Geophysical Journal International

Geophys. J. Int. (2011) **186,** 1220–1244 doi: 10.1111/j.1365-246X.2011.05107.x

Improved location procedures at the International Seismological Centre

István Bondár and Dmitry Storchak

International Seismological Centre, Thatcham, Berkshire, RG19 4NS, UK. E-mail: istvan@isc.ac.uk

Accepted 2011 June 5. Received 2011 April 4; in original form 2010 December 21

SUMMARY

The International Seismological Centre (ISC) is a non-governmental, non-profit organization with the primary mission of producing the definitive account of the Earth's seismicity. The ISC Bulletin covers some 50 yr (1960–2011) of seismicity. The recent years have seen a dramatic increase both in the number of reported events and especially in the number of reported phases, owing to the ever-increasing number of stations worldwide. Similar ray paths will produce correlated traveltime prediction errors due to unmodelled heterogeneities in the Earth, resulting in underestimated location uncertainties, and for unfavourable network geometries, location bias. Hence, the denser and more unbalanced the global seismic station coverage becomes, the less defensible is the assumption (that is the observations are independent), which is made by most location algorithms.

To address this challenge we have developed a new location algorithm for the ISC that accounts for correlated error structure, and uses all IASPEI standard phases with a valid *ak135* traveltime prediction to obtain more accurate event locations. In this paper we describe the new ISC locator, and present validation tests by relocating the ground truth events in the IASPEI Reference Event List, as well as by relocating the entire ISC Bulletin.

We show that the new ISC location algorithm provides small, but consistent location improvements, considerable improvements in depth determination and significantly more accurate formal uncertainty estimates. We demonstrate that the new algorithm, through the use of later phases and testing for depth resolution, considerably clusters event locations more tightly, thus providing an improved view of the seismicity of the Earth.

Key words: Computational seismology; Theoretical seismology.

INTRODUCTION

The International Seismological Centre (ISC), owing to its non-governmental status, represents a unique international collaboration in seismology. The ISC is an international non-profit organization supported by the U.S. National Science Foundation (NSF), the Royal Society in the United Kingdom, the Japan Meteorological Agency (JMA), the Russian Academy of Sciences, the China Earthquake Administration (CEA), the Indian Meteorological Department (IMD), the Department of Natural Resources of Canada and some 50 other institutions worldwide—seismological observatories, universities and governmental research and operational institutions responsible for producing national earthquake bulletins.

The primary mission of the International Seismological Centre is to produce the definitive account of the Earth's seismicity. To achieve this goal, the ISC collects bulletins (hypocentres, fault plane solutions, moment tensors, magnitudes, reports of shaking intensity, phase picks and amplitude readings) from seismological institutions all over the world in order to acquire the most complete set of station reports for each seismic event (Willemann & Storchak 2001). The data in the bulletins are merged and events meeting certain selection criteria (for instance, small events reported by just one agency are not reviewed) are relocated and carefully reviewed by the ISC analysts to produce the ISC Bulletin. Even though only approximately 20 per cent of all events are reviewed, every reported event is stored in the ISC database and accessible through the ISC website (www.isc.ac.uk). Before year 1999 small reported events were not included in the bulletin for operational reasons.

The reviewed ISC Bulletin is typically 24-months behind real time as the ISC has to wait until all bulletins and station reports are collected (for instance, some agencies publish their reviewed bulletins annually). Fig. 1 shows that in recent years both the number of reported events and the number of phases associated to events have increased exponentially. The phenomenal growth in data volume is attributed to the fact that in the past two decades the number of stations in local, national and regional networks has dramatically increased. The International Monitoring System (www.ctbto.org) network is also quite prolific in reporting phases. Recently extremely dense networks (e.g. USArray, IberArray) have been deployed and began reporting phases to the ISC. Hence, the ISC has to deal with

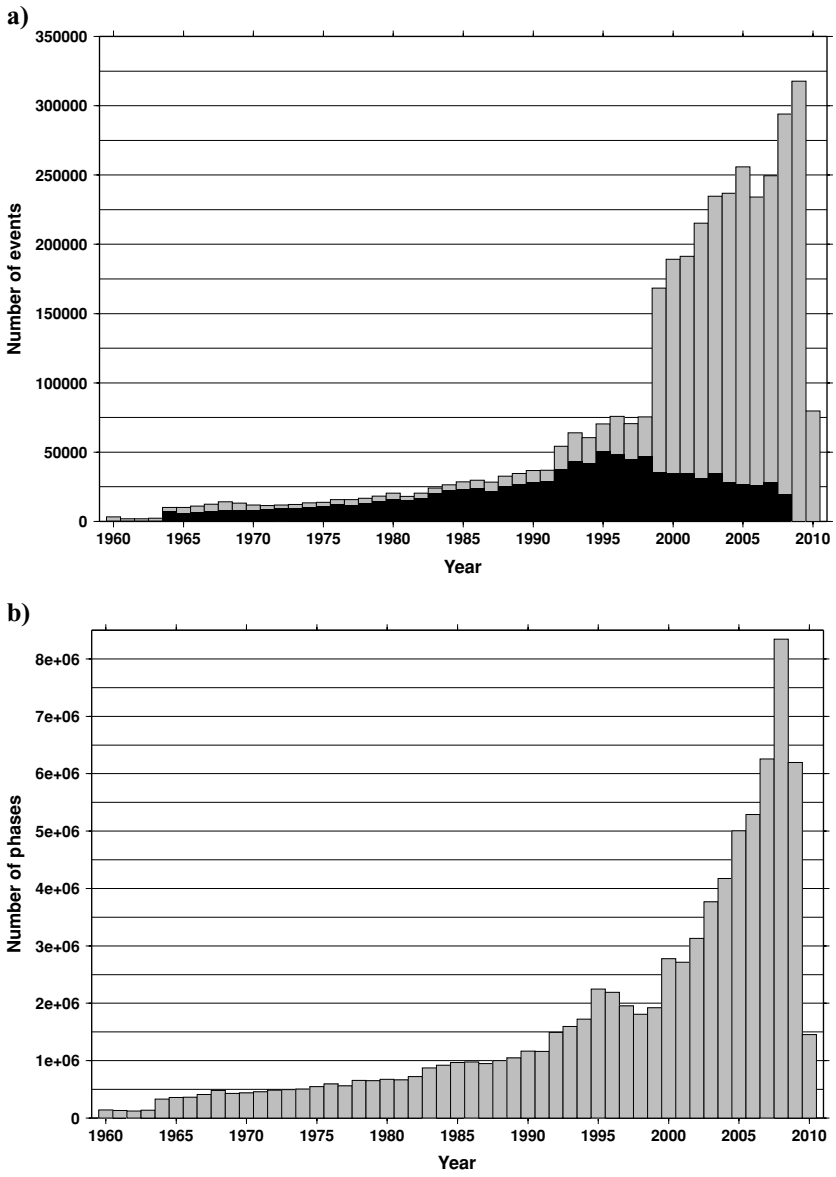

Figure 1. (a) Number of reported events per year (grey) and number of ISC solutions per year (black). (b) Number of associated phases per year. Both the number of reported events and associated phases increase exponentially with time. The apparent decrease in the number of the ISC hypocentre solutions reflect changes in the ISC procedures made to avoid relocation of small events where ISC analysis does not substantially improve locations reported by a local network operator.

an ever-growing data volume while maintaining the accuracy of the ISC Bulletin.

To illustrate the challenges posed by unbalanced, dense networks we consider the effect of the USArray on teleseismic event locations in the Chile-Bolivia border region. Fig. 2(a) shows the map of the region of interest as well as the locations of USArray stations. Because the USArray stations are all located in a very narrow azimuthal band with respect to the region of interest, many of the ray paths will be similar and thus produce correlated traveltime prediction errors due to unmodeled 3-D velocity structure in the Earth. We relocated 20 events that occurred there in 2007 and 2008 and were large enough to be recorded teleseismically. Fig. 2(b) shows the locations with and without using USArray stations. We assume that the local network locations are implicitly more accurate, although they cannot be considered ground truth. Hence, it is fair to say that including the phase picks from USArray stations makes the locations worse. When correlated systematic errors are ignored, they result in underestimated error ellipses (true locations do not lie within the ellipses) and introduce location bias. Thus, one ignores the correlated model error structure at his/her own peril. Indeed, as Fig. 2(c) indicates, when correlated errors are taken into account using the new ISC locator that we describe in this paper, the location bias is significantly reduced.

Ever since the ISC came into existence in 1964, it has been committed to providing a homogeneous bulletin that benefits scientific research. Hence the location algorithm used by the ISC, except for some minor modifications, remained largely unchanged for the past 40 yr (Bolt 1960; Adams *et al*. 1982). Nevertheless, minor changes were introduced especially in the past 10 years. From 2001 the ISC started using S, Sg, Sn and Sb arrivals in location (Storchak *et al*. 2000). Following the recommendations of several workshops organized by IASPEI (Schweitzer & Storchak 2006), the ISC started

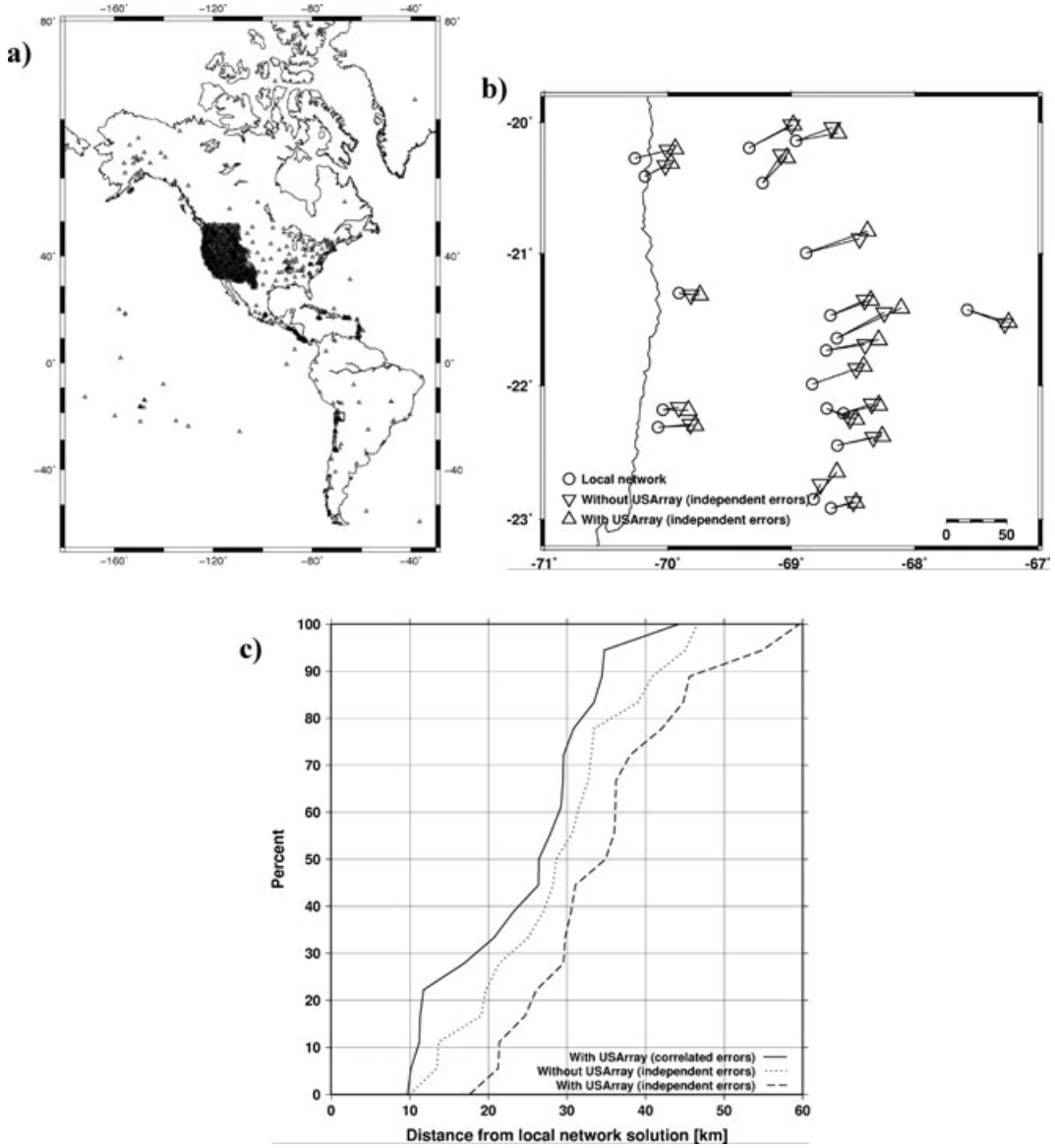

Figure 2. (a) The USArray stations dominate the network geometry when locating moderately sized events in the Chile-Bolivia border region. (b) Zoom-in on the target area. Location bias relative to local network solutions (circles) increases by about 10 km when USArray stations are included in the location (triangles). Inverted triangles indicate the locations without using USArray. (c) Cumulative distribution of mislocations relative to local network solutions. When assuming independent errors the locations get worse when USArray stations are used (dashed line). When correlated errors are accounted for (solid line) the location bias is significantly reduced.

using the *ak135* velocity model (Kennett *et al*. 1995), calculating *ak135* traveltime predictions using the tau-spline algorithm (Buland & Chapman 1983), instead of the Jeffreys–Bullen (Jeffreys & Bullen 1940) traveltime tables. This helped to mitigate the issue of baseline difference between *P*, *S* and *PKP* traveltimes in the Jeffreys-Bullen traveltime tables (Dziewonski & Anderson 1981; Engdahl *et al*. 1998). At the same time, for phase identification purposes the ISC started using the IASPEI Standard Seismic Phase List (Storchak *et al*. 2003) that is consistent with *ak135* velocity model. However, for the sake of comparison with the Jeffreys–Bullen tables, prior to changes proposed in this paper, the ISC has not introduced later phases (other than first-arriving *P* and *S* phases up to 100° epicentral distances) into the location algorithm.

To meet the challenge imposed by the ever-increasing data volume from heavily unbalanced networks we are introducing new ISC location procedures to ensure the efficient handling of data and to further improve the location accuracy of events reviewed by the ISC. The core of the new location procedures is the new ISC location algorithm. In this paper we describe the major components of the new ISC location algorithm and present the validation tests we performed by relocating a set of globally distributed ground truth events (Bondár & McLaughlin 2009b), and the entire ISC Bulletin. Note that discussing the ISC analyst review procedures is beyond the scope of this paper; here we focus on the new ISC location algorithm.

THE NEW ISC LOCATION ALGORITHM

Current ISC location procedures

The core of the current ISC locator is Jeffreys' uniform reduction method (Jeffreys 1932) that models the traveltime residual

distribution as a mixture of two Gaussian distributions, where the second Gaussian is described by a significantly larger variance that accounts for the observed heavy tails due to outliers (bad picks, phase identification errors, etc.). The method is an iteratively reweighted linearized least squares algorithm (Anderson 1982) that estimates hypocentre parameters by down-weighting outliers on the fly. The weighting function is recalculated at each iteration step and depends on the residuals themselves. Buland (1986) pointed out that the non-linearity of the weighting function introduces further non-linearity to an already non-linear problem.

Only Pg, Pb, Pn, Sg, Sb and Sn (up to 8°) and *P* and *S* (up to 100°, where *S* is downweighted by a factor of two beyond 20°) observations are considered in the location algorithm. Other phases (PKP branches, depth phases, etc.) are identified and their residuals are calculated, but they are not used in the location. Formal uncertainties are scaled to the 95 per cent confidence level. No provisions are made for model errors (systematic traveltime prediction errors due to unmodelled velocity heterogeneities in the Earth), and the observations are considered independent and normally distributed.

The reported hypocentres are ranked by their quality based on the increasing order of azimuthal gap provided by the stations contributing to the solution. During the location process the locator first attempts to find a free-depth solution by cycling through the reported hypocentres, that is, by taking each of the reported hypocentres, ordered by their rank, as an initial guess. If a convergent solution is found, the process moves on to calculate body and surface wave magnitudes. If no convergence is reached, the process is repeated, but now with the depth fixed to that of the trial hypocentre. If no solution is found, the process is repeated again, but now the depth is fixed to the default depth (10 or 35 km) assigned to each Flinn–Engdahl (Young *et al*. 1996) region number.

The ISC calculates body and surface wave magnitudes provided that there are a sufficient number of reported amplitude and period readings for an event. For body waves, station m_b is calculated by the Gutenberg & Richter (1956) formula from amplitudes reported for *P* waves of period $\leq$3 s in the distance range of 21–100°. Station MS is calculated by the Prague formula (Vanek *et al*. 1962) from surface wave amplitudes measured in the 10–60 s period range for events with focal depth $\leq$60 km in epicentral distances of 20–160°. Network magnitudes are calculated as the mean of corresponding station magnitudes. Note that even a single station magnitude could be used to produce a network magnitude and that no error estimates are computed.

During the validation tests described below we used the current ISC location algorithm as a baseline to measure the performance of the new location algorithm.

New ISC location procedures

While the current ISC location procedures have served the scientific community well in the past 50 yr, they can certainly be improved. The uniform reduction method works best when only first-arriving *P* (and possibly *S*) phases are used. Because it treats all phases as equal, incorporating later arriving phases would require further adjustments to ensure that later phases (observed less reliably) are down-weighted. Enforcing a relative weighting structure for the observations would not only introduce further complications in the algorithm but would also hamper the efficiency of the iteratively reweighted algorithm.

Linearized location algorithms are very sensitive to the initial starting point for the location. The current procedures make the assumption that a good initial hypocentre is available among the reported hypocentres. However, there is no guarantee that any of the reported hypocentres are close to the global minimum in the search space. Furthermore, attempting to find a free depth solution is futile when the data have no resolving power for depth (e.g. when the first arrival is not within the inflection point of the *P* traveltime curve). When there is no depth resolution, the algorithm would simply pick a point on the origin time–depth trade-off curve. The current ISC locator assumes that the observational errors are independent. As we pointed out before, accounting for correlated traveltime predictions errors is unavoidable if we wanted to improve (or simply maintain) location accuracy as station networks become progressively denser. Finally, publishing network magnitudes that may be derived from a single station measurement is rather prone to producing erroneous event magnitude estimates.

To address these issues we have developed a location algorithm that

(1) uses all *ak135* (Kennett *et al*. 1995) predicted phases (including depth phases) in the location with elevation, ellipticity (Dziewonski & Gilbert 1976; Kennett & Gudmundsson 1996; Engdahl *et al.* 1998) and depth-phase bounce point corrections (Engdahl *et al*. 1998);

(2) obtains the initial hypocentre guess via the neighbourhood algorithm (NA; Sambridge 1999; Sambridge & Kennett 2001);

(3) performs iterative linearized inversion using an *a priori* estimate of the full data covariance matrix to account for correlated model errors (Bondár & McLaughlin 2009a);

(4) attempts a free-depth solution if and only if there is depth resolution, otherwise it fixes the depth to a region-dependent default depth;

(5) scales uncertainties to 90 per cent confidence level and calculates location quality metrics for various distance ranges;

(6) obtains a depth-phase depth estimate based on reported surface reflections via depth-phase stacking (Murphy & Barker 2006), and

(7) provides robust network magnitude estimates with uncertainties.

Seismic phases

One of the major advantages of using the *ak135* traveltime predictions (Kennett *et al*. 1995) is that they do not suffer from the baseline difference between *P*, *S* and *PKP* phases compared to the Jeffreys–Bullen tables (Jeffreys & Bullen 1940). Furthermore, *ak135* offers an abundance of phases from the IASPEI Standard Seismic List that can be used in the location, most notably the PKP branches and depth-sensitive phases. In the course of this paper we refer to the phases that are used in the location as time-defining. Because we typically deal with many time-defining phases, it is not necessary to incorporate slowness and azimuth measurements in the location procedures. Phases are initially identified with respect to the median of the reported hypocentre parameters.

Elevation and ellipticity corrections (Dziewonski & Gilbert 1976; Kennett & Gudmundsson 1996; Engdahl *et al.* 1998), using the WG84 ellipsoid parameters, are added to the *ak135* predictions. For depth phases, bounce point (elevation correction at the surface reflection point) and water depth (for pwP) corrections are calculated by the algorithm of Engdahl *et al*. (1998). We use the ETOPO1 global relief model (Amante & Eakins 2009) to obtain the elevation or the water depth at the bounce point.

Phase picking errors are described by *a priori* measurement error estimates derived from the inspection of the distribution of ground

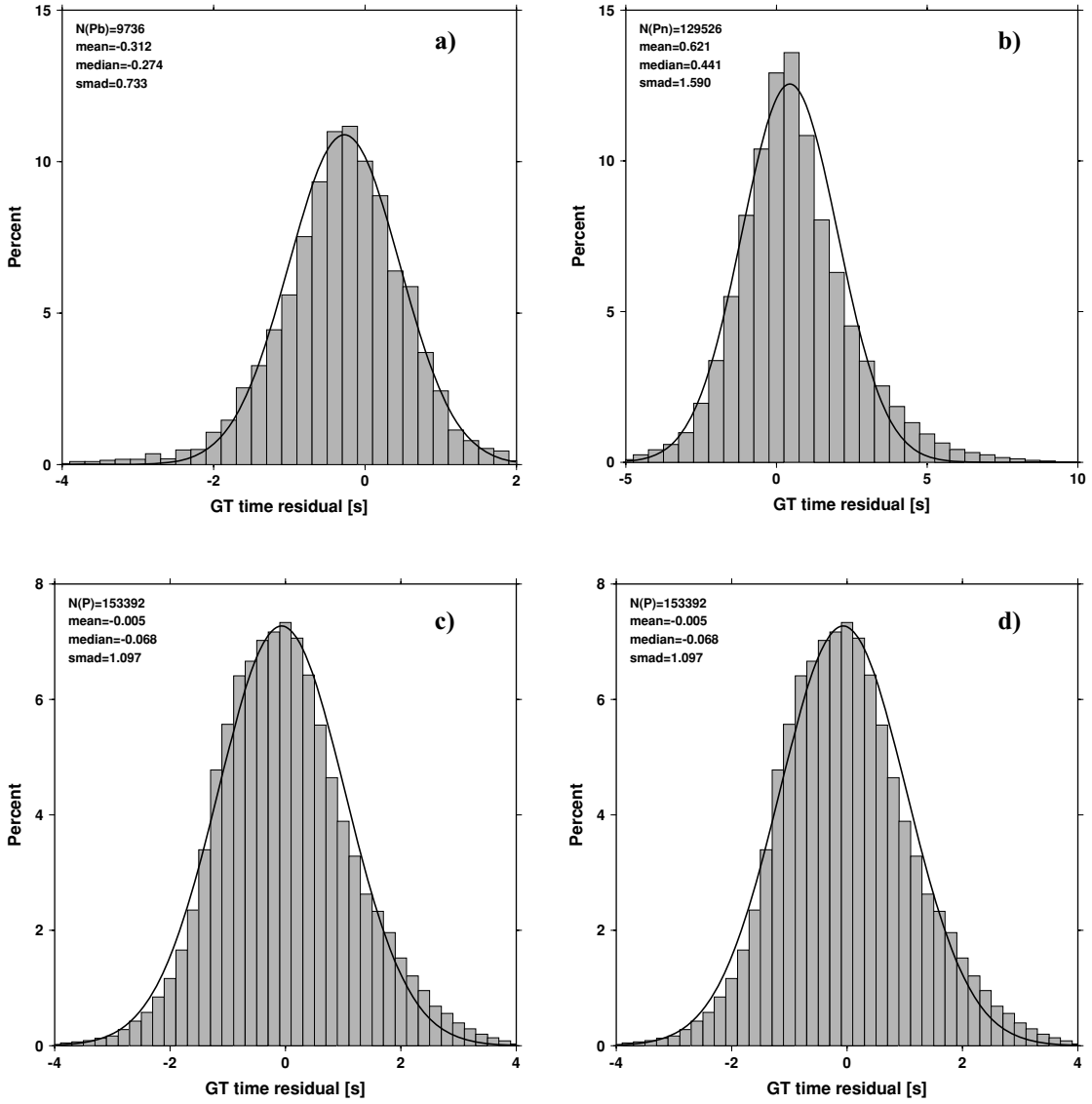

Figure 3. Distribution of ground truth residuals (residuals calculated w.r.t. the ground truth hypocentre in the IASPEI Reference Event List) for (a) Pb, (b) Pn, (c) P and (d) PKPdf phases. The solid line represents the best-fitting Gaussian distribution.

truth residuals (residuals calculated with respect to the ground truth location) from the IASPEI Reference Event List (Bondár & McLaughlin 2009b). Fig. 3 shows the histograms of ground truth residuals for Pb, Pn, P and PKPdf.

For phases that do not have sufficient number of observations in the ground truth database we established *a priori* measurement errors so that the consistency of the relative weighting schema is maintained. First-arriving *P*-type phases (P, Pn, Pb, Pg) are picked more accurately than later phases, so their measurement error estimates are the smallest, 0.8 s. The measurement error for first-arriving *S*-phases (S, Sn, Sb, Sg) is set to 1.5 s. Phases traversing through or reflecting from the inner/outer core of the Earth have somewhat larger (1.3 s for PKP, PKS, PKKP, PKKS and P'P' branches as well as PKiKP, PcP and PcS and 1.8 s for SKP, SKS, SKKP, SKKS and S'S' branches as well as SKiKP, ScP and ScS) measurement error estimates to account for possible identification errors among the various branches. Free-surface reflections and conversions (PnPn, PbPb, PgPg, PS, PnS, PgS and SnSn, SbSb, SgSg, SP, SPn and SPg) are observed less frequently and with larger uncertainty, and therefore suffer from large, 2.5 s, measurement errors. Similarly, a measurement error of 2.5 s is assigned to the longer period and typically emergent diffracted phases (Pdif, Sdif and PKPdif). The *a priori* measurement error for the commonly observed depth phases (pP, sP, pS, sS and pwP) is set to 1.3 s, while for the rest of depth phases (pPKP, sPKP, pSKS, sSKS branches and pPb, sPb, sSb, pPn, sPn, sSn) the measurement error estimate is set to 1.8 s. We set the measurement error estimate to 2.5 s for the less reliable depth phases (pPg, sPg, sSg, pPdif, pSdif, sPdif and sSdif). Note that we also allow for distance-dependent measurement errors. For instance, to account for possible phase identification errors at far-regional distances the *a priori* measurement error for Pn and P is increased from 0.8 to 1.2 s and for Sn and S from 1.5 to 1.8 s between 15 and 28°. The measurement errors between 40 and 180° are set to 1.3 and 1.8 s for the prominent PP and SS arrivals, respectively, but they are increased to 1.8 and 2.5 s between 25 and 40°.

The relative weighting scheme described above ensures that arrivals picked less reliably or prone to phase identification errors are down-weighted in the location algorithm. Optimally the

measurements errors should depend on the observed signal-to-noise ratio, but that, except for the IMS stations, is hardly ever reported. Since the ISC works with reported parametric data with wildly varying quality, we opted for a rather conservative set of *a priori* measurement error estimates.

Correlated model error structure

Most location algorithms, either linearized or non-linear, assume that all observational errors are independent. This assumption is violated when the separation between stations is less than the scale length of local velocity heterogeneities. When correlated travel-time prediction errors are present, the data covariance matrix is no longer diagonal, and the redundancy in the observations reduces the effective number of degrees of freedom. Thus, ignoring the correlated error structure inevitably results in underestimated location uncertainty estimates. For events located by an unbalanced seismic network this may also lead to a biased location estimate.

Chang *et al*. (1983) demonstrated that accounting for correlated error structure in a linearized location algorithm is relatively straightforward once an estimate of the non-diagonal data covariance matrix is available. To determine the data covariance matrix we follow the approach described by Bondár & McLaughlin (2009a). They assume that the similarity between ray paths is well approximated by the station separation. This simplifying assumption allows for the estimation of covariances between station pairs from a generic *P* variogram model (Fig. 4) derived from ground truth residuals. Because the overwhelming number of phases in the ISC bulletin is teleseismic *P*, we expect that the generic variogram model will perform reasonably well anywhere on the globe.

The estimates for the elements of the data covariance matrix are obtained as

$$C_D(i, j) = \sigma^2_{\text{sill}} - \gamma(h_{ij}) + \delta_{ij}\sigma^2_{\text{phase}}, \tag{1}$$

where σ^2_{sill} denotes the background variance where the variogram levels off (i.e. where the pairs become independent), $\gamma(h_{ij})$ is the variogram value for the distance, h_{ij}, between the ith and jth stations, δ_{ij} is the Kronecker delta and σ^2_{phase} is the *a priori* estimate of the measurement error covariance for an observed phase, respectively. The last term indicates that the measurement error variances add to the diagonal of the covariance matrix. We set the sill variance to $\sigma^2_{\text{sill}} = 1.2$, which is quite larger than Fig. 3 would suggest. Recall that we use a variogram model derived from teleseismic *P* observations. We found that it was necessary to increase the sill to account for unexplained model errors at local and regional distances in order to obtain adequate error ellipse coverage.

Note that the confidence error ellipse measures the *precision* of the epicentre estimate, but does not indicate the *accuracy* (bias) of the location. One could obtain a very precise estimate of the location parameters that could be still biased due to 3-D velocity structures unexplained by the velocity model. The accuracy of the location estimate can only be measured if ground truth (GT) information is available. We define the location coverage as

$$\kappa^2 = x^2 \big/ \left(s^2_{\text{maj}ax} + GTx^2\right) + y^2 \big/ \left(s^2_{\text{min}\,ax} + GTx^2\right), \tag{2}$$

where x and y are the Cartesian coordinates of the GTx epicentre (x stands for the GT accuracy, 0–5 km) in the coordinate system defined by the semi-axes of the error ellipse scaled to the required confidence level and centred on the seismic location. Since the GT epicentre also has an error term (GTx), when calculating the location coverage we inflate the confidence error ellipse to account for this uncertainty. We insert the coordinates of the GT epicentre into the standard equation of the error ellipse (obtained from the location algorithm and enlarged by the GT uncertainty) to calculate the location coverage parameter. If the coverage parameter is less than unity, the GT epicentre is inside the error ellipse; if it is larger than unity, the error ellipse does not cover the true location.

Because in this representation the covariances depend only on station separations, the covariance matrix (and its inverse) needs to be calculated only once. We assume that different phases owing to the different ray paths they travel along, as well as station pairs with a separation larger than 1000 km are uncorrelated. Hence, the data covariance matrix is a sparse, block-diagonal matrix. Furthermore, if the stations in each phase block are ordered by their nearest neighbour distance, the phase blocks themselves become block-diagonal. To reduce the computational time of inverting large matrices we exploit the inherent block-diagonal structure by inverting the covariance matrix block-by-block.

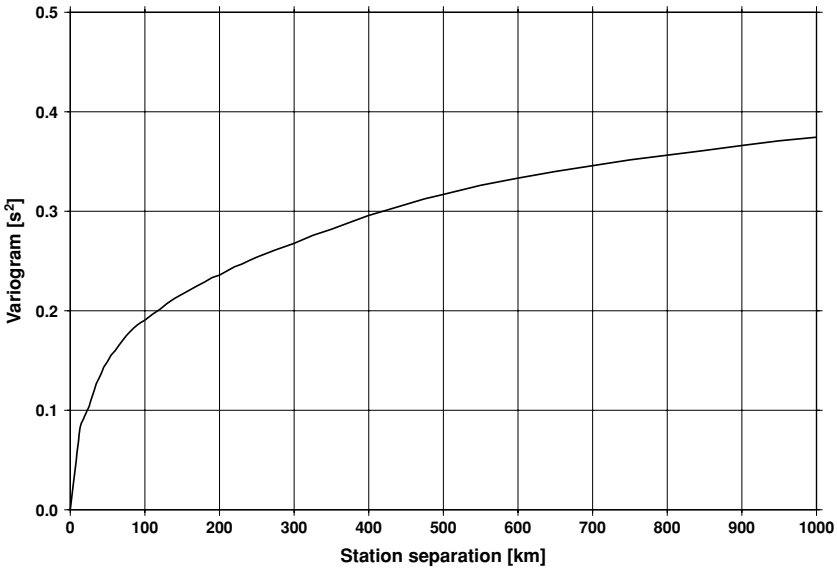

Figure 4. Generic variogram model derived from ground truth residuals for teleseismic *P* phases. The isotropic, stationary (c) variogram model is used to produce the *a priori* estimate of the data covariance matrix.

The singular value decomposition of the data covariance matrix is written as

$$\mathbf{C}_D = \mathbf{U}_D \mathbf{\Lambda}_D \mathbf{V}_D^T, \tag{3}$$

where $\mathbf{\Lambda}_D$ is the diagonal matrix of eigenvalues and the columns of $\mathbf{U}_D$ contain the eigenvectors of $\mathbf{C}_D$. Let $\mathbf{C}_D = \mathbf{B}\mathbf{B}^T$, with $\mathbf{B} = \mathbf{U}_D \mathbf{\Lambda}_D^{1/2}$, then the projection matrix

$$\mathbf{W} = \mathbf{B}^{-1} = \mathbf{\Lambda}_D^{-1/2} \mathbf{U}_D^T \tag{4}$$

will orthogonalize the data set and project redundant observations into the null space.

Depth resolution

In principle, depth can be resolved if there is a mixture of upgoing and downgoing waves emanated from the source, that is, if there are stations covering the distance range where the vertical partial derivative of the traveltime of the first-arriving phase changes sign (local networks), or if there are phases with vertical slowness of opposite sign (depth phases). Core reflections, such as PcP, and to a lesser extent, secondary phases (S in particular) could also help in resolving the depth.

We developed a number of criteria to test whether the reported data for an event have sufficient depth resolution:

(1) Local network: one or more stations within 0.2° with time-defining phases.

(2) Depth phases: five or more time-defining depth phases reported by at least two agencies (to reduce a chance of misinterpretation by a single inexperienced analyst).

(3) Core reflections: five or more time-defining core reflections (PcP, ScS) reported by at least two agencies.

(4) Local/near regional *S*: five or more time-defining *S* and *P* pairs within 5°.

We attempt a free-depth solution if any of the above criteria are satisfied; otherwise we fix the depth to a default depth dependent on the epicentre location. Note that a combination between local/regional phases and PKP observations may also help in resolving the depth, but we have not yet implemented this criterion. The depths of known anthropogenic events and landslides are fixed to the surface.

Depth-phase stack

While we use depth phases directly in the location, the depth-phase stacking method (Murphy & Barker 2006) provides an independent means to obtain robust depth estimates. Fig. 5 illustrates the method. The moveout curves (parametric on epicentral distance) are calculated for each depth-phase—first-arriving *P* pairs and plotted as a function of depth. The interval determined by the measurement error estimate of the depth-phase around the observed moveout time is projected to depth, producing a trace with a boxcar function at the depth corresponding to the observed moveout. The stack of the traces represents the distribution of depths consistent with all depth-phase observations. The depth-phase constrained depth is determined as the median of the stack and its uncertainty is characterised by the standard median absolute deviation (SMAD, the L1-norm equivalent of standard deviation). It should be noted that the depth uncertainty may be underestimated if it is correlated with the epicentre error, that is, when the hypocentre error ellipsoid has an appreciable component of plunge.

Because the depth obtained from the depth-phase stacking method implicitly depends on the epicentre itself, we perform the depth-phase stack only twice: first, with respect to the initial location in order to obtain a reasonable starting point for the depth in the grid search described in the following section; second, with respect to the final location to obtain the final estimate for the depth-phase constrained depth.

Initial hypocentre

For poorly recorded events the reported hypocentres may exhibit a large scatter and they could suffer from large location errors, especially if they are only recorded teleseismically. To obtain a good initial hypocentre guess for the linearized location algorithm we employ the NA (Sambridge 1999; Sambridge & Kennett 2001). NA is a non-linear grid search method capable of exploring a large search space and rapidly closing on the global optimum. Kennett (2006) discusses in detail the NA algorithm and its use for locating earthquakes.

We perform a search around the median of reported hypocentre parameters with a generously defined search region—within a 2° radius circle around the median epicentre, 10 s around the median origin time and 150 km around the median reported depth. These default search parameters were obtained by trial-and-error runs to achieve a compromise between execution time and allowance for gross errors in the median reported hypocentre parameters. Note that if our test for depth resolution fails, we fix the depth to the region-dependent default depth. The initial hypocentre estimate will be the one with the smallest L1-norm misfit among the NA trial hypocentres. Once close to the global optimum, we proceed with the linearized location algorithm to obtain the final solution and corresponding formal uncertainties.

Iterative linearized location algorithm

We adopt the location algorithm described in detail in Bondár & McLaughlin (2009a). We solve the matrix equation

$$\mathbf{G}_\mathrm{W}\mathbf{m} = \mathbf{W}\mathbf{G}\mathbf{m} = \mathbf{W}\mathbf{d} = \mathbf{d}_\mathrm{W}, \tag{5}$$

where $\mathbf{G}$ is the $(N \times M)$ design matrix containing the partial derivatives of N data by M model parameters, $\mathbf{m}$ is the $(M \times 1)$ model adjustment vector $[\Delta T, \Delta x, \Delta y, \Delta z]^T$, $\mathbf{d}$ is the $(N \times 1)$ vector of time residuals and $\mathbf{W}$ is the $(N \times N)$ projection matrix of eq. (4). Hence, we solve the inversion problem in the eigen coordinate system in which the transformed observations are independent, that is, $\mathbf{d}_\mathrm{W}$ represents linear combinations of the observed residuals, the 'eigen residuals'. Eq. (5) is solved by singular value decomposition, which yields the general inverse

$$\mathbf{G}_\mathrm{W}^{-1} = \boldsymbol{V}_\mathrm{W} \mathbf{\Lambda}_\mathrm{W}^{-1} \boldsymbol{U}_\mathrm{W}^T \tag{6}$$

and the model adjustment of

$$\mathbf{m}_\mathrm{est} = \mathbf{G}_\mathrm{W}^{-1} \mathbf{d}_\mathrm{W}. \tag{7}$$

After the *j*th iteration, the model vector is adjusted such that $\mathbf{m}_{j+1} = \mathbf{m}_j + \mathbf{m}_\mathrm{est}$. Once a convergent solution is obtained, the location uncertainty is defined by the *a posteriori* model covariance matrix

$$\mathbf{C}_\mathrm{M} = \mathbf{G}^{-1} \mathbf{C}_\mathrm{D} \mathbf{G}^{-1^T} = \boldsymbol{V}_\mathrm{W} \mathbf{\Lambda}_\mathrm{W}^{-2} \mathbf{V}_\mathrm{W}^T. \tag{8}$$

The model covariance matrix yields the 4-D error ellipsoid whose projections provide the 2-D error ellipse and 1-D errors for depth

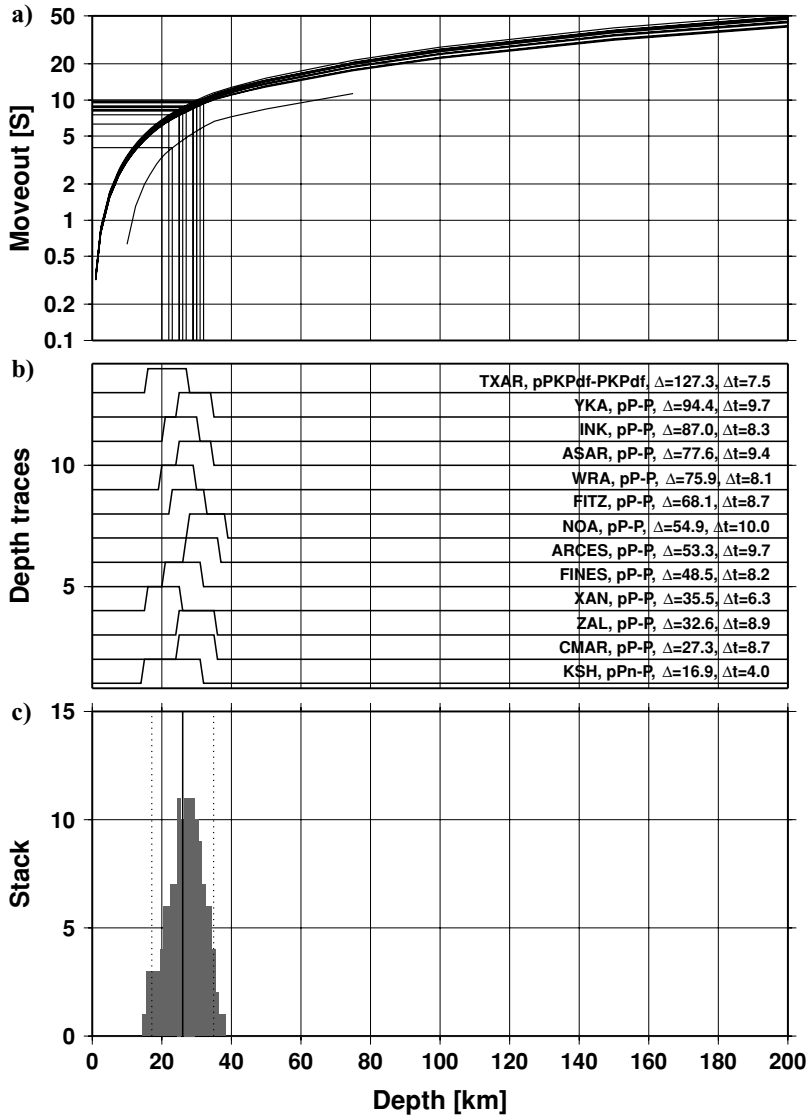

Figure 5. Illustration of the depth-phase stacking method of Murphy and Barker (2006). (a) The predicted moveout curves (depth phase – first arriving P arrival time) are generated for each station as a function of depth (they also depend on the epicentral distance). (b) For each observed moveout a depth trace is generated by putting a boxcar at the corresponding depth. The width of the boxcar is defined by the *a priori* measurement error for the depth phase; it is centred on the observed moveout and then projected to the *x*-axis (depth). (c) The depth traces are stacked and the depth-phase depth and its uncertainty are defined as the median and the SMAD of the stack, respectively.

and origin time. The error ellipse encompassing the confidence region at a given α percentile level is defined by

$$(\mathbf{r} - \mathbf{r}_{\rm loc})^T \mathbf{C}_M (\mathbf{r} - \mathbf{r}_{\rm loc}) = \kappa_\alpha^2, \tag{9}$$

where $\mathbf{r}_{\rm loc}$ denotes the location vector of the epicentre. We follow Jordan & Sverdrup (1981) to define κ_α^2 as

$$\kappa_\alpha^2 = M s^2 F_\alpha(M, K + N - M) \tag{10}$$

where the variance scaling factor s^2 is defined as

$$s^2 = \frac{K + \frac{1}{N}\sum d_{\rm W}^2}{K + N - M} \tag{11}$$

and F_α is an F statistic with M and $K + N - M$ degrees of freedom at the critical level $\alpha = 90$ per cent with $M = 2$ and with N independent observations, that is, the total number of observations less the number of observations projected to the null space. K is set to a large value (99 999) so that the formal uncertainty estimates approximate 'coverage' error ellipses (Evernden 1969), which assumes that the *a priori* error estimates are exactly known and therefore diminishes the role of *a posteriori* residuals. This is also the practice the Reviewed Event Bulletin produced by the International Data Center follows. Note that since we projected the system of equations into the eigen coordinate system, the number of independent observations is less than the total number of observations. Hence, the estimated location error ellipses necessarily become larger, providing a more realistic representation of the location uncertainties. The major advantage of this approach is that the projection matrix is calculated only once for each event location.

VALIDATION TESTS

To demonstrate improvements due to the new location procedures, we relocated some 7200 GT0–5 events in the IASPEI Reference Event List (Bondár & McLaughlin 2009b) both with the current

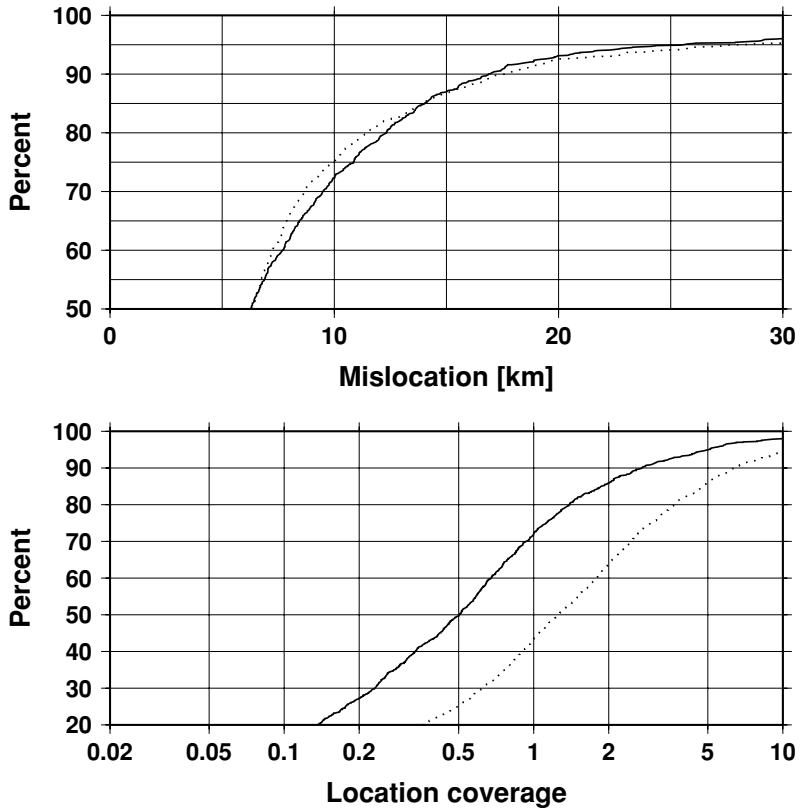

Figure 6. Cumulative distributions of mislocation and location coverage for explosions. Dotted lines represent the results from the current ISC locator; solid lines denote the results from the new locator. The location coverage is calculated according to eq. (2). The error ellipse covers the true location if the coverage parameter is less than or equal to unity. The actual coverage of the 95 per cent confidence error ellipse with the current locator is only about 45 per cent. When correlated errors are accounted for, the actual coverage of the 90 per cent confidence ellipse of the new locator covers more than 70 per cent of the true locations.

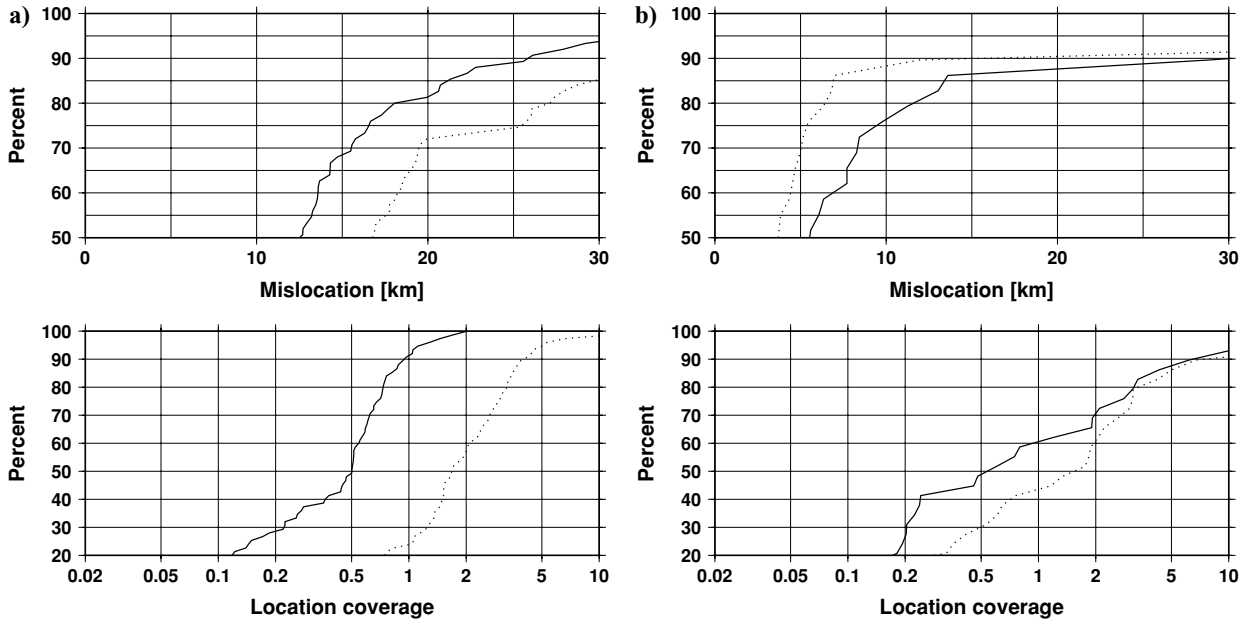

Figure 7. (a) Cumulative distributions of mislocation and location coverage for nuclear explosions conducted in French Polynesia. When correlated errors are accounted for (solid lines), we achieve significant improvements in both location and coverage. (b) Cumulative distributions of mislocation and location coverage for nuclear explosions conducted in the Novaya Zemlya test site. Accounting for correlated errors removes the conspiring effect of teleseismic stations, but reveals the inadequacy of regional *ak135* predictions for this region, which leads to a deterioration in location.

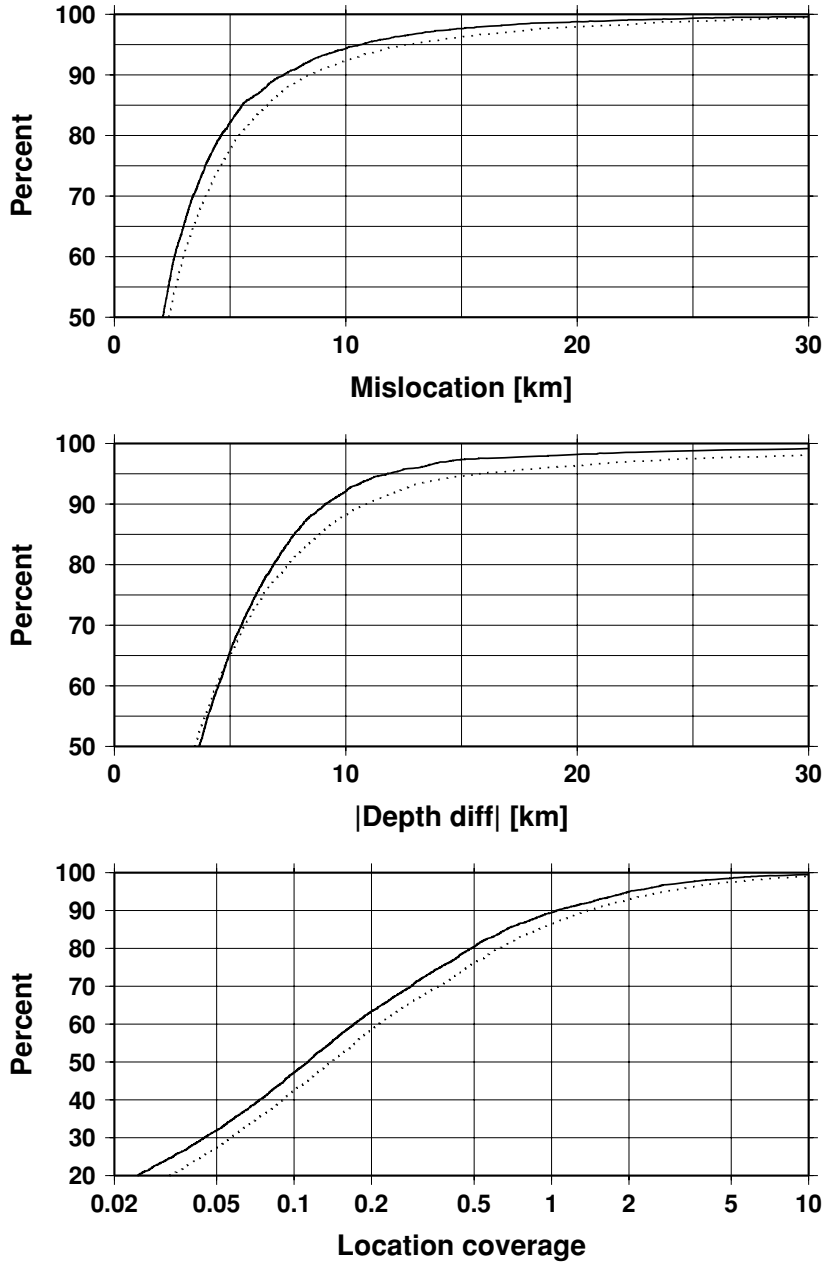

Figure 8. Cumulative distributions of mislocation, depth difference from GT and location coverage for free-depth solutions. The new locator (solid lines) provides systematic improvements in hypocentre and location coverage.

ISC locator (which constitutes the baseline) and with the new location algorithm. Since both locators use *ak135* predictions, we expect only minor location improvements. However, because we account for correlated model error structure, we expect significant improvements in location coverage (whether the confidence error ellipse contains the true epicentre). We also expect improvements in depth determination due to the test for depth resolution and the direct use of depth phases in the locator. In the figures that follow dotted lines represents the baseline (current ISC locator) and solid lines denote the results from the new locator.

Explosions

The explosions in the GT0–5 data set consist of both nuclear and chemical explosions. Some of the chemical explosions (refraction profile shots) are recorded by only a few stations with a large azimuthal gap, thus prone to large location errors. We included these chemical explosions in the validation tests because they represent extreme conditions (sparse, unbalanced networks) for the locator. Fig. 6 shows the cumulative distributions of mislocation and location coverage with the baseline and the new location algorithm for the explosions (both chemical and nuclear). In general, there are no significant differences in location accuracy, as in both cases 90 per cent of the locations are within 17 km of the true location. However, the location coverage is significantly improved when correlated errors are accounted for. Note that the location coverage parameter described by eq. (2) is less than one if the error ellipse, adjusted to the GT accuracy, covers the true location, and larger than unity when the true location falls outside the error ellipse. Hence, the actual coverage is given as the percentile level where the coverage

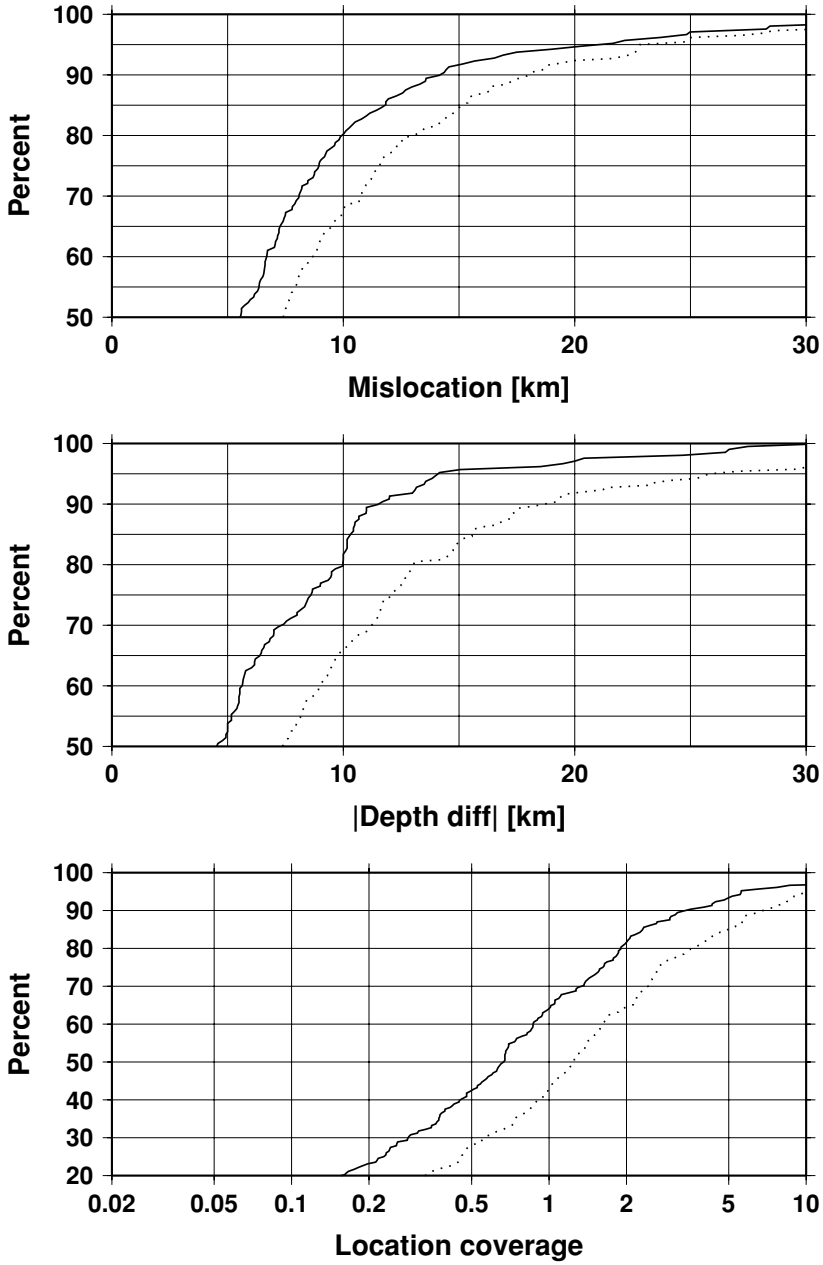

Figure 9. Cumulative distributions of mislocation, depth difference from GT and location coverage for earthquakes with depth-phase depth solutions. The new locator (solid lines) provides significant improvements in hypocentre and location coverage.

parameter is one. The baseline locator's 95 per cent confidence error ellipse covers the true epicentre only about 45 per cent of the time. When correlated errors are taken into account, the 90 per cent confidence error ellipse obtained from the new locator covers the true epicentre about 70 per cent of the time.

Fig. 7 shows the results for the French Polynesian and Novaya Zemlya test sites, which represent extremes in the spectrum. The nuclear explosions in the French Polynesian atolls of Tuamotou and Fangatafua were recorded by a heavily unbalanced teleseismic network, dominated by dense networks in Japan, Europe and California. Hence, the French Polynesian test sites represent a textbook case for demonstrating the effect of correlated errors. Indeed, as Fig. 7(a) shows, when correlated errors are accounted for, we reduce the location bias and achieve 90 per cent actual coverage, while the baseline algorithm provides an abysmal 25 per cent actual coverage. On the other hand, the nuclear explosions conducted at the Novaya Zemlya test site (Fig. 7b) were recorded by a mixture of teleseismic and regional stations, and our locations deteriorated compared to the baseline algorithm. This can be explained by the fact that once the effect of conspiring stations is taken out, the influence of regional data is increased and the poor *ak135* regional traveltime predictions for this region (e.g. Krementskaya *et al*. 2001; Hicks *et al*. 2004; Levshin *et al*. 2007) make the solutions worse. Nevertheless, we still get improvements in the location coverage.

Earthquakes: free-depth and depth-phase depth solutions

For some three-quarters of earthquakes in the IASPEI Reference Event List we had sufficient depth resolution to obtain free-depth solutions. Fig. 8 shows the cumulative distributions of

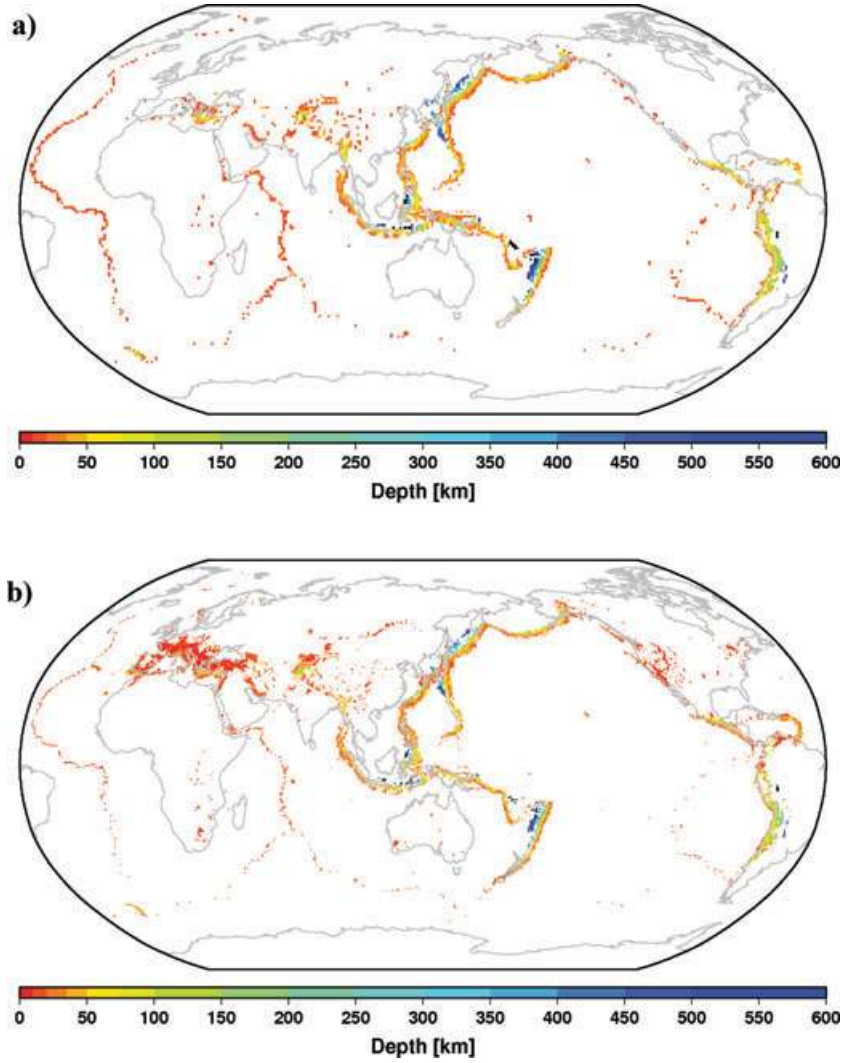

Figure 10. (a) Default depths on a 1 × 1 degree grid derived from EHB free-depth solutions. (b) Updated default depth grid obtained from EHB free-depth solutions and EHB events flagged as reliable depth, as well as free-depth solutions from the entire ISC bulletin relocated with the new locator. The significantly larger number of events allowed us to refine the grid to 0.5° × 0.5° resolution and to fill in seismically active regions with dominantly smaller magnitude events such as in North America and the Mediterranean as well as mapping out the southern boundary of the Australian plate. Note that the size of the coloured dots reflects the cell size in the grids.

mislocation and depth difference from the ground truth hypocentre as well as location coverage with the baseline and the new location algorithm. For free-depth solutions we obtain systematic improvements both in epicentre and depth with the new locator. In particular, we achieve 90 per cent actual location coverage. Furthermore, 95 per cent of the events are within 10 km of the true epicentre, and 92 per cent of the events deviate less than 10 km from the true depth. For the earthquakes where we were able to obtain depth-phase depth solutions (Fig. 9), the improvements are even more pronounced.

Earthquakes: fixed-depth solutions

As we mentioned earlier, we only attempt a free-depth solution if there is depth resolution. If the data have no resolving power for the depth, we fix the depth to a region-dependent default depth. Instead of fixing the depth to the conventional depth values (35 km or 10 km), we followed the approach of Bolton *et al*. (2006) and opted for generating default depth values on a global grid using the EHB bulletin (Engdahl *et al*. 1998), whose depths are superior to the ISC depth estimates. We considered the EHB free-depth solutions only, and generated a 1° × 1° grid of default depths. The default depth is defined as the median of all depths in a grid cell, provided that there were at least five events in the cell, and the 75–25 per cent quartile range was less than 100 km. The latter constraint is imposed to avoid regions with both shallow and deep seismicity. For locations where the default depth grid does not provide a default depth value, we fix the depth to the midpoint of the crust according to CRUST2.0 (Bassin *et al*. 2000).

Fig. 10(a) shows the default depth grid generated from EHB free-depth solutions. The EHB default depth grid represents well the depth distributions of major seismic regions. However, owing to the EHB event selection criteria the EHB bulletin is complete only above magnitude 5, and the default depth grid misses important regions, such as western North America and the Mediterranean region where the seismicity is dominated by relatively small events.

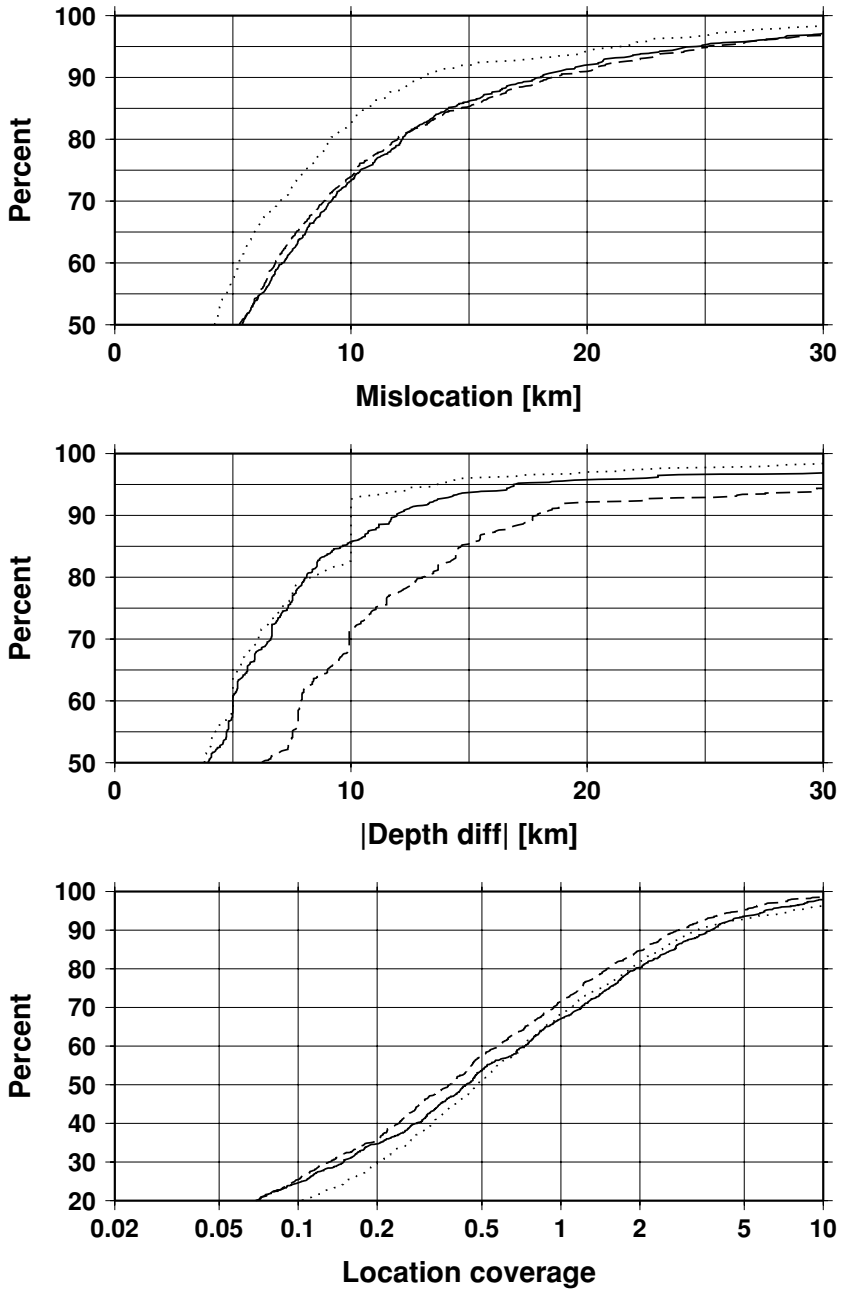

Figure 11. Cumulative distributions of mislocation, depth difference from GT and location coverage for earthquakes with fixed depth solutions. Dotted lines represent the solutions from the baseline algorithm, dashed lines are the solutions obtained with the EHB default depth grid and solid lines denote the solutions using the updated default depth grid. The updated default depth grid provides reasonable default depths where there is seismicity.

Initially, we used the default depth grid described above to get region-dependent default depth values, but we were not satisfied with the results for fixed-depth events. Thus, encouraged by the consistent improvements for free-depth solutions obtained by the new locator, we set out to create a new default depth grid which would provide default depths where seismicity is well established. We relocated the entire ISC bulletin with the new locator and used the free-depth solutions together with the EHB free-depth solutions, including the fixed-depth EHB earthquakes that were flagged as having reliable depth estimate (Bob Engdahl, personal communication, 2010), to produce a new default depth grid. The significantly increased number of events (to 815 000 from 90 000) allowed us to refine the grid to $0.5^\circ \times 0.5^\circ$ resolution and fill in the gaps not covered by the EHB. Fig. 10(b) shows the updated default depth grid. For locations where the default depth grid does not provide a default depth value, instead of using CRUST2.0, we now fix the depth to the median of reported depths to obtain a more robust default estimate for the depth. We provide a description of how the depth was obtained for each event in the ISC Bulletin.

Fig. 11 shows the cumulative distributions of mislocation, depth difference from GT and the location coverage for fixed-depth earthquakes obtained from the baseline algorithm (dotted lines), and those obtained from the new locator using the EHB default depth grid (dashed lines) and the updated default depth grid (solid lines). Recall that the current ISC locator fixes the depth to that of a trial hypocentre reported by an agency, and for the well-recorded GT

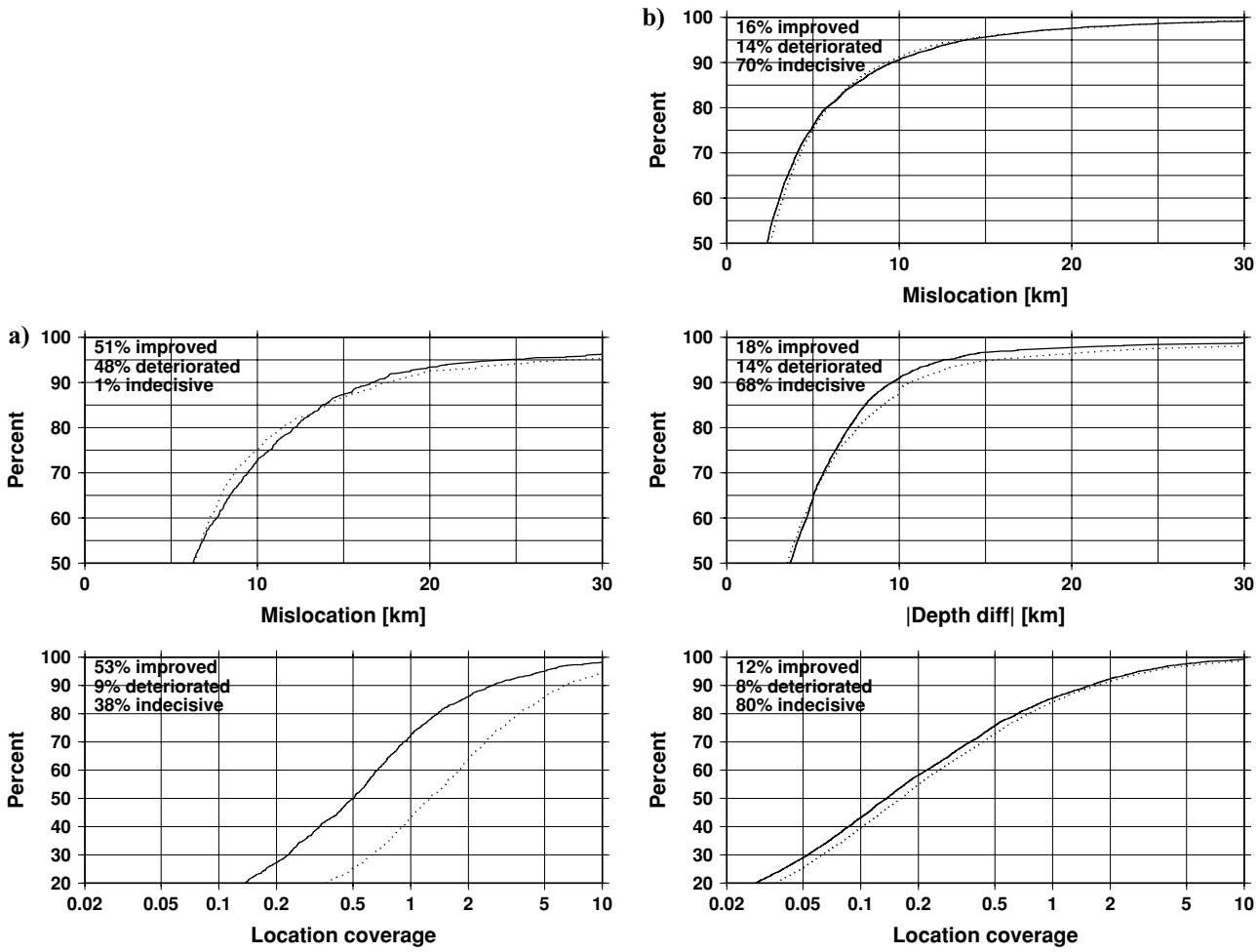

Figure 12. Cumulative distributions of mislocation, depth difference from GT and location coverage for (a) 1180 explosions and (b) 6003 earthquakes located by both the baseline (dotted lines) and the new locator (solid lines). More events are improved than deteriorated.

events there is a good chance that at least one of the reporting agencies gets the depth right. While we do not get location improvements for fixed-depth events, the updated default depth grid provides an adequate default depth for seismically active regions.

Pairwise comparisons

In the previous sections we discussed the results where both the old and new locator produced free-depth or fixed-depth solutions. Another way to look at the results is to plot the cumulative distributions for the events that were commonly located by both locators. This allows us to quantify the improvements. Fig. 12 shows the cumulative distributions of mislocation, depth difference and location coverage for the 1180 explosions and 6003 earthquakes from the IASPEI Reference Event List that were located both by the baseline and the new location algorithm. Since the majority of the GT0–5 events in the IASPEI Reference Event List are well-recorded, most of the location differences are indecisive, that is, the mislocation from the GT is within the GT accuracy. Nevertheless, more events are improved than deteriorated. For earthquakes we consistently improve the depth, and for explosions we get significant improvements in coverage.

Fig. 13 shows further details of the earthquake population. Fig. 13(a) shows the cumulatives for the 5243 earthquakes for which the baseline algorithm obtained a free-depth solution, and Fig. 13(b) shows the cumulative distributions for the 760 earthquakes for which the baseline algorithm fixed the depth. Note that the new locator, due to the depth resolution test, may or may not obtain a free-depth solution for the free-depth events by the old locator, or vice versa. For both cases, that is when the old locator obtained a free-depth solution or fixed the depth, Fig. 13 indicates that we improve the hypocentre parameters and the location coverage with the new locator.

The effect of correlated errors

As we stated before, accounting for correlated errors produces more reliable uncertainty estimates, and for heavily unbalanced networks, reduces bias. Hence, we expect overall improvements in location coverage, and for a minority of events, improvements in hypocentre parameters. In other words, we expect most improvements in the tails of the distributions.

Fig. 14 shows the median and 90 per cent percentile mislocation, origin time and depth difference from GT, as well as the error ellipse area, the location coverage and error ellipse eccentricity as a function of number of stations for the 7183 commonly located events by the old and new locators. At the median level (thin dotted and solid lines) the hypocentre parameters are barely distinguishable. Although sparse networks can also suffer from correlated model errors, it is more likely that dense networks have a sufficient number

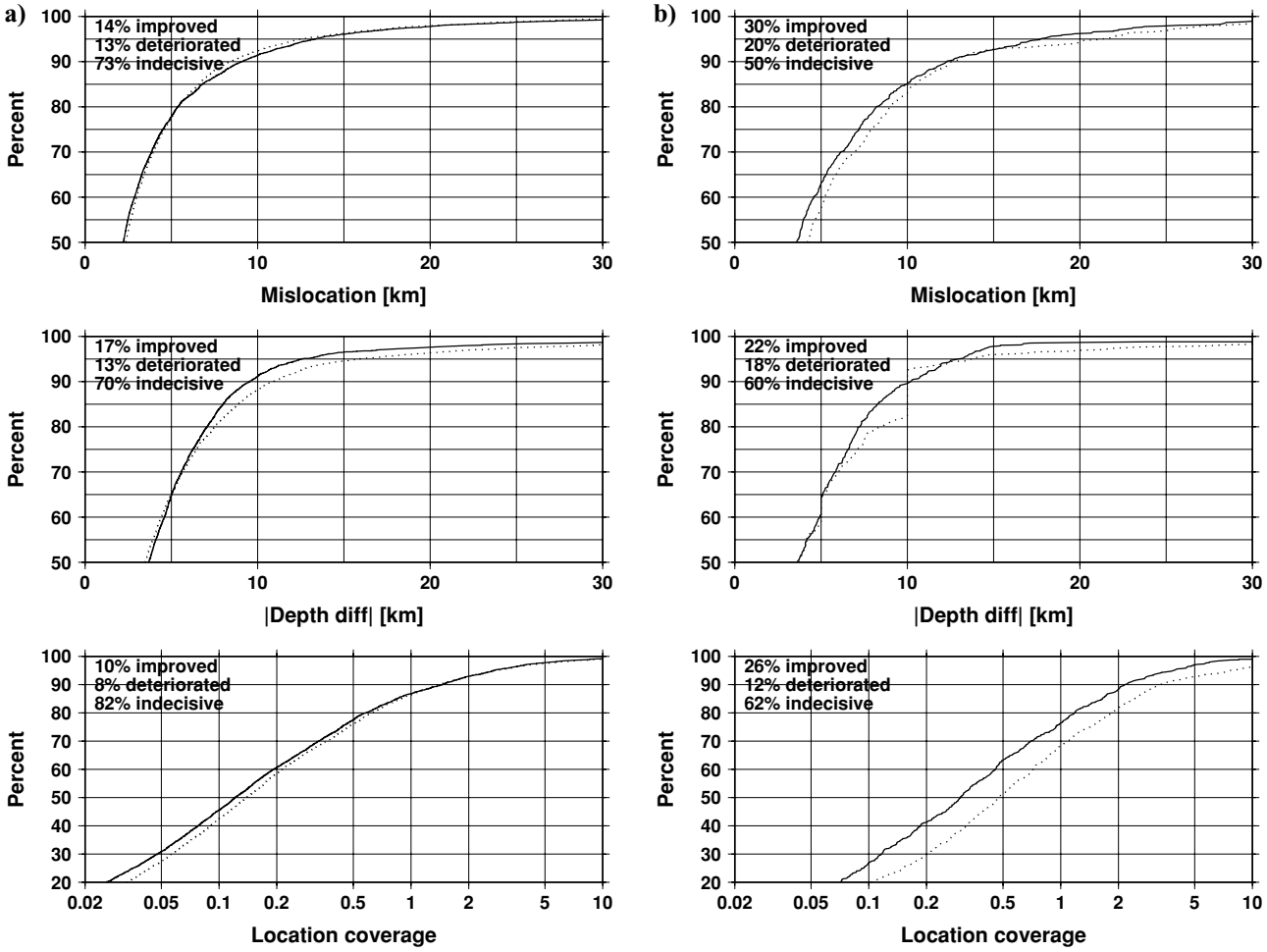

Figure 13. Cumulative distributions of mislocation, depth difference from GT and location coverage for earthquakes with (a) 5243 free-depth solutions and (b) 760 fixed-depth solutions by the baseline algorithm (dotted lines). The new locator (solid lines) improves more events.

of correlated ray paths to introduce location bias if the correlated error structure is ignored. Indeed, at the 90 per cent level (thick dotted and solid lines) we achieve consistent improvements in epicentre, depth and origin time once the number of stations exceeds 100 stations.

One of the major advantages of incorporating correlated error structure in the location algorithm through the estimate of the full data covariance matrix is that it decouples the 90 per cent error ellipse from the number of correlated observations. Indeed, the error ellipse area levels off both at the median and the 90 per cent percentile levels beyond 100 stations when correlated errors are incorporated in the location algorithm, thus maintaining coverage. Had we achieved 90 per cent coverage, the 90th percentile curve would run along unity. Since we have only about 85 per cent coverage, the curve runs just above unity. On the other hand, when correlated errors are ignored the area of error ellipse shrinks indefinitely with the number of stations, thus losing coverage. Recall that the *a posteriori* model covariance matrix depends on the number of independent observations. Because the number of equivalent uncorrelated data (linear combinations of residuals) is reduced, the formal location uncertainties described by the *a posteriori* model covariance matrix become larger, resulting in enlarged and more circular error ellipses as indicated by the eccentricity of the error ellipses (the smaller the eccentricity the more circular the error ellipse).

Comparison with the EHB

Fig. 15 shows the cumulative distributions of mislocation, origin time and depth difference from GT for the baseline algorithm (dotted lines) and the new locator (solid lines) for those EHB events (dashed lines) that are in the GT data set. The plots indicate that the solutions obtained with the new locator match or surpass the EHB location accuracy but not the EHB depth accuracy.

Station and network magnitudes

Body and surface wave station magnitudes are calculated using the same formulae as with the old procedure. As soon as the majority of reporting agencies adopted the IASPEI recommendations on measuring amplitude and period, the ISC will be able to compute the set of magnitudes according to the IASPEI standards (Bormann *et al*. 2009).

For each reported amplitude, we calculate the magnitude using the Gutenberg & Richter (1956) formula for m_b from amplitudes reported for *P* waves of period $\leq$3 s in the distance range of 21–100°. For MS we use the Prague formula (Vanek *et al*. 1962) from surface wave amplitudes measured in the 10–60 s period range for events with focal depth $\leq$60 km in epicentral distances of 20–160°. We set the station magnitude to the magnitude with the maximum A/T value among the amplitude and period pairs reported to that station.

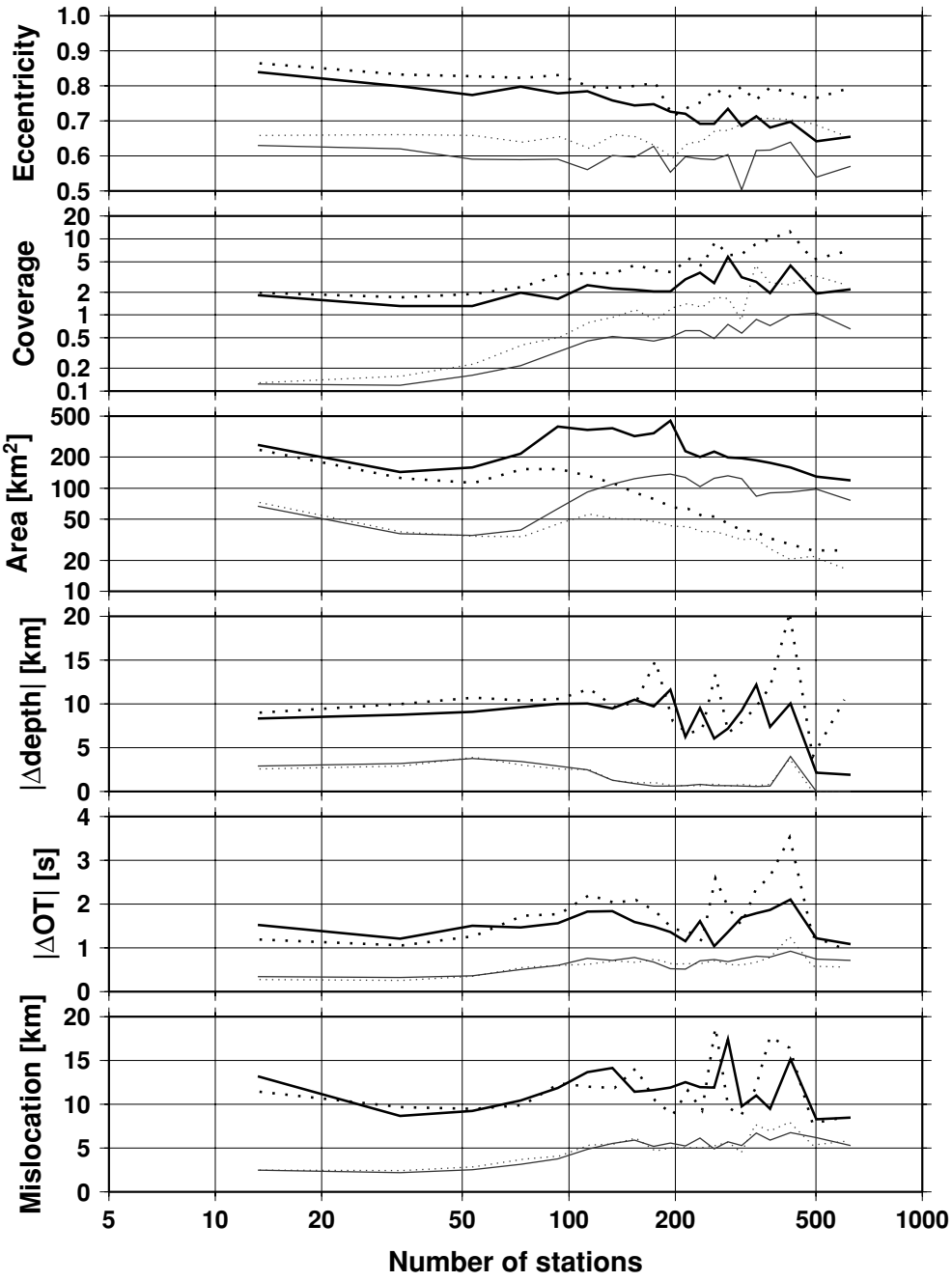

Figure 14. Median (thin grey) and 90 per cent percentile (thick black) mislocation, origin time and depth difference, error ellipse area, location coverage and error ellipse eccentricity as a function of number of stations with the old (dotted lines) and new (solid lines) locator. When correlated errors are accounted for, the size of error ellipse levels off, providing better coverage, and the hypocentre bias is reduced once the information contained in the network is exhausted. Because the (eigen)network is more balanced, the error ellipses become more circular.

In order to avoid blunders caused by outlier station magnitudes, we now require at least three valid station magnitudes and calculate the network magnitude as an alpha-trimmed median. The network magnitude is determined as the median of station magnitudes after discarding the lower and upper 20 percentiles of station magnitudes (the alpha-trimming procedure), and its uncertainty is given by the SMAD around the median. Since the new procedure requires at least three station magnitudes to produce a network magnitude, it mitigates the risk of introducing gross magnitude errors in the ISC Bulletin due to outlier amplitude observations.

AUTOMATIC RELOCATION OF THE ENTIRE ISC BULLETIN

As we mentioned before, we relocated the entire (1960–2010) ISC Bulletin, including four years of the International Seismological Summary (ISS, the predecessor of the ISC) catalogue (Villaseñor & Engdahl 2005, 2007), in order to obtain an updated default depth grid derived from the seismicity itself. When relocating the ISC Bulletin we again mimicked as much as possible (bulletin for most years already benefited from review by the ISC analysts) the automatic ISC location procedures, that is, we ignored the ISC, EHB

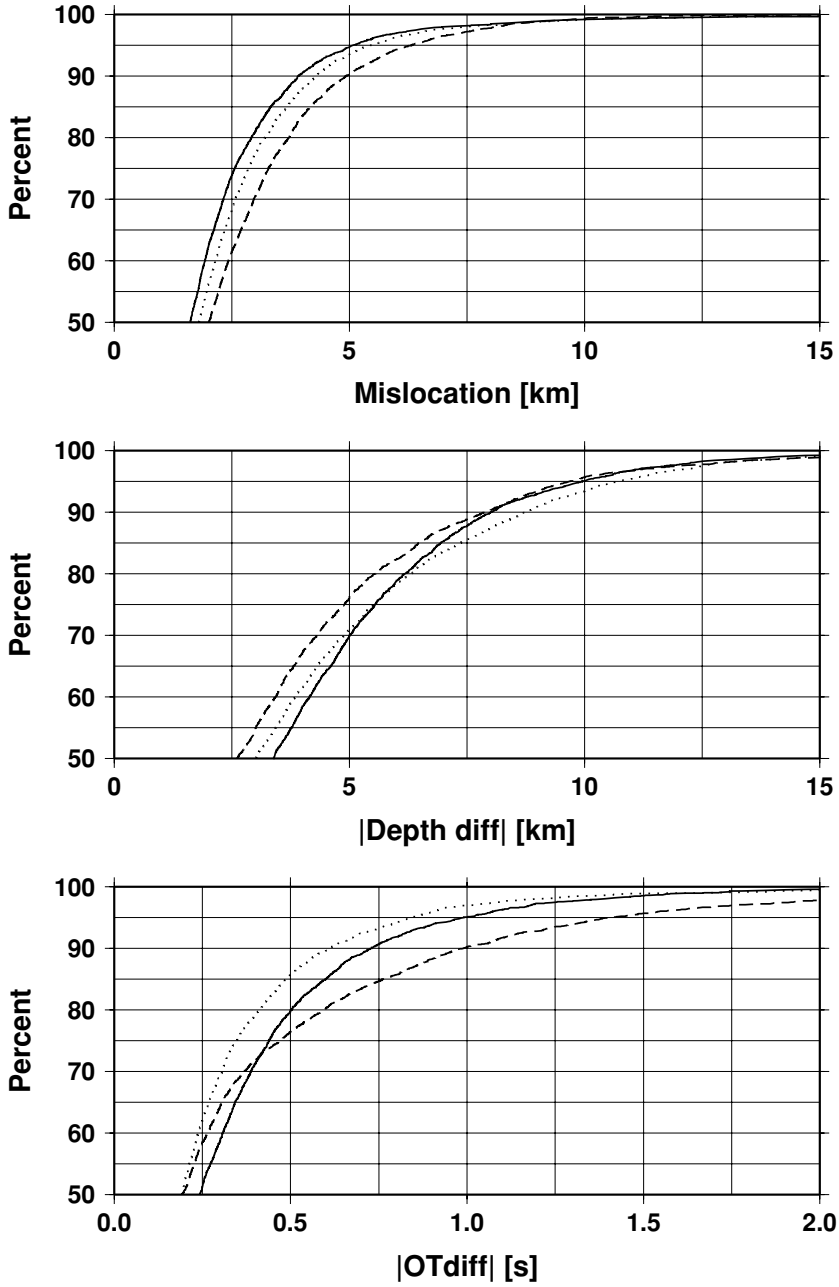

Figure 15. Cumulative distributions of mislocation, origin time and depth difference from GT for EHB earthquakes. Dotted lines represent the solutions from the baseline algorithm, dashed lines are the EHB solutions and solid lines denote the solutions with the new locator. The solutions obtained with the new locator match or surpass the EHB location accuracy but not the EHB depth accuracy.

and GT hypocentres and located all events in the bulletin, regardless of whether there was an ISC solution or not.

Fig. 16(a) shows the cumulative distribution of time needed to locate an event for the 2507 492 events relocated during this exercise. For 96 per cent of the events it took less than a second to obtain a location on a standard dual-core Linux box. This time includes the NA grid search, the inversion of the data covariance matrix, the iterative linearized location algorithm and magnitude calculations. Fig. 16(b) shows the median runtime needed to obtain a solution as a function of the number of associated phases. Not surprisingly, the execution time increases exponentially with the number of observations, and on average it takes more than a minute to locate an event when the number of observations exceeds 1500. This is an acceptable performance for an automatic location. For the ISC analyst review this performance can be further boosted because at that stage it is not necessary to perform the NA search, the slowest operation of all, as the editors have already been provided with a reasonable automatic location.

Finally, we present comparisons between the automatic locations obtained with the new locator and the reviewed ISC Bulletin. The reviewed ISC Bulletin currently (1960–2008 August) contains some 1 million earthquakes. The automatic location with the new locator provided some 2.3 million earthquake hypocentres (we relocated even those events that were not reviewed by the ISC). The number of earthquakes that are found both in the reviewed ISC Bulletin and in the relocated bulletin is 993 685. In

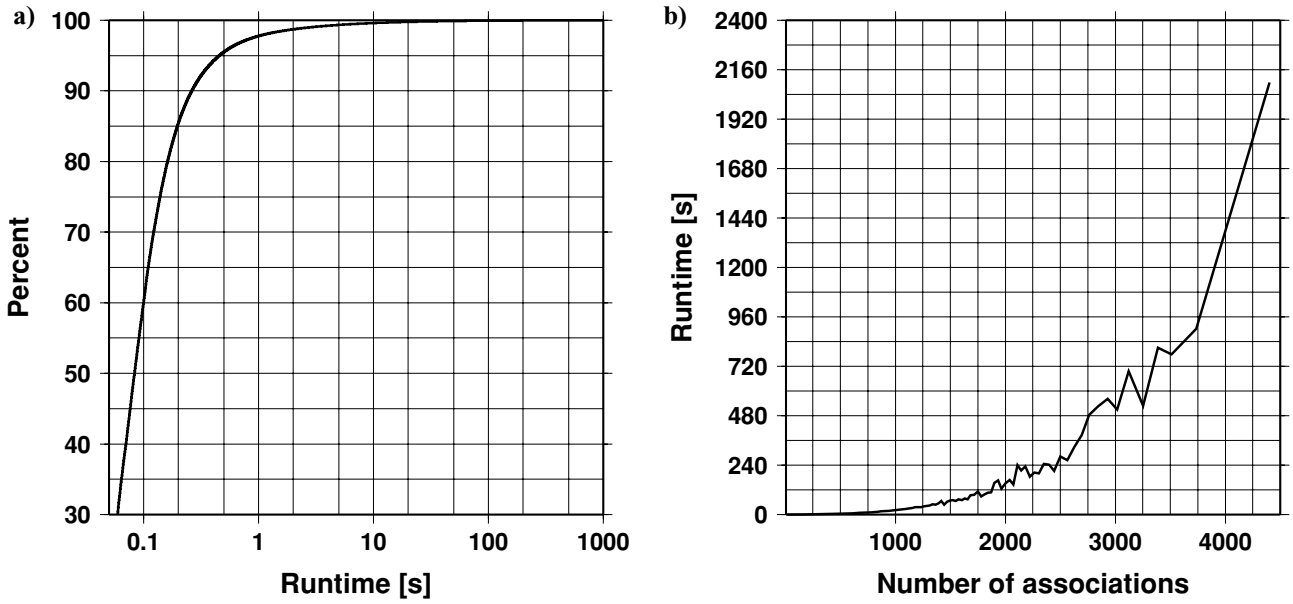

Figure 16. (a) Cumulative distribution of the time needed to locate an event. 96 per cent of the events are located within a second on a standard dual-core Linux box. (b) Median runtime of event location as a function of number of associated phases.

the followings we only consider this data set of commonly located earthquakes.

We expect that through the use of later phases and testing for depth resolution, the scatter in event locations will be reduced. The tightness of seismicity has been traditionally used as a qualitative measure of location improvements. To test our hypothesis we quantify the scatter in seismicity by calculating the entropy of epicentres. We adopt the definition of Nicholson *et al*. (2000) for the entropy of a point set. Their definition makes use of Voronoi diagrams (Shewchuk 1996). The area of the Voronoi cells is proportional to the density of events and thus can be used to estimate the probability density function of earthquake distribution in space. A Voronoi diagram (the geometric dual of Delaunay triangulation) is a unique partition of the domain where each cell contains one event and any point in the cell is closer to that event than to any other ones. Thus, Voronoi cells define a natural neighbourhood—the smaller the cells, the denser the events. We define the entropy of a point set as

$$S = \log(N) - \log(A) + \frac{1}{N}\sum_i \log(a_i), \tag{12}$$

where N is the number of events, a_i is the area of a Voronoi cell, A is the area of the convex hull (the minimal convex polygon that contains the point set). S is always non-positive and reaches its maximum value, zero, when the distribution of events is uniform. Hence, we accept our hypothesis (i.e. the new locator tightens the seismicity) if the entropy decreases relative to that obtained from the reviewed ISC Bulletin. To be correct, we should measure the volume of 3-D Voronoi cells derived from the point cloud of hypocentres, but doing that is far from trivial on a sphere. Therefore we resort to the approximation of measuring the entropy in two dimensions, on the surface of the Earth.

We calculated the entropy of epicentres in each seismic region for both the reviewed and relocated ISC Bulletin. Fig. 17(a) shows the entropy of epicentres by seismic region number obtained from the reviewed ISC bulletin and from the relocations with the new locator. Fig. 17(b) shows the entropy differences between the relocated and ISC bulletin earthquakes. For 44 of 50 seismic regions the new locator reduces the entropy of the data set, in some cases significantly, and overall by –0.5 on average. Hence, the new locator provides better clustering of the events, thus tightening the seismicity. Figs 18(a)–(g) further illustrate this point. We plotted the hypocentres of various seismically active regions for events from the reviewed ISC Bulletin (left-hand panels) and those obtained with the new locator (right-hand panels). The tightening of the seismicity with the new locator is visible even at the relatively large scale of the maps.

DISCUSSION AND CONCLUSIONS

We presented the new ISC location procedure that uses all reported phases with a valid *ak135* prediction, obtains the initial hypocentre guess via the NA algorithm, and takes into account the correlated traveltime prediction error structure in a linearized iterative least squares location algorithm.

We have shown that correlated model error structure can be accounted for if we solve the equations in the eigen system defined by the non-diagonal data covariance matrix. Following Bondár & McLaughlin (2009a), we construct the *a priori* estimate of the data covariance matrix using a generic variogram model for teleseismic *P*. The underlying assumptions in using a variogram are that the similarity between ray paths can be approximated by the distance between the two receivers, and an isotropic variogram adequately characterises the correlation structure. Owing to these simplifying assumptions, the data covariance matrix and its inverse needs to be calculated only once.

However, the isotropic assumption ignores the 3-D velocity structure and becomes violated for shorter ray paths where there is less averaging over various scale lengths of velocity perturbations. Admittedly, a generic variogram model derived from teleseismic phases cannot adequately capture all the details of the 3-D velocity variations at local and regional distances. Thus, for smaller events that recorded only by local or near-regional networks, our approach may under- or over-estimate the correlation between traveltime predictions along similar ray paths. One could attempt to derive variogram models for regional and local distances, but those would have to be region-specific and would only provide a poor substitute for a 3-D

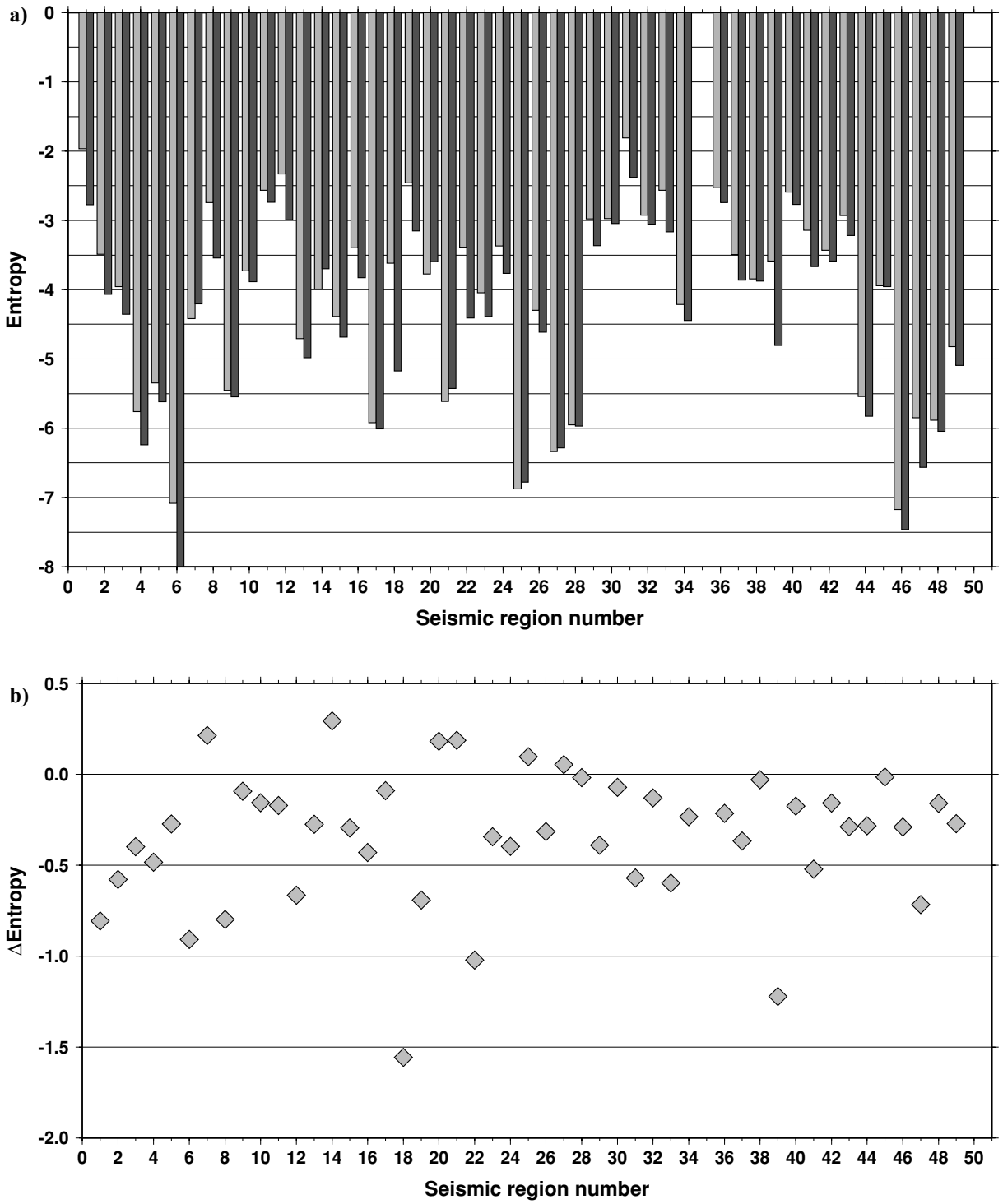

Figure 17. (a) Entropy of epicentres in each seismic region obtained from the reviewed ISC bulletin (grey) and from the relocations with the new locator (black). (b) Difference between entropies. For most regions the new locator reduces the entropy of the data set, thus tightening the seismicity.

velocity model. Furthermore, recall that the *a priori* measurement error variances add to the diagonal of the data covariance matrix derived from the variogram model. Large measurement errors will dominate the data covariance matrix and tend to dilute the effect of correlated structure of traveltime predictions errors, that is, they decrease the correlation strength between the off-diagonal elements. Increasing the sill value would counterweight this effect. Hence, the nearly threefold increase in the sill value we introduced allows for unexplained model errors at local and regional distances, and more importantly, enforces that the correlated error structure is taken into account despite the relatively large diagonal elements. Obviously, a global 3-D velocity model would be the true remedy for this problem. Indeed, perfect traveltime predictions would carry no model errors, and predictions along very similar ray paths would remain uncorrelated. Nevertheless, we have demonstrated that even using a 1-D model and a variogram model that fits teleseismic observations

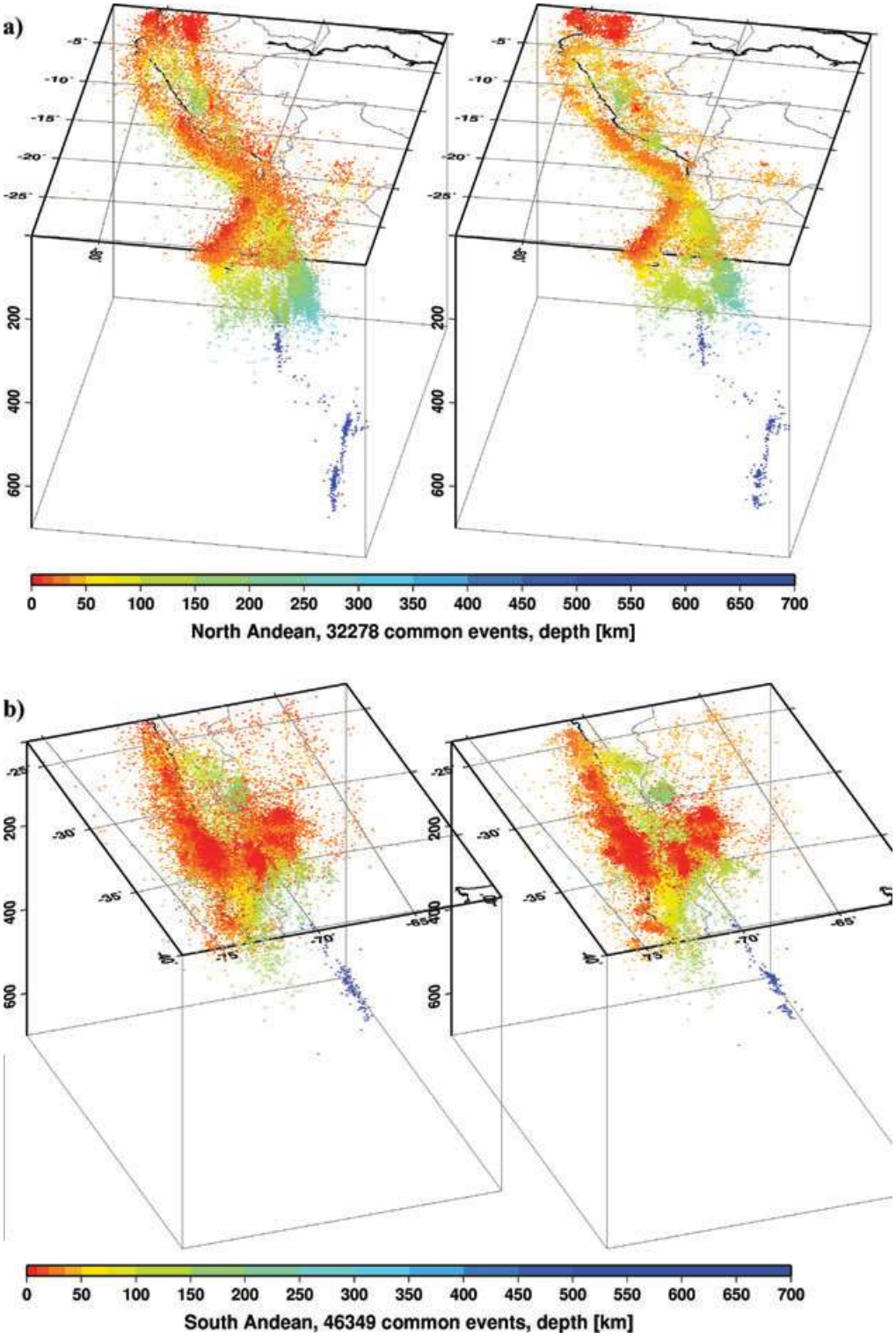

Figure 18. Comparison of seismicity maps for common events in the reviewed ISC Bulletin (old locator, left) and the relocated ISC bulletin (new locator, right) for various regions. (a) North Andean, (b) South Andean, (c) Central America, (d) Alaska, (e) Guam, Honshu, Ryukyu islands, (f) Hindu Kush – Pamir and (g) Southern Antilles. The events are better clustered when located with the new locator.

we could achieve realistic uncertainty estimates, as the 90 per cent confidence error ellipses cover the true locations 80–85 per cent of the time.

We noted above that the location improvements we demonstrated for the ground truth events are consistent, but minor. This is not surprising as most of the events in the IASPEI Reference Event List are very well-recorded with a small azimuthal gap and dominated by P-type phases. In these circumstances we could expect significant location improvements only for heavily unbalanced networks where large numbers of correlated ray paths conspire to

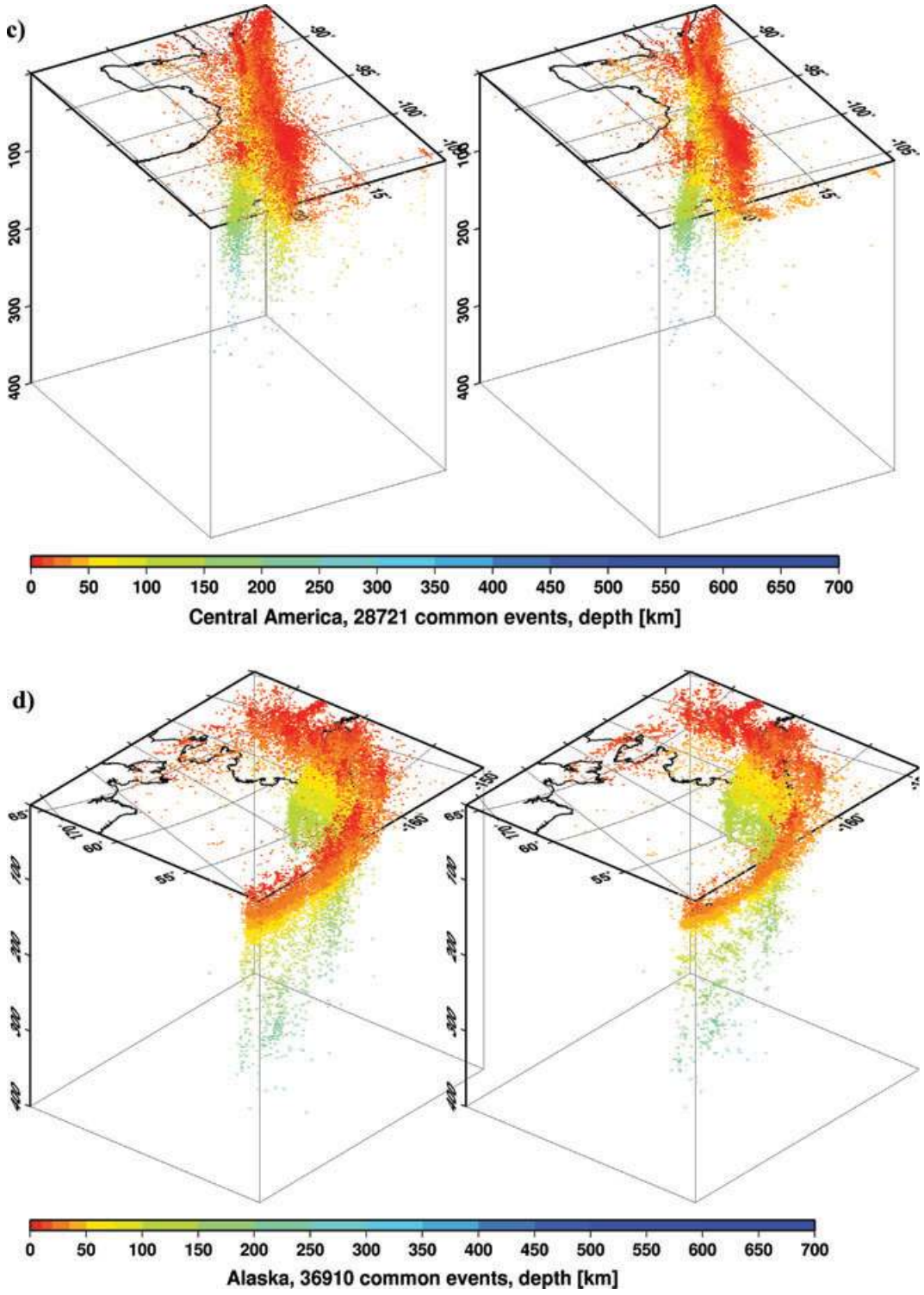

Figure 18. (*Continued*.)

introduce location bias. On the other hand, the ISC Bulletin represents a plethora of stations configurations ranging from reasonable to the most unfavourable network geometries. Hence, we could expect more dramatic location improvements when relocating the entire ISC Bulletin. Although in this case we cannot measure the improvement in location accuracy due to the lack of ground truth information, the better clustering of event locations suggests that the new locator indeed outperforms the baseline algorithm.

The improvements are attributed to a combination of several features in the new location algorithm. The NA grid search offers a better initial hypocentre for the linearized inversion to start with; accounting for the correlated traveltime prediction errors provides reliable formal uncertainty estimates, and for unbalanced networks reduces location bias; using depth-sensitive phases in the location algorithm results in considerably better depth estimates; for events with no depth-resolution the default depth grid gives a reasonable

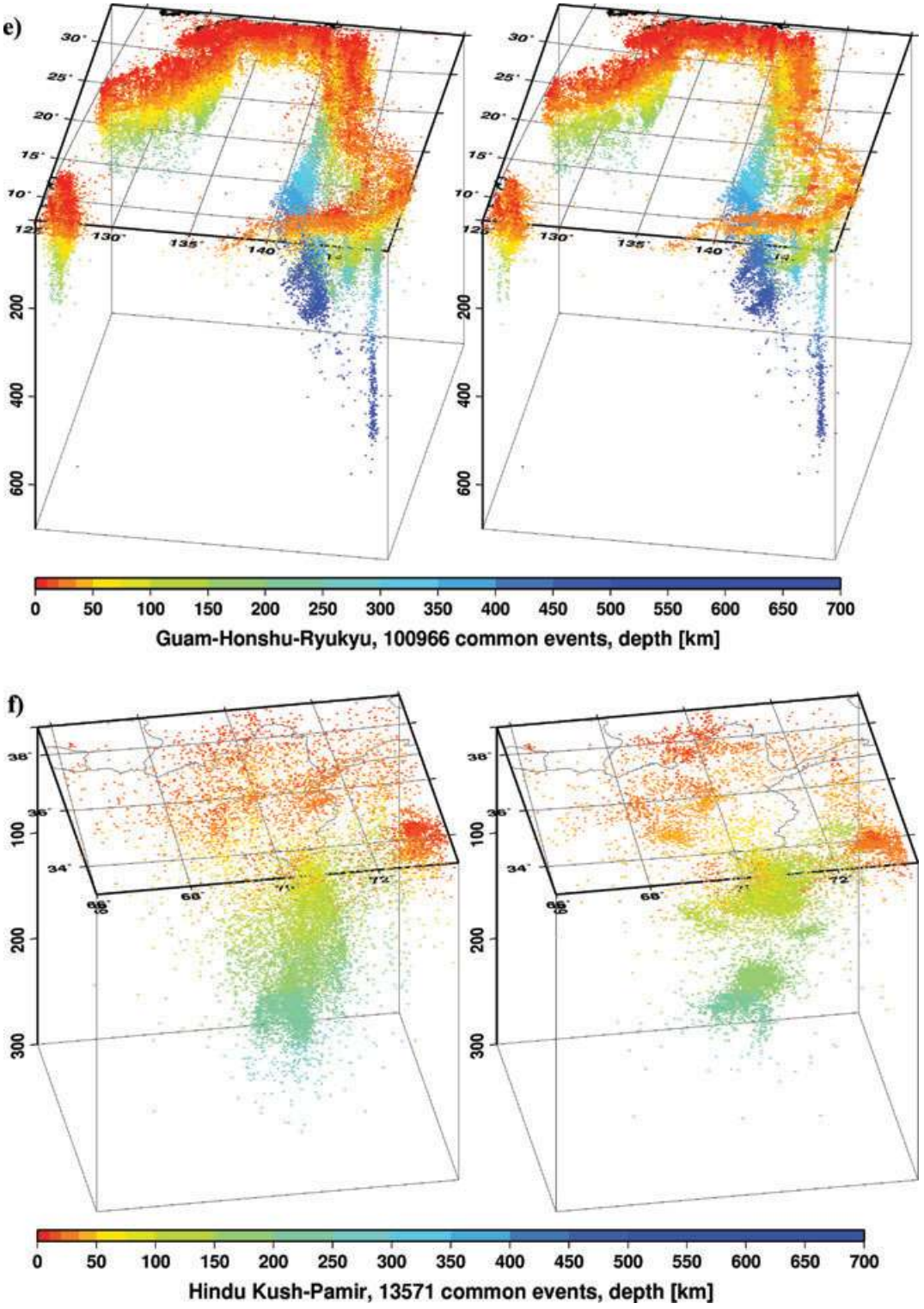

Figure 18. (*Continued*.)

default depth estimate. It is more difficult to characterise the effect of secondary and other later phases. They carry invaluable information about the earthquake hypocentre, but they have to be treated with caution. For instance, using far-regional phases in the location algorithm could make location worse due to 3-D heterogeneities unmodelled by the 1-D velocity model. This is why we have assigned larger *a priori* measurement errors to later arriving phases.

Significant location improvements could only be provided by better traveltime predictions, such as a global 3-D velocity model. Although the past years have seen considerable progress in developing global 3-D velocity models that give accurate traveltime predictions (e.g. Myers & Schultz 2000; Ritzwoller *et al*. 2003; Yang *et al*. 2004; Morozov *et al*. 2005; Murphy *et al*. 2005; Reiter *et al*. 2005; Flanagan *et al*. 2007; Zucca *et al*. 2009; Myers *et al*. 2010),

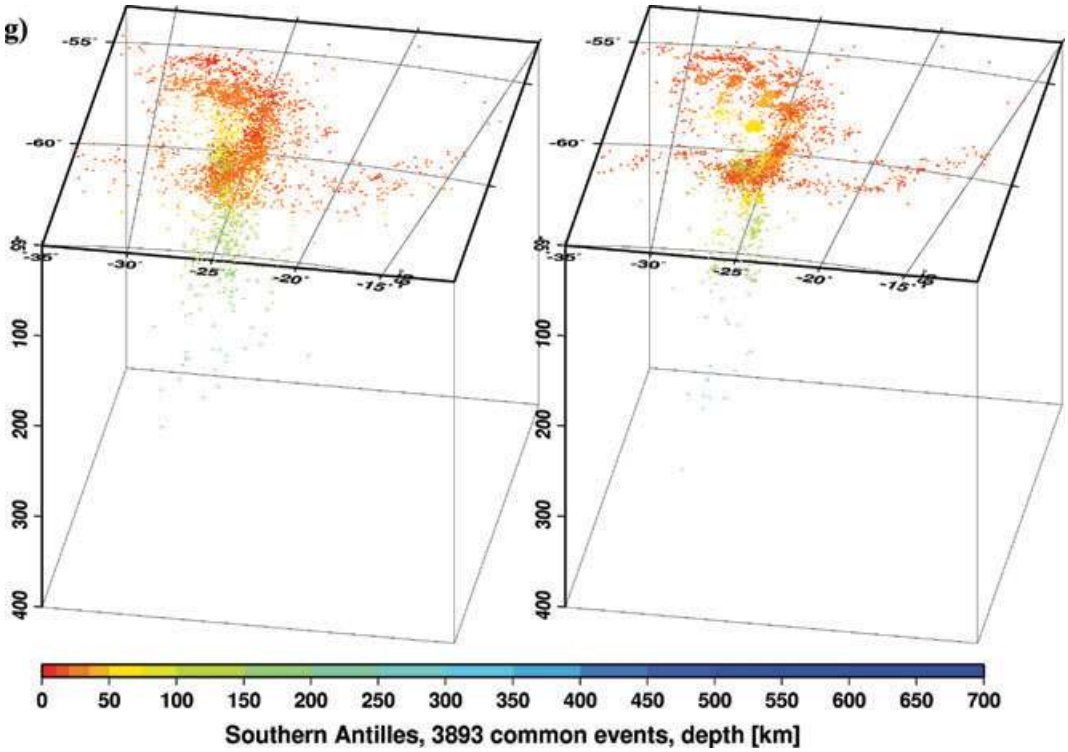

Figure 18. (*Continued*.)

they are not yet ready for routine operations. Furthermore, given the vast number of phases reported to the ISC, ray tracing through any 3-D models would be prohibitively expensive with the current state-of-the-art ray tracers. Therefore, we aimed for a more modest goal of obtaining the best possible hypocentre and corresponding uncertainty estimates using a global 1-D model.

Using the new ISC location procedures, we plan to rebuild the entire ISC Bulletin in order to provide a homogeneous bulletin of the seismicity of the Earth. Note that the current ISC Bulletin is not entirely homogeneous. For instance, the ISC used only first-arriving *P* phases in the location until 2001 when first-arriving S phases were introduced in the location procedures; the *ak135* model was introduced in 2006; and procedures that determined what reported events are selected for ISC review were changed in 1999, 2005 and 2006. Furthermore, the ISC only began publishing MS magnitudes in 1978, even though surface wave amplitudes have been regularly reported by agencies since 1971. Current efforts in collecting surface wave amplitude readings from historical station bulletins under the Global Earthquake Risk Model (www.globalquakemodel.org) project will further improve the continuity and uniformity of the magnitude assessments. Hence, rebuilding the ISC Bulletin will give us the opportunity to rectify known inconsistencies in the bulletin; to get rid of spurious events and blunders; to identify and mark time periods for individual stations with erroneous time stamps; and to reassign event type flags for anthropogenic events.

During the rebuild we will introduce essential data sets that were not available at the time of the original ISC Bulletin production, such as the ISS data covering 1960–1963 as well as data from permanent networks that were reported too late to be included in the ISC review, or that were accidentally not used by the ISC. We will also include data from permanent networks recovered retrospectively following political upheavals or administrative disputes. We will collect phase picks, if available, from temporary deployments including ocean-bottom seismometers. We anticipate that this effort will take several years to complete.

We have demonstrated that the new ISC location algorithm provides small, but consistent location improvements, considerable improvements in depth determination and significantly more accurate formal uncertainty estimates. The default depth grid provides reasonable depth estimates where there is seismicity. We have shown that the location and depth accuracy obtained by the new algorithm matches or surpasses the EHB accuracy. We have also demonstrated that the new algorithm, through the use of later phases and testing for depth resolution, considerably clusters event locations, thus providing an improved view of the seismicity of the Earth.

DATA AND RESOURCES

All data used in this paper came from published sources listed in the references. The IASPEI Reference Event List, the EHB and the ISC Bulletins are hosted by the ISC website, www.isc.ac.uk. The current ISC locator is available for download from the ISC website, and we will also make the new locator available to the public. We used the open source Cluster 3.0 library by de Hoon *et al*. (2004) and generated figures using the Generic Mapping Tool (GMT) software developed by Wessel & Smith (1991).

ACKNOWLEDGMENTS

This work was partly supported by the National Science Foundation grants EAR 0548649 and EAR 0949072, as well as by 56 ISC member-institutions. We are especially grateful to the NSF, the Indian Meteorological Department, the Japan Marine Science and Technology, the Earthquake Research Institute of University of Tokyo and the China Earthquake Administration for providing

additional funding that will help in re-building the ISC Bulletin for the entire period of its existence.

We thank Malcolm Sambridge and Brian Kennett (Australian National University) for making code of the NA available to us; Bob Engdahl (Colorado University at Boulder) for discussions and for providing the code for bounce point corrections; Jack Murphy and Brian Barker (Science Applications International Corp.) for the depth-phase stacking algorithm; Johannes Schweitzer (NORSAR) for discussions on the *libtau* software; we thank Louisa Tsang (Hong Kong University) for the relocation study in the Chile-Bolivia border region with and without USArray data she performed as an intern at the ISC; and last but not least we thank Richard Luckett (now at British Geological Survey) for converting the old Fortran ISC locator code into more readable C code during his tenure at the ISC. We are also grateful to Steve Myers and two anonymous reviewers whose comments helped to improve the paper.

REFERENCES

Adams, R.D., Hughes, A.A. & Gregor, D.M., 1982. Analysis procedures at the International Data Centre, *Phys. Earth planet. Int.,* **30**, 85–92.

Amante, C. & Eakins, B.W., 2009. ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis, *NOAA Technical Memorandum NESDIS NGDC-24*.

Anderson, K.R., 1982. Robust earthquake location using M estimates, *Phys. Earth planet. Int.,* **30,** 119–130.

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

Bolt, B.A., 1960. The revision of earthquake epicentres, focal depths and origin time using a high-speed computer, *Geophys. J. R. astr. Soc.*, **3,** 434–440.

Bolton, M.K., Storchak, D.A. & Harris, J., 2006. Updating default depth in the ISC bulletin, *Phys. Earth planet. Inter.*, **158,** 27–45.

Bondár, I. & McLaughlin, K., 2009a. Seismic location bias and uncertainty in the presence of correlated and non-Gaussian travel-time errors, *Bull. seism. Soc. Am.*, **99,** 172–193.

Bondár, I. & McLaughlin, K., 2009b. A new ground truth data set for seismic studies, *Seism. Res. Lett.*, **80,** 465–472.

Bormann, P., Liu, R., Xu, Z., Ren, K., Zhang, L. & Wendt, S. 2009. First application of the new IASPEI teleseismic magnitude standards to data of the China National Seismographic Network, *Bull. seism. Soc. Am.*, **99,** 1868–1891.

Buland, R. & Chapman, C.H., 1983. The computation of seismic travel times, *Bull. seism. Soc. Am.*, **73,** 1271–1302.

Buland, R., 1986. Uniform reduction error analysis, *Bull. seism. Soc. Am.*, **76,** 217–230,.

Chang, A.C., Shumway, R.H., Blandford,R.R. & Barker, B.W., 1983. Two methods to improve location estimates—preliminary results, *Bull. seism. Soc. Am.*, **73,** 281–295.

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference earth model, *Phys. Earth planet. Inter.,* **25,** 297–356.

Dziewonski, A.M. & Gilbert, F., 1976. The effect of small, aspherical perturbations on travel times and a re-examination of the correction for ellipticity, *Geophys. J. R. astr. Soc.*, **44,** 7–17.

Engdahl, E.R., van der Hilst, R. & Buland, R., 1998. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. seism. Soc. Am.*, **88,** 722–743.

Evernden, J., 1969. Precision of epicenters obtained by small numbers of world-wide stations, *Bull. seism. Soc. Am.*, **59,** 1365–1398.

Flanagan, M.P., Myers, S.C. & Koper, K.D., 2007. Regional travel-time uncertainty and seismic location improvement using a three-dimensional a priori velocity model, *Bull. seism. Soc. Am.*, **97,** 804–825.

Gutenberg, B. & Richter, C.F., 1956. Magnitude and energy of earthquakes, *Ann. Geof.*, **9,** 1–15.

Hicks, E.C., Kvaerna, T., Mykkeltveit, S., Schweitzer, J. & Ringdal, F., 2004. Travel-times and attenuation relations for regional phases in the Barents Sea region, *Pure appl. Geophys.*, **161,** 1–19.

de Hoon, M. J. L., Imoto, S., Nolan, J. & Miyano, S., 2004. Open source clustering software, *Bioinformatics*, **20**, 1453–1454.

Jeffreys, H., 1932. An alternative to the rejection of observations, *Proc. Roy. Soc. Lond.*, **187,** 78–87.

Jeffreys, H. & Bullen, K.E., 1940. *Seismological Tables*, British Association for the Advancement of Science, Gray-Milne Trust, London.

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,** 1105–1130.

Kennett, B.L.N., 2006. Non-linear methods for event location in a global context, *Phys. Earth planet. Int.*, **158,** 45–64.

Kennett, B.L.N. & Gudmundsson, O., 1996. Ellipticity corrections for seismic phases, *Geophys. J. Int.*, **127,** 40–48.

Kennett, B.L.N., Engdahl, E.R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, **122,** 108–124.

Kremenetskaya, E., Asming, V. & Ringdal, F., 2001. Seismic location calibration of the European Arctic, *Pure appl. Geophys.*, **158,** 117–128.

Levshin, A.L., Schweitzer, J., Weidle, C., Shapiro, N.M. & Ritzwoller, M.H., 2007. Surface wave tomography of the Barents Sea and surrounding regions, *Geophys. J. Int.*, **170,** 441–459.

Morozov, I.B., Morozova, E.A., Smithson, S.B., Richards, P.G., Khalturin, V.I. & Solodilov, L.N., 2005. 3D first-arrival regional calibration model of Northern Eurasia, *Bull. seism. Soc. Am.*, **95,** 951–964.

Murphy, J.R. & Barker, B.W., 2006. Improved focal-depth determination through automated identification of the seismic depth phases pP and sP, *Bull. seism. Soc. Am.*, **96,** 1213–1229.

Murphy, J.R., *et al*., 2005. Calibration of International Monitoring System (IMS) stations in Eastern Asia for improved seismic event location, *Bull. seism. Soc. Am.*, **95,** 1535–1560.

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,** 199–211.

Myers, S.C., *et al*., 2010. A crust and upper mantle model for Eurasia and North Africa for Pn travel-time calculation, *Bull. seism. Soc. Am.*, **100,** 640–656.

Nicholson T., Sambridge, M., & Gudmundsson, O., 2000. On entropy and clustering in earthquake hypocenter distributions, *Geophys. J. Int.*, **142,** 37–51.

Reiter, D., Rodi, W., & Johnson, M., 2005. Development of a tomographic upper-mantle velocity model beneath Pakistan and Northern India, *Bull. seism. Soc. Am.*, **95,** 926–940.

Ritzwoller, M.H., Shapiro, N.M., Levshin, E.A., Bergman, E.A., & Engdahl, E.R., 2003. Ability of a global three-dimensional model to locate regional events, *J. geophys. Res.*, **108**(B7), 2353, doi:10.1029/2002JB002167.

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

Sambridge, M. & Kennett, B.L.N., 2001. Seismic event location: non-linear inversion using a neighbourhood algorithm, *Pure appl. Geophys.*, **158,** 241–257.

Schweitzer, J. & Storchak, D.A., 2006. Modernizing the ISC location procedures, Editorial, *Phys. Earth planet. Inter.*, **158,** 1–3.

Shewchuk, J.R., 1996. Engineering a 2D quality mesh generator and Delaunay triangulator, *First Workshop on Computational Geometry*, Philadelphia, ACM, pp. 124–133.

Storchak, D.A., Chen, Q.F., Willemann, R.J., & Andrianirina, M., 2000. Improved locations for moderately large earthquakes using regional S and PKP. *EOS, Trans. Am. geophys. Un.*, **81**(48), S71C-01.

Storchak, D.A., Schweitzer, J., & Bormann, P., 2003. The IASPEI standard seismic phase list, *Seism. Res. Lett.*, **74,** 761–772.

Vanek, J., Zatopek, A., Karnik, V., Kondorskaya, N.V., Riznichenko, Y.V., Savarensky, Y.F., Solovev, S.L. & Shebalin, N.V., 1962. Standardization of magnitude scales, *Bull. (Izvest.) Acad. Sci. USSR, Geophys. Ser.*, **2,** 108–111.

Villaseñor, A. & Engdahl, E.R., 2005. A digital hypocenter catalog for the International Seismological Summary, *Seism. Res. Lett.*, **76,** 554–559.

Villaseñor, A. & Engdahl, E.R., 2007. Systematic relocation of early instrumental seismicity: earthquakes in the International Seismological Summary for 1960–1963, *Bull. seism. Soc. Am.*, **97,** 1820–1832.

Wessel, P. & Smith, W. H. F., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.*, **72,** 441.

Willemann, R.J. & Storchak, D.A., 2001. Data collection at the International Seismological Centre, *Seism. Res. Lett.*, *7***2,** 440–453.

Yang, X., *et al*., 2004. Validation of regional and teleseismic travel-time models by relocating GT events, *Bull. seism. Soc. Am.*, **94,** 897–919.

Young, J.B., Presgrave, B.W., Aichele, H., Wiens, D.A. & Flinn, E.A., 1996. The Flinn-Engdahl regionalization scheme: the 1995 revision, *Phys. Earth planet. Int.*, **96,** 223–297.

Zucca, J.J., *et al*., 2009. The prospect of using three-dimensional Earth models to improve nuclear explosion monitoring and ground-motion hazard assessment, *Seism. Res. Lett.*, **80,** 31–39.