Rapid estimation of rupture extent for large earthquakes: application to the 2004, M9 Sumatra-Andaman mega-thrust

Anthony Lomax

Anthony Lomax Scientific Software, Mouans-Sartoux, France

Submitted 14 January 2005, revised 21 March 2005, accepted 25 April 2005

Accepted for publication in Geophysical Research Letters.

Abstract

Rapid estimation of the rupture extent of large earthquakes is critical for tsunami warning and emergency response. The hypocenter of a distant earthquake is determined from the first seismic *P* waves within about 15 minutes after the event. However, methods for determining the size and rupture extent of very large earthquakes rely on long period recordings and aftershock locations, which are not available until several hours or more after the event. Here we introduce a method for rapid estimation of the rupture termination location for large earthquakes based on short period, *P*-wave recordings, available about 30 minutes after an event. The hypocenter and the rupture termination location provide estimates of the extent of rupture and the event size. Application to the 2004, M9 Sumatra-Andaman earthquake gives a rupture termination location near the Andaman Islands, about 1100 km north of the hypocenter, and a rupture duration of about 8 minutes.

Introduction

The 26 December 2004, M9 Sumatra-Andaman earthquake caused a tsunami that devastated coasts around the Eastern Indian Ocean within 3 hours. Tsunami hazard warning and emergency response for future large earthquakes would benefit greatly if knowledge of the extent of earthquake rupture were available within minutes after the event. Currently the size and extent of rupture of very large earthquakes are first estimated from moment tensor determinations based on long period, seismic surface-wave recordings [*Dziewonski et al.*, 1981; *Ekström*, 1994; *Kawakatsu*, 1995] and from examination of aftershock locations [e.g., *Scholtz*, 2002]; these estimates are not available until several hours or more after the event. Seismic *P* waves contain information on the earthquake rupture and are the earliest signal to arrive at distant recording stations. Within about 15 minutes after an event, the arrival times of the *initial P* wave are used routinely to locate the point of initiation of earthquake rupture, or *hypocenter*. The first available information about the termination of rupture is contained in the *last P*-wave energy radiated from the source. Here we introduce a simple, preliminary methodology to extract arrival times for this energy through analysis of the shape of the short period, *P*-wave signal. We use these arrival times to estimate the location of rupture termination in the same manner as for a hypocenter location. The required *P* wave recordings from global seismic stations are available about 20 to 30 minutes after an event. A few minutes later, this procedure provides an estimate of the rupture-termination location, and consequently the extent and duration of rupture for the earthquake. This information can aid in rapid assessment of tsunami hazard and damage distribution. We apply this methodology to the 2004 Sumatra-Andaman earthquake and other large earthquakes.

Method

We estimate of the position of rupture termination (R_{end}) using methods related to those of strong motion source studies [e.g., *Gusev and Pavlov*, 1991; *Cocco and Boatwright*, 1993; *Zeng et al.*, 1993]. We make three basic assumptions: 1) *P* waves radiated from the rupture, in particular at the last stage of rupture, contain higher frequencies than other wave types; 2) this signal can be isolated on the seismograms; 3) a meaningful time for the end of this signal can be determined. Observations and experience support the first two assumptions. For example, stacks of short period (< 2 s), vertical-component seismograms from large numbers of earthquakes [*Shearer*, 1999, Figure 4.18] shows that the direct *P* wave signal is dominant to a great-circle distance (GCD) of about 110°. And in the case of the 2004 Sumatra event, short period signals (~1 Hz) from a large aftershock (M7.5, 2004 Dec 26, 0421 UT) show little or no signal from later phases (e.g., *PcP*, *PP*, *S*) relative to the amplitude of the initial, direct *P* signal. However, in some cases the direct *S* wave or other phases can have overlapping high-frequency content with the *P* signal. The third assumption poses difficulties since the isolated *P* wave signal usually has an exponentially decaying coda caused by wave scattering that does not present a unique ending time for this signal.

We obtain arrival time estimates of the last *P*-wave energy (P_{end}) using vertical-component, 20 Hz, velocity seismograms and the following, preliminary procedure (Figure 1). 1) Convert the velocity seismograms to short-period using a narrow-band, Gaussian filter of the form $e^{-\alpha((f-f_{cent})/f)^2}$, where f is frequency, f_{cent} the filter center frequency, and α sets the filter width (here we use f_{cent} = 1.0 Hz and α = 10.0). 2) Convert the short-period seismograms to kinetic-energy density by squaring each of the velocity values. 3) Integrate the kinetic-energy density time-series over time to give a monotonically increasing, cumulative-energy time-series. 4) Determine a P_{end} arrival as the time where the cumulative-energy time-series reaches a specified fraction F of its maximum value (here we use F = 0.9). 5) Obtain an arrival time uncertainty from the time difference between the points where the cumulative-energy time-series reaches the fractions $F+\sigma$ and $F-\sigma$ of its maximum value (here we set a relatively large σ = 0.05). We exclude from this analysis seismograms that are clipped, truncated or otherwise corrupted and Gaussian-filtered seismograms that show *S* wave energy or strong background noise.

With the obtained P_{end} arrival-times, we locate R_{end} using the earthquake location program NonLinLoc [*Lomax et al.*, 2000; *Lomax et al.*, 2001; *Lomax*, 2005; http://www.alomax.net/nlloc]; (NLL). NLL combines the probabilistic methodology of *Tarantola and Valette* [1982] with efficient global sampling to obtain an estimate of the posterior probability density function (*pdf*) in 3D space for seismic event location. To build the *pdf* NLL uses a likelihood function based on the equal differential-time formulation (EDT) of *Font et al.* [2004], a generalization of the master-station method [*Zhou*, 1994] and the "method of hyperbolas" cited by *Milne* [1886]. NLL with EDT is highly robust in the presence of outliers in the data [*Lomax*, 2005]. This robustness is critical for the present problem, since our P_{end} arrival-time estimates are subject to large uncertainty and bias; many of the locations examined below do not converge with a standard L2 likelihood function. The EDT location determination is independent of origin time and reduces to a 3D search over latitude, longitude and depth. NLL performs this search with an efficient, cascading grid-search, importance-sampling method called Oct-tree [*Lomax and Curtis*, 2001; http://www.alomax.net/nlloc/octtree]. In this study, NLL searches over the whole earth from the surface to 600 km depth, and uses the ak135 model [*Kennett et al.*, 1995] to determine *P*-wave travel-times.

Application to the 2004 Sumatra-Andaman earthquake

We first apply our methodology to the 2004 Sumatra-Andaman mega-thrust. We obtain from the IRIS Data Management Center a set of BHZ component recordings at stations up to 90° GCD from the event; this data set is representative of that which would be available 30 minutes or less after a large earthquake. Applying the procedure outlined above, excluding data as necessary, we obtain a

set of 92 P_{end} arrival-times, with a mean uncertainty of about ±40 s, from stations at 15° to 89° GCD. Figure 1 illustrates the signal processing and arrival-time determination for station GRFO in Germany at 84° GCD. For this distant station most of the short-period signal arrives before the first *S* arrival, confirming that this signal contains direct *P* energy and little or no *S* energy.

Figure 2 shows the *pdf* for R_{end} given by NLL using the 92 P_{end} arrival-times. The *pdf* has near-constant density from the surface to 600 km depth indicating a lack of depth constraint, probably due to the predominance of steeply down-going rays from the source region to the available stations (i.e., there are no nearby stations) combined with the large uncertainty of the P_{end} arrival-time estimates. Thus in the following we only consider the horizontal position of R_{end} locations. The *pdf* indicates an uncertainty of about 200 km east-west and about 300 km north-south for the location of R_{end}. The *pdf* is centered near 13°N, 93°E in the Andaman Islands region, about 1100 km north of the hypocenter, and near the northernmost aftershocks. A line between the hypocenter and the center of the *pdf* provides an estimate of the rupture zone and its minimum length, and implies that rupture for this event propagated to the NNW. The NLL "origin-time" for shallow R_{end} locations indicates rupture duration of about 8 minutes and thus an average rupture velocity of about 2.3 km/s. These results agree with those of Vallée [2005] who is able to image the complete rupture with an established waveform procedure. Our results, available about 30 minutes after initiation of rupture, would aid monitoring organizations in rapidly understanding the size and geometry of an earthquake and in initiating tsunami early warning and modeling procedures.

To understand better these results, we examine the seismograms for two stations situated at the same distance from the hypocenter but at opposing positions along a line passing through the likely rupture zone (Figure 3). The first *P* signal arrives at about the same time at each station, consistent with both being at the same distance from the hypocenter. In contrast, the smaller duration of the strong, short period signal at OBN (~7 minutes) relative to that at CASY (~9 minutes) is consistent with rupture propagation from the hypocenter towards the NNW, and with CASY being further than OBN from R_{end}, as determined here.

Application to other large earthquakes

We next apply our methodology to other large earthquakes. For all events, we obtain IRIS data for stations at up to 90° GCD to form data sets representative of those available just after a large earthquake.

The results for a large subduction earthquake, 23 June 2001, M8.4, Peru, (Figures 4) shows a *pdf* for R_{end} centered over 400 km east-southeast of the hypocenter near the southeastern limit of the first week's aftershocks, along the down-dip direction of the subduction zone. The *pdf* overlaps the southeast end of the dense, coastal aftershock zone but not the mainshock hypocenter, implying rupture to the southeast, in agreement with other studies [e.g., *Giovanni et al.*, 2002]. In a warning scenario, our rapid Peru results would give early information that the event may be very large with rupture extending hundreds of kilometers to the southeast of the hypocenter.

The NLL location for a large transform event, 3 November 2003, M7.9, Denali, Alaska, produces a *pdf* for R_{end} centered a few hundred kilometers to the WNW of the hypocenter, opposite to the trend of aftershocks. This solution does not agree with the observed and inferred rupture [e.g., *Fuis and Wald*, 2003], but quality indicators in the NLL location such as residuals and RMS error clearly show that the solution is unreliable. Thus for this event one or more of our basic assumptions are incorrect or there is a weakness in our methodology. Close examination of the short-period seismograms shows at many stations to the NW a large, initial *P*-wave signal with duration much less than that expected for this event. This signal may be high-amplitude radiation from the identified initial thrust rupture preceding longer, strike-slip faulting [*Fuis and Wald*, 2003]. Other stations exhibit a *P* wave signal compatible with the expected rupture duration. Thus, the P_{end} arrival-times derived from our simple, cumulative-energy methodology at different stations may reflect different parts of the rupture and thus lead to a biased location.

The results for other large thrust events (e.g., 2003, M8, Tokachi-Oki, Hokkaido, Japan; 1999, M7.6 Chi-Chi, Taiwan) are reasonable, but indicate instability in our preliminary procedure for estimating P_{end} arrival-times. The results for several transform events (e.g., 1999, M7.6, Izmit, Turkey) are poor and indicate shortcomings in our simple, preliminary procedure.

Discussion

We have examined a method for rapid estimation of the rupture extent of very large earthquakes based on the form of short period, *P*-wave recordings. The method produces reasonable results for the 2004 Sumatra-Andaman mega-thrust and other large subduction earthquakes, but not for 2003 Alaska and other smaller transform events.

There are several possible reasons that our methodology works well for large subduction zone earthquakes. Firstly, these events may radiate strong signal around 1 Hz at all stages of rupture. Secondly, the recorded short period, *P*-wave signal for these events may typically be large relative to the short period *S*-waves. Thirdly, the very large rupture extent of these earthquakes leads to differential times at each station between the *P* and P_{end} arrival-times and variation of these differential times between stations much greater than the (large) uncertainty in our P_{end} arrival time estimations (e.g., Figure 3). In contrast, the method may be more difficult to apply to smaller events and to transform earthquakes. Rupture during transform earthquakes may radiate high frequency energy only from one or a few asperity patches, and may be smooth elsewhere. In addition, the average orientation of the *P*- and *S*-wave radiation patterns to distant stations may typically be less favorable for transform events than for thrust events for preserving short period *P* signal and for suppressing short period *S* signal.

It is likely that further study, development, testing and automation of this methodology will improve its robustness and performance, and may render it useful for smaller events and for transform and other event types. In particular, results for 2003 Alaska and other events indicate that correct estimation of a consistent set of P_{end} arrival-times requires more sophisticated processing and analysis of the short-period seismograms than that provided by the preliminary methodology outlined above. In addition, cases such as bi-lateral rupture, rupture termination along an extended front, and a large, independent event nearby in space or time will cause difficulty. However, this study shows that our approach gives promising results for large subduction earthquakes and could provide valuable information for rapid assessment of tsunami hazard and damage distribution.

Acknowledgements

I thank Stephan Husen and two reviewers for constructive comments. The SISMOS project, INGV, Rome (http://sismos.ingv.it) funded development of NLL for teleseismic location. The Java programs SeisGram2K and Seismicity Viewer (http://www.alomax.net/software) were used for processing and figures. The IRIS DMC provided access to waveforms used in this study.

References

Cocco, M., and J. Boatwright (1993), The envelopes of acceleration time histories, *J. Geophys. Res., 83*, 1095-1114.

Coffin, M.F., L.M. Gahagan, and L.A. Lawver (1998), Present-day Plate Boundary Digital Data Compilation. *Tech. Rep. 174,* Univ. of Tex. at Austin, Austin.

Dziewonski, A., 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*, 2825-2852.

Ekström, G. (1994) Rapid earthquake analysis utilizes the internet: *Computers in Physics*, *8*, 632-

638.

Font, Y., H. Kao, S. Lallemand, C.-S. Liu, and L.-Y. Chiao (2004), Hypocentral determination offshore Eastern Taiwan using the Maximum Intersection method. *Geophys. J. Int.*, *158*, 655-675.

Fuis, G.S., and L.A. Wald (2003), Rupture in South-Central Alaska—The Denali Fault Earthquake of 2002, *Fact Sheet 014-03*, U.S.G.S.

Giovanni, M.K., S.L. Beck, and L. Wagner (2002), The June 23, 2001 Peru earthquake and the southern Peru subduction zone, *Geophys. Res. Lett.*, *29*, 2018-2021.

Gusev, A. A., and V. M. Pavlov (1991), Deconvolution of squared velocity waveform as applied to study of a noncoherent short-period radiator in the earthquake source, *Pure Appl. Geophys. 136*, 235–244.

Kawakatsu, H., (1995), Automated near-realtime CMT inversion, *Geophys. Res. Lett.*, *22*, 2569-2572.

Kennett, B.L.N., E.R. Engdahl, and Buland R. (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int., 122*, 108-124

Lomax, A. (2005), A Reanalysis of the Hypocentral Location and Related Observations for the Great 1906 California Earthquake, *Bull. Seism. Soc. Am.*, (accepted, in press).

Lomax, A., J. Virieux, P. Volant, and C. Berge, (2000), Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations, in *Advances in Seismic Event Location*, C.H. Thurber and N. Rabinowitz (eds.), Kluwer, Amsterdam, 101-134.

Lomax, A., A. Curtis (2001), Fast, probabilistic earthquake location in 3D models using oct-tree importance sampling, *Geopys. Res. Abstr., 3*, 955.

Lomax, A., A. Zollo, P. Capuano, and J. Virieux (2001), Precise, absolute earthquake location under Somma-Vesuvius volcano using a new 3D velocity model, *Geophys. J. Int., 146*, 313-331.

Milne, J. (1886), *Earthquakes and Other Earth Movements*, Appelton, New York, 361pp.

Shearer, P. (1999), *Introduction to Seismology*, 260 pp, Cambridge Univ. Press, New York.

Scholtz, C.H. (2002), *The Mechanics of Earthquakes and Faulting*, 471pp, Cambridge Univ. Press, New York.

Tarantola, A. and B. Valette (1982), Inverse problems = quest for information, *J. Geophys., 50*, 159-170.

Vallée, M., (2005), Rupture extent and duration of the Mw=9. Sumatra earthquake, *Geophys. Res. Abs.*, 7, EGU General Assembly, Abstract EGU05-A-09675.

Zeng, Y., K. Aki, and T.-L. Teng (1993), Mapping of the high-frequency source radiation for the Loma Prieta earthquake, California, , *J. Geophys., Res., 98*, 11,981-11,993.

Zhou, H. (1994), Rapid 3-D hypocentral determination using a master station method, *J. Geophys., Res., 99*, 15439-15455.

Anthony Lomax

Anthony Lomax Scientific Software, Mouans-Sartoux, France; anthony@alomax.net, www.alomax.net

Figure Captions

Figure 1

Processing steps for estimating the P_{end} arrival time for the 2004 Sumatra-Andaman earthquake at station GRFO. Upper trace: raw, velocity seismogram; Upper middle: 1 Hz Gaussian filtered seismogram; Lower middle: kinetic-energy density time-series; Lower: cumulative-energy time-series. *P* and *S* indicate the ak135 predicted arrival times for the first *P* and *S* waves from the hypocenter. *Cumul_90* indicates the estimated P_{end} arrival-time and its uncertainty.

Figure 2

Map showing results for the 2004 Sumatra-Andaman earthquake. A density cloud (grey cluster of points) indicates the *pdf* for the termination of rupture location; this location is more likely to be in the denser regions of this *pdf* cloud. Also shown are the mainshock hypocenter (grey star), aftershocks located by NEIC within one week of the mainshock (black dots), and principal plate boundaries (gray lines, [*Coffin et al.,* 1998]).

Figure 3

Signal from the 2004 Sumatra-Andaman earthquake at two stations at 70° GCD from the hypocenter at opposing positions along the line of rupture: OBN in Russia to the NNW at 328° azimuth and CASY in Antarctica to the SSE at 173° azimuth. Upper two traces: raw, velocity seismograms; Lower 2 traces: 1-Hz Gaussian-filtered seismograms. Arrival labels as in Figure 1.

Figure 4

Map showing results for the 2001 Peru earthquake. Map elements as in Figure 2.

Figure 1

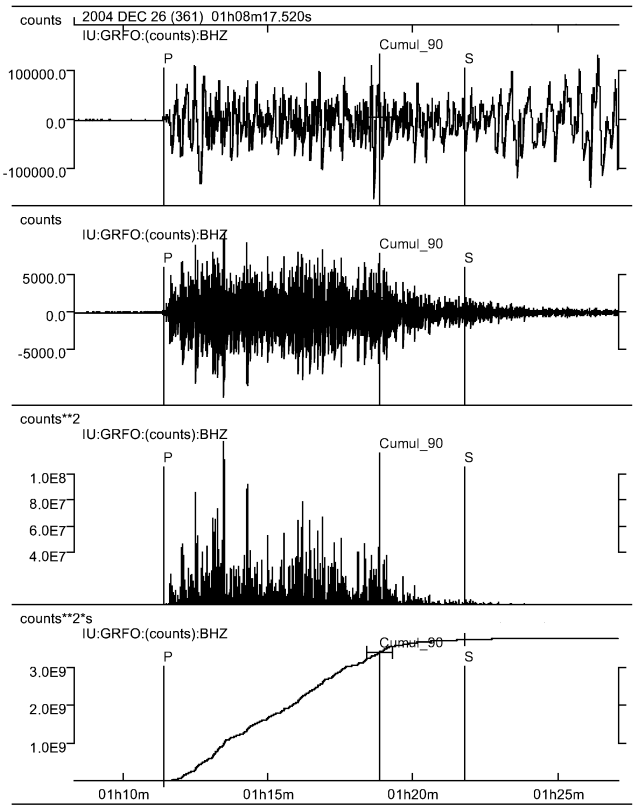
counts
2004 DEC 26 (361) 01h08m17.520s
IU:GRFO:(counts):BHZ
P
Cumul_90
S
100000.0
0.0
-100000.0
counts
IU:GRFO:(counts):BHZ
P
Cumul_90
S
5000.0
0.0
-5000.0
counts**2
IU:GRFO:(counts):BHZ
P
Cumul_90
S
1.0E8
8.0E7
6.0E7
4.0E7
counts**2*s
IU:GRFO:(counts):BHZ
P
Cumul_90
S
3.0E9
2.0E9
1.0E9
01h10m
01h15m
01h20m
01h25m

Figure 2

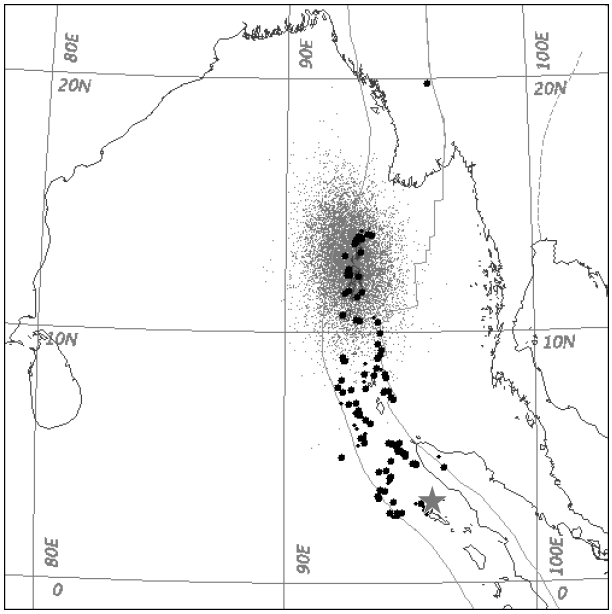

Figure 3

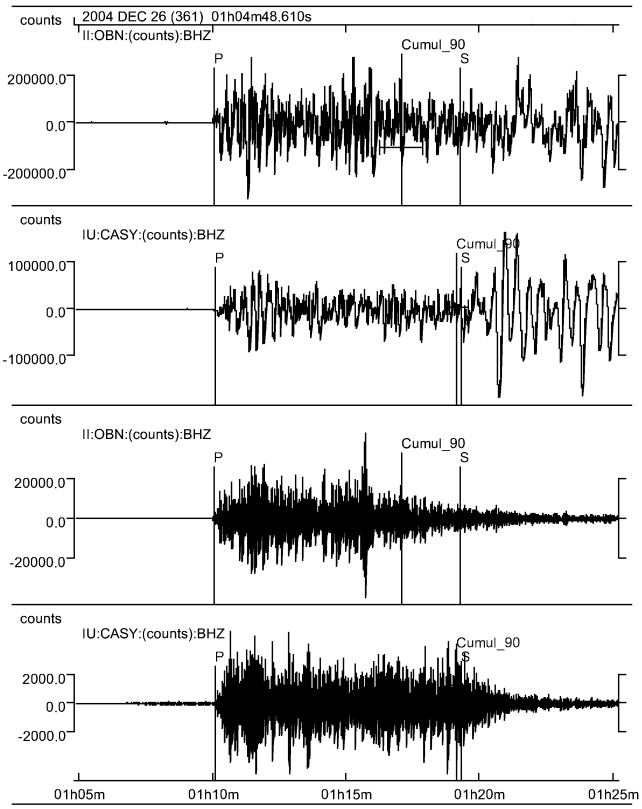
counts
2004 DEC 26 (361) 01h04m48.610s
II:OBN:(counts):BHZ
P
Cumul_90
S
200000.0
0.0
-200000.0
counts
IU:CASY:(counts):BHZ
P
Cumul_90
S
100000.0
0.0
-100000.0
counts
II:OBN:(counts):BHZ
P
Cumul_90
S
20000.0
0.0
-20000.0
counts
IU:CASY:(counts):BHZ
P
Cumul_90
S
2000.0
0.0
-2000.0
01h05m
01h10m
01h15m
01h20m
01h25m

Figure 4