

FULL PAPER

Open Access



Empirical relocation of large subduction-zone earthquakes via the teleseismic network correlation coefficient method

Ta-Wei Chang*  and Satoshi Ide

Abstract

The network correlation coefficient method is a powerful tool for relocating events under low signal-to-noise ratio conditions, as well as small ($M_W < 5.0$) near-field earthquakes, where the relocation is performed by maximizing the sum of the waveform correlation coefficients between pairs of events in the network. Here we extend the method to relocating large ($6.0 \leq M_W \leq 7.3$) teleseismic subduction-zone events. We apply a source duration correction to better correlate different-magnitude events. Four case studies are presented where we precisely locate interplate events near the rupture areas of giant megathrust earthquakes, namely the 2011 M_W 9.1 Tohoku, 2010 M_W 8.8 Maule, 2005 M_W 8.6 Nias, and 2007 M_W 8.5 Bengkulu earthquakes. We demonstrate the extensive links between these interplate events, with a clear alignment of earthquakes observed along each subduction interface.

Keywords: Relocation, Teleseismic, Cross-correlation, Network correlation coefficient

Introduction

Inferences on earthquake interactions and their associated seismicity naturally require knowledge of the accurate location and timing of the earthquakes in question. Even though the event catalogs of most global earthquakes are routinely determined by agencies such as the International Seismological Centre (ISC) Bulletin (International Seismological Centre 2016) and the Preliminary Determination of Epicenters (PDE) Bulletin by the U. S. Geological Survey (U.S. Geological Survey Earthquake Hazards Program 2017), the location differences between these catalogs can be up to several tens of kilometers. One of the challenges we face is that widely implemented 1-D global velocity models, such as the Preliminary Reference Earth Model (Dziewoński and Anderson 1981), IASP91 (Kennett and Engdahl 1991), and AK135 (Kennett et al. 1995), are inadequate for regions with complex structure, such as subduction zones (e.g., Engdahl et al. 1982; Herrin

and Taggart 1968) and volcanic regions (e.g., Lomax et al. 2001). Furthermore, these models can introduce potential baseline bias, such as the 1.2-s baseline shift of the S-wave arrivals at epicentral distances of 25°–99° in the IASP91 model (Kennett and Engdahl 1991; Kennett et al. 1995). Some studies have implemented 3-D velocity models to address these issues (e.g., Chen and Willemann 2001; Ritzwoller et al. 2003; Smith and Ekström 1996), yielding general improvements over the 1-D models for known explosion events and well-located earthquakes. Another challenge is the large depth uncertainty due to the limited arrival-time changes when only the P-wave arrival times are used (Havskov et al. 2012). The inclusion of additional phases, which include the crustal (P, S, pP, sP), ocean reflection (pwP), and core phases (PKiKP and PKPdf), can reduce this depth uncertainty, with a notable improvement reported for subduction-zone regions, as demonstrated by Engdahl et al. (1998) in generating the EHB catalog and its successor, the ISC-EHB catalog (Weston et al. 2018). Schöffel and Das (1999) included core–mantle boundary reflection phases (PcP and ScP) with the P, S, pP, and sP phases, yielding a narrower subduction seismic

*Correspondence: tomdee@eps.s.u-tokyo.ac.jp
Department of Earth and Planetary Sciences, School of Science, The University of Tokyo, Tokyo, Japan

zone over the EHB solution. The effect of including these additional phases can be more pronounced than employing a 3-D velocity model, as demonstrated by Chen and Willemann (2001).

Precise relative locations between events can also be obtained when the distances between the events and stations greatly exceed the inter-event distances, resulting in ray path similarity (Waldhauser and Ellsworth 2000). The master event relocation method (Jackson and Fitch 1979; Stoddard and Woods 1990) relates all of the events to a “master event”, which is an event of interest or a particularly well-located event. The joint hypocenter determination method (Douglas 1967; Dewey 1972) includes a station correction term and determines all of the relative event locations simultaneously in order to account for any bias in the velocity model arising from site-specific complexities. Waldhauser and Ellsworth (2000) proposed the double-difference method to increase the precision of relative event locations, which relocates the event hypocenters using phase-picking information. It can also locate the event centroids by cross-correlating the waveforms from pairs of events, making use of the waveform similarity between nearby events with similar focal mechanisms that are recorded at the same station. The algorithm has been extended to regional and teleseismic distances by Waldhauser and Schaff (2007). Cross-correlation methods have been further extended to sum the cross-correlation coefficients over the entire seismic network, such as the network correlation coefficient (NCC) method (Gibbons and Ringdal 2006), which is a promising approach for locating low-frequency earthquakes (Ohta and Ide 2008, 2011) and local repeating events (Okuda and Ide 2018a, b). Furthermore, Poiata et al. (2016) proposed a network-based cross-correlation of nonlinear waveform transformations via characteristic functions to detect and locate earthquakes and tremors.

Here we extend the NCC method proposed by Ohta and Ide (2008), which outperforms standard double-difference approaches when analyzing low signal-to-noise ratio events, to precisely relocate events using far-field seismic records. Analysis of the longer waveforms, as opposed to only the initial arrivals, yields the centroid locations instead of the initial rupture points. The NCC method (Ohta and Ide 2008) assumes waveform similarity between events with similar focal mechanisms in a given region, such that the events can be relocated by simply shifting their arrival times, with all signals that are not target events representing incoherent noise. Our revised NCC method implements a correction for the rupture duration difference between different-magnitude events, which is required to apply the NCC method to larger events. The correction is performed by convolving the assumed source-time functions with

the event waveforms. We present case studies for large ($6.0 \leq M_W \leq 7.3$) interplate earthquakes that in Fig. 1 are located within a few hundred kilometers of the centroids of four megathrust earthquakes: the 2011 M_W 9.1 Tohoku, the 2010 M_W 8.8 Maule, the 2005 M_W 8.6 Nias, and the 2007 M_W 8.5 Bengkulu earthquakes. The four study regions are hereafter referred to as the Tohoku, Maule, Nias, and Bengkulu regions. We demonstrate that even events separated by >600 km can be strongly correlated based on their waveform similarity. Our relocation results show a clear alignment of events along the subducting interface without incorporating depth phase information.

Method

The NCC method outlined by Ohta and Ide (2008, 2011) comprises two steps: The precise relative location between any two events is first determined, followed by an inversion using the statistically significant relative location pairs to retrieve the absolute locations.

The relative location difference $\Delta \mathbf{x}_{ij}$ and origin time difference Δt_{ij} between any two events i and j are determined via a grid search for the maximum network correlation coefficient value (NCC value), which is a summation of the cross-correlation coefficients between the waveforms from N available stations and L components for both events across the seismic network (Ohta and Ide 2008). We first set one event i as the reference event, with its assumed location and timing given by the original catalog, and perform a grid search for the target event j . We define the waveforms u from station n and component l for the two events in the cross-correlation as follows:

$$\begin{cases} u_m^{iln} = u^{iln}(\tau^i + t(x_i^0, x_n) - \Delta T_{\text{pre}} + (m-1)\delta t) \\ u_m^{jln} = u^{jln}(\tau^j + t(x_j^0 + \Delta \mathbf{x}_{ij}, x_n) + \Delta t_{ij} - \Delta T_{\text{pre}} + (m-1)\delta t) \end{cases} \quad (1)$$

where τ is the origin time of the event at the source, which is given by the catalog, $t(x_i^0, x_n)$ is the theoretical travel time from catalog location x_i^0 to station location x_n , ΔT_{pre} is the pre-signal time, δt is the waveform sample interval, M is the number of samples in the waveforms, and $m = 1, \dots, M$. The time shifts due to $\Delta \mathbf{x}_{ij}$ and Δt_{ij} are also included in the target event waveforms u^j .

The NCC value (Gibbons and Ringdal 2006) for a given set of $\Delta \mathbf{x}_{ij}$ and Δt_{ij} shifts is defined as follows:

$$\text{NCC}(\Delta \mathbf{x}_{ij}, \Delta t_{ij}) = \sum_{ln} \frac{\sum_m u_m^{iln} u_m^{jln}}{\sqrt{\sum_m (u_m^{iln})^2 \sum_m (u_m^{jln})^2}}. \quad (2)$$

The most appropriate $\Delta \mathbf{x}_{ij}$ and Δt_{ij} values between the two events are determined via a grid search over the entire model space, which yields the global maximum of the NCC values. A strong correlation among all of

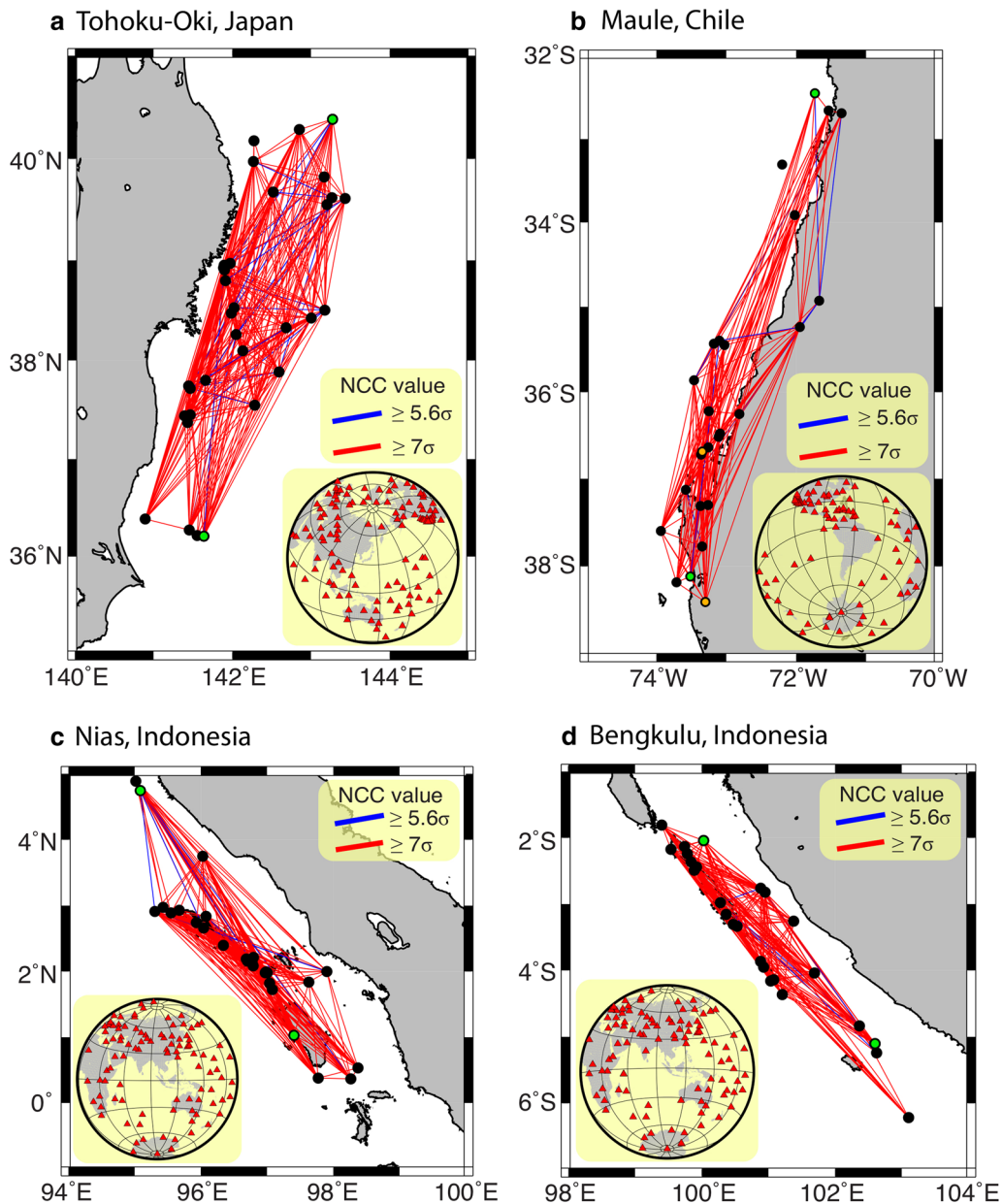


Fig. 1 Significant and consistent linkages for the four study regions: Tohoku, Japan (a), Maule, Chile (b), Nias, Indonesia (c), and Bengkulu, Indonesia (d) regions, respectively. The event linkages when both the forward and backward connections are significant and consistent in their relative location determinations in the four study regions are color-coded by their standard-deviation-normalized maximum NCC values in both directions. The inset figures show all of the stations used in each study region. The event pairs shown in orange in b and green in all subfigures are further examined in Fig. 2 and Additional file 1: Figures S1–S4, respectively

the station–component combinations between the two events is required to obtain a high NCC value, which is consistent with the assumption of waveform similarity. We create a network of relative location connections by performing grid searches to determine the maximum NCC value for all event pairs, which is then utilized to retrieve the absolute locations of the events. A two-step

grid search is implemented to save computational power, which is based on the statistical significance of the maximum NCC value. If the NCC values at each of the N_g grid points are independent and obey Gaussian distribution with zero mean and variance of σ^2 , the maximum NCC of $r\sigma$ is obtained with a probability of

Table 1 Grid-search parameters to determine the relative locations between events

	Search space (grid spacing): coarse				Search space (grid spacing): fine			
	N–S (km)	E–W (km)	Depth (km)	Time (s)	N–S (km)	E–W (km)	Depth (km)	Time (s)
Tohoku, Japan	1400 (10)	1000 (10)	100 (10)	40 (0.8)	100 (2)	100 (2)	100 (2)	10 (0.1)
Maule, Chile	1600 (10)	1000 (10)	100 (10)	40 (0.8)				
Nias, Indonesia	1400 (10)	1200 (10)	100 (10)	40 (0.8)				
Bengkulu, Indonesia	1400 (10)	1200 (10)	100 (10)	40 (0.8)				

$$P^{N_g}(r) = 1 - \left[\int_{-\infty}^r \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds \right]^{N_g}, \quad (3)$$

where r is the ratio between the maximum and standard deviation of the NCC values, and s is a variable for integration.

For each event pair i and j , we perform a grid search for the maximum of the NCC values and normalize the maximum NCC value by the standard deviation of the NCC values from the grid search to obtain the ratio r for the probabilities $P_{ij}^{N_g}$ of the event pair. Practically, we initially implement a coarser grid search over a 1000+ km horizontal, 100 km depth, and 40-s model space that encompasses $\Delta \mathbf{x}_{ij}$ and Δt_{ij} , with a 10-km grid spacing and 0.8-s time step (Table 1). When the statistical significance of the maximum NCC values is found for event pair i and j , namely $P_{ij}^{N_g} < 0.1$, we proceed to a finer grid search for that event pair, using a $100 \times 100 \times 100$ km and 10-s model space with a 2-km grid spacing and 0.1-s time step (Table 1). As approximately 8.0×10^6 to 9.6×10^6 coarser grids are employed as the model spaces in each study region, $P^{N_g} = 0.1$ correspond to a standard-deviation-normalized maximum NCC value of ~ 5.6 .

Following Ohta and Ide (2011), the inclusion of the relative locations between event pair i and j (i.e., forward connection from event i to j and reverse connection from j to i) to be used in the inversion for the absolute event locations is based on the statistical significance of their maximum NCC values to both be $P^{N_g} < 0.1$ ($\sim 5.6\sigma$). In addition, consistency screening that requires the relative locations between the forward and reverse connections to be similar to each other is applied to remove unreliable relative locations:

$$d_{ij} = \sqrt{(\Delta x_{ij}^{\text{EW}} + \Delta x_{ji}^{\text{EW}})^2 + (\Delta x_{ij}^{\text{NS}} + \Delta x_{ji}^{\text{NS}})^2 + (\Delta x_{ij}^{\text{UD}} + \Delta x_{ji}^{\text{UD}})^2}, \quad (4)$$

where d_{ij} is the difference in relative location between the forward and reverse connections, and EW, NS, and UD are E–W, N–S, and depth directions, respectively. Here we only include event pair i, j when $d_{ij} \leq 12$ km. When the consistency is not satisfied, an exception to

include one of the two connections is allowed when it is well determined (i.e., with low $P_{ij}^{N_g}$: high statistical significance), while the other is poorly determined (i.e., with high $P_{ji}^{N_g}$). Here, we allow $\Delta \mathbf{x}_{ij}^{\text{NCC}}$ when $P_{ij}^{N_g} < 0.00001$ ($\sim 7\sigma$) from i to j and $P_{ji}^{N_g} > 0.9$ ($\sim 5\sigma$) from j to i .

The absolute locations are calculated via a least-squares inversion that incorporates $\Delta \mathbf{x}_{ij}^{\text{NCC}}$ from the grid search and includes the original catalog location as a priori information. Assuming the previously determined relative locations are subjected to Gaussian errors, the absolute locations \mathbf{x} of all events are determined by minimizing the error function E (Ohta and Ide 2011):

$$E = \sum_{i,j,i \neq j} w_{ij} |\mathbf{x}_j - \mathbf{x}_i - \Delta \mathbf{x}_{ij}^{\text{NCC}}|^2 + \sum_i a^2 |\mathbf{x}_i - \mathbf{x}_i^0|^2, \quad (5)$$

where a is a hyperparameter that controls the weighting of catalog location \mathbf{x}_i^0 , which is determined by minimizing Akaike's Bayesian information criterion (Akaike 1980; Yabuki and Matsu'ura 1992) and w_{ij} is the weighting factor for $\Delta \mathbf{x}_{ij}^{\text{NCC}}$, which is calculated based on the probability of the relative location accuracy (Ohta and Ide 2008) (see Additional file 1: Text S1).

Waveform similarity is required to implement the NCC method since the relative locations are based on the correlation coefficient between waveforms. However, there are two noteworthy issues regarding the target events in this analysis: waveform differences due to different earthquake focal mechanisms (faulting types), and rupture duration differences due to different-magnitude earthquakes, even when they possess similar focal mechanisms. We address the former issue by applying this method to interplate events, given their relative abundance and similar focal mechanisms. Here we define interplate events based on the Kagan angle (Kagan 1991, 2007), the minimum angle of rotation between two double-couple focal mechanisms, which is described in more detail in Additional file 1: Text S2.

The latter issue can be difficult to overcome, and this limits the functionality of the method. If we assume unilateral (McGuire et al. 2002), circular rupture, the rupture duration t_d can be defined as follows:

$$t_d = \frac{2R}{V_R}, \quad (6)$$

where V_R is the rupture velocity and R is the radius of the circular rupture. Here we assume $V_R = 2.5$ km/s (Chouinet et al. 2018; Ye et al. 2016) and give R as follows (e.g., Shearer 2009):

$$R = \sqrt[3]{\frac{7}{16} \cdot 10^{(1.5M_W + 9.1)}} / \Delta\sigma, \quad (7)$$

where we assume $\Delta\sigma = 3$ MPa (Kanamori and Anderson 1975; Ye et al. 2016). For example, the pair of events at the lower and upper magnitude limits of this study ($M_W 6.0$ and $M_W 7.3$) will possess rupture durations of 4.6 and 20.3 s, respectively, via Eq. 6. Therefore, such a large difference in event duration makes it difficult to acquire a reasonable correlation between these different waveforms.

Here we assume that the source-time functions of all events are triangle functions, as suggested by a recent statistical study (Meier et al. 2017) and that their durations are defined by Eq. 6. We convolve the waveforms of one event with triangle functions whose duration equals that of the other event, and vice versa, to achieve a high degree of correlation between two event waveforms and determine their relative locations. Using the above-mentioned $M_W 6.0$ and $M_W 7.3$ events as an example, we convolve the waveforms for the $M_W 6.0$ event with triangle functions that have a 20.3-s duration, which is the duration of the $M_W 7.3$ event defined by Eq. 6, while also convolving the waveforms of the $M_W 7.3$ event with triangle functions that have a 4.6-s duration, which is the duration of the $M_W 6.0$ event. The amplitudes of the triangles are normalized such that the areas are equal to 1. For instance, if the duration of the triangle to be convolved with has duration of 10 s, the amplitude of that triangle would be 0.2. This approach adjusts the source-time functions inherited in the event waveforms of the two events, making them similar to each other. Since we choose events with similar faulting types and inter-event distances that are much shorter than the station–event distances, we can assume that the radiation pattern and propagation path effect are similar for two events recorded at the same station. Therefore, the event waveforms of the event pairs can be determined via cross-correlation once the source-time function correction is applied.

Data

Large, shallow events ($6.0 \leq M_W \leq 7.3$, depth ≤ 70 km) that were determined by the Global Centroid Moment Tensor (GCMT) catalog (Dziwowski et al. 1981; Ekström et al. 2012) and recorded in the four study regions

between 2001 and 2014 are initially considered for our analysis. Events that are clearly contaminated by other local events, such as those within 16 h of the mainshock, are discarded. The events are selected from the GCMT catalog, with the focal mechanism and moment magnitude information used to select interplate events and perform the duration correction, respectively. We limit the upper event size to $M_W 7.3$ due to the scarcity and potentially complex rupture processes of larger events, while setting $M_W 6.0$ as the lower event size limit to satisfy the far-field signal strength requirement and provide case studies using fewer events. We need to assume an initial catalog to utilize the NCC method. Here we employ the ISC-EHB catalog (Weston et al. 2018), which uses depth phases to improve the depth accuracy of the events, and exclude all of the events that are located outside of the study regions in the ISC-EHB catalog from the analysis. The dependence of the initial catalog is discussed in the next section.

Here we only use the teleseismic (30° – 95°) body waveforms to avoid contamination by later phases (e.g., PP and SS) and also ensure that the radiation pattern and propagation effect of the two events at each station remain similar to each other, as large event separations (> 600 km) are expected. Three-component broadband seismograms from the Global Seismographic Network (GSN) and International Federation of Digital Seismograph Network (FDSN) are used, where we assume the P and S waveforms are recorded on the vertical and horizontal components, respectively. The extracted waveforms are down-sampled to 10 Hz and detrended, with no instrument response removal or filtering employed in order to retain as much of the original event waveforms as possible. We then perform the above-mentioned convolution of the event waveforms from the two events in each event pair with triangle functions in their respective durations. We note that the event waveforms are subject to magnitude-dependent filtering during this procedure since the convolution step is effectively a low-pass filter. We then perform signal-to-noise ratio screening, where a given convoluted waveform is rejected if the ratio of the standard deviation of the signal to that of the noise is below 5. We also apply screening based on the mean level using waveform prior to the convolution to reject waveform including large irregular values as:

$$\bar{u} > 0.1 \max(u) \quad (8)$$

where u is the signal waveform prior to the convolution that is adjusted to the pre-signal noise level and \bar{u} is the mean of u . The criterion 0.1 is slightly smaller than the 99th percentile of the same value measured using all available data that pass the screening using standard deviation. These screening processes are independently performed

on each seismogram component at all stations. We do not explicitly consider the station distribution in the analysis; however, its effects on the resulting locations are discussed in the next section. Here the signal time window is empirically defined as the 80-s time period that begins 20 s prior to the catalog-predicted arrival time, assuming most waveform phases last 40 s. This allows for large catalog mis-locations that lead to arrival-time differences of up to 20 s in both directions to avoid removing high-quality data. Similarly, the noise time window is taken as the 80-s time period prior to the signal time window.

The source-to-station travel times are calculated using the tau-integral method (Buland and Chapman 1983) with the IASP91 velocity model (Kennett and Engdahl 1991) in this teleseismic modification of the NCC method. We select 44-s waveforms that begin 4 s prior to the catalog-predicted arrival times for the cross-correlations, assuming most waveform phases last 40 s, and allow for a shorter pre-event time than the signal-to-noise ratio screening to prevent the cross-correlation of unnecessary information. Compared with hypocentral location analysis, which only uses the arrival-time information provided by phase picking, this inclusion of slightly longer waveforms that cover most, if not all, of each corresponding phase yields the centroid locations. Discussions of the waveform duration for the cross-correlation procedure can be found in the next section. Event pairs with insufficient data, namely those with < 20 available station components, are discarded from the relative event relocation analysis.

Results and discussion

The NCC method requires a high degree of similarity among the far-field waveforms to yield the necessary correlation. Here we attempt to find the correlation between pairs of interplate events with $6.0 \leq M_W \leq 7.3$ that are more than 600 km apart. Figure 2 illustrates an example of the relative location determination via our modified NCC method for a pair of M_W 6.1 and M_W 7.1 events that are shown as orange points in Fig. 1b. Such event pairs that are connected under statistical significance of $P^{N_g} \leq 0.1$, which correspond to normalized NCC value of approximately 5.6σ , are herein referred to as being “linked” to each other. Events that are separated by > 450 km can be successfully linked in all four study regions since the analysis using far-field data ensures that the inter-event distances are still much shorter than the station–event distances; hence, similar radiation pattern and propagation effect from different events can still be expected despite the large separation. An example linkage from each region with event pairs shown in Fig. 1 as green points, which extends from about 450 to 650 km, is provided in Additional file 1: Figures S1–S4.

We find that 82.5% of the event pairs with sufficient data possess statistically significant linkages and pass our consistency screening for the inversion. This ratio, which we herein refer to as “approved ratio,” drops to 70.1% if the duration difference correction is not applied, thereby highlighting the importance of employing this correction to better constrain the relative locations. For event pairs with difference in magnitude 0.5 and more, the approved ratio increased from 65.7 to 83.0% after the duration difference correction. Similar increase in approved ratio from 43.5 to 64.5% and from 53.0 to 66.2% after the correction can be observed for events separated horizontally by 300 km or more and vertically by 15 km or more, respectively, using the ISC-EHB catalog location for consistency. From the approved ratios, we can infer that the event pairs with larger differences in magnitude or horizontal distances receive larger improvements after applying the duration difference correction, compared to all event pairs. By comparing the approved ratio of 83.0% for event pairs with magnitude difference 0.5 or more with that for all event pairs of 82.5%, we can infer that the effect of duration difference is successfully addressed by our duration difference correction. On the contrary, the approved ratios for events with larger separation in distance are systematically lower than that for all event pairs, as our correction is not designed to target the distance effect. The effect of distant event pairs is further investigated in the discussion below. Figure 1 shows the successful links between two events when both the forward and backward relative location determinations are significant and consistent. Approximately 96% of the events are connected by 10 or more significant and consistent linkages, with about 85% of the events constrained by 30 or more significant and consistent linkages, thereby providing an extensive network of linked connections covering each study region.

The relocated events generally follow the dip direction of the subducting interface, as shown in Figs. 3 and 4. We observe aligned earthquake distributions along each of the subducting interfaces in the four regions after the relocation. We perform simple 2-D polynomial fits to the relocated events and separately to the original ISC-EHB catalog (Weston et al. 2018) (see Additional file 1: Text S3 for details) to better assess the earthquake distributions, with the shallower and deeper subduction-zone locations given by the Slab 1.0 model (Hayes et al. 2012) used as constraints (Figs. 3b, d, 4b, d). We use the root-mean-square (RMS) difference to assess the deviation between the catalog and fitted plane, which is defined as follows:

$$\text{RMS difference} = \sqrt{\frac{\sum_{j=1}^n (\tilde{x}_j - \tilde{y}_j)^2}{n}}, \quad (9)$$

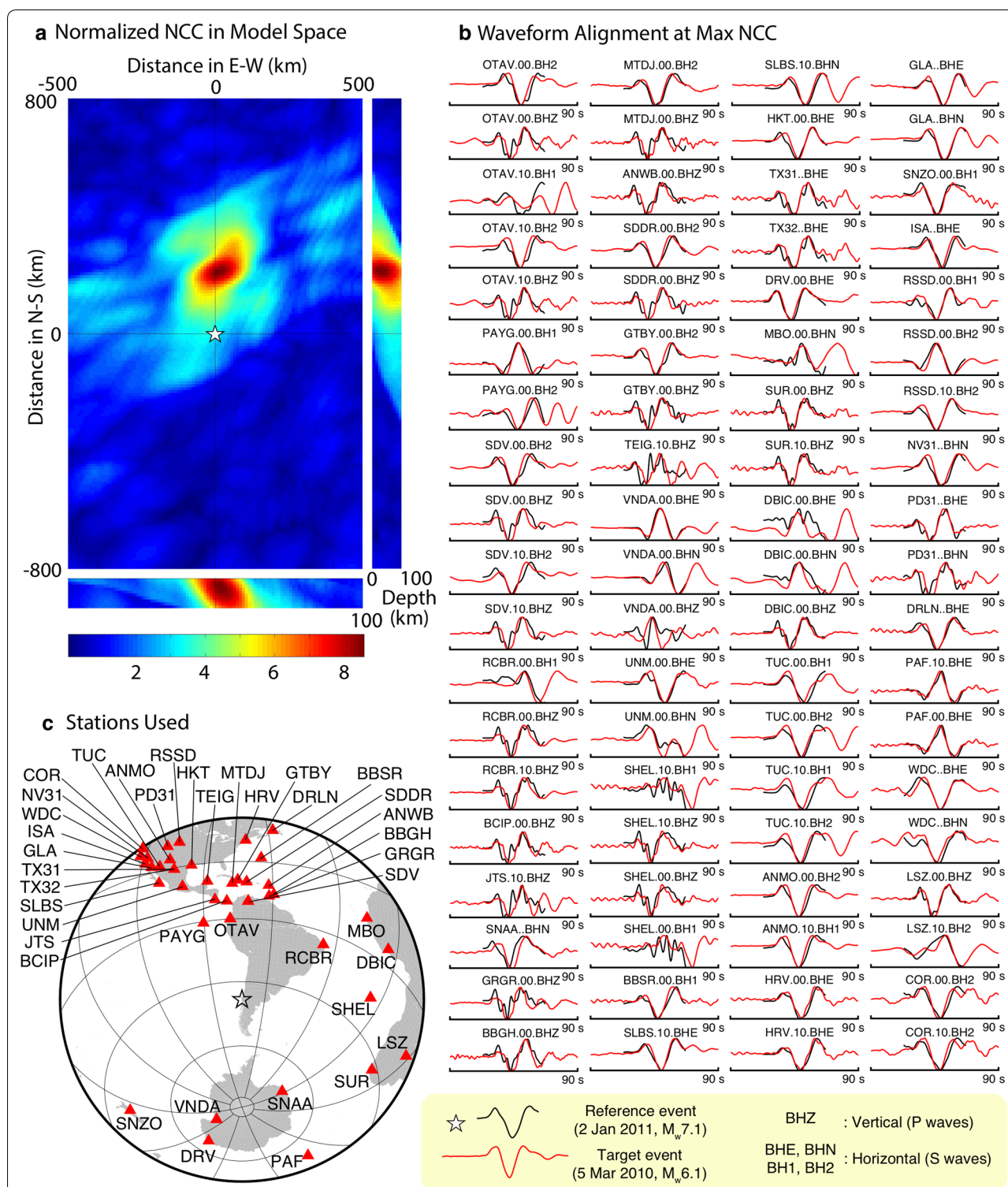


Fig. 2 Example grid-search analysis to determine the relative location in the Maule region, Chile. **a** Example grid search to determine the relative event location using the NCC values normalized by their standard deviation. **b** Alignment of the convolved waveforms for the two events when the maximum NCC value is reached. The event waveforms in **b** are normalized by their maximum amplitudes. The stations used in the analysis are shown in **c**. The black event is the reference event (January 2, 2011; M_w 7.1), and the red event is the target event (March 5, 2010; M_w 6.1). The two events are shown as orange points in Fig. 1b

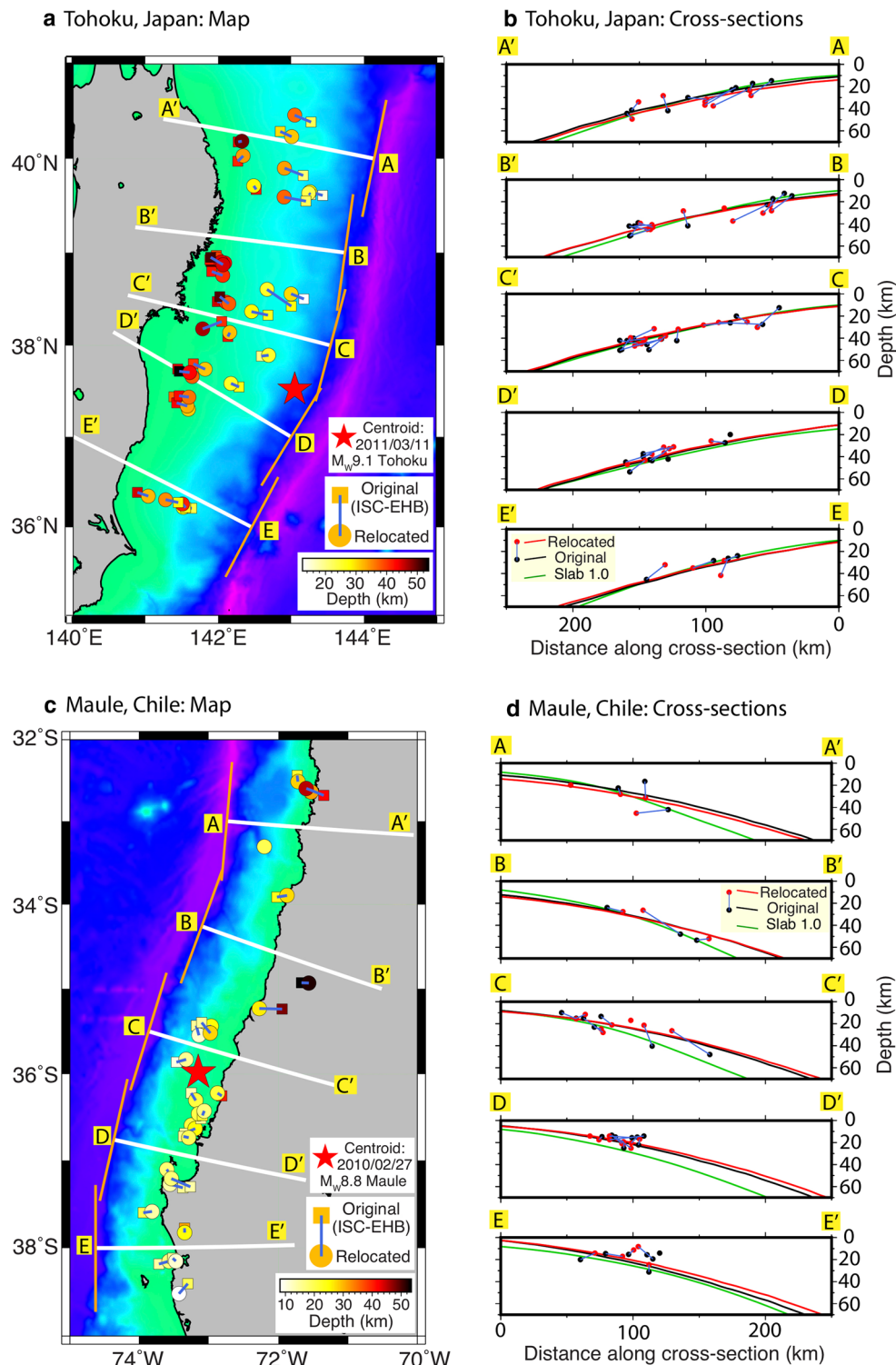


Fig. 3 Relocation results for the Tohoku and Maule regions. Original locations from the ISC-EHB catalog (squares; Weston et al. 2018) and relocated catalog (circles) in map view (**a**, **c**), and along the subduction-zone cross sections (**b**, **d**). The Slab 1.0 model (green; Hayes et al. 2012), as well as the fitted interfaces using the original (black) and relocated (red) catalogs, both of which incorporate the Slab 1.0 model (Hayes et al. 2012), is shown in the cross sections. See Additional file 1: Text S3 for details on the model fitting. The red stars in **a** and **c** represent the 2011 M_w 9.1 Tohoku and 2010 M_w 8.8 Maule earthquakes, respectively

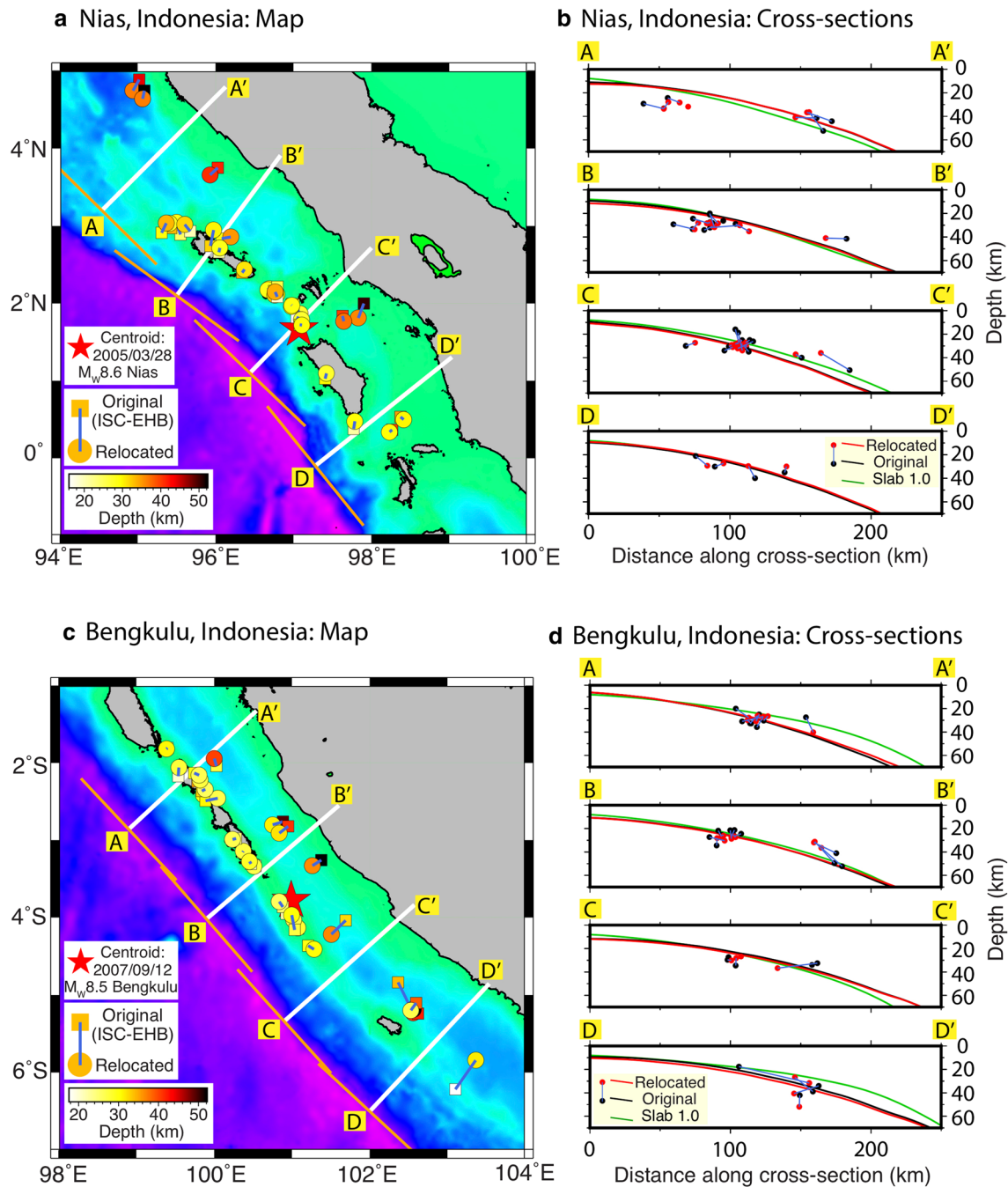


Fig. 4 Relocation results for the Nias and Bengkulu regions. The figure structure follows Fig. 3. The red stars in **a** and **c** represent the 2005 M_W 8.6 Nias and 2007 M_W 8.5 Bengkulu earthquakes, respectively

where \tilde{x}_j and \tilde{y}_j are the catalog and best-fit plane locations for event j , respectively, and n is the total number of events. We find that the RMS difference in depth between our relocated catalog and the best-fit planes for the four regions is 6.3 km, which is the same as the RMS difference in depth between the ISC-EHB catalog (Weston

et al. 2018) and its fitted planes. This result indicates that our relocated catalog is equally aligned along the fitted subduction model compared with the event locations in the original ISC-EHB catalog, even though the depth phases are not used to constrain the event depths, as was done by Weston et al. (2018). Furthermore, our method

provides the centroid locations, as opposed to the hypocentral locations in the ISC-EHB catalog. In addition to the above comparison with fitted planes, we further compare the two catalogs with the Slab 1.0 model, which is constructed from several earthquake catalogs (Hayes et al. 2012), yielding RMS differences in depth between the model, and our relocations and the ISC-EHB catalog of 7.4 and 6.8 km, respectively. These results indicate that the ISC-EHB catalog locations are slightly closer to the Slab 1.0 model (Hayes et al. 2012) than our relocations, which is reasonable since its sister EHB catalog (Engdahl et al. 1998) is one of the catalogs used to constrain the Slab 1.0 model (Hayes et al. 2012).

We examine the dependence of the earthquake locations on the initial catalog by conducting the same analysis across the Tohoku region, Japan, using the Japan Meteorological Agency (JMA) hypocenter catalog (Japan Meteorological Agency 2018) as the initial catalog, which is known for its accuracy due to the dense station distribution. The relocated results are shown in Additional file 1: Figure S5, with a clear shift observed between the two results. We could compensate for this by shifting the relocated catalog starting from the JMA catalog by 17.9 km WNW and 0.5 km shallower to align with that starting from the ISC-EHB catalog. This shift is the same when calculated using the difference between either the two original catalogs or the two relocated catalogs, reducing the RMS difference between the two relocated catalogs from 17.5 to 2.0 km, 4.7 to 1.2 km, and 3.3 to 3.2 km in E–W, N–S, and depth directions, respectively. This large reduction in RMS difference indicates that our method is sensitive to the relative location between events while preserving the centroid shift inherited from the initial catalog location.

We perform a bootstrap test (e.g., Efron and Tibshirani 1994) on the absolute location inversion results to further examine the robustness of the inter-event linkages, which is crucial for providing the final locations in this analysis. We randomly select the same number of relative locations that pass the significance and consistency test, as in the usual case where we utilize all of the available data, and perform the inversion 5000 times. We can retrieve the standard error (SE) in the locations via bootstrapping, as follows:

$$SE = \sqrt{\frac{\sum_{k=1}^K (x_k - \bar{x})^2}{K}}, \quad (10)$$

where K is total number of realizations, x_k is the location of the k th realization, and \bar{x} is the average of all realizations. The mean SEs are 1.1, 0.8, and 1.4 km in E–W, N–S, and depth directions, respectively, for events in the four regions, with maximum SEs of ≤ 5.3 km in all directions.

These low values indicate that the relative location linkages between events are in strong agreement with each other, thereby indicating a stable inversion for the absolute locations.

We perform another absolute location inversion, where we restrict the maximum horizontal distance between events to 300 km as given by the ISC-EHB catalog, to examine the effect of distant event pairs on the inversion results. The RMS differences between the two results are 1.5, 1.4, and 1.2 km in E–W, N–S, and depth directions, respectively, for the entire dataset, which are essentially negligible relative to the large inter-event distances. These low RMS differences indicate that the linkages between events provide reliable constraints on the relative locations even though they are separated by large distances.

Here we are able to obtain event centroid relocations along aligned features using 44-s waveforms, which cover most, if not all, of the corresponding phases, without requiring depth phase information. Our modified NCC method has several advantages over phase-picking methods. For example, it can be challenging to pick the arrival time in low signal-to-noise ratio conditions where the onset is unclear and/or when the event waveform is contaminated by other events. Furthermore, the depth phases of larger events might be hidden within the direct phases, which makes the picking of depth phases challenging. Finally, we obtain the centroid locations via our analysis, as opposed to the hypocentral locations via phase-picking analyses. We tested the effect of the waveform length by comparing our results (44-s waveforms) with the same analysis using 40- and 48-s waveforms, both of which include a 4-s pre-arrival time. We found that the RMS differences between the relocated catalogs using the 40- and 44-s waveforms were ≤ 1.7 km in all three directions, whereas they were ≤ 1.5 km between those using the 44- and 48-s waveforms. This indicates the robustness of the event relocations to slight changes in the waveform durations applied in the analysis.

There are some limitations to the modified NCC method that must be considered. The first is the calculation speed, as the relative locations between event pairs are determined via exhaustive grid-search methods. Furthermore, since waveform similarity is assumed, we can only apply the method to events with similar focal mechanisms, which yield similar radiation patterns at the stations. We also assume that the source-time functions of the events are simple triangle functions, meaning that the potential source complexity of larger-magnitude events might limit the feasibility of the method. Finally, since our method only targets the direct phases, which often dominate the energy in far-field body waveforms, the inclusion of depth phases could in turn act as an additional source of error.

Nevertheless, we found the signals of the waveform analyzed are dominated by the direct phases, which are also the most uniformly observed. Therefore, even as we do not exclude them from the analysis, we believe that the resulting location is not strongly affected by the depth phases.

We initially considered all of the available stations in our analysis, as shown in the inset figures in Fig. 1. However, the actual station distribution used in this analysis after the signal-to-noise ratio screening, as shown in Fig. 2c and Additional file 1: Figures S1c–S4c, was restricted and therefore affected the results. While a bootstrap test that randomly selects various combinations of stations and components for all event pairs is a viable option for evaluating the effect of the station distribution, it is a computationally intensive approach that is outside the scope of this study. Instead, we conduct a bootstrap test using the five pairs of events shown in Fig. 1 as orange and green points, and in Fig. 2 and Additional file 1: Figures S1–S4 for more detail, to gain insight into the variability in relative locations resulting from different combinations of data. We randomly select the same number of waveforms from all of the available stations and components as used in our regular analysis, and conduct a grid search for the relative locations, and repeat the process 400 times. The mean SEs are 3.4, 3.7, and 4.8 km in E–W, N–S, and depth directions, respectively, for the realizations that meet the consistency and significance thresholds. The SEs are higher for the event pair specified in Additional file 1: Figure S3, with SEs in one of the directions up to 3.8, 7.9, and 10.9 km in E–W, N–S, and depth directions, respectively, possibly reflecting the limited number of stations or the weaker correlation between some of the data. These values demonstrate that the station distribution, and therefore the data combination, can impact the event relocations. While our bootstrap analysis of the station distribution for the relative locations yields slightly higher mean SEs that are < 5 km in each direction, the resulting absolute locations are still robust, as shown by our bootstrap analysis using various relative locations for the final location that yields mean SEs of < 1.5 km in each direction. This might reflect the use of data from all available stations in the network when performing the latter bootstrap analysis.

Here we confine the method to events with $6.0 \leq M_W \leq 7.3$ to retain high far-field signal-to-noise ratios. However, since the applicability of this method to low signal-to-noise events has already been demonstrated (Ohta and Ide 2008), we expect the future application of this method to include smaller teleseismic events. This method might also be a strong candidate for the future centroid determination of events that are too small for stable moment tensor inversions if it is successful in relocating these smaller events.

Conclusion

We update the NCC method for teleseismic earthquake relocations and apply it to large interplate earthquakes in four subduction zones. We demonstrate that events with distinct rupture durations and magnitude differences, as well as distant events that are separated by several hundred kilometers, can be successfully linked after correcting for rupture duration. The centroid distribution after the relocation shows a clear alignment along the fitted subduction interface that slightly outperforms the original ISC-EHB catalog despite not including depth phase information. The robustness of our results is demonstrated via a bootstrap test and a comparison of our results with another initial catalog. We expect future analyses to extend this method to the centroid determination of smaller-magnitude events, as the NCC method was initially designed for events with poor signal-to-noise characteristics.

Additional file

Additional file 1. Additional file provides detailed explanations of the weighting scheme applied in the NCC relocation method (Text S1), interplate event selection criterion (Text S2), and approach for fitting a surface to the events (Text S3). Figures S1–S4 are examples of distant (about 450 to 650 km apart) events pairs that are significantly linked in each of the four study regions. Figure S5 illustrates the differences in the original and relocated event locations using two different initial catalogs.

Abbreviations

ISC: International Seismological Centre; PDE Bulletin: Preliminary Determination of Epicenters Bulletin; NCC method: network correlation coefficient method; NCC value: network correlation coefficient value; GCMT: Global Centroid Moment Tensor; GSN: Global Seismographic Network; FDSN: International Federation of Digital Seismograph Network; RMS difference: root-mean-square difference; JMA: Japan Meteorological Agency; SE: standard error.

Acknowledgements

All of the figures in this study were prepared using Generic Mapping Tools (GMT) (Wessel et al. 2013). The bathymetry shown in Figs. 3 and 4 and Additional file 1: Figure S5 is from the ETOPO1 global relief model (Amante and Eakins 2009). The authors thank Aitaro Kato and Pierre Romanet for valuable discussions and suggestions. The manuscript benefited greatly from comments and suggestions by two anonymous reviewers.

Authors' contributions

TC carried out the analyses, visualized the results, and wrote the majority of the manuscript. SI structured the study and prepared most of the tools for the analysis. Both authors read, modified, and approved the final manuscript.

Funding

This study is supported by the Earthquake and Volcano Hazards Observation and Research Program, Grant-in-Aid for Scientific Research on Innovative Areas 16H06477 and Grant-in-Aid for Scientific Research (A) 16H02219 from the Council for Science and Technology, Ministry of Education, Culture, Sports, Science and Technology, Japan, as well as the Science and Technology Research Partnership for Sustainable Development Projects from the Japan Science and Technology Agency. TC is additionally supported by the Graduate School of Science Scholarship for International Students, University of Tokyo.

Availability of data and materials

The GSN and FDSN waveforms used in this study were accessed from IRIS Data Services (http://ds.iris.edu/wilber3/find_event). The ISC-EHB (<http://www.isc.ac.uk/isc-ehb>), GCMT (<https://www.globalcmt.org>), and JMA (http://www.data.jma.go.jp/svd/eqev/data/bulletin/index_e.html) catalogs are available from their respective websites.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Received: 7 January 2019 Accepted: 5 July 2019

Published online: 19 July 2019

References

- Akaike H (1980) Likelihood and the Bayes procedure. *Trab Estad Y Investig Oper* 31:143–166. <https://doi.org/10.1007/BF02888350>
- Amante C, Eakins BW (2009) ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA technical memorandum NESDIS NGDC-24, Boulder, Colorado
- Buland R, Chapman CH (1983) The computation of seismic travel times. *Bull Seismol Soc Am* 73:1271–1302
- Chen Q, Willemann RJ (2001) Global test of seismic event locations using three-dimensional earth models. *Bull Seismol Soc Am* 91:1704–1716. <https://doi.org/10.1785/0120000300>
- Chounet A, Vallée M, Causse M, Courboulès F (2018) Global catalog of earthquake rupture velocities shows anticorrelation between stress drop and rupture velocity. *Tectonophysics* 733:148–158. <https://doi.org/10.1016/j.tecto.2017.11.005>
- Dewey JW (1972) Seismicity and tectonics of western Venezuela. *Bull Seismol Soc Am* 62:1711–1751
- Douglas A (1967) Joint epicentre determination. *Nature* 215:47–48. <https://doi.org/10.1038/215047a0>
- Dziwioński AM, Anderson DL (1981) Preliminary reference Earth model. *Phys Earth Planet Inter* 25:297–356. [https://doi.org/10.1016/0031-9201\(81\)90046-7](https://doi.org/10.1016/0031-9201(81)90046-7)
- Dziwioński AM, Chou T-A, Woodhouse JH (1981) Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J Geophys Res Solid Earth* 86:2825–2852. <https://doi.org/10.1029/JB086iB04p02825>
- Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. Chapman and Hall/CRC, Boca Raton
- Ekström G, Nettles M, Dziwioński AM (2012) The global CMT project 2004–2010: centroid-moment tensors for 13,017 earthquakes. *Phys Earth Planet Inter* 200–201:1–9. <https://doi.org/10.1016/j.pepi.2012.04.002>
- Engdahl ER, Dewey JW, Fujita K (1982) Earthquake location in island arcs. *Phys Earth Planet Inter* 30:145–156. [https://doi.org/10.1016/0031-9201\(82\)90099-1](https://doi.org/10.1016/0031-9201(82)90099-1)
- Engdahl ER, van der Hilst R, Buland R (1998) Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull Seismol Soc Am* 88:722–743
- Gibbons SJ, Ringdal F (2006) The detection of low magnitude seismic events using array-based waveform correlation. *Geophys J Int* 165:149–166. <https://doi.org/10.1111/j.1365-246X.2006.02865.x>
- Havskov J, Bormann P, Schweitzer J (2012) IS 11.1: seismic source location. In: *New manual of seismological observatory practice (NMSOP-2)*, Potsdam, Germany, pp 1–36
- Hayes GP, Wald DJ, Johnson RL (2012) Slab1.0: a three-dimensional model of global subduction zone geometries. *J Geophys Res Solid Earth* 117:B01302. <https://doi.org/10.1029/2011jb008524>
- Herrin E, Taggart J (1968) Source bias in epicenter determinations. *Bull Seismol Soc Am* 58:1791–1796
- International Seismological Centre (2016) On-line bulletin. <http://www.isc.ac.uk>. Accessed 24 Dec 2018
- Jackson J, Fitch TJ (1979) Seismotectonic implications of relocated aftershock sequences in Iran and Turkey. *Geophys J R Astron Soc* 57:209–229. <https://doi.org/10.1111/j.1365-246X.1979.tb03781.x>
- Japan Meteorological Agency (2018) The seismological bulletin of Japan, Tokyo, Japan. http://www.data.jma.go.jp/svd/eqev/data/bulletin/index_e.html. Accessed 24 Dec 2018
- Kagan YY (1991) 3-D rotation of double-couple earthquake sources. *Geophys J Int* 106:709–716. <https://doi.org/10.1111/j.1365-246X.1991.tb06343.x>
- Kagan YY (2007) Simplified algorithms for calculating double-couple rotation. *Geophys J Int* 171:411–418. <https://doi.org/10.1111/j.1365-246X.2007.03538.x>
- Kanamori H, Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bull Seismol Soc Am* 65:1073–1095
- Kennett BLN, Engdahl ER (1991) Traveltimes for global earthquake location and phase identification. *Geophys J Int* 105:429–465. <https://doi.org/10.1111/j.1365-246X.1991.tb06724.x>
- Kennett BLN, Engdahl ER, Buland R (1995) Constraints on seismic velocities in the Earth from traveltimes. *Geophys J Int* 122:108–124. <https://doi.org/10.1111/j.1365-246X.1995.tb03540.x>
- Lomax A, Zollo A, Capuano P, Virieux J (2001) Precise, absolute earthquake location under Somma–Vesuvius volcano using a new three-dimensional velocity model. *Geophys J Int* 146:313–331. <https://doi.org/10.1046/j.0956-540x.2001.01444.x>
- McGuire JJ, Zhao L, Jordan TH (2002) Predominance of unilateral rupture for a global catalog of large earthquakes. *Bull Seismol Soc Am* 92:3309–3317. <https://doi.org/10.1785/0120010293>
- Meier M-A, Ampuero JP, Heaton TH (2017) The hidden simplicity of subduction megathrust earthquakes. *Science* 357:1277–1281. <https://doi.org/10.1126/science.aan5643>
- Ohta K, Ide S (2008) A precise hypocenter determination method using network correlation coefficients and its application to deep low-frequency earthquakes. *Earth Planets Space* 60:877–882. <https://doi.org/10.1186/BF03352840>
- Ohta K, Ide S (2011) Precise hypocenter distribution of deep low-frequency earthquakes and its relationship to the local geometry of the subducting plate in the Nankai subduction zone, Japan. *J Geophys Res* 116:B01308. <https://doi.org/10.1029/2010JB007857>
- Okada T, Ide S (2018a) Hierarchical rupture growth evidenced by the initial seismic waveforms. *Nat Commun* 9:3714. <https://doi.org/10.1038/s41467-018-06168-3>
- Okada T, Ide S (2018b) Streak and hierarchical structures of the Tohoku–Hokkaido subduction zone plate boundary. *Earth Planets Space* 70:132. <https://doi.org/10.1186/s40623-018-0903-8>
- Poiata N, Satriano C, Vilotte J-P et al (2016) Multiband array detection and location of seismic sources recorded by dense seismic networks. *Geophys J Int* 205:1548–1573. <https://doi.org/10.1093/gji/ggw071>
- Ritzwoller MH, Shapiro NM, Levshin AL et al (2003) Ability of a global three-dimensional model to locate regional events. *J Geophys Res Solid Earth* 108:2353. <https://doi.org/10.1029/2002JB002167>
- Schöffel H-J, Das S (1999) Fine details of the Wadati–Benioff zone under Indonesia and its geodynamic implications. *J Geophys Res Solid Earth* 104:13101–13114. <https://doi.org/10.1029/1999JB900091>
- Shearer PM (2009) *Introduction to seismology*, 2nd edn. Cambridge University Press, Cambridge
- Smith GP, Ekström G (1996) Improving teleseismic event locations using a three-dimensional Earth model. *Bull Seismol Soc Am* 86:788–796
- Stoddard PR, Woods MT (1990) Master event relocation of Gorda block earthquakes: implications for deformation. *Geophys Res Lett* 17:961–964. <https://doi.org/10.1029/GL017i007p00961>
- U.S. Geological Survey Earthquake Hazards Program (2017) Preliminary Determination of Epicenters (PDE) bulletin: U.S. Geological Survey. <https://doi.org/10.5066/f74t6gjc>
- Waldhauser F, Ellsworth WL (2000) A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California. *Bull Seismol Soc Am* 90:1353–1368. <https://doi.org/10.1785/0120000006>
- Waldhauser F, Schaff D (2007) Regional and teleseismic double-difference earthquake relocation using waveform cross-correlation and global

- bulletin data. *J Geophys Res* 112:B12301. <https://doi.org/10.1029/2007JB004938>
- Wessel P, Smith WHF, Scharroo R et al (2013) Generic mapping tools: improved version released. *EOS Trans Am Geophys Union* 94:409–410. <https://doi.org/10.1002/2013EO450001>
- Weston J, Engdahl ER, Harris J et al (2018) ISC-EHB: reconstruction of a robust earthquake data set. *Geophys J Int* 214:474–484. <https://doi.org/10.1093/gji/ggy155>
- Yabuki T, Matsu'ura M (1992) Geodetic data inversion using a bayesian information criterion for spatial distribution of fault slip. *Geophys J Int* 109:363–375. <https://doi.org/10.1111/j.1365-246X.1992.tb00102.x>

- Ye L, Lay T, Kanamori H, Rivera L (2016) Rupture characteristics of major and great ($M_w \geq 7.0$) megathrust earthquakes from 1990 to 2015: 1. Source parameter scaling relationships. *J Geophys Res Solid Earth* 121:826–844. <https://doi.org/10.1002/2015JB012426>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
