Bulletin of the Seismological Society of America, Vol. 96, No. 4A, pp. 1201–1212, August 2006, doi: 10.1785/0120050175

Global Detection and Location of Seismic Sources by Using Surface Waves

by Göran Ekström

Abstract We develop an algorithm for the detection and location of seismic sources using intermediate-period (35–150 sec) surface waves recorded on a global array of stations. Continuous vertical seismic waveforms from the global network are collected and a $4^\circ \times 4^\circ$ global grid of target locations is defined. For each target location and each station, a surface-wave propagation operator is deconvolved from the seismogram to restore any source pulse present. The envelope of the seismogram is calculated and cross correlated with a theoretical source-pulse shape. The resulting waveforms are stacked to improve signal-to-noise characteristics, and the quality, strength, and timing of the potential detection are determined. When a successful event detection is made, a finer grid is applied to locate the event with greater precision. We apply the algorithm systematically for the period 1993–2003 and catalog the events. Approximately 2000 events are detected and located for each year and 98% of shallow events in the Harvard Centroid Moment Tensor (CMT) catalog are detected and located by the new algorithm. A comparison of 9482 events common to the two catalogs allows the detection strength to be calibrated against the CMT moment magnitude. All detected events have estimated moment magnitudes $M_w > 4.5$. In each year, approximately 100 events not listed in other global seismicity catalogs are detected and located. Many of these events lie along the ridge-transform plate boundaries in the Southern Hemisphere and may be regular earthquakes that have gone undetected because of poor station coverage. Other events, located in areas where global and regional networks provide good coverage, are potentially anomalous and may have escaped detection as a result of their unusual source properties.

Introduction

Earthquakes and other seismic events are traditionally detected and located using the arrival times of short-period body-wave phases. In common practice, seismic phases are first detected and identified in seismograms at individual stations. Similar arrival times from several stations are grouped and associated with a postulated event, which is then located by using iterative triangulation methods. In these location algorithms, the detection of arrivals of seismic energy in individual seismograms precedes the determination of an event location. Small-aperture seismic-monitoring arrays are sometimes used in a different mode (e.g., Ringdahl and Husebye, 1982), in which a target location is prescribed, the array is tuned for the specific location, and the phase and event detection follow. The tuned-array technique has several advantages: it allows for the application of signal-processing techniques that enhance the signal-to-noise ratio in the seismograms and it reduces the complexities of association when several events occur at nearly the same time in different locations, for example. A significant drawback of this technique for global monitoring is the prohibitively large number of computations that would be needed to monitor the volume of the Earth where earthquakes may occur.

In this article, we explore the possibility of implementing a routine, automated global event-detection and event-location algorithm based on array analysis of teleseismic intermediate-period surface waves. The use of surface waves reduces the grid to be searched to the surface of the Earth, and the use of longer-period waves makes it sufficient to search a relatively coarse grid in both space and time. The success of the method rests on the observation that the background seismic-noise spectrum has a pronounced low in the period range 30–500 sec, which makes it possible to observe long-period surface waves at teleseismic distances even for earthquakes of moderate size. For shallow sources, surface waves are frequently the largest-amplitude signals in long-period seismograms. Despite these characteristics, surface waves have so far not been routinely exploited for event detection or location. This is due in part to the large variations in phase and group velocity exhibited by surface waves at shorter periods. Also, the detection of surface waves and the determination of an arrival time is more complex than for body waves, owing to the dispersive nature of surface waves. Instead, surface waves have mainly been used to determine the surface-wave magnitude M_S after an event has

been located.

A motivation for developing an event-detection algorithm for surface waves is the possibility that events with unusual source spectra may elude detection by traditional methods, but could be detected using longer-period surface waves. All routine determinations of global earthquake locations, such as those published in catalogs and bulletins by the National Earthquake Information Center (NEIC), the International Seismological Centre (ISC), and the International Data Center (IDC) of the International Monitoring System (IMS) are based on detecting and determining arrival times of seismic phases, primarily short-period *P* waves. It has been observed that some earthquakes radiate lower-amplitude *P* waves than would be predicted from their seismic moment using average global relationships. In some cases, the reduced *P*-wave amplitudes can be explained by the radiation pattern of the earthquake and the distribution of recording stations. For example, for strike-slip earthquakes along the fracture zones in the southern oceans all recording stations are at teleseismic distances, and the *P*-wave amplitudes are small owing to the radiation pattern. In so-called slow earthquakes, the rupture characteristics of the events are such that the short-period portion of the source spectrum is depleted. Events with *P*-wave amplitudes that are small for either reason may not generate body-wave phases that are easily detected, and they may therefore be missed in the compilation of global seismicity catalogs.

The algorithm developed here utilizes the global network of seismic stations in an array-processing mode. A global grid of target locations is monitored continuously for detections of coherent surface-wave energy. The tuning of the array for each target location is accomplished by calculation of path corrections using global phase-velocity maps (Ekström *et al.*, 1997). The quality of the detection and the size of the event are determined by using a matched-filter technique. Several of the components in the current algorithm build on experience gained in earlier studies. Shearer (1994) used a stack of long-period seismograms from known earthquakes recorded on the International Deployment of Accelerometers (IDA) network to develop a matched filter for surface-wave detection. He then applied the matched filter to seismograms from the same network to locate events on a grid of points separated by approximately 10° on the Earth's surface. Ihmlé and Jordan (1996) used a similar grid-search technique to match long-period ($T > 80$ sec) Rayleigh-wave waveforms or their envelopes in a search for slow earthquakes. Both of these earlier studies reported successful detections of some moderate ($M_S \sim 5.0$) earthquakes. In the method developed here, shorter-period surface waves, with larger signal-to-noise ratios, are used, allowing us to detect smaller events and to determine the locations and origin times of detected events with greater precision.

In this article, we describe the theoretical background and technical implementation of the algorithm. We then describe the results from a systematic application of the method to 11 years of data from the Global Seismographic Network (GSN) and various other seismographic networks. A central result of our study is the detection and location of ~1300 seismic events that do not appear in other seismicity catalogs. A subset of these events, characterized by their association with glaciers, has been discussed in an earlier publication (Ekström *et al.*, 2003). The focus here is on the performance and calibration of the detection and location algorithm. To illustrate the quality of the detections, we present the data for two previously undetected earthquakes as examples. A comprehensive presentation of the event catalog resulting from the application of the method is the topic of a follow-up article (G. Ekström and M. Nettles, unpublished manuscript).

Theory

Consider the propagation of surface waves from a source to an array of globally distributed stations. Each surface-wave seismogram $u(\omega)$ can be written

$$u(\omega) = A(\omega) \exp[i\Phi(\omega)], \tag{1}$$

where $A(\omega)$ and $\Phi(\omega)$ are the amplitude and phase, respectively, of the wave as functions of the angular frequency ω. For a given minor-arc source–receiver geometry, the phase Φ is the sum of three terms,

$$\Phi = \Phi_S + \Phi_R + \Phi_P, \tag{2}$$

where Φ_S is the source phase calculated from the source mechanism and geometrical ray takeoff azimuth, Φ_R is the receiver phase, and Φ_P is the propagation phase

$$\Phi_P(\omega) = \int \frac{\omega}{c(\omega)}\, ds, \tag{3}$$

where c is the local phase velocity and the integration follows the ray path. The amplitude A can be expressed as

$$A = A_S A_R A_\Delta A_Q, \tag{4}$$

where A_S is the excitation at the source, A_R is the receiver amplitude, A_Δ is the geometrical spreading factor, and A_Q is the decay factor due to attenuation along the ray path. The operator $O_P(\omega)$ that describes the distortion of the source pulse due to propagation (Φ_P, A_Δ, and A_Q) can be written:

$$O_P(\omega) = (\sin\Delta)^{-1/2} \exp\left[\frac{i\omega\Delta R}{c(\omega)}\right] \exp\left[\frac{-\omega\Delta R}{2Q(\omega)U(\omega)}\right], \tag{5}$$

where R is the radius of the Earth, Δ is the epicentral distance, $Q(\omega)$ is the quality factor, and $U(\omega)$ is the group velocity.

Propagation effects can be removed from the seismograms by deconvolution of the propagation operator, written in the frequency domain as

$$u_P(\omega) = u(\omega) O_P^{-1}(\omega), \tag{6}$$

where $u_P(\omega)$ is the waveform corrected for propagation effects. The resulting waveform will contain the source pulse centered at the origin time. The source pulse corresponds to the convolution of the source time function of the earthquake with the complex radiation pattern. In general, the source pulse will therefore be a filtered, variably phase-shifted version of the source time function, with a shape that depends on the takeoff azimuth. The amplitude spectrum of this pulse will depend on the depth and focal mechanism of the earthquake, as well as the Earth structure in the source region. The normalized amplitude spectrum commonly varies little with azimuth, and the dominant variation of the source pulse with azimuth is in the phase.

Although the source phase cannot be directly removed from the source pulse without additional knowledge of the source parameters, the pulse shapes in different azimuths can be equalized by calculation of the envelope $u_E(t)$ of the propagation-corrected waveform as

$$u_E(t) = [u_P(t)]^2 + \tilde{u}_P(t)]^2]^{1/2}, \tag{7}$$

where $\tilde{u}_P(t)$ is the Hilbert transform of $u_P(t)$ (Dziewonski and Hales, 1972). The envelope of the source pulse will be very similar in all azimuths. This allows for phase-coherent time-domain processing of the waveform envelopes $u_E(t)$ from a set of globally distributed stations.

Method

The implementation of the theoretical approach outlined in the preceding section was guided by the objective of monitoring the entire surface of the Earth continuously with available computer resources. This goal is reflected in the adoption of several practical and *ad hoc* strategies for reducing the number of locations on the Earth that must be processed and for selecting subsets of seismograms for inclusion in the analysis.

We use data from permanent networks equipped with broadband instrumentation, primarily the GSN, Geoscope, Mednet, GEOFON, and the Canadian Seismograph Network. Continuous vertical-component data from the long-period channel (LHZ) sampled at one sample per second are extracted in 9-hr-long segments. The data are resampled at 8 sec per sample using a fast Fourier transform (FFT), and the instrument response is deconvolved. The resampled displacement waveforms are stored in daily files containing six 9-hr-long partially overlapping records with start times offset by 4 hr. That is, the records will have start times of 0, 4, 8, 12, 16, and 20 hr UT for each day. Typically, records from 100–200 stations are available at a given time. The daily files are indexed for rapid access.

The detection and location occurs in two main processing stages. In the first stage, detection strengths are determined by sequential analysis of points on a global grid of potential source locations. Four-hour-long time windows are processed at a time, using the corresponding data record from each of the available stations. In the second stage of the analysis, the space-time array of detection strengths resulting from the first stage is analyzed to determine whether an event has been detected and, if so, the optimal time and location for the event. The quality of the event detection and an event magnitude are also assigned in the second processing stage.

Stage 1: Detection

All available deconvolved seismograms for a 4-hr period are read from a computer disk and band-pass filtered between a 35- and 150-sec period. An important element in the algorithm is the selection of a subset of stations that will contribute to the determination of a detection. The selection is based on a ranking of the probable utility of each seismogram for contributing to a detection at a particular location. The ranking is based on three criteria: observed noise level, probable signal level, and contribution to azimuthal coverage. The noise level can be determined without reference to a particular source location. It is estimated by calculating, for each station i, the root-mean-square (rms) signal amplitude r_i not exceeded 20% of the time in the 9-hr-long record. A lower value indicates a quieter and therefore more useful station for detecting signals.

The processing then enters a loop over test locations distributed on the surface of the Earth. Through trial and error, we determined that the coarsest grid that was sufficiently dense not to miss events due to decorrelation of the signal was $4^\circ \times 4^\circ$, which means that 4050 locations are analyzed. Before the seismograms are corrected for propagation from a given location, two additional trace parameters are calculated and used to establish an overall ranking of the available stations. Surface-wave amplitudes vary geometrically as $(\sin \Delta)^{-1/2}$. At a given period, amplitudes decay monotonically with distance due to intrinsic attenuation. For each location we modify the noise estimate by dividing it by a signal estimate s_i to which we assign the value

$$s_i = \min\left[(\sin\Delta)^{-1/2} \frac{\pi - \Delta}{\pi}, 3.333\right]. \tag{8}$$

The maximum value of 3.333 leads to constant values of s_i for distances $\Delta < 5^\circ$. The noise-to-signal ratio r_i' is then

$$r_i' = r_i / s_i, \tag{9}$$

In addition to factors related to the noise level at the station and the distance from the potential source location,

we include a factor based on station coverage in our selection of stations. It is clear that to constrain well the location of a detected event, it is necessary to include stations at different azimuths. We define a coverage function c^N to represent the quality of the azimuthal distribution of N stations for the postulated event location. Consider the distribution of N stations around a potential source location. The azimuths to the stations from the source location are α_i, ordered such that $\alpha_1 < \alpha_2 < \ldots < \alpha_N$. We can then define angular gaps between stations, $g_i = \alpha_{i+1} - \alpha_i$, $i = 1, 2, \ldots, N-1$, and $g_N = 2\pi + \alpha_1 - \alpha_N$. We now define

$$\frac{1}{c^N} = \sum_{i=1}^{N} g_i^2/(2\pi)^2. \tag{10}$$

It can easily be verified that if all stations lie in one direction from the source location, $c^N = 1$. Instead, if N stations are evenly distributed in azimuth, $c^N = N$. In fact, c^N can be thought of as an effective number of stations.

We use the quantities r_i' and c^N to determine the importance of each station and to select the stations to be included in the detection calculation. The first station to be included is the station that has the lowest r_i' To select the second station to be included, we calculate $c_i^{N=2}$ for all stations i combined in a pair with the first-selected station. The $N-1$ coverage functions $c_i^{N=2}$ are then ranked. The rankings based on noise and signal (r_i') and coverage ($c_i^{N=2}$) are then summed. The station with the highest combined ranking is then included. The process is continued until the desired number of stations have been identified. After many experiments, we chose $N = 30$ for subsequent processing.

The records from the N selected stations are then processed in the following manner. Dispersion curves between each location and station have been precalculated based on the phase-velocity maps of Ekström *et al.* (1997) and stored on disk. Propagation effects are deconvolved from each record following equation (6). The envelope function is calculated as in equation (7). If a source pulse is present, it will appear as a localized peak in the record. To determine whether a peak is present in the data, we calculate the correlation $u_i^C(t)$ of the time series $u_i^E(t)$ with a template peak shape $s(t)$ corresponding to the envelope of the impulse response of the bandpass filter $u_i^C(t) = u_i^E(t) * s(t)$. The correlation is calculated over a 400-sec-wide time window. For each time point in the record we then have a correlation value between -1 and 1. The larger the correlation value for a given time, the more likely it is that signal from an event has been detected.

Figure 1 shows the original seismograms $u_i(t)$ and the corresponding envelope waveforms $u_i^E(t)$ for a successful detection. When the trial location coincides with the true source location, the peaks in the $u_i^E(t)$ seismograms will occur at the same time. This is the criterion for detection. We quantify this coincidence in time by summing the correlation functions $u_i^C(t)$ to form a stacked correlation function $u^S(t)$,

$$u^S(t) = \frac{1}{N} \sum_{i=1}^{N} u_i^C(t). \tag{11}$$

The criterion for declaring a detection is that $u^S(t)$ for a given test location and a given time exceeds a threshold value. After extensive testing of the algorithm, we chose $u^S(t) > 0.7$ as the detection criterion. Additional criteria are applied in the second stage of processing to determine whether an event has been detected. All times and grid locations for which the detection criterion is satisfied are written to a file, together with additional information that describes the quality of the detection. This information includes the number of traces N_D for which a single-trace detection criterion was satisfied (we used $u_i^C(t) > 0.6$), the azimuthal-coverage value c^N, and an estimate of the magnitude of the detection M_{SW}. The magnitude is determined from the median value of $u_i^E(t)$, calculated based on the N_D traces for which the single-trace detection criterion was satisfied. This value, which reflects the amplitude of the source pulse, is converted to a magnitude by

$$M_{SW} = c + \frac{2}{3} \log u^E_{\mathrm{median}}(t), \tag{12}$$

where c is a constant chosen such that M_{SW} agrees with the moment magnitude M_{w} for a set of calibration sources (see the Results section).

If a detection is found for a given grid location, a refined grid surrounding the initial location is analyzed. Figure 2 illustrates the process. We use four levels of refinement, with the smallest having a grid spacing of $0.5^\circ \times 0.5^\circ$. A search on the finer grid is also performed if the average of the peak values of $u_i^C(t)$ calculated over a time window T_W exceeds a critical value. The time window is determined by consideration of the travel time of a surface wave across a cell in a grid with spacing D as

$$T_W = \frac{D\sqrt{2}}{3.5}, \tag{13}$$

where D is measured in kilometers and 3.5 km sec^{-1} is a typical surface-wave phase velocity.

The computation time for processing one 4-hr-long time window is approximately 15 min on a Sun UltraSparc processor. Processing time is roughly proportional to the number of test locations, stations, and the sampling frequency.

Stage 2: Location

The end result of the detection stage is, for each 4-hr-long time window, a list of the locations and times of a set of potential seismic sources. If a seismic event is present in the data, it will typically generate more than one hundred detections. This occurs because several adjacent grid locations can generate correlations that meet the detection cri-

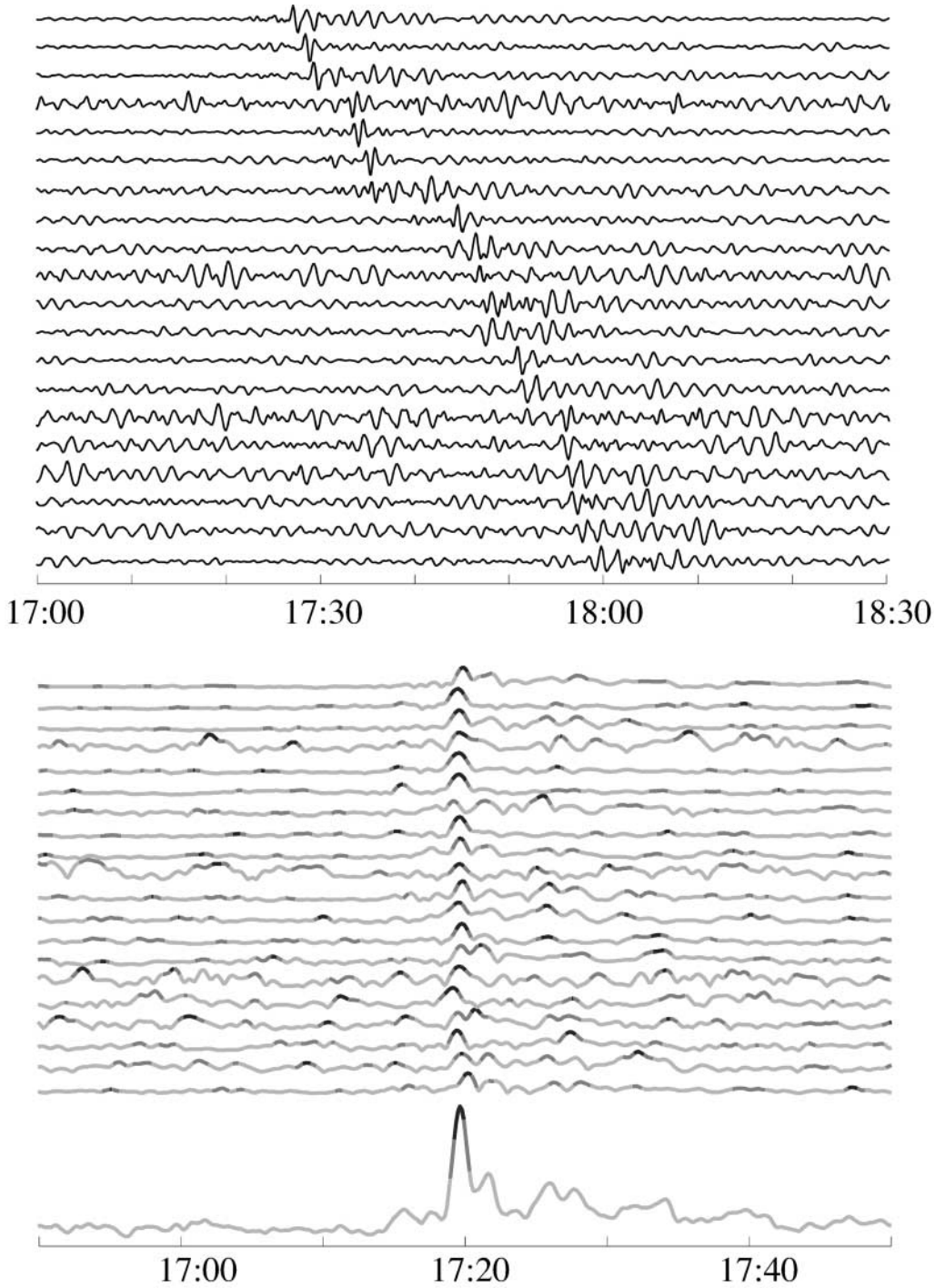

Figure 1. Teleseismic long-period seismograms for an M_{SW} 5.0 event that occurred in Greenland on 28 December 2001. Top panel shows long-period seismograms $u(t)$ for stations ranging in distance from 19° (top trace) to 154° (bottom trace) from a test location at 68.75° N, 33.25° W. In the bottom panel, the propagation effects from the test location to each station have been deconvolved and the envelope $u_E(t)$ of each seismogram has been calculated. The shading of each seismogram depends on the correlation $u_C(t)$ of the envelope $u_E(t)$ with the template peak shape, a measure of the probability that Rayleigh wave energy has been detected. A darker shading indicates higher correlation. The bottom trace is a stack of the individual station traces. The alignment of the individual detections indicates the presence of Rayleigh wave energy emanating from the test location. The time of the peak correlation in the stack gives the event origin time. Note that the bottom panel has a narrower timescale than the top panel.

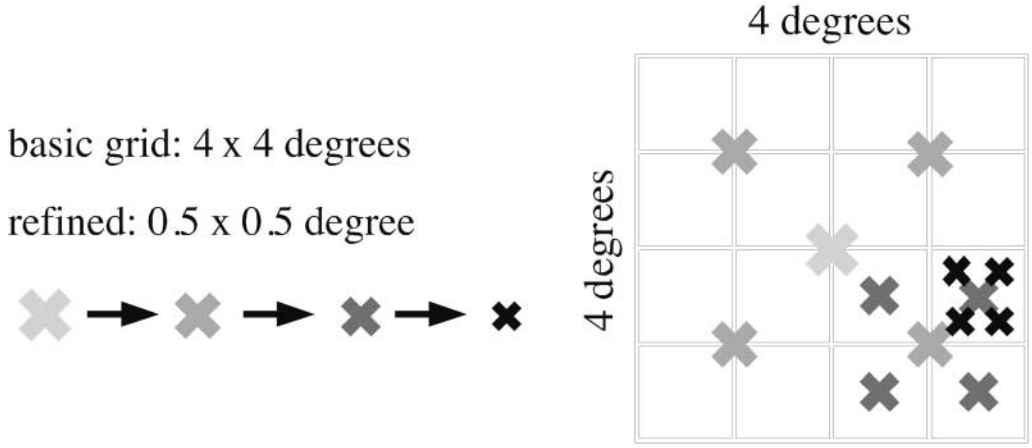

Figure 2. Diagram describing the nested search. A global search is performed on a 4° × 4° grid. If a signal is detected, the grid is gradually refined so that the smallest grid sampled is 0.5° × 0.5°.

terion. The same is true for adjacent times. The objective of the location stage is to sift the detections and determine the optimal times and locations for the detected events.

The location stage involves processing all the detections for a given period, usually one month. All the detections are read in (50,000 or more for a typical month). The detection with the largest value for the stacked correlation function $u^S(t)$ defines the first event. Additional detections associated with this event are identified based on proximity in time and space: if they are within 300 sec and 15 degrees they are considered associated and cannot form a new event. An exception is made for detections more than 90 sec before or after the event occurring at exactly the same location. Once the first event has been identified, and all the associated detections have been excluded, the process is repeated until there are no unassociated detections left. About 200 events are typically identified in a given month.

The located events are assigned a quality, A, B, or C, based on the value of the stacked correlation function and the number of records that satisfy the single-record detection criterion. Quality A is assigned to events with $u^S(t) > 0.75$ and for which at least 25 of 30 records show a detection. Quality B is assigned to events with $u^S(t) > 0.725$ that do not satisfy the A criteria. Quality C is assigned to events with $u^S(t) > 0.7$ that do not satisfy the A or B criteria.

When the event list for a month has been determined, two additional criteria are applied to identify potentially false events. First, we have found that detections are occasionally generated based on major-arc arrivals. The result is that events become defined close to the antipode of an earlier, true earthquake. This is by itself surprising, because it suggests that low attenuation or strong focusing makes it possible for intermediate-period surface waves to be observed at unusually large distances. Here, we are interested in removing these false events; we therefore apply the criterion that events close to the antipode ($\Delta > 170°$) occurring 5,000 to 6,000 sec after a larger-magnitude event are flagged as antipodal. Similarly, we have found that a false event of low quality is occasionally defined based on the waveforms from an earlier event. We therefore flag all quality-C events that occur 60 to 5000 sec after another event as false. Clearly, these criteria may exclude some true events. However, our objective here is not to maximize completeness, but rather to minimize the number of false events.

Results

We have applied the detection and location algorithm to 11 years of data, from 1993 to 2003. This analysis leads to the detection of 24,412 events. We refer to this collection of events as the "surface-wave catalog." The locations are shown in Figure 3. Of these events, 16,795 are of quality A; 5232 are of quality B; and 2385 are of quality C. Most of the events are known earthquakes that have been reported in traditional hypocenter catalogs. A subset of these events have also been analyzed by the Harvard Centroid Moment Tensor (CMT) project (Dziewonski *et al.*, 1981; Ekström *et al.*, 2005) and therefore have associated moment tensors and scalar moments. A smaller number of the events are not found in any traditional catalog and are therefore referred to here as "new" earthquakes.

Calibration

We associate the events in the surface-wave catalog with events in other catalogs by requiring that the events (1) have event times in the two catalogs that do not differ by more than 120 sec and (2) epicenters not separated by more than 5°. We first associate the events in the surface-wave catalog with events in the CMT Catalog to calibrate the magnitudes. We find that 9482 of 10,159 (93.3%) CMT earthquakes for this period are detected and located. Most of the CMT earthquakes not detected are deep and therefore do not generate large surface waves. More than 98% of all

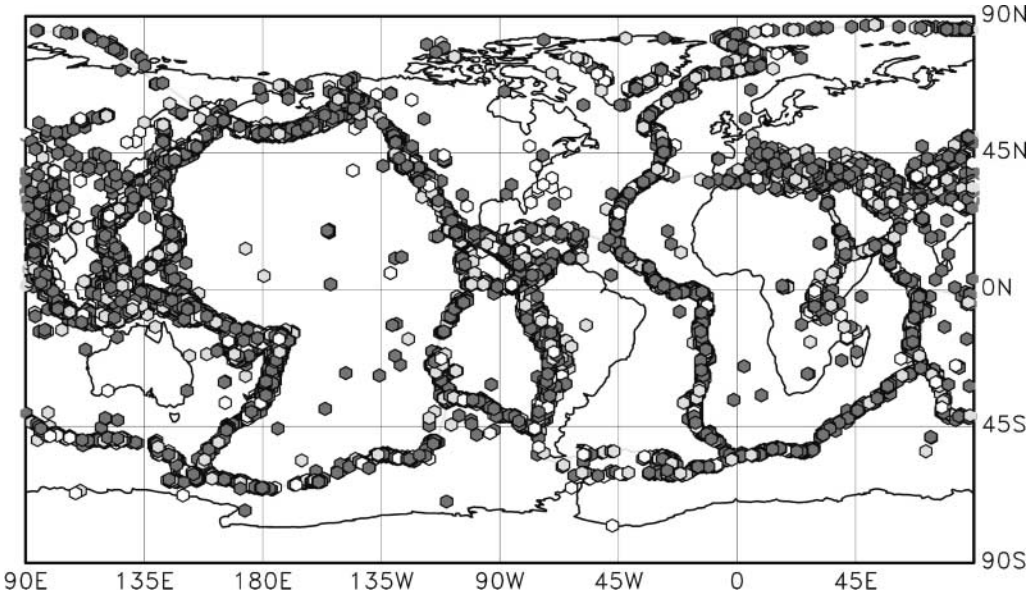

Figure 3. Map showing the locations of the 24,412 events detected and located in this study. The quality of the detection is indicated by the shading of the symbol. Dark-gray symbols indicate the best quality (A), light-gray symbols intermediate quality (B), and open symbols the lowest quality (C).

shallow ($h < 70$ km) earthquakes in the CMT Catalog are represented in the surface-wave catalog.

We use the events present in both the CMT Catalog and the surface-wave catalog to calibrate the magnitude M_{SW} estimated from the amplitude of the surface-wave source pulse. Figure 4 shows a comparison of the moment magnitude M_w determined in the CMT analysis and the surface-wave magnitude determined here. The parameter c in equation (12) was determined to minimize the misfit between the two magnitudes in the range $5.0 \leq M_w \leq 7.0$. The magnitude M_{SW} correlates very well with M_w with a standard deviation of $\sigma = 0.14$. It is clear that M_{SW} saturates with respect to M_w for earthquakes greater than M_w 7.0. This is to be expected, because the corner periods for these large events will be longer than the shortest period used in the detection algorithm. The fact that deep events generate smaller surface waves than shallow events is reflected in the M_{SW} frequently being smaller than M_w for deeper earthquakes (Fig. 4).

Figure 5 shows the distance between events common to the CMT Catalog and the surface-wave catalog. The average separation between epicenters is 48 km; the median separation is 43 km. Because this calculation does not include the estimated uncertainty of approximately 25 km in the CMT locations (Smith and Ekström, 1997), these values can be used as conservative uncertainties in the locations. Approximately 95% of the separations are smaller than 100 km, and nine events have separations greater than 200 km. The origin times show a smaller scatter. Figure 6 shows the distribution of origin-time differences between the CMT centroid time and the event time obtained in the surface-wave algorithm. The distribution is nearly symmetric with a mean near zero (0.6 sec). Such a distribution is consistent with the surface-wave origin time reflecting the center of moment release, as does the centroid location of the CMT analysis.

The results we obtain from calibration using events in the CMT Catalog indicate that the surface-wave algorithm can provide a reliable estimate of moment magnitude, with a standard deviation of only 0.14 magnitude unit. The accuracy in epicentral location and origin time is limited, with uncertainties of ~50 km and 6 sec. This is not surprising, given the long period of the signals used and the coarseness of the spatial and temporal grid. We do not observe any clear correlations of the accuracy of the epicenter, origin time, or magnitude estimate with event size.

Using the calibrated magnitude scale M_{SW}, we assess the completeness of the surface-wave catalog by construction of a magnitude-frequency diagram. Figure 7 shows the number of events of different magnitudes contained in the surface-wave catalog. The cumulative and noncumulative distributions are strikingly consistent, with a slope of $b = -1$ over a wide range of magnitudes (M_{SW} 4.9–7.4). The catalog rapidly becomes incomplete below magnitude 4.9, and no event with magnitude smaller than 4.6 was detected. A comparison of magnitude-frequency diagrams for events in the Southern and Northern Hemispheres shows a difference in the completeness threshold magnitude of approximately 0.1 magnitude unit, with small events in the Northern Hemisphere being detected more frequently, presumably as a result of the denser station coverage there. It appears likely that the sharp termination of the distribution at small magnitudes reflects the relatively constant noise levels in the Earth around 40 sec period, and the existence of a threshold

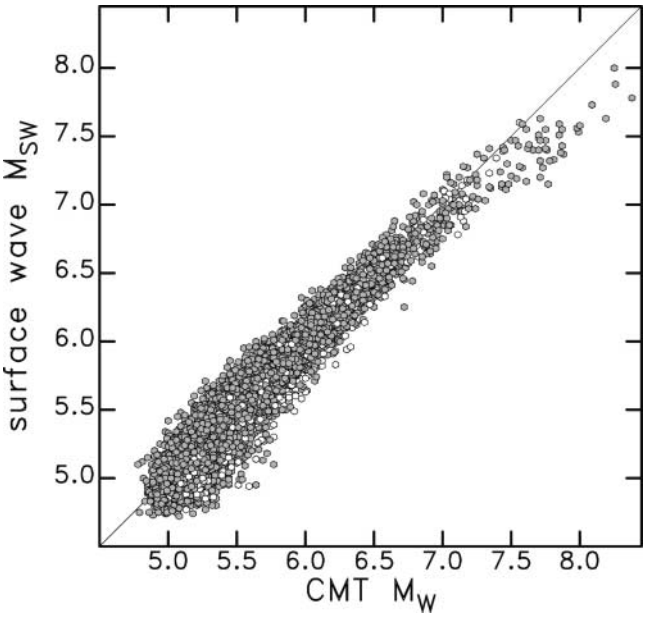

Figure 4. Correlation of the magnitude M_{SW} determined by the surface-wave event-detection algorithm with the moment magnitude M_w calculated by CMT analysis for 9482 earthquakes. Deep earthquakes, identified by open symbols, generate smaller-amplitude surface waves, and their sizes are therefore slightly underestimated by the detection algorithm. The M_{SW} scale is saturated for events of $M_w > 7.0$.

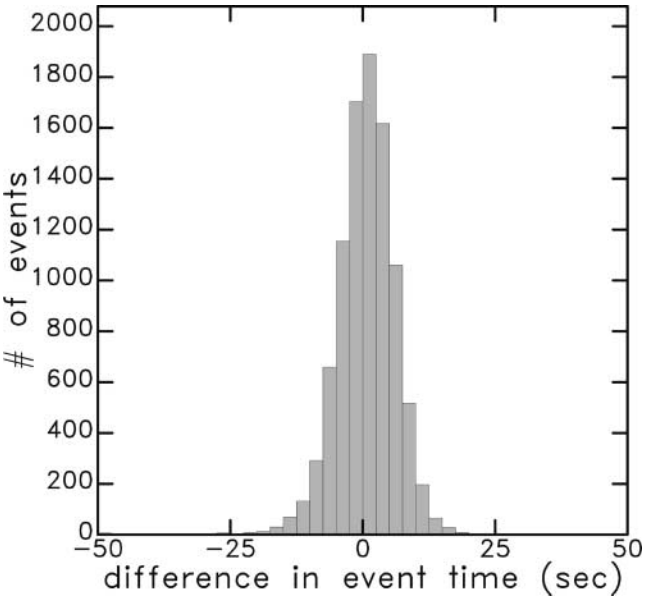

Figure 6. Histogram showing the difference between the event time determined in the surface-wave location algorithm and the CMT centroid time for the 9482 events common to the two catalogs. The symmetry of the distribution and the small mean separation (0.6 sec) are consistent with both event-time estimates depending on the center of moment release, as distinct from the event onset. The small spread of the distribution ($\sigma = 5.8$ sec) indicates that the event time can be estimated well despite the 8-sec sample rate in the original waveforms.

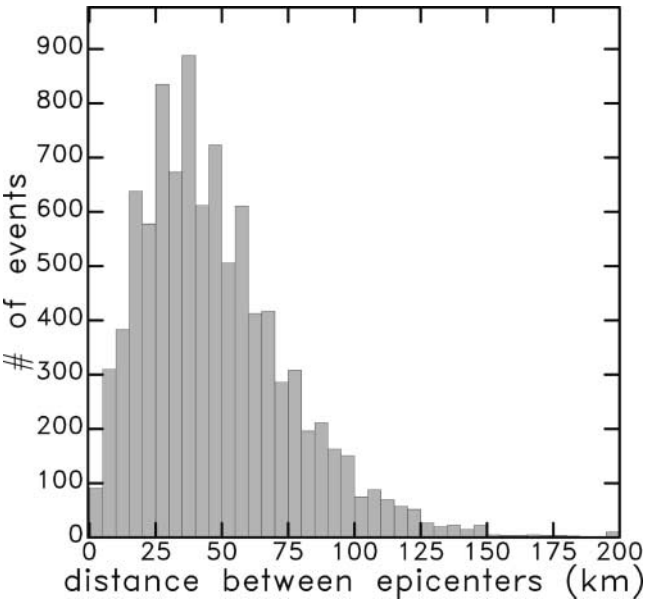

Figure 5. Histogram showing the distance between epicenters determined by the surface-wave location algorithm and CMT analysis for the 9482 events common to the two catalogs. The mean separation is 48 km; the median separation is 43 km. Nine events with separations greater than 200 km are included in the rightmost (195–200 km) bin.

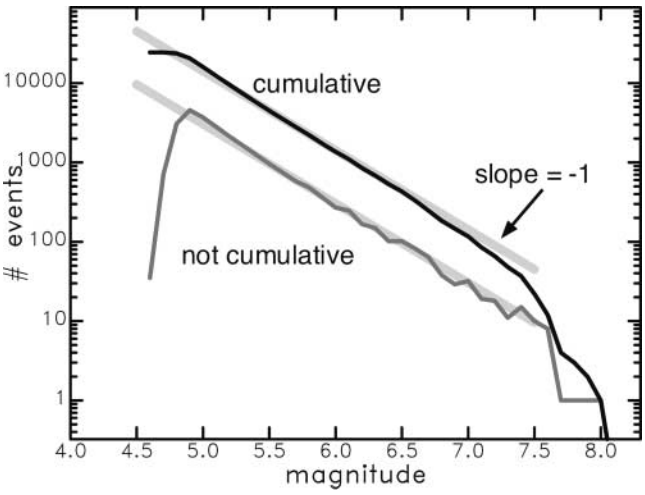

Figure 7. Magnitude-frequency diagram showing the size distribution of the 24,412 events detected in this study. The bin width is 0.1 magnitude unit. Both the cumulative and noncumulative distributions are shown, as well as reference lines with a slope of $b = -1$.

signal-to-noise value below which the algorithm cannot find a coherent signal. The lack of events with smaller magnitude also suggests that completely spurious detections, generated by noise, are rare.

New Earthquakes

We associate the events in the surface-wave catalog with those found in standard catalogs to determine which of them have not previously been reported. We limit the comparison to three standard catalogs: (1) for the period 1993–2003, we use the Preliminary Determination of Epicenters (PDE) compiled by the NEIC, (2) for the period 1993–2001 we also use the global catalog of the ISC, and (3) for 2002–2003, we use the Reviewed Event Bulletin (REB) of the IMS. For each year, we therefore associate the surface-wave events with two catalogs: PDE and ISC, or PDE and REB.

Using the event criteria described previously, we find that, for the period 1993–2003, the surface-wave catalog contains 1301 "new" earthquakes not previously reported in the global catalogs. For the period 1993–2001, the surface-wave catalog contains 1118 events not found in the combined PDE and ISC catalogs; for the period 2002–2003, the surface-wave catalog contains 183 events not reported in the combined PDE and REB catalogs. For the complete 11-year period, 550 of the new earthquakes have $M_{SW} \geq 5.0$. The geographical distribution of the new earthquakes is shown in Figure 8. Most of the new earthquakes are located near plate boundaries or in other known seismogenic areas. This clear association with known seismogenic areas supports the notion that few of the events represent spurious detections, because random detections would be likely to have a random geographical distribution. A preponderance of the events are located in the southern oceans along the ridge-transform plate boundaries. Previous studies have also found earthquakes missing from standard catalogs (e.g., Rouland *et al.*, 1992; Shearer, 1994; Ekström, 2001) in these areas. It is not clear at this time whether the nondetection of these events using high-frequency *P* waves is caused entirely by the remoteness of the southern oceans, or whether these events have source spectra depleted in high-frequency radiation. Other new events are located in areas where seismic activity is unexpected (e.g., on Greenland and off the East Coast of North America), or where the detection of new earthquakes is surprising (e.g., near California, Japan, or Europe) given the presence of regional networks that provide monitoring of small earthquakes. The new earthquakes on Greenland have been the subject of a focused investigation (Ekström *et al.*, 2003), and their existence and anomalous character have been verified. We are currently investigating other groups of new events and a complete catalog of these events is in preparation for publication.

Event Verification

To illustrate the character of the data that lead to the detections of new earthquakes and to illustrate one way in which we verify that a new event has been detected, we present record sections for two new earthquakes. Figure 9 shows long-period seismograms for one of of the larger previously undetected events. This earthquake (10 December 2001) occurred on the Pacific-Antarctic ridge; the magnitude estimate for the event is M_{SW} 5.6. The amplitudes of both long-period surface waves and body waves are clearly well

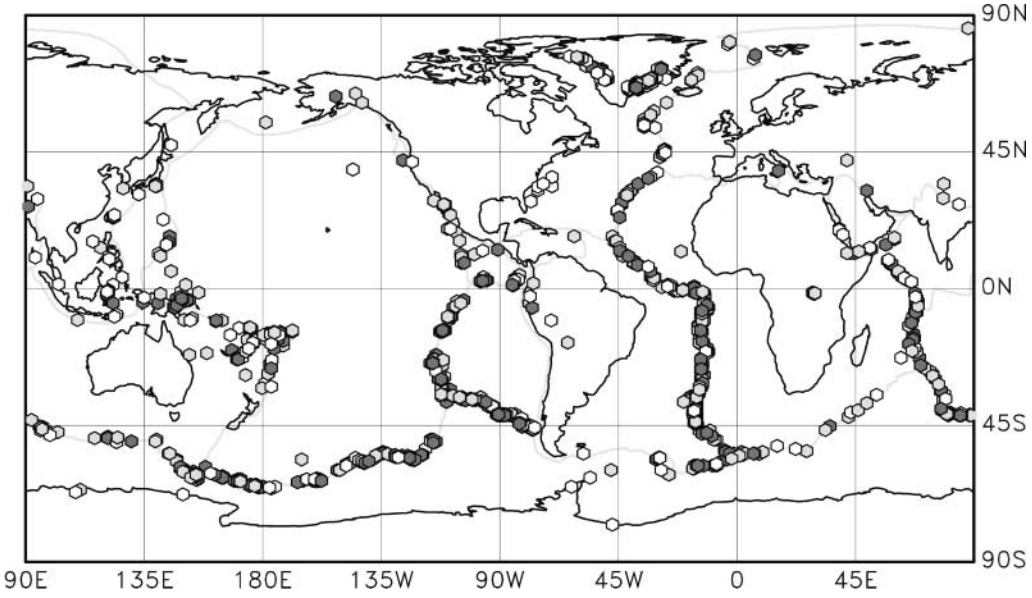

Figure 8. Map showing the locations of 1301 previously undetected earthquakes detected and located by using the surface-wave algorithm. The smallest earthquake has an estimated magnitude of M_{SW} 4.6 and the largest M_{SW} 5.9. The quality of the detection is indicated by the shading, as in Figure 3.

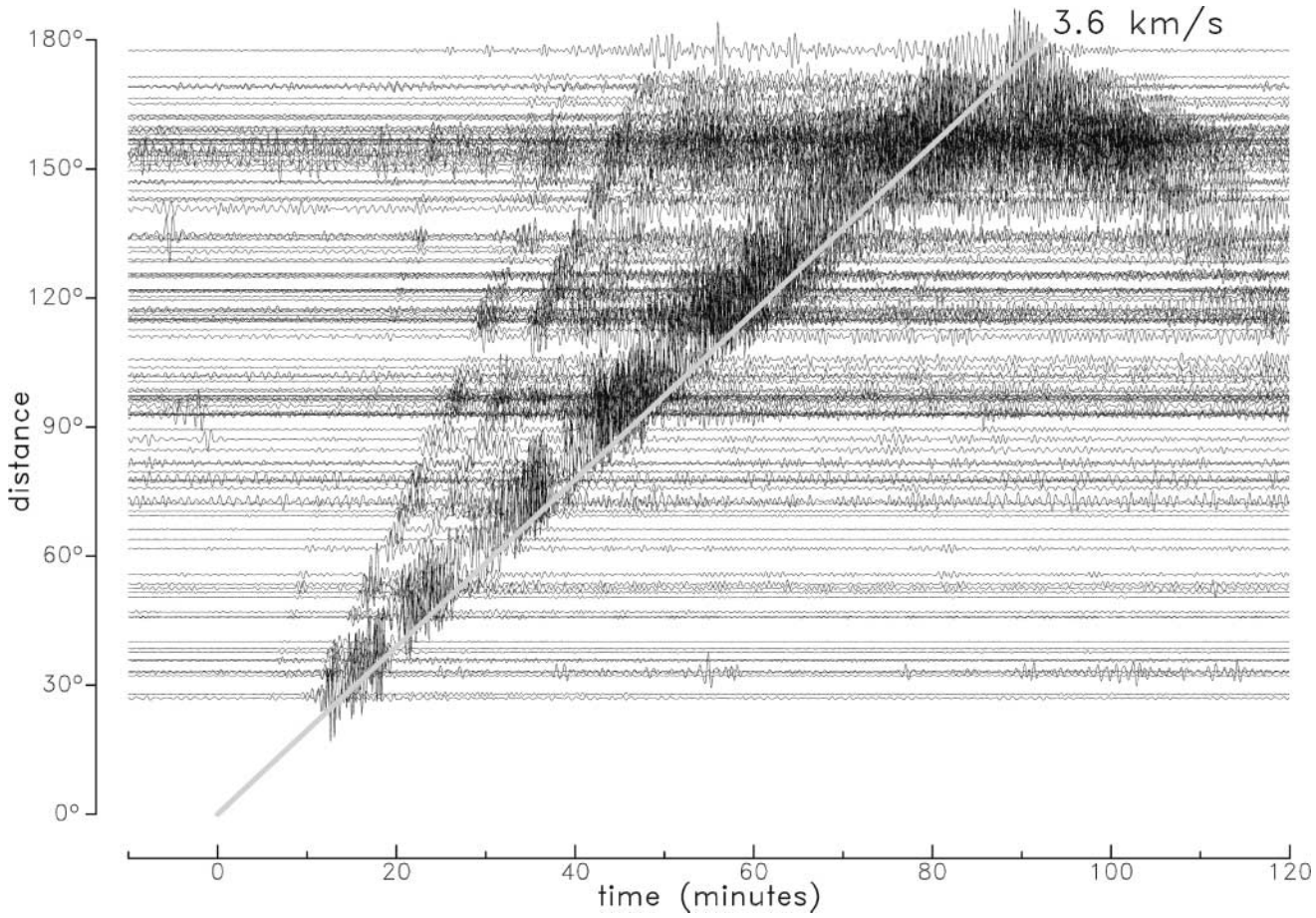

Figure 9. Record section showing seismograms for a previously unreported M_{SW} 5.6 earthquake on the Pacific-Antarctic ridge (10 December 2001, 14:38:00; 57.25° S, 141.25° W). The seismograms have been bandpass filtered between a 25- and 75-sec period. Clear surface-wave and body-wave arrivals can be seen, with a moveout consistent with the location and time of the event. The gray line corresponds to a velocity of 3.6 km/sec.

above the noise level. We are able to calculate a standard CMT focal mechanism for the event and obtain a strike-slip mechanism, consistent with the tectonic setting, and a moment magnitude M_w 5.8. Further analysis of high-frequency *P* waves would be needed to determine why this event is not included in the PDE or ISC catalogs.

Waveforms for a smaller new earthquake (26 January 2000) are shown in Figure 10. This event is located in the Gulf of California; the surface-wave analysis leads to a magnitude M_{SW} 4.8. This event is assigned quality B. While a traditional record section shows surface waves consistent with the event epicenter and time, we show here the envelopes of dispersion-corrected waveforms ($u_E(t)$) to illustrate better the complexity of the event. This migrated record section shows that the new event is, in fact, a sequence of at least three events, with two smaller events following the first, larger event with delays of approximately 4 and 26 min. There is also some suggestion in the waveforms that the main event was preceded by a smaller event approximately 3 min earlier.

Discussion

The event-location algorithm developed in this report is based on basic signal- and array-processing concepts. The technical advance introduced here is the combination of surface-wave phase equalization, waveform-envelope calculation, and cross correlation with the envelope of a synthetic source pulse. In addition, we have chosen to stack correlation functions rather than phase-equalized envelopes to enhance the robustness of the stack.

The systematic application of our new method shows that it is successful at detecting and locating earthquakes of $M_w > 4.6$. It is difficult to estimate the completeness of the catalog generated by the algorithm, but comparison with the CMT Catalog indicates that more than 98% of shallow earthquakes of $M_w > 5.5$ are detected. The magnitude-frequency relationship for the surface-wave catalog suggests that this completeness level may extend to magnitudes as low as M_w 4.8, but this remains to be verified.

More than 1000 earthquakes not appearing in traditional

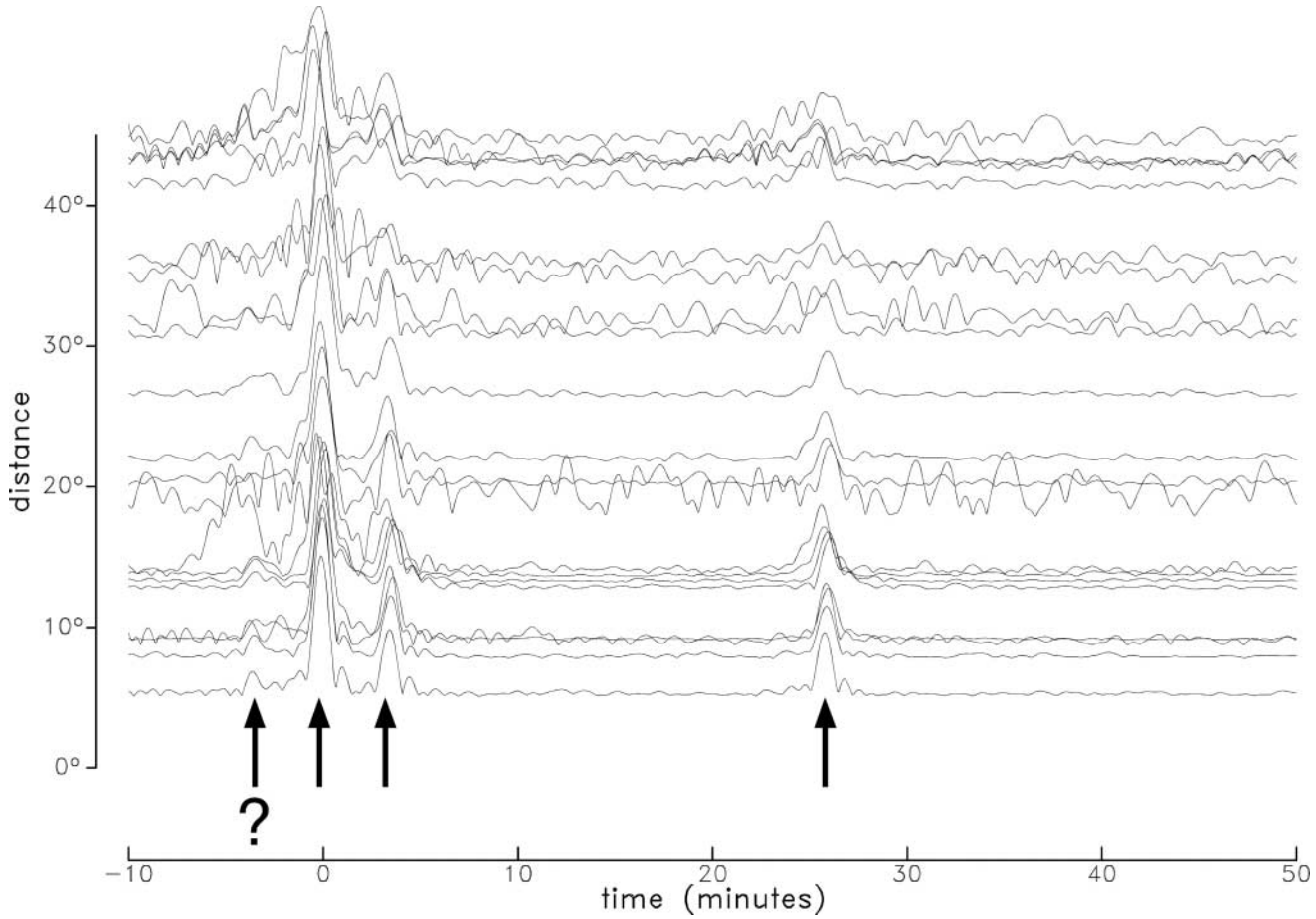

Figure 10. Record section showing seismograms for a previously unreported M_{SW} 4.8 earthquake in the Gulf of California (26 January 2000 12:08:08; 27.25° N, 111.25° W). The seismograms have been bandpass filtered between 25- and 75-sec period. The propagation phase has been removed using phase velocities from the preliminary reference Earth model (Dziewonski and Anderson, 1981), and the envelopes of the resulting waveforms have been calculated. Note the prominent source pulse near $t = 0$ defining the event, but also the smaller peaks 4 and 26 min later and, possibly, 3 min earlier, suggestive of additional earthquakes (aftershocks and foreshock) at approximately the same location.

seismicity catalogs have been detected and located. The causes for their earlier nondetection are varied and additional analysis will be needed to determine which of the new events have anomalous source characteristics.

Acknowledgments

The seismic waveforms used in this study were obtained from the Global Seismographic Network operated by Incorporated Research Institutions for Seismology and the U.S. Geological Survey and the Geoscope, Mednet, GEOFON, and Canadian seismic networks. I am grateful to everyone involved in the collection and distribution of data from these networks. I also thank Jeff McGuire and Meredith Nettles for many helpful comments on the manuscript. This work was funded by National Science Foundation Award EAR-0207608.

References

Dziewonski, A. M., and D. L. Anderson (1981). Preliminary reference Earth model, *Phys. Earth Planet. Interiors* **257,** 297–381.

Dziewonski, A. M., and A. L. Hales (1972). Numerical analysis of dispersed seismic waves, in *Methods in Computational Physics*, Vol. 11, Academic Press, New York.

Dziewonski, A. M., T.-A. Chou, and J. H. Woodhouse (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.* **86,** 2825–2853.

Ekström, G. (2001). Time domain analysis of Earth's long-period background seismic radiation, *J. Geophys. Res.* **106,** 26,483–26,493.

Ekström, G., A. M. Dziewonski, N. N. Maternovskaya, and M. Nettles (2005). Global seismicity of 2003: Centroid-moment tensor solutions for 1087 earthquakes, *Phys. Earth Planet. Interiors* **148,** 327–351.

Ekström, G., M. Nettles, and G. A. Abers (2003). Glacial earthquakes, *Science* **302,** 622–624.

Ekström, G., J. Tromp, and E. W. F. Larson (1997). Measurements and global models of surface wave propagation, *J. Geophys. Res.* **102,** 8137–8157.

Ihmlé, P. F., and T. H. Jordan (1996). A space-time grid search for infraseisms, *EOS Trans. AGU* **77** (suppl.), F501.

Ringdahl, F., and E. S. Husebye (1982). Application of arrays in the detection, location, and identification of seismic events, *Bull. Seism. Soc. Am.* **72,** S201–S224.

Rouland, D., C. Condis, C. Parmentier, and A. Souriau (1992). Previously undetected earthquakes in the southern hemisphere located using long-period Geoscope data, *Bull. Seism. Soc. Am.* **82,** 2448–2463.

Shearer, P. M. (1994). Global seismic event detection using a matched filter on long-period seismograms, *J. Geophys. Res.* **99,** 13,713–13,735.

Smith, G. P., and G. Ekström (1997). Interpretation of earthquake epicenter and CMT centroid locations, in terms of rupture length and direction, *Phys. Earth Planet. Interiors* **102,** 123–132.

Dept. of Earth and Planetary Sciences
Harvard University
20 Oxford Street
Cambridge, Massachusetts 02138
ekstrom@seismology.harvard.edu

Manuscript received 15 August 2005.