Bulletin of the Seismological Society of America, Vol. 92, No. 6, pp. 2277–2296, August 2002

Regional Seismic-Event Characterization Using a Bayesian Formulation of Simple Kriging

by Steven Bottone, Mark D. Fisk, and Gary D. McCartor

Abstract An approach is presented to calibrate and use regional *P-S* amplitude ratios to improve seismic-event characterization capabilities with regard to monitoring the Comprehensive Nuclear-Test-Ban Treaty. Data for presumed earthquakes are used to estimate distance corrections for *Pn-Sn* and *Pn-Lg* ratios in the 6–8-Hz passband for tectonic and stable-region types. The regional phase-amplitude ratios are further corrected for path variations using simple kriging. Simple kriging is derived using a Bayesian approach. A correction surface is determined for each type of amplitude ratio at each station as an optimal linear combination of existing amplitude-ratio data at the station, giving greater weight to calibration data nearer to the correction location. A corresponding uncertainty surface is also estimated in terms of the residual variance of the data and a calibration variance. For well-calibrated locations, the correction converges to the mean of nearby data, and the uncertainty converges to the residual variance. For locations far from calibration data, the correction surface converges to the worldwide average, with larger uncertainty. With these correction and uncertainty surfaces, corrected values of *Pn*/*Smax*(6–8 Hz) are obtained and used to define a hypothesis test that fixes the significance level with respect to misclassifying explosions. The criterion is applied to 140 explosions at known nuclear-test sites and to 4173 Reviewed Event Bulletin (REB) events above m_b 3.5 (presumed to be mostly earthquakes) with regional recordings between 3° and 17°, *Pn* signal-to-noise ratio (SNR) >2.0, and *Sn* or *Lg* SNR >1.3. At a 0.005 significance level, none of the 140 explosions at any of the known nuclear-test sites are screened out, whereas about 78% of the REB events are screened out. Correcting regional *P-S* ratios for spatial variations improves the screening performance by about 25% over just correcting for distance. The screening results are fairly insensitive to estimates of parameters (correlation length, calibration variance, and residual variance) that are used, along with data, to compute the correction and uncertainty surfaces at each station.

Introduction

Although depth and M_s:m_b are essential methods for characterizing seismic events, the depth and M_s:m_b screening criteria, currently being tested at the prototype International Data Center (pIDC), screen out <50% of the presumed earthquakes >m_b 3.5 in the Reviewed Event Bulletin (REB) (e.g., Fisk *et al.*, 1999b). However, a significant number of REB events have useful regional seismic data. For example, about 20% of onshore events in the REB >m_b 3.5 have regional data with adequate signal-to-noise ratio (SNR). This percentage is expected to increase dramatically as the International Monitoring System (IMS) is completed. Over half of these events do not have M_s:m_b and cannot be screened out as deep. The situation is much more pronounced for REB events near known nuclear-test sites, where none of the events we examined could be confidently screened out as deep and only 24% could be screened out by M_s:m_b (Bottone *et al.*, 2001). In contrast, 89% of the events could be analyzed by using regional data. Thus, regional data must be used to supplement depth and M_s:m_b methods for characterizing/identifying seismic events, particularly those <m_b 4.5.

Numerous studies (e.g., Bennett *et al.*, 1989; Taylor *et al.*, 1989; Baumgardt *et al.*, 1992; Kim *et al.*, 1993; Walter *et al.*, 1995; Fisk *et al.*, 1996; Taylor, 1996; Hartse *et al.*, 1997; Kim *et al.*, 1997; Taylor and Hartse, 1997; and others) have demonstrated that high-frequency regional *P-S* ratios, as well as spectral and cross-spectral ratios, provide useful separation of earthquakes and explosions. However, their effective use is complicated by the fact that regional seis-

mograms and corresponding *P-S* ratios often exhibit significant variations due to path and station effects. This requires that these effects are calibrated on a region/station-specific basis. One approach that has been investigated is regression analysis of topographic and crustal parameters to characterize path effects on *P-S* ratios (e.g., Zhang and Lay, 1994; Zhang *et al.*, 1994, 1996; Lay, 1997; Fan and Lay, 1998; Hartse *et al.*, 1998; Rodgers *et al.*, 1998). Another approach is spatial prediction or interpolation and, more specifically, kriging (e.g., Phillips *et al.*, 1998; Phillips, 1999; Rodgers *et al.*, 1999; Fan *et al.*, 2002). In this article, we develop and apply a Bayesian calibration method that uses prior information and reference data to obtain correction and uncertainty surfaces for path and station effects (for the prior distributions that we use, the correction and uncertainty surfaces are identical with those obtained using simple kriging), for the purpose of event screening.

The basic concept of event screening is to screen out those events that may be considered as natural or nonnuclear, manmade phenomena with high confidence, without screening out any explosions that may correspond to potential violations of the Comprehensive Nuclear-Test-Ban Treaty (CTBT). Events that are not screened out can indicate either that they have explosion-like characteristics or simply that the uncertainties associated with the screening criteria are too large for the event to be screened out with sufficient confidence. No effort is made, within the context of event screening at the pIDC, to further distinguish or identify such events that are not screened out. It should be emphasized that we wish to develop and implement at the pIDC an algorithm that will efficiently screen out events worldwide with little chance of screening out an explosion. Although many studies have shown that more optimal methods exist at particular stations, we attempt to develop here a more conservative procedure that will safely screen out events at stations with few earthquake and no explosion data.

We first discuss data sets that we have compiled and processed, including 268 regional recordings of 161 underground explosions (mostly nuclear) conducted in diverse geological regions and representing a very broad range of magnitudes (M_L 2.4–m_b 6.2) and regional distances (2°–20°). We believe that this is the most comprehensive data set, assembled to date, of processed regional seismograms for nuclear explosions worldwide. We compare these explosions to 4173 presumed earthquakes $>m_b$ 3.5 in the REB, recorded at regional distances by existing IMS seismic stations and processed at the pIDC. We describe these data sets, the regional seismic phase-amplitude measurements, and the SNR restrictions used in this study.

Local calibration of regional amplitude ratios is obtained by a sequence of three steps. First, we apply distance corrections for *Pn-Sn* and *Pn-Lg* amplitude ratios. The corrections are estimated using REB data and depend on region type (tectonic or stable). These distance corrections are used to remove trends from the data before treating region-specific variations and to determine a prior distribution for a background model. Next, we present our Bayesian method of spatial prediction, equivalent to simple kriging with our choice of prior distributions, and apply it at each station to further correct *Pn*/*Sn* and *Pn*/*Lg* measurements for path variations. The algorithm also provides corresponding uncertainty surfaces at each station that treat both the calibration uncertainty and the residual variance observed in the data.

Third, to characterize the events, we use *Pn*/*Smax* in the 6–8-Hz band, where $Smax = \max(Sn, Lg)$, which has been corrected for distance and for path variations by using the local correction surface. We present a hypothesis test that accounts for the varying uncertainty associated with the correction surface to assess whether an event is consistent with the explosion population at a fixed significance level with respect to incorrectly screening out an explosion. We apply the test to the explosion and earthquake data sets acquired under this project. A score is computed such that events with positive scores are screened out, that is, are inconsistent with the hypothesis that the event belongs to the explosion population. Events with scores ≤ 0 are not screened out.

At the 0.005 significance level, none of the available explosions at any of the known nuclear-test sites are screened out, whereas about 78% of the presumed earthquakes in the REB, $>m_b$ 3.5, are screened out. We find that local calibration significantly improves the event-screening performance (by about 25%) over simply applying distance corrections. In addition, over half of these REB events do not have an M_s measurement and cannot be screened out by depth, indicating the potential complementary benefit of including an event-screening criterion based on high-frequency regional *P-S* amplitude ratios.

We also examine the robustness of the calibrated *P*/*S* screening technique to variations in the model parameters, namely, the correlation length, the calibration variance, and the residual variance, that are estimated from available REB data. The results of this robustness study indicate that the method is not very sensitive to variations in these parameters.

Finally, we provide concluding remarks regarding the utility of high-frequency regional *P-S* amplitude ratios for improving CTBT monitoring capabilities, particularly when they are calibrated for region-specific variations.

Data Sets

Figure 1 shows the locations of the explosions and presumed earthquakes used in this study, as well as the locations of 52 IMS stations (reporting to the pIDC) and four non-IMS stations (KEV, WMQ, BRVK [now an IMS auxiliary], and KNB) with regional recordings of these events. Here we summarize the data sets. See Fisk *et al.* (2002) for further details of these data sets and additional plots of the data.

Regional Explosion Data

We have compiled 268 regional seismic recordings of 134 underground nuclear explosions (UNEs), 20 peaceful nuclear explosions (PNEs), and seven underground chemical

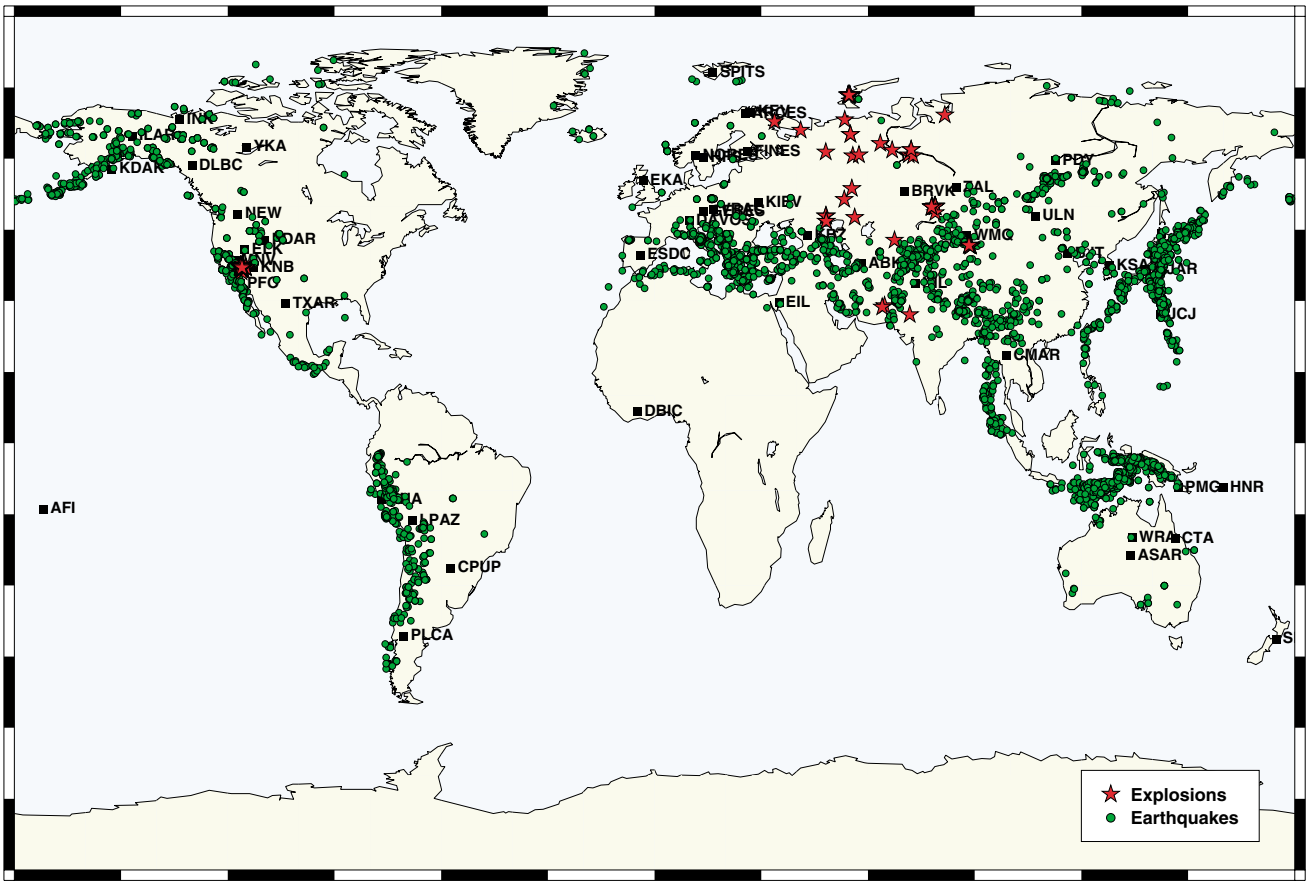

Figure 1. Locations of 161 explosions (stars) and 4173 REB events (circles) > m_b 3.5. Locations of 52 IMS stations and 4 non-IMS stations with regional recordings of these events are also shown.

explosions (UCEs). Table 1 summarizes our current database of regional explosion recordings, listed by station. They represent explosions in very diverse geological regions and cover a magnitude range of about M_L 2.4–m_b 6.2 and a distance range of about 2°–20° to the recording station. *Pn*/*Lg* measurements for UNEs at the Nevada Test Site (NTS), recorded by MNV and KNB, were provided by Patton and Walter (pers. comm., 1994). Regional waveforms from station WMQ in China were provided by Baumgardt (pers. comm., 1993). BRVK recordings of 16 PNEs, conducted by the former Soviet Union (FSU), were provided by Murphy (pers. comm., 1999).

Maximum amplitudes of *Pn, Pg, Sn*, and *Lg* were measured on vertical channels in 2–4, 4–6, 6–8, and 8–10-Hz bands (only the 6–8-Hz measurements are used in the screening analysis presented later). Root-mean-square (rms) beams of vertical channels were used for arrays. All but the MNV and KNB data were processed off-line at the pIDC in the same manner as the earthquake data described later. Measurement of MNV and KNB amplitudes are described in detail in Walter et al., (1995). Not all events have complete sets of amplitude measurements. For example, for the six UNEs at Novaya Zemlya (NZ) recorded by KEV in Finland, three of which were also recorded by ARCES in Norway, *Lg* was not measured owing to blockage.

Regional Earthquake Data

More than 5300 regional waveforms, recorded by 52 IMS seismic stations and processed at the pIDC, are used in this study. These recordings correspond to 4173 distinct REB events with m_b >3.5 and depths <40 km that are presumed to be earthquakes (see Fig. 1). The restriction to m_b 3.5 and above is used to reduce the possibility of mining blasts contaminating the earthquake data sets, based on an assessment by Khalturin *et al.* (1997) that few regions worldwide conduct blasting above this level. We also limit these recordings to distances beyond 3° to avoid problems of distinguishing *Pn* and *Pg* and to those <17°, although regional phase amplitudes are computed at distances up to 20° at the pIDC. In addition, we limit this study to data from IMS stations that have at least 20 recordings so that useful training sets may be formed.

Regional phase time-domain amplitudes were measured for the earthquake data, as for the explosions, using the Detection and Feature Extraction code at the pIDC (see IDC5.2.1, 1999). Amplitude-measurement types include ab-

Table 1
Summary of Regional Seismic Data for Underground Explosions

Station Code	Distance (°)	Magnitude (m_b or M_L)	Nuclear Test Site	Date of Explosion (yr/mo/day)
KEV	9.4	5.52–5.90	NZ UNEs	821011, 841025, 870802, 880507, 881204, 901024
	3.1–19.5	4.70–5.50	FSU PNEs	820904, 840811, 840825, 840827, 850718
ARCES	9.9	5.60–5.90	NZ UNEs	880507, 881204, 901024
ZAL	12.2	4.71	Lop Nor UNEs	960729
	6.2, 5.5, 5.5, 6.0, 6.0, 6.0	3.95, N/A, N/A, 3.80, 3.67, 4.04	STS UCEs	970803, 970831, 970928, 980822, 990925, 000729
NIL	14.4–14.7	5.73, 5.54, 5.69, 4.71	Lop Nor UNEs	950515, 950817, 960608, 960729
	6.7	5.00	India UNE	980511
ULN	14.3	5.73, 5.54	Lop Nor UNEs	950515, 950817
WMQ	8.6	4.6–6.1	STS UNEs	870606, 870629, 870717, 870802, 871115, 871213, 871220, 871227, 880206, 880213, 880403, 880422, 880504, 880914, 881018, 881123, 881217, 890212, 890217, 890708, 891004
	2.2	4.7	Lop Nor UNE	880929
BRVK	16.8–16.9	6.00, 5.73, 5.54, 5.69	Lop Nor UNEs	941007, 950515, 950817, 960608
	6.2	5.70–7.20	STS UNEs	880213, 880403, 880504, 880614, 880914, 881217
	5.8–6.2	3.95, N/A, N/A, 3.80, N/A, 3.67, 4.04	STS UCEs	970803, 970831, 970928, 980822, 980917, 990925, 000729
	7.2–17.2	4.4–5.8	FSU PNEs	730815, 731026, 770930, 781017, 790714, 791004, 791024, 801008, 801210, 810525, 810902, 821016, 840811, 870419, 871003, 880906
MAKZ	6.9–7.2	6.00, 5.73, 5.69, 4.71	Lop Nor UNEs	941007, 950515, 960608, 960729
MNV	1.7–2.2	2.4–5.5	NTS UNEs	93 total (1979–1992)
KNB	2.5–2.8	2.6–5.5	NTS UNEs	90 total (1979–1992)

solute maximum amplitude, maximum peak to trough, maximum peak to peak, (rms) amplitude, and so on. Beam types include coherent, steered, incoherent, and rms measured on vertical, radial, or transverse components. Time intervals for the measurements can be based on predicted travel times of seismic phases or on fixed group-velocity windows, or they may be computed from observed arrival times in the database.

The following time-domain regional-phase amplitude measurements are computed for each seismic station that is associated to and within 20° of an analyst-reviewed event:

(1) absolute maximum amplitude on 2–4, 4–6, 6–8, 8–10, 10–12, and 12–14-Hz rms beams for predicted time–velocity windows around *Pn, Pg, Sn*, and *Lg* phases: *Pn*: 8 sec before the theoretical arrival time of *Pn* to a group velocity of 6.4 km/sec; *Pg*: 6.3–5.8 km/sec group velocity; *Sn*: 5 sec before the theoretical arrival time of *Sn* and a 20-sec duration; and *Lg*: 3.7–3.0 km/sec group velocity;

(2) absolute maximum amplitude on 2–4, 4–6, 6–8, 8–10, 10–12, and 12–14-Hz rms beams for predicted time and/or velocity windows around pre-*Pn*, pre-*Pg*, pre-*Sn*, and pre-*Lg* noise: pre-*Pn*: 13 sec before the theoretical arrival time of *Pn* and a 5-sec duration; pre-*Pg*: 6.4–6.3 km/sec group velocity; pre-*Sn*: 10 sec before the theoretical arrival time of *Sn* and a 5-sec duration; and pre-*Lg*: 3.8–3.7 km/sec group velocity.

The windows for the crustal modes (*Pg* and *Lg* as well as pre-*Pg* and pre-*Lg*) are defined in terms of the well-defined group-velocity bounds that characterize such phases. In contrast, the windows for the body waves (*Pn* and *Sn* as well as pre-*Pn* and pre-*Sn*) are defined (with a single exception) in terms of the arrival time, which delineates such phases with less ambiguity. The instrument-response correction at the band-center frequency is applied to all of the amplitude measurements.

For this study, we use *Pn*/*Sn* and *Pn*/*Lg* in the 6–8-Hz band that satisfy the SN criteria of *Pn* signal divided by pre-*Pn* noise >2.0 and *S* (*Sn* or *Lg*) signal divided by pre-*S* noise >1.3, which follows Jenkins *et al.* (1998). As described there, a lower threshold is used for *S* phases because pre-*S* noise includes coda from earlier arrivals. Without this lower threshold, *S* phases for the Lop Nor explosions would not

have met the SNR criterion. We also applied an outlier test to eliminate many events that had anomalously high or low values of *P*/*S*. Waveforms for many suspicious events at stations near test sites were visually inspected and eliminated from the data set if spikes or other problems were observed.

In addition, we have acquired BRVK recordings of 61 regional earthquakes and WMQ recordings of 24 earthquakes in or near China. These waveforms were processed off-line at the pIDC in the same manner as the IMS data. Lastly, our data set includes 60 recordings by MNV and 75 by KNB of earthquakes at NTS. *Pn*/*Lg* values were processed by Patton and Walter (pers. comm., 1994) in the same way as for the 99 NTS UNEs (see Walter *et al.*, 1995). These data are needed to properly calibrate MNV and KNB.

Distance Corrections

Because regional *P* and *S* phases typically attenuate and spread at different rates, *P*-*S* ratios for events recorded at varying distances must be corrected if they are to be compared (e.g., Sereno, 1990; Bottone *et al.*, 1997; Fisk *et al.*, 1998; Jenkins *et al.*, 1998; Taylor and Hartse, 1998; Fisk *et al.*, 1999a, 2000). The station-by-station calibration for path variations described subsequently will account for distance dependence, and therefore, removing a distance-dependent trend is not absolutely necessary. However, most spatial-prediction methods work better if detrending is done first (Cressie, 1993). Because we wish to calibrate for path variations at stations with little earthquake data and no explosion data, we choose to have the correction surfaces that approach the worldwide average in regions or at stations with no calibration data. To establish a prior distribution to be used as the background model in the Bayesian approach described later that is independent of station, it is necessary, for consistency, to use station-independent distance corrections to detrend the data. If it could be justified to use a background model that was station dependent, where the correction surface would approach the station average in regions with no calibration data, then station-dependent distance corrections would be appropriate and ordinary kriging could be used to correct for path variations.

To detrend the data, the distance dependence is modeled by a three-parameter equation of the form

$$\log(P/S) = a + b \log\Delta + c\Delta \qquad (1)$$

where Δ is the epicentral distance from the event to the recording station, and *a, b*, and *c* are parameters depending on the particular type of amplitude ratio (e.g., *Pn*/*Lg* or *Pn*/*Sn*), frequency band, and region type. Following Jenkins *et al.* (1998), we divide the world into tectonic and stable-region types and estimate the three coefficients for each region type by using all available REB data in such regions. Dividing the world into tectonic and stable-region types offers a simple way of detrending the data with respect to attenuation and geometric spreading. We then correct the *Pn*-*Lg* and *Pn*-*Sn* amplitude ratios for a given event by using equation (1) with the appropriate coefficients for the specific type of amplitude ratio, frequency band, and region type in which the event occurred.

Figure 2 shows the distance dependence of the REB events (circles) in tectonic (top) and stable region (bottom) types for *Pn*/*Lg* (left) and *Pn*/*Sn* (right) in the 6–8-Hz band. Also shown are the corresponding *P*/*S* values for the explosions in our data set, with the same marker types as those defined in the legend of Figure 3. The data in these plots exhibit considerable scatter owing to regional path variations, station dependence, and other effects (e.g., focal mechanism and near-source effects). This scatter will be reduced by using the local calibration method, described later, to treat station dependence and region-specific path variations.

We compute *Pn*-*Lg* and *Pn*-*Sn* phase ratios by using *P* and *S* amplitude measurements in the same frequency band to minimize source size effects. A problem with using all *Pn*/*Lg* and *Pn*/*Sn* values is that *Lg* or *Sn* may be severely attenuated or blocked for some paths. For example, *Lg* is typically blocked along many oceanic or other thin crustal paths, and *Sn* is severely attenuated in the western United States and the Middle East. *Lg* is also blocked along continental paths with crustal thickness variations and severely attenuated where crustal attenuation is high because of hot, actively deforming crusts (e.g., Tibet, Turkish–Iranian plateau). Because it is difficult to map all such paths, we will use *Pn*/*Smax* values in the screening procedure defined later, where *Smax* = max(*Sn, Lg*); that is, *Smax* is the *Lg* or *Sn* phase that propagates more efficiently in a given region.

Higher-frequency bands typically provide better separation of explosion and earthquake populations (e.g., Blandford, 1995; Walter *et al.*, 1995; Fisk *et al.*, 1996; Kim *et al.*, 1997). However, signals at frequencies >8 Hz are often below noise levels. We have found that the best overall performance of the hypothesis test presented later is obtained by using log[*Pn*/*Smax*(6–8 Hz)]. Taking the logarithm results in populations that are more normally distributed (Bottone *et al.*, 1997), thus allowing a simpler statistical treatment.

Figure 3 shows a scatterplot of log[*Pn*/*Smax*(6–8 Hz)], after applying distance corrections, for 140 explosions (134 of which were NEs) and 4173 earthquakes. Figure 4 shows a histogram of the same data. As can be seen in either plot, if the explosion with the smallest value of distance-corrected log[*Pn*/*Smax*(6–8 Hz)] were to be used as a screening threshold, then only about 50% of the earthquakes would be screened out. It will be seen subsequently that local calibration, using a Bayesian calibration technique, improves the screening performance; that is, the earthquake screening rate is increased while keeping the explosion screening rate at a fixed significance level.

Regional Corrections Using a Bayesian Formulation of Simple Kriging

The global-average distance corrections for tectonic and stable-region types, described previously, do not depend on

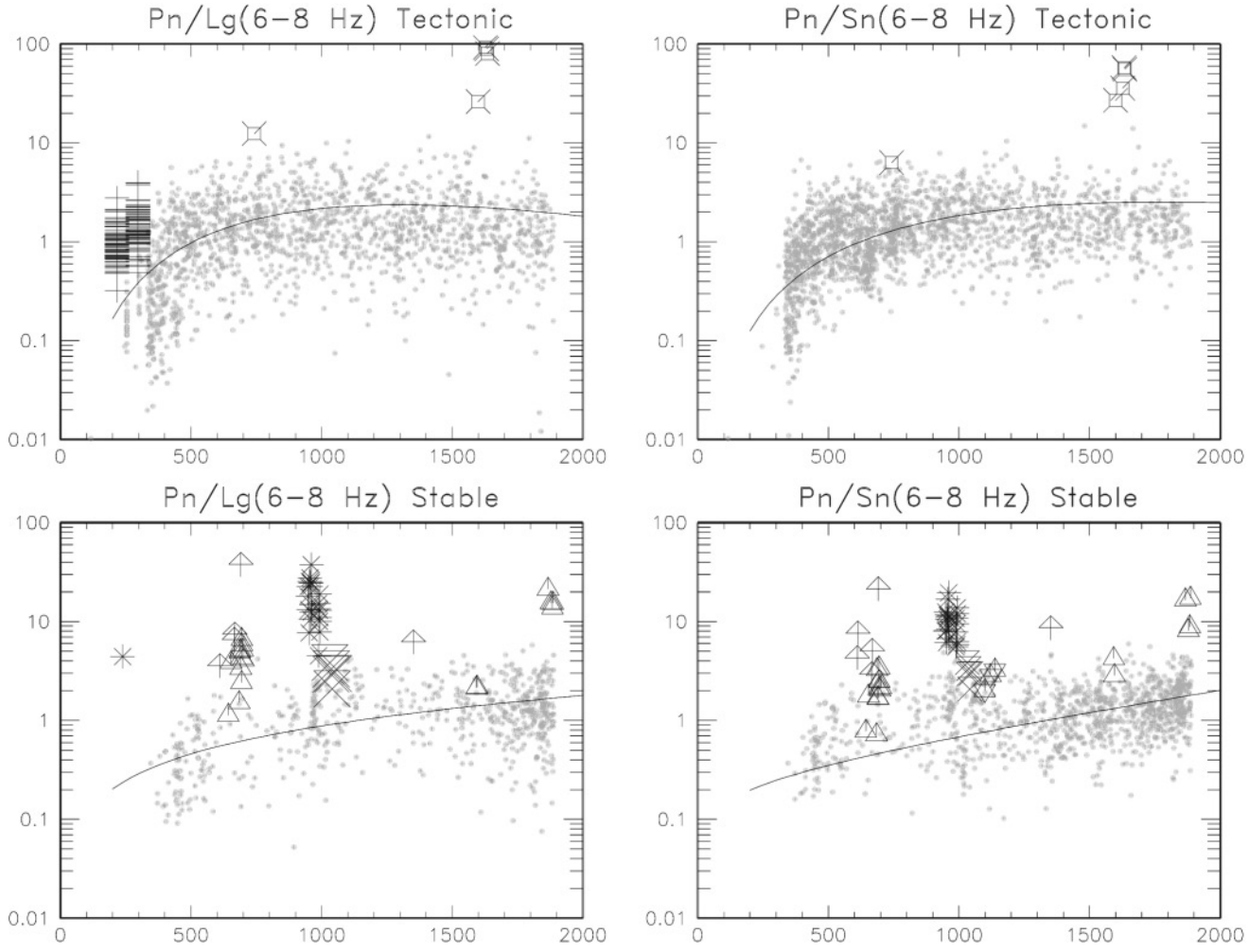

Figure 2. Uncorrected *Pn*/*Lg* (left) and *Pn*/*Sn* (right) in the 6–8-Hz band versus distance (in km) for REB events and explosions worldwide in tectonic (top) and stable (bottom) region types. The curves in each plot are the best-fit estimates of the distance dependence for the earthquakes. See Fig. 3 for legend.

the station and do not account for region-specific (or local) variations in the *P*-*S* amplitude ratios due to path effects. One way of treating path variations is to use optimal spatial prediction, which in various forms is often called kriging (Cressie, 1993). Among other geophysical applications, kriging has been used to calibrate and treat uncertainties of seismic travel times (e.g., Schultz *et al.*, 1998), regional *P*-*S* ratios (e.g., Phillips, 1999; Rodgers *et al.*, 1999; Fan *et al.*, 2002), and 20-sec Rayleigh-wave group velocities (Pasyanos, 2000).

Optimal spatial prediction generally refers to making inferences at a new location, given previously measured data at N separate locations, which minimizes the uncertainty of that prediction under a given set of statistical assumptions. The set of predictions at all spatial locations may then be used as a correction surface. If calibration data for a station are limited to only certain subregions, possibly with anomalous propagation properties, a prediction in another subregion, far removed from these data, may be incorrect. Thus, in subregions with limited or no calibration data, we wish for the prediction to converge to a value that is the average over all data worldwide, with an appropriately large uncertainty. To achieve this limiting behavior, we develop an optimal spatial predictor using a Bayesian method. For the prior distributions we use (normal distributions), it will turn out that our predictor is identical to simple kriging. However, we present the method in some detail to demonstrate certain advantages. Choices of prior distributions other than normal will lead to nonlinear predictors (and, therefore, not to kriging), although the integrations then become nontrivial. This method is also easily generalized to have a prior for the variance as well as the mean (even a normal distribution for a prior on the variance would lead to nontrivial integrations and nonlinear predictors). The use of a background model to describe limiting behavior for each station, estimated from global data, is more transparent in this formulation, and for any distribution other than normal, it would be necessary to use such a formulation. The resulting predictor in this

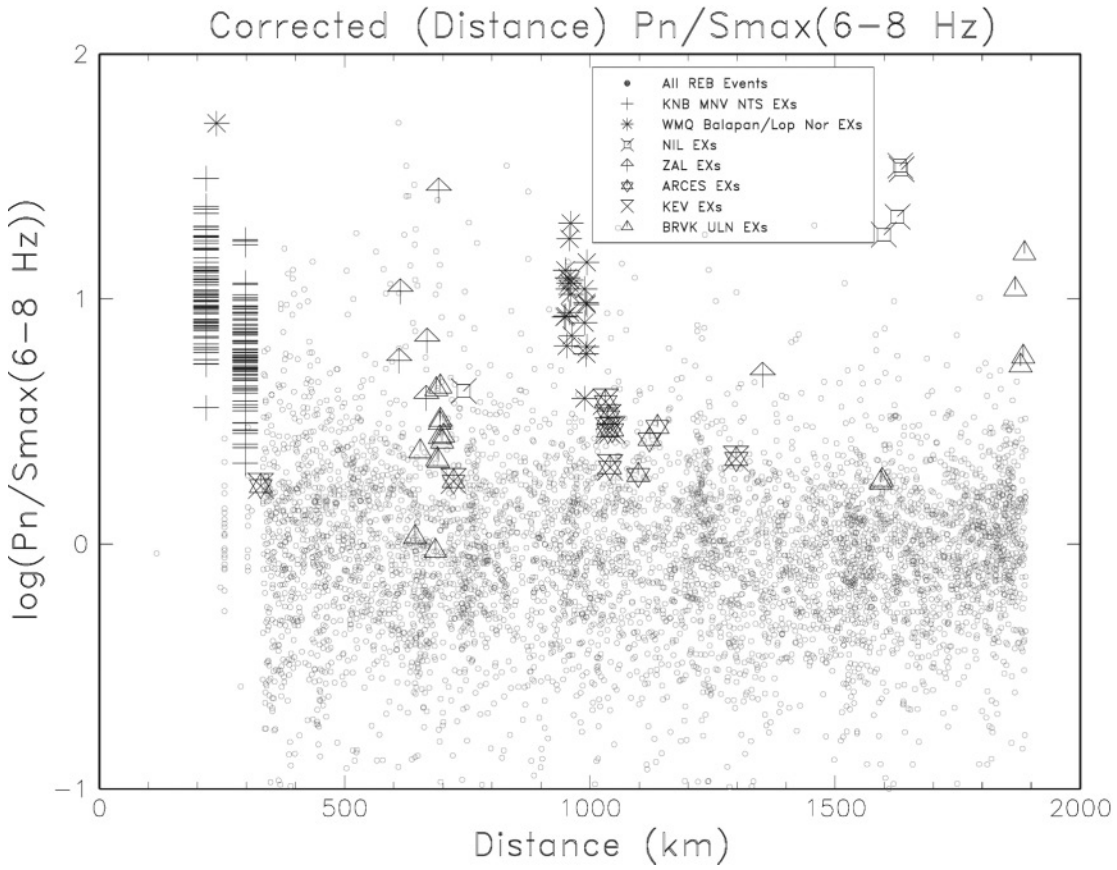

Figure 3. Distance-corrected log[*Pn*/*Smax*(6–8 Hz)] for 140 explosions and 4173 REB events.

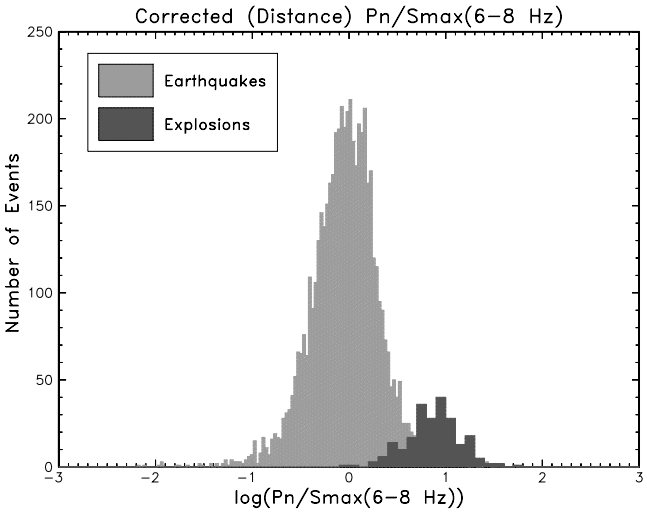

Figure 4. Histograms of distance-corrected log[*Pn*/*Smax*(6–8 Hz)] for 140 explosions (right) and 4173 earthquakes (left). If the explosion with the smallest value were used as a screening threshold, only about 50% of the earthquakes would be screened out.

method is a statistical distribution, so that statistical inferences, such as confidence intervals, can be made naturally. To make statistical inferences using kriging, distributional assumptions must be made independent of the prediction. The advantage of using simple kriging in the usual formulation is an easier, well-known, calculation. The generalization to simple kriging developed by Schultz *et al.* (1998) could be termed locally weighted kriging (Shumway, pers. comm., 2000). It that method, prediction at points far from data approach a background model with a rate depending on damping functions, which are determined by geophysical knowledge of the region. If the damping functions are equal to one, the method reduces to simple kriging. Because the Bayesian predictor derived subsequently will turn out to be equivalent to simple kriging, those interested only in the application to the regional data set can skip to the Examples section.

Theory

Suppose there are N data values at locations $\mathbf{s}_1, \ldots, \mathbf{s}_N$ for a particular regional amplitude ratio (e.g., *Pn*/*Sn* or *Pn*/*Lg*) recorded by a given station. Let $\mathbf{x} = (x_1, \ldots, x_N)'$ be an N-dimensional data vector, where the x_i, $i = 1, \ldots, N$ are distance-corrected values of log(*P*/*S*). That is,

$$x_i = x(\mathbf{s}_i) = \log[P/S]_i - \hat{a} - \hat{b}\log\Delta_i - \hat{c}\Delta_i \quad (2)$$

where $\hat{a}$, $\hat{b}$, and $\hat{c}$ are the best-fit estimates of a, b, and c in equation (1) for the appropriate amplitude ratio, frequency, and region type, and Δ_i is the epicentral distance between location $\mathbf{s}_i$ and the station. We normalize the data so that the distance-corrected worldwide average is zero (merging data from different region types may lead to a nonzero mean).

If a given station receives data from a location, $\mathbf{s}_i$, we assume that the data (x_i) can be modeled by a random variable with a normal distribution. We assume that the mean, u_i, depends on the location (and the station) but that the variance, σ_r^2 , does not. That is,

$$x_i \sim N(u_i, \sigma_r^2) \quad (3)$$

We are assuming that the mean at a given location, u_i, is a deterministic quantity; however, if we choose a random point on the Earth, the mean at that point, u_i, is a random variable because that point is a random variable. We assume that this random variable is also normally distributed, with zero mean and constant variance, σ_c^2 . We further assume that the correlation between any two u_i and u_j depends only on the distance separating the two locations, that is,

$$\rho_{ij} = \mathrm{corr}(u_i, u_j) = f[\Delta(\mathbf{s}_i, \mathbf{s}_j)] \quad (4)$$

where $\Delta(\mathbf{s}_i, \mathbf{s}_j)$ is the epicentral distance between locations $\mathbf{s}_i$ and $\mathbf{s}_j$. The function f must be a positive definite function, ensuring that the covariance matrix of the random vector of means is positive-definite. The form of this function is usually taken to be a known positive-definite function with unknown parameters estimated from the available data. We generally use

$$f(\Delta) = \exp(-\Delta/\alpha) \quad (5)$$

where α, the correlation length, is estimated from the data.

We now wish to use the N data values, $\mathbf{x}$, located at $\mathbf{s}_1, \ldots, \mathbf{s}_N$ to calibrate another location, $\mathbf{s}_0$. To do this, we can determine the distribution, $p(u_0|\mathbf{x})$, for the mean at a location $\mathbf{s}_0$, given the N data values, $\mathbf{x}$, located at $\mathbf{s}_1, \ldots, \mathbf{s}_N$. The expectation value of this distribution can be used as an estimator of u_0 at that location, which will be the correction value. The variance of u_0 can be used as the uncertainty of the correction. Under our assumptions, we know $p(x_0, \mathbf{x}|u_0, \mathbf{u})$ and $p(u_0, \mathbf{u})$, where $\mathbf{u} = (u_1, \ldots, u_N)'$, so we can use Bayes's theorem to calculate $p(u_0|\mathbf{x})$.

The conditional probability density $p(u_0|\mathbf{x})$ may be written

$$p(u_0|\mathbf{x}) = \frac{p(\mathbf{x}, u_0)}{p(\mathbf{x})} = \frac{\int p(\mathbf{x}, u_0, \mathbf{u})\, d\mathbf{u}}{p(\mathbf{x})} = \int p(u_0, \mathbf{u}|\mathbf{x})\, d\mathbf{u} \quad (6)$$

Bayes's theorem for probability distributions (Papoulis, 1984) gives

$$p(u_0, \mathbf{u}|\mathbf{x}) = \frac{p(\mathbf{x}|u_0, \mathbf{u})\, p(u_0, \mathbf{u})}{p(\mathbf{x})} \quad (7)$$

However,

$$p(\mathbf{x}|u_0, \mathbf{u}) = \frac{p(\mathbf{x}, u_0, \mathbf{u})}{p(u_0, \mathbf{u})} = \frac{\int p(x_0, \mathbf{x}, u_0, \mathbf{u})\, dx_0}{p(u_0, \mathbf{u})} = \int p(x_0, \mathbf{x}|u_0, \mathbf{u}) dx_0 \quad (8)$$

Substituting equations (7) and (8) into equation (6) then gives

$$p(u_0|\mathbf{x}) \propto \int [\int p(x_0, \mathbf{x}|u_0, \mathbf{u}) dx_0]\, p(u_0, \mathbf{u})\, d\mathbf{u} \quad (9)$$

Because integrals of normal distributions are normal distributions, $p(u_0|\mathbf{x})$ will be also be normally distributed, so the proportionality constant need not be determined explicitly. Substituting the appropriate distributions for $p(x_0, \mathbf{x}|u_0, \mathbf{u})$ and $p(u_0, \mathbf{u})$ as described earlier and performing the integral over x_0 gives

$$p(u_0|\mathbf{x}) \propto \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{v})' A_1 (\mathbf{y} - \mathbf{v})\right] \exp\left[-\frac{1}{2}\mathbf{v}'\Sigma^{-1}\mathbf{v}\right] d\mathbf{u} \quad (10)$$

where $\mathbf{y}$ and $\mathbf{v}$ are $(N + 1)$-dimensional vectors given by $\mathbf{y} = (0, \mathbf{x}')'$ and $\mathbf{v} = (u_0, \mathbf{u}')'$. A_1 is an $(N + 1)$-by-$(N + 1)$-dimensional matrix given by

$$A_1 = \frac{1}{\sigma_r^2}\begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & I_N \end{bmatrix} \quad (11)$$

where I_N is the identity matrix in N dimensions. Σ is an $(N + 1)$-by-$(N + 1)$-dimensional covariance matrix such that

$$\Sigma_{ij} = \sigma_c^2 \rho_{ij},\ i, j = 0, 1, \ldots, N \quad (12)$$

To compute the integral in equation (10), we combine the arguments of the exponentials, giving

$$(\mathbf{y} - \mathbf{v})' A_1 (\mathbf{y} - \mathbf{v}) + \mathbf{v}'\Sigma^{-1}\mathbf{v} = (\mathbf{v} - \tilde{\mathbf{v}})' S^{-1} (\mathbf{v} - \tilde{\mathbf{v}}) + \text{const}, \quad (13)$$

where

$$S^{-1} = \Sigma^{-1} + A_1 \quad (14)$$

and

$$\tilde{\mathbf{v}} = S\,A_1\mathbf{y} \tag{15}$$

The integral now becomes

$$p(u_0|\mathbf{x}) \propto \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(\mathbf{v}-\tilde{\mathbf{v}})'S^{-1}(\mathbf{v}-\tilde{\mathbf{v}})\right]d\mathbf{u} \tag{16}$$

This integral is computed using the theorem that, if a random vector, $\mathbf{v}$, is distributed as $N(\tilde{\mathbf{v}},S)$, then the marginal distribution of any set of components of $\mathbf{v}$ is multivariate normal with means and covariance matrix obtained by taking the corresponding components of $\tilde{\mathbf{v}}$ and S. This gives

$$p(u_0|\mathbf{x}) \propto \exp\left[-\frac{1}{2}(S_{00})^{-1}(u_0-\tilde{u}_0)^2\right] \tag{17}$$

that is, the distribution of u_0, given $\mathbf{x}$, is the normal, $N(\tilde{u}_0, \sigma^2)$, where

$$\tilde{u}_0 = (S\,A_1\mathbf{y})_0 \tag{18}$$

and

$$\sigma^2 = S_{00} = [(\Sigma^{-1}+A_1)^{-1}]_{00} \tag{19}$$

It can be shown that equations (18) and (19) are equivalent to simple kriging, although the variance given in equation (19) is smaller by σ_r^2, since we are predicting the mean. Our covariance matrices are automatically chosen to avoid exact interpolation and will not give singular matrices if there are co-located data points.

A correction surface is obtained by computing $\tilde{u}_0(\mathbf{s}_0)$ for all points of $\mathbf{s}_0$ within regional distances of the station. An uncertainty surface is given by $\sigma^2(\mathbf{s}_0)$. In principle, such surfaces can be estimated at each station from either earthquake or explosion data. However, owing to the relatively limited amount and spatial distribution of explosion data at most stations, it is only practical to estimate the calibration surfaces from earthquake data. We assume that the calibration surfaces to treat station and path effects are independent of source type, probably not true, but necessary because of the lack of explosion data.

An explosion data value, x, equal to distance-corrected log[P/S(6–8 Hz)] at location $\mathbf{s}_0$ can be compared with the earthquake surface by computing the difference $x(\mathbf{s}_0) - \tilde{u}_0(\mathbf{s}_0)$. This should be a large positive number if explosions are to be distinguished from earthquakes. Any unknown event at $\mathbf{s}_0$ can also be compared with the surface in the same way. However, if we wish to compare any of the N earthquakes used to determine the surface with a prediction at that location, we must remove that earthquake from the data set and compute a new surface value at that location by using the $N-1$ remaining earthquakes (the leave-one-out procedure). This is also called cross-validation and has been used by Rodgers *et al.* (1999) to validate correction surfaces.

Parameter Estimation

There are three parameters that must be estimated by using all available data: σ_c^2, the calibration variance; σ_r^2, the residual variance; and α, the correlation length. One way of estimating these parameters is using the variogram. Let the data, x_i, at location $\mathbf{s}_i$, be modeled by

$$x_i = x(\mathbf{s}_i) = u_i + e_i \tag{20}$$

where the e_i are independent of each other and of the u_i. By equations (3) and (11), we have

$$\text{var}(e_i) = \sigma_r^2,\ \text{var}(u_i) = \sigma_c^2,\ \text{cov}(u_i, u_j) = \sigma_c^2\,\rho_{ij} \tag{21}$$

where ρ_{ij} depends only on the distance, $\Delta(\mathbf{s}_i, \mathbf{s}_j)$, between $\mathbf{s}_i$ and $\mathbf{s}_j$, and is given by equations (4) and (5). From these relationships, it follows that the variance of x_i is

$$\text{var}(x_i) = \sigma_c^2 + \sigma_r^2 \tag{22}$$

and the covariance of x_i and x_j, $i \neq j$, is,

$$\text{cov}(x_i, x_j) = \text{cov}(u_i, u_j) = \sigma_c^2\rho_{ij} = \sigma_c^2\exp(-\Delta(\mathbf{s}_i, \mathbf{s}_j)/\alpha),\ i \neq j \tag{23}$$

The variogram, 2γ, is defined by

$$2\gamma(h) = \text{var}(x_i - x_j) \tag{24}$$

where the variogram is assumed to depend on only the distance, h, between the locations of x_i and x_j. Expanding the variance gives

$$2\gamma(h) = \text{var}(x_i) + \text{var}(x_j) - 2\,\text{cov}(x_i, x_j) = 2\sigma_r^2 + 2\sigma_c^2\,[1-\exp(-h/\alpha)] \tag{25}$$

A plot of the semivariogram, $\gamma(h)$, as a function of h, asymptotically approaches $\sigma_c^2 + \sigma_r^2$ as h gets large, and intersects the $h = 0$ axis at a value of σ_r^2.

To estimate $2\gamma(h)$, one can average the squares of the difference of all pairs of data values with locations separated by h, for various values of h. In practice, there are no pairs of points separated by exactly h, so one averages over points with separations between $h - \Delta h$ and $h + \Delta h$. The classic estimator is given by

$$2\bar{\gamma}(h) = \frac{1}{|N(h)|}\sum_{N(h)}[x(\mathbf{s}_i) - x(\mathbf{s}_j)]^2 \tag{26}$$

where

$$N(h) \equiv \{(\mathbf{s}_i, \mathbf{s}_j); |\Delta(\mathbf{s}_i, \mathbf{s}_j) - h| \leq \Delta h; i, j = 1, \ldots, N\} \quad (27)$$

and $|N(h)|$ is the number of distinct pairs in $N(h)$. A more robust estimator (to contamination by outliers), given by Cressie and Hawkins (1980), is

$$2\bar{\gamma}(h) = \frac{\left(\frac{1}{|N(h)|}\sum_{N(h)} |x(\mathbf{s}_i) - x(\mathbf{s}_j)|^{1/2}\right)^4}{0.457 + 0.494/|N(h)|} \quad (28)$$

Estimated semivariograms are computed using equation (28) by summing over all distinct pairs with lag h, measured at the same station. Figure 5 shows a plot of an estimated semivariogram computed using the robust estimator given by equation (28), for the case of distance-corrected log[*Pn*/*Lg*(6–8 Hz)] for events in tectonic regions. Semivariograms for log [*Pn*/*Sn*(6–8 Hz)] in tectonic regions and for log[*Pn*/*Lg*(6–8 Hz)] and log[*Pn*/*Sn*(6–8 Hz)] in stable regions have also been computed and show similar properties. In the figure, the squares are the estimates at 1° intervals (or lags), with Δh set at 0.5°. The solid line represents a nonlinear best fit of a model semivariogram of the form

$$\gamma(h) = c_r + c_c[1 - \exp(-h/c_\alpha)] \quad (29)$$

with best-fit coefficients $\hat{c}_r$, $\hat{c}_c$, and $\hat{c}_\alpha$ determined by minimizing the square of the residuals and corresponding to σ_r^2, σ_c^2, and α in equation (25). Table 2 summarizes the fits for each of the four cases. The best-fit equation for log[*Pn*/*Lg*(6–8 Hz)] in tectonic regions is shown on Figure 5. As shown in Table 2, the correlation length is typically about 6° for all cases. The calibration and residual variances are closer to being equal in the tectonic cases, which represent many more distinct pairs than the stable cases. The standard deviations (square root of variance) are between 0.2 and 0.3. Owing to the uncertain nature of these parameters (not enough pairs are available at each lag, so the uncertainty is large), we set $\sigma_c = \sigma_r = 0.25$. Sensitivity of the final results to changes in these parameters is limited and will be examined later.

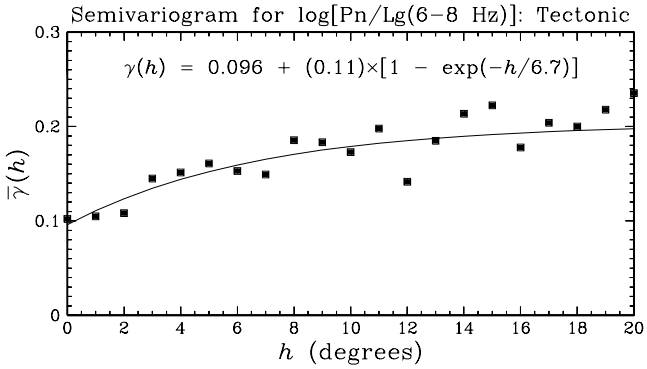

Figure 5. Estimated semivariogram for distance-corrected log[*Pn*/*Lg*(6–8 Hz)] for all events in tectonic regions (squares) and best-fit semivariogram using an exponential model (solid line).

Table 2
Best-Fit Coefficients for Semivariograms

Amplitude Ratio	Region Type	$\hat{c}_r$	$\hat{c}_c$	$\hat{c}_\alpha$
log[*Pn*/*Lg*(6–8 Hz)]	Tectonic	0.096	0.11	6.7
log[*Pn*/*Sn*(6–8 Hz)]	Tectonic	0.036	0.051	5.1
log[*Pn*/*Lg*(6–8 Hz)]	Stable	0.043	0.11	5.0
log[*Pn*/*Sn*(6–8 Hz)]	Stable	0.018	0.085	6.0

Examples

Figure 6a shows a typical correction surface and Figure 6b the corresponding uncertainty (variance) surface for distance-corrected log[*Pn*/*Sn*(6–8 Hz)] at station NIL. Blue colors correspond to values less than the worldwide average (green) and red to values greater than the worldwide average. At locations far from available data, the correction surface approaches the distance-corrected worldwide average (zero value). Locations of NEs (at the Indian and Lop Nor test sites) and earthquakes recorded by NIL are depicted by triangles and crosses, respectively. Black (white) markers indicate values that are greater (less) than the local value of the calibration surface. The size of the marker is proportional to the absolute value of the distance from the correction surface. Thus, very explosion-like values correspond to large black markers. This surface is similar to that given in Fan *et al.* (2002), although that surface was computed without a distance trend removed. Figure 6b shows that the uncertainty is lower in areas with many calibration data points, whereas it is higher in areas where there are no calibration data.

As another example, Figure 7a shows the correction surface and Figure 7b the uncertainty surface for distance-corrected log[*Pn*/*Sn*(6–8 Hz)] at ARCES. Included in Figure 7 are the locations of six UNEs at the NZ test site that were recorded by ARCES and/or KEV and five Soviet PNEs that were recorded by KEV (events recorded at KEV, about 50 km away from ARCES, are shown for comparison). Note that near the NZ test site, where the large black triangles represent the historic UNEs, the uncertainty is relatively large owing to the lack of earthquake calibration data in this area. As will be seen later, in regions of large calibration uncertainty, it is less likely that an event will be screened out than for an event in a well-calibrated region that is the same distance from the correction surface.

Event-Screening Criterion and Score

Methodology

Using the distance- and path-corrected data and associated uncertainties described previously, an event-screening criterion can be developed as a hypothesis test, with a fixed significance level with respect to screening out an explosion.

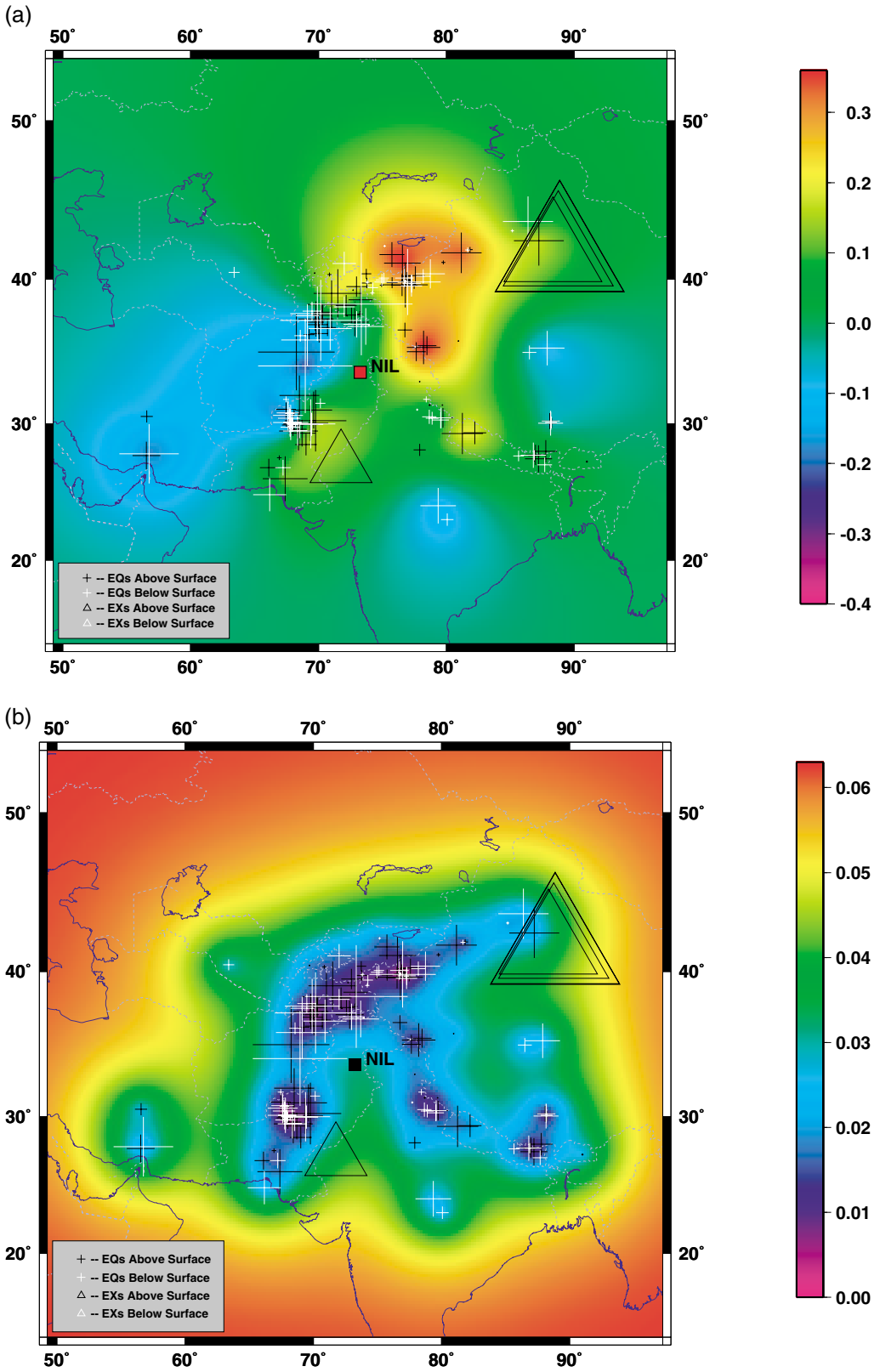

Figure 6. Correction (a) and uncertainty (b) surfaces for log[*Pn*/*Sn*(6–8 Hz)] at NIL. Triangles and crosses represent explosions and earthquakes, respectively. Black (white) markers have values greater (less) than the local value of the correction surface. Marker size is proportional to the absolute distance from the correction surface.

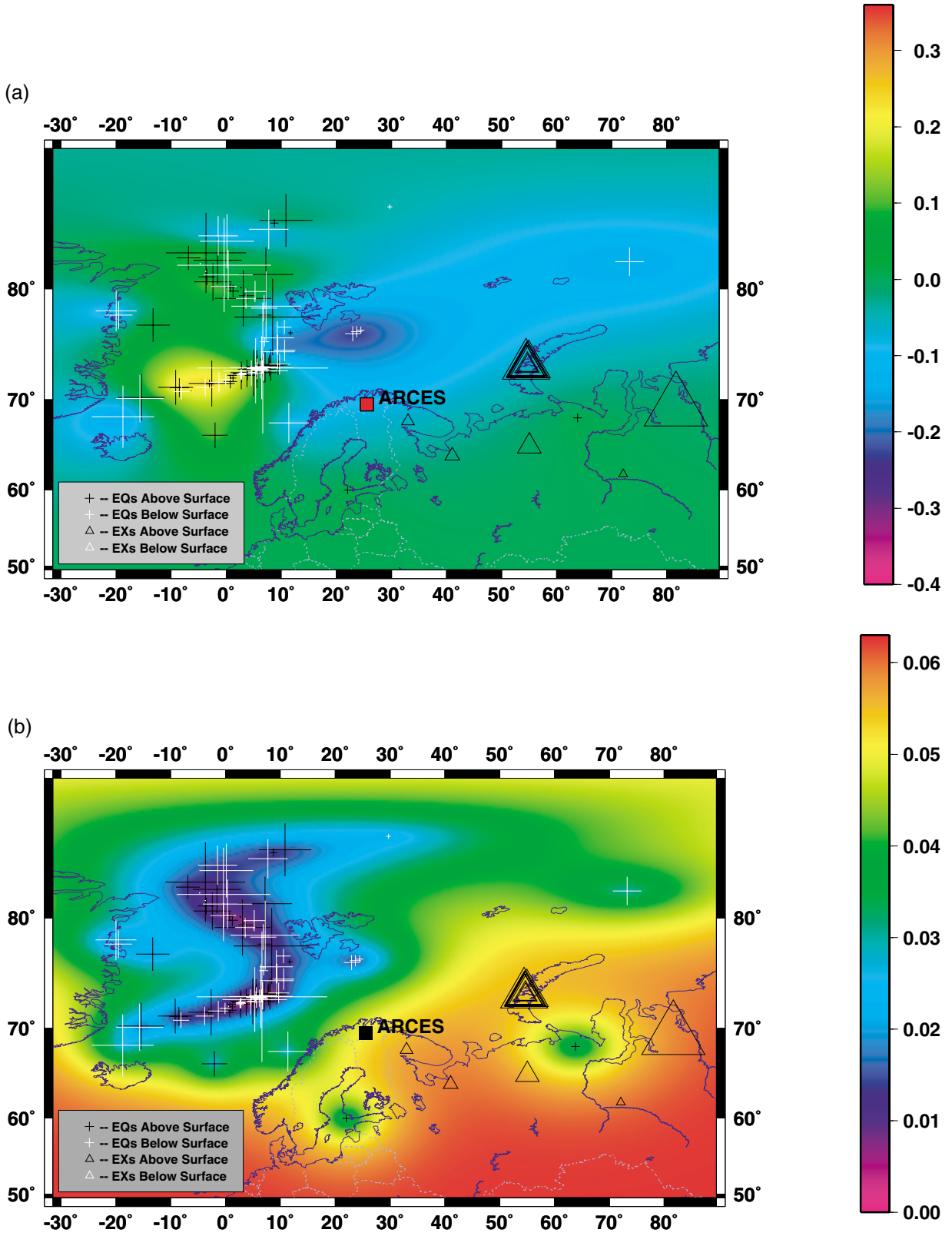

Figure 7. Correction (a) and uncertainty (b) surfaces for log[*Pn*/*Sn*(6–8 Hz)] at ARCES. Triangles show locations of explosions and crosses show the locations of earthquakes. Black (white) markers have values greater (less) than the local value of the correction surface. Marker size is proportional to the absolute distance from the correction surface. Some explosions recorded at KEV are included for comparison.

For each station, correction and uncertainty surfaces are calculated for both log[*Pn*/*Sn*(6–8 Hz)] and log[*Pn*/*Lg*(6–8 Hz)]. To construct a hypothesis test, we define a scaled variable, λ, given by

$$\lambda = \frac{y - \mu_{EX}}{\sqrt{\sigma^2 + \sigma^2_{r,EX}}}, \tag{30}$$

where

$$y = x(\mathbf{s}_0) - \hat{u}_0(\mathbf{s}_0) \tag{31}$$

In equations (30) and (31), $x(\mathbf{s}_0)$ is the distance-corrected value of log[*Pn*/*Smax*(6–8 Hz)] for a test event located at $\mathbf{s}_0$, where *Pn*/*Smax* is *Pn*/*Sn* if *Sn* is larger than *Lg* or *Pn*/*Lg*, otherwise; $\hat{u}_0(\mathbf{s}_0)$ is the value of the correction surface at $\mathbf{s}_0$ for *Pn*/*Smax;* μ_{EX} is the mean value of y for all explosions in the data set; σ^2 is the value of the uncertainty surface at $\mathbf{s}_0$ (including the calibration and earthquake residual variances), and $\sigma^2_{r,EX}$ is the residual variance for explosions.

Under the assumptions described in the preceding paragraphs, λ is approximately normally distributed with zero mean and unit variance. With this in mind, we define the screening criterion to be

$$\lambda < -z_\alpha \tag{32}$$

where z_α is the $(1 - \alpha)$-percentile of the standard normal distribution, which is equal to 2.576 for a 0.005 significance level. We also define a regional score such that an event is screened out at the α significance level if the score is greater than zero:

$$\text{Score}_R = -\frac{\lambda}{z_\alpha} - 1 \tag{33}$$

In more basic terms, an event is screened out if its corrected value of log[*Pn*/*Smax*(6–8 Hz)] is significantly different from the mean of the explosion population, accounting for the calibration uncertainty, and the residual earthquake and explosion variances.

Results

Here we apply the *P*/*S* screening criterion to 140 explosions at known nuclear-test sites and 4173 REB events $>m_b$ 3.5. For each explosion, the correction and uncertainty values are computed at the explosion location for each station with regional recordings, from all earthquake data available at that station. The value of y, given by equation (31), is then computed, from which the explosion mean, μ_{EX}, can be estimated from all explosion data. Using the explosion data, we also estimate the explosion residual variance, $\sigma^2_{r,EX}$, to be $(0.22)^2$. The estimates used for the other parameters are $\sigma_{r,EQ} = 0.25$, $\sigma_c = 0.25$, and $\alpha = 6.0°$. For each earthquake, the correction and uncertainty values are computed at the location of the earthquake, by using the remaining earthquakes recorded by that station (i.e., using the leave-one-out procedure). The value of λ and the score are then computed for each event. For an event recorded by more than one station, the average score is used.

Figure 8 is a scatterplot of log[*Pn*/*Smax*(6–8 Hz)] after applying worldwide distance corrections and local Bayesian calibration versus distance from the station for 140 explosions and 4173 presumed earthquakes. Figure 9 shows histograms of the scores for the same events. Events with scores <0 are not screened out, and those with positive scores are screened out. None of the 140 explosions are screened out at the 0.005 significance level, whereas about 78% of the 4173 REB earthquakes are screened out. Figure 8 and 9 can be compared with Figure 3 and 4, which plot the same events by using only distance corrections. As can be seen from Figures 3 and 4, if the explosion with the smallest amplitude ratio had been used as the screening threshold, only about 50% of the presumed REB earthquakes would be screened out compared with 78% from using the Bayesian calibration method. This indicates the significant improvement resulting from local Bayesian calibration.

Application to Soviet PNE Data

In the preceding section, we applied the regional *P*/*S* criterion to 140 explosions at known nuclear-test sites. In addition to these explosions, we have recently compiled and processed regional recordings of 20 Soviet PNEs conducted between 1973 and 1988, 16 of which were recorded by BRV and five by KEV; one was recorded by both stations (see Table 1). Sultanov *et al.* (1999) provide a summary of these and many other Soviet PNEs. Note that the station code BRV is used here to indicate the older instrumentation (with 11-bit digitizer) at Borovoye, dating back to 1966 (e.g., Richards *et al.*, 1992; Kim and Ekström, 1996), whereas BRVK is used for the broadband three-component instrument installed in July 1994 (Kim, pers. comm., 2000). The PNEs were all recorded by the older instrumentation, and some data-quality issues are still under investigation.

We have also acquired and processed regional BRVK recordings of 61 presumed earthquakes ($m_b > 3.5$) listed in the REB. Forty-one of these events have *P*/*S* measurements in the 6–8-Hz band that satisfy the SNR criteria described previously. These events are used to compute correction and uncertainty surfaces for the region surrounding Borovoye.

Figure 10 shows locations of the 41 earthquakes (circles) and 32 explosions (triangles) recorded at Borovoye. The explosions include 16 PNEs, four Lop Nor UNEs, six UNEs, and six UCEs at the STS. Marker size is proportional to *Pn*/*Sn*(6–8 Hz). Two of the PNEs (Kraton-1 on 17/10/1978 and Angara on 10/12/1980), located about 1000 km north of Borovoye, have *Pn*/*Sn*(6–8 Hz) values lower than typical explosions. Murphy *et al.* (1996, 1997) and Baumgardt (1998) have discussed these events in detail. Note that an earthquake located at 58.7° N, 68.2° E with similar path but closer distance to Borovoye has a *Pn*/*Sn*(6–8 Hz) value so

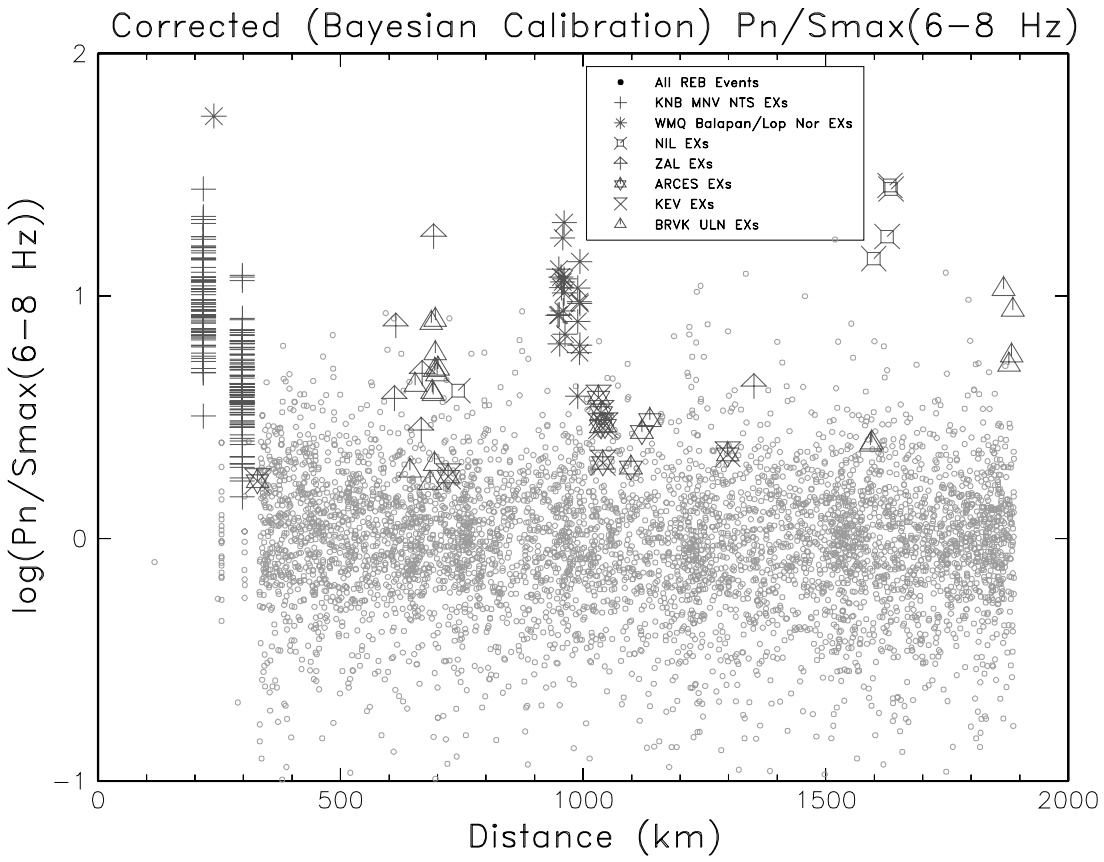

Figure 8. Values of log[*Pn*/*Smax*(6–8 Hz)], after applying worldwide distance corrections and path corrections, versus distance from the station for 140 explosions and 4173 earthquakes.

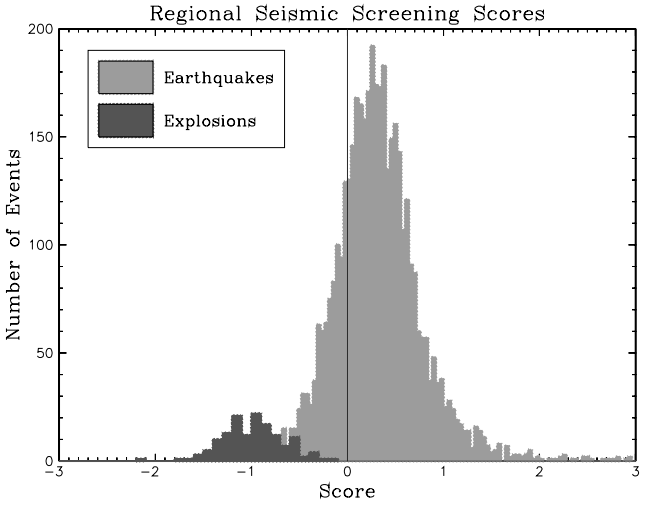

Figure 9. Histograms of regional screening scores for 140 explosions and 4173 presumed earthquakes. None of the explosions are screened out (all have scores less than zero at a 0.005 significance level), whereas 78% of the REB earthquakes are screened out (scores greater than zero).

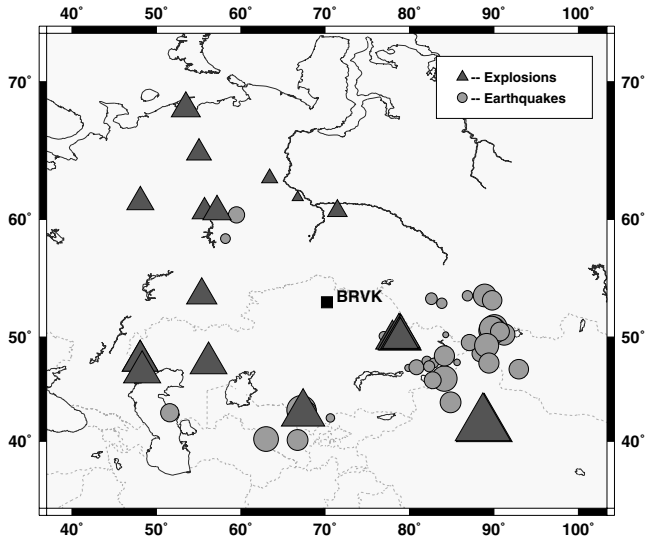

Figure 10. Locations of 32 explosions (16 PNEs, 10 Lop Nor and STS UNEs, 6 STS UCEs—triangles) and 41 earthquakes (circles) recorded at Borovoye. Marker size is proportional to *Pn*/*Sn*(6–8 Hz).

small that the size of the circle representing its value appears as a dot in Figure 10. In the calibration and subsequent screening analyses, these PNEs are less likely to be screened out for two reasons. First, earthquakes near these explosions have even smaller *Pn*/*Sn* values. Second, there are not many earthquakes in this aseismic region, and therefore, the calibration uncertainty will be relatively large.

Figure 11a shows the correction surface and Figure 11b the uncertainty surface for log[*Pn*/*Sn*(6–8 Hz)] at BRVK. The entire correction surface is less than or equal to the worldwide average, which the surface approaches at distances far from data points. At locations near the two PNEs with small *Pn*/*Sn*(6–8 Hz) values (depicted by white triangles), the correction surface is much smaller than the worldwide average, owing to earthquake data in the vicinity with values that are much smaller than average. Examination of the uncertainty surface near these PNEs shows that the uncertainty is relatively large, because of the limited amount of nearby calibration data. Large uncertainties will be present in such aseismic regions; thus, it is less likely that events in such regions will be screened out.

Figure 12 shows plots of log[*Pn*/*Smax*(6–8 Hz)] versus distance for the earthquakes and explosions recorded at Borovoye at various stages of the calibration process. The upper left-hand plot shows the data before any corrections have been made. The best-fit distance-correction curve is also shown. The standard deviation of the earthquake values before correction is $\sigma = 0.29$, and the separation of the explosion and earthquake means is $\Delta\mu = 0.41$. The lower left-hand plot shows the data after applying worldwide distance corrections. The standard deviation has been decreased to $\sigma = 0.25$, and the separation of means has increased to $\Delta\mu = 0.44$. However, if the explosion with the lowest value of *Pn*/*Smax*(6–8 Hz) is used as a screening threshold, very few earthquakes would be screened out at this stage. The upper right-hand plot shows the data after corrections are made using the Bayesian correction surface. The standard deviation has again been decreased to $\sigma = 0.19$, and the separation of means has increased further to $\Delta\mu = 0.49$.

In the lower right-hand plot of Figure 12, the locally corrected data have been scaled by the local uncertainty, as in equation (30). The horizontal line in this plot indicates the screening threshold, derived previously from 140 nuclear and chemical explosions. All but the 10/12/1980 PNE is above the line (i.e., are not screened out). About 70% of the earthquakes are below the line (i.e., are screened out), a rate consistent with the previous results. This case illustrates that local calibration has improved the screening performance dramatically by decreasing the variance of the data and increasing the separation of explosions and earthquakes. Note that although local calibration nearly always reduces the earthquake variance, it does not always further separate the explosion and earthquake means. However, for all stations for which we have explosion and earthquake data, the screening performance, which depends on both the separation of the populations and the variance, is improved.

If the 20 Soviet PNEs are included with the other 140 explosions at known nuclear-test sites, then one of 161 explosions would be screened out, consistent with the 0.005 significance level of the test. The screening criterion can be set more conservatively so that even the 10/12/1980 PNE is not screened out, at the expense of screening out about 15% fewer earthquakes. We plan to revisit this issue and make appropriate changes, after completing our examination of data-quality issues associated with the waveforms recorded by the older BRV instrumentation.

Dependence on Parameter Estimates

Here we investigate the sensitivity of the *P/S* screening test to the three unknown parameters, the correlation length, α, the calibration variance, σ_c^2 and the earthquake residual variance, $\sigma_{r,EQ}^2$. These parameters affect the correction and uncertainty surfaces, which also affect the results of the hypothesis test through the correction in equation (31). The explosion residual variance, $\sigma_{r,EX}^2$, also affects the results of the test more directly through equation (30).

The three parameters, which determine how the calibration surfaces are computed from data, were estimated by fitting the semivariogram given by equation (29). To determine these values accurately and to show that the assumed functional form of the semivariogram adequately describes the data require much more data at each distance lag than are available. Thus, the parameter estimates have large uncertainty. If the final results depend strongly on the precise values of these parameters, then their large uncertainty presents a problem. However, if the results do not depend strongly on the precise values of these parameters, then we gain confidence in the robustness of our approach.

To study this dependence, we estimate the power of the hypothesis test (i.e., the probability that an earthquake is screened out, with a fixed significance level of screening out an explosion) by the earthquake screening rate, from our explosion and earthquake data sets, as the three parameters vary about their estimated values of $\alpha = 6.0°$, $\sigma_c = 0.25$, and $\sigma_{r,EQ} = 0.25$. Because of computational constraints, we will vary each parameter about its estimated value while holding the values of the other two parameters fixed and observe the resulting changes in the screening rate. We will also fix the explosion residual variance, which does not affect the correction and uncertainty surfaces, at its estimated value of $\sigma_{r,EX} = 0.22$.

Figure 13 plots the screening rate (power) as a function of the correlation length, α, between 3° and 10°, while holding the other parameters constant. The screening rate varies between about 74% and 80%—a 6% swing. This dependence on correlation length is not particularly strong, providing confidence that the screening results are not overly sensitive to the precise value of this parameter.

Similarly, Figure 14 plots the screening rate (power) as a function of the calibration variance, σ_c, between 0.2 and 0.3, while holding the other parameters constant. The screening rate varies between about 81% and 75%, which is also

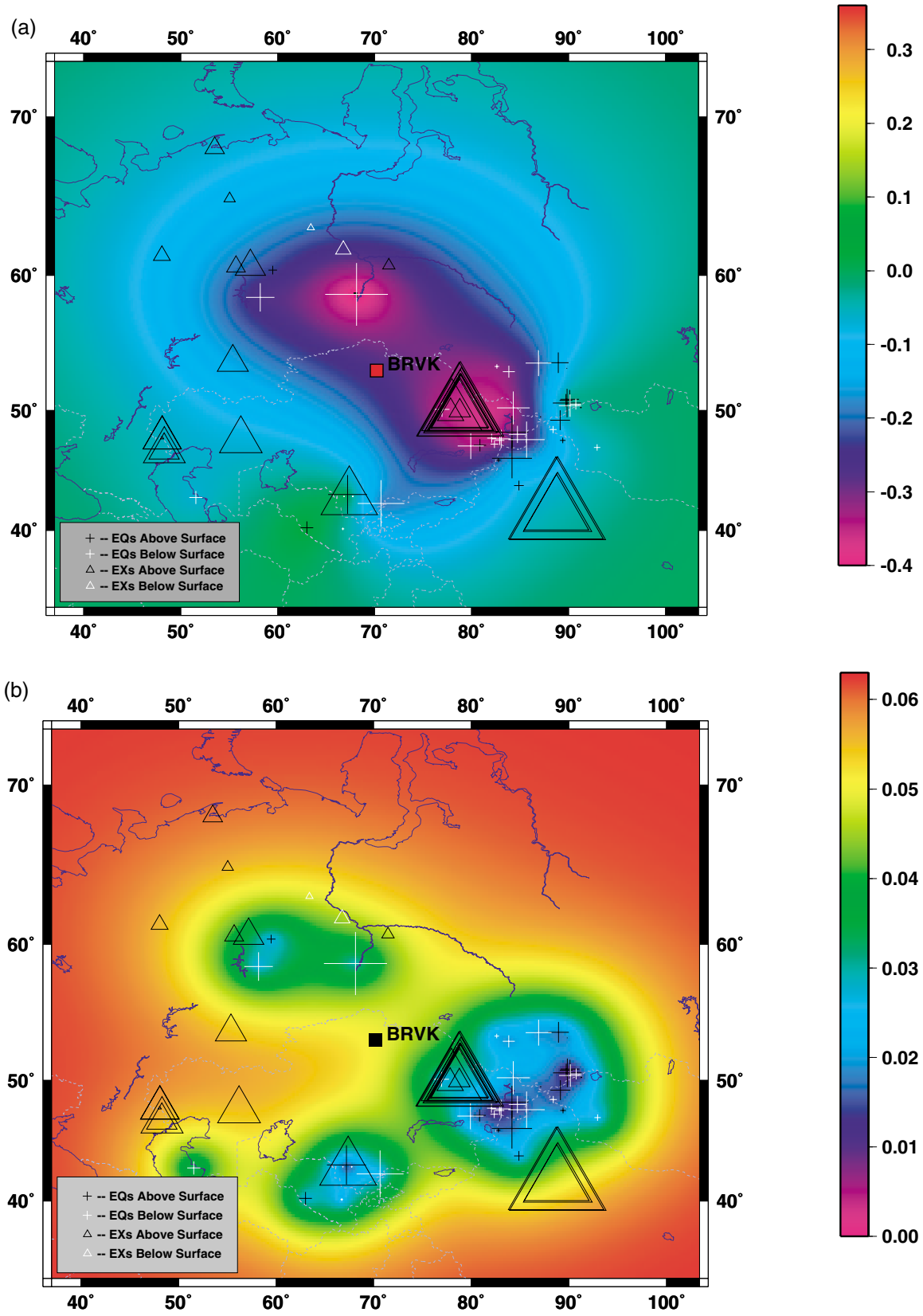

Figure 11. Correction (a) and uncertainty (b) surfaces for log[*Pn*/*Sn*(6–8 Hz)] at BRVK. Triangles represent explosions and crosses represent earthquakes. Black (white) markers have values greater (less) than the local value of the correction surface. Marker size is proportional to the absolute distance from the correction surface.

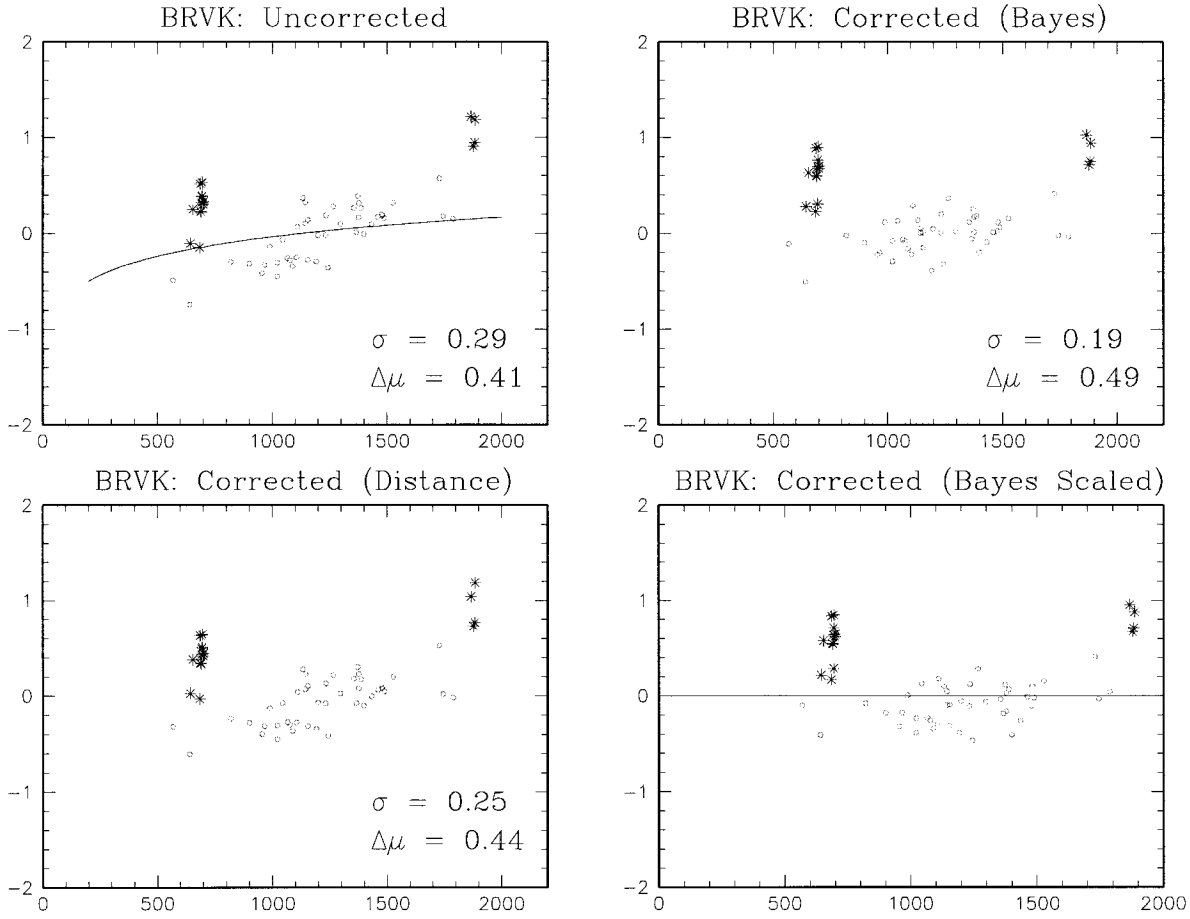

Figure 12. *Pn*/*Smax* (6–8 Hz) versus distance for explosions (asterisks) and earthquakes (circles) recorded by BRVK before corrections (upper left), after distance corrections (lower left), after local Bayesian corrections (upper right), and after scaling by the local uncertainty (lower right).

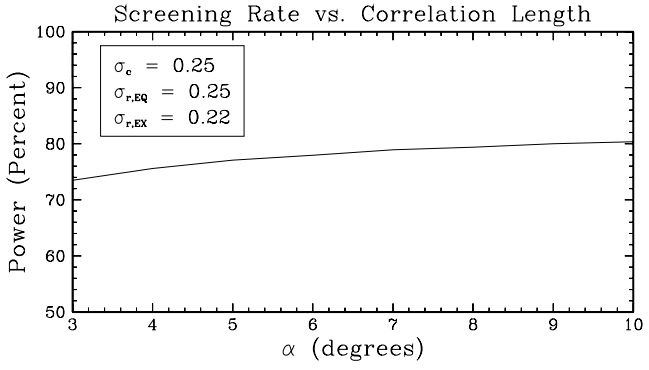

Figure 13. Screening rate (power) of the hypothesis test as a function of the correlation length, keeping the calibration variance, earthquake residual variance, and explosion residual variance fixed.

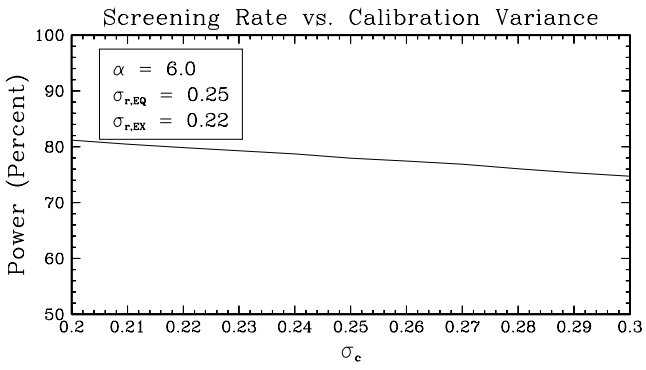

Figure 14. Screening rate (power) of the hypothesis test as a function of the calibration variance, keeping the correlation length, earthquake residual variance, and explosion residual variance fixed.

about a 6% swing. As for the correlation length, this dependence is not particularly strong.

Last, Figure 15 plots the screening rate (power) as a function of the earthquake residual variance, $\sigma_{r,EQ}$, between 0.2 and 0.3, while holding the other parameters constant. The screening rate varies between about 79% and 78%, which is less than the swing due to similar changes in the calibration variance. Thus, the performance of the hypothesis test is least sensitive to the earthquake residual variance.

In general, the earthquake screening rate is not overly sensitive to changes in any of the parameters that are used to characterize the calibration surfaces from data. Recall that the significance level of the test (i.e., the probability of screening out an explosion) has been fixed for all cases. This robustness study provides greater confidence in results that depend on parameters that are not well known.

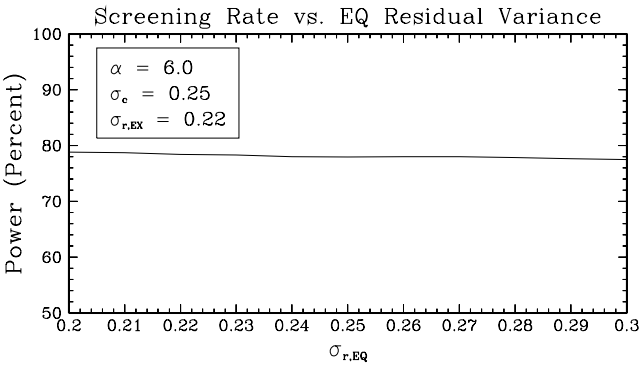

Figure 15. Screening rate (power) of the hypothesis test as a function of the earthquake residual variance, keeping the correlation length, calibration variance, and explosion residual variance fixed.

Conclusions

We presented an approach using regional *P/S* to supplement existing event-screening criteria based on depth and M_s:m_b. We described how the *P/S* (*Pn*/*Sn* and *Pn*/*Lg*) measurements are corrected for distance dependence. We then presented a Bayesian calibration technique, equivalent to simple kriging when normal prior distributions are used, to treat path- and station-dependent variations and to account for local uncertainty in such a way that, in regions far from calibration data, the correction surface approaches the global average and the uncertainty surface approaches maximum uncertainty. We presented an event-screening criterion as a hypothesis test with fixed significance level with respect to screening out an explosion by using corrected values of *Pn*/*Smax* in the 6–8-Hz band. We also defined a score that indicates numerically the degree to which an event is either screened out or not.

We applied this criterion to 140 explosions at known nuclear-test sites and to 4173 REB events $> m_b$ 3.5 (presumed to be mostly earthquakes) with regional recordings between 3° and 17°, *Pn*-SNR > 2.0, and *S*-SNR > 1.3. At a 0.005 significance level, none of the 140 explosions at the NTS, Lop Nor, Indian, Pakistan, Semipalatinsk, and NZ test sites were screened out, whereas about 78% of the REB events were screened out. We showed that correcting regional *P/S* ratios for spatial variations improves the screening performance by about 25%, over just correcting for distance, and that the screening results are not very sensitive to estimates of parameters (correlation length, calibration variance, and residual variance) that are used, along with data, to compute the correction and uncertainty surfaces at each station.

As a separate test, we applied the *P/S* criterion to 20 Soviet PNEs recorded at regional distances by BRV and/or KEV. One PNE (Angara on 10/12/1980) would be screened out by the existing criterion, although some data-quality issues are still under investigation. Overall, one of 268 recordings for 161 explosions would be screened out, consistent with the 0.005 significance level of the test. The screening criterion can be set more conservatively so that even the 10/12/1980 PNE is not screened out, at the expense of screening out about 15% fewer earthquakes. We plan to revisit this issue and make appropriate changes, after completing our examination of data-quality issues associated with the waveforms recorded by the older BRV instrumentation.

There are several potential ways to further improve the regional screening performance. First, although *P/S* measurements of vertical-component seismograms provide useful separation of earthquakes and explosions, various studies indicate that better separation can be obtained by using three-component (3-C) data (Kim *et al.*, 1997; Bottone *et al.*, 2002; Bowers *et al.*, 2001). Preliminary work in this area shows promise. We plan to process and evaluate additional 3-C data for explosions and earthquakes to examine this more thoroughly. Because all of the IMS seismic stations are (or will be) equipped with 3-C short-period or broadband sensors, such data could be utilized.

Second, the hypothesis test presented in this article uses only *Pn*/*Smax*(6–8 Hz). Extending the test to include additional event characteristics, for example, *Pn*/*Smax*(4–6 Hz) and/or *Pn*/*Smax*(8–10 Hz), generally improves screening performance, although not as many measurements are available in the 8–10 Hz band (see Taylor, 1996; Taylor and Hartse, 1998; Fisk *et al.*, 2002). Third, further calibration work is needed. As the IMS seismic network is completed, more events will have useful regional data. However, region-specific calibration will need to be extended to these new stations and improved at existing stations as more regional data are collected. Fourth, future work is needed to develop an approach to combine the results of the depth, M_s:m_b, and/or regional *P/S* criteria. Last, robust capabilities need to be developed to characterize the large number of events $< m_b$ 3.5, including mining blasts.

Acknowledgments

We are grateful to David Jepsen of the Australian Geological Survey Organization for his help in processing regional waveforms used in this study. We also thank Doug Baumgardt, Jack Murphy, Bill Walter, and Howard Patton for providing their data sets and for many useful discussions. Many useful suggestions for improving this article were made by Bill Walter, Artie Rodgers, and Scott Phillips. Special thanks to Sean Moore for providing plotting software. This work was sponsored by the Defense Threat Reduction Agency, under Contract Number DSWA01-98-C-0152.

References

Baumgardt, D. R. (1998). Anomalies in high frequency *P/S* ratios: the 16 October 1997 underwater explosion near Murmansk and PNEs recorded at Borovoye, technical note, ENSCO, Springfield, Virginia.

Baumgardt, D. R., J. Carney, M. Maxson, and S. Carter (1992). Evaluation of regional discriminants using the intelligent seismic event identification system, semi-annual technical report SAS-TR-93-38, ENSCO, Inc., Springfield, Virginia.

Bennett, T. J., B. W. Barker, K. L. McLaughlin, and J. R. Murphy (1989). Regional discrimination of quarry blasts, earthquakes, and underground nuclear explosions, GL-TR-89-0114, Hanscom AFB, Massachusetts.

Blandford, R. R. (1995). Discrimination of mining, cratering, tamped, and decoupled explosions using high frequency *S*-to-*P* ratios, AFTAC-TR-95-002, Air Force Technical Applications Center, Patrick AFB, Florida.

Bottone, S., M. D. Fisk, H. L. Gray, G. D. McCartor, and R. J. Carlson (1997). Initial regionalization efforts for the IMS seismic network, PL-TR-97-2090, Phillips Laboratory, Hanscom AFB, Massachusetts.

Bottone, S., M. D. Fisk, and G. D. McCartor (2002). Regional seismic event characterization using Bayesian calibration, MRC-R-1621, Mission Research Corporation, Santa Barbara, California (in press).

Bowers, D., P. D. Marshall, and A. Douglas (2001). The level of deterrence provided by data from the SPITS seismometer array to possible violations of the Comprehensive Test Ban in the Novaya Zemlya region, *Geophys. J. Int.* **146,** 425–438.

Cressie, N. (1993). *Statistics for Spatial Data*, Revised Edition, Wiley, New York.

Cressie, N., and D. M. Hawkins (1980). Robust estimation of the variogram, I, *J. Int. Assoc. Math. Geol.,* **12,** 115–125.

Fan, G.-W., and T. Lay (1998). Regionalized waveguide effects on seismic discriminants in western China, *Bull. Seism. Soc. Am.* **88,** 1260–1274.

Fan, G.-W., T. Lay, and S. Bottone (2002). Path corrections for source discriminants: a case study at two international seismic monitoring stations, *Pure App. Geophys.* Special Edition on Monitoring a Comprehensive Test Ban Treaty (CTBT), **159,** 651–678.

Fisk, M. D., S. Bottone, R. J. Carlson, H. L. Gray, and G. D. McCartor (1998). Event characterization development and analysis at the prototype IDC, DTRA-TR-99-5, MRC-R-1548, Mission Research Corporation, Santa Barbara, California.

Fisk, M. D., S. Bottone, H. L. Gray, and G. D. McCartor (1999a). Event characterization using regional seismic data, *Proc. 21st Seismic Res. Symp.*, 427–437.

Fisk, M. D., S. Bottone, and G. D. McCartor (2002). A seismic event-screening approach using regional *P/S* amplitude ratios, MRC-R-1603, Mission Research Corporation, Santa Barbara, California (in press).

Fisk, M. D., H. L. Gray, and G. D. McCartor (1996). Regional discrimination without transporting thresholds, *Bull. Seism. Soc. Am.* **86,** 1545–1558.

Fisk, M. D., D. Jepsen, and J. R. Murphy (1999b). Experimental seismic event-screening criteria at the prototype international data center, *Pure Appl. Geophys.*, Special Edition on Monitoring a Comprehensive Test Ban Treaty (CTBT).

Hartse, H. E., R. Flores and P. Johnson (1998). Correcting regional seismic discriminants for path effects in western China, *Bull. Seism. Soc. Am.* **88,** 596–608.

Hartse, H. E., S. R. Taylor, W. S. Phillips, and G. E. Randall (1997). Regional seismic discrimination in central Asia with emphasis on western China, *Bull. Seism. Soc. Am.* **87,** 551–568.

IDC5.2.1 (1999). *IDC Processing of Seismic, Hydroacoustic, and Infrasonic Data*, SAIC-99/3023, Science Applications International Corporation.

Jenkins, R. D., T. J. Sereno, and D. A. Brumbaugh (1998). Regional attenuation at PIDC stations and the transportability of the *S/P* discriminant, AFRL-VS-HA-TR-98-0046, Air Force Research Laboratory, Hanscom AFB, Massachusetts.

Khalturin, V. I., T. G. Rautian, P. G. Richards, and W. Y. Kim (1997). Evaluation of chemical explosions and methods of discrimination for practical seismic monitoring of a CTBT, AFRL-VS-HA-TR-98-0012, Air Force Research Laboratory, Hanscom AFB, Massachusetts.

Kim, W.-Y., V. Aharonian, A. L. Lerner-Lam, and P. G. Richards (1997). Discrimination of earthquakes and explosions in southern Russia using regional high-frequency three-component data from IRIS/JSP Caucasus network, *Bull. Seism. Soc. Am.* **87,** 569–588.

Kim, W.-Y., and G. Ekström (1996). Instrument responses of digital seismographs at Borovoye, Kazakhstan, by inversion of transient calibration pulses, *Bull. Seism. Soc. Am.* **86,** 191–203.

Kim, W.-Y., D. W. Simpson, and P. G. Richards (1993). Discrimination of earthquakes and explosions in eastern United States using regional high-frequency data, *Geophys. Res. Lett.* **20,** 1507–1510.

Lay, T. (1997). Crustal waveguide effects on regional phases in China and southeast Asia, PL-TR-97-2163, Phillips Laboratory, Hanscom AFB, Massachusetts.

Murphy, J. R., D. D. Sultanov, B. W. Barker, I. O. Kitov, and M. E. Marshall (1996). Application of Soviet PNE data to the assessment of the transportability of regional discriminants, PL-TR-96-2290, Phillips Laboratory, Hanscom AFB, Massachusetts.

Murphy, J. R., D. D. Sultanov, B. W. Barker, I. O. Kitov, and M. E. Marshall (1997). Further analysis of regional seismic data recorded from the Soviet PNE program: implications with respect to CTBT monitoring, PL-TR-97-2141, Phillips Laboratory, Hanscom AFB, Massachusetts.

Papoulis, N. (1984). *Probability, Random Variables, and Stochastic Processes*, Second Edition, McGraw-Hill, New York.

Pasyanos, M. E. (2000). Predicting geophysical measurements: testing a combined empirical and model-based approach using surface waves, *Bull. Seism. Soc. Am.* **90,** 790–796.

Phillips, W. S. (1999). Empirical path corrections for regional seismic phases, *Bull. Seism. Soc. Am.* **89,** 384–393.

Phillips, W. S., G. E. Randall, and S. R. Taylor (1998). Path correction using interpolated amplitude residuals: an example from central China, *Geophys. Res. Lett.* **25,** 2729–2732.

Richards, P. G., W. Y. Kim, and G. Ekström (1992). The Borovoye geophysical observatory, *EOS* **73,** 201–206.

Rodgers, A. J., T. Lay, G. Fan, and W. R. Walter (1998). Calibration of distance and path effects on regional *P/S* discriminants at station ABKT (Alibek, Turkmenistan): statistical analysis of crustal waveguide effects, UCRL-JC-129165, Lawrence Livermore National Laboratory, Livermore, California.

Rodgers, A. J., W. R. Walter, C. A. Schultz, S. C. Myers, and T. Lay (1999). A comparison of methodologies for representing path effects on regional *P/S* discriminants, *Bull. Seism. Soc. Am.* **89,** 394–408.

Schultz, C. A., S. C. Myers, J. Hipp, and C. J. Young (1998). Nonstationary Bayesian kriging: a predictive technique to generate spatial corrections for seismic detection, location and identification, *Bull. Seism. Soc. Am.* **88,** 1275–1288.

Sereno, T. J. (1990). Attenuation of regional phases in Fennoscandia and estimates of arrival time and azimuth uncertainty using data recorded by regional arrays, SAIC-90/1472, Science Applications International Corporation, San Diego, California.

Sultanov, D. D., J. R. Murphy, and K. D. Rubinstein (1999). A seismic source summary for Soviet peaceful nuclear explosions, *Bull. Seism. Soc. Am.* **89,** 640–647.

Taylor, S. R. (1996). Analysis of high-frequency *Pg*/*Lg* ratios from NTS explosions and western U.S. earthquakes, *Bull. Seism. Soc. Am.* **86,** 1042–1053.

Taylor, S. R., and H. E. Hartse (1998). A procedure for estimation of source and propagation amplitude corrections for regional seismic discriminants, *J. Geophys. Res.* **103,** 2781–2789.

Taylor, S. R., M. D. Denny, E. S. Vergino, and R. E. Glaser (1989). Regional discrimination between NTS explosions and western U.S. earthquakes, *Bull. Seism. Soc. Am.* **79,** 1142–1176.

Taylor, S. R., and H. E. Hartse (1997). An evaluation of generalized likelihood ratio outlier detection to identification of seismic events in western China, *Bull. Seism. Soc. Am.* **87,** 824–831.

Walter, W. R., K. M. Mayeda, and H. J. Patton (1995). Phase and spectral ratio discrimination between NTS earthquakes and explosions. part I: empirical observations, *Bull. Seism. Soc. Am.* **85,** 1050–1067.

Zhang, T.-R., and T. Lay (1994). Analysis of short-period regional phase path effects associated with topography in Eurasia, *Bull. Seism. Soc. Am.* **84,** 119–132.

Zhang, T.-R., T. Lay, S. Schwartz, and W. Walter (1996). Variation of regional seismic discriminants with surface topographic roughness in the western United States, *Bull. Seism. Soc. Am.* **86,** 714–725.

Zhang, T.-R., S. Schwartz, and T. Lay (1994). Multivariate analysis of waveguide effects on short-period wave propagation in Eurasia and its applications in seismic discrimination, *J. Geophys. Res.* **99,** 21,929–21,945.

Mission Research Corporation
735 State St.
P. O. Drawer 719
Santa Barbara, California 93102
(S.B.)

Mission Research Corporation
8560 Cinderbed Rd., Suite 700
Newington, Virginia 22122
(M.D.F.)

Southern Methodist University
Department of Physics
P. O. Box 0175
Dallas, Texas 75275
(G.D.M.)

Manuscript received 21 March 2001.