THE 2004 SUMATRA EARTHQUAKE AND INDIAN OCEAN TSUNAMI: WHAT HAPPENED AND WHY

Seth Stein and Emile A. Okal

Dept of Geological Sciences, Northwestern University, Evanston Illinois 60208, seth@earth.northwestern.edu

On December 26, 2004 the world saw yet again how strains built up over hundreds of years by slow and almost imperceptible motions of tectonic plates are released with devastating effect. A giant earthquake beneath the Indonesian island of Sumatra generated a massive sea wave that crossed the Indian Ocean in a few hours, wreaking destruction along seacoasts and causing at least 300,000 deaths.

Geological Cause

The geologic causes of this event can be traced back over 120 million years ago, when the southern supercontinent of Gondwanaland broke up. The subcontinent of India separated from Antarctica and started its steady motion northward. 50 million years ago it collided with Asia, raising the Himalayas and forming the Tibetan plateau. The plate collision continues today as the Indian plate moves northward, forcing pieces of China and southeast Asia eastward.

Part of the plate boundary extends along the trench on the west coast of Sumatra. Here, an oceanic part of the Indian plate subducts beneath the Burma plate (Figure 1). The Burma plate is a small sliver or microplate between the Indian plate and the Sunda plate that contains much of southeast Asia. The east-dipping Indian plate can be identified by earthquakes that occur within it, down to a depth of about 300 km. However, most of the time little seems to be happening along the great thrust fault, sometimes called a mega-thrust fault that forms the plate boundary interface.

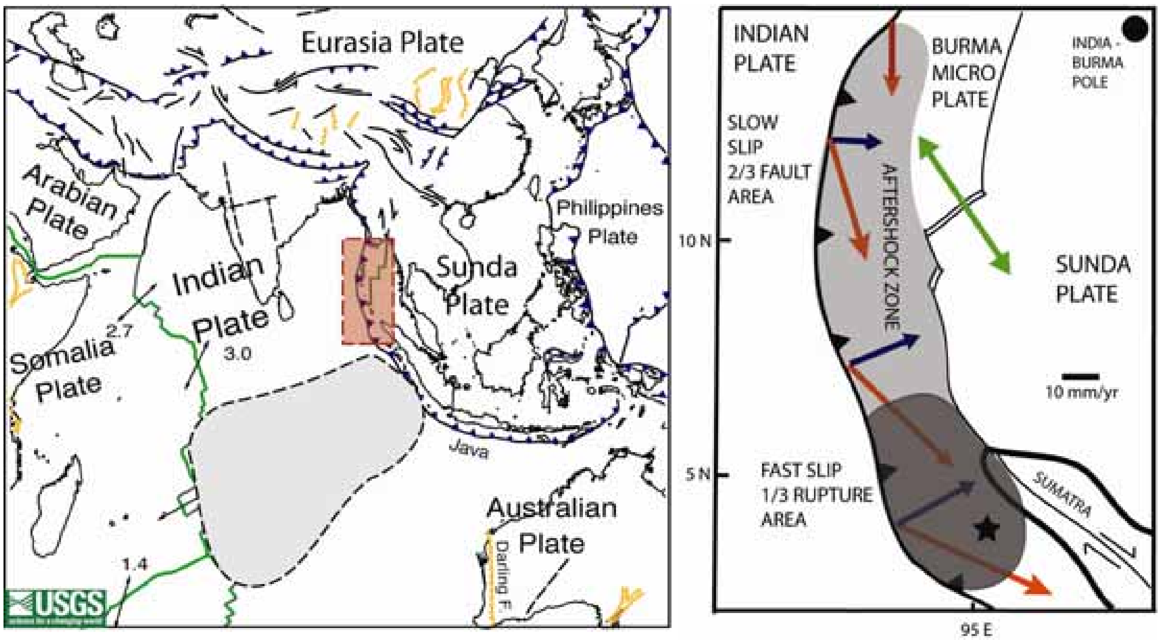

Fig. 1a. Approximate plate tectonic boundaries in the region of the Sumatra earthquake. Dashed region is broad boundary zone between India and Australia. The box shows the area of the map in Figure 1b. Fig. 1b. Schematic illustration of the regional tectonics and slip process in the earthquake. Studies based on high frequency seismic waves find that fast slip was concentrated on the southern part of the aftershock zone (dark grey) whereas the normal mode study (discussed below) shows a much larger possible area of slow slip (light grey). Star denotes epicenter where rupture started. The earthquake resulted from the Indian plate subducting beneath the Burma microplate due to motions about the rotation pole. Total (red arrows) and orthogonal (blue arrows) convergence between the plates is shown.

In reality, a lot is going on. Every year, about 20 mm of convergence occurs between the Indian and Burma plates. However, the mega-thrust fault is locked, so strain builds up on it (Figure 2). Eventually the accumulated strain exceeds the frictional strength of the fault, and it slips in a great earthquake like December's.

Such plate boundary thrust fault earthquakes can be very large – by far the largest that occur.

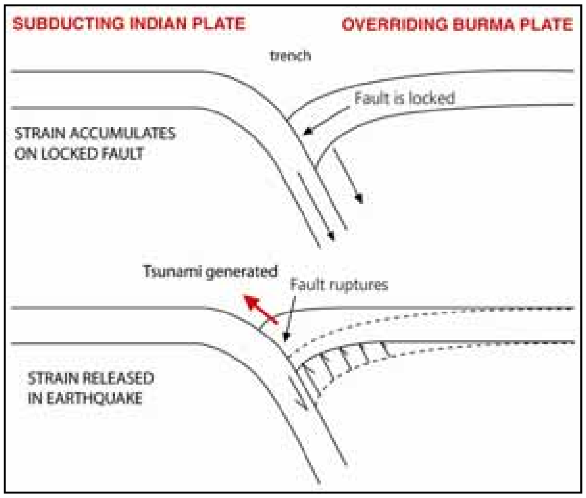

Fig. 2. Illustration of the cycle of strain accumulation and release that causes great thrust fault earthquakes at a subduction zone.

A huge area of the plate interface slips, generating seismic waves that can do great damage near the earthquake. Moreover, because this typically occurs at an underwater trench, the overriding plate that had been dragged down since the last earthquake rebounds and displaces a great volume of water, causing a tsunami that can have devastating effects far away.

Measuring Earthquake Size

The huge size of this earthquake has consequences for the fault rupture process and generation of the tsunami. To understand this issue requires understanding the concept of earthquake magnitude, a measure of earthquake size based on the amplitude of the resulting waves recorded on a seismogram. The earliest magnitude scale, introduced by Charles Richter in 1935 for Southern California earthquakes, is the local or "Richter" magnitude. This scale has been replaced by other magnitude scales that use seismic waves of different periods. These give more information, because an earthquake radiates different amounts of seismic energy at different periods.

To see why different measurements yield different magnitudes, consider the spectrum of the earthquake source, or how much energy is radiated at different periods. Figure 3 shows the logarithm of amplitude of the radiated waves versus the logarithm of the wave frequency (1 over the period). Ideally the plot is flat at low frequency (long period) and then decays for frequencies above (periods shorter than) “corner” frequencies proportional to 1 over the times needed for the rupture to propagate along the length of the fault and for slip to be completed at a point on the rupture. The larger the earthquake, the more the corner frequencies move to the left.

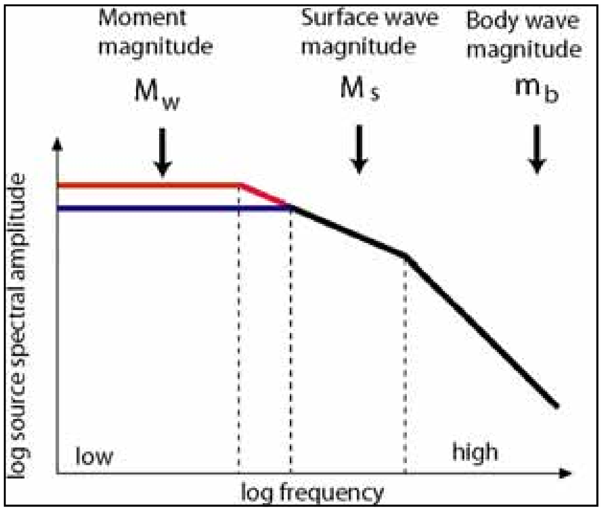

Fig. 3. Illustration of earthquake spectra showing corner frequencies (dashed vertical lines) and different magnitude determinations. The earthquake whose spectrum is shown in red has larger moment magnitude than the one with spectrum shown in blue, even though they have the same surface and body wave magnitudes, as shown by the black part of the spectra that are the same for both earthquakes.

Typically three different magnitudes are used, each of which measures the seismic energy radiated at a different period. The body wave magnitude mb is determined from the amplitude of waves that travel through the earth’s interior, with a period of 1 second. Similarly, the surface wave magnitude Ms is determined from the amplitude of waves that travel along the earth’s surface, with a period of 20 seconds. A problem with both these magnitudes is that they saturate or remain constant once earthquakes exceed a certain size. This happens because the added energy release in the very large earthquakes is all at longer periods than are measured by the 20 sec period surface waves. No matter how big an earthquake is, its body and surface wave magnitudes do not get above about 6.5 and 8.4, respectively. Hence for very large earthquakes these magnitude measurements underestimate the earthquake’s size. This issue is crucial for tsunami warning, as we will see.

To surmount this difficulty, we use the seismic moment that can be calculated by measuring the energy in the longest periods of the seismogram. The seismic moment also relates directly to the physical properties of the fault, so the moment can be determined either from seismograms or from the fault dimensions. In terms of the fault dimensions, the seismic moment is found by

Mo = [fault rigidity] X [fault area] X [fault slip]

The rigidity is the strength of the fault and is an approximate value determined from lab experiments. The moment magnitude Mw is calculated from the seismic moment using the relation Mw = (log Mo / 1.5) - 10.73. The constants in the equation have been chosen so that the moment magnitude scale correlates with the other magnitudes when they do not saturate.

Size of the Sumatra Earthquake from the Earth's Normal Modes

The Sumatra earthquake was a gigantic event. The aftershock zone extended 1200 km northward along the trench. Studies using body waves show that rupture started at the epicenter at the south end of this zone and propagated northward, with most of the rapid slip on the southern third of the rupture. Initial estimates based on surface waves with periods less than 300 s found a seismic moment of 4 x 1029 dyn-cm, corresponding to Mw = 9.0.

Additional insight into the size of the event comes from the earth's longest period normal modes. These are vibrations in which the earth rings like a bell (or more precisely rattles like a garbage can at many frequencies) for days and even weeks after a gigantic earthquake. Analysis of long seismograms shows distinct energy peaks whose height reflects the earthquake's seismic moment. The modes are standing seismic waves on a spherical earth analogous to standing waves on a string that add up to form traveling waves. The longest period modes occur in groups or multiplets consisting of singlets or peaks that are split - have distinct periods or frequencies - because the standing waves are affected by the rotation and shape of the earth. Seismic waves traveling in the direction of the rotation travel faster than those going the other way and the effect varies with latitude since a piece of the earth at the equator is traveling faster than a piece near the poles. In addition, waves traveling across the poles travel a shorter distance than ones traveling around the equator because of the shape of the earth.

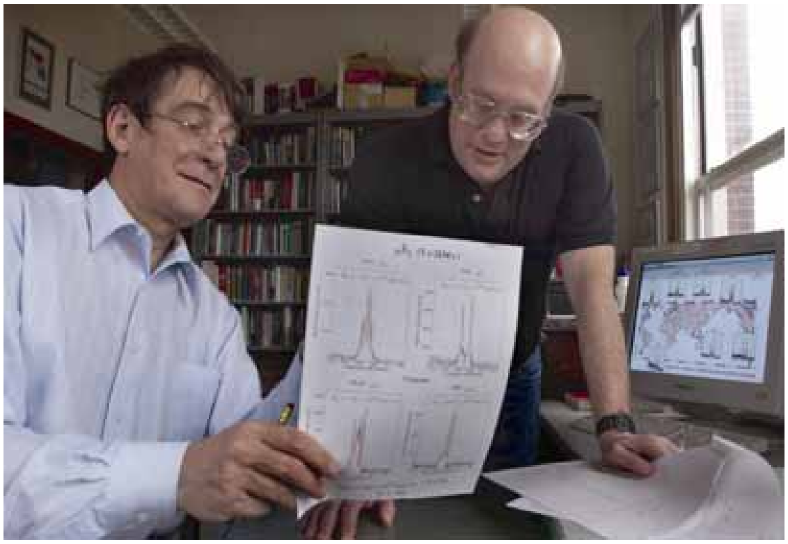

Fig. 4. The authors discussing data from the Sumatra earthquake.

The Sumatra earthquake excited the earth's normal modes beautifully. We analyzed them using techniques we developed with Robert Geller (now at the University of Tokyo) as graduate students almost 30 years ago. However, because such gigantic earthquakes are rare, these methods had been essentially unused until records of the Sumatra earthquake became available on modern digital seismometers of the Global Seismographic Network operated by IRIS. Hence immediately after the earthquake, we exhumed computer programs (some so old that they were originally on punch cards) and set to work (Figure 4). The computer programs calculate numerical seismograms, based on a model of how an earthquake will generate normal modes in the earth. Comparing the seismograms from seismic stations around the world and the modeling results are done in terms of the relative energy at different frequencies, called the frequency spectra (Figure 5).

Matching the amplitudes of the peaks shows that the earthquake had seismic moment of 1 x 1030 dyn-cm, or moment magnitude Mw = 9.3, approximately 2.5 times larger than shown by the

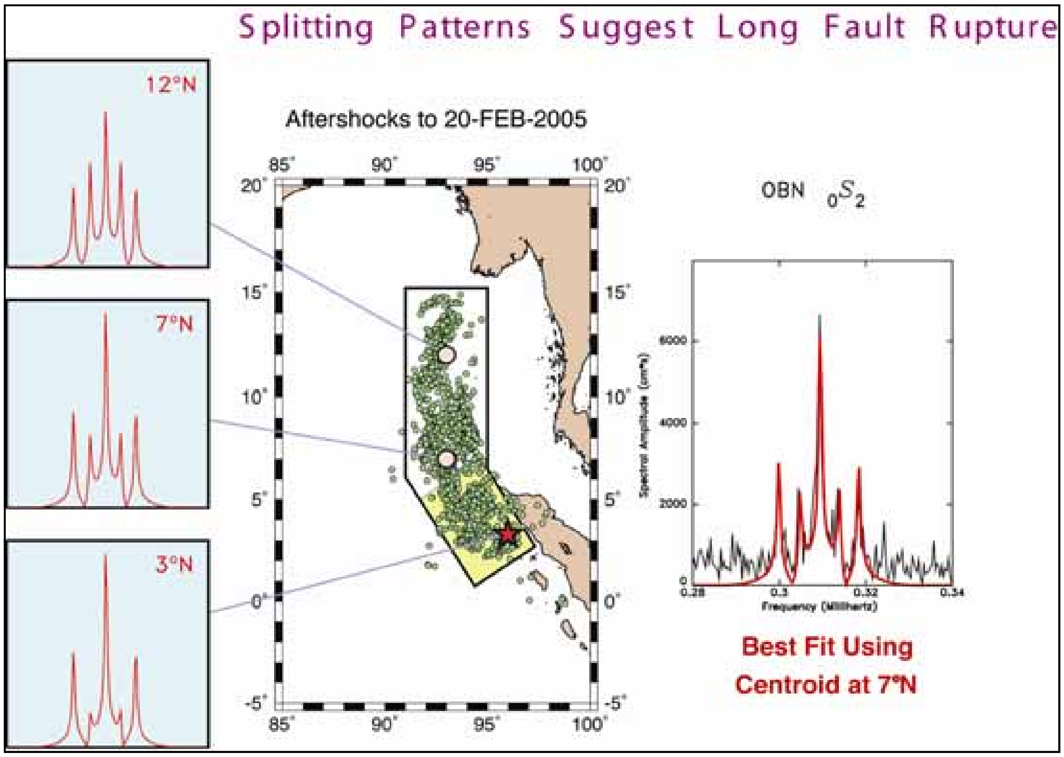

Fig. 5. Comparison of data (black) and model (red) at different seismic stations for the earth's longest period mode, which has a period of 52 minutes. Note the similarity of the peaks for stations with similar latitude (compare BFO and YSS) and the different shapes of the peaks for stations to the north and south. The pattern of the peaks depends only on the latitude of the seismometer because it reflects the earth's rotation and ellipticity, which are symmetric about the North Pole.

surface waves. This difference arises because the earthquake is so large that even the 300 second surface waves used in the initial Mw calculation did not record the very long period energy.

This larger magnitude likely reflects slow slip along the entire rupture zone suggested by aftershocks. The larger moment as calculated from the seismograms can be fit to the moment calculated using the fault dimension equation by 11 m of slip on a fault 1200 km long and 200 km wide (down-dip dimension). A larger rupture area is consistent with the fact that relative amplitudes of modes are better fit by a source with average position (known as the centroid) at 7°N than by one at the epicenter (Figure 6). Thus while the epicenter of the earthquake as determined from high frequency waves is at the south end of the fault rupture (the star in Figure 6), the centroid indicated by the normal modes is the center of the aftershock zone.

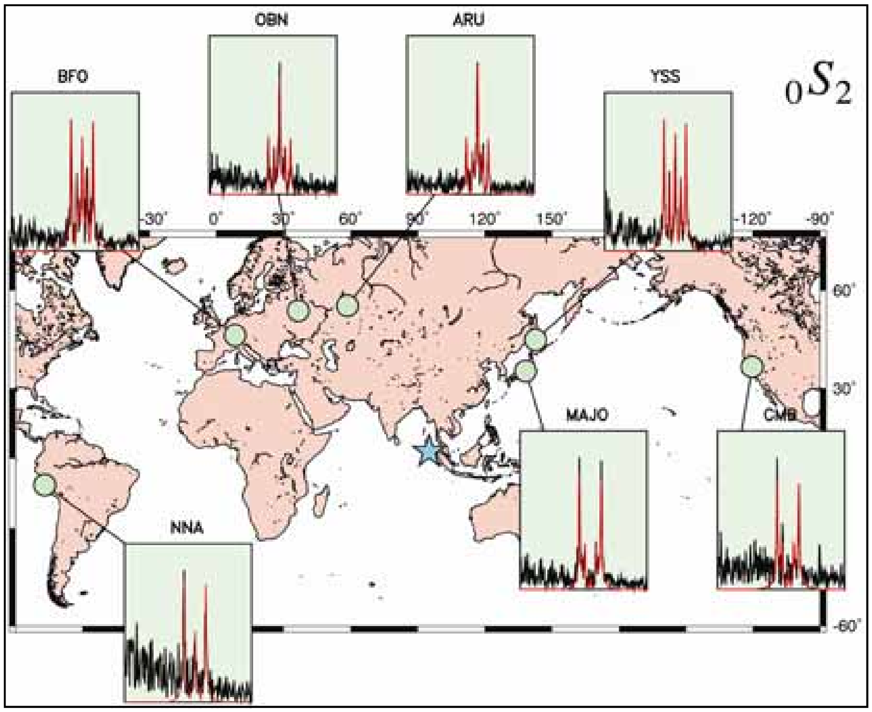

Fig. 6. Comparison of the peaks of the fundamental mode (black line) at seismic station OBN to theoretically predicted values (red lines) for different latitudes of the earthquake centroid.

Implications of Magnitude 9.3

The Sumatra earthquake appears to be the second largest earthquake (after the 1960 Chilean earthquake) recorded since the invention of the seismometer in the late 1800s. Its size is illustrated in Figure 7 where its fault area is compared to some California earthquakes. Relative to the 1906 San Francisco earthquake (known locally as "the big one") the Sumatra earthquake had about three times the slip on a fault three times longer along strike and about 20 times wider (down dip). This difference illustrates the general principle that the largest earthquakes are at subduction zones because of their geometry. The San Francisco earthquake ruptured a long segment of the San Andreas transform fault, which dips vertically, so the down-dip width is controlled by the fact that rocks deeper than about 20 km are weak due to high temperatures and so slide rather than accumulate elastic strain for future earthquakes. In contrast, subduction zone earthquakes on the shallow-dipping plate interfaces have much larger rupture areas at depths shallow enough for strain to build up. Moreover, larger fault dimensions give rise to greater slip, so the combined effects of larger fault area and more slip yield the largest earthquakes.

For the same reason, great subduction zone earthquakes cause the largest tsunamis. In the case of Sumatra, the long rupture played a key role in generating the devastating tsunami. In particular, the large tsunami amplitudes in Sri Lanka and India result from rupture on the northern, north-trending, segment because tsunami amplitudes are largest perpendicular to the fault. This effect is shown by comparison of snapshots from two tsunami animations (Figure 8).

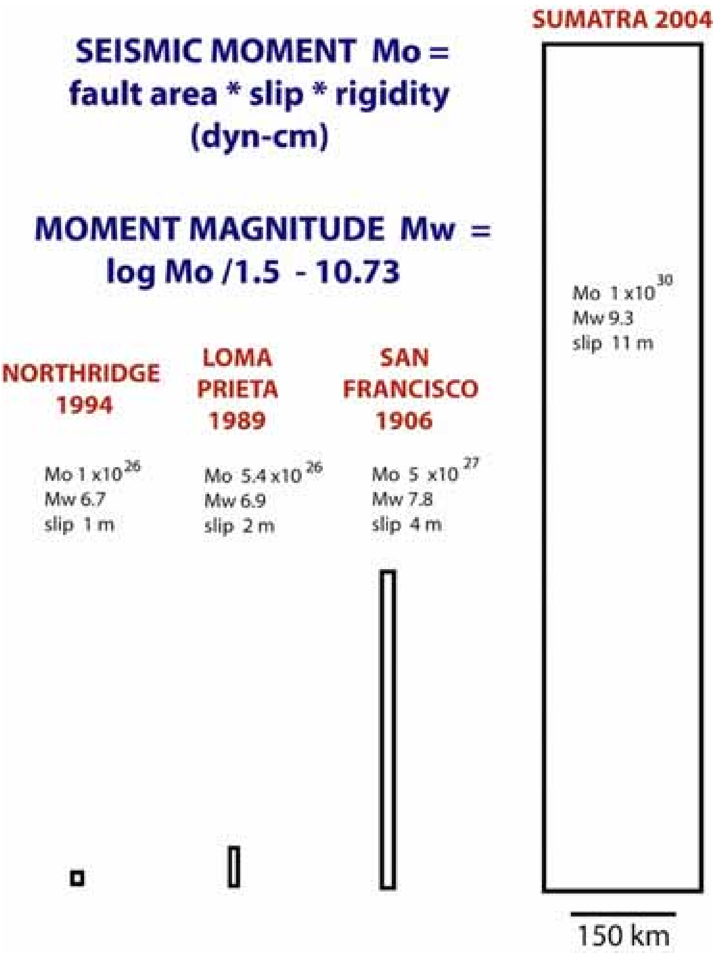

Fig. 7. Comparison of fault areas, seismic moment, slip, and magnitude for the Sumatra earthquake and some California earthquakes.

Tectonic and Hazard Implications

The normal mode analysis indicates that a much larger fault ruptured than found by the earlier body and surface wave analysis, and that there was a very slow rupture in the northern segment. This view is consistent with the regional tectonics. Although the plate geometry and motions are not precisely known, Figure 1b shows estimates of India's motion with respect to Burma. Plate motions between two plates are described by rotations about a pole. Since the pole is nearby, the convergence direction varies along the rupture zone and motion becomes strike-slip at the north end of the rupture, presumably explaining why rupture ceased. The thrust faulting in earthquake reflects the arc-normal component of convergence.

If the entire aftershock zone slipped, strain accumulated from subduction of India beneath

Burma on the northern part of the rupture has also been released. This leaves no immediate danger of a large tsunami being generated by slip on this segment of the plate boundary, since such earthquakes should be at least 400 years apart. However, the danger of a large tsunami resulting from a great earthquake on segments to the south, or a local tsunami due to a large aftershock, remains.

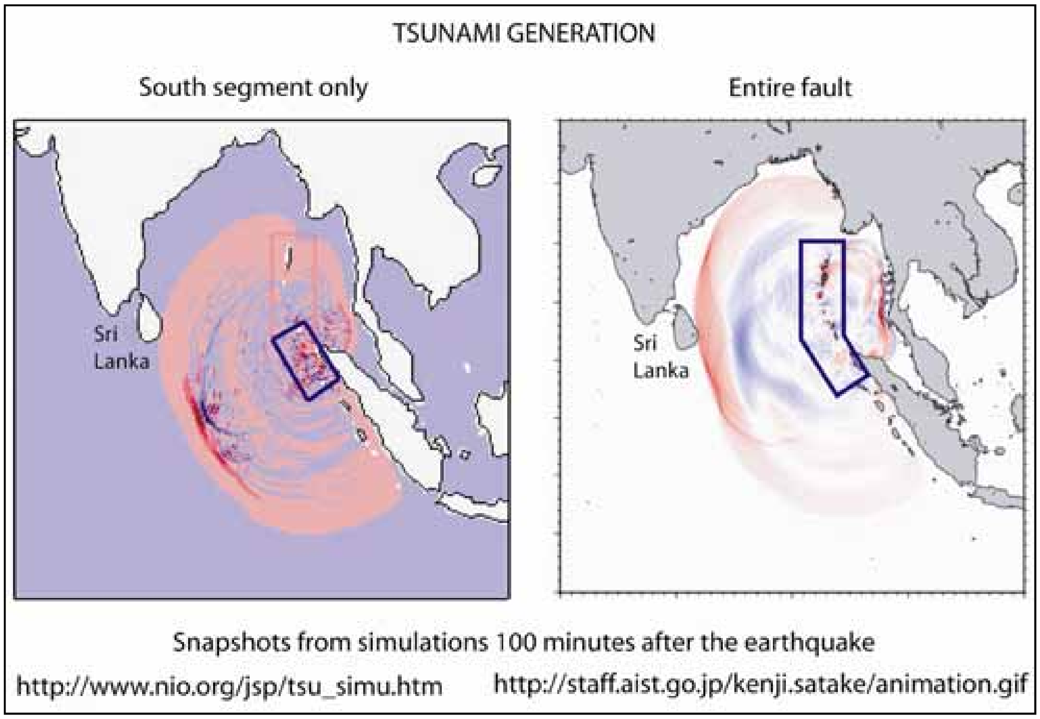

Fig. 8. Comparison of predicted tsunami amplitudes assuming the entire fault ruptured or only the southern segment did. The largest tsunami waves would have missed Sri Lanka if only the southern segment of the fault had ruptured.

Finally, the Sumatra earthquake illustrates the challenge in tsunami warning, namely rapidly determining whether a large earthquake will generate a destructive oceanwide tsunami. The problem is that this must be done quite rapidly, because the water wave travels across the ocean at jet plane speeds. For example, the December tsunami hit Sri Lanka only two hours after the earthquake. Seismic waves from earthquakes travel much faster, giving a very short time window for seismologists to locate the earthquake, decide if a major tsunami will result, and start the warning process. Because false alarms would be enormously expensive and destroy the credibility of the warning system, a difficult decision must be made quickly. The problem is that the tsunami is generated by the long period part of the slip, so – as shown in Figure 3 - the body and surface wave magnitudes do not show whether an earthquake is large enough to generate a major oceanwide tsunami. However surface and body wave magnitudes are much quicker and easier to determine and thus had been used as the basis for tsunami estimates. As a result, algorithms are now being developed to more rapidly assess the seismic moment and decide if a warning should be issued. These approaches together with sea floor sensors that detect the tsunami are the key elements in warning systems.