SCARDEC: a new technique for the rapid determination of seismic moment magnitude, focal mechanism and source time functions for large earthquakes using body-wave deconvolution

M. Vallée,[1] J. Charléty,[1] A. M. G. Ferreira,[2,3] B. Delouis[1] and J. Vergoz[4]

[1]*Geoazur, Observatoire de la Côte d'Azur, IRD, CNRS, Université de Nice–Sophia Antipolis, Valbonne, France. E-mail: vallee@geoazur.unice.fr*
[2]*School of Environmental Sciences, University of East Anglia, Norwich, UK*
[3]*ICIST, Instituto Superior Técnico, Lisboa, Portugal*
[4]*Laboratoire de Détection Géophysique, CEA, Bruyères le Chatel, France*

SUMMARY

Accurate and fast magnitude determination for large, shallow earthquakes is of key importance for post-seismic response and tsunami alert purposes. When no local real-time data are available, which is today the case for most subduction earthquakes, the first information comes from teleseismic body waves. Standard body-wave methods give accurate magnitudes for earthquakes up to $M_w = 7$–7.5. For larger earthquakes, the analysis is more complex, because of the non-validity of the point-source approximation and of the interaction between direct and surface-reflected phases. The latter effect acts as a strong high-pass filter, which complicates the magnitude determination. We here propose an automated deconvolutive approach, which does not impose any simplifying assumptions about the rupture process, thus being well adapted to large earthquakes. We first determine the source duration based on the length of the high frequency (1–3 Hz) signal content. The deconvolution of synthetic double-couple point source signals—depending on the four earthquake parameters strike, dip, rake and depth—from the windowed real data body-wave signals (including *P, PcP, PP*, *SH* and *ScS* waves) gives the apparent source time function (STF). We search the optimal combination of these four parameters that respects the physical features of any STF: causality, positivity and stability of the seismic moment at all stations. Once this combination is retrieved, the integration of the STFs gives directly the moment magnitude. We apply this new approach, referred as the SCARDEC method, to most of the major subduction earthquakes in the period 1990–2010. Magnitude differences between the Global Centroid Moment Tensor (CMT) and the SCARDEC method may reach 0.2, but values are found consistent if we take into account that the Global CMT solutions for large, shallow earthquakes suffer from a known trade-off between dip and seismic moment. We show by modelling long-period surface waves of these events that the source parameters retrieved using the SCARDEC method explain the observed surface waves as well as the Global CMT parameters, thus confirming the existing trade-off. For some well-instrumented earthquakes, our results are also supported by independent studies based on local geodetic or strong motion data. This study is mainly focused on moment determination. However, the SCARDEC method also informs us about the focal mechanism and source depth, and can be a starting point to study systematically the complexity of the STF.

Key words: Inverse theory; Earthquake source observations; Body waves; Surface waves and free oscillations; Wave propagation; Subduction zone processes.

1 INTRODUCTION

Most major earthquakes ($M > 7.5$) occur in subduction zones, often in places where there is sparse local seismological or geodetical instrumentation. In these cases, the knowledge that we can obtain about these events depends mainly on our ability to analyse the teleseismic wavefield. Efficient methods are important both to give accurate information in the near-real time (tsunami alert, post-seismic reaction) and to provide later precise and systematic information on the seismicity (tectonics, seismic source understanding and seismic

hazard...). Current methods to analyse teleseismic waves usually involve two main steps. First, simplified source models are used to determine the earthquake's focal mechanism, magnitude and depth. Then, detailed analyses can be done to retrieve further information about the seismic source process (location of major slip zones, average rupture velocity...). A refinement of moment magnitude can also be done in this second step.

However, for major earthquakes, the possibility to decouple fault geometry and source processes has to be questioned. When using classical body-wave (*P* and/or *SH*) point-source approaches (e.g. Nabelek 1984; Ruff & Miller 1994; Goldstein & Dodge 1999), we intrinsically impose that the source time function (STF) is the same at all stations. This assumption is reasonable for moderate earthquakes, at least if high frequency waves are discarded, but becomes increasingly invalid as the magnitude and source dimension increase; extended source effects cause the STFs to be dependent on the recording station. Methods incorporating source complexity in the definition of the focal mechanism exist, but generally require some tuning, as for example, the iterative approach of Kikuchi & Kanamori (1991) or the slip patch method of Vallée & Bouchon (2004). For large shallow earthquakes, another complication arises because the low-frequency part of the signal, which controls the seismic moment determination, is strongly attenuated by destructive interferences between direct wave (*P*) and surface reflected phases (*pP, sP*).

Apart from simple body-wave point-source methods, the other main class of semi-automatic methods used to determine focal mechanism, depth and magnitude is the centroid approach. This technique, based on the work of Dziewonski *et al.* (1981), is today routinely implemented in the Global Centroid Moment Tensor (GCMT) catalogue, which is extensively used in tectonic studies. Based on low-frequency body and/or surface waves, the method simultaneously optimizes the location and timing of the centroid of the source (which can be seen as the spatial and temporal barycentre of the earthquake) and the seismic moment tensor. The method is very attractive because it incorporates in the centroid location the major complexities of the source. The high-frequency STF variability is also a minor problem, because low-frequency surface waves (generally periods of about 150 s) strongly control the solution. However the method presents a few drawbacks. First, its interest for tsunami alert is limited because it requires the recording of slow surface waves at teleseismic distances. Second, when the earthquake is large ($M_w \geq 7.8$) and shallow, the GCMT solution is obtained using mostly low-frequency surface waves. As a consequence, it suffers from a well-known trade-off between the fault's dip δ and the seismic moment M_0 (e.g. Kanamori & Given 1981). For dip-slip earthquakes, the method precisely retrieves the quantity ($M_0 \sin 2\delta$), but cannot accurately resolve the two parameters separately. Because large subduction earthquakes often occur on shallow-dipping planes (where $\sin 2\delta \sim 2\delta$), the effect of the trade-off is large for this type of earthquakes. For example, values of dip of $6°$ or $12°$ would lead to an uncertainty of a factor of 2 for M_0, or an uncertainty of 0.2 in moment magnitude M_w. The latter problem also occurs for the recently developed *W*-phase approach (Kanamori 1993; Kanamori & Rivera 2008), which uses the low-frequency information of the beginning of the seismic signals (between *P* and *S* waves). A last minor problem with GCMT is the empirical determination of the source half-duration. Low values of this parameter make the wave amplitudes larger, which implies that lower values of the earthquake moment are required to explain the data. In the Global CMT (GCMT) routine, the half-duration is not inverted but is fixed as a function of the magnitude. However there is a large duration diver-

sity, even for earthquakes of the same magnitude. As an example for earthquakes given with $M_w = 7.7$ in the GCMT catalogue, we can take the 2001/01/13 El Salvador earthquake and the 2006/07/17 Java earthquake. The first one is a short and impulsive earthquake (duration of about 15 s; Vallée *et al.* 2003), while the second one is a slow tsunami earthquake with duration around 150 s (Ammon *et al.* 2006). Consistently, the latter study determines a moment magnitude 0.1 larger than that reported in the GCMT catalogue for the 2006 Java earthquake.

Because magnitude is a decisive information for alert purposes, some studies aim at determining the moment magnitude without resolving the focal mechanism or the depth. One of these methods is known as the M_{wP} method (Tsuboi *et al.* 1995). It directly integrates the *P*-wave displacement to estimate the associated moment magnitude. The method first requires an azimuthal average of the displacements to take into account the radiation pattern. Another greater problem arises if reflected phases *pP* or *sP* arrive before the end of the direct *P* radiation (which is always the case for large shallow earthquakes); arrival of these waves strongly pollute the measured amplitude displacements. Other methods, based on semi-empirical considerations, analyse the high-frequency part of the *P* radiation to determine the source duration (Ni *et al.* 2005; Lomax 2005), and then use a refined M_{wP} approach (Lomax & Michelini 2009), energy considerations (Lomax *et al.* 2007) or amplitude measurements (Hara 2007) to retrieve the moment magnitude. These approaches can be very useful to get a first idea of the size of a major earthquakes, but lack a physical basis to better understand the characteristics of these events.

The goal of this study is to provide a fast and reliable determination of the main characteristics of major earthquakes, without using empirical relationships or oversimplifications of the source process. The objective is to provide both rapid information and reliable source characteristics, useful for further analyses of the earthquakes. We present here a way to do so, based on a deconvolutive approach of a broad range of body waves (*P, PcP, PP, S, ScS*, along with all the associated surface reflected phases). The STF can have an arbitrary complexity and the apparent STFs may differ from station to station, as expected for large earthquakes. This approach, that we will name the SCARDEC method, is applied to most subduction earthquakes with $M_w \geq 7.8$ in the period 1990–2010. Results are generally found close to GCMT parameters. However, for half of the earthquakes, the fault's dip angle is found steeper and the seismic moment is smaller (by up to a factor of 2) than in the GCMT catalogue. In these cases, we check by forward modelling that our proposed model explains surface wave data as well as the GCMT model. We show in the following sections that the SCARDEC method reliably determines the first-order characteristics of large earthquakes, using seismic data arriving in the 30 min following the earthquake origin time. Moreover, the method provides as a by-product the apparent STFs, which are valuable for further analyses of the source process.

2 SCARDEC METHOD

2.1 Wave modelling and data selection

In the teleseismic range ($30° < \Delta < 90 − 95°$), the modelling of direct *P* and *SH* Green's functions along with the associated local surface reflections (*pP, sP, sS*) can be carried out accurately using standard ray techniques. We use here the method of Bouchon (1976), which includes the reflectivity method (Fuchs & Müller

Table 1. Teleseismic data used for each subduction earthquake in this study. Index, name, date and GCMT moment magnitude of each event are first given. *P* stations and *SH* stations indicate the number of stations used in the analysis of compressive and transverse body waves, respectively. *P* gap and *SH* gap are the maximum azimuthal gaps (°) between stations for compressive and transverse body waves, respectively.

n^0	Name	Date	Mw GCMT	*P* stations	*P* gap	*SH* stations	*SH* gap
1	Java	02/06/1994	7.76	14	72.8	13	91.7
2	Chile	30/07/1995	8.00	11	84.2	14	62.3
3	Jalisco	09/10/1995	7.98	10	98.2	11	88.4
4	Kuril	03/12/1995	7.88	18	74.0	18	74.0
5	Minahassa	01/01/1996	7.87	18	42.3	18	41.7
6	IrianJaya	17/02/1996	8.19	13	65.7	13	65.7
7	Andreanof	10/06/1996	7.88	19	65.6	20	65.6
8	Kamtchatka	05/12/1997	7.76	20	62.0	18	62.3
9	Peru	23/06/2001	8.39	15	59.8	15	70.6
10	Hokkaido	25/09/2003	8.26	20	59.9	22	59.9
11	Sumatra	28/03/2005	8.62	23	39.4	26	33.7
12	Kuril	15/11/2006	8.30	21	47.9	17	77.7
13	Solomon	01/04/2007	8.07	16	72.5	17	74.6
14	Peru	15/08/2007	7.97	15	75.9	19	42.6
15	Sumatra	12/09/2007	8.49	18	62.2	21	40.2
16	NewZealand	15/07/2009	7.78	18	71.8	18	71.8
17	Chile	27/02/2010	8.79	18	49.4	18	57.6

1971; Müller 1985) for both source and receiver crusts. The mantle propagation is simply taken into account by geometrical spreading and attenuation (t*) factors. Take-off angles below the crust and the geometrical spreading factor are deduced from the global traveltime model IASP91 (Kennett & Engdahl 1991). Simple modifications of the same technique allow us to model the core-reflected (*PcP* and *ScS*) and surface-reflected (*PP* and *SS*) phases. In both cases, take-off angles and geometrical spreading have to be computed from the traveltime derivatives of the corresponding phases. For the core-reflected phases, the computed Green's function has to be multiplied by the reflection coefficient at the core surface (1 for *ScS*, because we use only the transverse component). For the surface-reflected phases, we multiply the Green's function by the reflection coefficient at the Earth's surface and Hilbert-transform the resulting wavefield. Modelling of surface-reflected phases is imprecise for distances shorter than 60°, because these waves remain in the heterogeneous upper mantle. Thus the Green's function including direct, core and surface-reflected phases can be computed in the range from 60° to 90–95°. Currently, even in this restrained distance range, the station distribution of the seismic global network (FDSN) insures a suitable azimuthal coverage (see e.g. Table 1).

For the scope of our method, the *PcP*, *PP* and *ScS* phases have to be used because for large earthquakes with long source durations, one of these phases interferes with the direct *P* or *SH* wave. For a 100 -s-long superficial source, this occurs with the *PcP* phase for distances larger than 40° and with *ScS* for distances larger than 60°. For a 150 -s-long superficial source, this occurs with the *PcP* phase for distances larger than 35°, with the *PP* phase for distances shorter than 70° and with *ScS* for distances larger than 50°. The integration of the *SS* phase in our method is less useful, because in the 60–95° distance range, it arrives at least 240 s after the *S* wave. Moreover, its arrival time can be close (150 s) to the Love waves arrival at distances around 60°, causing significant wave interference. Using the combination of *P*, *PcP* and *PP* in the 60–90° distance range and of *SH* and *ScS* in the 60–95° distance range, we can analyse earthquakes with a source duration up to 250 s ($M_{\rm w} = 8.7$–9). For longer—but very rare—earthquakes, some mixing between phases would still occur, which impedes the precise analysis of giant earth-

quakes. We call hereafter 'compressive waves' the three phases *P*, *PcP* and *PP*, and 'transverse waves' the two phases *S* and *ScS*.

We propose here to check our method for the major interplate subduction earthquakes of the last 20 yr. Specifically, we select earthquakes occurring between 1990 and 2010, with moment magnitude larger than 7.8, with a thrust mechanism and with depth smaller than 50 km. Such a request from the GCMT catalogue (http://www.globalcmt.org/CMTsearch.html) gives a list of 23 earthquakes. In this list, we do not consider the Sichuan earthquake (continental intraplate). In addition, we do not include the 2000 November 16 New Ireland, 2000 November 17 New Britain, 2007 September 12 (23h49) Sumatra and 2009 October 7 (22h18) Santa Cruz earthquakes, because they were preceded within a day by a similar or larger earthquake, which makes the waveforms noisy. We finally discard the 2004 Sumatra earthquake because the source duration is much longer than 250 s. For such an earthquake, we believe that its giant character is most efficiently identified by its very long high-frequency duration (Lomax 2005; Ni *et al.* 2005). The remaining 17 earthquakes are presented in Table 1 and on the map of Fig. 1. For each of these earthquakes, we automatically retrieve FDSN broad-band data using the IRIS Wilber interface (http://www.iris.edu/wilber). When several stations are present in a 10° azimuthal range, we only select the one with the best signal-to-noise ratio. The number of stations selected for compressive and transverse waves, along with the largest azimuthal gap, are shown in Table 1.

2.2 Source duration determination

The first step in our method is to estimate the earthquake source duration. This can be sometimes directly read on the *P*-wave seismograms, but some subjective interpretation is necessary, in particular when the earthquake is long and little impulsive, or when the *pP* and *sP* phases lengthen the signal. For an automated approach, we follow the methods based on the high-frequency *P*-wave duration (e.g. Lomax 2005; Ni *et al.* 2005). These methods use the simple observation that at high frequency (around 2 Hz), the vertical component teleseismic waveform is mostly dominated by the
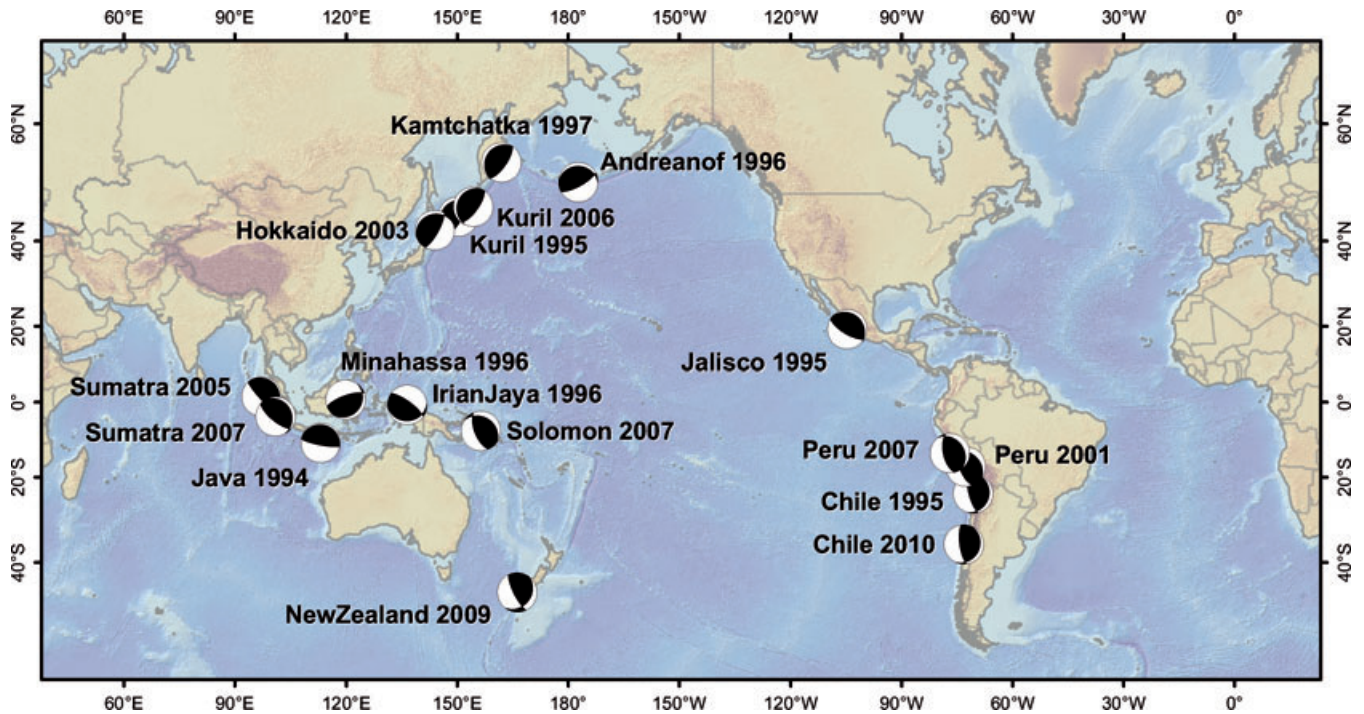
Figure 1. Location of the studied subduction earthquakes. The focal mechanisms determined in this study are presented at the epicentral location of each earthquake.

direct P wave. Therefore a measurement of the duration of the signal in this frequency range gives a good estimation of the source duration.

In practice, some care has to be taken to automatically determine the end of the high-frequency signal. In particular, some noisy stations can lead to a large overestimation of the P-wave duration. Moreover, even for stations with good signal-to-noise ratio, a com-

plex P-wave coda lengthens the high-frequency signal (Fig. 2). As in previous studies (Lomax 2005; Ni *et al.* 2005), we thus have to tune the duration measurement's criteria. We use the following procedure, based on systematic tests with a large earthquake catalogue (about 50 earthquakes with magnitude larger than 7): for each of the n vertical component signals, we select the time of the first P-wave arrival (T_0) as the origin time. After bandpass filtering between 1
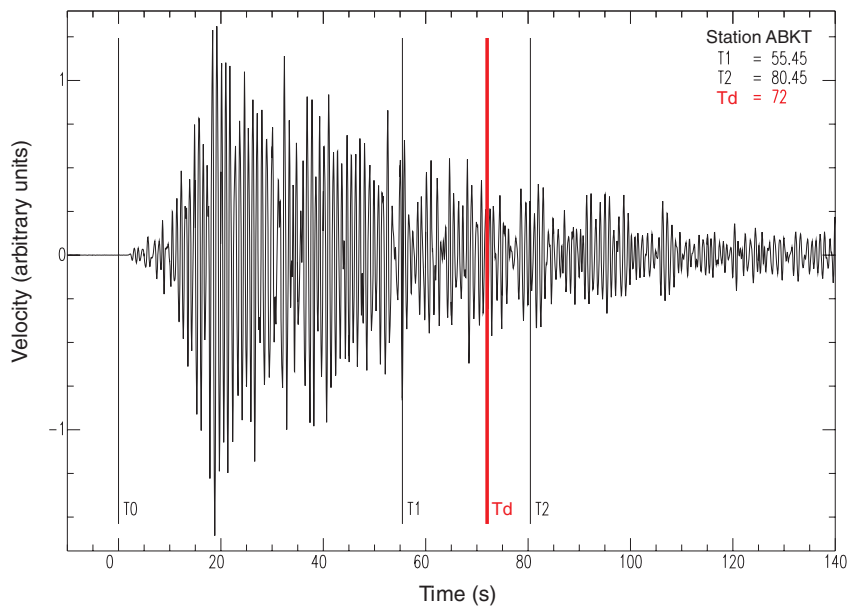

Figure 2. Source duration determination by high-frequency analysis of vertical teleseismic waveforms. The origin time T_0 is the time of the first P-wave arrival. After defining the times T_1 for the n vertical teleseismic waveforms and classing them by ascending order, we extract the station corresponding to the index $n/4$ (see main text). We show an illustrative example for the 2003 Hokkaido earthquake, for which ABKT is the selected station. The vertical waveforms bandpass filtered between 1 Hz and 3 Hz, along with the times T_1, T_2 and T_d—estimate of the rupture duration—are shown in this figure.

and 3 Hz, we locate the time of the last signal point which is above 50 per cent of the maximum of the signal (T_1). We class the times T_1 by ascending order, and select the time T_1 with index $n/4$ in the ordered list. This reduces the chance of using stations which underestimate (rare) or overestimate (more common) the signal duration. The choice of using the time with index $n/4$ comes from extensive tests with our large earthquake catalogue, after having tried a variety of different criteria. As an estimation of the robustness of the measurement, we have also checked that stations corresponding to indices neighbours of $n/4$ give a very similar estimation. The signal corresponding to this index, in the case of the 2003 Hokkaido earthquake, is presented in Fig. 2.

We now consider the time T_2 equal to $(T_1 + 25)$ s. The time T_1 is lengthened for three reasons. First, given the criterion used to define T_1 (last point above 50 per cent of the signal maximum), it is very likely that we miss the final part of the source emission. Second, we aim at defining a source duration which does not underestimate the source duration seen at any station. In fact, directivity effects may cause the source duration to be apparently longer in some azimuths. Third, it is better to slightly overestimate the source duration than to underestimate it. Overestimation of the source duration results in the introduction of some low-amplitude noise signal while underestimation implies that a part of the real source emission is not considered. The choice of the 25 s value mainly comes from this third criterion: we have checked with our test catalogue that this additional time prevents us from underestimating the source duration. Finally, we subtract to T_2 the $(pP - P)$ time to take into account that for shallow and intermediate-depth earthquakes, the pP phase also contributes to the high-frequency, vertical component seismogram, lengthening the signal. This final time, noted T_d, is presented in Table 2 for all the earthquakes of this study.

For compressive waves (P, PcP, PP), this time T_d is directly used as an estimate of the source duration. For transverse waves (SH, ScS), directivity effects are expected to be larger. Simple calculations for a unilateral rupture with a fast 3.5 km s^{-1} rupture velocity show us that these directivity effects may lead to an apparent duration 15 per cent longer for transverse waves than for compressive waves. We thus take the value $1.15.T_d$ as an estimate of the transverse waves source duration.

Table 2. Source duration T_d determined by high-frequency analysis (1-3 Hz) of vertical teleseismic waveforms.

n^0	Name	Date	T_d (s)
1	Java	02/06/1994	110.3
2	Chile	30/07/1995	96.3
3	Jalisco	09/10/1995	71.9
4	Kuril	03/12/1995	63.8
5	Minahassa	01/01/1996	66.3
6	IrianJaya	17/02/1996	105.4
7	Andreanof	10/06/1996	64.5
8	Kamtchatka	05/12/1997	55.8
9	Peru	23/06/2001	121.0
10	Hokkaido	25/09/2003	72.0
11	Sumatra	28/03/2005	105.8
12	Kuril	15/11/2006	117.7
13	Solomon	01/04/2007	91.6
14	Peru	15/08/2007	121.4
15	Sumatra	12/09/2007	105.2
16	NewZealand	15/07/2009	66.2
17	Chile	27/02/2010	127.4

2.3 Deconvolutive approach

Most body-wave methods use strong *a priori* constraints on the source process for the fast determination of the earthquake's magnitude and focal mechanism. Generally, the absolute STF is represented by discrete points and the methods optimize the value of these points together with the depth and the focal mechanism parameters to determine the focal mechanism and magnitude (Nabelek 1984; Ruff & Miller 1994; Goldstein & Dodge 1999). Such approaches do not give a complete freedom to the STF, and, most importantly, impose that the STF is the same for all stations. This is not a serious concern for moderate-to-large earthquakes (up to $M_w = 7-7.5$) because directivity effects, which cause changes in the STF at each station, are generally weak. However, for larger earthquakes, these effects increase and using a unique STF for all stations becomes a poor approximation. Modifications of the method of Nabelek (1984) and Ruff & Miller (1994) have been introduced to take into account a very simple directivity (i.e. unilateral propagation with a constant rupture velocity), but they cannot fully represent the diversity of directivity effects (due e.g. to bidimensional propagation or changes in rupture velocity). An alternative could be to low-pass filter the body waves, for example below 0.01 Hz, to reduce the high-frequency directivity effects. However this is not a solution either because the body waves would interfere with other low-frequency waves, such as the W phase (Kanamori 1993).

Another difficulty arises for large and shallow earthquakes. It is well known that the direct P-wave displacement is directly the STF, if we correct for focal mechanism and propagation constants (e.g. Lay & Wallace 1995, p. 337). Therefore, for deep earthquakes (or, more precisely, for depths such that the end of the earthquake occurs before the arrival of pP wave), resolving the seismic moment is relatively straightforward because it only requires an integration of the direct P wave, after correcting for the required constants. For shallower earthquakes, the direct P wave interferes with pP and sP waves. It creates a more complex P wave train and causes a reduction of its low-frequency content because one of the pP or sP waves generally have an opposite polarity (high-pass filter effect). When optimizing the agreement between synthetics and such complex P wave train, the fit will thus be much more influenced by some high-frequency features (little affected by the destructive interferences between P, pP and sP wave) than by the reduced-amplitude low-frequency features. The obtained STF is likely peaky, reproducing the impulsive parts of the P wave train, and lacks some long-period trend. This last effect explains why there is a tendency of underestimating the seismic moment of large earthquakes when using classical P-wave methods.

The basic idea of this study is to propose a method able to retrieve the first-order characteristics of earthquakes (seismic moment, depth and focal mechanism) without imposing constraints on the source process. We begin with the classic representation theorem (e.g. Aki & Richards 2002, p. 51) of the teleseismic displacement U, which depends on the source term f and the propagation term $G_{\phi,\delta,\lambda}$ (where ϕ, δ, λ are respectively the strike, dip and rake of the earthquake). Neglecting the along-dip extension of the source (line-source approximation), we have

$$U(\omega) = \int_{L_1}^{L_2} f(x, \omega)\, G_{\phi,\delta,\lambda}(x, z_c, \omega)\, \mathrm{d}x, \tag{1}$$

where L_1 and L_2 are the lateral edges of the fault, and z_c is an average depth of the earthquake. For an individual body wave in a spherical Earth, $G_{\phi,\delta,\lambda}$ can be easily modelled

as

$$G_{\phi,\delta,\lambda}(x, z_c, \omega) = G^0_{\phi,\delta,\lambda}(z_c, \omega)\, e^{i\vec{k}.\vec{x}} \quad \forall x \in [L_1 L_2], \tag{2}$$

where $\vec{k}$ is the wave vector of the considered body wave. G^0 represents the teleseismic wavefield generated by a double-couple point source located at the earthquake hypocentre. This term can be numerically evaluated using the techniques explained in Section 2.1. For a propagating rupture along the fault, the source term f may be written as

$$f(x, \omega) = s(x, \omega) e^{-i\omega T_r(x)}, \tag{3}$$

where s is the local STF describing the shape of the movement of each point x of the fault and T_r is the rupture propagation time. We can now rewrite (1) as

$$U(\omega) = G^0_{\phi,\delta,\lambda}(z_c, \omega) \int_{L_1}^{L_2} s(x, \omega) e^{i(\vec{k}.\vec{x} - \omega T_r(x))} \, dx. \tag{4}$$

In the time domain, (4) may be written as

$$U(t) = G^0_{\phi,\delta,\lambda}(z_c, t) * F(t) \tag{5}$$

where F, often called the apparent or relative source time function (RSTF), is

$$F(t) = \int_{L_1}^{L_2} s\left(x, t + \frac{x \sin\theta \cos(\phi - \alpha)}{v_\phi} - T_r(x)\right) dx. \tag{6}$$

In this last equation, θ, v_ϕ are respectively the take-off angle and phase velocity of the considered body wave, and α is the azimuth of the recording station. These last three parameters, which depend on the body wave type and/or the location of the station, explain why F is called an apparent or relative STF. However, F has an important integral property, independent of the wave type or station location

$$\int_0^\infty F(t)\, dt = M_0 \quad \forall \alpha, \theta, v_\phi, \tag{7}$$

where M_0 is the seismic moment of the earthquake. F has also three other important properties, which directly come from the properties of the local STF s: F is a positive, causal and bounded function. As we have an estimate of the global source duration T_d, we can be more precise on this last property and assert that F has to be bounded at T_d. The causality property comes from the fact that for body waves the directivity term $\Gamma = \frac{x \sin\theta \cos(\phi - \alpha)}{v_\phi}$ is shorter than $T_r(x)$, even in the intrasonic rupture propagation regime. Finally, because θ is small and v_ϕ is high (particularly for the faster P wave), the directivity term Γ remains moderate for body waves. This implies that the function F cannot differ a lot from station to station. Therefore, when deconvolving $G^0_{\phi,\delta,\lambda}(z_c, t)$—for a given set of parameters $(\phi, \delta, \lambda, z_c)$—from $U(t)$ at all recording stations, the tested set of parameters is realistic only if the deconvolution result F^1 verifies the five following conditions:

  (i) F^1 is positive;
  (ii) F^1 is causal;
  (iii) F^1 is bounded to T_d;
  (iv) the time integral of F^1 is constant for all stations and
  (v) F^1 varies moderately from station to station, particularly for P waves.

Respecting all these conditions at all stations and for all body wave types puts strong constraints on the set of four parameters on which depend the deconvolution. The idea of this study is therefore that even if we do not know what really happens inside the source (function s, rupture propagation T_r), we have enough information on F to constrain the focal mechanism and depth of the

earthquake. Clearly, these constraints are stronger when a maximum of stations and wave types are taken into account, because it better samples the focal sphere. Here, we compute $G^0_{\phi,\delta,\lambda}(z_c, t)$ separately for compressive body waves and for transverse body waves, using the epicentral location given in the NEIC catalogue (http://neic.usgs.gov/neis/epic/). For compressive body waves, we include the direct P wave, the PcP and PP waves. For transverse body waves, we include the direct SH wave and the ScS (transverse) wave. In both cases, all the refracted and reflected waves in the source and receiver crust are considered. Because we use a Moho depth of 35 km with a simple linear wave velocity increase (between $6\,km\,s^{-1}$ and $8\,km\,s^{-1}$ for P waves), the only energetic waves generated in the crust are the local surface reflected waves (i.e. pP, sP, sS, and similarly $pPcP$, $sPcP$, $sScS$, pPP, sPP). We show in Fig. 3(b) an example of the term $G^0_{\phi,\delta,\lambda}(z_c, t)$ for the compressive waves. There is an approximation in deconvolving in this way the compressive and transverse wavefield. In fact, the take-off angles—between P, PcP and PP waves on one hand and between SH and ScS waves on the other hand—vary while the derivation between eqs (1) and (6) is theoretically exact only if all the waves share the same wave vector. However, the changes remain moderate (no more than 20° variation) and the gain obtained in integrating the PcP, ScS and PP waves, both for the better sampling of the focal sphere and for the analysis of long earthquakes, justifies this approximation.

It would however be difficult to follow exactly the methodology explained above to determine the optimal set of parameters ($\phi, \delta, \lambda, z_c$). First, an unconstrained deconvolution is well known to be unstable and second it would be very difficult to build a misfit function that simultaneously takes into account the five conditions. A more efficient way to do is to constrain the deconvolution result F^1 to respect the conditions, and then to estimate the misfit by reconvolving F^1 with $G^0_{\phi,\delta,\lambda}(z_c, t)$ and comparing with U. Conditions (i), (ii) and (iii) can be integrated in the deconvolution process with the method of Bertero *et al.* (1997). Condition (iv) can be taken into account with the method of Vallée (2004). We present in Fig. 3(c) the result of the constrained deconvolution, for the compressive body waves recorded at station NOUC during the 2003 Hokkaido earthquake. In this example, $(\phi, \delta, \lambda, z_c) = (251°, 22°, 129°, 35\,km)$, $T_d = 72s$, and $M_w = 8.15$. Such parameters are shown here to be realistic because when reconvolving the stabilized deconvolution result with $G^0_{\phi,\delta,\lambda}(z_c, t)$ the agreement with the observed waveforms is good (Fig. 3d).

2.4 Optimization strategy

2.4.1 Optimal source model

Before analysing seismic body waves to determine the earthquake focal mechanism and depth, we first have to define the suitable body-wave frequency band. In fact, both very low and very high frequencies have to be discarded. The lower limit is constrained by the existence of the low-frequency W phase (Kanamori 1993), which becomes predominant for frequencies lower than 0.005 Hz (Kanamori & Rivera 2008). The upper limit is governed by several factors. First, we model the earthquake depth extension by its average depth. This is clearly not exact at high frequency, and imposes us to reject the high-frequency signal content to keep the deconvolutive approach robust. Such filtering also allows us to reduce the influence of local variations of focal mechanism. Second, while the direct P and SH waves can be precisely modelled for short periods (down to a few seconds), this is not the case for the PcP, ScS and PP waves included in this study. The first two waves interact with the
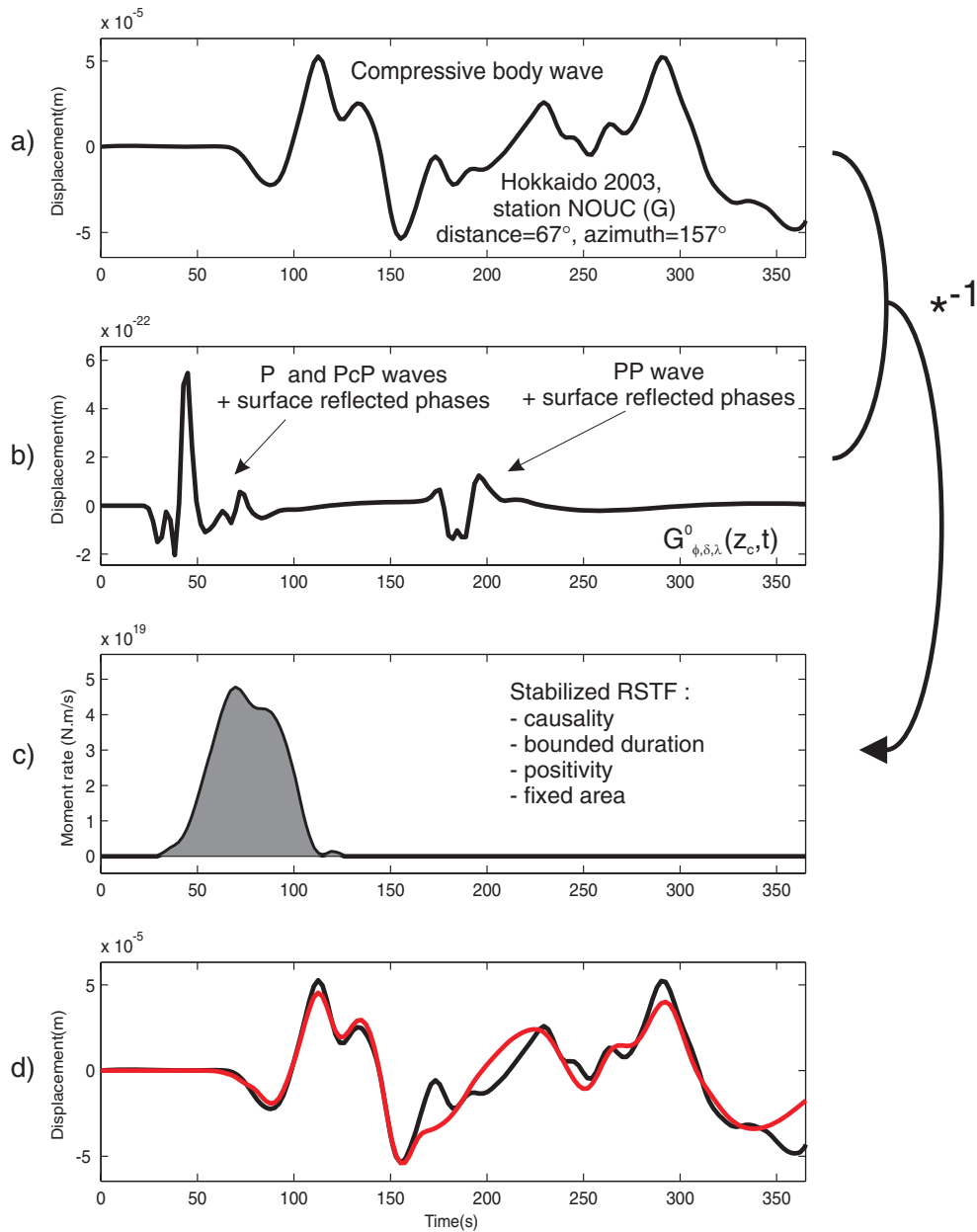
Figure 3. Principle of the deconvolutive approach. (a) Example of teleseismic compressive waveform. The waveform shows the vertical displacement recorded at station NOUC (Geoscope) after the 2003 Hokkaido earthquake. We show the first 300 s after the *P*-wave arrival, bandpass filtered between 0.005 Hz and 0.03 Hz (see filter types in the main text). (b) Theoretical propagation function ($G^0_{\phi,\delta,\lambda}(z_c, t)$) for compressive waves, including *P*, *PcP* and *PP* waves. The seismic source is represented by a double-couple point-source of moment 1 N m s^{-1}. $G^0_{\phi,\delta,\lambda}(z_c, t)$ is computed for $(\phi, \delta, \lambda, z_c) = (251°,\ 22°,\ 129°,\ 35\ \text{km})$, and high-pass filtered at 0.005 Hz. (c) Stabilized deconvolution of (b) from (a), using conditions (i)–(iv) (see main text). Moment magnitude used to constrain the seismic moment (condition iv), is $M_w = 8.15$. The obtained function is the RSTF smoothed at 0.03 Hz. Note that an advance shift has been introduced in $G^0_{\phi,\delta,\lambda}(z_c, t)$ (b), so that the beginning of the RSTF is not too close from the origin time. (d) Comparison between observed waveforms [black; same signal as in (a)], and reconstructed waveforms [red; by convolution between (b) and (c) signals].

complex D'' region, and the latter one crosses two additional times the heterogeneous lithosphere and crust. As a result, these waves have a high-frequency content both less energetic and more difficult to model than the direct waves. Finally we also have a practical constraint, because the computing time for the stabilized deconvolutions depends directly on the number of samples. Considering only low frequencies allows us to reduce the number of samples and to accelerate the deconvolution process.

We take into account the high-frequency limitation by filtering the frequencies higher than 0.03 Hz. To do so, we convolve the data

with f_g, defined as a time-shifted Gaussian function of standard error 4.4 s (which leads to a corner frequency at -3 dB of 0.03 Hz) and time integral equal to 1. The time-shift is selected so that only negligible energy arrives before origin time, making f_g very close to a causal function. Eq. (5) can be written as

$$U(t) \times f_g(t) = G^0_{\phi,\delta,\lambda}(z_c, t) * F(t) * f_g(t). \tag{8}$$

Deconvolving $G^0_{\phi,\delta,\lambda}(z_c, t)$ from $(U(t) * f_g(t))$ therefore gives a more reliable smoothed RSTF. The conditions for the RSTFs defined in Section 2.3 remain valid, as f_g is a causal positive function with

time integral equal to 1. Only the condition (iii) has to be slightly modified, because the obtained RSTF is now bounded at a time larger than T_d, due to the duration of f_g. For the low-frequency limit, a six-pole Butterworth high-pass filter at 0.005 Hz is applied both to the data and to the computed $G^0_{\phi,\delta,\lambda}(z_c, t)$, so that the conditions derived in Section 2.3 remain unchanged.

To optimize the set of parameters $(\phi, \delta, \lambda, z_c)$, we first deconvolve the computed function $G^0_{\phi,\delta,\lambda}(z_c, t)$ for transverse body waves, using stabilizing conditions (i), (ii) and (iii). By integration of the obtained RSTFs at each station, we have independent estimates of the seismic moment. There are several advantages in estimating the seismic moment from transverse body waves rather than from compressive body waves. First, S waves have a lower frequency content than P waves, which make them more sensitive to the zero-frequency seismic moment. Then, transverse S waves have only one local surface reflected phase (sS), which can be of the same polarity as the direct SH wave. Therefore, compared to the compressive waves, they suffer less from the high-pass filtering effect described before. Finally, when looking at the propagation coefficients which relate the focal mechanism to the radiated wavefield, there is no apparent trade-off between focal parameters and seismic moment (see e.g. coefficients $b1$ and $b2$ in Bouchon 1976, p. 523). For compressive body waves, there is a factor (called $a2$ in Bouchon 1976) which depends only on $\sin \lambda \sin 2\delta$. This term becomes predominant when take-off angles approach the vertical direction. In this case, compressive waves suffer from a similar trade-off as low-frequency surface waves, the seismic moment becoming strongly dependent on the focal mechanism parameters.

Once estimated the seismic moment at all stations for transverse body waves, we select its median value (called M_{0m}) and now deconvolve both transverse and compressive waves, using stabilizing conditions (i), (ii), (iii) and (iv). For this last condition, the moment at all stations is constrained to be equal at M_{0m}. The obtained RSTFs are then reconvolved with $G^0_{\phi,\delta,\lambda}(z_c, t)$, and we call the result of this operation U^1. The misfit ϵ_1 between data U and synthetics U^1 is evaluated using the classical variance reduction

$$\epsilon_1 = 1/N \sum_{i=1}^{N} C(i) \frac{\int_{t_0}^{t_0+t_f} \left(U_i^1(t) - U_i(t)\right)^2 \, \mathrm{d}t}{\int_{t_0}^{t_0+t_f} (U_i(t))^2 \, \mathrm{d}t}, \tag{9}$$

where N is the number of stations, C is a weighting factor accounting for the non-homogeneity of the station azimuth distribution, t_f is the fitting duration and t_0 refers to the arrival time of the direct P or SH wave. We evaluate ϵ_1 separately for compressive and transverse waves. In the case of compressive waves, t_f is fixed to the differential time between direct P arrival and PPP arrival, because this latter wave is not taken into account in the analysis. For transverse waves, it is fixed to the differential time between direct S and SS wave. This insures that a duration of at least 210 s is used to determine the fit for each station and each wave type. To take into account condition (v), we first estimate the average F_m of the obtained RSTFs noted F_i^1 for each station i.

$$F_m(t) = 1/N \sum_{i=1}^{N} F_i^1(t). \tag{10}$$

Then we define ϵ_2, measuring the non-similarity of the RSTFs.

$$\epsilon_2 = 1/N \sum_{i=1}^{N} \frac{\int_0^{T_d} \left(F_i^1(t) - F_m(t)\right)^2 \, \mathrm{d}t}{\int_0^{T_d} (F_m(t))^2 \, \mathrm{d}t}. \tag{11}$$

The computation of ϵ_2 is also done separately for compressive and transverse waves. Calling ϵ_1^P and ϵ_1^S, the misfit ϵ_1 computed

for compressive and transverse waves, respectively, and ϵ_2^P and ϵ_2^S, the misfit ϵ_2 computed for compressive and transverse waves, respectively, we define the global misfit ϵ as

$$\epsilon = \left[\epsilon_1^P \left(1. + a^P \epsilon_2^P\right) + W_{PS} \left(\epsilon_1^S \left(1. + a^S \epsilon_2^S\right)\right)\right] / [1 + W_{PS}]. \tag{12}$$

a^P and a^S are chosen, respectively, equal to 2. and 1., to take into account that transverse RSTFs are expected to vary more than compressive RSTFs. Using larger values for a^P and a^S (up to 10 and 5, respectively) has a negligible effect on the results. W_{PS} is taken equal to 0.5, because a precise analysis of transverse waves is more difficult (in particular because the beginning of the signal may be noisy and because a part of the strong SV component may contaminate the signal). The chosen misfit function logically gives more weight to the ϵ_1 terms. The ϵ_2 terms, quantifying the similarity of the RSTFs, are only used as second-order stabilizing constraints. This makes the misfit function very different from most classical source inversions, where the RSTFs are intrinsically the same at each station. Because ϵ_2 terms have a small weight in the computation of ϵ, ϵ can be seen as the weighted average of ϵ_1^P and ϵ_1^S. This makes the values of ϵ directly interpretable as classical variance reduction values (i.e. $\epsilon = 0$ corresponds to a perfect reconstruction of the waveforms and $\epsilon = 1$ to the null hypothesis). Using the misfit function ϵ, and $(\phi, \delta, \lambda, z_c)$ as inversion parameters, the optimal set of parameters is determined by the Neighbourhood Algorithm (NA, Sambridge 1999). ϕ, δ and λ are, respectively, allowed to vary in the $[0°–360°]$, $[0°–90°]$ and $[−180°–180°]$ ranges. z_c can freely vary between $(z_n − 50)$ km and $(z_n + 50)$ km, where z_n is the event depth (in kilometres) retrieved in the NEIC catalogue. If $z_n − 50$ is smaller than 12, the minimal depth considered in NA is fixed at 12 km, as in the GCMT method. The main steps of the optimization procedure are summarized in Fig. 4. We hereafter refer to this approach as the SCARDEC method (from 'Seismic source ChAracteristics Retrieved from DEConvolvolvin g teleseismic body waves').

2.4.2 Dip, depth and moment uncertainties

Body wave analysis is expected to have a good dip and depth resolution because the take-off angles sample well the central part of the focal sphere and because the time arrival of surface-reflected phases are directly related to depth. We can verify this by computing the misfit variation when dip and depth vary around their optimal values. Fixing the strike and rake to their optimal values, we compute the misfit corresponding to depths at ±30 km around the optimal value and dips at $\pm 15°$ around the optimal value. Examination of this bi-dimensional misfit function for a broad range of earthquakes has shown us that in general the misfit varies little close to the optimal parameter set. However, when parameter values significantly differ from the optimal combination, the misfit value begins to increase sharply. We have observed that the change between these two behaviours occurs when the misfit function is about 10 per cent larger than its optimal value (see also the next section for actual examples). We thus consider that the acceptable parameters are those leading to misfit values not exceeding the optimal value by more than 10 per cent. The parameter range defined by this uncertainty analysis gives us information on the resolution of the SCARDEC method. Additionally, this analysis allows us to assess the sensitivity of the seismic moment to these acceptable variations of dip and depth.
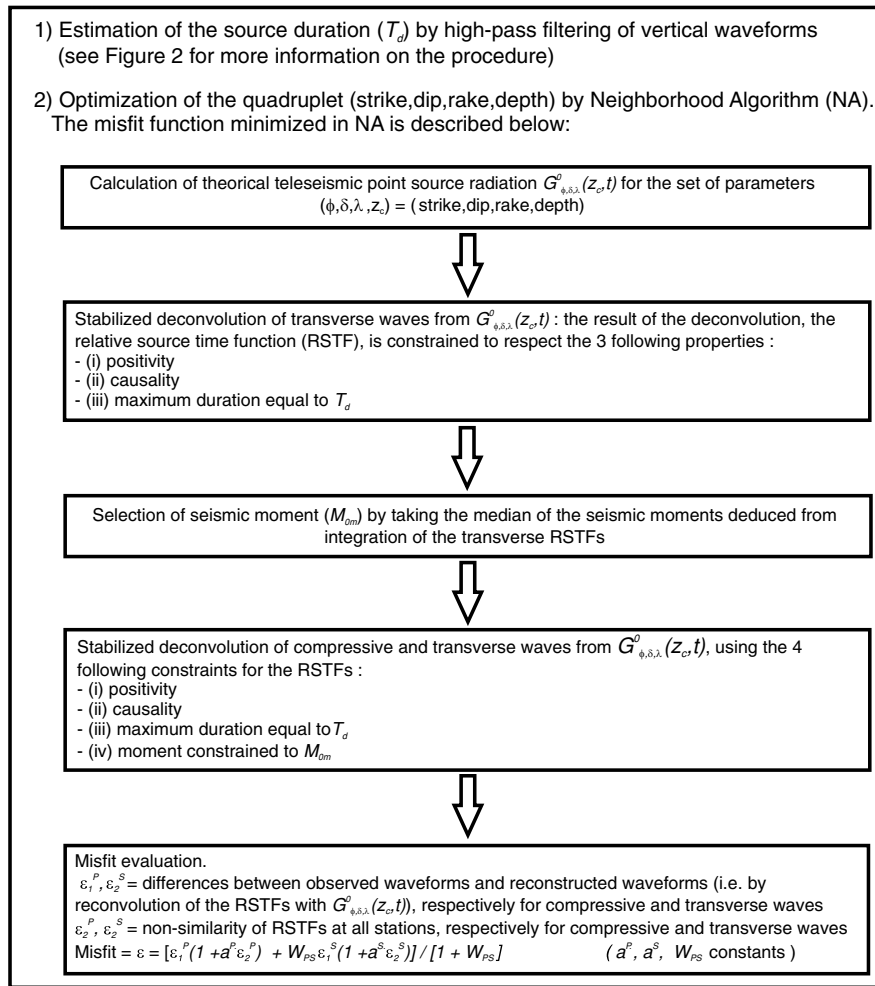
1) Estimation of the source duration (T_d) by high-pass filtering of vertical waveforms
   (see Figure 2 for more information on the procedure)

2) Optimization of the quadruplet (strike,dip,rake,depth) by Neighborhood Algorithm (NA).
   The misfit function minimized in NA is described below:

Calculation of theorical teleseismic point source radiation $G^0_{\phi,\delta,\lambda}(z_c,t)$ for the set of parameters
$(\phi,\delta,\lambda,z_c) = ($strike,dip,rake,depth$)$

⬇

Stabilized deconvolution of transverse waves from $G^0_{\phi,\delta,\lambda}(z_c,t)$: the result of the deconvolution, the
relative source time function (RSTF), is constrained to respect the 3 following properties :
- (i) positivity
- (ii) causality
- (iii) maximum duration equal to T_d

⬇

Selection of seismic moment (M_{0m}) by taking the median of the seismic moments deduced from
integration of the transverse RSTFs

⬇

Stabilized deconvolution of compressive and transverse waves from $G^0_{\phi,\delta,\lambda}(z_c,t)$, using the 4
following constraints for the RSTFs :
- (i) positivity
- (ii) causality
- (iii) maximum duration equal to T_d
- (iv) moment constrained to M_{0m}

⬇

Misfit evaluation.
$\varepsilon_1^P, \varepsilon_1^S$ = differences between observed waveforms and reconstructed waveforms (i.e. by
reconvolution of the RSTFs with $G^0_{\phi,\delta,\lambda}(z_c,t)$), respectively for compressive and transverse waves
$\varepsilon_2^P, \varepsilon_2^S$ = non-similarity of RSTFs at all stations, respectively for compressive and transverse waves
Misfit = $\varepsilon = [\varepsilon_1^P(1 + a^P\varepsilon_2^P) + W_{PS}\varepsilon_1^S(1 + a^S\varepsilon_2^S)] / [1 + W_{PS}]$ (a^P, a^S, W_{PS} constants)

Figure 4. Flowchart explaining the principles of the SCARDEC method: diagram of moment magnitude, focal mechanism and depth optimization.

3 APPLICATION TO MAJOR SUBDUCTION EARTHQUAKES IN THE PERIOD 1990–2010

3.1 Detailed results for one event: the 2003 Hokkaido earthquake

We first detail the results for the 2003 September 25 Hokkaido earthquake. This earthquake is particularly interesting, because it is one of the very few major subduction earthquakes which was recorded and analysed with a large amount of seismological and geodetical data (see following sections).

The results obtained for this earthquake are presented in Fig. 5 for the source model and its uncertainties, and in Fig. 6 for the agreement between data and synthetics. The optimization process of minimizing ϵ has lead to determine $\phi = 251°$, $\delta = 22°$, $\lambda = 129°$ and $z_c = 35$ km. The magnitude associated with this mechanism and depth is $M_w = 8.15$. The figures show that, with this optimal focal mechanism and depth, the RSTFs respecting the physical conditions (i), (ii), (iii) and (iv) are able to explain well the teleseismic displacement data ($\epsilon = 0.104$). The RSTFs for the various stations are similar, but clearly not identical. For example, a clear feature is that RSTFs in southeastern azimuths (i.e. stations PPT, RAR, NOUC and CTAO) are less impulsive than in northwestern azimuths (i.e.

ABKT, GNI and MLR). This characteristic agrees well with detailed studies of this earthquake (Koketsu *et al.* 2004; Yagi 2004), which have shown that the rupture propagation of the Hokkaido earthquake was mainly in the downdip direction. This observed variability also gives an insight of the interest of the SCARDEC method compared to classical point source techniques. The use of these latter methods, which intrinsically impose the equality of the RSTFs, are expected to introduce biases in the determination of focal mechanism and magnitude. In fact, the use of a unique RSTF would reduce the agreement between data and synthetics. Because of this reduced fit, the reliability of the solution should decrease.

The estimation of dip and depth uncertainties can be seen in the bottom-left-hand side of Fig. 5. Considering that the acceptable solutions are inside the area where misfit is smaller than 1.1 times its optimal value (see Section 2.4.2), we determine that dip and depth are respectively equal to $22 \pm 3°$ and 33 ± 8 km. The extreme values of magnitude associated with the acceptable dip and depth variability are 8.12 and 8.16.

Strike and rake are found very close to GCMT values ($\phi = 250°$, $\lambda = 132°$). Depth for the optimal model is deeper than GCMT (35 km versus 28 km), but if we take into account the uncertainties, we see that the depth of 28 km is acceptable. However, even with the uncertainties, we find that dip and magnitude differs from
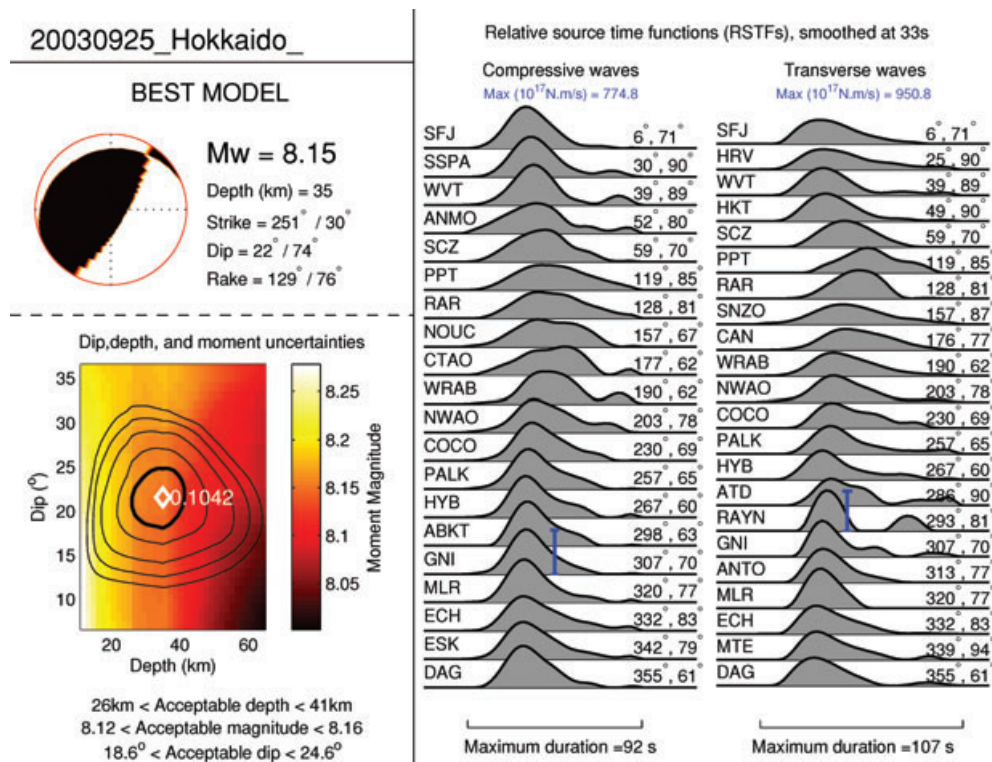
Figure 5. Source parameters, uncertainties and RSTFs. (Top left-hand side) Optimal values of moment magnitude, depth and focal mechanism. (Bottom left-hand side) Uncertainty analysis: misfit and moment magnitude changes as a function of dip and depth variations around their optimal values. Optimal dip and depth are indicated by the white diamond (the best misfit value is also shown). The thick line is the iso-misfit contour (noted C1) joining points with misfit 10 per cent larger than the best value. The four thin lines are the iso-misfit contours joining points with misfit 25 per cent, 50 per cent, 75 per cent and 100 per cent larger than the best value. Note that the observation of these misfit contours shows well the bell-shaped form of the misfit function, with a flat minimum surrounded by a sharp increase of the misfit. Moment magnitude associated with each (dip-depth) couple is shown with the colour scale. Acceptable values of dip, depth and magnitude are those which are inside the C1 contour. (Right) Relative source time functions (RSTFs) for compressive and transverse waves. These RSTFs are smoothed at 33 s (see main text) so that their durations are longer than the actual ones. The indicated maximum values correspond to the absolute maximum of all the moment rates, respectively, for compressive and transverse RSTFs. The corresponding scale is indicated by the blue bars, which are plotted next to the location of the maximal RSTF. For each RSTF, the name of the station, its azimuth and epicentral distance are shown.

GCMT. Dip is found 8–14° steeper than CMT and moment magnitude 0.11–0.15 smaller than GCMT. We show in the following paragraphs that other earthquakes share this property of a steeper dip associated with a smaller magnitude.

3.2 Global results

Results for the 17 studied earthquakes are presented in Table 3. Individual results—presented in a similar way as in Figs 5 and 6 for the 2003 Hokkaido earthquake—can be found in the Supplementary Figs 1 to 16. Considering the uncertainties, we observe a good depth agreement with GCMT. On the other hand, there are differences in strike and rake, up to 30° (event 7, Andreanof 1996 and event 13, Solomon 2007), for some earthquakes. The variations of these two parameters are not uncorrelated because the value $(\phi - \lambda)$ is much more consistent between GCMT and SCARDEC method. This is expected as body waves, having their take-off angle close to the vertical, cannot detect very accurately if there is a small strike-slip component in these shallow-dip thrust earthquakes. Sensitivity tests show however that the uncertainty should not be larger than ±15° for a 20° dipping fault. Differences larger than this uncertainty are thus thought to be meaningful, which is consistent with detailed studies of the 1996 Andreanof and 2007 Solomon earthquakes. In the first case, both the trench geometry and the study of Kisslinger & Kikuchi (1997) indicate that the fault strike is between the GCMT

strike and the strike retrieved here. In the second case, the studies of Furlong *et al.* (2009) and Chen *et al.* (2009), as well as the trench geometry, show a fault strike very close to our determination.

The other clear difference with GCMT concerns the moment magnitude and dip. This latter parameter is reliably retrieved by body wave analysis because it is very sensitive to waves with take-off angles close to vertical. For about half of the studied earthquakes (Jalisco 1995, Kuril 1995, Minahassa 1996, Andreanof 1996, Peru 2001, Hokkaido 2003, Sumatra 2005 and Sumatra 2007), we clearly determine a steeper dip, associated with a smaller magnitude, than GCMT. Other earthquakes also indicate a similar behaviour, but given the uncertainties, the solutions remain consistent with GCMT. Dip angle comparisons, including uncertainties, are presented in Fig. 7. The observed differences may be due to the well-known trade-off between magnitude and dip affecting the GCMT results. We recall that the product $M_0 \sin 2\delta$ can be accurately resolved but that the relative weight of the two factors remains much less known. This means that a larger M_0 (and thus a larger M_w) with a smaller δ, or reciprocally a smaller M_w with a larger δ are plausible solutions.

To quantitatively evaluate if SCARDEC solutions are consistent with the expected trade-off, we can compare the obtained magnitude with a corrected GCMT magnitude, called $M_w'^c$ and expressed as

$$M_w'^c = 2/3\log\left(\frac{M_0^c \sin 2\delta^c}{\sin 2\delta^d}\right) - 6.06\,, \tag{13}$$

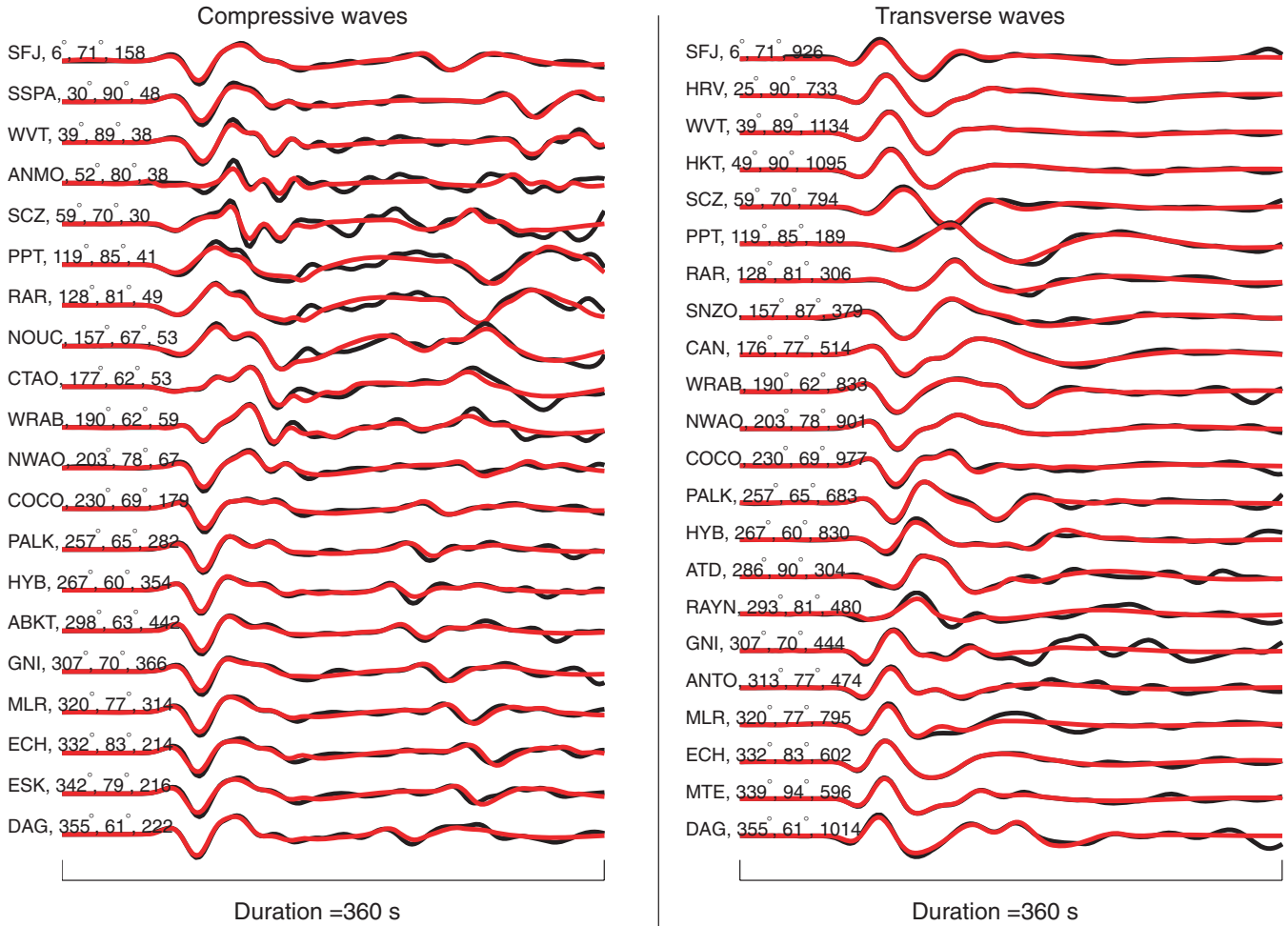## Agreement between displacement data (black) and synthetics (red)

Figure 6. Agreement between data (black) and synthetics (red) for compressive waves (left-hand side) and transverse waves (right-hand side). For each station and wave type, synthetics are obtained from the convolution between $G^0_{\phi,\delta,\lambda}(z_c, t)$ and the obtained RSTF. The name of the station, its azimuth and distance, and the displacement maximum absolute value (in micrometres) of each signal are also shown.

where M_0^c is the GCMT seismic moment in N.m and δ^c and δ^d are the dips retrieved by GCMT and SCARDEC method, respectively. To be consistent with the $M_0 \sin 2\delta$ dependency, we should have $M_w^d = M_w^{\prime c}$, where M_w^d is the magnitude found in the present analysis. As there is a clear magnitude dependency on the earthquake's depth (see Fig. 5 and Supplementary Figures), it is more consistent to compare M_w^c, M_w^d and $M_w^{\prime c}$ for the same depth. Because the GCMT depths are inside or very close to the error bars of the depths determined in this study, we select M_w^d as the moment magnitude calculated at the GCMT depth (keeping the other three optimal parameters of the deconvolution, namely ϕ, δ, λ). Fig. 8(a) first shows the direct magnitude comparison between M_w^d and M_w^c. We see that there is some dispersion around the $x = y$ line, particularly for high magnitudes (>8.1), where $M_w^c > M_w^d$. In Fig. 8(b), where M_w^d is now plotted against $M_w^{\prime c}$, the dispersion is much smaller, and earthquakes are well aligned along the $x = y$ line. While the average difference between M_w^d and M_w^c is 0.095, the average difference between M_w^d and $M_w^{\prime c}$ is only 0.044. This indicates that a large part of the differences between the GCMT and the SCARDEC method can be explained by the trade-off affecting the low-frequency surface wave analysis. We note that M_w^d tends

to slightly overestimate $M_w^{\prime c}$ (average overestimation equal to 0.03) and attribute this effect to the slight overestimation of the source duration (see Section 2.2), which may cause some late signals in the RSTFs.

After correction of the $M_w - \delta$ trade-off, the main remaining differences may also be explained. Only two earthquakes show a difference between M_w^d and $M_w^{\prime c}$ larger than 0.09: the 1996 Minahassa earthquake (event 5) and the 2007 Peru earthquake event 14. For the first one, the dip determined by GCMT is very small (6°) so that $M_w^{\prime c}$ is very sensitive to δ^d. Taking δ^d equal to 10°, which is a value inside the uncertainties we estimated, would make M_w^d and $M_w^{\prime c}$ consistent. The 2007 Peru earthquake is a long-duration earthquake with respect to its magnitude (see Table 2). This suggests that the choice of a magnitude-dependent half-duration causes the GCMT solution to underestimate the moment magnitude.

The usual explanation of the underestimation of seismic moment by body-wave analysis invokes low-frequency source processes, which would be better resolved by the lower frequency surface waves. However, there is no real theoretical reason for this assertion, at least when source duration is significantly shorter than the longest

Table 3. Comparison between SCARDEC results and GCMT source parameters. The first column shows the index of each earthquake (see Table 1). Strike (°), dip (°), rake (°), depth (km) and moment magnitude (ϕ, δ, λ, z_c, M_w) are given for both approaches, respectively. We also provide the acceptable ranges for dip, depth and moment magnitude (respectively $\Delta\delta$, Δz_c, ΔM_w) determined by our uncertainty analysis.

|	Global CMT					SCARDEC							
n^0	ϕ	δ	λ	z	M_w	ϕ	δ	λ	z	M_w	$\Delta\delta$	Δz	ΔM_w
1	278	7	89	15	7.76	291	10	105	30	7.63	8-12	13-42	7.57–7.70
2	354	22	87	29	8.00	17	24	115	30	8.07	22-25	24-36	8.07–8.07
3	302	9	92	15	7.98	312	20	99	13	7.80	18-23	0-17	7.77–7.82
4	225	12	95	26	7.88	240	21	115	19	7.82	17-25	13-28	7.79–7.86
5	36	6	54	15	7.87	38	15	59	27	7.67	9-19	18-36	7.66–7.71
6	103	11	69	15	8.19	84	15	53	12	8.10	11-18	0-18	8.06–8.14
7	248	17	84	29	7.88	273	25	116	18	7.82	22-31	13-27	7.80–7.85
8	202	23	74	34	7.76	215	20	88	32	7.81	17-23	21-41	7.79–7.83
9	310	18	63	30	8.39	307	29	59	35	8.36	26-33	26-43	8.34–8.37
10	250	11	132	28	8.26	251	22	129	35	8.15	19-25	26-41	8.12–8.16
11	333	8	118	26	8.62	327	14	105	30	8.46	12-17	21-39	8.44–8.47
12	215	15	92	14	8.30	205	17	83	12	8.25	13-19	0-12	8.25–8.28
13	333	37	121	14	8.07	304	33	65	19	8.06	29-35	15-28	8.04–8.09
14	321	28	63	34	7.97	324	28	69	33	8.12	22-33	21-44	8.10–8.14
15	328	9	114	24	8.49	331	16	112	19	8.35	12-20	13-31	8.33–8.39
16	25	26	138	23	7.78	37	29	147	35	7.72	24-34	22-41	7.67–7.74
17	19	18	116	23	8.79	24	21	119	35	8.74	18-25	25-40	8.72–8.74
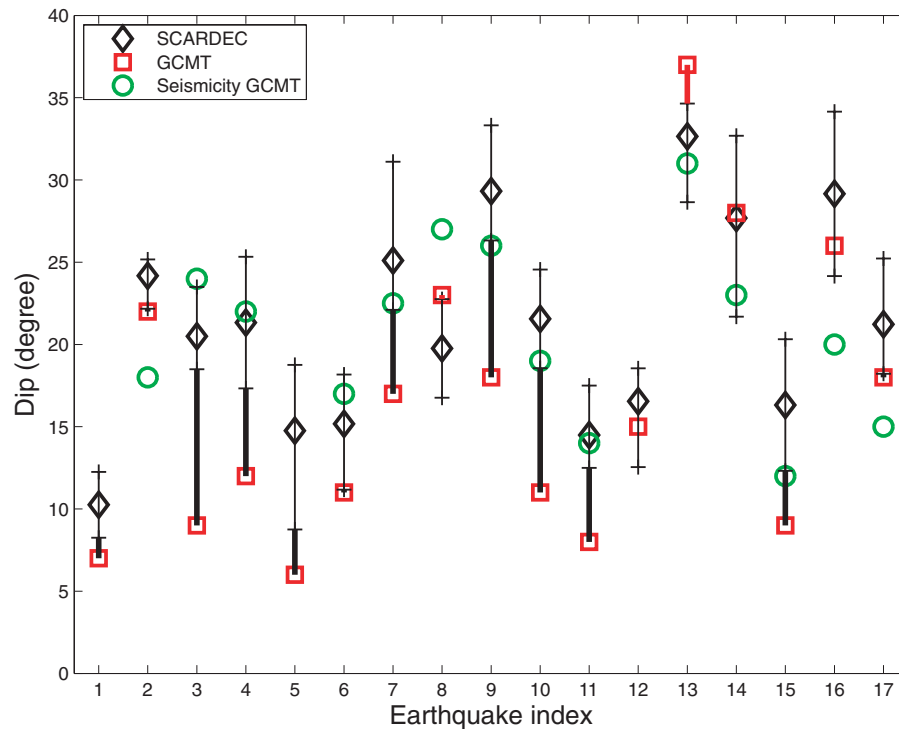
Figure 7. Earthquake fault dip comparisons. For each earthquake (see correspondence between indices and earthquakes in Tables 1 or 2), we show the best dip found by SCARDEC method (black diamond) and by GCMT (red square). Extreme values determined by our uncertainty analysis are shown by the '+' signs, so that the possible dips are along the thin black line joining these '+' signs. When existing, the thick lines indicate the discrepancy between GCMT and SCARDEC dip values; black lines indicate that we retrieve a dip steeper than GCMT, whereas red lines indicate the opposite. Green circles show the median dip values inferred by GCMT for moderate-to-large seismicity in the same region and period of occurrence as the main shock (see Section 5). Three earthquakes do not have enough foreshocks or aftershocks to define this independent information.

period present in the seismograms. If $G_{\phi,\delta,\lambda}(x, z_c, \omega)$ is correctly estimated, the deconvolution of this term from U gives the broad-band RSTF, from which the moment can be directly calculated. Moreover, if this intrinsic underestimation of seismic moment by body waves was true, it would subsist even after the sin 2δ factor correction.

4 AGREEMENT BETWEEN SCARDEC BODY-WAVE SOLUTIONS AND LONG-PERIOD SURFACE WAVE DATA

To further validate the moment magnitudes and focal mechanisms determined in this study it is important to test if they can explain data
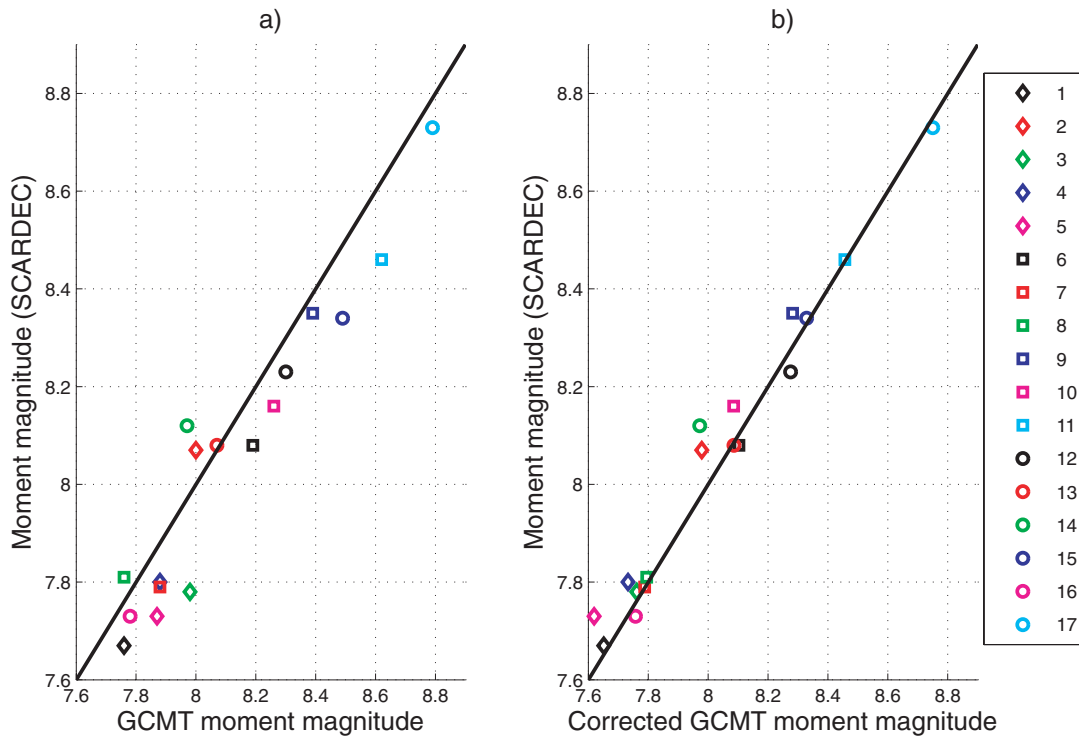
a)   b)

Figure 8. Effect of the $M_w - \delta$ trade-off on the differences in moment magnitude between the GCMT and SCARDEC methods. (a) Direct comparison between the SCARDEC and GCMT moment magnitudes. (b) Comparison between the SCARDEC moment magnitude and the corrected GCMT magnitude, taking into account the $M_w - \delta$ trade-off (see the expression of the corrected magnitude in the main text). The SCARDEC moment magnitude is the magnitude computed for the same depth as GCMT, as explained in the text. In both cases, the black line shows the $x = y$ line, where there is a perfect agreement between both magnitude estimates. The agreement clearly improves when we take into account the trade-off. Each earthquake is represented by a symbol referring to the indices shown in the right part of the figure (see correspondence between earthquakes and indices in Tables 1 or 2).

that were not used to constrain them, notably long-period surface wave data. In this section we compare real long-period surface wave seismograms with theoretical seismograms calculated using our new seismic source parameters.

We calculate synthetic seismograms for long-period ($T \geq 40$ s), three-component fundamental mode, minor-arc, surface waves using a full ray theory approach (e.g. Ferreira & Woodhouse 2007). We use the 3-D mantle model S20RTS (Ritsema *et al.* 1999) combined with the global crust model CRUST2.0 (Bassin *et al.* 2000). We calculate seismograms using different point source models: (i) GCMT source parameters and (ii) the centroid latitude, longitude and origin time reported by the GCMT, the depth as determined in this study and a moment tensor calculated from the seismic moment and fault geometry determined in this study, assuming a pure double-couple mechanism; we consider a variety of possible source models by taking into account the determined uncertainties in depth, dip and moment magnitude (see Section 2.4.2) and refer to them as SCARDEC models. In both cases, a triangular STF is used with a half-duration as reported in the GCMT catalogue.

To test how well these different seismic source models explain long-period surface waves, we compare the synthetic seismograms with real broad-band data from the FDSN. Instrument response deconvolution is conducted on the seismograms and the horizontal components are rotated into longitudinal and transverse directions for each earthquake. The data are convolved with the response of an SRO instrument and low-pass cosine tapered to capture the low-frequency characteristics of the signal (typically between $T = 150$–200 s, depending on the particular earthquake).

Figs 9 and 10 compare synthetic seismograms (red, green) with real data (black) recorded at various stations from the FDSN, following the 2003 September 25 Hokkaido earthquake (see earthquake number 10 in Table 3 of this paper). For this earthquake, SCARDEC predicts a steeper fault than in the GCMT model by $8°-14°$ and a moment magnitude of $M_w = 8.12 - 8.16$ rather than the magnitude $M_w = 8.26$ reported in the GCMT catalogue. The synthetics in red are calculated for the GCMT source model, whereas the synthetics in green correspond to a SCARDEC model with the optimal strike, dip and rake, with a depth of 41 km and a magnitude of 8.12. The synthetic seismograms calculated using the SCARDEC source model explain the phase of the long-period Rayleigh waves as well as the GCMT model. Moreover, for Rayleigh waves, the SCARDEC model explains the amplitude data slightly better than the GCMT model, notably for stations WVT, KIP, PPT, COCO and PALK (Fig. 9). For Love waves, the GCMT model explains the data slightly better than the SCARDEC model, particularly for stations RAR, ARU, MLR and MORC (Fig. 10).

We quantify the fit between synthetics and data by measuring both phase shifts and amplitude ratios between synthetic and real surface wave data in the time domain. A time window is selected centred on the maximum amplitude of the desired wave train, with its edges at zero-crossings of the seismograms, to minimise errors in the measurements. A non-linear least-squares algorithm calculates the phase shift and amplitude factor that best fits the synthetic waveform to the real seismogram. Moreover, we calculate the waveform misfit $m^2 = \frac{(\mathbf{s}-\mathbf{d})^2}{\mathbf{d}^T \mathbf{d}}$ also in the time domain, where $\mathbf{s}$ are the theoretical seismograms and $\mathbf{d}$ are the data. Table 4 shows the average phase,
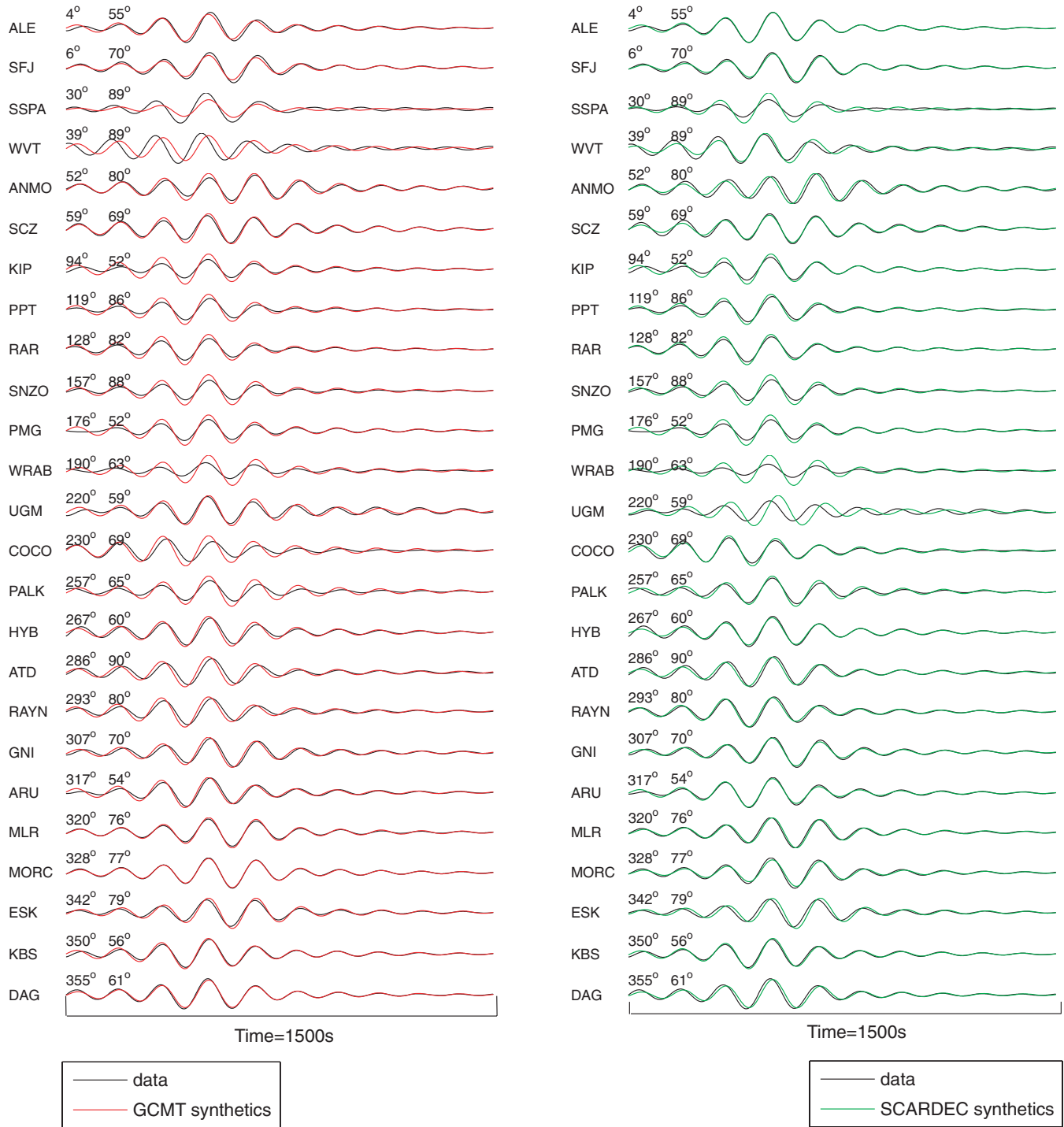
Figure 9. Comparison of vertical component observed Rayleigh waves (black) with theoretical seismograms (red, green) at various stations of the FDSN, following the 2003 Hokkaido earthquake. The name of each station is shown in the left of the waveforms and the corresponding source–receiver azimuth and epicentral distance are shown in the top, respectively. The synthetic seismograms are calculated for the earthquake source parameters in the GCMT catalogue (red) and for the parameters in the SCARDEC model (green; see main text for details). All traces have been deconvolved from instrumental response followed by convolution with the response of an SRO instrument and low-pass cosine tapered around T = 150 s.

amplitude and waveform misfits between data and synthetics over all the stations, for the GCMT and SCARDEC source models for the 2003 Hokkaido earthquake. It is clear that the differences in misfits are small, so that overall the GCMT and SCARDEC source models

explain the long-period surface wave data equally well. Thus, for the Hokkaido earthquake, a source model with a fault dip angle of 11° and moment magnitude $M_w = 8.26$ (as in the GCMT catalogue) is as compatible with long-period surface waves as a fault

LHT
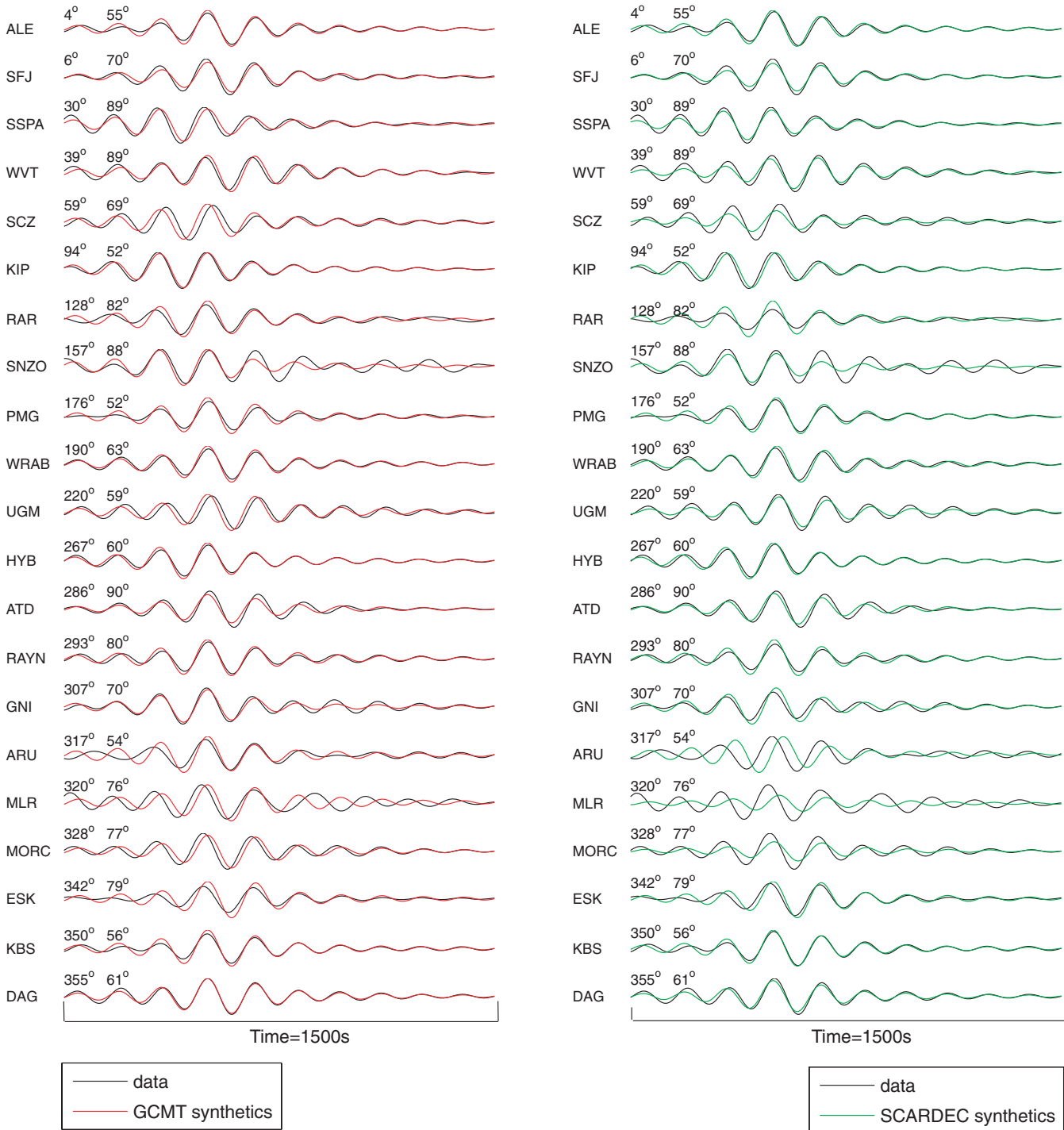
Figure 10. Same as in Fig. 9, but for transverse component Love waves.

dip angle of 22° together with a moment magnitude of $M_{\rm w} = 8.12$. This clearly illustrates the trade-off between the fault's dip angle and the seismic moment for shallow earthquakes when determining these parameters using long-period surface waves, as explained in previous sections. To further verify our comparisons, we also calculated theoretical seismograms using the spectral element method (Komatitsch & Tromp 2002) for the GCMT and SCARDEC source models and compared them with real data, obtaining very similar results to those for full ray theory synthetics.

We carried out these comparisons between real data and synthetics for all the studied earthquakes for which the GCMT parameters are not within the range of acceptable moment magnitude and/or fault dip determined in this study. We found that in all cases the conclusions were similar to those for the 2003 Hokkaido earthquake,

Table 4. Average of phase ($\delta\psi$), amplitude (δA) and waveform (m^2) misfits between three-component long-period surface wave synthetic seismograms and data for the stations in Figs 9–10, for the source models GCMT and SCARDEC for the 2003 September 25 Hokkaido earthquake (see text for details). Perfect fit corresponds to $\delta\psi = 0$ s, $\delta A = 1$ and $m^2 = \frac{(\mathbf{s}-\mathbf{d})^2}{\mathbf{d}^T\mathbf{d}} = 0$, where **s** are the theoretical seismograms and **d** are the data.

	$\delta\psi$ (s)		δA		m^2	
	GCMT	SCARDEC	GCMT	SCARDEC	GCMT	SCARDEC
Z	6.7	6.3	0.89	0.90	0.15	0.16
L	6.8	5.9	0.91	0.97	0.10	0.08
T	8.1	8.6	0.98	1.18	0.12	0.13

Table 5. Same as in Table 4, but for the 2005 March 28 Sumatra earthquake (see text for details).

	$\delta\psi$ (s)		δA		m^2	
	GCMT	SCARDEC	GCMT	SCARDEC	GCMT	SCARDEC
Z	11.4	10.8	0.79	0.98	0.23	0.20
L	12.0	10.4	0.81	0.89	0.35	0.25
T	7.7	8.2	0.81	0.99	0.30	0.32

that is, overall the earthquake source parameters determined in this study explain long-period surface waves as well as the parameters reported in the GCMT catalogue. We show a second example of long-period surface wave comparisons for the 2005 March 28 Sumatra earthquake (see earthquake number 11 in Table 3 of this paper). Supplementary Figs 17 and 18 show waveform comparisons between GCMT synthetics and those calculated using the best-fitting SCARDEC model, and the corresponding misfits are presented in Table 5. The SCARDEC model explains the phase data as well as the GCMT model, with a slight overall improvement in the amplitude fit as shown in Table 5. This better agreement can be seen, for example, for Rayleigh waves recorded at stations ANTO, ECH and ESK and for Love waves recorded at stations ESK, OBN and KBS (Supplementary Figs 17 and 18). This shows that the optimal moment magnitude $M_\mathrm{w} = 8.46$ for the 2005 March 28 Sumatra earthquake determined in this study is as compatible with long-period surface wave data as the larger moment magnitude $M_\mathrm{w} = 8.62$ reported in the GCMT catalogue, despite of the fact that long-period surface waves are not used in this study to retrieve earthquake moment magnitude and focal mechanism.

5 DIP AND MAGNITUDE OF MAJOR SUBDUCTION EARTHQUAKES

We have shown in the two previous sections that the source parameters deduced from a broad range of body waves (including *P, PcP, PP, SH* and *ScS* waves) explain long-period surface waves as well as the GCMT source parameters. In this section, we compare our results with other sources of information available for these major earthquakes.

In the list of the studied earthquakes, the 2003 Hokkaido earthquake is by far the best instrumented event. A dense array of accelerometers and GPS, located along the Japan coast, recorded well the local ground motion. Several studies used these data to provide independent estimates of magnitude and focal mechanism. Yagi (2004) used both teleseismic and strong motion data to determine a moment magnitude $M_\mathrm{w} = 8.0$ associated with a dip of 20°. Using

only strong motion data, Honda *et al.* (2004) have found a similar mechanism, with a dip of 18°. Koketsu *et al.* (2004) have successfully modelled both strong motion and GPS data using the 20° dip retrieved by Yamanaka & Kikuchi (2003). Miyazaki *et al.* (2004) analysed only high rate GPS data and have also found a dip equal to 20° and a moment magnitude of 8.1. In all these studies, only Honda *et al.* (2004) found a moment magnitude close to GCMT ($M_\mathrm{w} = 8.25$). All the other analyses have determined a moment magnitude between 8 and 8.15. We also have information on the interplate geometry based on aftershock relocation. Using OBS data, Machida *et al.* (2009) have simultaneously estimated the aftershock hypocentres and the local 3-D velocity model. This analysis reveals that the angle of the dipping plate is equal or steeper than 16° in the source area of the 2003 earthquake. Gathering the available information, we find a magnitude-dip couple closer to the SCARDEC results ($M_\mathrm{w} = 8.14 \pm 0.02$; $\delta = 22 \pm 3°$) than to the GCMT parameters ($M_\mathrm{w} = 8.26$; $\delta = 11°$).

To a lesser extent, there is also interesting independent information for the 2005 Sumatra (Nias) earthquake. This earthquake was recorded by continuous GPS located in Sumatra and in islands (Simeulue, Nias) above the rupture plane. There are also data coming from coral uplifts. Konca *et al.* (2007) used GPS and coral data together with teleseismic waves (body waves and normal modes) to determine the rupture process of the 2005 Sumatra earthquake. These authors suggest that the combination of normal mode and geodetic data gives a good resolution on the magnitude-dip couple. Once possible ranges of magnitude and dip angle are estimated by normal-mode data analysis, geodetic data are used to determine the most appropriate magnitude value, which suppresses the $M_\mathrm{w} - \delta$ trade-off. A drawback of this approach is that the rigidity structure around the earthquake fault must be well known, which is generally difficult in remote subduction zones. Konca *et al.* (2007) report that a fault plane with dip equal or steeper than 12° would lead to a too small magnitude to explain the geodetic data. However when looking at their selected rigidity structure, we observe that most part of the coseismic slip is located below 22 km depth, in a region where the rigidity is high (68.5 GPa, typical of upper-mantle values). However, it is very likely that for a major interplate earthquake, the rigidity is actually between crustal ($\sim$30 GPa) and upper-mantle values. Thus, the rigidity selected by Konca *et al.* (2007) is probably an upper bound of the realistic rigidity. Choosing smaller rigidity values would make steeper dips acceptable. Interestingly, Kreemer *et al.* (2006) have also analysed the coseismic GPS displacements to retrieve the coseismic slip on the fault. In their fault geometry model, they allow the dip to vary from 8° at the surface to 23° at 50 km depth. They can explain well the GPS vectors with a moment magnitude of 8.37, calculated in a medium with a crustal rigidity of 30GPa. This moment magnitude would be equal to 8.61 in a 68.5 GPa rigidity structure, which agrees with the results of Konca *et al.* (2007). These two studies show that SCARDEC results for the 2005 Sumatra earthquake ($M_\mathrm{w} = 8.45 \pm 0.02$; $\delta = 15 \pm 3°$) are realistic.

For the other earthquakes (Jalisco 1995, Kuril 1995, Minahassa 1996, Andreanof 1996, Peru 2001 and Sumatra 2007) where we obtain clear differences with GCMT, there are fewer independent estimates of the moment magnitude. For the first five ones, we can mainly compare our results with other studies analysing teleseismic body waves. Interestingly, most studies that refine the GCMT mechanism using their own modelling generally obtain steeper dips than GCMT. This is the case of the study of Mendoza & Hartzell (1999) for the 1995 Jalisco earthquake in which they found that a dip of 14° explains data better that the 9° GCMT value. Similarly, Shao

& Ji (2007) have modelled the 1995 Kuril earthquake with a dip of 18° (to be compared to the 12° GCMT value). The optimal focal mechanism of Kisslinger & Kikuchi (1997) for the 1996 Andreanof earthquake also shows a steeper dip than GCMT (21° versus 17°). For the 1996 Minahassa earthquake, the difference between the dip determined by Gomez *et al.* (2000), equal to $7 \pm 3°$, and GCMT (6°) is small. The 2001 Peru earthquake dip was found steeper than GCMT by Kikuchi & Yamanaka (2001) and Bilek & Ruff (2002) (respectively, by 3° and 5°). A counterexample exists for this 2001 Peru earthquake, where Giovanni *et al.* (2002) have assumed a dip of 14° (compared to the GCMT value of 18°), but without detailing the reason of this choice. The 2007 Sumatra earthquake has been analysed both with geodetic and teleseismic data. Yagi (2007) used teleseismic data, obtaining a dip of 18°, which is twice the GCMT dip value. Konca *et al.* (2008) have successfully modelled teleseismic and geodetic data with a 15° dip plane. The same dip value has been retrieved by Yamanaka (2007). Among the studies of these six earthquakes, the study of Bilek & Ruff (2002) for the 2001 Peru earthquake is the only one to find a moment magnitude very close to GCMT. All other analyses have determined a moment magnitude 0.05–0.28 smaller than GCMT.

A last external information comes from the focal mechanisms of moderate-to-large earthquakes ($5.5 < M_w < 7.2$) occurring in the vicinity of the main shocks. In this magnitude range, GCMT makes also use of body waves so that the $M_w - \delta$ trade-off reduces. Hjörleifsdóttir & Ekström (2010) have recently confirmed, using synthetic data computed in a realistic Earth, that GCMT results are close to the real source parameters when both body and surface waves are used. Assuming that thrusting foreshocks and aftershocks occur on the same interplate plane as the main shock, we get another independent information on the fault geometry. For each of the large subduction earthquakes studied, we retrieve in the GCMT catalogue the earthquakes satisfying the following criteria (z_m is the centroid depth of the main shock):

(1) thrust mechanism,
(2) moment magnitude between 5.5 and 7.2,
(3) origin time between 1 month before the main shock and 3 months after the main shock,
(4) epicentral location within two degrees in latitude and longitude compared to the main shock's centroid and
(5) depth larger than $(z_m - 20)$ km and smaller than $(z_m + 5)$ km.

This last criterion has been selected to exclude earthquakes considerably deeper than the main shocks, for which it can be argued that their steeper dips are simply due to the bending at depth of the subducting plate. Considering this same bending plate hypothesis, we would expect that this dissymmetric depth criterion would lead to some underestimation of the main shock dip. If, for a given earthquake, the selection includes at least two earthquakes, we take the median dip value (noted δ_a) as an estimate of the local fault geometry. Three earthquakes (Java 1994, Minahassa 1996 and Kuril 2006) have at most one suitable foreshock or aftershock and thus cannot be considered here. The median values δ_a for all other studied earthquakes have been represented in Fig. 7 (green circles), along with the GCMT main shock dip (red squares). For nine over 14 earthquakes, δ_a is found steeper than the GCMT main shock dip, in spite of the dissymmetric depth criterion. The average difference between δ_a and GCMT dip is 6.4°, while the difference between δ_a and the SCARDEC dip is only 3.9°. These independent sources of information support the idea that the fault's dip angle determination in this study is more precise than the one of GCMT.

6 DISCUSSION AND PERSPECTIVES

6.1 Advantages of the SCARDEC method

Our body-wave deconvolutive approach allows us to determine both quickly and reliably the moment magnitude of major earthquakes. The method is automated, with two main steps. First the source duration is estimated based on the high-frequency content of teleseismic body waves, and then the optimization process of stabilized RSTFs gives us access to the moment magnitude, as well as to the focal mechanism and depth of the earthquake. The resolution of these earthquake parameters is enhanced by using a broad range of teleseismic waves (*P, PcP, PP, S, ScS*). These waves also have the advantage of arriving within a 30-min interval following the event origin. The entire inversion process requires less than 30 min on a simple computer with a 2.33 GHz processor. The parallelized version of the SCARDEC method, done on a 16-core machine, reduces this time to less than 5 min. Using the real-time transmission available for most of the FDSN data, a SCARDEC solution can therefore be obtained 35 min after the earthquake's occurrence.

As the SCARDEC method does not make the assumption that the STF is the same at each teleseismic station, it is better adapted to large earthquakes than most of the automated techniques used to analyse source parameters of distant events. Compared to extended source methods (Olson & Apsel 1982; Hartzell & Heaton 1983), it presents the advantage that no constraints are imposed on the spatio-temporal complexity of the rupture process. For example, the rupture velocity regimes, the shape of the local STF or the slip roughness do not enter in the parametrization of the inversion. This last point can also explain why the method should not be subject to underestimation of the moment magnitude. Because the shape of the STF is free for each station, the deconvolution transfers the whole waveform energy to the STF. The method does not suffer from inappropriate parametrization of the source process which could impede the modelling of some features of the waveforms and could result in a smaller moment magnitude.

In addition to arriving faster than surface waves, body waves are not sensitive to the magnitude-dip trade-off that affects shallow earthquake determinations using surface waves. This explains why we have found for some earthquakes values of magnitude and dip different from GCMT. Though different, we show that the values agree well after correcting the GCMT parameters for the existing trade-off. We have confirmed by forward modelling that SCARDEC parameters explain long-period surface waves as well as GCMT parameters. Other independent information, including studies analysing geodetic data or focal mechanisms of moderate seismicity, also support our findings. For about half of the large subduction earthquakes studied here, the magnitude-dip trade-off seems to cause the GCMT method to preferentially underestimate the dip and overestimate the seismic moment than the opposite. The two earthquakes for which we obtain the most convincing evidence of this behaviour are the 2003 Hokkaido and the 2005 Sumatra earthquakes. In both cases our estimate of the seismic moment is smaller (M_w reduced by 0.1–0.18) than the GCMT value.

The accurate determination of seismic moment of major earthquakes provides valuable information both for a better anticipation of the consequences of these events (e.g. for tsunami alert) and as a first-order parameter for more detailed earthquake source process studies. It also has an important role in assessing the balance between seismic and aseismic deformation in the Earth, because this balance is strongly influenced by the largest earthquakes. Considering all the earthquakes analysed here, we find that their cumulative

seismic moment deduced from SCARDEC solutions is about 25 per cent smaller than the one inferred from GCMT catalogue. As a result, the part of the aseismic processes (creep, silent earthquakes) in the global deformation processes is expected to be larger.

In this study, we have applied the SCARDEC method to the major subduction earthquakes. However the use of the method is not limited to this tectonic setting or to very large earthquakes. Without any modification, we apply the SCARDEC method to the recent earthquakes with magnitudes larger than $M_w = 6.8$. Results for the most significant events can be seen in the webpage http://geoazur.oca.eu/spip.php?rubrique57. For smaller earthquakes (down to magnitude 6), the method can also be used but requires two modifications. First, data filtering has to be changed because low frequencies are little excited by moderate earthquakes. As an example, we have analysed the $M_w = 6.3$ Aquila earthquake (Italy, 2009 April 06) in the 0.0125–0.1 Hz frequency band. Second, the automatic determination of the source duration T_d (see Section 2.2) gives values longer than the actual duration. This is due to the larger number of noisy stations and also to the fact that P-coda affects proportionally more the short duration source signals than the longer ones. This duration has thus to be determined either by signal inspection or as a function of the earthquake magnitude.

6.2 Source time function properties

In the SCARDEC method, we use primarily the physical constraints of the RSTFs as efficient criteria to optimize the focal mechanism and depth. However once the optimal parameter set is retrieved, the obtained RSTFs themselves provide valuable information on the earthquake rupture process. At the first order, we can observe for each earthquake the common features of all the RSTFs. In the analysed frequency band (0.005–0.03 Hz), some of the earthquakes have a simple moment release distribution (e.g. the Hokkaido earthquake; see Fig. 5), while other earthquakes are shown to be more complex, such as, for example, the Peru 2001 and 2007 earthquakes (Supplementary Figs 9 and 13). Both these earthquakes show two main episodes of moment release, which is confirmed by other analyses. More precisely, it is well known that the RSTFs give robust information on the preferential direction of the rupture propagation (e.g. Velasco *et al.* 1994); RSTFs tend to have shorter durations and higher amplitudes in the rupture propagation direction. This shows us, for example, that the 1995 Jalisco earthquake propagated in the northwest direction (Supplementary Fig. 3), while the 2001 Peru earthquake propagated in the southeast direction (Supplementary Fig. 9).

These source characteristics can be analysed more quantitatively when looking at higher frequency RSTFs. This can be done by a simple extension of the SCARDEC method. Once the optimal parameter set is determined, we can deconvolve both compressive and transverse body waves in a broader frequency range. To do so, we reduce the standard error of the Gaussian function f_g (eq. 8) and keep the same high-pass filtering corner (0.005 Hz). Using a standard error of 0.27 s, we can now retrieve RSTFs in a broad frequency range (0.005–0.5 Hz). Fig. 11 shows these broad-band RSTFs obtained for the 2003 Hokkaido earthquake. These RSTFs can be seen
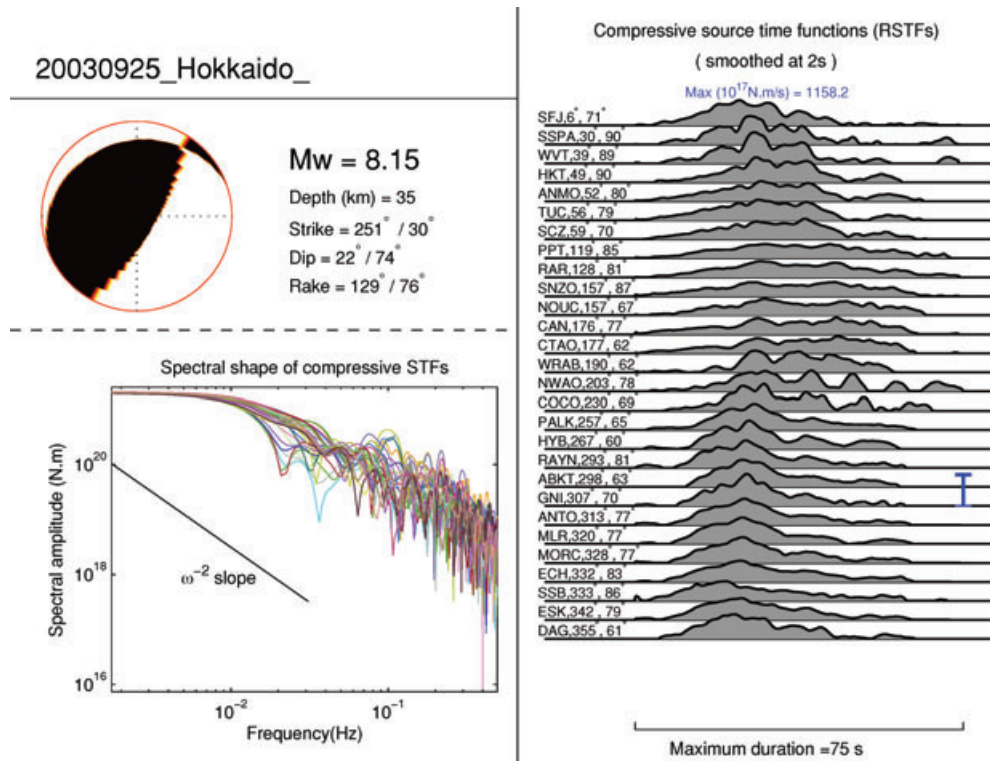

Figure 11. Broad-band RSTFs for the 2003 Hokkaido earthquake, in the time and frequency domains. (Top left-hand side) Optimal values of moment magnitude, depth and focal mechanism. (Bottom left-hand side) Spectrum of the broad-band compressive waves RSTFs (0.005–0.5 Hz). The classical ω^{-2} slope is shown in the left part of the figure. (Right-hand side) Broad-band RSTFs, in the time domain, for compressive waves. Compared to the RSTFs obtained in Fig. 5, the time properties can now be directly interpreted, because the smoothing time (2 s) is much smaller than the source duration (about 60 s). The indicated maximum value corresponds to the absolute maximum of all the moment rates. The corresponding scale is indicated by the blue bar, which is plotted next to the location of the maximal RSTF. For each RSTF, the name of the station, its azimuth and epicentral distance are presented.

as 'simplified' seismograms, because the source term is still present while most of the propagation term has been removed. These indirect data are thus well adapted for the application of extended source methods (Olson & Apsel 1982; Hartzell & Heaton 1983), to retrieve the rupture process on the earthquake fault. Because the method is automated, another perspective is to systematically analyse the rupture complexity on a large earthquake catalogue. This complexity can be estimated from the shape of the temporal RSTFs (right-hand part of Fig. 11), or from their spectral characteristics (left-hand part of Fig. 11). In the frequency domain, we can compare in particular the frequency decay with the classical ω^{-2} law (Brune 1970). Future applications of the SCARDEC model include the analysis of the diversity of earthquake complexity as a function of the earthquake location, the tectonics environment or the nature of the faults.

ACKNOWLEDGMENTS

REFERENCES

Aki, K. & Richards, P.G., 2002. *Quantitative Seismology,* 2nd Edn, University Science Books, Sausalito, CA.

Ammon, C.J., Kanamori, H., Lay, T. & Velasco, A.A., 2006. The 17 July 2006 Java tsunami earthquake, *Geophys. Res. Lett.,* **33,** L24308, doi:10.1029/2006GL028005.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS, Trans. Am. geophys. Un.,* **F897,** 81.

Bertero, M., Bindi, D., Boccacci, P., Cattaneo, M., Eva, C. & Lanza, V., 1997. Application of the projected Landweber method to the estimation of the source time function in seismology, *Inverse Probl.,* **13,** 465–486.

Bilek, S.L. & Ruff, L.J., 2002. Analysis of the June 23, 2001 Mw=8.4 Peru underthrusting earthquake and its aftershocks, *Geophys. Res. Lett.,* **29**(20), 1960, doi:10.1029/2002GL015543.

Bouchon, M., 1976. Teleseismic body wave radiation from a seismic source in a layered medium, *Geophys. J. R. astr. Soc.,* **47,** 515–530.

Brune, J.N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. geophys. Res.,* **75,** 4997–5009.

Chen, T., Newman, A.V., Feng, L. & Fritz, H.M., 2009. Slip distribution from the 1 April 2007 Solomon Islands earthquake: a unique image of near-trench rupture, *Geophys. Res. Lett.,* **36,** L16307, doi:10.1029/2009GL039496.

Dziewonski, A.M., Chou, T.A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Ferreira, A.M.G. & Woodhouse, J., 2007. Source, path and receiver effects on seismic surface waves, *Geophys. J. Int.,* **168,** 109–232.

Fuchs, K. & Müller, G., 1971. Computation of synthetic seismograms with the reflectivity method and comparison with observations, *Geophys. J. R. astr. Soc.,* **23,** 417–433.

Furlong, K., Lay, T. & Ammon, C., 2009. A great earthquake rupture across a rapidly evolving three-plate boundary, *Science,* **324**(5924), 226–229.

Giovanni, M.K., Beck, S.L. & Wagner, L., 2002. The June 23, 2001 Peru earthquake and the southern Peru subduction zone, *Geophys. Res. Lett.,* **29**(21), 2018, doi:10.1029/2002GL015774.

Goldstein, P. & Dodge, D., 1999. Fast and accurate depth and source mechanism estimation using P-waveform modelling: a tool for special event analysis, event screening, and regional calibration, *Geophys. Res. Lett.,* **26,** 2569–2572.

Gomez, J.M., Madariaga, R., Walpersdorf, A. & Chalard, E., 2000. The 1996 earthquakes in Sulawesi, Indonesia, *Bull. seism. Soc. Am.,* **90,** 739–751.

Hara, T., 2007. Measurement of duration of high-frequency energy radiation and its application to determination of magnitudes of large shallow earthquakes, *Earth Planets Space*, **59,** 227–231.

Hartzell, S.H. & Heaton, T.H., 1983. Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake, *Bull. seism. Soc. Am.,* **73,** 1553–1583.

Hjörleifsdóttir, V. & Ekström, G., 2010. Effects of three-dimensional Earth structure on CMT earthquake parameters, *Phys. Earth planet. Int.,* **179,** 178–190, doi:10.1016/j.pepi.2009.11.003.

Honda, R., Aoi, S., Morikawa, N., Sekiguchi, H., Kunugi, T. & Fujiwara, H., 2004. Ground motion and rupture process of the 2003 Tokachi-oki earthquake obtained from strong motion data of K-NET and KiK-net, *Earth Planets Space,* **56,** 317–332.

Kanamori, H., 1993. W phase, *Geophys. Res. Lett.,* **20**(16), 1691–1694.

Kanamori, H. & Given, J., 1981. Use of long-period surface waves for rapid determination of earthquake-source parameters, *Phys. Earth planet. Int.,* **27,** 8–31.

Kanamori, H. & Rivera, L., 2008. Source inversion of W phase: speeding up seismic tsunami warning, *Geophys. J. Int.,* **175,** 222–238.

Kennett, B.L.N. & Engdahl, E.R., 1991. Travel times for global earthquake location and phase association, *Geophys. J. Int.,* **105,** 429–465.

Kikuchi, M. & Kanamori, H., 1991. Inversion of complex body waves III, *Bull. seism. Soc. Am.,* **81,** 2335–2350.

Kikuchi, M. & Yamanaka, Y., 2001. EIC Seismological Note N105, Earthquake Information Center, University of Tokyo, http://wwweic.eri.u-tokyo.ac.jp/EIC/EIC_News/105E.html (last accessed 2010 October 29).

Kisslinger, C. & Kikuchi, M., 1997. Aftershocks of the Andreanof Islands Earthquake of June 10, 1996, and local seismotectonics, *Geophys. Res. Lett.,* **24**(15), 1883–1886.

Koketsu, K., Hikima, K., Miyazaki, S. & Ide, S., 2004. Joint inversion of strong motion and geodetic data for the source process of the 2003 Tokachi-oki, Hokkaido, earthquake *Earth Planets Space,* **56**(3), 329–334.

Komatitsch, D. & Tromp, J., 2002. Spectral-element simulations of global seismic wave propagation II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.,* **150,** 308–318.

Konca, A.O. *et al.*, 2007. Rupture kinematics of the 2005 Mw 8.6 Nias-Simeulue Earthquake from the joint inversion of seismic and geodetic data, *Bull. seism. Soc. Am.,* **97**(1A), S307–S322.

Konca, A.O. *et al.*, 2008. Partial rupture of a locked patch of the Sumatra megathrust during the 2007 earthquake sequence, *Nature,* **456,** 631–635.

Kreemer, C., Blewitt, G. & Maerten, F., 2006. Co- and postseismic deformation of the 28 March 2005 Nias Mw 8.7 earthquake from continuous GPS data, *Geophys. Res. Lett.,* **33,** L07307, doi:10.1029/2005GL025566.

Lay, T. & Wallace, T.C., 1995. *Modern Global Seismology,* Academic Press, San Diego, California.

Lomax, A., 2005. Rapid estimation of rupture extent for large earthquakes: application to the 2004, M9 Sumatra-Andaman mega-thrust, *Geophys. Res. Lett.,* **32,** L10314, doi:10.1029/2005GL022437.

Lomax, A. & Michelini, A., 2009. Mwpd: A duration-amplitude procedure for rapid determination of earthquake magnitude and tsunamigenic potential from *P* waveforms, *Geophys. J. Int.,* **176,** 200–214.

Lomax, A., Michelini, A. & Piatanesi, A., 2007. An energy-duration procedure for rapid determination of earthquake magnitude and tsunamigenic potential, *Geophys. J. Int.,* **170,** 1195–1209, doi:10.1111/j.1365-246X.2007.03469.x

Machida, Y. *et al.*, 2009. Heterogeneous structure around the rupture area of the 2003 Tokachi-oki earthquake (Mw = 8.0), Japan, as revealed by aftershock observations using Ocean Bottom Seismometers, *Tectonomphysics,* **465**(1-4), 164–176, ISSN 0040-1951, doi:10.1016/j.tecto.2008.11.009.

Mendoza, C. & Hartzell, S., 1999. Fault-slip distribution of the 1995 Colima-Jalisco, Mexico, earthquake, *Bull. seism. Soc. Am.,* **89,** 1338–1344.

Miyazaki, S. *et al.*, 2004. Modeling the rupture process of the 2003 September 25 Tokachi-Oki (Hokkaido) earthquake using 1-Hz GPS data, *Geophys. Res. Lett.,* **31,** L21603, doi:10.1029/2004GL021457.

Müller, G., 1985. The reflectivity method: a tutorial, *J. Geophys.,* **58,** 153–174.

Nabelek, J.L., 1984. Determination of earthquake source parameters from inversion of body waves, *Ph.D. Dissertation*, 361 pp., Mass. Inst. of. Technol., Cambridge

Ni, S., Kanamori, H. & Helmberger, D., 2005. Energy radiation from the Sumatra earthquake, *Nature*, **434**(7033), 582–582.

Olson, A.H. & Apsel, R.J., 1982. Finite fault and inverse theory with applications to the 1979 Imperial Valley earthquake, *Bull. seism. Soc. Am.,* **72,** 1969–2001.

Ritsema, J., van Heijst, H. & Woodhouse, J., 1999. Complex shear wave velocity structure imaged beneath Africa and Iceland, *Science*, **286,** 1925–1928.

Ruff, L.J. & Miller, A.D., 1994. Rupture process of large earthquakes in the northern Mexico subduction zone, *Pure appl. Geophys.,* **142,** 101–172.

Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm. I. Searching a parameter space, *Geophys. J. Int.,* **138,** 479–494.

Shao, G. & Ji, C., 2007. Preliminary result of the Dec 3, 1995 Mw 7.81 Kuril Earthquake, http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/1995/12/smooth/1995kuril.html (last accessed 2010 October 29).

Tsuboi, S., Abe, K., Takano, K. & Yamanaka, Y., 1995. Rapid determination of Mw from broadband P waveforms, *Bull. seism. Soc. Am.,* **83,** 606–613.

Vallée, M., 2004. Stabilizing the empirical Green function analysis: development of the projected Landweber method, *Bull. seism. Soc. Am.,* **94,** 394–409.

Vallée, M. & Bouchon, M., 2004. Imaging coseismic rupture in far field by slip patches, *Geophys. J. Int.,* **156,** 615–630.

Vallée, M., Bouchon, M. & Schwartz, S.Y., 2003. The 13 January 2001 El Salvador earthquake: a multidata analysis, *J. geophys. Res.,* **108**(B4), 2203, doi:10.1029/2002JB001922.

Velasco, A.A., Ammon, C.J. & Lay, T., 1994. Empirical Green function deconvolution of broadband surface waves: rupture directivity of the 1992 Landers, California (Mw = 7.3) earthquake, *Bull. seism. Soc. Am.,* **84,** 735–750.

Yagi, Y., 2004. Source rupture process of the 2003 Tokachi-oki earthquake determined by joint inversion of teleseismic body wave and strong ground motion data, *Earth Planets Space,* **56** (3), 311–316.

Yagi, Y., 2007. The 2007/09/12 Sumatra earthquake, http://www.geo.tsukuba.ac.jp/press_HP/yagi/EQ/20070912 (last accessed 2010 October 29).

Yamanaka, Y., 2007. The South Sumatra (07/09/12) earthquake, http://www.seis.nagoya-u.ac.jp/sanchu/Seismo_Note/2007/070912.jpg (last accessed 2010 October 29).

Yamanaka, Y. & Kikuchi, M., 2003. Source process of the recurrent Tokachi-oki earthquake on September 26, 2003, inferred from teleseismic body waves, *Earth Planets Space,* **55,** e21–e24.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Results for the 1994 Java earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S2. Results for the 1995 Chile earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S3. Results for the 1995 Jalisco earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details

Figure S4. Results for the 1995 Kuril earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S5. Results for the 1996 Minahassa earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S6. Results for the 1996 Irian-Jaya earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S7. Results for the 1996 Andreanof earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S8. Results for the 1997 Kamtchatka earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S9. Results for the 2001 Peru earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S10. Results for the 2005 Sumatra earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S11. Results for the 2006 Kuril earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S12. Results for the 2007 Solomon earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S13. Results for the 2007 Peru earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S14. Results for the 2007 Sumatra earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S15. Results for the 2009 New-Zealand earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S16. Results for the 2010 Chile earthquake. (Top panel) Focal mechanism, depth, magnitude, uncertainties and RSTFs. See Fig. 5 for more details. (Bottom panel) Agreement between data and synthetics, see Fig. 6 for more details.

Figure S17. Comparison of vertical component observed Rayleigh waves (black) with theoretical seismograms (red, green) at various stations of the FDSN, following the 2005 March 28 Sumatra earthquake. The name of each station is shown in the left of the waveforms and the corresponding source–receiver azimuth and epicentral distance are shown in the top, respectively. The synthetic seismograms are calculated for the earthquake source parameters in the GCMT catalogue (red) and for the parameters in the SCARDEC model (green; see main text for details). All traces have been low-pass cosine tapered around $T = 200$ s.

Figure S18. Same as in Supplementary Fig. 17, but for transverse component Love waves.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.