broad high-frequency band 5 – 25 Hz at regional distances in the eastern U. S.: the mean ratio for 230 explosion records is about 1.25. The opposite is true for signals from earthquakes with magnitudes 1.3 – 3.5: the mean $P_g/L_g$ ratio for 255 earthquake records in the same frequency band is about 0.5.

In a lower frequency band, 1-10 Hz, $P_g/L_g$ ratios show similar results but the separation between explosions and earthquakes is less clear. We note that observations in the western U. S. have been reported as having $P_g/L_g$ higher for earthquakes than explosions [e.g., Bennett & Murphy, 1986]. The key to our method and conclusions is the observability of frequencies up to about 20 Hz at regional distances in the eastern U. S.

One basis for the empirical success of our proposed high-frequency $P_g/L_g$ ratio method is that, if selected spectral amplitude reinforcement occurs at the source (as is the case for ripple-fired blasts), it will affect both early $P$ phases as well as later-arriving $L_g$ phases. Such spectral scalloping will largely cancel in the ratio we have used. The ratio is robust as a discriminant to the extent that there is cancellation of this and other characteristics, such as event size, corner frequency, some effects of focal depth, and instrument response. However, the $P_g/L_g$ ratio discriminant shows some dependence on near source conditions which may excite $P_g$ and $L_g$ signals with different efficiency (e.g., explosion source in low velocity rocks).

Strong $P_g$ but weak $L_g$ excitation from a vertical strike-slip earthquake may diminish the discrimination power of the $P_g/L_g$ ratio method at some station azimuths [Kim, 1987]. The dependence of frequency content of $P$ and $S$ waves on specific propagation paths and the receiver structure may explain why the $P_g/L_g$ ratio method applied to single station records had mixed results in several previous studies [Taylor, 1989]. We believe the key to avoidance of such problems of source and path is use of a network averaged $P_g/L_g$ ratio.

We have shown that the $P_g/L_g$ spectral amplitude ratio is an adequate discriminant for explosions from earthquakes (magnitude smaller than 4) in the eastern U. S. The stability of the $P_g/L_g$ ratio in discriminating ripple-fired explosions from small regional earthquakes is significant for earthquake monitoring in the eastern U. S., since the region is characterized by low seismicity with frequent occurrences of such explosions. Our results demonstrate the importance of data at frequencies up to at least 20 Hz, with implications for the design of networks in regions that have observable high-frequency signals.

**Acknowledgments.** D. Johnson, D. Lentricehia, R. Such of the Lamont-Doherty Earth Observatory have assisted in collecting the data used in this study. This project was supported by USGS Grant 1434-92-A-069, and by the Air Force Phillips Laboratory under F19628-90-K-0048. Lamont-Doherty Earth Observatory contribution number 5083.

**References**


---


(Received March 4, 1993; revised May 10, 1993; accepted May 10, 1993)