REGIONAL WAVE PROPAGATION IN SIBERIA

Keith D. Koper and Terry C. Wallace, Department of Geosciences, University of Arizona

Sponsored by U.S. Department of Defense
Defense Threat Reduction Agency
Grant DSWA01-97-C-0123

ABSTRACT

Monitoring of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) requires the ability to discriminate between low-yield nuclear explosions and small-magnitude earthquakes. Standard discriminants are generally based on the properties of regional seismic phases such as L_g and P_n; however, the character of these phases depends strongly on the geology and tectonic setting along the ray path. Thus, for a discrimination scheme to be successful, the propagation characteristics of regional phases in the area of interest must be well known. Although previous work on the propagation characteristics of regional phases has concentrated mainly on former test site areas, here we present results for Siberia, a region of potential importance that has been under studied.

We use data recorded at 12 broad band stations located either in or adjacent to Siberia. Four of these stations were installed in 1995, and an additional six have been deployed only since 1992-1993. This fact, coupled with the low natural seismicity of the cratonic Siberian platform, leaves us with a sparse data set. Nevertheless, we have assembled waveforms from 427 events, yielding 618 viable ray paths, for analysis. Although the majority of paths cover the border regions of southern Russia, northeastern China, northern Kazakstan, and Mongolia (latitude 50°N) a significant number sample the eastern half of Siberia (longitude 120°E), and a few sample the Siberian platform itself. The events range in size from 3.1 m_b to 7.5 M_s.

Initially, we calculate the magnitude of the L_g phase, $m_b(L_g)$, for each ray path assuming a constant Q_o value of 800. The $m_b(L_g)$ values are compared to values of teleseismic m_b reported by the PDE. A linear regression of the regional and teleseismic magnitudes yields a correlation coefficient of 0.96 and a slope of 0.85. Using different Q_o values in the calculations leaves these values largely unchanged, with only the intercept having significant variation. The difference between the two magnitudes for a given ray path provides a simple estimate of the relative efficiency of L_g propagation. Coherent azimuthal trends in the magnitude differences are observed to correlate well with regional tectonic features. Our results also compare favorably with previous large-scale tomographic models of Q_o for the region, while offering higher resolution and greater spatial sampling in Siberia.

Key Words: CTBT, L_g, magnitude, Siberia

OBJECTIVE

The success of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) depends on the ability of seismologists to monitor the global occurrence of seismic events; of particular importance is the ability to discriminate low-yield nuclear explosions from small-magnitude earthquakes, meteorite impacts, mining blasts, and other man-made chemical explosions. Such discrimination will rely heavily upon the character of the regional distance (source-receiver distances less than 1500-2000 km) wavefield and seismic phases such as P_n, P_g, and L_g. However, these phases travel in the most heterogeneous portions of the Earth, the crust and uppermost mantle, thus their properties have significant regional variation.

Our objective is to characterize the propagation of such regional seismic phases in Siberia [*Wallace and Tinker*, 1998]. By constraining the Earth structure signature in the waveforms we will be better able to infer the source properties of anomalous events in the region. Only in the past decade have enough "open", broadband instruments been deployed in this region to permit such an undertaking. Our specific objectives are (1) to construct a high-quality data base of regional waveforms from the area, (2) to define the spatial and temporal trends of seismicity in the region, and (3) to describe the variation in amplitude and frequency content of the regional phases. The data base will be freely available to researchers in the discrimination community.

RESEARCH ACCOMPLISHED

Construction of Database

We have identified 427 earthquakes that produced well-recorded, regional distance waveforms at one or more of the 12 broadband stations located in or near Siberia (Figure 1). The raw data are freely available and the majority were obtained via the Incorporated Research Institutions of Seismology (IRIS) data management system, the exception being 7 waveforms from station PDY that were obtained via the autodrm at the Prototype International Data Center (PIDC). The earthquakes occurred between December of 1989 and March of 1999, and range in magnitude from 3.1 m_b to 7.5 M_s. The data are currently stored as 600 s-long waveforms in Seismic Analysis Code (SAC) format. We have installed header information, such as event coordinates and origin time, for each waveform and we have constructed a catalog that contains (1) the event hypocenters, (2) the corresponding teleseismic body wave magnitude, $m_b(P)$, (3) the distance, azimuth, and backazimuth of each station which recorded each event, (4) an average Q_o for each ray path based on the model of *Mitchell et al.* [1997], and (5) the corresponding $m_b(L_g)$ values for each ray path. We discuss the calculation and interpretation of the $m_b(L_g)$ values in a later section.

Seismicity Trends

Figure 2 illustrates the time-rate accumulation of "useful" seismic events in Siberia. The number of events is defined as those events that produced at least one high-quality, regional distance ray path to a broadband seismometer in the region. Thus we implicitly account for factors such as station geometry and downtime as well as the actual rate of seismicity in the area. A large number (157) of the cataloged events occurred along the eastern edge of the Kamchatka peninsula where the Pacific plate is subducting beneath the Sea of Okhotsk microplate; ray paths from these events to station PET sample only a small slice of the peninsula (Figure 1) and so they are not included in Figure 2. Since the addition of four stations in 1995 the earthquake total has grown at a linear rate of approximately 50 events per year. From that time period there are ~ 1.5 ray paths per event, so equivalently the data base has been growing by about 75 ray paths per year.

Most of the 461 non-PET ray paths are concentrated in southern Russia and northern Mongolia, however a significant number sample Siberia and eastern Russia. The least sampled tectonic province is the Siberian platform. This is expected since the old Cratonic lithosphere in this region has extremely low seismicity and no stations are located in the deep interior of the platform. We have increased coverage of this region by including paths with lengths of up to 2000 km from events occurring near Lake Baikal and recorded at station NRIL. Although these paths do not produce usable P_n observations, the L_g phases are generally observable.

Unfortunately the coverage of ray paths in the Siberia data base is not expected to increase dramatically in the near future. We anticipate the coverage to continue to be dominated by earthquakes occurring in the Kurile-Kamchatka subduction zone recorded at PET and BILL, and earthquakes occurring in the Inner Asian belt (southern Russia and northern Mongolia) recorded at KURK, TLY, ULN, and HIA. The region most likely to have a significant increase in coverage in the near future is eastern Russian north of ~ 60 N. This region should become much better sampled as events occurring in the Chersky Seismic Belt (CSB, located at the intersection of the North American, European, and Okhotsk plates, see *Mackey et al.*, [1998]) are recorded at stations TIXI, YAK, MA2, and BILL. The northernmost tip of Siberia (latitude > 70°N, 90°E < longitude < 120°E) will also become significantly better sampled as events from the Gakkel Ridge in the Arctic are recorded at NRIL.

The region most likely to remain undersampled given the current station geometry is the western 2/3 of the Siberian platform. The best way to increase coverage of this region would be the installation of a broadband station approximately 800-1000 km due south of NRIL. Such a station would record regional distance events from the Inner Asian belt to the south, the Lake Baikal rift zone to the southeast, the CSB to the east, and possibly the Gakkel rift system to the northeast. As an alternative to the installation of a new instrument in Siberia, increasing the availability of data from station PDY could dramatically improve the coverage of this region. Although this station has been deployed since August of 1993, and is positioned to record regional distance waveforms from the Inner Asian belt, the CSB, and the landward extension of the Gakkel rift system, we were able to obtain only 7 records from PDY.

Calculation of L_g Magnitudes

For each ray path in the data base we calculate a magnitude, $m_b(L_g)$, based on the amplitude of the L_g phase using the method of *Nuttli* [1973] as described by *Priestly and Patton* [1997]. We first deconvolve the instrument response from the observed waveform and convolve in the response for a short period WWSSN instrument. This has the effect of a narrow bandpass filter near 1 Hz. Next we window the L_g phase by combining the source-receiver distance with a velocity bracket of 3.3-3.6 km/s. We take the average of the three largest maxima in this time window, from each of the three components, as the observed amplitude (A) of the L_g phase. Since the frequency content of the data is sharply peaked at ~ 1Hz, estimates of the frequency dependence of Q are unneeded and only Q_o (Q at 1 Hz) is required for the calculation of $m_b(L_g)$.

Given values for A and Q_o it is straightforward to apply the formulas in *Priestly and Patton* [1997] to generate $m_b(L_g)$ estimates; however, it is also important to quantify how errors in A and Q_o are mapped into $m_b(L_g)$ values. We expand the appropriate formula to first order, rewriting $m_b(L_g)$ as m_{bLg} for notational clarity,

$$\Delta m_{bLg}(A, Q_o) = \frac{\partial m_{bLg}}{\partial A}\Delta A + \frac{\partial m_{bLG}}{\partial Q_o}\Delta Q_o$$

and find,

$$?m_{bLg} = 0.434*\left(\frac{?A}{A}\right) - 0.434*\left(\frac{1}{Q_o}\right)*\left(\frac{(d-10.0)}{UT}\right)*\left(\frac{?Q_o}{Q_o}\right),$$

where U is the group velocity of the L_g phase, T is the period, and d the source-receiver distance. The expression is dominated by the second term since even an error as large as 25% in the measurement of the amplitude results in an error of only ~ 0.11 in $m_b(L_g)$.

We illustrate the effect of the second term, inaccurate Q_o values, on $m_b(L_g)$ in Figure 3 using U = 3.4 km/s, T = 1 s, and a Q_o of 600. Because Q_o appears in the exponent for the $m_b(L_g)$ expression it is necessary to specify a theoretical value of Q_o, in addition to the percent error of Q_o, when considering $m_b(L_g)$ error estimates. Thus for low-Q_o regions a given percent error in Q_o will have much larger impact than in high-Q_o regions. As an example, if the true Q_o value along a ray path is 600 but a value of 400 is inadvertently used, then at a distance of 1000 km, $m_b(L_g)$ will be overestimated by ~ 0.25 units.

One method that has been used to account for this well-known dependence of $m_b(L_g)$ on Q is that of *Nuttli* [1988]. In this approach the L_g-coda is analyzed to determine Q as a function of frequency along a given ray path, and the Q_o from this relationship is used in the $m_b(L_g)$ computation. A weakness of this method, however, is that the calculated Q_o value may not be a robust estimate of anelastic Earth structure because of source effect biases. Here we adopt two different approaches to the problem of Q_o estimation.

In the first approach, we trace each Siberia ray path through the 3° by 3° L_g coda-Q_o model of *Mitchell et al.* [1997], presented in Figure 4. We take a weighted average of the Q_o values from the cells that are sampled and assume Q_o to be 600 for path segments outside of the model bounds. *Mitchell et al.'s* [1997] Q_o model is the result of a large tomographic inversion and so should provide a robust estimate of Q_o along a given ray path. Second, we assume Q_o is constant over the entire region and use a value of 800 in our $m_b(L_g)$ calculations. Although there will clearly be significant Q_o variations over an area as large as Siberia, by comparing the corresponding $m_b(L_g)$ values to the reported body wave magnitudes ($m_b(P)$ values, which are presumably much less biased by shallow Earth structure) we hope to map out anelastic variability on station by station basis. This will provide empirically determined, azimuthally dependent, relations for $m_b(L_g)$ calculations at a given station.

The resulting $m_b(L_g)$ values for the data set as a whole are plotted against the corresponding body wave magnitudes, $m_b(P)$, in Figure 5. For both Q_o models the $m_b(L_g)$ values are well correlated with the $m_b(P)$ values, having correlation coefficients of 0.97 and 0.96. Linear regression of the binned data indicates that the slopes of the best-fitting lines for the $m_b(L_g)$-$m_b(P)$ relationships are similar (0.94 and 0.85); however, the intercept values are significantly different. Using the regional Q_o model of *Mitchell et al.* [1997] we find the intercept to be 0.18 and for the constant Q_o model we find the intercept to be 0.83.

The fact that the $m_b(L_g)$ values determined with the constant Q_o model are systematically higher than those determined with the *Mitchell et al*. [1997] model is expected since it is simply a reflection of the difference in mean Q_o between the two models. It is more surprising that (1) the two Q_o models provide similar slope values with respect to $m_b(P)$ and (2) the correlation between $m_b(L_g)$ and $m_b(P)$ is not increased by using a laterally varying Q_o model. It may be that in the Siberia region *Mitchell et al.'s* [1997] model lacks the resolution to improve the magnitude correlation compared to a constant Q_o model; however, it may also be that improvement is seen when the magnitudes are compared on a station by station basis as opposed to the data set as a whole.

We investigate this by performing linear regressions between $m_b(L_g)$ and $m_b(P)$, for both Q_o models, on a station-by-station basis. The results of this are shown in Table 1. Just as in the previous case the intercepts from the different Q_o models are substantially different but the correlation coefficients and slope are quite similar. This confirms the idea that model of *Mitchell et al.* [1997], developed to infer continental scale structures, is not sensitive to the fine scale anelastic variations in the Siberia region.

Several of the stations in Table 1 show poor correlation between $m_b(L_g)$ and $m_b(P)$. One possibility for the scatter is that the $m_b(P)$ values are poorly determined in certain cases, either because of anomalous Earth structure at depth, or a sparse number of teleseismic observations. A more important factor, however, is that the $m_b(L_g)$ values are being influenced by laterally varying Earth structure. We present an example of this in Figure 6 for TIXI and NRIL, two stations which show very poor magnitude correlations (Table 1). In the case of TIXI there is an azimuthal trend to the magnitude residuals resulting from ray paths which straddle the eastern edge the Siberian platform. As the earthquake location moves to the west (increasing in azimuth) the ray path to TIXI samples increasingly more of the high-Q_o Siberia platform, generating positive magnitude residuals. In the case of NRIL, all the magnitude residuals from events occurring at the Gakkel ridge (azimuth < 40°) are strongly negative. This is indicative of the low-Q_o nature of the oceanic portion of the ray path for these events. Identifying and accounting for factors such as these greatly increases the correlation between $m_b(L_g)$ and the teleseismic body wave magnitudes.

CONCLUSIONS AND RECOMMENDATIONS

We have assembled a high quality data set of regional distance waveforms which sample the areas in and around Siberia. All the waveforms show clear P_n and L_g arrivals, while the existence of P_g is more sporadic.

The source-receiver paths of the waveforms are concentrated in southern Russia, northern Mongolia, and the eastern portion of the Kamchatka peninsula. Significant coverage of eastern Russia and Siberia is also obtained, however the western 2/3 of the Siberian platform is extremely undersampled. Coverage of this region is not likely to improve in the near future with the current station geometry. The installation of a single broadband seismometer deep in the interior of the Siberia platform, 800-1000 km south of NRIL, would be the ideal way to enhance coverage of this region, since it would be positioned to record data from the Inner Asia seismic belt, including the highly active Lake Baikal region to the south and southeast, the Chersky seismic belt to the east, and the Gakkel Ridge to the north and northeast. An alternative idea, that would in theory be simpler and cheaper, is to increase the availability of data recorded at PDY to the general academic community.

We have calculated $m_b(L_g)$ for all of the paths in the Siberia data base using two distinct Q_o models. The resulting magnitude values are highly correlated with the body wave magnitude estimates as a whole, and, to a lesser degree, on a station by station basis. The effect of the assumed Q_o value is subtle with regard to the slope of the $m_b(L_g)$ vs. $m_b(P)$ relation, however it is more pronounced as a static offset (change in intercept) of the magnitude values. In some cases the scatter in the magnitude residuals can be directly mapped to tectonic features in the region. By analyzing the regional variations in magnitude residuals, we can empirically determine $m_b(L_g)$ station corrections that depend on the azimuth of the ray path as well as a static offset. These relations will provide a higher-resolution view of the anelastic structure of the region than is currently provided by large-scale tomographic models.

REFERENCES

Mackey, K. G., K. Fujita, and L. J. Ruff, Crustal thickness of northeast Russia, *Tectonophysics, 284*, 283-297, 1998.

Mitchell, B. J., Y. Pan, J. Xie, and L. Cong, L_g coda Q variation across Eurasia and its relation to crustal evolution, *J. Geophys. Res., 102*, 22767-22779, 1997.

Nuttli, O. W., Seismic wave attenuation and magnitude relations for Eastern North America, *J. Geophys. Res., 78*, 876-855, 1973.

Nuttli, O. W., L_g magnitudes and yield estimates for underground Novayan Zemlya nuclear explosions, *Bull. Seismol. Soc. Am., 78*, 873-884, 1988.

Priestly, K. F., and H. J. Patton, Calibration of $m_b(P_n)$, $m_b(L_g)$ scales and transportability of the M_o:m_b discriminant to new tectonic regions, *Bull. Seismol. Soc. Am., 87*, 1083-1099, 1997.

Wallace, T. C. and M. A. Tinker, Seismic characterization of Siberia, in *Proceedings of 20th Annual Symposium on Monitoring a Comprehensive Test Ban Treaty (CTBT)*, Santa Fe, New Mexico, 536-541, 1998.

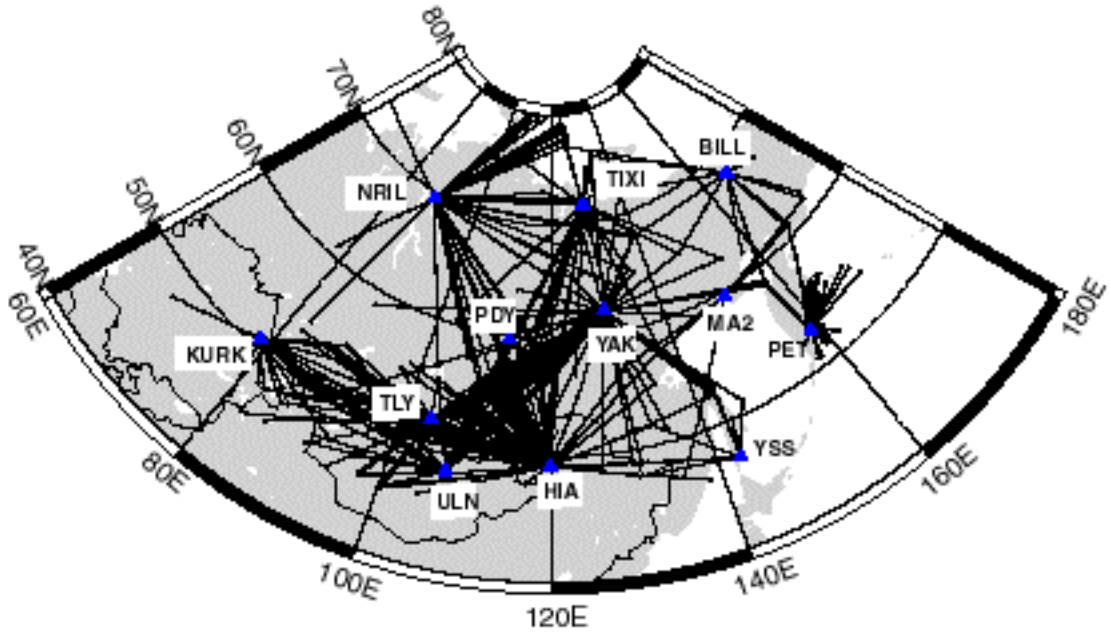

Figure 1. Event-receiver paths of waveforms currently archived in the Siberia data base. We have identified a total of 618 regional paths, from 427 earthquakes, that were well-recorded at one or more of 12 broadband stations which have been installed in the previous decade.

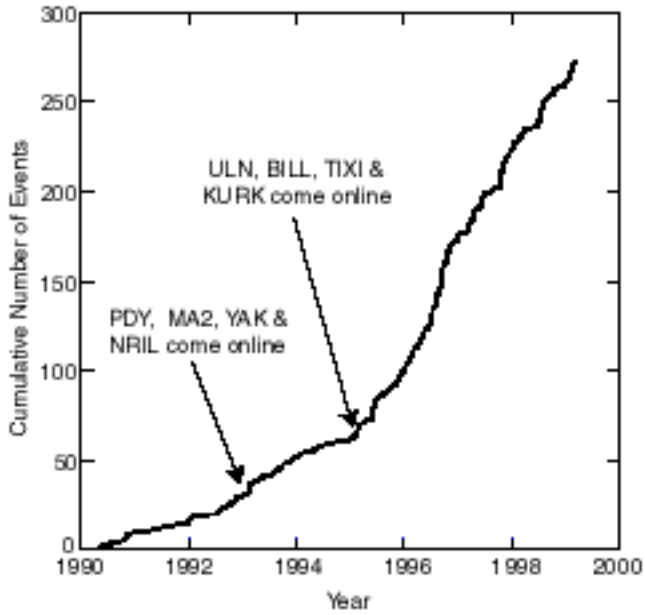

Figure 2. The cumulative number of useful events as a function of time. Events from the Kurile-Kamchatka subduction zone to PET are not included as they sample only a small slice of the Kamchatka peninsula. Since 1995 approximately 50 events per year have occurred which provide useful data.

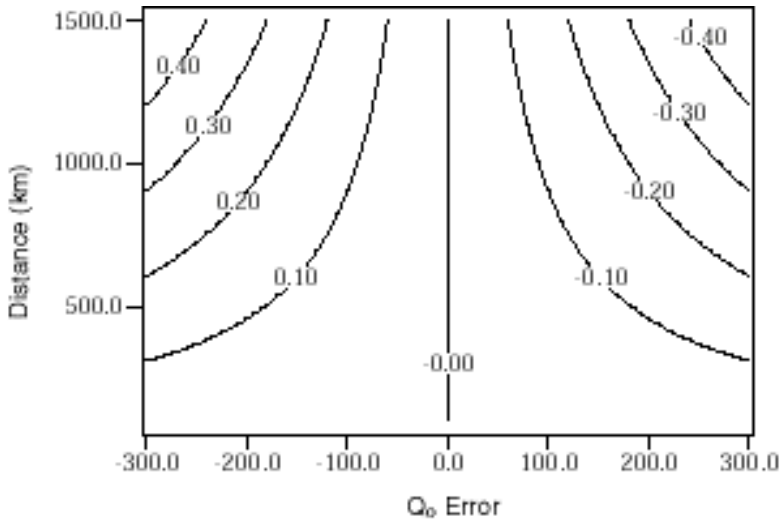

Figure 3. The error in $m_b(L_g)$ as a function of distance and uncertainty in the average, along-path Q_o value. The contours of $m_b(L_g)$ error are in intervals of 0.1. The magnitude errors are negatively correlated with Q_o errors and positively correlated with distance. We assume a theoretical Q_o value of 600 for this calculation.

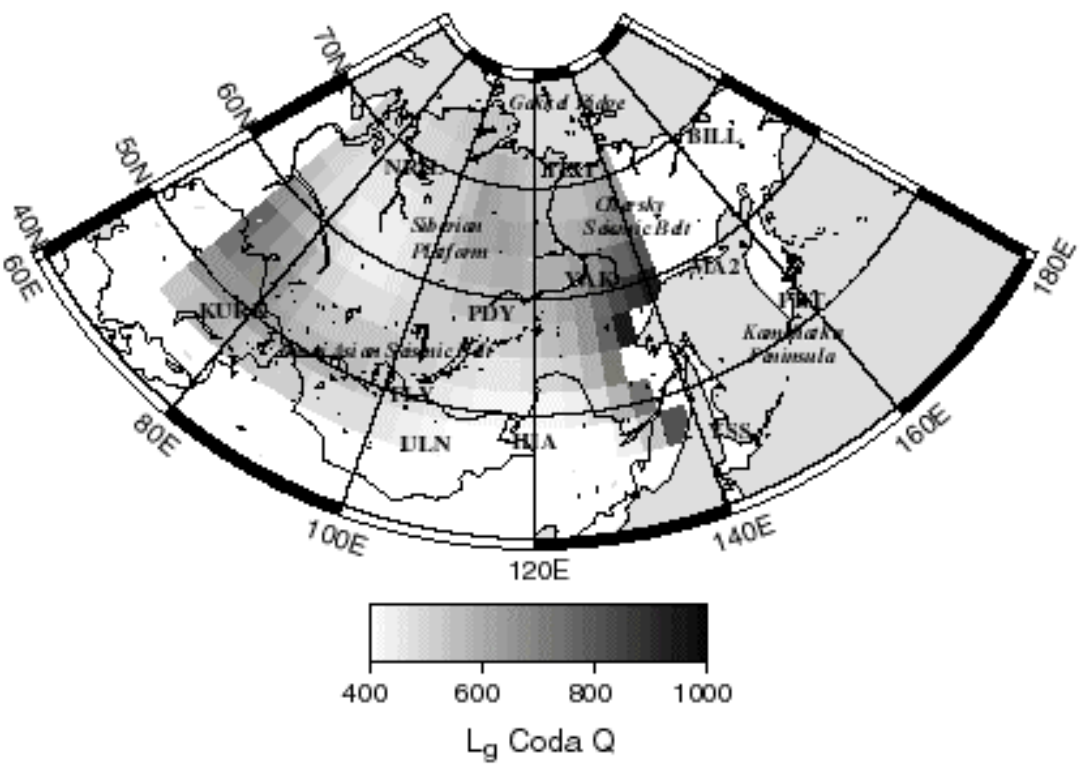

Figure 4. A portion of the coda Q model developed by *Mitchell et al.* [1997]. The model continues to the south and west but is bounded on the east by the 140°E meridian. This model facilitates calculations of L_g magnitudes.

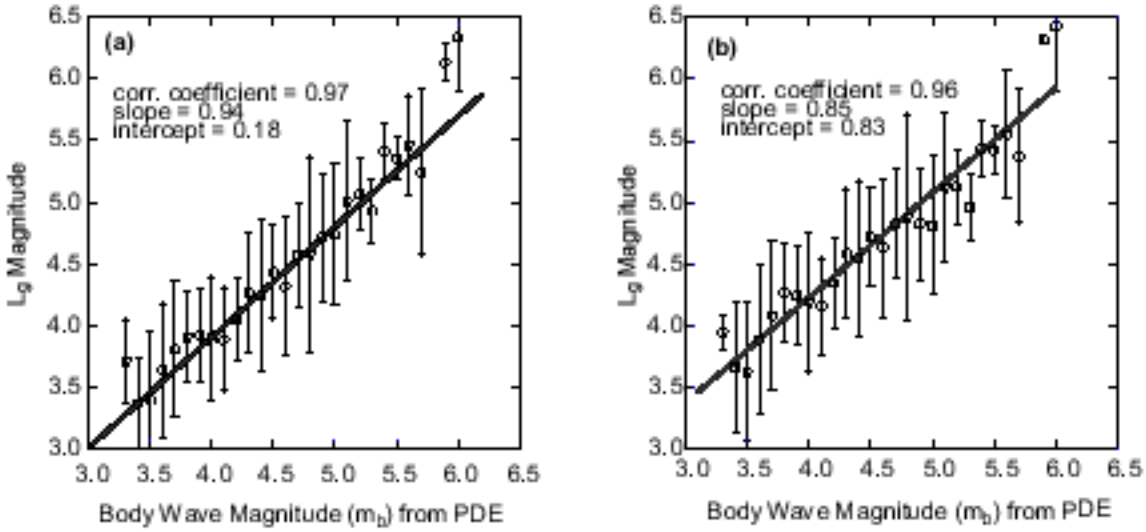

Figure 5. Correlation between magnitudes determined from teleseismic body waves and magnitudes determined from the L_g pulse, for the entire Siberian database, using **(a)** the Q_o model of *Mitchell et al.* [1997], and **(b)** a constant Q_o value of 800. The circles represent the average L_g magnitude for 0.1 bins of m_b; the error bars represent +/- one standard deviation of each bin. In both cases the magnitude values are highly correlated and the best-fitting lines have similar slopes.

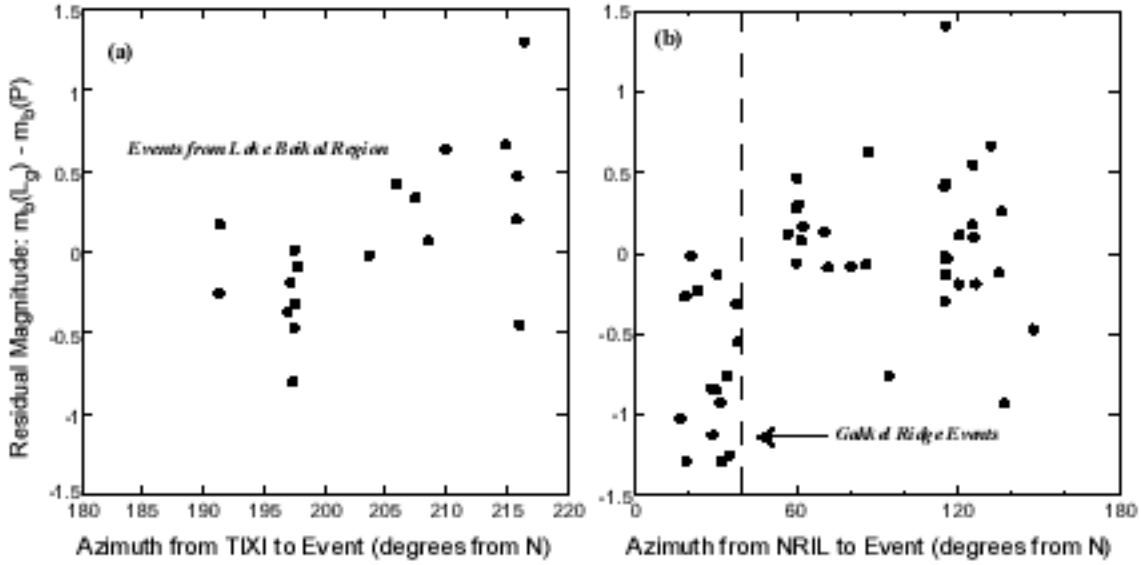

Figure 6. The azimuthal dependence of $m_b(L_g)$ - $m_b(P)$ residuals for **(a)** TIXI and **(b)** NRIL. In both cases we compute $m_b(L_g)$ using the constant Q_o model. In the former case magnitude residuals are shown from a string of events occurring in the Lake Baikal region. The increase of residual with increasing azimuth is due to enhanced sampling of the Siberian platform (a high-Q_o region) as the event location moves to the west. In the latter case, NRIL, events from the Gakkel Ridge (azimuth < 40°) show strongly negative residuals, that are related to the sampling of oceanic lithosphere along the ray path. These effects contribute to the poor $m_b(L_g)$ vs. $m_b(P)$ correlation reported in Table 1.

Station	Number of Observations	Correlation Coefficients	Slopes	Intercepts
BILL	15	0.80,0.78	1.14,1.12	-0.69,-0.44
HIA	91	0.80,0.80	0.98,0.95	0.12,0.69
KURK	25	0.87,0.88	1.13,1.13	-0.21,-0.07
MA2	6	—	—	—
NRIL	49	0.43,0.41	0.42,0.43	2.14,2.55
PDY	7	—	—	—
PET	131	0.43,0.45	1.10,1.06	-0.72,-0.44
TIXI	43	0.47,0.45	0.92,0.97	0.15,0.19
TLY	78	0.69,0.73	0.92,0.94	-0.19,-0.04
ULN	37	0.83,0.84	1.02,1.02	0.00,0.38
YAK	72	0.77,0.68	0.83,0.73	1.02,1.60
YSS	15	0.52,0.53	0.81,0.77	0.13,0.42

Table 1. Linear regressions of magnitudes for each station. The magnitude relationship is of the form $m_b(L_g) = slope \times m_b(body\ wave) + intercept$. We calculate $m_b(L_g)$ using two separate Q_o models: the L_g-coda model of *Mitchell et al.* [1997] and a constant model with Q_o=800. The magnitudes retain the same degree of correlation and slope values independent of which Q_o model is used, though the intercept is much more dependent on the assumed Q_o structure.