Nooshiri, N., Saul, J., Heimann, S., Tilmann, F., Dahm, T. (2016): Revision of Earthquake Hypocentre Locations in Global Bulletin Data Sets using Source-Specific Station Terms. - *Geophysical Journal International*, *208*, 2, pp. 589—602.

DOI: http://doi.org/10.1093/gji/ggw405

Geophysical Journal International

Geophys. J. Int. (2017) **208,** 589–602
Advance Access publication 2016 October 27
GJI Seismology

doi: 10.1093/gji/ggw405

Revision of earthquake hypocentre locations in global bulletin data sets using source-specific station terms

Nima Nooshiri,[1,2] Joachim Saul,[1] Sebastian Heimann,[1] Frederik Tilmann[1,3] and Torsten Dahm[1,2]

[1] *Deutsches GeoForschungsZentrum GFZ, Potsdam, Germany. E-mail: nima.nooshiri@gfz-potsdam.de*
[2] *Universität Potsdam, Potsdam, Germany*
[3] *Freie Universität Berlin, Berlin, Germany*

Accepted 2016 October 26. Received 2016 October 24; in original form 2016 April 12

SUMMARY

Global earthquake locations are often associated with very large systematic travel-time residuals even for clear arrivals, especially for regional and near-regional stations in subduction zones because of their strongly heterogeneous velocity structure. Travel-time corrections can drastically reduce travel-time residuals at regional stations and, in consequence, improve the relative location accuracy. We have extended the shrinking-box source-specific station terms technique to regional and teleseismic distances and adopted the algorithm for probabilistic, nonlinear, global-search location. We evaluated the potential of the method to compute precise relative hypocentre locations on a global scale. The method has been applied to two specific test regions using existing *P*- and *pP*-phase picks. The first data set consists of 3103 events along the Chilean margin and the second one comprises 1680 earthquakes in the Tonga-Fiji subduction zone. Pick data were obtained from the GEOFON earthquake bulletin, produced using data from all available, global station networks. A set of timing corrections varying as a function of source position was calculated for each seismic station. In this way, we could correct the systematic errors introduced into the locations by the inaccuracies in the assumed velocity structure without explicitly solving for a velocity model. Residual statistics show that the median absolute deviation of the travel-time residuals is reduced by 40–60 per cent at regional distances, where the velocity anomalies are strong. Moreover, the spread of the travel-time residuals decreased by ∼20 per cent at teleseismic distances (>28°). Furthermore, strong variations in initial residuals as a function of recording distance are smoothed out in the final residuals. The relocated catalogues exhibit less scattered locations in depth and sharper images of the seismicity associated with the subducting slabs. Comparison with a high-resolution local catalogue reveals that our relocation process significantly improves the hypocentre locations compared to standard locations.

Key words: Seismicity and tectonics; Computational seismology; Subduction zone processes; Pacific Ocean; South America.

1 INTRODUCTION

The use of a 1-D velocity model for seismic event location is often associated with significant travel-time residuals. Particularly for regional stations in subduction zones, where the velocity structure strongly deviates from the assumed 1-D model, residuals of up to ±10 s are observed even for clear arrivals, resulting in systematically biased locations. Since these residuals are normally not due to measurement errors but regional travel-time anomalies, arrival times at regional and near-regional stations do not match the location obtained with teleseismic picks, and vice versa (Fig. 1). If the earthquake is weak and only recorded regionally, or if fast location based on regional stations is needed, the location may be far off the corresponding teleseismic location.

A proper solution to this problem requires knowledge of the dominant factors causing uncertainties of earthquake locations, which are (Pavlis 1986; Husen & Hardebeck 2010): assumed velocity model-dependent errors of calculated travel times, timing and picking errors of seismic phase arrival times, and inherent nonlinearity of the earthquake location problem. The latter is only relevant to linearized, least-squares location methods. Direct, global-search location algorithms are not affected by errors due to nonlinearity of the inverse problem since they do not involve linearization of the problem. The other two groups of errors are relevant for both linearized

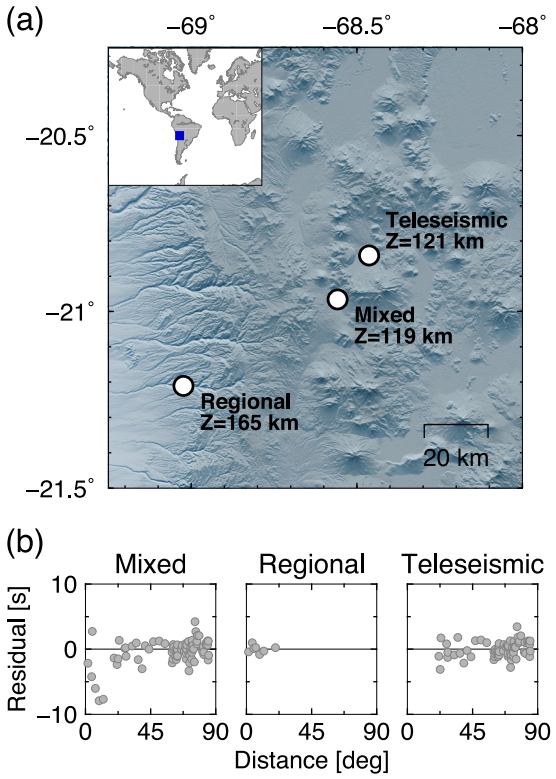

Figure 1. Example illustrating an earthquake that is differently located using either regional or teleseismic arrival times. Note the discrepancies of more than 50 km horizontally and ∼45 km vertically between regional and teleseismic epicentres. (a) Map view and (b) travel-time residuals corresponding to the three solutions. Combining regional and teleseismic stations results not only in another different hypocentre location but also large travel-time residuals at regional stations. Travel times have been calculated using the global 1-D velocity model ak135.

and nonlinear methods. Inaccuracies in the assumed velocity model introduce systematic bias (Billings *et al.* 1994), and hence they are often much more significant than random, uncorrelated errors caused by arrival time picking, which average out given a sufficient number of picks. Indeed, the accuracy of an earthquake location depends on the modelling errors of calculated travel times. Fig. 1 shows how the location of an earthquake in the Chilean subduction zone is biased due to inaccuracies and limitations in the assumed velocity model. Note the discrepancy of more than 50 km between the epicentral solutions based on teleseismic and regional arrival times, respectively.

There are different strategies that can be adopted to lower the effects of unmodelled velocity structure and improve the seismic locations. One approach is to use tomographic velocity models to predict the travel times of seismic phases more accurately. Several studies have demonstrated improved travel-time prediction accuracy using a global 3-D model (Smith & Ekström 1996; Antolik *et al.* 2001; Chen & Willemann 2001; Myers *et al.* 2015). However, although the obtained velocity information provides enhanced absolute hypocentre locations by reducing the biasing effects of large-scale velocity structure, usually only a smooth model can be obtained, which cannot effectively account for small-scale velocity heterogeneities, particularly when applied to global earthquake location. In addition to improved velocity models, several studies showed that applying corrections to travel-time predictions to account for travel-time bias can lead to improvement in location. Travel-time corrections can be computed empirically based on ground-truth (GT) event locations. Travel-time residuals for the GT events are calculated relative to a velocity model and then these calibration points are statistically interpolated to form continuous travel-time correction surfaces (Schultz *et al.* 1998; Myers & Schultz 2000). Myers & Schultz (2000) demonstrated that applying such empirical corrections significantly reduces the bias in location and improve the ability to locate the lower magnitude events. In another technique, the travel-time corrections can be refined by determining a set of station corrections. The use of station corrections along with differential arrival times is the basic concept behind most of the relative location techniques (e.g. Douglas 1967; Jordan & Sverdrup 1981; Pavlis & Booker 1983; Poupinet *et al.* 1984; Got *et al.* 1994; Richards-Dinger & Shearer 2000; Waldhauser & Ellsworth 2000; Myers *et al.* 2007). When the travel-time-prediction errors are dominated by the biasing effects of lateral velocity heterogeneities, the relative location techniques can be highly effective for events within a localized region, provided the local source region is small compared to the heterogeneity. This is because nearby events share common ray-paths to network stations and therefore the residuals to a given station are spatially correlated among these events.

The double-difference algorithm (Waldhauser & Ellsworth 2000) and the shrinking-box source-specific station terms (SSST) method (Richards-Dinger & Shearer 2000; Lin & Shearer 2005) are both techniques that can improve the relative location accuracy among nearby events in the case of distributed seismicity. Despite differences in how the methods work, they solve the same underlying problem. Synthetic tests have shown that these methods yield very similar results for the relative locations among nearby events when applied to identical data sets but that the shrinking-box SSST is slightly better in terms of absolute location accuracy (Lin & Shearer 2005). The shrinking-box SSST has some further advantages: since SSST technique separates the location part from the station correction calculation, any desired single-event location algorithm can be used, including L1-norm misfit measures or grid-search methods. Moreover, SSST is computationally more efficient for very large data sets (e.g. >400 000 events (Lin *et al.* 2007)), because large matrix inversions are not required (Lin & Shearer 2006). Furthermore, large to small scales of velocity heterogeneity are handled through the shrinking-box SSST approach.

In this study, the SSST method has been applied to obtain improved locations of earthquake hypocentres using existing *P* and *pP* phase picks at regional and teleseismic distances. We use this global SSST technique to relocate two different data sets: 3103 shallow to intermediate events in the Chilean margin; and 1680 shallow to very large depth (∼700 km) earthquakes in the Tonga-Fiji subduction zone (Fig. 2). For this purpose, both automatic and manual phase picks provided by the GEOFON bulletin data have been used. GEOFON (http://geofon.gfz-potsdam.de/) is the seismological infrastructure operated by the GFZ German Research Center for Geosciences. Using real-time seismological data from their own broadband network GEOFON Data Centre (1993) and many open and selected restricted station networks around the world, GEOFON determines rapid automatic location estimates for all globally recorded and most larger regional earthquakes, and provides manually revised solutions for at least the largest earthquakes.

2 RELOCATION METHOD

Travel-time residuals to a single seismic station contain the effect of unmodelled 3-D velocity structure and random errors of arrival time

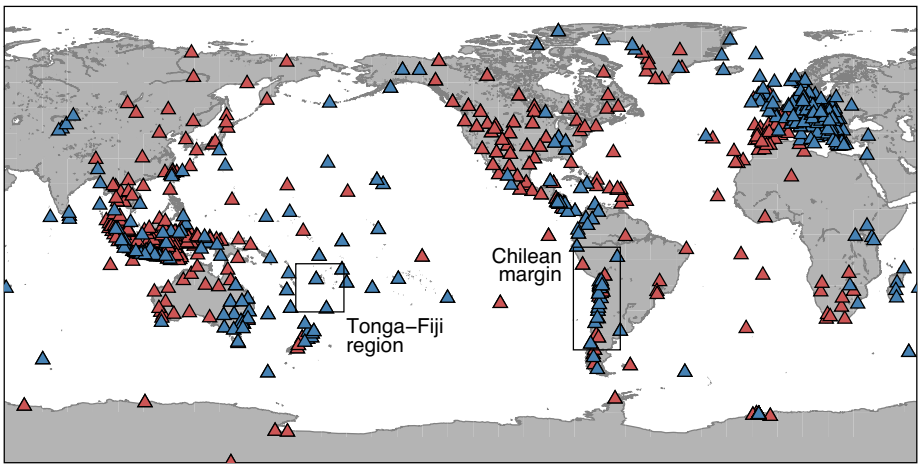

Figure 2. Distribution of seismic stations used to relocate 3103 Chilean events and 1680 earthquakes in the Tonga-Fiji region. Seismic stations recording at least one *pP* depth phase are shown by red triangles. Solid boxes indicate the boundaries of the study regions.

measurements. The latter is uncorrelated and the former is spatially correlated in the source region because close-by events will be affected by similar structure. The goal of the SSST method is to reduce the effect of spatially correlated residuals caused by 3-D velocity structure (Lin *et al.* 2007). By using this technique, a set of station corrections is computed iteratively for each receiver. Therefore, the time corrections for each station vary as a function of source position. The calculated station terms are added to the travel times computed from the velocity model to take unmodelled structure into account. Thus, the travel-time residual misfit function to be minimized is a function of station terms in addition to earthquake hypocentre parameters:

$$\phi(t_i, \boldsymbol{x_i}, s_{ij}) \propto \| T_{ij}^{\text{obs}} - t_i - \tau_j(\boldsymbol{x_i}) - s_{ij} \|_p, \quad (1)$$

where t_i and $\boldsymbol{x_i}$ are the origin time and location of the ith event respectively, s_{ij} is the SSST for the ith event at the jth station, T_{ij}^{obs} is the observed arrival time of the ith event at the jth station, $\tau_j(\boldsymbol{x_i})$ gives the model-predicted travel time from event location $\boldsymbol{x_i}$ to the station j, and $\|.\|_p$ denotes the p-norm. An additional constraint is imposed that the station terms must only be varying slowly as a function of source position. This is a coupled system of equations that can be solved by using a two-step iterative procedure of Frohlich (1979). In each iteration, first, ϕ is minimized with respect to hypocentre parameters (t, $\boldsymbol{x}$) while the station terms s are held fixed. Then the station terms are updated by appropriate averaging of residuals while holding the hypocentres fixed. The basic algorithm is described by the following steps (Richards-Dinger & Shearer 2000):

(i) First, all the earthquakes are located with the actual arrival times.

(ii) Then, an SSST is calculated for each source-receiver pair as the mean of the residuals from at most N neighbouring events which are located within a sphere of radius R_{max} around the target event and have picks at the station in question.

(iii) The original arrival time picks are corrected by the calculated SSSTs ($T_{ij}^{'\text{obs}} = T_{ij}^{\text{obs}} - s_{ij}$).

(iv) Then, the earthquakes are relocated with these corrected arrival times.

(v) The above steps are repeated iteratively.

As mentioned earlier, the so-called shrinking-box SSST technique has been applied in this study. In this method, obtained station corrections vary smoothly as a function of earthquake source location in both position and depth, with the degree of smoothing determined by the number of neighbouring events N and cut-off distance R_{max}. The number N and cut-off distance R_{max} were set to case specific initial values in the first iteration and then were gradually reduced to predefined minimum values in the final iteration. In this study, we chose different values of N and R_{max} for different regions considering the seismicity and the extent of the study area. To relocate events in the Chilean margin (Tonga-Fiji subduction zone), the procedure started with 600 (400) neighbouring events and a cut-off radius of 350 km (450 km) and ended after 10 (7) iterations with 20 (12) nearby events and a maximum inter-event distance of 35 km (55 km). For intermediate iterations, N and R_{max} are set to values spaced evenly in logarithmic scale. In this way, large to small scales of lateral velocity heterogeneities have been modelled during the relocation procedure. It should be pointed out that the iterative approach described above provides good relative locations but does not allow to quantify the absolute location error.

Here, we adopted the SSST relocation procedure for the probabilistic, nonlinear, global-search location method implemented in the software package NonLinLoc (Lomax *et al.* 2000). In a probabilistic framework, a solution to the earthquake location problem is achieved by constructing the *a posteriori* probability density function (PDF) and directly searching it to find optimal solutions (Lomax *et al.* 2009). The *a posteriori* PDF is formulated as a combination of independent states of information including prior information on model parameters obtained independently of data, information obtained from (uncertain) observations, and a PDF describing the uncertainty of the forward problem (travel-time calculation). In NonLinLoc, the posterior PDF can be constructed by either least-squares, L2-norm (LS-L2; Tarantola & Valette 1982) or equal-differential-time (EDT; Font *et al.* 2004; Lomax 2005) likelihood functions. While the former seeks to best satisfy all of the observations simultaneously, the latter tries to achieve its highest value (lowest misfit) by satisfying the most pairs of observations and thus is much more robust in the presence of outliers (Lomax 2005). Using the EDT likelihood function, the location PDF has the following form

(Lomax 2005):

$$\mathrm{PDF}(\boldsymbol{x}) \propto k \left[\sum_{\mathrm{obs}_a, \mathrm{obs}_b} \frac{1}{\sqrt{\sigma_a^2 + \sigma_b^2}} \exp \right.$$
$$\left. \times \left(- \frac{\{[T_{\mathrm{obs}_a}(\boldsymbol{x}) - T_{\mathrm{obs}_b}(\boldsymbol{x})] - [TT_{\mathrm{calc}_a}(\boldsymbol{x}) - TT_{\mathrm{calc}_b}(\boldsymbol{x})]\}^2}{\sigma_a^2 + \sigma_b^2} \right) \right]^N \quad (2)$$

where T_{obs_a} and T_{obs_b} are observed arrival times and TT_{calc_a} and TT_{calc_b} are calculated travel times for stations a and b. σ_a and σ_b are assigned errors for the two observations, k is a normalization factor and N is the total number of observations. The exponential term in eq. (2) has its maximum value at points $\boldsymbol{x}$ where the two differential times (the terms in the braces) are equal. Such points $\boldsymbol{x}$ best satisfy the two observations obs_a and obs_a. In general, sets of points $\boldsymbol{x}$ where the exponential is non-zero construct a finite width, curved surface in 3-D space. In eq. (2), since the sum taken over differential times of all pairs of observations is outside the exponential, the EDT likelihood function has its maximum value for those points $\boldsymbol{x}$ where the most pairs of observations are satisfied and, therefore, is less sensitive to outlier data (Lomax 2005). Furthermore, Lomax *et al.* (2009) performed comprehensive tests with synthetic data comparing the EDT approach with the standard LS-L2 method in the presence of outlier data and incorrect velocity model confirming the robustness of EDT likelihood function. This becomes especially important if automatic picks are used in the location procedure, since in an automated picking process, occasional phase misidentification and picking errors are inevitable.

The resulting posterior PDF is typically irregular and multimodal but can be efficiently sampled by using the oct-tree algorithm (Lomax & Curtis 2001), which is a stable, global-search method. It is based on a recursive subdivision and sampling of cells in 3-D space and importance sampling of the PDF with a high density of sampled cells in the areas of higher PDF values. In our study, we used the oct-tree method to sample the complex EDT PDF in 3-D space to find the maximum likelihood hypocentres of earthquakes. The location procedure thus is less affected by local optima in the model space. Moreover, it has been shown that the horizontal and vertical errors estimated by NonLinLoc are more realistic compared to those derived by linear algorithms (Lippitsch *et al.* 2005). In addition, 3-D velocity models can be used by NonLinLoc although we did not make use of this possibility in the present study.

We solved the spherical forward problem of calculating travel times of *P*- and *pP*-wave arrivals by assuming the global 1-D earth model ak135 (Kennett *et al.* 1995) also used for routine location by GEOFON.

3 APPLICATION TO SUBDUCTION EARTHQUAKES IN THE CHILEAN MARGIN

3.1 Data set

Our first data set includes the arrival times of *P* and *pP* waves from 3103 events in the Chilean subduction zone. The earthquakes occurred from 2010 to 2015 between 15°S–45°S and 62°W–90°W (Fig. 4a). The depth of earthquakes ranges from 2 to 624 km and event magnitudes are between M3.3 and M8.4. The events were recorded at more than 400 global seismic stations. Fig. 2 shows the distribution of seismic stations recording the earthquakes used in this study. The final data set consists of 125 236 *P* and 5327 *pP* arrival time picks. Approximately, two-thirds of the Chilean catalogue had been revised and picked manually by GEOFON analysts. The events were originally located using *scautoloc*, the location module of the SeisComP3 package (https://www.seiscomp3.org/), producing an earthquake catalogue for all available station networks (hereafter referred to as GEOFON catalogue). Rather than using the hypocentres in the GEOFON catalogue as starting locations for our relocation procedure, we relocate all of the events with NonLinLoc as described above. These locations are referred to as single-event locations in the following section.

3.2 Relocation results and residual analysis

Based on the initial locations, first a set of *P*-wave static corrections was calculated for each station and then used as a starting point for the SSST calculation. Our tests showed that this leads to some improvement in the absolute locations of the events because the shallow velocity heterogeneities beneath each station are accounted for by the static terms. For similar reason, the same static station terms were applied to both *P* and *pP* arrival times. Then, in an iterative manner, we estimated the *P*-wave SSSTs for each source-receiver ray paths to reduce the biasing effect of the true 3-D earth structure and then relocated the events by applying the obtained correction terms to *P* phases (at this step, we did not apply the same *P* SSSTs to *pP* phases due to their ray-geometries). Fig. 3(a) shows examples of final calculated *P*-wave SSSTs (relative to the 1-D reference model ak135) for two particular regional stations (GE-LVC and GT-LPAZ) and a teleseismic station (US-GOGA). The station terms for each event recorded at the receiver are plotted at each event location. Spatial coherence of the pattern of the SSSTs indicates the existence of heterogeneities in the real velocity structure. Positive correction terms indicate low velocities and negative values reflect high velocity anomalies on the way of the seismic rays travelling from sources to the stations. Comparison between histograms of *P* travel-time residuals at regional stations observed from locations before (Fig. 3b) and after (Fig. 3c) application of the computed SSSTs reveals significant reduction of their misfit. The median absolute deviation (MAD), mean (μ) and standard deviation (σ) of each distribution are indicated on each subplot. The MAD is a measure of spread that is less sensitive to outliers in the data set than the standard deviation and so often more appropriate in the presence of non-Gaussian errors. Applying SSSTs leads to the MAD reduction of 63 per cent and 56 per cent for stations GE-LVC and GT-LPAZ, respectively. The SSST values for teleseismic station are quite small, since in the lower mantle, where most of the teleseismic P-wave ray path resides, lateral heterogeneities are relatively weak and hence the velocity distribution is smooth.

The final relocated catalogue is shown in map view and in selected east–west cross-sections in Fig. 4. Visual inspection of the SSST hypocentre locations in the cross-sections presents a sharper image of the seismicity associated with the subducting slab compared to the nonlinear, single-event locations. Particularly, a narrower Wadati–Benioff zone is imaged with fewer outlier locations. However, it should be noted that although SSST method reduces the relative errors among nearby events, the absolute locations remain biased since they require knowledge of the true 3-D velocity structure.

In most cases, accurately resolving depth is the most challenging part of earthquake location. Therefore, some indication of the

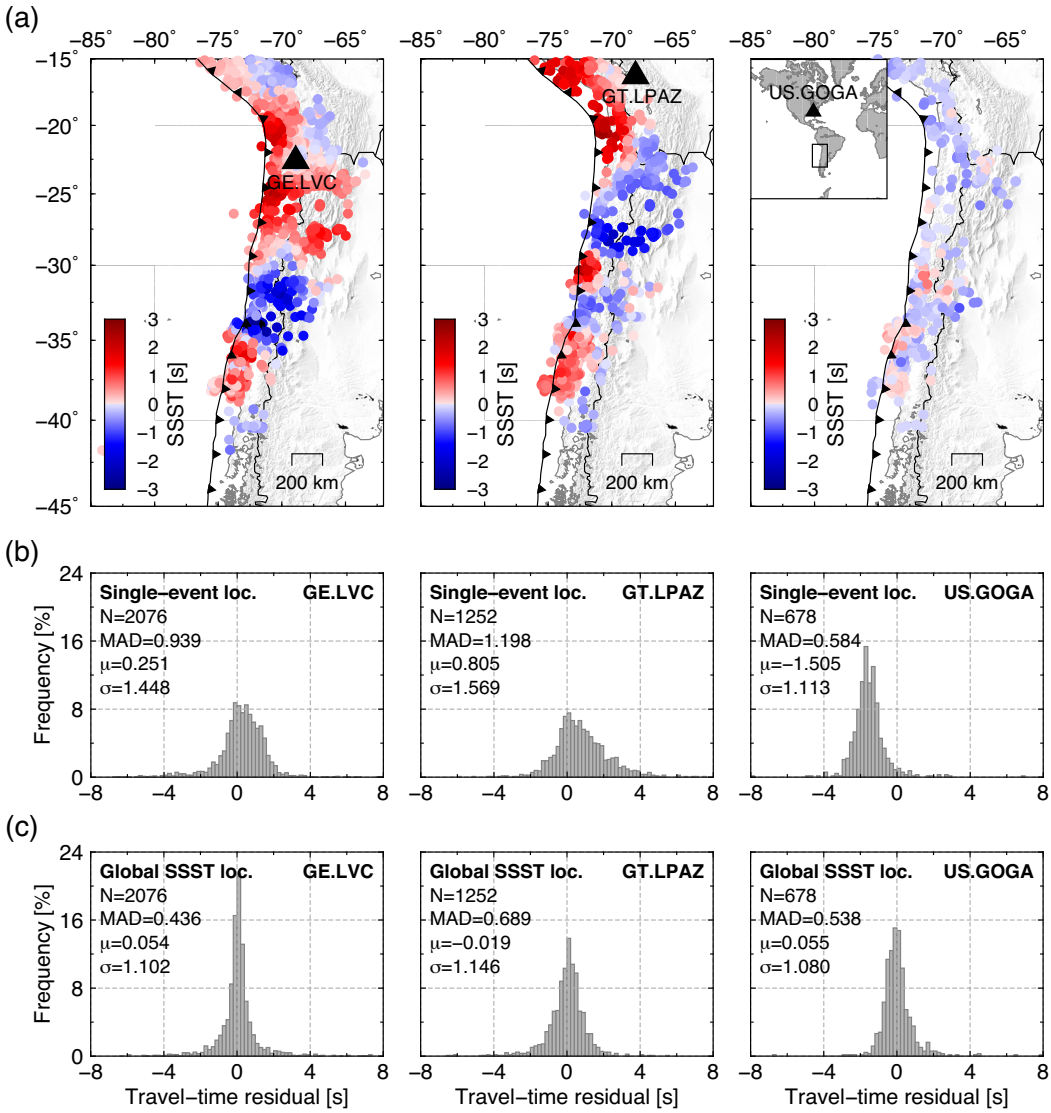

Figure 3. (a) Final *P*-wave source-specific station terms (SSSTs) for two selected regional stations and a teleseismic station (the right-most subplot) calculated from travel-time residuals for events along the Chilean margin recorded by these stations. The SSST values are plotted at the event locations. (b) Histograms of the travel-time residuals from single-event locations and (c) the SSST relocations for the stations shown in Fig. 3(a). The median absolute deviation (MAD), mean (μ) and sample standard deviation (σ) of each distribution are indicated on each subplot.

reliability of an earthquake location method may be achieved by studying the overall distribution of depths in an earthquake catalogue. We compare the histograms of depth distributions from the original GEOFON catalogue, nonlinear single-event locations, and our final SSST locations in Fig. 5. In the GEOFON catalogue, there is a sharp peak at 10–15 km depth, because for some events the depth is fixed to this level when it is poorly constrained. There is also a strong peak at 0–5 km depth in the single-event locations, indicating that there may be no or weak depth control and the maximum likelihood hypocentres calculated by NonLinLoc stuck at the surface. The depth distribution in our final SSST catalogue is more uniform and smooth, especially for shallower depths, with peaks likely reflecting actual trends in the seismicity. By looking at Fig. 4 we can identify the peak at 30–35 km with seismogenic zone activity, and the peak at a little over 100 km with the flat slab in this depth range.

Fig. 6(a) shows a comparison of the histograms of the *P*-wave travel-time residuals for the single-event locations and the SSST locations. A substantial reduction of spread is observed after global SSST relocation. The MAD of the residuals drops from 1.144 s before relocation to 0.528 s after relative location; the root-mean-square (RMS) residual decreases from 1.506 to 1.003 s. Fig. 6(b) presents the residuals before and after SSST relocation as a function of recording distance. Strong variations in initial residuals are smoothed out in the final residuals, especially in the regional distance range where the effect of lithospheric velocity anomalies is strong.

3.3 Comparison with high-resolution earthquake locations

To assess the effectiveness of our relocation procedure, we compared our global SSST locations in two selected areas of the Chilean margin for which earthquake catalogues derived from a dense, local seismic network or by a temporary deployment of seismic sensors for aftershock studies are available. We consider these data sets as

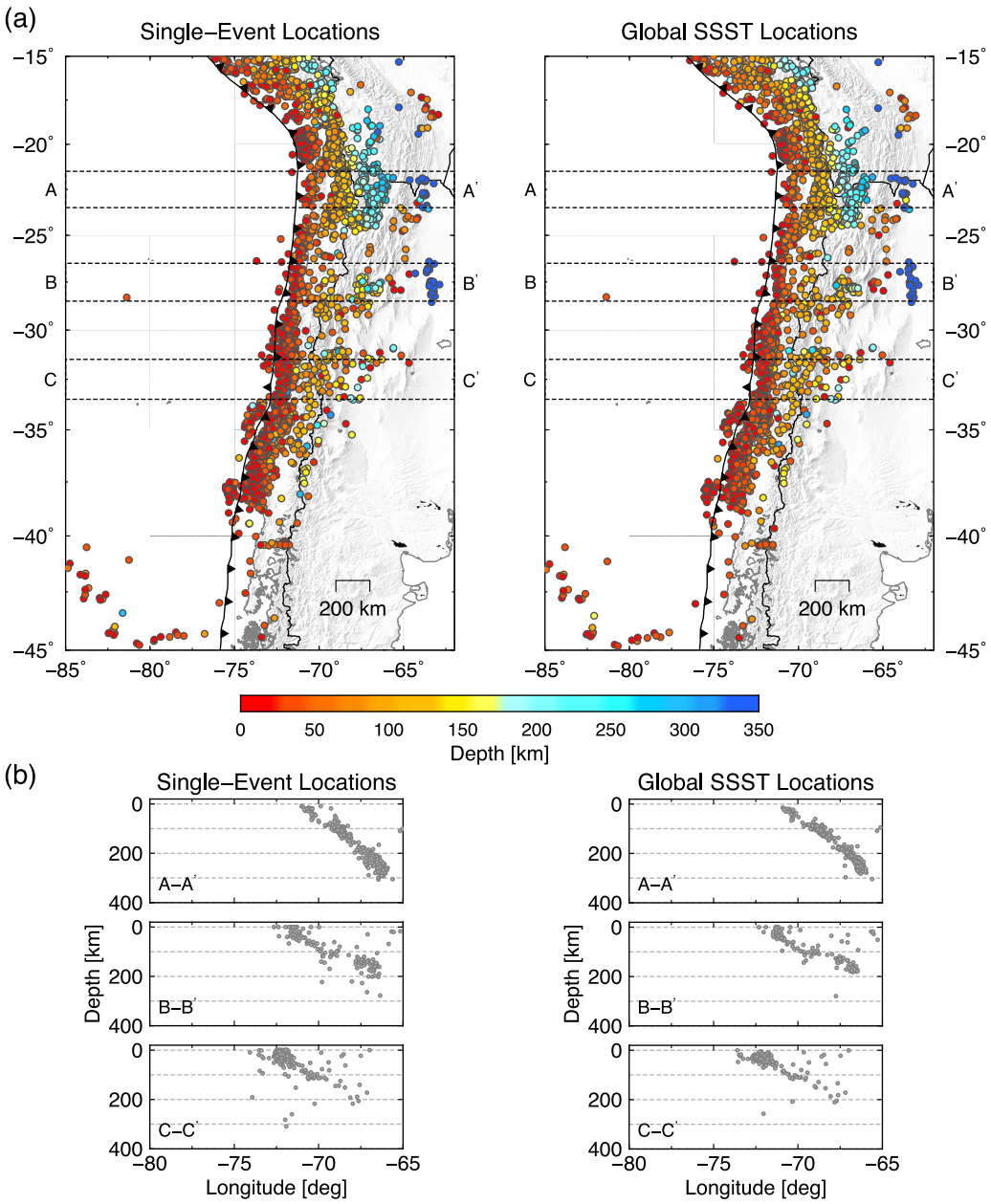

Figure 4. (a) Map view and (b) east–west cross-sections of the 3103 earthquakes in the Chilean subduction zone (left) before and (right) after source-specific station terms relocation. Hypocentre locations in the left panels are from the single-event relocation. Boxes in the map views indicate the locations of the cross-sections. Same events are shown in corresponding subplots.

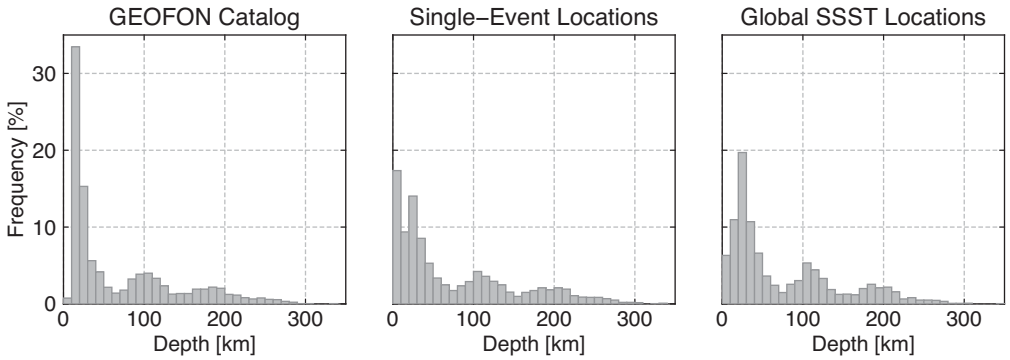

Figure 5. Histograms of depth distributions for different hypocentre locations indicated above each subplot.

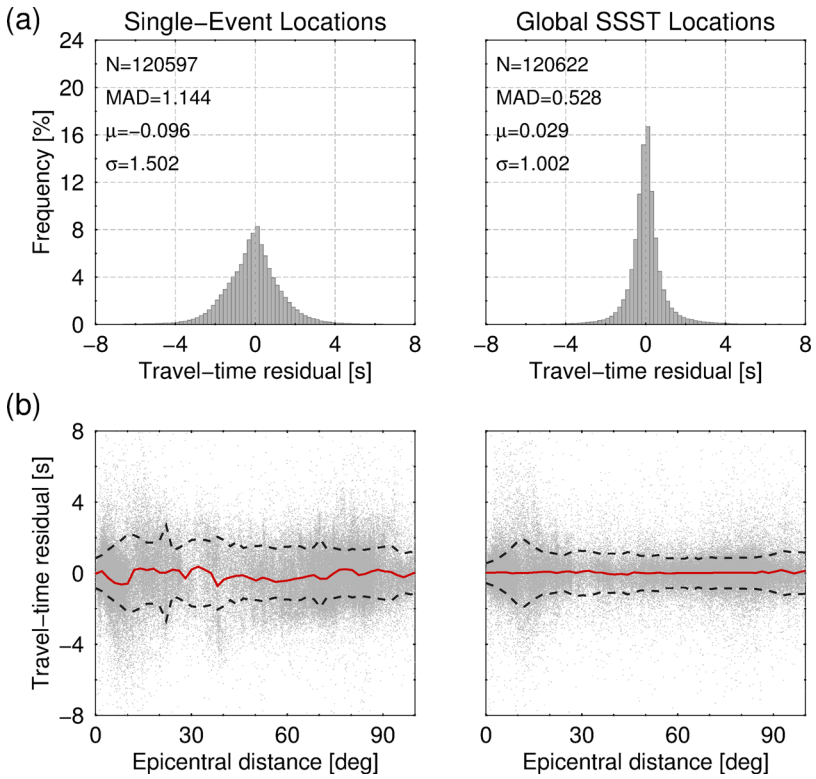

Figure 6. (a) Histograms of travel-time residuals from single-event locations (left) and our global SSST relocations (right). (b) Residuals as a function of epicentral distance. Solid red lines and dashed black lines indicate mean and one standard deviation, respectively. Mean and standard deviation are computed within bins of 2° distance.

the most accurate locations available. They are the closest we can get to actual GT known locations, even though they do not fulfil the formal criteria for GT events (Bondár *et al.* 2004).

The first database is provided by the IPOC (Integrated Plate Boundary Observatory Chile) network (GFZ German Research Centre for Geosciences; Institut des Sciences de l'Univers-Centre National de la Recherche CNRS-INSU 2006) shown by triangles in Figs 7(a) and (b) [catalogue provided by Bernd Schurr as personal communication, 2015; a subset of this catalogue was published in Schurr *et al.* (2014)]. Events roughly within the network and with magnitudes $M \geq 4$ are used as comparison data set. None of the IPOC stations were used in the SSST relocation procedure, such that the data underlying both catalogues are (nearly) independent. The single-event locations (without applying any travel-time corrections) and global SSST locations are visually compared with GT hypocentres in Fig. 7(a). Our results show the importance of the relocation procedure to constrain event depths when using only regional and teleseismic data. Quantitatively, the global SSST application improves the reduction of the mean relative mislocation to 52 per cent horizontally (Fig. 7b) and 55 per cent vertically (Fig. 7c). We thus confirm that our relocation procedure using SSSTs results in global-network hypocentre locations that are more consistent with dense, local-network locations.

We used the aftershocks of the Maule earthquake (M_w8.8 on 2010 February 27) recorded by the dense temporary seismological network IMAD (International Maule Aftershock Deployment; Beck *et al.* 2014) as a second reference set to further evaluate the validity of the global SSST results. The station spacing of ∼30 km provides a good resolution of hypocentral coordinates, including depth, for this database [catalogue published in Lange *et al.* (2012)], and none of the IMAD stations contributed to the SSST locations. Fig. 8(a) illustrates how the relocation procedure recovers the hypocentre locations better. Qualitatively, it can be seen that the global SSST improves the agreement with the local hypocentre locations and reduces systematic bias. Moreover, statistical measures of differences in hypocentre locations (Figs 8b and c) show a reduction of horizontal and vertical differences in location after applying SSST corrections.

4 APPLICATION TO DEEP EARTHQUAKES IN THE TONGA-FIJI SUBDUCTION ZONE

The Tonga-Fiji subduction zone is the most seismically active subduction zone worldwide and one of the best regions for imaging of the down-going slab and mantle wedge structure. It stretches from the North Island of New Zealand to the north–northeast and is about 2600 km long. The depth of earthquakes in the Tonga subduction zone ranges from very shallow (∼5 km) to very deep (∼700 km) and this region accounts for about two thirds of the globally recorded deep seismicity (>300 km depth). Many seismological studies identified complex morphology of the subducted lithosphere and anomalous mantle beneath the Tonga Trench, such as lateral mantle displacements (Gurnis *et al.* 2000), anomalously hot mantle (Chen & Brudzinski 2001), and mantle upwelling (Pysklywec *et al.* 2003). The consequence of this structural complexity is large uncertainty in hypocentre locations in the Tonga-Fiji subduction zone listed in global catalogues like the GEOFON bulletin that use 1-D velocity models for computing earthquake locations.

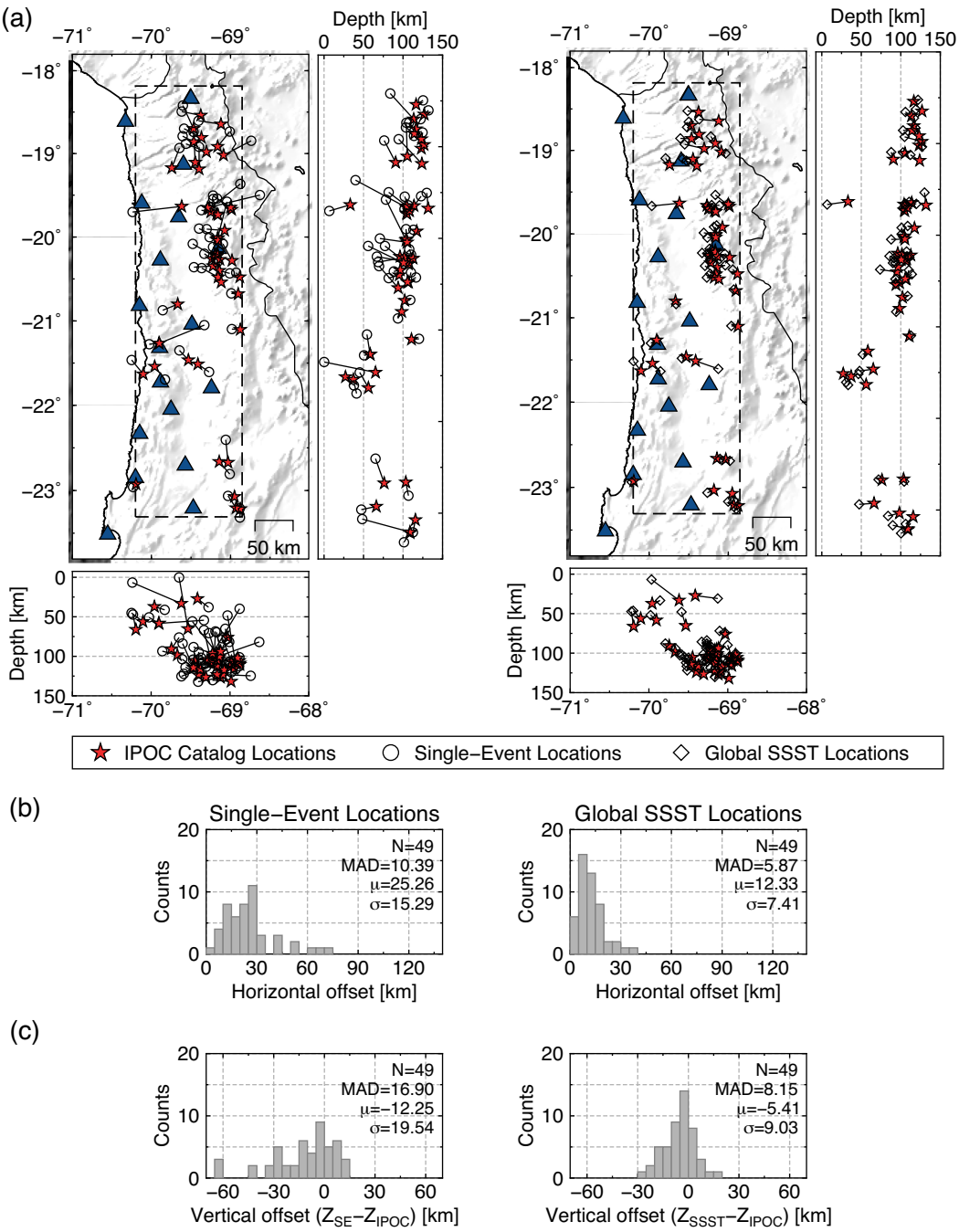

Figure 7. (a) Map views and cross-sections of the seismicity in northern Chile, where high-resolution earthquake locations (stars) are available from the local IPOC network (triangles). Events roughly within the network (dashed box) and with magnitudes $M \geq 4$ are used as 'ground-truth' (GT) events (in spite of not fulfilling formal criteria for GT events). Nonlinear, single-event locations (left) and global SSST relocations obtained in this study (right) are compared with these GT events. Solid lines indicate mislocations between corresponding hypocentres. (b) Horizontal and (c) vertical differences between local hypocentres and single-event locations (left) and between local hypocentres and global SSST locations (right) are shown as histograms. MAD, mean relative mislocation (μ) and sample standard deviation (σ) of each distribution are shown in each subplot.

4.1 Data set

We use 1680 events that occurred in the Tonga-Fiji subduction from 2010 to 2014 to evaluate the ability of global SSST method to relocate hypocentres in this region (Fig. 2). Body-wave magnitudes range between 4.1 and 7.5 with 685 events having $M \geq 5$. The events span a depth range of 5-682 km, making them well suited to investigate the effect of upper mantle and transition zone on global SSST solutions. The original data set is composed of 113 931 *P* and 1726 *pP* phases included in the GEOFON bulletin. Less than 1 per cent of phases recorded at local distances (0°–2.5°), ∼13 per cent at regional and upper-mantle distances (2.5°–28°), and ∼86 per cent at teleseismic distances (28°–95°) (Fig. 9). Consequently, earthquake location inversion leads to large travel-time residuals at the

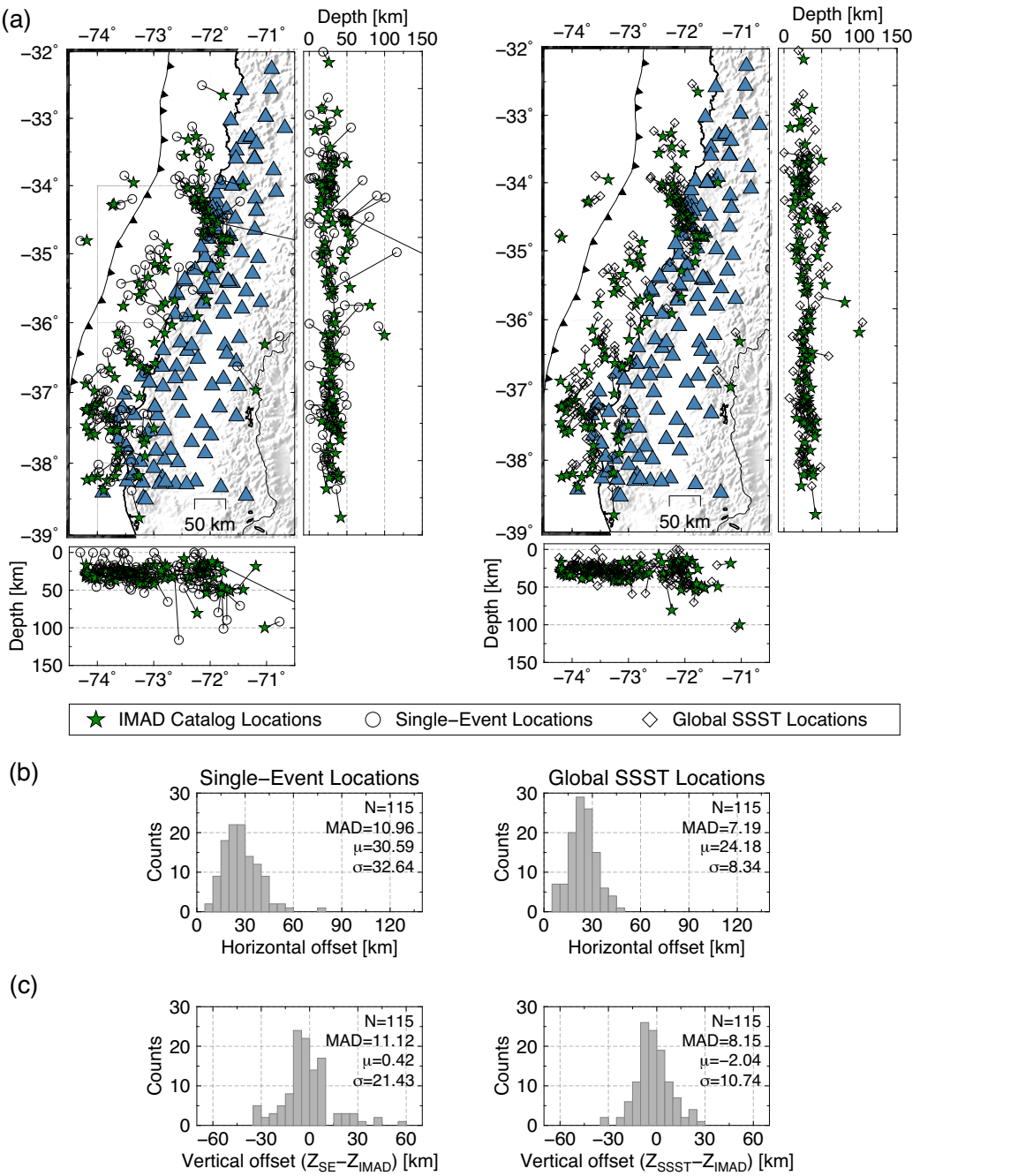

Figure 8. (a) Comparison of locations for 115 earthquakes in the 2010 Maule aftershock sequence (Lange *et al.* 2012) that were both locally and teleseismically recorded. The locations determined using local data (stars) available from a dense seismological network (triangles) are used as GT data set. Single-event locations (left) and our global SSST relocations (right) are compared with these GT events. Solid lines indicate mislocations between corresponding hypocentres. (b) Horizontal and (c) vertical differences between local hypocentres and single-event locations (left) and between local hypocentres and global SSST locations (right) shown as histograms.

few near stations to better fit the abundance of distant data as in the first example. These large travel-time residuals also could result from inadequate representation of heterogeneous upper-mantle structure in 1-D global velocity models. Like for the Chilean example case, instead of using initial locations from the GEOFON bulletin, the whole data set was first relocated by using the NonLinLoc program and then a set of static station terms was calculated based on these starting locations and applied before SSST relocation procedure.

4.2 Source-specific station terms and statistical analysis of residuals

Using the algorithm described in Section 2, the *P*-wave SSSTs were iteratively computed for each source-receiver pair recorded at a given station. The station term values are a smoothed version of the event-specific residual field for each station. Figs 10(a) and (c) show examples of final SSSTs for two different regional stations both of which show coherent patterns. The most obvious features in these

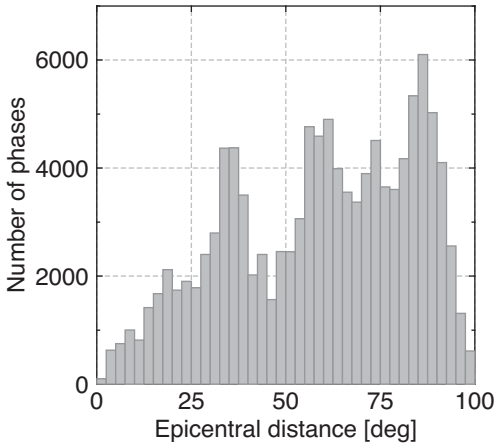

Figure 9. Histogram of number of recorded phases as a function of epicentral distance for earthquakes in the Tonga-Fiji region.

figures are quite large positive SSST values at station II-MSVF (Fig. 10a) and large negative correction terms at station AU-NIUE (Fig. 10c) from sources located at shallow to intermediate depths. This can easily be interpreted by considering that these stations are located on opposite sides of the Tonga Trench. Along the Tonga-Fiji subduction zone, the Pacific Plate to the east is subducting beneath the Tonga Plate to the west. Several studies have determined tomographic models of this region, which delineate the Tonga Wadati-Benioff zone as a fast subducted slab with a slow mantle wedge structure in the back arc region (van der Hilst 1995; Gorbatov & Kennett 2003). Westward propagating seismic waves produced by earthquakes that occur at shallow to intermediate depths in the subduction zone will travel mostly through high attenuation and low velocity mantle wedge to arrive at the station II-MSVF (Fig. 10a). In contrast, seismic waves from the same source but propagating eastwards towards station AU-NIUE will pass through the high velocity slab (Fig. 10c). This effect is smaller for deeper earthquakes

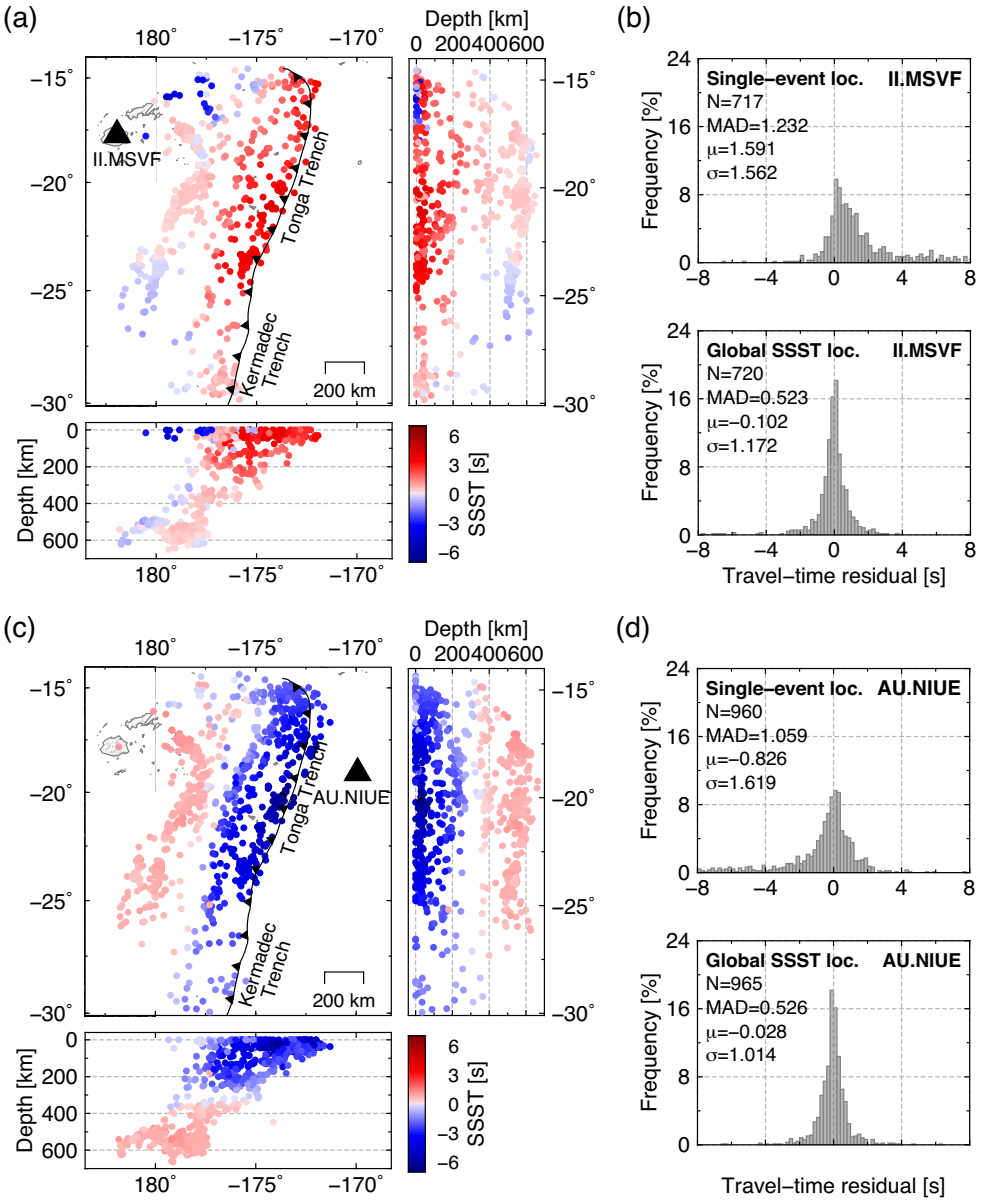

Figure 10. *P* SSSTs for particular stations (a) II-MSVF and (c) AU-NIUE, calculated from travel-time residuals for earthquakes in the Tonga-Fiji subduction zone recorded by these stations. The SSST values are plotted at the event locations. (b,d) Histograms of the travel-time residuals for stations II-MSVF and AU-NIUE before and after global SSST relocation.

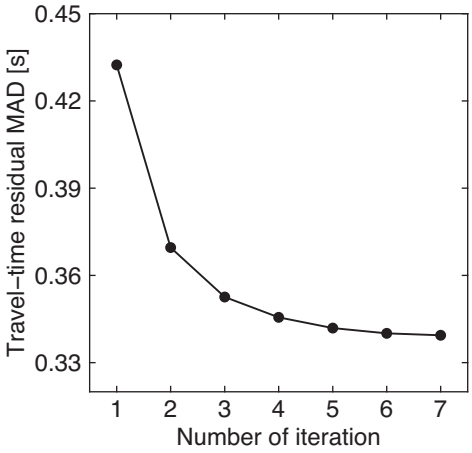

Figure 11. Reduction of travel-time residual MAD for the 1680 events in the Tonga-Fiji subduction zone with iteration numbers.

where the slab steepens up. Roughly similar SSST patterns are thus found for deep events (Figs 10a and c).

Significant improvements to residual reduction are achieved by using our global relocation procedure (Figs 10b and d). The MAD of the travel-time residuals decreases by 58 per cent and 51 per cent at regional stations II-MSVF and AU-NIUE, respectively. For the whole catalogue a gradual reduction of travel-time residual MAD from the 1680 events with iteration number in our global SSST calculation can be observed (Fig. 11). For comparison, the RMS residual decreased from 1.050 s to 0.912 s.

4.3 Relocated seismicity; close-ups and cross-sections

The differences between the single-event locations and our SSST relocations are most obvious in some depth cross-sections and close-ups of spatial clusters of earthquakes. Here, we select a few examples for detailed comparisons. Figs 12(a) and (b) show in map view and cross-sections a comparison of global SSST locations and

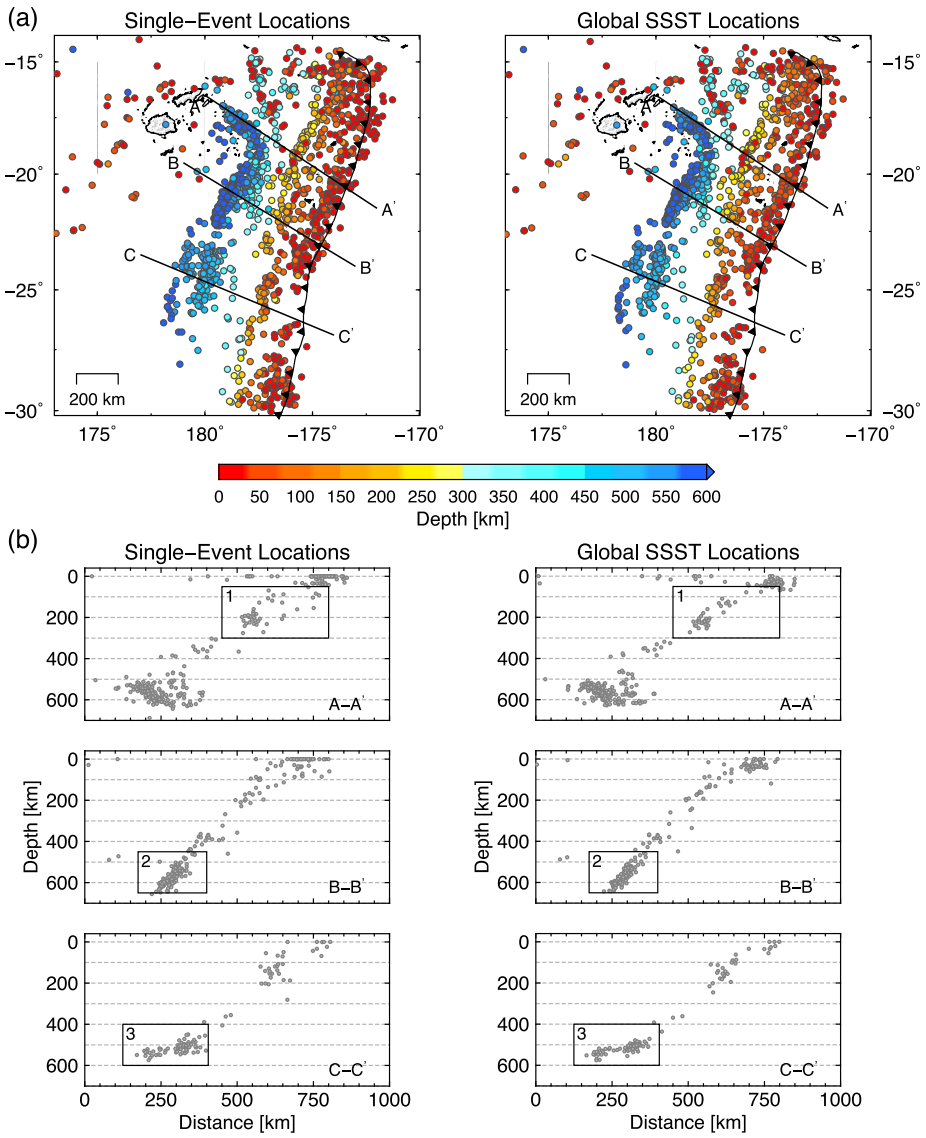

Figure 12. (a) Map view and (b) depth cross-sections of the 1680 earthquakes in the Tonga-Fiji region determined using (left) single-event location method and (right) global SSST relocation procedure. Boxes in panel (b) include events shown in panel (c). Same events are plotted in corresponding subplots. (c) Single-event (left) and global SSST (right) hypocentres of the events within the boxes shown in panel (b).

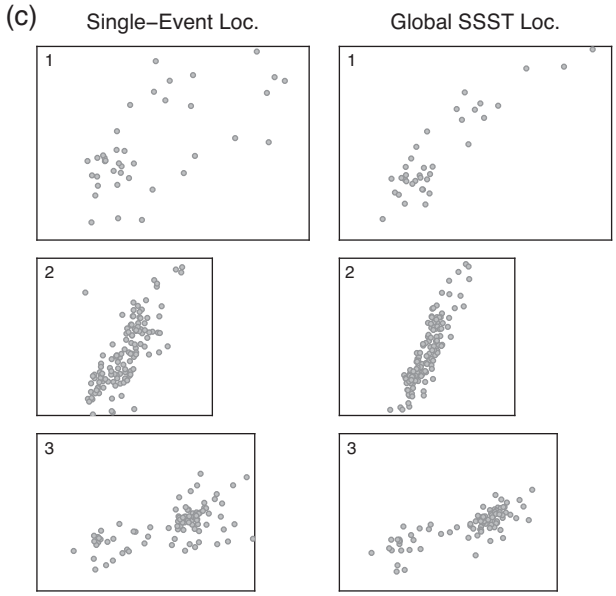

Figure 12 (*Continued*).

single-event locations. Because we do not have access to any high-resolution local catalogue of events in this region, we consider the spatial seismicity pattern as a qualitative measure of the improvement in global SSST relocations. Relative locations of the events appear to have visibly improved after the SSST relocation. This is visible in the decreased scatter in the new hypocentre locations, which better define the slab geometry. Additionally, events whose depths were poorly constrained by the single-event location and were thus previously stuck at the surface now have more plausible depths, removing the horizontal stripe artefacts at shallow depths seen in some cross-sections for the single-event locations (Fig. 12b). At a smaller scale our SSST locations are improved compared to the single-event locations, as measured by the degree of spatial scattering and alignment of seismicity into planar features, in particular in the deep slab (Fig. 12c) where improved clustering can be observed. Finally, we tried to simulate improvement in regional earthquake location that can be achieved by using pre-computed travel-time corrections. Events with at least 100 picks were selected and first located using only teleseismic observations (28°–95°). Then, they were located using only regional and near-regional phases (2.5°–20°) once without applying any time corrections and once refined with station terms determined by SSST method. These two sets of regional solutions are separately compared with teleseismic locations and shown in Fig. 13. Although it leaves room for improved locations, it shows that relocation solutions based on regional phases refined by pre-computed SSSTs are more consistent with teleseismic solutions. This promises the feasibility of using pre-calculated SSSTs to locate small magnitude events recorded regionally, or where fast location based on regional stations is needed.

5 SUMMARY AND CONCLUSION

We have successfully implemented and tested a global SSST technique based on the procedures outlined in Richards-Dinger & Shearer (2000) and Lin & Shearer (2005). We adopted the method to a global setting and employed a nonlinear, global-search method. We tested the method in two regions with strongly heterogeneous structure, relocating 3103 earthquakes along the Chilean margin and 1680 events in the Tonga-Fiji subduction zone using primary and depth phases recorded at regional and teleseismic stations. For our relocated catalogues of events in the Chilean margin (Tonga-Fiji subduction zone), average shifts relative to single-event locations are ∼18 (12) km in epicentres and ∼10 (13) km in depth. In general, the relocation procedure results in sharpening seismicity features associated with subducting slabs. Substantial improvement of focal depths is one of the highlights of our relocation results, demonstrating how effectively the global SSST approach can handle the model errors associated with regional structures such as subducting slabs, which otherwise can severely bias depth estimation. Furthermore, residual statistics show a strong reduction (∼40–60 per cent) of travel-time residual MAD at regional stations, reducing regional biases. In addition to better defining slab geometry by decreasing location scatter and reducing travel-time residuals, comparison with high-resolution earthquake locations reveal the potential of the SSST method to improve hypocentre locations on a global scale.

Our results represent the first step in an effort to relocate global bulletin data sets in a systematic way. In addition to applying the algorithm to larger catalogues of events and updating the catalogues on a rolling basis in the future as more data become available, some other improvements can be done to the extension of the relocation algorithm. It should be noted that our relocation results at this point represent our best estimates based on phase pick data alone. Further improvements are possible by extending the global SSST method and using differential times obtained via waveform cross correlation

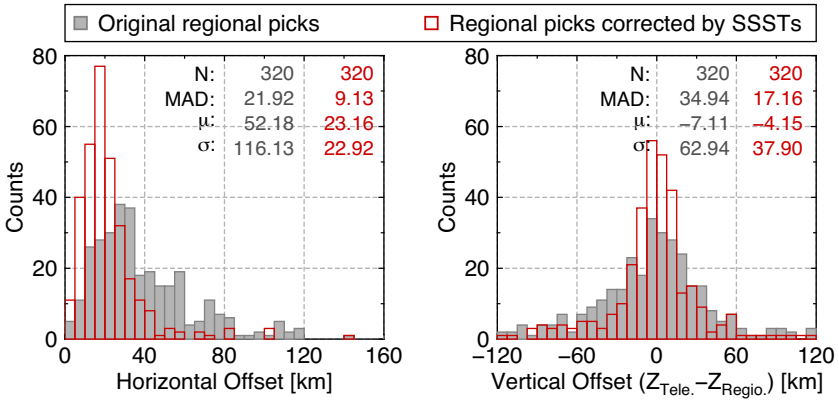

Figure 13. Horizontal and vertical relative mislocation of regional solutions without station terms and teleseismic hypocentres (grey) and regional locations corrected by the pre-calculated SSSTs and teleseismic hypocentres (red) are shown as histograms. The event mislocation is reduced when SSST corrections are applied.

to achieve high precision earthquake locations (Shearer *et al.* 2005; Lin *et al.* 2007; Matoza *et al.* 2013). Our future work will focus on developing a global SSST algorithm using cross correlation methods applied to regional and teleseismic phases.

Since our relocation procedure is based on a global-search location method, it could be implemented for a global 3-D tomography model [e.g. LLNL-G3Dv3 model (Simmons *et al.* 2012)] to achieve further improvement in absolute location accuracy. Travel times from stations to the model spatial grid points can be computed once and saved as travel-time grid files and retrieved during the location procedure.

One of the goals of earthquake relocation studies is to integrate the relocation algorithms to near-real-time processing to locate new events with the accuracy of the relocated reference catalogue (e.g. Got *et al.* 2002; Waldhauser 2009). Rapid estimation of spatial location and origin time of a new event has particularly useful applications in earthquake and volcano monitoring and geohazard mitigation. For these purposes, relocation of events in near-real-time is a high priority. It is clear that a reprocessing of the whole catalogue is not feasible in this context. Instead, pre-computed SSSTs must be stored in a database at appropriate granularity. The generation of this database requires not only an *a priori* known high-resolution catalogue of past seismicity and fast access to an archive of past seismic data, but also an automated SSST procedure that can update the pre-computed terms at regular intervals. Moreover, as discussed by (Shearer *et al.* 2005), since relative location solutions are dependent on all nearby events, adding a new event to the catalogue would affect the locations of older events in the catalogue and the whole data set would need to be relocated for consistency.

ACKNOWLEDGEMENTS

We thank Anthony Lomax for developing and supporting the NonLinLoc software package (http://alomax.free.fr/nlloc/), and Bernd Schurr for providing us with his catalogue of events for the IPOC area. We further thank all the global networks which have contributed their data. We also thank Prof. Egill Hauksson, Prof. Eric Engdahl, and one anonymous reviewer for their detailed and thoughtful comments that helped to improve the manuscript. The fist author expresses his thanks to Mehdi Nikkhoo for reviews and constructive comments. All Maps and figures in this article were created using Generic Mapping Tools (GMT) software (Wessel *et al.* 2013).

REFERENCES

Antolik, M., Ekström, G. & Dziewonski, A., 2001. Global event location with full and sparse data sets using three-dimensional models of mantle *P*-wave velocity, *Pure appl. Geophys.*, **158**(1), 291–317.

Beck, S. *et al.*, 2014. Advancing subduction zone science after a big earthquake, *EOS, Trans. Am. geophys. Un.,* **95,** 193–194.

Billings, S.D., Sambridge, M.S. & Kennett, B.L.N., 1994. Errors in hypocenter location: picking, model, and magnitude dependence, *Bull. seism. Soc. Am.,* **84**(6), 1978–1990.

Bondár, I., Myers, S.C., Engdahl, E.R. & Bergman, E.A., 2004. Epicentre accuracy based on seismic network criteria, *Geophys. J. Int.,* **156**(3), 483–496.

Chen, Q.-f. & Willemann, R.J., 2001. Global test of seismic event locations using three-dimensional earth models, *Bull. seism. Soc. Am.,* **91**(6), 1704–1716.

Chen, W.P. & Brudzinski, M.R., 2001. Evidence for a large-scale remnant of subducted lithosphere beneath Fiji, *Science,* **292,** 2475–2479.

Douglas, A., 1967. Complex morphology of subducted lithosphere in the mantle beneath the Tonga trench, *Nature,* **215,** 47–48.

Font, Y., Kao, H., Lallemand, S., Liu, C.-S. & Chiao, L.-Y., 2004. Hypocentre determination offshore of eastern Taiwan using the Maximum Intersection method, *Geophys. J. Int.,* **158**(2), 655–675.

Frohlich, C., 1979. An efficient method for joint hypocenter determination for large groups of earthquakes, *Comput. Geosci.,* **5**(3–4), 387–389.

GEOFON Data Centre, 1993. *GEOFON Seismic Network,* Deutsches GeoForschungsZentrum GFZ.

GFZ German Research Centre for Geosciences; Institut des Sciences de l'Univers-Centre National de la Recherche CNRS-INSU 2006. *IPOC Seismic Network,* doi:10.14470/PK615318.

Gorbatov, A. & Kennett, B., 2003. Joint bulk-sound and shear tomography for Western Pacific subduction zones, *Earth planet. Sci. Lett.,* **210**(3-4), 527–543.

Got, J.-L., Fréchet, J. & Klein, F.W., 1994. Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea, *J. geophys. Res.,* **99**(B8), 15 375–15 386.

Got, J.L., Okubo, P., Machenbaum, R. & Tanigawa, W., 2002. A real-time procedure for progressive multiplet relative relocation at the Hawaiian Volcano Observatory, *Bull. seism. Soc. Am.,* **92**(5), 2019–2026.

Gurnis, M., Ritsema, J., Van Heijst, H.-J. & Zhong, S., 2000. Tonga slab deformation: the influence of a lower mantle upwelling on a slab in a young subduction zone, *Geophys. Res. Lett.,* **27**(16), 2373–2376.

Husen, S. & Hardebeck, J.L., 2010. Earthquake location accuracy, *Community Online Resource for Statistical Seismicity Analysis,* **10,** doi:10.5078/corssa-55815573.

Jordan, T.H. & Sverdrup, K.A., 1981. Teleseismic location techniques and their application to earthquake clusters in the south-central pacific, *Bull. seism. Soc. Am.,* **71**(4), 1105–1130.

Kennett, B. L.N., Engdahl, E.R. & Buland, R., 1995. Constraints on seismic velocities in the earth from travel times, *Geophys. J. Int.,* **122**(1), 108–124.

Lange, D. *et al.*, 2012. Aftershock seismicity of the 27 February 2010 Mw 8.8 Maule earthquake rupture zone, *Earth planet. Sci. Lett.,* **317–318,** 413–425.

Lin, G. & Shearer, P., 2005. Tests of relative earthquake location techniques using synthetic data, *J. geophys. Res.,* **110**(B4), B04304, doi:10.1029/2004jb003380.

Lin, G. & Shearer, P., 2006. The COMPLOC earthquake location package, *Seismol. Res. Lett.,* **77**(4), 440–444.

Lin, G., Shearer, P.M. & Hauksson, E., 2007. Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *J. geophys. Res.,* **112**(B12), B12309, doi:10.1029/2007JB004986.

Lippitsch, R., White, R.S. & Soosalu, H., 2005. Precise hypocentre relocation of microearthquakes in a high-temperature geothermal field: the Torfajökull central volcano, Iceland, *Geophys. J. Int.,* **160**(1), 371–388.

Lomax, A., 2005. A reanalysis of the hypocentral location and related observations for the great 1906 California earthquake, *Bull. seism. Soc. Am.,* **95**(3), 861–877.

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

Lomax, A., Virieux, J., Volant, P. & Berge-Thierry, C., 2000. Probabilistic earthquake location in 3-D and layered models, in *Advances in Seismic Event Location,* Vol. 18 of Modern Approaches in Geophysics, pp. 101–134, eds Thurber, C.H. & Rabinowitz, N., Springer.

Lomax, A., Michelini, A. & Curtis, A., 2009. Earthquake location, direct, global-search methods, in *Encyclopedia of Complexity and Systems Science,* pp. 1–33, ed. Meyers, R.A., Springer.

Matoza, R.S., Shearer, P.M., Lin, G., Wolfe, C.J. & Okubo, P.G., 2013. Systematic relocation of seismicity on Hawaii Island from 1992 to 2009 using waveform cross correlation and cluster analysis, *J. Geophys. Res.,* **118**(5), 2275–2288.

Myers, S.C. & Schultz, C.A., 2000. Improving sparse network seismic location with Bayesian kriging and teleseismically constrained calibration events, *Bull. seism. Soc. Am.,* **90**(1), 199–211.

Myers, S.C., Johannesson, G. & Hanley, W., 2007. A Bayesian hierarchical method for multiple-event seismic location, *Geophys. J. Int.,* **171**(3), 1049–1063.

Myers, S.C., Simmons, N.A., Johannesson, G. & Matzel, E., 2015. Improved regional and teleseismic *P*-wave travel-time prediction and event location using a global 3D velocity model, *Bull. seism. Soc. Am.,* **105**(3), 1642–1660.

Pavlis, G.L., 1986. Appraising earthquake hypocenter location errors: a complete, practical approach for single-event locations, *Bull. seism. Soc. Am.,* **76**(6), 1699–1717.

Pavlis, G.L. & Booker, J.R., 1983. Progressive multiple event location (PMEL), *Bull. seism. Soc. Am.,* **73**(6A), 1753–1777.

Poupinet, G., Ellsworth, W.L. & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. geophys. Res.,* **89**(B7), 5719–5731.

Pysklywec, R.N., Mitrovica, J.X. & Ishii, M., 2003. Mantle avalanche as a driving force for tectonic reorganization in the southwest Pacific, *Earth planet. Sci. Lett.,* **209**(1-2), 29–38.

Richards-Dinger, K.B. & Shearer, P.M., 2000. Earthquake locations in southern California obtained using source-specific station terms, *J. geophys. Res.,* **105**(B5), 10 939–10 960.

Schultz, C.A., Myers, S.C., Hipp, J. & Young, C.J., 1998. Nonstationary Bayesian kriging: a predictive technique to generate spatial corrections for seismic detection, location, and identification, *Bull. seism. Soc. Am.,* **88**(5), 1275–1288.

Schurr, B. *et al.*, 2014. Gradual unlocking of plate boundary controlled initiation of the 2014 Iquique earthquake, *Nature,* **512**(7514), 299–302.

Shearer, P., Hauksson, E. & Lin, G., 2005. Southern California hypocenter relocation with waveform cross-correlation, part 2: results using source-specific station terms and cluster analysis, *Bull. seism. Soc. Am.,* **95**(3), 904–915.

Simmons, N.A., Myers, S.C., Johannesson, G. & Matzel, E., 2012. LLNL-G3Dv3: global P wave tomography model for improved regional and teleseismic travel time prediction, *J. geophys. Res.,* **117**(B10), B10302, doi:10.1029/2012JB009525.

Smith, G.P. & Ekström, G., 1996. Improving teleseismic event locations using a three-dimensional Earth model, *Bull. seism. Soc. Am.,* **86**(3), 788–796.

Tarantola, A. & Valette, B., 1982. Inverse problems = quest for information, *J. Geophys.,* **50,** 159–170.

van der Hilst, R.D., 1995. Complex morphology of subducted lithosphere in the mantle beneath the Tonga trench, *Nature,* **374,** 154–157.

Waldhauser, F., 2009. Near-real-time double-difference event location using long-term seismic archives, with application to northern California, *Bull. seism. Soc. Am.,* **99**(5), 2736–2748.

Waldhauser, F. & Ellsworth, W.L., 2000. A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California, *Bull. seism. Soc. Am.,* **90**(6), 1353–1368.

Wessel, P., Smith, W. H.F., Scharroo, R., Luis, J. & Wobbe, F., 2013. Generic Mapping Tools: improved version released, *EOS, Trans. Am. geophys. Un.,* **94**(45), 409–410.