Supplementary Material

The supplement consists of three parts. The first concerns the numerical scheme used to solve the heat flow equation in the sediments, the first step. It differs from that used by M\textsuperscript{c}Kenzie et al. (2005) because it includes an advective, as well as a diffusive, term, and takes account of the discontinuity at the Moho. The second part uses the waveforms of earthquakes beneath the Bengal Fan to determine their depths. The third part uses the waveforms of earthquakes from a large earthquake on the Antarctic plate to determine its depth and details of its rupture propagation.

1. Numerical Scheme

The variation of $\rho$ and $C_p$ with temperature is small compared with that of $\kappa$ and $k$. It is therefore convenient to rewrite equation (3) as

\[
\frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left( \kappa \frac{\partial T}{\partial z} \right) - u \frac{\partial T}{\partial z} + \frac{H}{\rho C_p} + E(z, t) \tag{A1}
\]

and

\[
E(z, t) = \frac{1}{\rho C_p} \left[ -T \frac{\partial (\rho C_p)}{\partial t} + \left( \kappa \frac{\partial T}{\partial z} - uT \right) \frac{\partial (\rho C_p)}{\partial z} \right] \tag{A2}
\]

and to treat $E$ as a small perturbation, solving equation (A1) by iteration. This correction produced trivial changes to the thermal evolution and was therefore ignored. The nonlinear term in equation (A1) can be written as

\[
\frac{\partial^2 F}{\partial z^2}
\]

using the expression for $\kappa(T)$

\[
F = \int \kappa(T) dT = aT - cb \exp(-T/c) - de \exp(-T/e) \tag{A3}
\]

It is also convenient to rewrite equation (A1) in dimensionless form, using the values of $\kappa$, $\rho$ and $C_p$ when $T = 0^\circ C$, $\kappa_0$, $\rho_0$ and $C_0$, the thickness of the oceanic lithosphere $a$, and the temperature at the base of the lithosphere $T_p$

\[
z = az', \quad t = \frac{a}{\kappa_0} t', \quad T = T_p T', \quad H = \frac{\kappa_0 \rho_0 C_0 T_p}{a^2} H', \quad F = \kappa_0 T_p F'
\]

giving

\[
\frac{\partial T'}{\partial t'} = \frac{\partial^2 F'}{\partial z'^2} - Pe_c \frac{\partial T'}{\partial z'} + H' \tag{A4}
\]

where the thermal Peclet number $Pe_c$ is $ua/\kappa_0$. Equation (A4) can then be solved using a modified Crank-Nicholson scheme. Omitting primes, an implicit finite difference approximation to equation (A4) is

\[
\frac{T_{i+1}^{n+1} - T_i^n}{\Delta t} = \left[ \left( F_{i+1}^{n+1} - 2F_i^{n+1} + F_{i-1}^{n+1} \right) + \left( F_{i+1}^n - 2F_i^n + F_{i-1}^n \right) \right] / (2\Delta z^2)
\]

\[
- Pe_c \left[ \left( T_{i+1}^{n+1} + T_{i-1}^{n+1} \right) - \left( T_{i+1}^n + T_{i-1}^n \right) \right] / (4\Delta z) + H_i \tag{A5}
\]

where $T_i^n$ is the value of $T(n\Delta t, i\Delta z)$ and similarly for $F$, $\kappa$, $k$ and $H$. Equation (A5) is centred on $(n+1/2)\Delta t$ in time and $i\Delta z$ in space. If $\kappa$ is constant this equation is stable for all values
of $\Delta t$, though it is only accurate if $\Delta t < \Delta z^2$ and $\Delta z/Pe_c$. $F_i^{n+1}$ can be written in terms of $T$ and $\kappa$ using a Taylor expansion

$$F_i^{n+1} = F_i^n + (T_i^{n+1} - T_i^n)\kappa_i^n$$

(A6)

Equation (A5) can be then rewritten in a form that can be solved by tridiagonal elimination

$$AT_i^{n+1} + BT_i^{n+1} + CT_i^{n+1} = R$$

(A7)

where

$$f_a = \frac{\Delta t}{2\Delta z^2}, \quad f_b = \frac{Pe_c\Delta t}{4\Delta z}$$

$$A = -f_a\kappa_i^n - f_b, \quad B = 1 + 2f_a\kappa_i^n, \quad C = -f_a\kappa_i^{n+1} + f_b,$$

$$R = T_i^n + 2f_a(F_i^{n+1} - 2F_i^n + F_i^{n-1}) - f_a(\kappa_i^{n+1}T_i^{n+1} - 2\kappa_i^nT_i^n + \kappa_i^{n-1}T_i^{n-1})$$

$$-f_b(T_i^{n+1} - T_i^n) + H_i\Delta t$$

(A8)

and $H_i = 0$ except in the sediments. The boundary conditions $T(0) = 0$ and $T(1) = 1$ were imposed by setting the first and the last values of $T$ on the grid to these values. Once sedimentation stops, both the Peclet number and $f_b$ are set to zero. These expressions are valid when the three points $i - 1$, $i$, $i + 1$ are all either in the mantle or in the crust. Dealing with the discontinuity in material properties at the Moho is more complicated. It was handled by placing the discontinuity at the nearest mesh point $i$ and using the conservative form of equation (A4) to calculate the conductive and advective flows across boundaries at $i - 1/2$ and $i + 1/2$, centred in time at $n + 1/2$. Defining

$$G^n(c) = \left[G^n_{i-1}(c) + G^n_i(c)\right]/2, \quad G^n(m) = \left[G^n_i(m) + G^n_{i+1}(m)\right]/2,$$

$$\bar{G}^n = \left[G^n_i(c) + G^n_i(m)\right]/2$$

$$k^n(c) = \left[k^n_{i-1}(c) + k^n_i(c)\right]/2, \quad k^n(m) = \left[k^n_i(m) + k^n_{i+1}(m)\right]/2,$$

where $(c)$ is the value in the crust and $(m)$ that in the mantle, and

$$g_a = \frac{\Delta t}{2\bar{G}^n\Delta z^2}, \quad g_b = \frac{Pe_c\Delta t}{4\bar{G}^n\Delta z}$$

gives

$$A = -g_ak^n(c) - g_bG^n(c), \quad B = 1 + g_a[k^n(c) + k^n(m)] + g_b[G^n(m) - G^n(c)],$$

$$C = -g_ak^n(m) + g_bG^n(m),$$

$$R = T_i^n + g_a\left[k^n(m)(T_i^{n+1} - T_i^n) - k^n(c)(T_i^n - T_i^{n-1})\right]$$

$$+ g_b\left[G^n(c)(T_i^n + T_{i-1}^n) - G^n(m)(T_{i+1}^n + T_i^n)\right]$$

(A9)

If $\kappa$ and $G$ are constant and are the same in both the crust and mantle, then $\kappa = k/G$ and $F = \kappa T$. Substitution into equations (A8) and (A9) then gives the same expressions for $A$, $B$, $C$ and $R$ in each case, showing that equation (A9) neglects the variation of $\kappa$ with temperature for the mesh point centred on the Moho. A value of $\Delta z$ of 0.005 was used, partly to minimise the resulting error.
2. Earthquake source parameters

The fault-plane solutions and depths of earthquakes shown in Fig. 2, and their sources, are listed in Table 1. Body-wave modeling was carried out in the manner described by Kumar et al. (2015). New solutions for three earthquakes are shown in Figs. 1–3.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Long.</th>
<th>Lat.</th>
<th>Strike</th>
<th>Dip</th>
<th>Rake</th>
<th>$M_w$</th>
<th>Depth (km)</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>DaMoYr</td>
<td>HrMnSc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>080482</td>
<td>024119</td>
<td>86.35</td>
<td>18.66</td>
<td>325</td>
<td>51</td>
<td>131</td>
<td>5.2</td>
<td>38</td>
<td>CMT</td>
</tr>
<tr>
<td>040782</td>
<td>183427</td>
<td>90.62</td>
<td>19.53</td>
<td>240</td>
<td>73</td>
<td>0</td>
<td>5.6</td>
<td>34</td>
<td>CMT</td>
</tr>
<tr>
<td>010785</td>
<td>022352</td>
<td>87.19</td>
<td>18.34</td>
<td>60</td>
<td>58</td>
<td>1</td>
<td>5.4</td>
<td>40</td>
<td>CMT</td>
</tr>
<tr>
<td>120689</td>
<td>000410</td>
<td>89.77</td>
<td>21.86</td>
<td>343</td>
<td>75</td>
<td>167</td>
<td>5.6</td>
<td>47</td>
<td>PSH/Fig.S3</td>
</tr>
<tr>
<td>090792</td>
<td>213401</td>
<td>89.96</td>
<td>20.98</td>
<td>79</td>
<td>59</td>
<td>13</td>
<td>5.3</td>
<td>30</td>
<td>CMT/Fig.S2</td>
</tr>
<tr>
<td>260703</td>
<td>231818</td>
<td>92.31</td>
<td>22.85</td>
<td>338</td>
<td>32</td>
<td>82</td>
<td>5.6</td>
<td>13</td>
<td>PSH/Kumar (2015)</td>
</tr>
<tr>
<td>270703</td>
<td>120729</td>
<td>92.34</td>
<td>22.83</td>
<td>2</td>
<td>16</td>
<td>88</td>
<td>5.4</td>
<td>13</td>
<td>PSH/Kumar (2015)</td>
</tr>
<tr>
<td>110506</td>
<td>172254</td>
<td>94.32</td>
<td>23.31</td>
<td>15</td>
<td>42</td>
<td>102</td>
<td>5.6</td>
<td>31</td>
<td>PSH/Kumar (2015)</td>
</tr>
<tr>
<td>300506</td>
<td>194345</td>
<td>91.93</td>
<td>20.63</td>
<td>63</td>
<td>85</td>
<td>-2</td>
<td>4.9</td>
<td>21</td>
<td>CMT</td>
</tr>
<tr>
<td>100910</td>
<td>172416</td>
<td>90.66</td>
<td>23.41</td>
<td>149</td>
<td>77</td>
<td>171</td>
<td>5.1</td>
<td>18</td>
<td>CMT</td>
</tr>
<tr>
<td>210514</td>
<td>162154</td>
<td>88.03</td>
<td>18.20</td>
<td>319</td>
<td>84</td>
<td>180</td>
<td>6.0</td>
<td>50</td>
<td>PSH/Fig.S1</td>
</tr>
<tr>
<td>120219</td>
<td>013227</td>
<td>85.73</td>
<td>14.40</td>
<td>310</td>
<td>58</td>
<td>150</td>
<td>4.9</td>
<td>40</td>
<td>CMT</td>
</tr>
</tbody>
</table>

Table 1: Source parameters and depths of earthquakes whose fault-plane solutions are shown in Fig. 2. Depths are determined from P and SH body-wave modeling (PSH) or from the updated catalogue of Engdahl et al. (1998). Focal mechanisms are determined either from body-wave modeling (PSH) or are taken from the gCMT catalogue (www.globalcmt.org)
Figure 1:  P (top) and SH (bottom) nodal planes and waveforms (observed, solid; synthetic, dashed) for the 21 May 2014 earthquake in the Bay of Bengal. Numbers beneath the heading are strike/dip/rake/centroid depth/moment. The station code for each waveform is accompanied by a letter identifying its position on the focal sphere, arranged clockwise alphabetically. The time window used for the inversion is marked by vertical bars on each waveform. P and T axes are solid and open circles on the P focal sphere. STF is the Source Time Function.
Figure 2: P (top) and SH (bottom) nodal planes and waveforms for the 9 July 1992 earthquake in the Bay of Bengal. There is an insufficient distribution of stations to independently constrain the fault-plane solution, which is that of the gCMT. The P waves are not very sensitive to the depth because the pP and sP raypaths depart along nodal planes. But the SH waveforms are more sensitive, and the separation of sS is clearly visible, particularly at MAJO, constraining the depth to ~30 km.
Figure 3: P (top) and SH (bottom) nodal planes and waveforms for the 12 June 1989 earthquake in the Bay of Bengal. The depth is well constrained at $\sim47$ km by the clearly visible separation of sS, particularly at INU, MDJ and OBN.
3. The Antarctic Plate earthquake of 4 December 2015 (Mw 7.1)

On 4 December 2015 an earthquake of Mw 7.1 occurred in the Antarctic plate 500 km SW of the Southeast Indian Ridge in lithosphere 20–25 Ma old. Andrews et al. (2020) claimed this event had a centroid depth of 34 km, deeper than the 800°C isotherm. They based their argument on a finite-source fault model, using teleseismic P and SH waveforms. Several aspects of the analysis they presented caused us to also examine this earthquake ourselves. Our concerns were, briefly:

(a) Aspects of the finite-source method and results:

- The fault grid in the inversion of Andrews et al. consists of 8 × 8 km² squares, and they solve for the slip at the nodes. It is unclear whether this really gives the resolution that is claimed: especially for an event of only Mw 7.1, which is small for such a finite-source analysis and only likely to have a source dimension of about 40–50 km.

- Several of the results they presented were puzzling: (i) the source time function was 40 seconds long (their Fig. 4c), which is very long for an event of only Mw 7.1; (ii) the maximum slip was only 1.06 m and the average only 0.12 m: these are very low values for an earthquake of Mw 7.1.

- Superficially, the misfit plot (their Fig. 3d) is quite convincing, but it is unclear how shallow the minimum really is.

- There is little discussion of the trade-offs between source parameters that are inevitable in such finite-source inversions, especially between depth, rupture velocity, source dimension and duration.

- The results of the inversion are dominated by a single circular slip patch on the fault, but it is unclear how well its dimensions are resolved, given the smoothing inherent in both the inversion routine and the frequency content of the data.

(b) Other technical considerations:

- Andrews et al. used stations out to 100° epicentral distance: yet beyond 90° (80° for S) interference from core phases that arrive within a few seconds of direct P and SH is important when the source duration is likely to be 10 s or more. Indeed, of their 58 P waveforms, 25 of them were beyond 90°, and of their 39 SH waves, at least 2 were beyond 90° (and these 2 are the only stations at a crucial azimuth: see below).

- Their inversion for a source time function (STF), using 10 triangles of 3 s overlapping by 1.5 s, gives a total possible duration of 16.5 s. It is therefore unclear how they obtained a duration of 40 s (their Fig. 9d).

- With 58 P and 39 SH waveforms there would normally be a good chance of breaking the usual trade-offs between depth, STF and Mo. But close inspection of their Fig. 5 and of the azimuths to the stations (the number after the station code), reveals that of the 39 SH: 7 are in the azimuth range 086–100 (west Pacific); 28 are in the azimuth range 144–190 (Antarctica); 2 are in direction ~060 (both beyond 90° distance); 1 (Macquarie) at azimuth 127, and 1 at 274 (Africa). There are no SH waveforms at all in the azimuth
range 277–127: i.e. all the SH stations are in the south and none are to the north. This matters, because the trade-offs will be bad, and as for resolving the claimed N-S rupture propagation, impossible: the biggest effects by far are in the azimuth of SH and if only one direction is illuminated, you just can’t tell.

- The SH are down-weighted relative to P by 0.1. This is an unusually large difference, demoting the influence of the phase best able to resolve azimuthal directivity.

- Their Fig. 5, which shows the fits of synthetic to observed seismograms, is scaled to a uniform peak-amplitude, so all information about real relative amplitudes and sensitivity is lost.

We therefore performed some analysis ourselves, using the MT5 inversion program used in section 2 of the Supplementary Material above and described in detail by Kumar et al. (2015).

3.1 The best centroid solution (Fig. 4)

This has a much better station distribution, especially for SH (and all in the distance range 30-90°), than the analysis of Andrews et al. The source time function was defined by 10 elements with half-duration 3 s, overlapping by 3 s, giving a total possible duration of 33 s; but the last few elements were all near-zero, and the STF is only 20 s long. We used a velocity model with 7 km of crust (Vp=6.8 km/s, Vs=3.9 km/s) over a mantle half-space of Vp=8.1 km/s, Vs=4.5 km/s, with 3.5 km water depth. The best-fit centroid depth is 15 km, and is measured from the sea surface (11.5 km below the sea bed). This depth, and the normal-faulting mechanism, are both extremely stable, converging to this result from widely differing starting positions. The nodal planes are very well constrained, and the mechanism (034/44/288 for strike/dip/rake) is not significantly different from the Global CMT solution (042/47/302). The relative amplitudes are real (though are displayed, corrected for geometrical spreading, as though at the same epicentral distance to reflect true amplitude variations related to the radiation pattern). The advantage of using real, not scaled, amplitudes is illustrated by the very small SH at CHTO in the north, requiring it to be near a nodal plane for SH and its surface reflection. (Recall that Andrews et al. had no SH at all in the northern hemisphere). The fit for P and SH is reasonable, but there are some systematic effects. In particular the SH wavelets in the NNE e.g. (JAGI, MMRI, PMG, WRAB) are all slightly too narrow (the downward peak is too early). In the other direction, to the SSW, the wavelet at SNA is too broad. This might be effect of rupture propagation, since whichever plane is the fault plane, the fault must strike approximately NNE-SSW. So we looked at this next.

3.2 Effects of rupture propagation. (Figs. 5 and 6)

We added the effect of a line source propagating either to the SSW or to the NNE at 2.7 km/s (a fairly typical value, and the one used by Andrews et al.). Fig. 5 illustrates the effect using 6 stations (3 P in black, 3 SH in red) at azimuths to the NNE (KAPI, MMRI), SSW (SNA) or to the side (TAU, ATD). The top line is the centroid solution from Fig. 4. We fixed the strike, dip, rake, and left the STF and Mo free in the next inversions.

Line 2 is for a rupture propagating to the SSW. As expected, the SH wavelet is broadened at MMRI (NNE) and narrowed at SNA (SSW), improving the fits, and there is no effect to the side (ATD). The effect is greater for SH than for P, but notice the fit also improves for P, because the broader wavelet to the NNE (KAPI) shows a double-upward pulse, which is better separated than in the narrower pulse to the SSW (SNA), and both azimuths are better fitted. Again, there is no effect to the side (TAU).
Line 3, by contrast, includes a rupture propagating NNE, where the same effects produce dramatically worse fits.

Note also that each inversion produces a slightly different depth, which was left as a free parameter. The propagation to the SSW (line 2) needs a slightly shorter time function overall, and so has a slightly deeper depth to make up the wavelet width: a classic trade-off. The moment is about the same. The depth moves from 15 km (centroid) to 21 km (line source to the SSW). The fit for all stations, from the inversion of line 2 showing the best SSW-propagating source, is shown in Fig. 6.

3.3 Resolution of centroid (or line source) depth (Fig. 7)

The SSW-propagating line source (Fig. 6) is clearly the best fit to the waveforms. This is not a surprise: this size of Mw 7.1 is about when such effects are expected, because the source no longer looks like a point source in the bandpass we use, which is the response of the old WWSSN 15-100 instrument (see e.g. Yielding et al. 1981). The best test for depth resolution is to fix the depth at various values, leaving all other parameters free, and see if the trade-offs can manage the fit as well. This is shown in Fig. 7, using 3 P stations (black) and 3 SH (red), chosen carefully so their reflected phases do not lie on nodal planes. The value at which depth was fixed is shown in red (in km), varying from 11 to 35 (Andrews et al. got 34 km). The formal best fit, with the smallest misfit, is at 22 km (Fig. 6), and waveform fits at other depths that are noticeably worse are marked with a cross. Note that as depth increases (top to bottom) the STF gets shorter (from 35 to 16 s), and the moment decreases (from $6.3 \times 10^{19}$ to $3.0 \times 10^{19}$ Nm): again, a classic trade-off, with the moment increasing at shallow depths because P and pP have opposite polarities and start to cancel each other out for dip-slip faulting. The inversion tries to maintain the wavelet (pulse) width by adjusting either the STF or the depth, but the trade-off is not complete, because whereas the STF duration affects P and SH the same, the depth does not, because Pp-P and sS-S delays do not change in quite the same way with depth. The misfit minimum is fairly broad and a range of 22 ± 5 km is acceptable. These are values below the sea surface (water depth 3.5 km): so the depth is about 19 ± 5 km below the sea bed. The values of 34 km from Andrews et al., and 29 km from global CMT, are both too high.

3.4 Source dimensions.

An estimate of the fault length is the STF duration times the rupture velocity, Vr (with obvious trade-offs between the two). In this case the best STF (Fig. 6) is about 20 s with Vr of 2.7 km/s; which gives a fault length of 54 km — quite reasonable for Mw 7.1.

The moment value of $6 \times 10^{19}$ Nm is quite stable, and the same as the global CMT solution. From the expression for Mo, we can now estimate the fault width W, where:

$$Mo = \mu A \bar{u} = \mu LW(\alpha L),$$

with $L =$ fault length, $A =$ fault area, $W =$ down-dip fault width, $\bar{u} =$ average slip, $\mu = 3 \times 10^{10}$ Nm$^{-2}$, and $\alpha = \bar{u}/L = 5 \times 10^{-5}$ (a typical value, equivalent to a stress drop of 1.5 MPa). This gives a value for $W$ of 12 km, and an expected average slip ($\alpha L$) of about 3 metres: all perfectly sensible, ordinary values for an earthquake of Mw 7.1.

By contrast, the solution of Andrews et al. had an average slip of 0.12 m on a square fault of dimension $48 \times 48$ km$^2$, which would have a stress drop of 0.08 MPa. Those slip and stress-drop values would be extraordinary for a Mw 7.1 earthquake.
3.5 The 800°C isotherm.

The down-dip width of 12 km on a fault dipping 45° gives a vertical depth extent of faulting of about 10 km: i.e. ±5 km either side of the centroid (line) source depth. For the best-fit depth of 22 km this means a depth extent of faulting between 17 and 27 km. To this we should add ±5 km for errors. Again these are values below the sea surface; so this means 14–24 km (±5) below the sea bed. Those ranges are shown in Fig. 8, adapted from Craig et al. (2014): the yellow circle is the best depth; white is the vertical extent of faulting, with black errors). The centroid is well above the 600°C isotherm and, even allowing for errors, the deepest penetration of rupture is above the 700°C isotherm. Note that, with SH only to the south, and ray paths all downward along the fault, Andrews et al. had no realistic chance of resolving vertical propagation effects or the vertical extent of slip from directivity in their finite-source model. Directivity effects are always far greater with azimuth (i.e. horizontal propagation), and that could not be resolved either, since all their illumination is from one direction.

3.6 Conclusion.

We conclude that there is no evidence that the centroid of the 2015 Antarctic plate earthquake is deeper than the 600°C isotherm, and that rupture is unlikely to have penetrated deeper than about 600-650°C. This earthquake is consistent with the worldwide pattern documented (with much more data) by Craig et al. (2014).

Additional reference:
Figure 4: Centroid solution for the 2015 Antarctic plate earthquake. Format and symbols as in Fig. 1.
Figure 5: Illustration of the effects of rupture propagation to the SSW (line 2) and NNE (line 3), compared with the centroid solution of Fig. 3 (line 1). See text for details.
Figure 6: Solution for the SSW-propagating line source. See text for details.
Figure 7: Tests to show the sensitivity to depth of the SSW-propagating line source, and the trade-offs between source time function (STF), depth and seismic moment. See text for details.
Figure 8: Depths of intraplate oceanic earthquakes as a function of lithosphere age and temperature, adapted from Craig et al. (2014). The top panel is a detail of age range 0–30 Ma. Our determination of the centroid of the 2015 Antarctic plate earthquake, for the SSW-propagating line source, is the yellow circle. The white bar shows the probable depth extent of faulting, with black error bars. The circled event marked C at 50 Ma is the centroid redetermination by Craig et al. (2014) of the 1964 Chile outer-rise earthquake, which had been similarly mis-assigned.