Source Effects on Regional Seismic Discriminant Measurements
by Jiajun Zhang, Thorne Lay, Julia Zaslow, and William R. Walter

Abstract  Source effects on high-frequency regional seismic discriminant measurements involving $P/S$-type ratios ($P_n/L_g$, $P_g/L_g$) are isolated for earthquakes in southern California and southern Nevada. The scatter in earthquake population discriminant measurements for fixed propagation paths is statistically analyzed to assess the spatial and frequency-dependent contributions from source radiation pattern and source depth variations along with variable near-source scattering. This is useful because procedures for reducing the variance caused by propagation effects must be constrained by knowledge of the intrinsic scatter resulting from independent source effects, for which corrections are not available without a priori knowledge of the source type and depth. Broadband waveforms for earthquakes near the Nevada Test Site (NTS) and in southern California recorded at stations of the Livermore NTS Network (LNN), the Caltech TERRAscope Network, and the Berkeley Digital Seismic Network (BDSN) are used to obtain $P_g/L_g$ and $P_n/L_g$ measurements and to assess source depth and mechanism contributions to measurement variance. $P_n/L_g$ and $P_g/L_g$ spectral amplitude ratios for several frequency bands between 0.5 and 10 Hz from vertical velocity waveforms are computed for a total of 18 tightly located clusters of earthquakes. For each cluster, we examine the correlations between observed $P_n/L_g$ and $P_g/L_g$ ratios and estimated source depths or predictions based on previously determined source mechanisms. Focal mechanism predictions do not match the observations well, likely due to strong near-source scattering and directivity effects that overwhelm the point-source radiation pattern. However, the data do show significant scatter even for sources only a few kilometers apart. Source depth produces some coherent trends, which can be emphasized by averaging results from multiple stations. These trends vary from region to region, presumably due to specific local crustal layering, but most of the variance is not deterministically predictable and appears to arise from strong variations in near-source scattering. Variance estimates for logarithmic amplitude ratios normalized by mean amplitude ratios vary from 0.01 to 0.05 for different phase pairs. Frequency-dependent variations of the mean, variance, minimum, and maximum amplitude ratio parameters show strong distance dependence out to about 300 km and then appear to become independent of event-station distance. These variance parameters appear to be suitable for use in spatial interpolations of observed $P/S$-type ratios for a specific unidentified event path, as needed for nuclear test monitoring.

Introduction

Seismic monitoring of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) depends in part on the ability to seismically detect and discriminate between clandestine underground nuclear tests and other seismic sources, including earthquakes. Most discrimination challenges are associated with low-yield underground nuclear explosions, given that such events may only be recorded at regional distances (100–1500 km) by a few monitoring stations. Seismic recordings at regional distances tend to be very complex because of the superposition of energy traveling on multiple paths within the crustal-wave guide, compounded by source radiation pattern and source depths effects. The waveform complexities result in substantial scatter in various regional discriminant measurements designed to extract attributes diagnostic of the source type. This becomes a particular problem for isolated stations in tectonically complex areas with great variation in path and earthquake properties. For global CTBT monitoring, identification of rare explosion signals becomes a statistical problem of recognizing measurements reflecting explosion attributes as outliers in a population of
corresponding measurements for earthquake signals. Reduction of the scatter in the reference earthquake population is essential for high confidence outlier detection (e.g., Fisk, 1994).

Seismic-wave discriminants typically involve relative amplitudes of different types of seismic waves that provide sensitivity to the source radiation differences between explosions and earthquakes. For large magnitude seismic events, there are well-established measurements such as $M_c$; $m_b$ that provide reliable source discrimination by utilizing relative excitation of long-period surface waves to short-period body waves. For small magnitude events ($m_b < 4$) the surface waves are usually masked by background noise, so high-frequency ($f > 0.5$ Hz) seismic discriminants are required. The high-frequency discriminants commonly involve regional waves such as $Pn$, $Pg$, $Sn$, and $Lg$ because relative amplitudes of these waves are sensitive to the differences in explosion and earthquake radiation. However, our ability to extract information about the source type from regional waves is limited because of the intrinsic complexity of the regional signals and our ignorance of detailed crustal structure on each path. Despite this, several regional high-frequency seismic discriminants have substantial promise for source identification.

$P/Lg$ measurements are inversely proportional to relative shear-wave excitation by the source, which tends to be lower for explosions (giving higher $P/Lg$ ratios) than for earthquakes. High-frequency $(3–8 +$ Hz) $P/Lg$ amplitude or spectral ratios are presently among the most promising regional distance seismic discriminants (e.g., Bennett and Murphy, 1986; Evernden et al., 1986; Taylor et al., 1989; Baumgardt and Der, 1994; Walter et al., 1995; Goldstein, 1995; Taylor, 1996; Harts et al., 1997; Kim et al., 1997). Walter et al. (1995) show that for $Pg/Lg$ and $Pn/Lg$ ratios in the passband 6–8 Hz for explosions and earthquakes at the Nevada Test Site (NTS) recorded at two broadband stations, the explosion and earthquake populations separate quite well at magnitudes above 3.5; however, there is significant overlap at lower magnitudes despite all the sources sharing common propagation paths to a given station. The overlap tends to be more severe at lower frequencies (e.g., Bennett and Murphy, 1986; Taylor et al., 1988, 1989; Lynnes and Baumstark, 1991). The major obstacle to reliable source identification with these discriminants is the scatter in the two event populations, which results in overlap at smaller magnitudes and potentially overlap at larger magnitudes if more earthquake signals were available. Some of the scatter in the explosion population is recognized to be the result of variable source emplacement properties, such as water-table proximity and percent gas-filled porosity (Taylor and Denny, 1991; Walter et al., 1995). The source contributions in the earthquake population are generally not well understood, but there appears to be some influence of source depth, with very shallow events looking more explosion-like (Walter et al., 1995). It is generally believed that earthquake radiation patterns may influence discriminant measures, especially single-station observations (e.g., Bennett and Murphy, 1986), but there has been no quantification of what portion of the scatter is due to mechanism variability, or what the frequency dependent behavior of this scatter is. These factors are impossible to anticipate or correct for in any monitoring effort, but a sound understanding of the causes of scatter is still needed. The NTS may provide some of the more dramatic near-source variations likely to be encountered, other than in decoupling situations (Taylor and Denny, 1991), so the scatter observed by Walter et al. (1995) may be unusually large for events on a given path. However, the CTBT must address identification of events in all locations, not just as historical test sites, and this means that variable propagation effects will be superimposed on the source effects, adding much more scatter to explosion and earthquake observations. The scatter of $P/S$-type ratios is ubiquitous and often ranges over an order of magnitude for events in a given region (e.g., Kim et al., 1997; Harts et al., 1997; Fan and Lay, 1998a).

Scatter in regional discriminant measurements associated with varying path propagation effects is typically suppressed by empirical means such as distance corrections (either based on direct data interpolations or models constructed to match data), attenuation corrections, spatial averaging of event populations by cap averaging or kriging (e.g., Harts et al., 1998; Phillips et al., 1998; Schultz et al., 1998; Rodgers et al., 1999; Phillips, 1999), or regression on wave-guide parameters such as crustal thickness, sediment thickness, and elevation measures (e.g., Zhang et al., 1994; Fan and Lay, 1998a,b,c). When such corrections are estimated, there is always a question of how much of the scatter in the reference population should be attributed to propagation effects versus source effects? Corrections that reduce all of the variance in the reference population would very likely map source-induced scatter into erroneous path corrections, potentially giving poor performance of the discriminant for the next unknown event in the region. Similarly, path corrections based on sampling of a limited range of source mechanisms on a given path may prove erroneous when applied to a new event on that path with a distinct source mechanisms (such as an explosion). A firm statistical basis for separating correctable path contributions from intrinsically noncorrectable (other than by simple methods such as azimuthal averaging, which is seldom viable in the CTBT monitoring context) near-source contributions is needed.

This study investigates source-related scatter in regional seismic discriminants using well-controlled earthquake populations. We characterize earthquake radiation pattern and source depth effects on high-frequency regional $P/S$-type ratios using earthquakes near NTS and in southern California recorded by broadband seismic stations of the Livermore NTS Network (LNN), the Caltech TERRAscope Network, and the Berkeley Digital Seismic Network (BDSN). We exploit the dense distribution of well-located seismicity with diverse faulting geometries in southern California to provide re-
ngional data on nearly equivalent paths, so that the variable source contribution to the signal fluctuations can be isolated. For each cluster of events that share common propagation paths to a given station, regional seismic discriminants are computed and their characteristics estimated; the differences between measurements for events in the same cluster can be ascribed to near-source effects. We focus on characterizing the variation of scatter of discriminant measurements with distance and its dependence on source depth and focal mechanism variations. While deterministic characterization of the variations arising from near-source effects proves challenging, robust statistical estimates of the variations are obtained.

Data Analysis

For quantifying source radiation pattern and source depth effects on high-frequency regional discriminants, we isolate the variable source contributions by using closely located events recorded by a given station (Zaslow et al., 1996). We consider three sets of events; each comprises one or more distinct clusters of earthquakes with epicentral locations within a few kilometers of one another. For the first and second sets of events, we analyzed clusters of small earthquakes with independently obtained focal mechanisms (clearly these are subject to error), allowing us to consider source radiation pattern effects deterministically. The events in the first set involve a single cluster of aftershocks of the 29 June 1992 Little Skull Mountain earthquake ($M_w = 5.7$) near NTS. The events in the second and third sets are earthquakes in southern California located in 3 and 17 clusters, respectively. The focal mechanisms and depths for all events in the second set were determined from relatively robust surface-wave spectral inversions by Thio and Kanamori (1995). The third set of events is an extended event set, which includes all of the second set of events plus additional events for which focal mechanisms are not available (or only less reliable first-motion solutions are available), along with some large magnitude events omitted from the second set. The hypocentral parameters of the 17 clusters of events in the third set are taken from the Southern California Earthquake Center (SCEC) catalog. Hereinafter we refer to these three sets of events as the LSM events (set 1), TK events (set 2), and SCEC events (set 3).

Clusters are defined as groups of events having epicentral locations within a few kilometers from the center of the cluster. Thus, regional recordings at stations from 100 to 1500 km away from a given cluster essentially share common propagation paths (and scattering volumes). The convolutional nature of seismic signals suggests that one can potentially separate source and propagation effects, but in practice this is actually a very tricky proposition, and we must be careful to define what we mean by source effects. Gross structural effects on the signals, such as geometric spreading, attenuation, receiver effects, and wave-guide controls on energy partitioning and layer reverberations are to a large extent shared by all observations for a cluster at a specific station; however, the precise illumination of these fixed effects from source to source does depend on the variable source depth and radiation patterns of events in the cluster. Regional distance signals such as $L_g$ involve many arrivals that take off from the source at different angles and that interact with the structure in complex manner. Usually spectral amplitudes or root mean square (rms) amplitudes for a given group velocity are measured, providing averaging of the suite of arrivals in the signal. The so-called source effects that we can hope to isolate by considering signals that share a common path actually involve interactions of the source radiation with details of the wave-guide structure along the entire path; what we can detect is the sensitivity of the transmitted energy to the changes in the depth and focal mechanism of the events. This may involve nonlinear interactions between the source attributes (particularly depth) and the crustal transfer function and may be enhanced by subtle aspects of the source rupture directivity and frequency content. Given that the sources are never exactly in the same location, strong near-source scattering effects may also account for variations in the signals. When earthquake mechanisms are identical, event locations within one-fourth wavelength spatial separation can provide nearly identical waveforms, but larger separations result in rapid waveform decorrelation, varying exponentially with distance separating the two events, even when the source geometries are the same (e.g., Geller and Mueller, 1980; Poupinet et al., 1984; Menke, 1999). In order to provide a sufficient data set to analyze focal mechanism and source depth variations, we must tolerate some degree of spatial separation of the sources; this means that near-source scattering is not fully suppressed. In the global CTBT monitoring context, relative location accuracy within a few kilometers can seldom be attained, so discrimination will always be practiced on events that have significant location uncertainty, typically exceeding the size of our clusters.

We expect a range-dependent pattern in the waveform complexity caused by varying source depth and radiation pattern as the station distance increases from the range of simple direct crustal phases (within a few tens of kilometers) to reverberative, multipathed phases (beyond 100 km). In terms of amplitude variance, as the range increases, the focal sphere averaging effects of the regional phase measurements should progressively reduce the influence of source attributes. The averaging effects intrinsic to spectral or rms amplitudes of overlapping signals should further tend to subdue but not eliminate the source sensitivity. We analyze a very large data set in order to provide thorough sampling of the source/path coupling, but our interpretations must be tempered by the recognition that we are constrained in the range of focal mechanisms and crustal types sampled. Other regions may have a higher or lower sensitivity to source variations depending on the nature of the crustal transfer functions and regional earthquake processes.

Although there is some evidence suggesting magnitude
dependence of $Pn/Lg$ ratios (e.g., Xie and Patton, 1999), we do not attempt any correction for $P$ and $S$ corner frequency scaling with magnitude because it is not well constrained by theory or observation. The extent to which focal mechanism and source depth effects fail to account for variations in the observations leaves room for source spectrum scaling, but our data do not clearly indicate (nor preclude) such magnitude dependence.

All of our measurements are made from vertical component seismograms because the data with suitable horizontal components are much sparser. Many applications of discriminants for CTBT monitoring utilize the vertical components for phases such as $Lg$ because data are obtained by vertical-component arrays, so this is not incompatible with typical data availability. The behavior of transverse-component $Lg$ versus vertical-component $Lg$ is usually quite similar, which suggests that $Lg$ sensitivity to double-couple radiation patterns is subdued. Nonetheless, future work using three-component data may extend the results of this investigation.

Figure 1. Geometry for analyzing effects of focal mechanism and source depth for events in the aftershock sequence of the Little Skull Mountain earthquake. The paths to stations MNV and KNB are almost identical for the events. Observations for MNV are shown below, normalized by the mean behavior for each set of ratios and plotted against computed ratios for the range of focal mechanisms in the data. A straight line with a slope of unity would indicate perfect prediction of the data. The correlations are low for both phases, but there is a factor of 4 spread in the $Pg/Lg$ and a factor of 2.5 spread in the $Pn/Lg$ measurements.
LSM Events Set

We begin with analysis of the first set of events, for which radiation pattern variations can be characterized using previously determined focal mechanisms. Our procedure is illustrated in Figure 1 for aftershocks of the Little Skull Mountain earthquake. Focal mechanisms and relative source depths are available for these events from Harmsen (1994) and Meremonte et al. (1995), and discriminant measurements are taken from Walter et al. (1995). Figure 1 shows that there is moderate variability in mechanism type for events in this cluster. Most events involve normal faulting with the exception of two strike-slip events; however, note that the dip and strike of the nodal planes are variable, which predicts substantial variability for Pn. For station MNV, Pg/Lg measurements range over a factor of 4, whereas Pn/Lg measurements vary over a factor of 2.5; the scatter is larger for both ratios for station KNB. The magnitude of the scatter does not vary much with frequency as shown, and correlations among measurements in different frequency bands are relatively high for Pn/Lg at station KNB but low for other ratios at KNB and MNV.

Predicted P/S ratios were obtained by integration of the focal mechanisms over appropriate takeoff angles and azimuths. Only the source radiation term is computed, not the propagation-induced weighting of each raypath, as that is assumed to be relatively constant for the cluster. The procedure is schematically illustrated in Figure 2. For a range of takeoff angles and azimuths, we average the P-wave or SV radiation patterns for the average relative source radiation for the principal arrivals contributing to the Pn, Pg, and Lg phases. This effectively estimates the relative strength of S versus P energy going toward the station, which should then be acted on by the common propagation path effects. Raypaths in a local crustal model were used as a guide to appropriate takeoff angle sampling; however, synthetic computations for the associated one-dimensional crustal model do not replicate the overall high-frequency energy distribution of the regional signals very well. We integrate over rather generous azimuth and takeoff angle ranges for each phase to accommodate both uncertainty in the structure and contributions from scattered energy radiated in the general direction of the station. The simplest phase is Pn, which has a relatively well-defined principle set of arrivals contributing energy to the phase window (Pn, pPn, sPn), with takeoff angles that are critically incident on the Moho. Pg and Lg arrivals sample wider ranges of takeoff angles from the source than Pn, with an increasing range of takeoff angles contributing arrivals as the distance increases. As our distances are large relative to the crustal thickness, we average the radiation pattern for Pg spanning the range from the Moho reflection, PmP, up to the surface reflection, pPmP, whereas for Lg we average the SV radiation pattern for the full range of postcritical angles of incidence. This captures the main suite of upgoing and downgoing P arrivals that contributes to the Pg group velocity window and the multiple-SmS arrivals that contribute to Lg. Although this is a relatively crude procedure, it provides a good first-order characterization of the relative amplitudes of the suite of arrivals contributing to each of our phase windows. This is demonstrated by comparison with synthetic calculations in a following section. Note that the full radiation pattern term is considered so that azimuthal patterns are accounted for; these cause some of the strongest relative variations for Pg and Lg, whereas Pn is quite sensitive to the precise focal mechanism because energy is concentrated near localized critical angles. Comparison with radiation patterns for high-frequency Rayleigh-wave overtones indicates that our relative Lg-amplitude predictions are quite reasonable, because SV energy dominates in the vertical component of Lg. As shown subsequently, we further tune the integration of focal mechanism using synthetics when we are dealing with data at varying ranges. For the LSM events, which are all at nearly the same distance, this is not as important.

We find a weak correlation of predictions with observed Pg/Lg amplitude variations for MNV (Fig. 1) as well as for KNB (not shown here). Perfect prediction of the observations would align the data on a straight line with unity slope in
these plots. The correlation is stronger for frequencies from 0.5 to 2 Hz, especially for $P_g/L_g$, and negligible for higher frequencies. For MNV the total variation of predicted ratios for $P_g/L_g$ is comparable to that for observed ratios, whereas for $P_n/L_g$, the variation of predicted ratios is a factor of 2 larger than that for measurements. This is due in part to the greater focal mechanism sensitivity of predicted $P_n$ amplitudes; the variations in dip amongst the sources may actually be less than in the estimated focal mechanisms, leading us to overpredict the range of amplitude variation. Little correlation with depth was found for the $P_g/L_g$ and $P_n/L_g$ measurements at either station. There is substantial uncertainty in the focal mechanism and depth estimates for these events, and only a few available stations, so data from additional clusters were sought for the more densely instrumented region of Southern California. However, it is worth noting that the quality of the depth estimates and focal mechanism estimates for the LSM events is probably better than what will be available for earthquakes in most regions, so isolation of focal mechanism effects is likely to prove difficult in many cases.

Southern California Events

Southern California provides several additional earthquake clusters for which the focal mechanisms have been determined by broadband waveform inversions (e.g., Romanowicz et al., 1993; Ritsema and Lay, 1995; Thio and Kanamori, 1995), along with many additional clusters lacking focal mechanisms. Figure 3 shows locations of a total of 17 clusters of southern California earthquakes and locations of 18 broadband stations used in this study. The data are broadband waveforms recorded at the eleven BDSN and seven TERRAscope stations indicated in Figure 3. For each cluster we calculate $P_n/L_g$ and $P_g/L_g$ spectral amplitude ratios for various frequency bands between 0.5 and 10 Hz (the frequency passbands span 0.5–1, 1–2, 2–4, 4–6, 6–8, and 8–10 Hz) using vertical-component velocity waveforms (no instrument correction is applied as it cancels in the spectral

Figure 3. Map showing locations of broadband stations (triangles) in California (11 BDSN and 7 TERRAscope stations with their names labeled here are used in this study) and 17 clusters of southern California earthquakes (crosses indicate central locations of various clusters); all the events in a cluster were located within 10 km to the center of the cluster. The small insert map at bottom left shows locations of cluster A1 events (each open circle represents one or more events with the same epicentral location).
rations which only involve common frequency bands). To characterize the scatter of the observed $Pn/Lg$ and $Pg/Lg$ measurements for a given cluster we compute the variance, mean, minimum, and maximum of the amplitude ratios for all events in the cluster. The events in the clusters occurred from 1992 to 1998, and we collected signals recorded at regional distances up to 1100 km. We use the same phase isolation and amplitude estimation procedure as Walter et al. (1995). In our calculation of spectral amplitudes, the $Pn$-phase window typically involves energy from $Pn$, $pPn$, and $sPn$; the $Pg$ window spans from the end of $Pn$ window to a 5 km/sec group arrival; the $Lg$ window spans from the 3.0 to 3.6 km/sec group arrivals.

Altogether, the southern California clusters comprise 359 earthquakes with magnitudes ranging between 3.5 and 7.0 (there are 210, 93, 38, 15, and 3 events with magnitudes less than 4.0, 4.0–4.5, 4.5–5.0, 5.0–5.5, 5.5 and larger, respectively). Each cluster includes 20–70 earthquakes that were located within no more than 10 km from the center of the cluster. Some of the clusters are grouped in close proximity and are named accordingly (Fig. 4); events in A clusters are associated with the Northridge earthquake ($M_w 6.7$) sequence commencing 17 January 1994; B clusters are for the Joshua Tree earthquake ($M_w 6.1$) sequence commencing on 23 April 1992 and the Landers earthquake ($M_w 7.3$) sequence beginning 28 June 1992; C clusters are for the Big Bear earthquake ($M_w 6.2$) sequence beginning 28 June 1992.

**TK Events Set**

The subset of clusters with focal mechanism information is from clusters A1, B1, and C1; these comprise the TK events set. Figure 5 shows the focal mechanisms for the TK events set in cluster A1 (Fig. 4) and ray paths from the events to various TERRAscope stations at regional distances. These events have magnitudes of 3.3–4.4 and variable focal depths and faulting mechanisms. For each station the regional distance recordings show frequency-dependent variation in waveforms, which presumably have some contributions from differences in source depth and focal mechanism. The variations are least pronounced for frequencies above 5 Hz (Fig. 5), but there is substantial scatter in all regional phase amplitude ratios on these paths.

One of the main difficulties encountered in this analysis involves the windowing of regional phases at the relatively close distances of less than 200 km, particularly when source depths vary from 5 to 15 km. The problem is that as the source depth increases depth phases may move out of one group velocity window (say, for $Pn$) and into another (say, for $Pg$), thus causing contamination of the energy measures computed for the separate phases. This is difficult to account for in the prediction of amplitudes by focal mechanism integration.

In order to tune the integrations to better reflect the actual contributions from phases for different source depths and path lengths, we constructed regional distance synthetics for various source/receiver geometries using the ranges in source depths and focal mechanisms of the observations. Then we found ranges of radiation pattern integrations that provide reasonably good predictions of the observed relative amplitudes in the $Pn$, $Pg$, and $Lg$ windows of the synthetics (Fig. 6). This guided us to incorporate $sPn$ contributions to the $Pn$ window for shallow events and $Sg$ contributions to the $Pg$ window ($S$ waves that come out from the source, convert to $P$, and then arrive in the $Pg$ window). We then found quite good predictability of the synthetic amplitudes at distances of 220–400 km (Fig. 6), although shorter distance windows remain very problematic due to the short time windows and overlapping of depth phases. Again, the synthetics do not closely match any of the observations, so this provides only a guide as to the appropriate angles over which we perform focal mechanism integrations.

Using our refined radiation pattern integration criteria, we analyzed the three TK events clusters; each cluster comprises events located within a 10-km radius, recorded at several TERRAscope and BDSN stations, and having diverse focal mechanisms. Figure 7 displays the results for low-frequency ratios (0.5–1.0 Hz) found using the focal mechanisms and depths determined by Thio and Kanamori (1995), plotted for each of the five station-cluster pairs for stations ISA, CMB, and BAR. Note that for cluster A1 events, the path to ISA is too short to measure $Pn$ reliably for most observations. The correlations with focal mechanism predictions and source depth are generally weak, just as for LSM events, certainly much lower than for synthetic cases. One might expect the deterministic structure of the wave field to be more predictable in this low-frequency band than in higher-frequency bands, but only a few isolated cases show...
much trend (e.g. ISA for cluster C1). Yet the scatter in the observations is comparable to the predicted ranges, and it presumably does reflect either a source radiation effect or near-source scattering that is remarkable sensitivity to the precise travel path taken. The depth range among the sources is about 20 km, which is the maximum extent found in southern California, but any dependence of the ratios on depth is very subtle. The results for higher-frequency ratios are very similar.

Figure 8 shows the $P_g/L_g$ and $P_n/L_g$ measurements and corresponding synthetic predictions for all the TK event station-cluster pairs, plotted for events in several depth ranges. The scatter in the observations is generally comparable with that in the synthetic predictions for events at various depth ranges, and the scatter is quite uniform with depth. Slightly stronger correlations are found for $P_n/L_g$, but there is not a strong variation in the correlation as a function of event depth.
SCEC Events Set

Given that deterministic predictability of the discriminant ratios appears to be very limited, we proceed to emphasize the statistical attributes of the variance in the discriminant measures using a much larger data set. This is the behavior that needs to be understood for practical applications of path corrections anyway; generally focal mechanisms and precise depths will not be known for the events used to calibrate regional discriminants. We examine characteristics of $P_g/L_g$ and $P_n/L_g$ measurements for the SCEC events, which comprise all the events in the 17 southern California earthquake clusters (Fig. 4). This data set is extended from the TK events, with focal mechanisms for most events not being known and only the catalog depths available. Additional events are included in the three clusters defined for the TK set.

Figure 9a shows the $P_g/L_g$ measurements for the 0.5- to 1-Hz passband as a function of SCEC catalog source depth for the expanded clusters A1, B1, and C1 recorded at station ISA. In comparison with Figure 7, Figure 9a includes many more events, but it still indicates little correlation between observed $P_g/L_g$ measurements and estimated source depth. Weak correlations of $P_g/L_g$ and $P_n/L_g$ measurements with source depth are also found for other passbands, and the level of scatter in the observations is comparable to that shown in Figure 9.

With the exception of $P_n/L_g$ measurements for cluster A1 events, all of the logarithmic $P_n/L_g$ and $P_g/L_g$ measurements (normalized by the separate cluster mean values) for ISA lie in the range $-0.5$ to $0.5$, corresponding to an order of magnitude total scatter in the data for a given path. As shown subsequently, this range of scatter decreases as prop-
agation distance increases (the ranges from each cluster to ISA are shown in Figure 3). The lack of systematic trend between observed $P/Lg$ ratios and source depth suggests a complex behavior that is not easily related to the layered crustal structure. Although the bulletin depth estimates are not believed to be very precise (even in this extremely heavily instrumented region), there is no clear pattern of energy partitioning between $P$ and $S$ due to, for example, varying mode excitation for the $Lg$ phase. This robustness of the measurements is an asset for CTBT discrimination practices, but the large scatter in observations on a fixed path is a major concern; path corrections cannot, or at least should not, reduce this scatter in the population of earthquake measurements.

Figures 9b and 9c show the various parameters (mean, variance, minimum, and maximum) used to characterize the $Pg/Lg$ measurements for the three clusters at ISA as functions of the center frequencies of the different passbands. In the following we refer to these parameters as distribution parameters and use them to characterize $P/Lg$ measurements for many station/cluster pairs. Note that in Figure 9 all the means are negative due to skew in the logarithmic normalized amplitude ratio distributions. There are moderate variations for each distribution parameter over the frequency passband used here; the mean, variance, minimum, and maximum range between $-0.06$ and $-0.02$, $0.02$ and $0.06$, $-0.6$ and $-0.2$, and $0.2$ and $0.6$, respectively.

Figure 9 shows that ISA $Pg/Lg$ measurements for clusters B1 and C1 events have similar scatter in the depth distribution (Fig. 9a), variance (Fig. 9b), and mean (Fig. 9b), all being slightly larger than for cluster A1. Similarity is also observed in the scatter of $Pn/Lg$ measurements for clusters

![Figure 7. $Pg/Lg$ and $Pn/Lg$ measurements (0.5–1 Hz) at stations ISA (left three columns), BAR (second column from the right), and CMB (first column from the right) for three clusters of earthquakes, plotted as a function of radiation pattern predictions and source depth. The earthquakes for each cluster had diverse focal mechanisms and source depths and were located on common great-circle paths to the station. Note that for cluster A1 events, the ISA path is too short to measure $Pn$.](image-url)
B1 and C1 events. This may reflect the proximity of wave paths to station ISA from cluster C1 events and from cluster B1 events versus paths from cluster A1 events (see Fig. 3). We expect scatter in the measurements will depend both on the particular distribution of source mechanisms and depth range of events in a cluster as well as on the specific path traversed and the range to the station (cluster A1 is closer to ISA than the other two clusters). We must consider the scatter as functions of distance for many path combinations to get a realistic assessment of variance due to source attributes.

Our data set samples a large range of regional distances, with each station spanning as much as 200-km range from the 17 clusters. For station ISA, our clusters are located in the distances of 100–300 km, and for station YBH, the clusters are located in the distance range 800–1300 km. Figure 10 shows the distribution parameters (variance, mean, minimum, and maximum) for \( \frac{P_g}{L_g} \) and \( \frac{P_n}{L_g} \) ratios for the 0.5- to 1-Hz passband at stations ISA, CMB, and YBH, as a function of station-cluster distance. The scatter parameters appear to vary slowly with event-station distances over limited ranges, but overall there are clear reductions in the scatter as distance increases. The question then arises as to whether this is a station signal or truly a range dependence.

Figure 11a and b shows the distribution parameters for \( \frac{P_g}{L_g} \) and \( \frac{P_n}{L_g} \) ratios, respectively, for the 0.5- to 1-Hz passband for all the station-cluster pairs used in this study, plotted as functions of path length. Figure 12a and b shows corresponding results for the 2- to 4-Hz passband. Results are plotted using various symbols, identifying the corresponding station (coded by color) and cluster (coded by symbol shape; note that groups of proximate clusters are given the same symbol). Figures 11 and 12 show that the range dependence of the scatter does dominate over station dependence; distinct station-cluster pairs contribute to the overall decrease of scatter with distance. The scatter for episcopal distances of 100–300 km is much larger than that for longer distances; the distribution parameters tend to become small in absolute value at large distances. In southern California, \( P_n \) is well isolated from \( P_g \) beyond about 150 km and for deeper sources a few measurements can be made as close as 120 km, albeit with a very short \( P_n \) window prior to \( P_g \) (e.g. Hearn, 1984; Hearn and Clayton, 1986b). The short \( P_n \) window and lack of development of the \( L_g \) phase undoubtedly contributes to the variance in the ratios at distances less than 200 km in Figures 11b and 12b. \( P_g \) is robust at these close distances (e.g. Hearn and Clayton, 1986a), but \( P_n \) weakens, becoming a scattered suite arrivals beyond 300 km. From consideration of Figure 3, it is clear that our path sampling of the region is far from isotropic, so there is some potential for biasing of these distributions, but considering that each point in Figures 11 and 12 corresponds to statistical values for 20 or more seismograms, this is a vastly larger data set than has ever been assessed for this question. There are a few outliers from the general pattern, which involve stations with particularly high noise levels; more stringent signal-to-noise criteria than ours would reduce the occurrence of these outliers substantially. Progressive development of the \( L_g \) phase from isolated \( S_mS \) reverberations to a superposition of multiple reverberations combined with scattered contributions may be the dominant factor in the strong trend from 100- to 300-km distance. The strong distance dependence of variance motivates incorporation of this behavior into path
correction interpolation methods rather than use of a single number based on either the worst case (close-in) or best case (most distant) data behavior.

In the preceding analysis of LSM and TK events, we found little correlation between \( P/L_g \) measurements and source depth at a given station. Either the scattering and radiation pattern effects overwhelm the source depth dependence, or it is truly negligible. One approach to suppressing the radiation pattern effect is to average ratios for stations at different azimuths and distances; this may extract a subtle common behavior with source depth for a given cluster. We thus perform a ratio-averaging procedure that allows us to examine the effects of source depth in more detail for each cluster. Ideally, the stations would be well distributed azimuthally, but our stations tend to be distributed to the north or northwest, and there are usually not many stations with adequate signal levels for events less than magnitude 4. Therefore, we do not hope to fully isolate the depth effect from other contributions, particularly for the \( Pn/L_g \) ratios that have poorer sampling, but other effects should be suppressed to some extent.

Figure 13a and b shows the depth distributions of the averaged \( P_g/L_g \) and \( Pn/L_g \) measurements for events in each cluster. The ratios are normalized by the overall average for all events in the cluster. Each average value at a given depth was calculated from at least two (Fig. 13a) or four (Fig. 13b) stations. The difference between Figure 13a and b should reflect the progressive suppression of contributions from source radiation pattern, which is significant for the B and C clusters.

The cluster A results are quite stable because of the large number of observations and stations available for these events. There are low \( P_g/L_g \) ratio values very near the surface with a consistent high ratio in the 1- to 3-km-depth
range. Enhanced $L_g$ excitation for events in the shallow sedimentary basins near the surface could account for the low $P_g/L_g$ ratios, but the cause of the peak near 2 km is unclear. $P_n/L_g$ shows high values in the upper few kilometers in most cases. The A clusters show low ratios from 4 to 7 km for both $P_g/L_g$ and $P_n/L_g$, with rapid depth variations among different clusters at depths from 9 to 12 km. The common features that emerge are possibly controlled by local crustal layering, affecting the $P/S$ energy partitioning for the tightly spaced clusters of events. Low $P_g/L_g$ ratios near the surface are found for C clusters as well as the D cluster. The A, C, and D clusters all occur in basin structures, unlike the B clusters, which do not show low $P_g/L_g$ ratios at the surface. The B clusters, which spread over a larger region than the A clusters, do show coherent features, with a strong decrease in $P_g/L_g$ ratios near 3–4 km, and again from 9 to 10 km, whereas $P_n/L_g$ increases from 6 to 8 km (although this is not apparent for the reduced data set constrained to four station averages). It is likely that variable crustal layering accounts for the distinct depth-dependent features among clusters, and this feeds into the overall variance for a given cluster. Differences in the distance sampling, numbers of available data, and precision of earthquake depth determinations make interpretation of the detailed patterns problematic, but it does appear that there are coherent depth effects of a complex nature affecting the data that are obscured in the raw ratios.
Discussion and Conclusions

This study has demonstrated that the scatter in high-frequency regional P/S ratio discriminant measures for fixed propagation paths is substantial and depends on the distance range of the observations. Our coverage is such that we expect these results to be rather robust for California and likely to hold as typical of active tectonic regions with strongly heterogeneous crust (although regions with a greater range in source depths may have accentuated variations). One consequence of the acute variability in observations on a given path is that deterministic prediction of the observations using focal mechanism integration or modeling approaches is intrinsically very limited. A region such as southern California has far better event locations and focal mechanism determinations that is typically available anywhere else in the

Figure 11. (a) Characteristics of $P_g/L_g$ measurements at 0.5–1 Hz, plotted as a function of station-cluster distance for earthquakes in various clusters. Each point is for a particular cluster and station, with clusters identified by symbol and stations by color. Groups of clusters have the same symbol. (continued on next page)
world, so this is one region where the database is sufficient to explore this issue in detail.

The lack of correlation with predictions is probably due to the fact that the crustal models typically available even in well-studied regions are highly simplified and do not fully capture the coupled sensitivity of source and propagation effects for frequencies above 1 Hz. In addition, high-frequency radiation from earthquake sources is complicated by source finiteness, directivity, and scattering effects that are difficult to predict. Although there has been excellent success in modeling signals for periods of 5 sec and longer to determine focal mechanisms by regional waveform inversion, analysis of the isolated higher frequency information is more challenging. It is possible that regions with much simpler crustal structures will exhibit a higher degree of predictability of discriminant variations based on focal mechanism effects, but whether this is the case for secondary arrivals such as $P_g$ and $L_g$ for large numbers of small events is yet to be determined.

Because regional phases such as $P_g$ and $L_g$ intrinsically involve energy with a wide range of takeoff angles (and azimuths) that sample the radiation pattern in a complicated fashion, it may be useful to explore alternate measures of faulting orientation to seek systematic sensitivity to source

Figure 11. (continued). (b) As in (a) but for $P_n/L_g$ measurements at 0.5–1 Hz.
Source Effects on Regional Seismic Discriminant Measurements

Empirical measures such as low-frequency Love/Rayleigh wave ratios may serve as a useful gauge of mechanism variation against which high-frequency discriminant behavior can be compared. Ultimately, quantification of the specific scatter due to radiation pattern would be satisfying from the point of view of comprehensive understanding of regional signals, but it may also be valuable in developing a high-fidelity reference baseline for regional discrimination in the situation of problem events for which conventional methods fail. This could also help in confidence building for path corrections, which could be determined without contamination from source effects.

Zhang (1997) noted that the fully coupled source-propagation effect for a given event-station geometry for a seismic waveform or spectrum may be represented as a tensor of six components rather than as a single complex number characterizing the time delay and phase shift. For several earthquake clusters, each with sufficiently large number of

Figure 12. (a) Characteristics of $P_g/L_g$ measurements at 2–4 Hz, plotted as a function of station-cluster distance for earthquakes in various clusters, as in Figure 11a.

(continued on next page)
events, it may be possible to examine the degree of freedom (1, 6, or larger?) that can be extracted to characterize the propagation effects. This involves the question of, for a given source-region and station geometry, how many waveforms are required to fully estimate the characteristics of the propagation effects or the characteristics of the seismic discriminants for the event-station combination?

Whether deterministic approaches to understanding regional discriminants fully succeed, the results of this article have substantial practical value for discrimination applications. For example, some of the most promising discrimination path correction techniques involve spatial interpolation techniques such as kriging (e.g., Phillips, 1999; Rodgers et al., 1999). Kriging in turn requires the proper estimation of several variances in order to provide proper uncertainty propagation. One such key parameter is the nugget value or variance of earthquake amplitudes with nearly coincident epicenters. These estimates are commonly determined through variogram modeling of existing data, but this is dependent on adequate sampling. To the extent that a region

Figure 12. (continued). (b) As in (a) but for Pn/Lg measurements at 2–4 Hz.
Figure 13. (a) Curves showing the $P_g/L_g$ (left column) and $P_n/L_g$ (right column) measurements for events in various clusters, plotted as a function of source depth. The measurements were calculated from the observed $P_g/L_g$ or $P_n/L_g$ ratios in the 0.5 to 1-Hz passband over various stations for each event, for which the ratios are available from at least two stations. The cluster identification is shown above each curve and the zero reference line.

(b) Curves showing the $P_g/L_g$ (left column) and $P_n/L_g$ (right column) measurements for events in various clusters, plotted as a function of source depth. The measurements were calculated from observed $P_g/L_g$ or $P_n/L_g$ ratios in the 0.5- to 1-Hz passband over various stations for each event, for which the ratios are available from at least four stations.
is aseismic or has long repeat intervals for earthquakes compared with its data history there is a concern that variogram modeling of the existing data could underpredict the true variance. For example, a few-year history of data at a new IMS station might fail to capture the actual variability in depth and focal mechanism expected over the next 30 or more years. The combination of extensive aftershock studies in seismically active areas and explorations of focal mechanism and depth effects in this article provide minimum variances for these effects that can be applied in such cases.

Similarly, the issue of the amplitude variance as a function of range is also poorly determined for the purposes of assessing the uncertainty of kriging path corrections. Near duplication zones such as at the $Pn-Pg$ crossover, or at upper-mantle duplication distances, the amplitude variance would be expected to increase. Nonstationary Bayesian kriging (Schultz et al., 1998) can handle these cases if an a priori model of this variance is available. One could, for example, use the upper bound of the variance estimates in Figures 11 and 12 to map out a variance range model. The transportability of these numbers needs to be assessed, but the large sampling provided by this study probably assures the general utility of these results, particularly to the extent that California heterogeneity is stronger than elsewhere. We believe that aftershock studies of the type presented here hold much promise for providing practical determinations of the true uncertainties in parameters associated with discrimination path correction techniques such as kriging.

Acknowledgments

We appreciate the ready data availability from the TERRA scope and BGS data centers as well as IRIS DMS. We thank the associate editor and two anonymous reviewers for their helpful comments on the manuscript. This research was supported by IGPP/Lawrence Livermore National Laboratory (LLNL) UCRP Grant 99 GS-022 (funded by the GNEM program at LLNL) and the Defense Threat Reduction Agency under Contract Number DSWA 01 98-C-0161. This research was performed under the auspices of the U.S. Department of Energy under Contract Number W-7405-ENG-48 to the Lawrence Livermore National Laboratory. This is CSIDE Contribution Number 453, IGPP-CSIDE, University of California, Santa Cruz, and Contribution Number UCRL-JC-145341 from LLNL.

References


IGPP Center for the Study of Imaging and Dynamics of the Earth
Earth Science Department
University of California–Santa Cruz
Santa Cruz, California 95064
(J.Z., T.L., J.Z.)

Geophysics and Global Security Division
Lawrence Livermore National Laboratory
Livermore, California 94550
(W.R.W.)

Manuscript received 5 September 2001.