JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 99, NO. B7, PAGES 13,713–13,725, JULY 10, 1994

Global seismic event detection using a matched filter on long-period seismograms

Peter M. Shearer
Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography
University of California, San Diego, La Jolla

Abstract. An image derived by stacking long-period seismograms is used as an empirical matched filter to detect and locate earthquakes. Records from 564 events ($m_b \geq 6$) recorded at very long periods ($T \geq 60$ s) by the 15 to 20 stations of the International Deployment of Accelerometers (IDA) network are stacked using a method that emphasizes the surface wave arrivals. The first 3 hours of this time versus range image are used to construct a matched filter for continuous application to the IDA data. The output of this filter contains spatial and temporal peaks that define the location and origin time of probable seismic events. Implementation of this technique to 11 years of IDA data from 1981 to 1991 identifies 4061 events. These include 65% of cataloged events of $m_b \geq 5.5$. Earthquakes which appear anomalously strong relative to their surface wave magnitudes are mainly located on oceanic transform faults and probably represent unusually slow ruptures. The method successfully detects two eruptions of the El Chichón volcano in southern Mexico, from the anomalous low-frequency energy radiated during the eruptions. In addition, 32 earthquakes are detected which are not in the standard global catalogs. These events appear to be about $M_s = 5$ and are mainly located in the southern oceans, where there are gaps in the coverage of the high-frequency networks. Although these earthquakes are probably "slow" since they occur mostly on oceanic transform faults, they are not "silent" as they also can be observed in higher-frequency data.

Introduction

Global seismic events have traditionally been detected through the identification and timing of P wave arrivals on individual seismograms. Analyses of these times from a variety of stations then permits the detection and location of events and the generation of an earthquake catalog. Various algorithms have been devised for automatic event detection, using both small-aperture arrays and regional or global networks [e.g., *Ringdal et al.*, 1977; *Blandford*, 1982; *Ringdal and Husebye*, 1982; *Blandford and Goncz*, 1985; *Bratt and Bache*, 1988; *Roberts et al.*, 1989; *Sereno and Bratt*, 1989; *Bache et al.*, 1990; *Bratt et al.*, 1990; *Cassidy et al.*, 1990; *Joswig*, 1990; *Kushnir et al.*, 1990; *Ruud and Husebye*, 1992; *Tsvang et al.*, 1993]. In most cases, these schemes are based on associating observed arrival times of individual seismic phases with those predicted by a theoretical model.

In this paper, I describe experiments with a more empirical approach to this problem, which utilizes a matched-filter technique to automatically detect and locate earthquakes. This method uses a time versus distance image of the wave field from known events as a matched filter to identify additional events which generate a similar pattern of arrivals. The output of the filter contains spatial and temporal peaks that define the location and origin time of probable seismic events. The fit is based on the entire seismic wave field and not simply on the arrival times of particular seismic phases. For this reason the method is fairly insensitive to noise spikes and other data artifacts and requires no hand picking of the seismograms.

This technique is particularly effective at low frequencies due to the coherence and multiplicity of long-period seismic phases. Here, I apply the method to 11 years of very long period data recorded by the International Deployment of Accelerometers (IDA) network. A comparison of the results with standard event catalogs demonstrates the effectiveness of the method, even with a relatively sparse station distribution, and reveals the occurrence of ~3 events per year that had not previously been detected.

IDA Network

The IDA network [*Agnew et al.*, 1986] is a global array of ~20 digital vertical component accelerometers that record at very long periods ($T \geq 60$ s). Its primary purpose has been to record Earth's normal modes, but the data also contain direct surface and body wave arrivals. A significant advantage of the IDA data set for this study is the 10-s sample interval: 10 years of con-

Paper number 94JB00498.
0148-0227/94/94JB-00498$05.00

tinuous data consume less than a gigabyte of storage. Figure 1 shows the distribution of IDA stations, not all of which were operational at the same time. Although IDA began recording in 1976, the early data are from relatively few stations and were digitized at a sample interval of 20 s. This analysis will be restricted to times after 1980, when the 10-s sample interval became standard. The number of operating stations increased gradually during the early 1980s with an average of about 15 stations recording events during 1985.

Figure 2 shows the first 3 hours of a stacked record section of 8097 IDA seismograms recorded from 564 events of $m_b \geq 6$ from 1981 to 1991. The automatic gain control stacking procedure described by *Shearer* [1991a] is used, in which a running short-term average to long-term average (STA/LTA) ratio is computed for each seismogram prior to stacking. In this case, the STA time is set to 5 min and the LTA time to 30 min. The data are high-pass filtered before applying the STA/LTA algorithm to exclude frequencies below 3 mHz but are not edited to remove clipped data or noisy records. To avoid the largest effects of depth phases, only earthquakes shallower than 200 km are used in the stack. The data are binned in 10-s time increments and 0.5° range increments, and a nine-point median filter [see *Shearer*, 1991a] is applied to smooth the image.

This plot shows a somewhat blurry image of the P and S body wave arrivals in the region before the first Rayleigh wave. Much sharper images of these phases can be obtained by using shorter time intervals for the STA/LTA filter. However, these improvements degrade the appearance of the surface waves and lead to computational problems in the implementation of the matched filter (more about this later). Polarity information is discarded in this stack, since the abolute value is taken of the seismograms prior to applying the STA/LTA filter. Reference phase stacking techniques preserve polarity information [e.g., *Shearer*, 1991b] and potentially could achieve greater sensitivity in detecting events. However, the detection problem would become considerably more complicated since radiation pattern differences between events would need to be taken into account, and a finer time sampling in the matched filter would be required to resolve the changing polarities.

The second Rayleigh wave ($R2$) can be seen coming back toward the source at times between 90 and 180 min (the Rayleigh waves are visible for over 10 hours after the event origin time using this stacking technique). High-order S surface multiples are visible (faintly) between 0° and 90° at about 100 min. Some energy can be seen arriving before the P wave, particularly at close ranges. This results from a slight acausality in the high-pass filter applied to large impulsive arrivals. The image is blank beyond 174°, reflecting a lack of data near the event antipodes. No editing of the data were performed prior to stacking; the effects of noise spikes, clipping, and other data artifacts are minimized simply by averaging over a large number of traces.

Matched Filter

The stacked image described above shows the observed behavior of seismograms as a function of range for known seismic events. We can use this image to

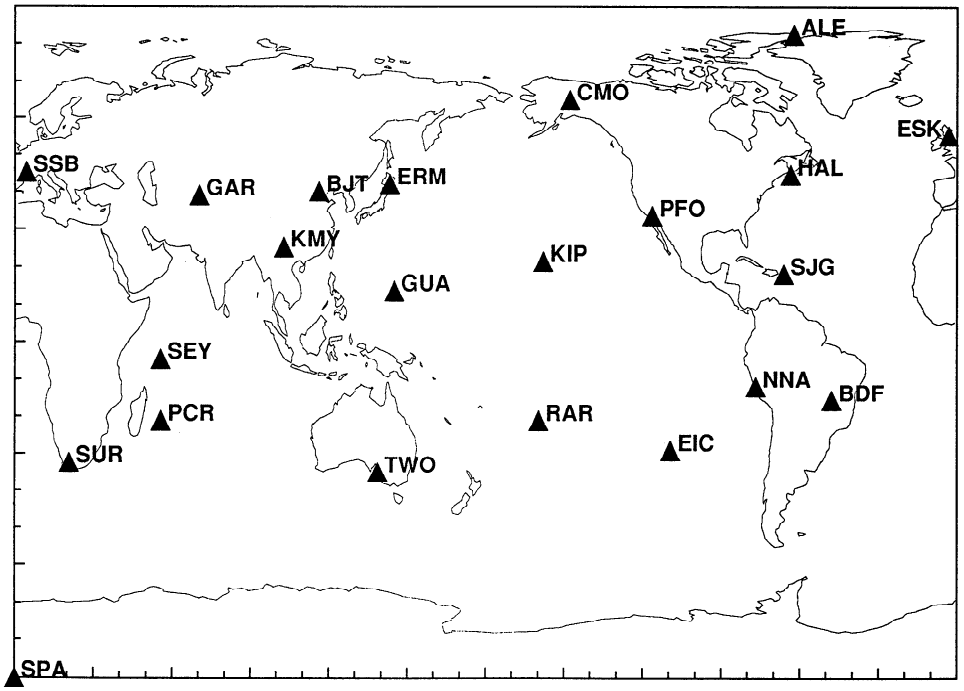

Figure 1. Locations of the 22 IDA stations used in this study. Not all of these stations were operational at the same time; in 1985 about 15 stations were recording data. Station SPA is located at the south pole.

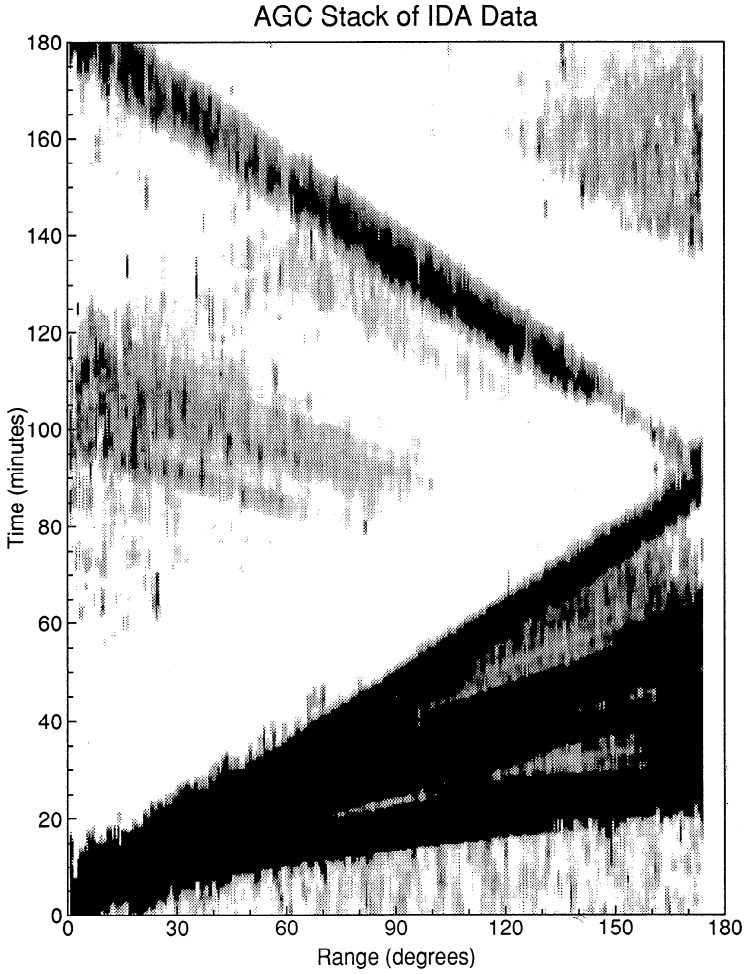

Figure 2. A stacked record section of 8097 IDA seismograms from 564 shallow events ($m_b \geq 6$, depth ≤ 200 km) from 1981 to 1991. Prior to stacking, the data are high-pass filtered to exclude frequencies below 3 mHz and then are processed using an automatic gain control algorithm. The data are binned in 10-s time increments and 0.5° range increments, and a nine-point median filter [see *Shearer*, 1991a] is applied to smooth the image. The STA/LTA amplitude scale used in this image ranges from 1.0 (white) to 1.7 (black). There are no data available near the antipode (beyond 174°).

search for unknown events by looking for correlations between the image and observed seismograms for a candidate earthquake. This is done by devising a matched filter based on the stacks. Figure 3 illustrates how this procedure works.

Assume we have seismograms $a_i(t)$ at stations $\mathbf{s}_i = (s_x, s_y)_i$. We run the seismograms through a processing filter $P[a_i(t)]$. This is any arbitrary function, i.e., a picking algorithm, or, in our case, an STA/LTA filter. We then use a series of known events, $\mathbf{e}_j = (e_x, e_y, e_z, e_t)_j$, and calculate the source-receiver ranges Δ_{ij}. We compute the global stack $\mathbf{S}$ as

$$\mathbf{S}(\Delta, T, e_z) = B\{P[a_i(t)], \mathbf{e}_j\} \quad (1)$$

where T is the time from the event origin time and B is the stacking method. For the image shown in Figure 2, B is simply an average of $P[a_i(t)]$ computed for all seismograms with source-receiver ranges within a small interval around Δ.

For a given candidate location $\mathbf{q} = (q_x, q_y, q_z)$ and the filtered seismograms $P[a_i(t)]$, we compute the dot product

$$\mathbf{S}(\Delta, T, e_z) \cdot B(P[a_i(t)], \mathbf{q}) = Q(t, \mathbf{q}). \quad (2)$$

The function $Q(t, \mathbf{q})$ will have maxima which define the origin time and location of likely events. This matched-filtering scheme is completely general and can be applied to any target image. For example, $\mathbf{S}(\Delta, T)$ could be defined using synthetic seismograms based on a reference Earth model or simply by using an envelope around a travel time curve for a target seismic phase. The advantage of using observations to image $\mathbf{S}(\Delta, T)$ directly is that uncertainties associated with the construction of the reference Earth model and computation of the synthetic seismograms are removed. The precision of the locations is limited by the scatter in the data itself.

The method is computationally intensive, since for each time step, $Q(t, \mathbf{q})$ must be sampled at closely

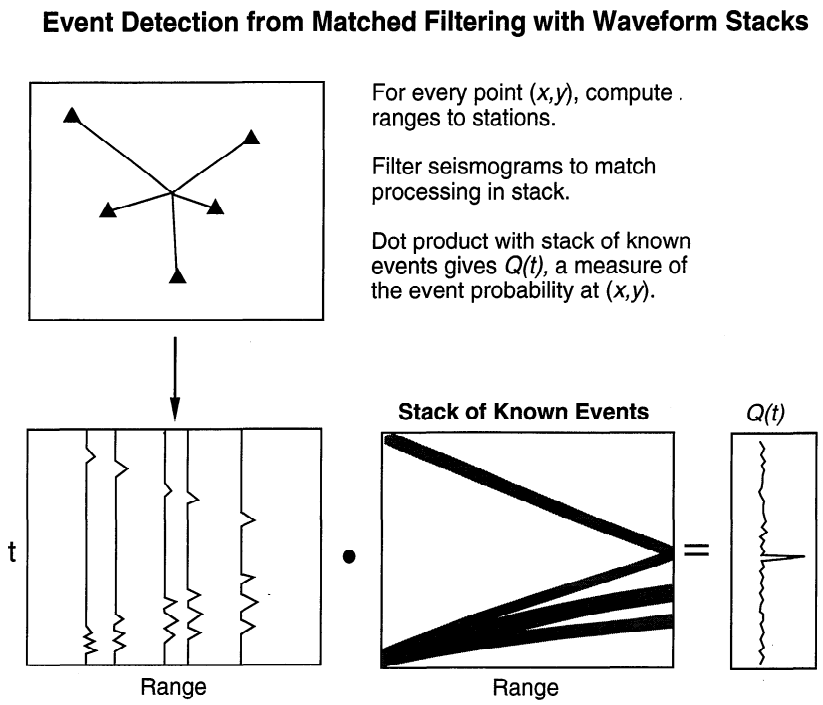

Figure 3. A cartoon illustrating the operation of the matched-filter procedure. The method searches for origin times and locations at which the observed range dependence of the waveforms correlates with the pattern obtained from known events.

spaced points on the Earth's surface at a range of possible source depths. If the $Q(t, \mathbf{q})$ maxima are sharply peaked, then more points in space and time are required in order to ensure that a peak is not missed. Thus a fairly blurry image such as that shown in Figure 2 is efficient for event detection purposes since it will lead to broad maxima in $Q(t, \mathbf{q})$. If more accurate locations are required for particular events, a sharper image could be applied during subsequent processing.

Some experimentation was necessary to develop a matched-filter algorithm from the image shown in Figure 2 that was reliable in detecting earthquakes at a reasonable computational cost. The method used for the examples in this paper has the following characteristics:

1. A set of 416 candidate earthquake locations are used, spaced approximately 10° apart on the Earth's surface. This gives a nominal location accuracy of 5°. All events are assumed to occur at the surface; however, in practice the method successfully detects deep events as well.

2. Ranges from these locations to the IDA station locations are computed, and the appropriate index number (for the range bin in the image) is saved. Since the 416 candidate locations are fixed, these source-receiver ranges need only be computed once.

3. The real values for the stack shown in Figure 2 are replaced by three integers. Values below 0.7 are assigned -1, values between 0.7 and 1.5 are assigned 0, and values above 1.5 are assigned 1. This scheme is also used for the processed seismograms $P[a_i(t)]$. In this way, all of the multiplications required to compute the dot product can be performed using a simple 3x3 lookup table. In addition, this normalization of the amplitudes has the effect of equalizing the influence of seismograms at different ranges, leading to more reliable earthquake locations. Without the normalization, the large amplitudes at close ranges tend to dominate and the information from the more distant seismograms is lost.

4. For 10-day time intervals, the IDA seismograms are high-pass filtered and processed using the same STA/LTA method applied to the traces used to generate Figure 2.

5. The dot product, $Q(t, \mathbf{q})$, is computed for each of the 416 locations at time increments of 2 min. This represents a factor of 12 decimation from the original sample interval (10 s) and increases the speed of the method greatly since the speed goes up as the square of the decimation rate. Note that the phases imaged in Figure 2 are generally 5 min or longer in width, so the 2-min sampling interval is sufficient.

6. At each time step, the maximum value of $Q(t, \mathbf{q})$ is determined from the 416 possible earthquake locations. This value is saved, together with the index number of the earthquake location, in a time series for subsequent processing. The time series, $Q(t, \mathbf{q})_{max}$, typically has nonzero mean since the maximum value of $Q(t, \mathbf{q})$ will tend to be positive. We remove this offset and any other long-period variations by applying a high-pass filter to $Q(t, \mathbf{q})_{max}$.

The method is computationally feasible since $P[a_i(t)]$

need only be computed once (for a time interval limited only by the computer memory); the dot product must be computed multiple times but this is a fast computation involving only addition (see point 3 above) which is suited for vectorization schemes. On a Sun ELC, 1 year of data could be processed in slightly more than a day (using the f77 compiler and the -O option). The speed was limited by the dot product sums, rather than by the i/o or filtering operations. Increasing the spatial or temporal resolution of the filter output is costly, since the computation time increases linearly with the number of candidate locations and as the square of the sample rate. If more accurate locations are desired, it would be much more efficient to implement a two-stage procedure in which a coarse grid search is used to detect the events, and a finer grid (perhaps based on a sharper stacked image) is used to obtain an improved location and origin time.

This algorithm was applied to 11 years of continuous IDA data (1981 to 1991). Figure 4 shows $Q(t, \mathbf{q})_{max}$ (high-pass filtered) for the first 11 days of 1987, compared to the preliminary determination of epicenters (PDE) earthquake magnitudes ($m_b > 5$) for the same time period. Peaks in $Q(t, \mathbf{q})$ can generally be associated with cataloged events, but not all of the cataloged events can be seen in $Q(t, \mathbf{q})$. For the larger events, $Q(t, \mathbf{q})$ will often show a peak followed by a series of secondary peaks at 90-min intervals. This is a result of surface wave aliasing where the technique will associate events with different parts of the surface wave train. For example, the first secondary peak at 90 min following the main peak is caused by the pattern formed by the $R2$ and $R3$ arrivals (which the method fits to $R1$ and $R2$, locating the event at the antipode). The second peak (at 3 hours) is caused by the $R3$ and $R4$ pattern, and is located at the same place as the actual earthquake. This aliasing continues until the surface waves die out (which can last 10 hours or more for the largest events).

In addition to these peaks following the main peak, there often is a precursory peak 90 min before the event, which is located at the antipode and results from the method associating $R2$ in the data with $R1$ in the matched filter. One approach to dealing with the surface wave aliasing would be to redesign the matched filter to reject these secondary features. This could be done, for example, by adding 90 min to the front of the filter to actively discriminate against any surface wave energy arriving before the event (to eliminate the trailing peaks) and by requiring that the $R2$ amplitudes be less than the $R1$ amplitudes (to eliminate the precursory peaks). However, this would complicate the matched-filter algorithm and reduce its efficiency. As an alternative, these secondary peaks can usually be identified and removed in postfilter processing since they occur at diagnostic 90-min intervals at locations defined by the main peak. This approach is used here, and the 11-year-long $Q(t, \mathbf{q})_{max}$ function is converted into a series of probable event times and locations.

Association With Known Events

The matched-filter procedure described above identifies several thousand possible events during the 1981–1991 period. By comparing origin times and locations, these events can be associated with known events in the PDE and International Seismological Centre (ISC) catalogs. An event is considered associated if the com-

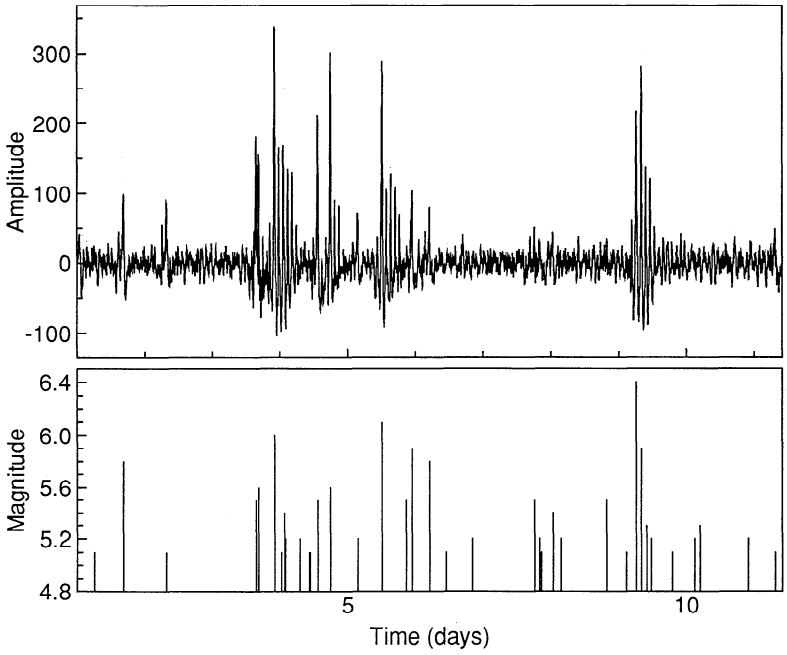

Figure 4. The matched-filter output ($Q(t, \mathbf{q})_{max}$) for the first 11 days of 1987 (top), compared to PDE seismicity ($m_b > 5$) for the same time period. Peaks in $Q(t, \mathbf{q})$ are generally associated with the occurrence of earthquakes.

puted origin time is within 7 min of the catalog time and the location is within 30°. These numbers are conservative in order to ensure identification of marginal events; in most cases the times and locations agree to within 2 min and 10° (i.e., the grid spacing).

Figure 5*a* shows a comparison between peak $Q(t, \mathbf{q})$ values and body wave magnitudes for 4061 identified events. The majority of the earthquakes are above $m_b = 5$, although a few events are seen at $m_b = 4.5$ and below. Some events are identified which are only in the ISC catalog; these are shown as pluses. The larger events, as one might expect, are associated with generally higher values of $Q(t, \mathbf{q})$. Figure 5*b* shows a comparison to the surface wave magnitude M_s. In this case, only 2870 points are plotted since the PDE catalog often does not list M_s. Figure 5*c* shows peak $Q(t, \mathbf{q})$ values plotted against M_W obtained for 3718 events from the Harvard centroid-moment tensor (CMT) catalog [e.g., *Dziewonski and Woodhouse*, 1983], using the *Hanks and Kanamori* [1979] scaling of moment to M_W.

Although the comparisons shown in Figure 5 exhibit a roughly linear trend between peak $Q(t, \mathbf{q})$ and event magnitude, there is considerable scatter among the individual points. This scatter may be attributed to two

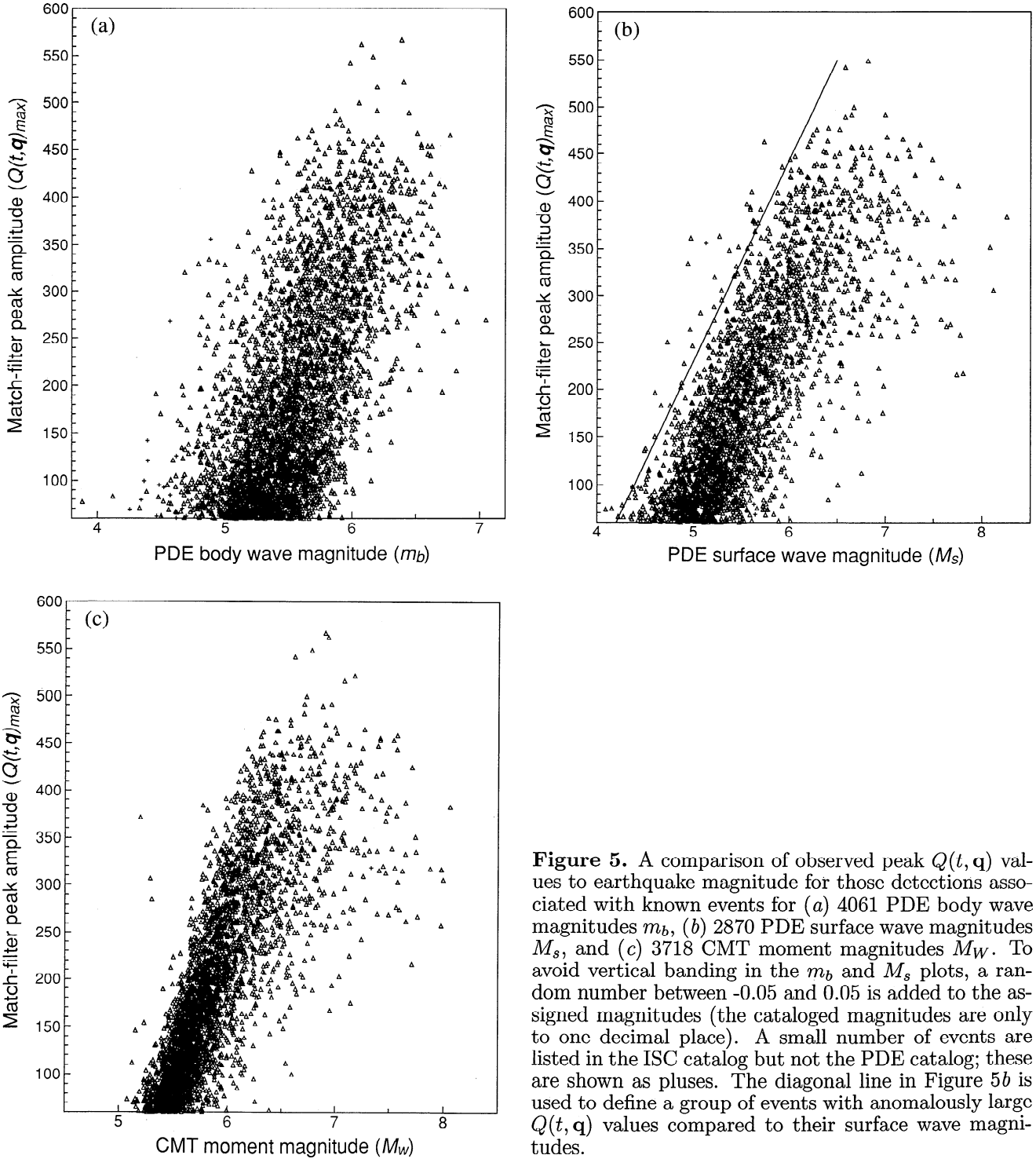

Figure 5. A comparison of observed peak $Q(t, \mathbf{q})$ values to earthquake magnitude for those detections associated with known events for (*a*) 4061 PDE body wave magnitudes m_b, (*b*) 2870 PDE surface wave magnitudes M_s, and (*c*) 3718 CMT moment magnitudes M_W. To avoid vertical banding in the m_b and M_s plots, a random number between -0.05 and 0.05 is added to the assigned magnitudes (the cataloged magnitudes are only to one decimal place). A small number of events are listed in the ISC catalog but not the PDE catalog; these are shown as pluses. The diagonal line in Figure 5*b* is used to define a group of events with anomalously large $Q(t, \mathbf{q})$ values compared to their surface wave magnitudes.

factors: First, $Q(t, \mathbf{q})_{max}$ is not a direct measure of surface wave amplitude but is more a measure of the clarity with which an event can be seen. Thus, for example, an event which occurs during a quiet period of background noise will be assigned a higher value of $Q(t, \mathbf{q})$ than an identical event occurring during a noisy period. In addition, the peak $Q(t, \mathbf{q})$ values saturate for the largest events. A $M_s = 7.5$ event cannot be seen more clearly than a $M_s = 6.5$ event (in fact, clipping and instrument nonlinearity are greater problems for the largest events). Second, apparent event magnitude varies with frequency. The matched-filter results are obtained at relatively long periods ($T \geq 60$ s) compared to the data used to assign the catalog magnitudes. Thus it is not surprising that the comparisons with the body wave magnitudes m_b show more scatter than the M_s and M_W comparisons, since M_s and M_W are obtained from lower-frequency data than m_b.

Figure 6 plots the fraction of PDE events detected as a function of event magnitude. The algorithm detects about 85% of all events greater than magnitude 6 and more than 60% of those events with $M_s \geq 5$. The detection efficiency is much less for events of small body wave magnitude (i.e., about 20% for $m_b \geq 5$); this probably reflects the difference in the frequency sensitivity of the M_s and m_b measurements and the lack of assigned M_s values for many of the smaller quakes. A number of quite large events are missed; for example, 10% of $M_s \geq 6.5$ events are not detected. The majority of these missed events are aftershocks or occurred during the coda of a large earthquake, during the time period that surface waves from the main event are contaminating the records. It is probable the performance of the algorithm could be improved substantially during these periods through some form of waveform stripping procedure.

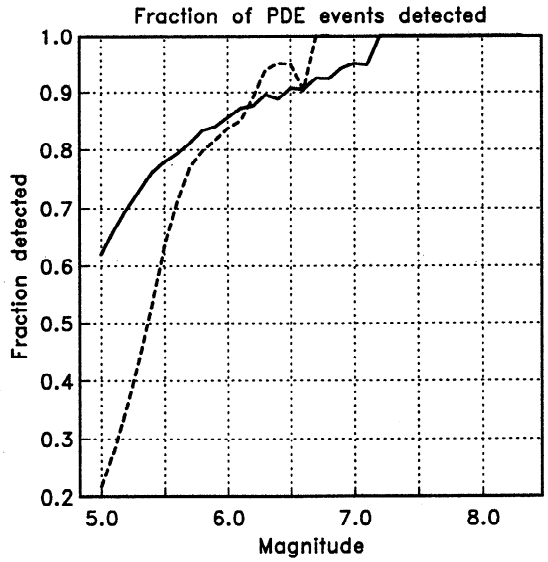

Figure 6. The fraction of PDE events detected as a function of minimum body wave magnitude m_b (dashed line) and surface wave magnitude M_s (solid line). Approximately 60% of PDE events with $M_s \geq 5.0$ are detected.

Slow Earthquakes

The term slow earthquake has been given to earthquakes that have slower-than-average rupture speeds and which radiate anomalously large amounts of low-frequency energy compared to their output at high frequencies. A review of this topic is given by *Beroza and Jordan* [1990]. Owing to the relatively low frequency content of the IDA seismograms, slow events might be expected to produce anomalously large $Q(t, \mathbf{q})$ values compared to their surface wave magnitudes. The diagonal line drawn on Figure 5*b* is used to define a small number of such anomalous events. Figure 7 shows PDE locations for these 40 events, almost all of which occurred on oceanic transform faults (the Challenger and Eltanin fracture zones are particularly well represented). This is consistent with previous studies which have indicated that earthquakes on oceanic transforms tend to have slow ruptures [e.g., *Kanamori and Stewart*, 1976; *Okal and Stewart*, 1982; *Stewart and Okal*, 1983; *Beroza and Jordan*, 1990; *Jordan et al.*, 1991; *Stein and Pelayo*, 1991; *Pelayo and Stein*, 1991].

These results suggest that some of the scatter shown in Figure 5 is due to variations in the frequency content of the events, and that this could be used as a way to identify potential slow events. *Beroza and Jordan* [1990] found 14 slow events by using an analogous method applied to their study of normal mode excitation in the IDA data set. However, the general matched-filter technique described here is not a particularly efficient way to study the frequency content of known events, for which more direct analysis techniques can be applied. A more profitable use of the matched-filter algorithm lies in its ability to detect previously unidentified events.

Unassociated Events

Not all of the peaks in the computed $Q(t, \mathbf{q})$ function can be associated with known CMT, PDE or ISC events. Figure 8 plots unassociated $Q(t, \mathbf{q})$ peaks versus time. For values of $Q(t, \mathbf{q})$ above 250, all peaks correspond to cataloged earthquakes. However, at smaller values, increasing numbers of unassociated peaks are seen For all peaks of $Q(t, \mathbf{q}) \geq 80$, I directly examined the IDA seismograms to see what was causing the peak. In many cases, particularly at low $Q(t, \mathbf{q})$ values, the algorithm triggered incorrectly. Most of these false triggers were caused by aliasing from large events (see above), which were not properly removed by the postfilter processing. In some cases the triggers resulted from multiple events producing overlapping arrival patterns. All of the false triggers were caused by earthquake generated signals, to which the algorithm assigned spurious locations or origin times. In no cases were $Q(t, \mathbf{q})$ peaks caused by uncorrelated, random noise on the seismograms.

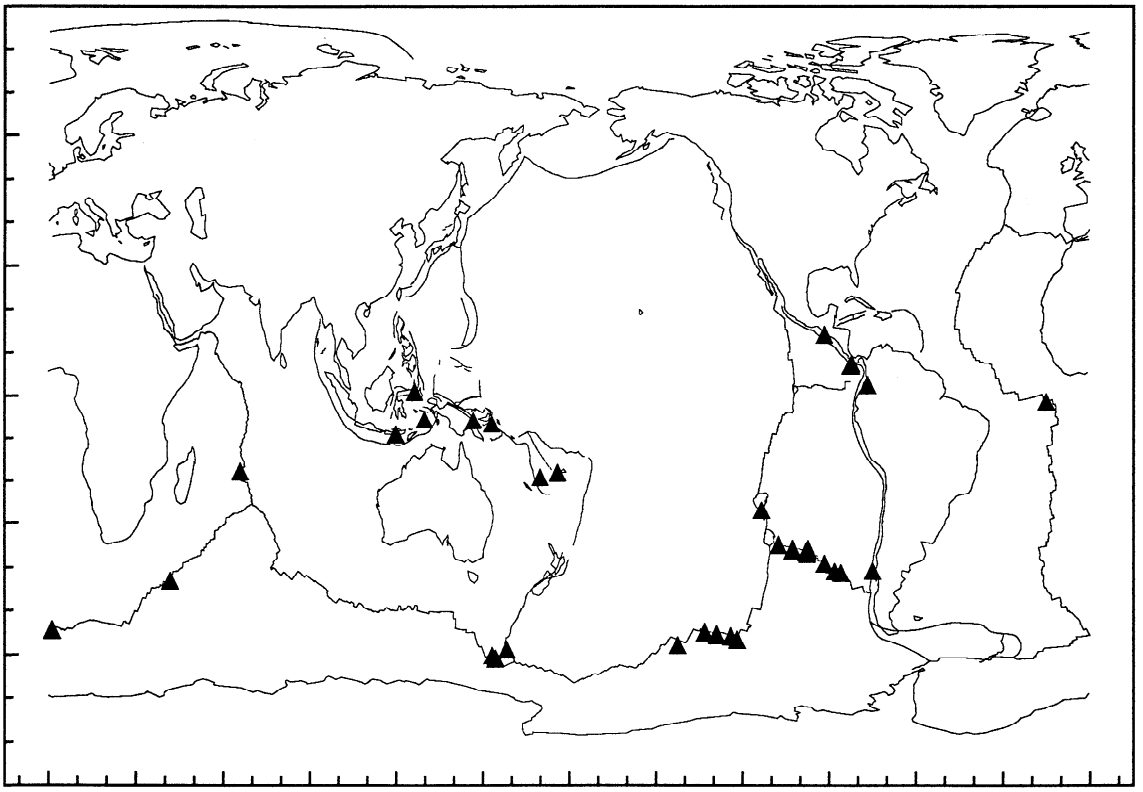

Figure 7. PDE locations for the 40 events with anomalously large $Q(t, \mathbf{q})$ values compared to their surface wave magnitudes (the events above the diagonal line in Figure 5*b*). Since the matched-filter results are obtained at longer periods than are used to determine the surface wave magnitudes, these are probable slow events which radiate less energy at high frequencies. Most of these earthquakes occurred on oceanic transform faults, consistent with previous studies which have shown that events on these faults tend to have slower-than-average ruptures. The Challenger and Eltanin fracture zones are particularly well represented (10 events and 6 events, respectively, some nearly colocated events do not appear separately on this plot).

In 32 cases, shown as the triangles in Figure 8, the technique correctly detected a previously unidentified earthquake. Figure 9 plots seismograms, in record section format, for three typical examples of these events, an earthquake on August 8, 1981, located near the Scott Fracture Zone (south of New Zealand), a December 3, 1982, event near the Tristan Da Cunha Fracture Zone (south Atlantic) and a September 7, 1988, event near the Eltanin Fracture Zone (south Pacific). $R1$, part of $R2$, and hints of body wave phases are clearly visible for these examples, which are high-pass filtered to above 300 s and plotted at the full 10-s sample rate. Table 1 lists peak $Q(t, \mathbf{q})$ values, times and locations for the 32 events; these locations are plotted in Figure 10. Note that the times and locations are those obtained directly from the matched-filter algorithm and are only accurate to within about 2 min and 5–10° (in most cases the signals are clear enough that much more accurate locations could be obtained). Most of the events appear to occur on transform faults in the southern Pacific and Atlantic Oceans. They are about $M_s = 4.5$ to 5.5, based on the scaling shown in Figure 5*b*.

These events are probably slower than average since they occur on oceanic transform faults, but they are not exceptionally slow since their Rayleigh waves do not

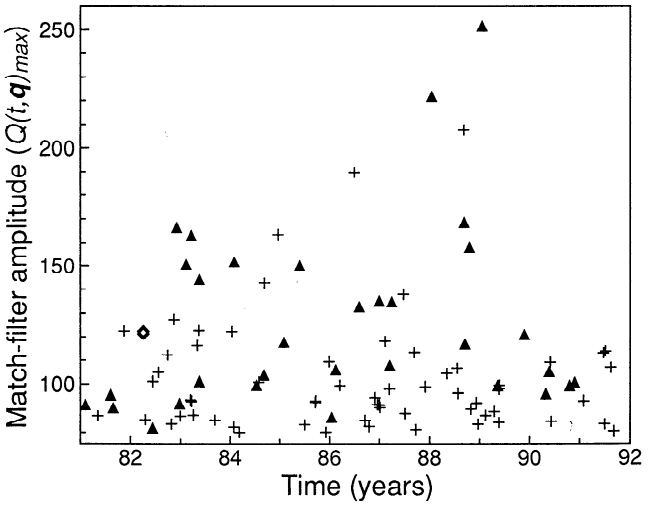

Figure 8. Peak $Q(t, \mathbf{q})$ values versus time for matched-filter detections which are not associated with cataloged events. The triangles show 32 previously unidentified earthquakes (~3 per year), while the pluses indicate false triggers (see text). Origin times and locations for these events are listed in Table 1. The diamonds in early 1982 show detections of two separate eruptions of the El Chichón volcano in southern Mexico.

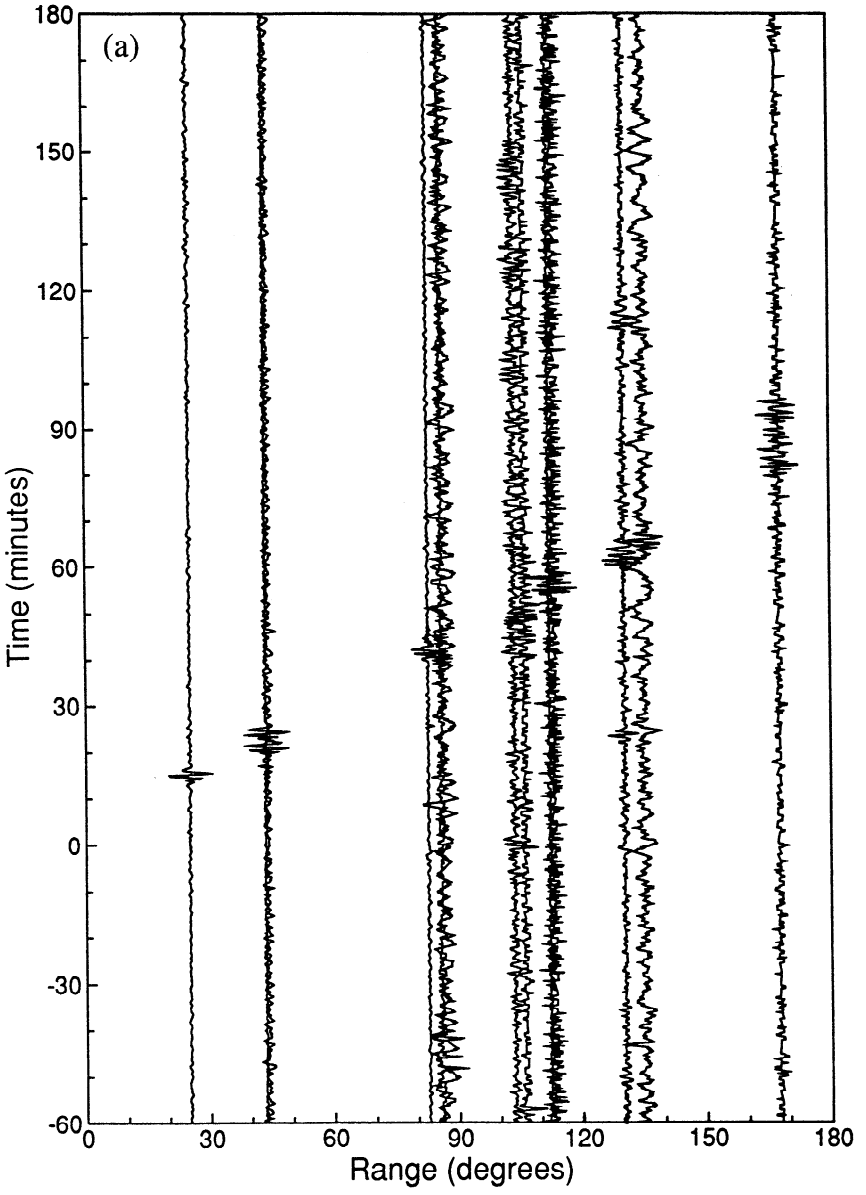

Figure 9. Examples of IDA seismograms plotted as a function of range for three previously unidentified earthquakes, including (*a*) an August 8, 1981, event near the Scott Fracture Zone, (*b*) a December 3, 1982, event near the Tristan Da Cunha Fracture Zone, and (*c*) a September 7, 1988, event near the Eltanin Fracture Zone.

appear to be of unusually low frequency (compared to other IDA recorded events). In order to confirm this, I obtained long-period records from the Global Digital Seismograph Network (GDSN) for several of these events and found clear Rayleigh wave signals on the higher-frequency GDSN records (1-s sample rate). Thus the reason that these earthquakes are not identified in global earthquake catalogs is probably primarily due to their locations in the southern oceans, in regions of known gaps in the coverage of the high-frequency networks. Most, if not all, of these events could have been identified and located using standard methods, provided proper attention were paid to the surface wave arrivals in long-period records. The usefulness of the matched-filter technique lies in its ability to quickly scan large data sets to flag likely events, without the necessity for any manual analysis of the data.

The matched-filter algorithm also triggered for three volcanic events, the April 4, 1982, El Chichón eruptions *B* (0135 UT) and *C* (1122 UT) [e.g., *Carey and Sigurdsson*, 1986], and the June 15, 1991, Mount Pinatubo eruption. The El Chichón Rayleigh waves, first noted by *Eissler* [1986], are anomalously low frequency in character, and neither event is listed in the PDE or ISC catalogs. The IDA records from the El Chichón *C* eruption were first examined by *Widmer and Zürn* [1992], who found that the source radiation lasted for about an hour and that the radiated energy was strongly peaked at 3.76 and 5.14 mHz. A similar phenomenon has been observed for Mount Pinatubo, which excited bichro-

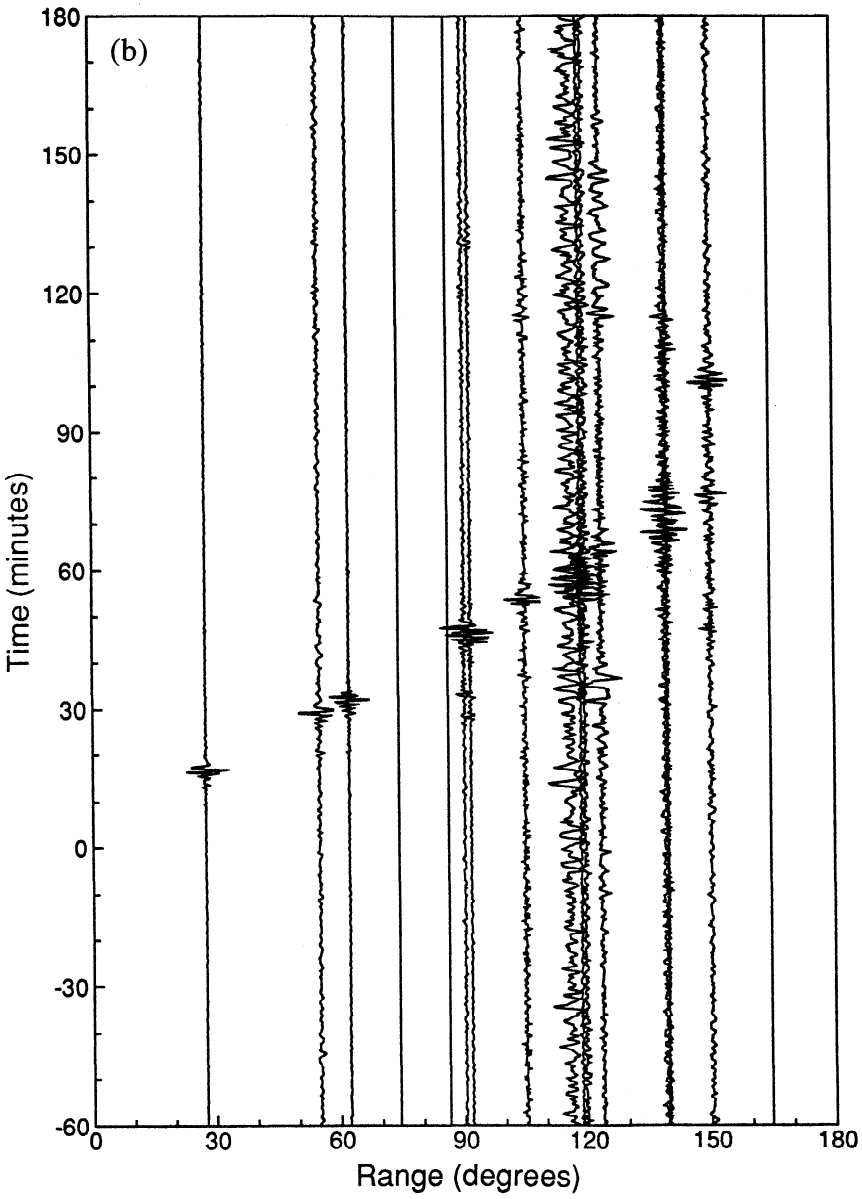

Figure 9. (continued)

matic Rayleigh waves for about 7 hours during its eruption [*Widmer and Zürn*, 1992; *Kanamori and Mori*, 1992]. However, in this case, the matched-filter algorithm did not trigger on the extended wave train but on a $m_b = 5.5$ earthquake (2037 UT) which occurred many hours into the eruption.

Discussion

This study is similar in many ways to the pioneering work of *Beroza and Jordan* [1990], who also applied an event detector to IDA seismograms. The Beroza and Jordan technique is based on examining the Earth's normal mode frequencies at discrete periods between about 200 and 700 s (1.4–5 mHz) to see if they are anomalously excited relative to the background noise. On the basis of statistical tests applied to a theoretical noise model, Beroza and Jordan computed the probability that observed excitation events were caused by far-field seismic signals. Because their study used only the mode amplitudes within 3-hour windows, discarding phase information, origin times are approximate and no locations could be assigned to observed excitation events. In contrast to the Beroza and Jordan technique, the matched-filter method described here is a time-domain approach applied to a higher-frequency part of the IDA records (60–300 s). An empirical approach is used, based on searching for correlations to a waveform pattern obtained by stacking known events. Origin times and locations are provided directly, facilitating the unambiguous association of events with those in earthquake catalogs and permitting the direct identification of false triggers. The method is robust with respect to noise glitches and clipped records, requiring no preediting of the data.

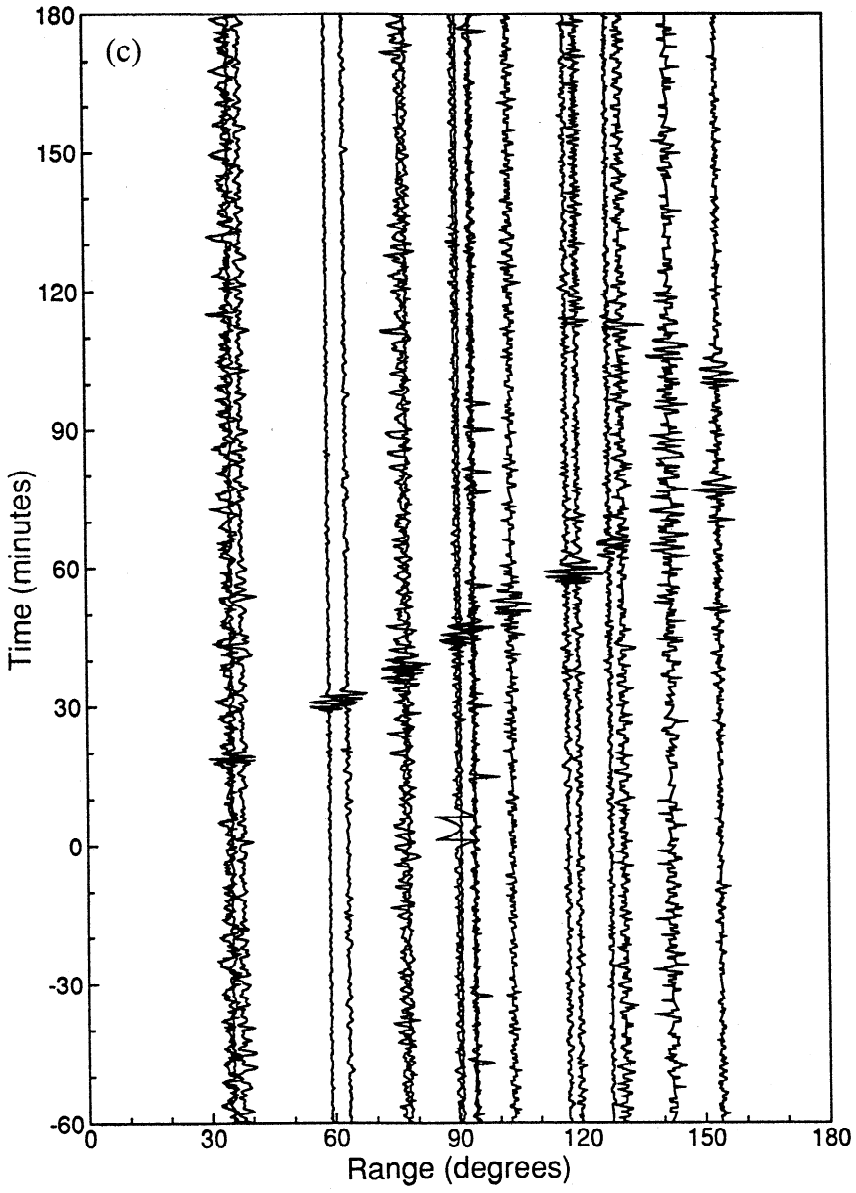

Figure 9. (continued)

Beroza and Jordan applied their technique to 2 years of IDA data (1978–1979) and found evidence for 27 episodes of mode excitation which could not be associated with significant earthquakes. They speculated that these might represent "silent" earthquakes which do not generate significant high-frequency energy. Application of the matched-filter technique to 11 years of IDA data at a later time period (1981-1991) than the Beroza and Jordan study identifies 32 previously undetected events, some of which may be slow, but none of which appear silent since they generate clear surface wave arrivals. However, the Beroza and Jordan anomalies were obtained at longer periods than those examined here and occurred at times without any obvious seismic signals on the IDA records. Such anomalies would probably not be detectable by the matched-filter algorithm described in this paper. A goal of future research will be to extend the matched-filter method to longer periods to search more directly for events invisible at higher frequencies.

Conclusions

These results demonstrate the practicality of an empirical matched-filter approach to global earthquake detection that searches for correlations between observed seismic wavefields and those generated by known events. Such a filter, applied to 11 years of very long period data of the IDA network, succeeds in detecting and locating a large fraction of PDE events of $M_s \geq 5.5$, despite the relatively sparse IDA station distribution ($\sim$15 stations per event). Earthquakes which appear anomalously strong in the IDA matched-filter output relative to their surface wave magnitudes are located mainly on oceanic transform faults and probably represent unusually slow ruptures. There are 32 events found during

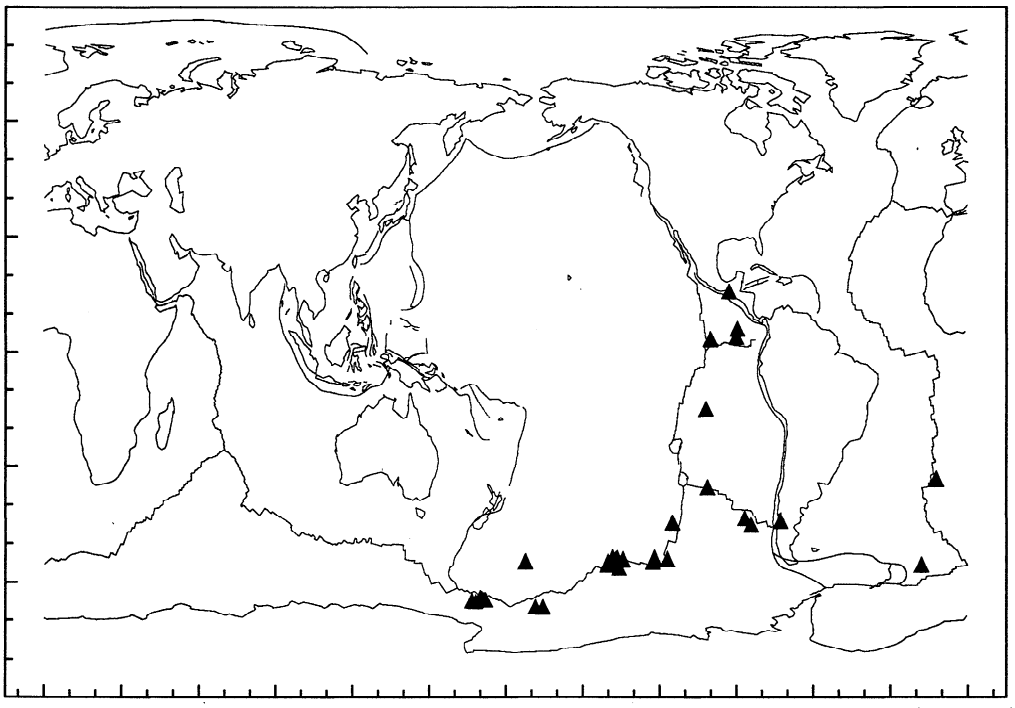

Figure 10. Locations of 32 previously unidentified events which were detected by the matched-filter algorithm. The locations were obtained directly from the matched-filter method, which assigns the events to one of 416 predefined sites, located about 10° apart. The probable location errors are up to about 10° for these events. For plotting purposes, an additional random location uncertainty of 2° is added to split apart the many colocated events.

Table 1. Previously Unidentified Events

Date	Time (UT)	Latitude	Longitude	Qmax
1981 Feb. 5	1947	-65.0	168.0	91.7
1981 Aug. 8	2117	-65.0	-168.0	95.8
1981 Aug. 27	0422	-65.0	168.0	90.5
1982 June 10	1960	-45.0	-72.0	81.9
1982 Dec. 3	0518	-35.0	-12.4	166.5
1982 Dec. 24	1018	-65.0	168.0	92.2
1983 Feb. 10	1258	-55.0	-137.1	150.7
1983 March 19	1120	-65.0	-168.0	163.2
1983 May 18	1104	-55.0	-17.1	101.1
1983 May 18	1824	-55.0	-137.1	144.1
1984 Jan. 30	0754	-55.0	-137.1	151.8
1984 July 11	2137	-15.0	-102.9	99.6
1984 Sept. 3	1660	-55.0	-137.1	103.8
1985 Jan. 30	1649	-45.0	-115.2	117.8
1985 May 26	1118	-55.0	-171.4	150.3
1986 Jan. 15	1130	-65.0	168.0	86.5
1986 Feb. 13	2328	-55.0	-120.0	106.0
1986 Aug. 6	2135	-55.0	-120.0	132.7
1986 Dec. 29	1911	-45.0	-86.4	135.1
1987 March 16	0650	5.0	-100.0	107.9
1987 March 30	1908	-55.0	-137.1	134.7
1988 Jan. 15	0224	-65.0	168.0	221.7
1988 Sept. 7	2302	-55.0	-137.1	168.5
1988 Sept. 14	1750	5.0	-90.0	117.1
1988 Oct. 17	1704	-45.0	-86.4	158.2
1989 Jan. 18	1921	-55.0	-137.1	251.6
1989 May 13	0748	-55.0	-137.1	99.4
1989 Nov. 23	1657	-55.0	-120.0	121.0
1990 April 27	1902	-65.0	168.0	96.2
1990 May 23	0505	-55.0	-137.1	105.2
1990 Oct. 18	1660	-35.0	-99.3	99.4
1990 Nov. 24	0820	-55.0	-137.1	100.6

the 11-year period that were not identified previously in the global earthquake catalogs. Although computationally intensive, the matched-filter technique has the advantages of simplicity and robustness. The method is very general and could readily be adapted to other data sets and realtime applications.

Acknowledgments. The Project IDA team deserves credit for the successful operation of the IDA network. Guy Masters and Ruedi Widmer helped make the data conveniently available to me. Paul Richards and Dave Simpson wrote helpful reviews. This research was partially supported by National Science Foundation grant EAR91-18309. The IDA instruments were purchased with funds provided by Cecil and Ida Green.

References

Agnew, D.C., J. Berger, W.E. Farrell, J.F. Gilbert, G. Masters, and D. Miller, Project IDA: A decade in review, *Eos Trans. AGU, 67*, 203–212, 1986.

Bache, T.C., S.R. Bratt, J. Wang, R.M. Fung, C. Kobryn, and J.W. Given, The intelligent monitoring system, *Bull. Seismol. Soc. Am., 80*, 1833–1851, 1990.

Beroza, G.C., and T.H. Jordan, Searching for slow and silent earthquakes using free oscillations, *J. Geophys. Res., 95*, 2485–2510, 1990.

Blandford, R.R., Seismic event discrimination, *Bull. Seismol. Soc. Am., 72*, S69–S87, 1982.

Blandford, R.R., and J.H. Goncz, Automatic association, in The VELA Program: A Twenty-Five Year Review of Basic Research, edited by A.U. Kerr, pp. 559–564, Executive Graphic Services, Arlington, Virginia, 1985.

Bratt, S.R., and T.C. Bache, Locating events with a sparse network of regional arrays, *Bull. Seismol. Soc. Am.*, *78*, 780–798, 1988.

Bratt, S.R., H.J. Swanger, R.J. Stead, F. Ryall, and T.C. Bache, Initial results from the intelligent monitoring system, *Bull. Seismol. Soc. Am.*, *80*, 1852–1873, 1990.

Carey, S., and H. Sigurdsson, The 1982 eruptions of El Chichon volcano, Mexico, 2, Observations and numerical modelling of tephra-fall distribution, *Bull. Volcanol.*, *48*, 127–141, 1986.

Cassidy, F., A. Christoffersson, E.S. Husebye, and B.O. Ruud, Robust and reliable techniques for epicentral location using time and slowness observations, *Bull. Seismol. Soc. Am.*, *80*, 140–149, 1990.

Dziewonski, A.M., and J.H. Woodhouse, An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981, *J. Geophys. Res.*, *88*, 3247–3271, 1983.

Eissler, H.K., Investigations of earthquakes and other seismic sources in regions of volcanism, Ph.D. dissertation, Calif. Inst. of Technol., Pasadena, 1986.

Hanks, T.C., and H. Kanamori, A moment magnitude scale, *J. Geophys. Res.*, *84*, 2348–2350, 1979.

Jordan, T.H., P. Ihmlé and P. Harabaglia, A global survey of the low-frequency characteristics of large earthquakes, *Eos Trans. AGU, 72 (17) Spring Meeting suppl.*, 201, 1991.

Joswig, J., Pattern recognition for earthquake detection, *Bull. Seismol. Soc. Am.*, *80*, 170–186, 1990.

Kanamori, H., and J. Mori, Harmonic excitation of mantle Rayleigh waves by the 1991 eruption of Mount Pinatubo, Phillipines, *Geophys. Res. Lett.*, *19*, 721–724, 1992.

Kanamori, H., and G.S. Stewart, Mode of strain release along the Gibbs Fracture Zone, Mid-Atlantic Ridge, *Phys. Earth Planet. Inter.*, *11*, 312–332, 1976.

Kushnir, A.F., V.M. Lapshin, V.I. Pinsky, and J. Fyen, Statistically optimal event detection using small array data, *Bull. Seismol. Soc. Am.*, *80*, 1934–1950, 1990.

Okal, E.A., and L.M. Stewart, Slow earthquakes along oceanic fracture zones: Evidence for asthenospheric flow away from hotspots?, *Earth Planet. Sci. Lett.*, *57*, 75–87, 1982.

Pelayo, A.M., and S. Stein, Variation in source parameters for oceanic ridge, transform and intraplate earthquakes, *Eos Trans. AGU, 72 (17) Spring Meeting suppl.*, 202, 1991.

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

Ringdal, F., E.S. Husebye, and J. Fyen, Earthquake detectability estimates for 478 globally distributed seismograph stations, *Phys. Earth Planet. Inter.*, *15*, P24–P32, 1977.

Roberts, R.G., A. Christoffersson, and F. Cassidy, Real time event detections, phase identification and source location estimation using single-station three-component seismic data, *Geophys. J. Int.*, *97*, 471–480, 1989.

Ruud, B.O., and E.S. Husebye, A new three-component detector and automatic single-station bulletin production, *Bull. Seismol. Soc. Am.*, *82*, 221–237, 1992.

Sereno, T.J., and S.R. Bratt, Seismic detection capability at NORESS and implications for the detection threshold of a regional network in the Soviet Union, *J. Geophys. Res.*, *94*, 10,397–10,414, 1989.

Shearer, P.M., Imaging global body wave phases by stacking long-period seismograms, *J. Geophys. Res.*, *96*, 20,353–20,364, 1991a.

Shearer, P.M., Constraints on upper mantle discontinuities from observations of long-period reflected and converted phases, *J. Geophys. Res.*, *96*, 18,147–18,182, 1991b.

Stein, S., and A. Pelayo, Seismological constraints of stress in the oceanic lithosphere, *Philos. Trans. R. Soc. London, Ser. A*, *337*, 53–72, 1991.

Stewart, L.M., and E.A. Okal, Seismicity and aseismic slip along the Eltanin fracture zone, *J. Geophys. Res.*, *88*, 10,495–10,507, 1983.

Tsvang, S.L., V.I. Pinsky, and E.S. Husebye, Enhanced seismic source discrimination using NORESS recordings from Eurasian events, *Geophys. J. Int.*, *112*, 1–14, 1993.

Widmer, R., and W. Zürn, Bichromatic excitation of long-period Rayleigh and air waves by the Mount Pinatubo and El Chichón volcanic eruptions, *Geophys. Res. Lett.*, *19*, 765–768, 1992.

P. M. Shearer, IGPP 0225, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0225.

(Received October 8, 1993; revised February 3, 1994; accepted February 16, 1994.)