doi:10.26443/seismica.v4i1.1536

SEISMICA

# Revising the Seismic Ground Truth Reference Event Identification Criteria

**Ryan Gallacher** * [1], **Tom Garth** [1], **James Harris** [1], **István Bondár** [2], **Keith McLaughlin** [3], **Dmitry A. Storchak** [1]

[1]International Seismological Centre, Thatcham, UK, [2]Seismic Location Services, Lagos, Portugal, [3]Leidos Inc, Reston, VA, USA

Author contributions: *Conceptualization*: Ryan Gallacher, Tom Garth, James Harris. *Methodology*: Ryan Gallacher, Tom Garth, István Bondár. *Data Curation*: Ryan Gallacher, James Harris. *Formal Analysis*: Ryan Gallacher, Tom Garth, James Harris, Keith McLaughlin. *Funding Acquisition*: Dmitry Storchak. *Writing – original draft*: Ryan Gallacher, Tom Garth. *Writing – review & editing*: Ryan Gallacher, Tom Garth, István Bondár, Keith McLaughlin, Dmitry Storchak.

**Abstract** Accounting for improvements in the location algorithms used to relocate seismic events, we investigate the criteria used to identify well constrained seismic event locations, particularly seismic Ground Truth events. By relocating explosions with known epicentres using modern location techniques, we have determined that the parameter measuring unbalanced station distributions, ΔU, can be replaced by a more general measure of azimuthal station coverage defined as the Cyclic Polygon Quotient. This is the ratio of the area of a cyclic polygon formed by connecting event to station azimuths on a unitary circle, and the area of the unitary circle. We demonstrate that the semi-major axis of the error ellipse after relocation is a strong discriminant of location quality. We show that hypocentre depth can be resolved where multiple seismic stations report both P & S phase arrivals, providing an alternative to the previous Ground Truth identification criteria, which requires a station within 10 km of the event. These findings are incorporated into an updated set of criteria for identifying well constrained seismic event locations. We show that the updated criteria increases both the number and geographic distribution of Ground Truth events that can be added to the IASPEI Reference Events List.

Production Editor:
Yen Joe Tan
Handling Editor:
Hongyu Sun
Copy & Layout Editor:
Ethan Williams

Signed reviewer(s):
Anthony Lomax, Yifan Yu

Received:
16 December 2024
Accepted:
19 May 2025
Published:
2 July 2025

## 1 Introduction

Seismic events with well constrained locations and origin times, referred to as "Ground Truth" (GT) events, are an important resource allowing seismologists to test and validate new techniques relating to seismic event locations, such as location algorithms and velocity models (e.g. Begnaud et al., 2020). As we demonstrate in this study GT events are of particular use for calibrating the absolute event locations (e.g. Belinić and Markušić, 2017) and additionally GT events can be used as the seed locations for multi-event relocation techniques (e.g. Bergman et al., 2022; Bondár et al., 2023).

GT events can be defined in many ways: by explosions or other anthropogenic events, where the location and timing of the event is known or well documented by nearby observations (e.g. Bennett et al., 2010; Bittner et al., 2022); seismic events that are well located by a dense and well distributed near-field seismic network (e.g. Boomer et al., 2010; Bondár and McLaughlin, 2009); seismic events well located by a multi-event location technique (e.g. Bondár et al., 2008; Bergman et al., 2022). Additionally, there are less established methods of obtaining GT events including using ambient seismic noise (e.g. Zeng et al., 2014) and InSAR (e.g. Zhu et al., 2021). In this study we are concerned with the second of these cases, where GT events are well located earthquakes defined by single event GT criteria. Single Event GT events (referred to here after simply as GT events) are defined as events that have an acceptable station distribution, i.e. there are plentiful well distributed stations within 150 km of the proposed GT event, contributing to defining the event origin, with a well-defined hypocentre and a relatively small error ellipse after relocation.

The criteria currently used to identify GT5 events (ground truth events where the epicentre is known to within 5 km accuracy), described by Bondár and McLaughlin (2009), requires at least one station within 10 km of the candidate GT event, a maximum secondary gap (defined as the largest azimuthal gap in the station coverage if one station is removed) of 160° and a ΔU (network quality metric, describing station distribution as proposed by Bondár and McLaughlin (2009) of less than or equal to 0.35 (See Table 1). Identification and certification of GT events occurs in two parts, GT candidate events are identified with the above criteria after which the event is relocated using only stations within 150 km (1.36°) of the event. The location output is certified as a GT5 event if the semi-major axis of the error ellipse (henceforth called semi-major axis) is less than 5km, the earthquake depth has not been fixed to a set value and the candidate criteria remain met.

Many potential GT candidate events fail to meet these criteria, significantly limiting the geographic distribution of GT events. We investigate the GT criteria with the intention of increasing the geographic distribution of GT events and evaluating the applicability of the criteria given the advances in earthquake location methods

*Corresponding author: ryan@isc.ac.uk

| Current Ground Truth Criteria |
|---|
| Event Magnitude < 6.1 |
| At least one station within 10 km of event |
| Secondary Azimuthal Gap < 160° |
| Balanced station distribution, ΔU < 0.36 |
| Event recorded at a station at a distance of at least 2° |
| *Semi-major axis of the error ellipse $\leq$5 km* |
| *Earthquake depth is not fixed to a set value* |

**Table 1** The current Ground Truth criteria (Bondár & McLaughlin, 2009a), with the criteria in italics applied after relocation. Other criteria are applied before and after relocation.

since the most recent GT identification criteria were defined (Bondár and McLaughlin, 2009).

## 1.1 History of Ground Truth Events

The standardisation of GT event criteria was achieved by an iterative process of criteria refinement and publication (e.g. Sweeney, 1996, 1998; Bondár et al., 2001, 2004a,b) leading to a broadly accepted set of definitions (Bondár et al., 2004a). This set of criteria is based on the number and distribution of stations local to a candidate GT event. Further development of the GT criteria was designed to avoid correlated travel time errors at nearby stations. This was achieved through the addition of a new criteria, ΔU, which quantifies the azimuthal distribution of the station locations in the local area around the event (Bondár and McLaughlin, 2009). This criteria is designed to eliminate GT candidate events with unbalanced station distributions as they may be disproportionally affected by correlated travel time errors or more generally by a majority of stations in one particular azimuthal direction. The revised set of criteria described by Bondár and McLaughlin (2009) is used to populate a searchable database of ground truth events, the IASPEI Ground Truth Reference Event list, which is maintained by the ISC (International Seismological Centre). The IASPEI Ground Truth Reference Event list is supervised by the Commission on Seismological Observation and Interpretation (CoSOI) working group on Reference Events for Improved Locations. This database currently contains 12,278 ground truth events, with all events added since 2009 being based on the criteria given by Bondár and McLaughlin (2009).

## 1.2 Considerations for Modern Ground Truth

The ever-increasing numbers of seismic stations reporting seismic phase arrivals, and dense seismic deployments in particular, have resulted in an increased number of widely recorded events with unbalanced station distributions. This has led to a greater proportion of well constrained events being rejected based solely on the value of ΔU. Given the relative paucity of ground truth events in some regions, the rejection of events due to ΔU alone leaves some areas of the globe underrepresented by the GT list, limiting the ability to verify seismic velocity models and earthquake location techniques in these areas.

Improvements in earthquake location methods, such as correcting for bias in the distribution of seismic stations (Bondár and Storchak, 2011; Bondar and McLaughlin, 2009) suggest that the strictest criteria of station distribution, ΔU, may have been superseded. Specifically, ISCloc (Bondár and Storchak, 2011), which is used in the ISC-Bulletin and IASPEI GT reference list locations for seismic events occurring since January 2011, accounts for correlated errors in closely related stations following Bondár and McLaughlin (2009). The correlated errors are accounted for by the non-diagonal elements of the covariance matrix as defined by Bondár and Storchak (2011), with the covariance for a given station pair depending on the station separation. The covariance matrix can vary at every iteration of the linear relocation, as phases are redefined or even rejected.

The posterior data covariance matrix, calculated for the final converged hypocentre is used to define the error ellipse. The remaining phase residuals are combined with the posterior data covariance matrix to inform the 4D error ellipse (e.g. equation 8 of Bondár and Storchak (2011)). This 4D error ellipse (latitude, longitude, depth and origin time) is then used to define the 2D horizontal error ellipse, described by the semi-major axis, semi-minor axis, and the orientation of the semi-major axis, as well as 1D errors for event depth and origin time. The reported error ellipse is scaled to a 90% confidence level, through benchmarking with the original GT list (Bondár and Storchak, 2011).

The single event GT earthquakes considered in this study are seeded with well recorded events from the ISC Bulletin, that fulfil the criteria discussed and refined in this paper (see Table 1 and 2). The events are then relocated using ISCloc which employs an iterative linear relocation procedure, where the hypocentre is refined from a given starting point through the linearised reduction of travel time residuals between the observed seismic phases, and those predicted by ak135 (Kennett et al., 1995). ISCloc attempts to solve for the event depth where one of the following is true; there is at least one reported station within 0.2° distance, there are at least five stations reporting P & S phases within 3° distance, there are at least five reported depth phases or there are at least five core reflection phases. We note that the last two of these are irrelevant for constraining depths for GT events defined using local data. If the linear inversion fails to converge using a resolved depth from the above criteria then the inversion is repeated with a "fixed depth" which is taken from a geographic grid of user defined depths. In the case of GT qualifying events, the free depth criteria within ISCloc will almost certainly be met for all events considered, as the equivalent GT criteria are much stricter. Fixed, or unresolved depths can still occur however when the linear relocation procedure fails to converge. This may occur if the available phases have elevated degrees of error (e.g. pick errors resulting from noisy waveform data), or if the travel times predicted from the 1D velocity model account for the arrival times of the observed phases poorly. If the depth is unresolved and thus set to a fixed depth, the event is rejected as a GT event.

The improvements to the location procedures imple-

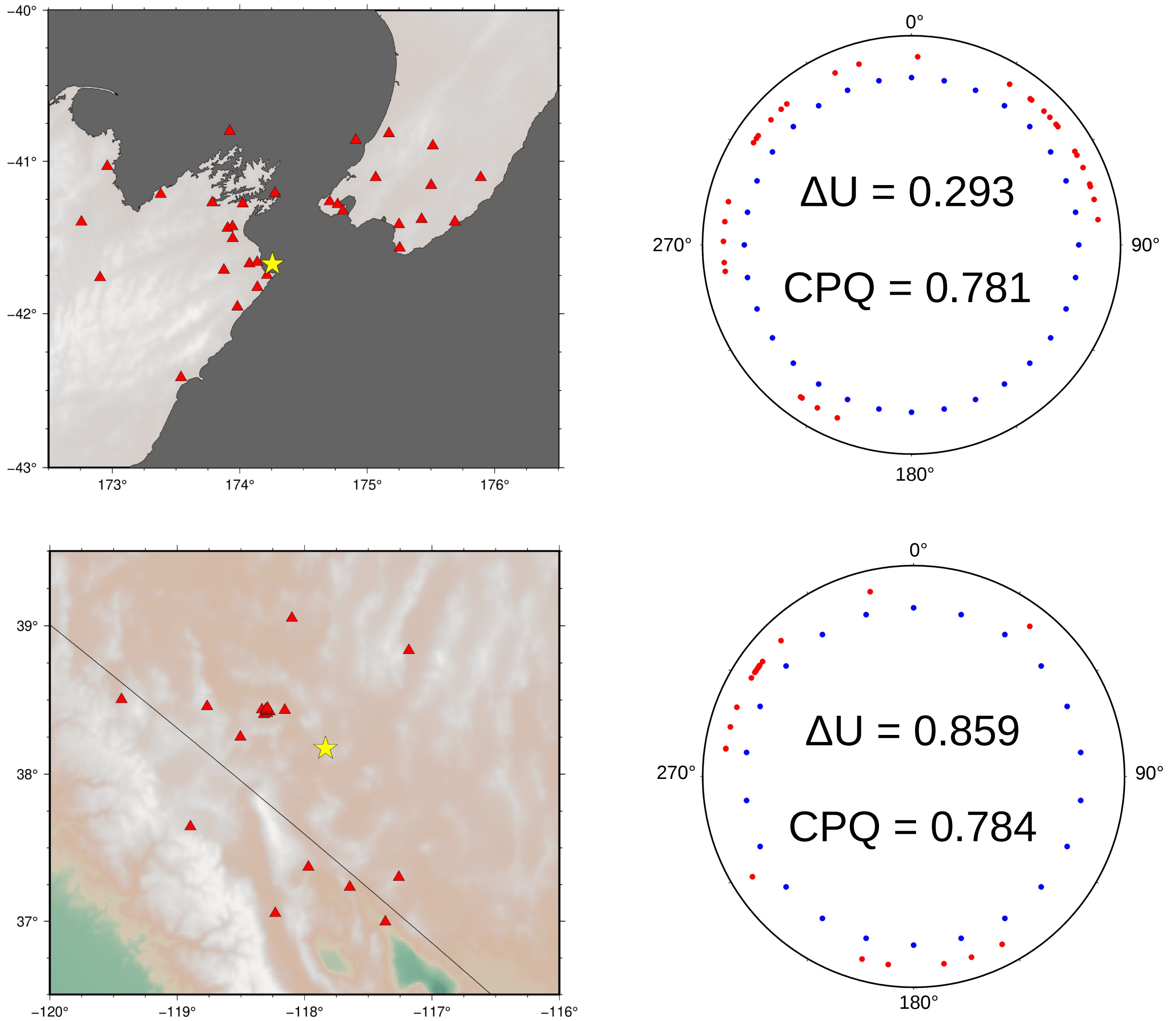


**Figure 1** Top left: Plot of event (yellow star) and stations (red triangles) for a ML 3.4 in New Zealand on 2nd May 2018. Top right: plot showing the distribution of event to station azimuths (red circles) and the same number uniformly distributed event to station azimuths (blue circles) on a unitary circle (right) for the New Zealand event. Bottom left: Plot of event and stations for a ML 3.3 in Nevada on 15th May 2020. Bottom right: plot of event to station azimuths and the same number uniformly distributed event to station azimuths on a unitary circle (right)

mented in ISCloc by Bondár and Storchak (2011) provide an opportunity to update the GT criteria. If we can rely on the error ellipse produced by the location algorithm, including accounting for bias arising from uneven station distributions, to identify GT events that have a location known to within 5 km, then the pre-relocation GT criteria can be more permissive. The purpose of these criteria therefore becomes filtering out events that cannot be expected to produce a sufficiently small error ellipse, as well as ensuring that the event is well enough recorded to exclude the possibility of an error ellipse being underestimated.

The ΔU criteria is a measure of the deviation of the azimuthal station distribution from a perfect azimuthal distribution. For example, a perfect distribution of five points would be described by the corners of a pentagon. This means that adding a single station can decrease the value of ΔU by altering the perfect distribution that the station azimuths are compared to, thus causing a GT candidate event to fail while all other criteria are improved. We show that ΔU is highly sensitive to seismic networks that deviate even slightly from a perfect geometrical distribution, for example if there is a cluster of stations on one azimuth (e.g. Figure 1, bottom). This has the effect that as the density of seismic stations increases, it becomes ever more implausible that the station distribution required for a GT can be achieved, especially in well instrumented areas (e.g. Japan). While such a stringent criteria may well have been necessary to guard against highly uneven or clustered station distributions in the past, when using location algorithms that do not account for correlated station errors, we show that modern location algorithms that correct for uneven station distributions (e.g. Bondár and Storchak,

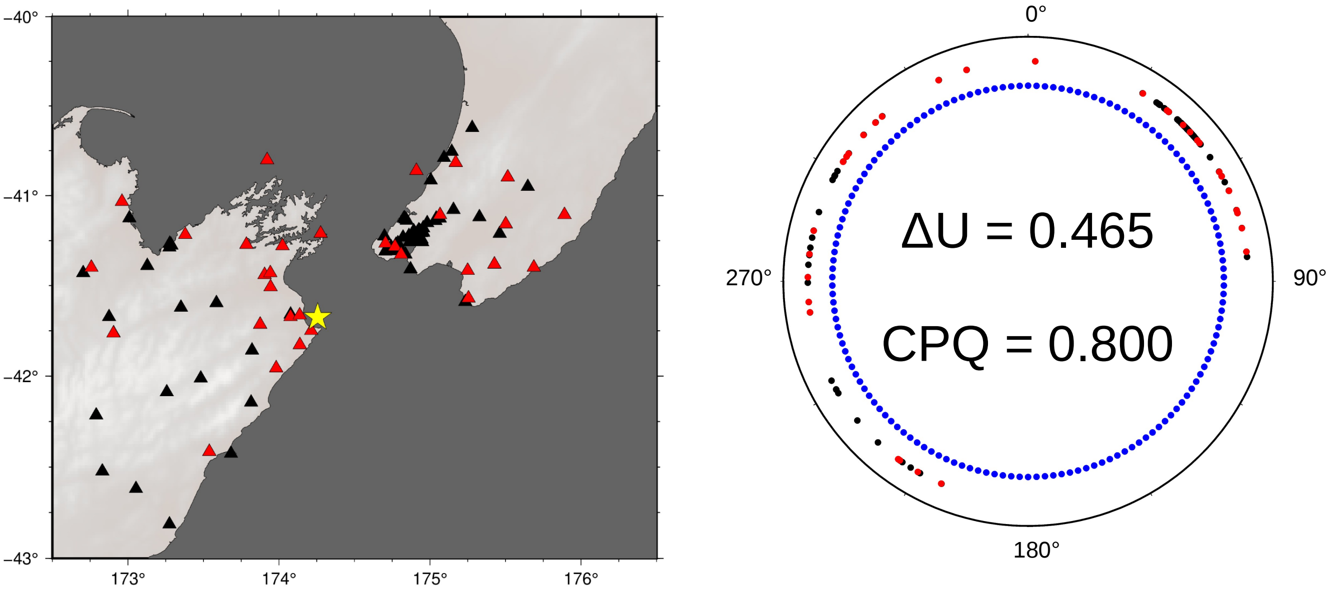


**Figure 2** Left: Event (yellow star), reported stations (red triangles) and unreported stations (black triangles) for the ML 3.4 in New Zealand on 2nd May 2018. Right: Event to station azimuths for reported stations (red circles), unreported stations (black circles) and the same total number of uniformly distributed event to station azimuths (blue circles) on a unitary circle (right).

2011) negate the requirement for such a non-inclusive criteria.

We propose a replacement criteria referred to as the Cyclic Polygon Quotient (CPQ), which better describes the azimuthal station coverage for an event. We demonstrate that CPQ and secondary gap are the most useful GT candidate criteria and show that the semi-major axis of the error ellipse derived in the relocation procedure is a robust criteria for selecting relocated GT events. Additionally, we demonstrate that the requirement for a station to be within 10 km of a GT event can be relaxed when sufficient P & S phase pairs are reported for nearby stations.

## 2 Quantifying Unbalanced Station Distributions

The ΔU parameter is a normalised measure of the difference between a uniform azimuthal distribution of local stations and the actual station distribution. Only stations within 150 km of the event are defined as local and are used to determine ΔU, defined by Bondár and McLaughlin (2009) as:

$$\Delta U = 4\frac{\sum |esaz_i - (unif_i + b)|}{360N}, \quad 0 \leq \Delta U \leq 1 \qquad (1)$$

where $esaz_i$ is the $i^{th}$ event to station azimuth in the sorted list of $esaz_i$, $unif_i$ is each uniform event to station azimuth ($unif_i = 360i/N$ for $i = 0, \dots, N-1$), the $b$ value calculates the offset between the average actual and uniform station distributions where $b = \text{avg}(esaz_i) - \text{avg}(unif_i)$ and $N$ is the total number of stations.

To demonstrate the calculation of ΔU we consider two well recorded earthquakes and the associated station distributions within 150 km of the events. A map of the locations of these example earthquakes, a ML 3.4 (reported by GNS Science) in New Zealand on 2nd May 2018 and a ML 3.3 (reported by National Earthquake Information Center) in Nevada on 15th May 2020, alongside the reported stations for the events is shown in Figure 1. The New Zealand event has a ΔU of 0.293 and the Nevada event has a ΔU of 0.859. Figure 1 also shows the same station distribution for both events represented by plotting event to station azimuth at a unitary distance, along with the associated value of ΔU. While both example earthquakes have a good distribution of nearby stations, the stations are azimuthally clustered. This means that despite the prevalence of well distributed stations, the Nevada event would not qualify as a GT event due to the value of ΔU, whereas the New Zealand event would qualify as a GT event. Removing the majority of the clustered stations from the Nevada event could improve the value of ΔU such that the event passed the ΔU criteria. In contrast, the values of CPQ are very similar for the two events and requires no stations to be removed.

### 2.1 Drawbacks of ΔU

An issue with using ΔU as a constraint on the selection of GT events is the possible increase in ΔU as additional stations are added to an event. This is demonstrated in Figure 2, by taking the example event from New Zealand, which is currently part of the IASPEI Ground Truth Reference Events list, and adding picks from nearby unreported stations. The example event considered is reported at 32 stations within 150 km of the event. This event has a ΔU of 0.293, a secondary gap of 127.7° and multiple stations within 10 km. After relocating the event with ISCloc, the semi-major axis for this event is 4.58 km.

Figure 2 shows the location and azimuth of 114 additional unreported stations within 150 km of the event and a plot of the event to station azimuth including these additional stations. If phases from all of these stations were picked and made available, the revised ΔU for this event would be 0.465, with a secondary gap of 120.9°. Although the azimuthal coverage of the event recordings improves, and the secondary azimuthal gap improves, the increase in ΔU means that the event no longer meets the ground truth criteria. This also implies that ground truth events can be "created" by removing reported phases with no reference to whether the phases are accurate or not. With increasingly dense seismic deployments, both permanent and temporary, it is possible that regions which previously provided ground truth events may no longer produce them as the modern network is too dense (on certain azimuths). We therefore propose that a low value of ΔU is not necessary to provide a well constrained epicentre, and that without ΔU a greater number of GT events with a wider geographic distribution become possible.

### 2.2 Cyclic Polygon Quotient

Assuming that GT criteria are not required to account for unbalanced station distributions we propose an alternative criteria to ΔU, that is focused only on characterising the azimuthal distribution of stations around an event. This replacement uses the area of the cyclic polygon formed by plotting the event to station azimuths on a unitary circle and connecting them together. We define the ratio between the area of the polygon and the area of the unitary circle as the Cyclic Polygon Quotient (CPQ). Given that there are multiple methods to calculate the area of a cyclic polygon there are thus multiple correct ways to calculate CPQ, we show one below (Braden, 1986):

$$CPQ = \frac{\left| \sum_{i=1}^{n-1} x_i y_{i+1} + x_n y_1 - \sum_{i=1}^{n-1} x_{i+1} y_i - x_1 y_n \right|}{2\pi} \quad (2)$$

where x and y are the Cartesian coordinates of the event to station azimuths ($esaz_i$) ordered from 0 to 360 degrees, the subscript refers to the number of the vertex in clockwise order. The Cartesian coordinates are found using $x = \cos(esaz_i)$ and $y = \sin(esaz_i)$ with a radius of 1. Once the vertices of the cyclic polygon are in Cartesian coordinates, we can apply the "Shoelace Formula" where in a clockwise (or counter-clockwise) direction we calculate the sum of the product of the x coordinate value with the y coordinate value of the next vertex and the subtraction of the product of the y coordinate value with the x coordinate value of the next vertex. We can subtract these two sums and divide by two to get the area of the cyclic polygon. By dividing this value by π, we get the ratio of the area of the cyclic polygon to the area of the unitary circle.

Larger values of CPQ imply a larger area of the unitary circle is covered/enclosed by the polygon, indicating a better azimuthal coverage. Figure 3 shows a simplified synthetic example of how CPQ is calculated. Adding stations will always result in an increased value of CPQ, as demonstrated in Figures 1 & 2 for the New Zealand example event.

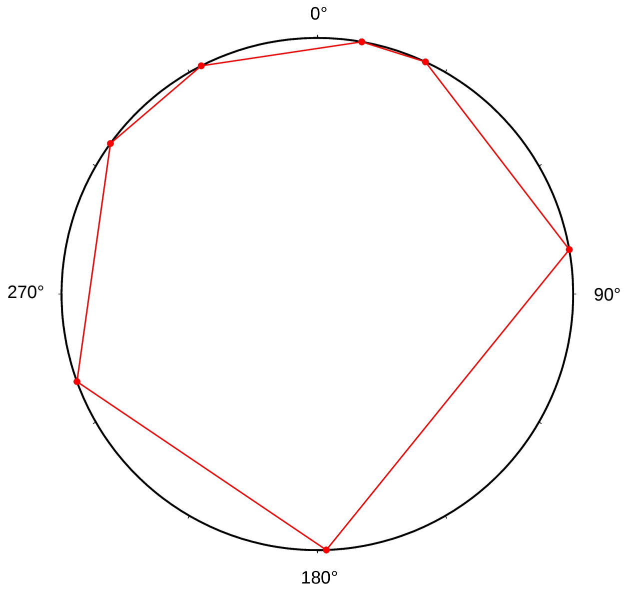


**Figure 3** Synthetic example of a Cyclic Polygon formed by connecting event to station azimuths on a unitary circle. Red dots are event to station azimuths, the red line shows the Cyclic Polygon. CPQ is given by the ratio enclosed by the red polygon to the area enclosed by the black circle. This example has a CPQ of 0.781.

## 3 Investigating Mislocation & Depth for Seismic Event Relocation

So far, we have explored existing methods of identifying well located seismic events and proposed a new formulation to quantify the azimuthal distribution of reporting seismic stations. We now test the proposed methods of identifying well located seismic events using two data sets. The first data set we consider is comprised of anthropogenic explosions that have a known location and origin time (section 3.1). The second data set consists of natural earthquakes taken from the 2018-2020 Reviewed ISC Bulletin that are recorded on five or more seismic stations within 150 km of the seismic event epicentre (section 3.2).

### 3.1 Mislocation of Anthropogenic Explosions

One element of the GT data set are anthropogenic explosions where the location and origin time is known to a very high degree of certainty from near-field non-seismic observations (i.e. an anthropogenic test explosion that was placed in a known position and the detonation was precisely timed). Here we use a subset of anthropogenic explosions with known origins to test the how well CPQ, ΔU and the semi-major axis of the error ellipse relate to the accuracy of the event location once it is relocated using seismic observations and the GT relocation procedure. To quantify the mislocation we follow the methodology of Bondár and McLaughlin (2009), who define the mislocation as the distance between a known explosion site and the hypocentre arrived at through the relocation procedure, and use this as a metric to evaluate ΔU.

In this analysis Bondár and McLaughlin (2009) take a number of explosions with known locations and relocate each anthropogenic explosive event multiple times with a randomly defined subset of the seismic observations. These subsets of seismic stations (and the associated phase observations) then provide a test data set with which to quantify which features of the near-field seismic network most strongly control the event mislocation. Bondár and McLaughlin (2009) performed this analysis for 15,000 possible permutations of station distributions for 47 explosions. Here we replicate this test using 133 explosions that are distributed globally, each with an independently constrained location, resulting in 124,811 usable random permutations of station combinations and mislocation. Only station distributions which include five or more stations are evaluated, as opposed to the less restrictive three or more stations used by Bondár and McLaughlin (2009). Approximately 50,000 additional permutations failed to produce a value for mislocation as the number of randomly selected stations was insufficient for ISCloc to produce a stable hypocentre. Figure 4 shows the mislocation of each permutation with respect to ΔU and CPQ for all 124,811 qualifying and successful permutations.

The majority of events have a ΔU of between 0.1 and 0.6 with events at the 95$^{th}$ percentile are mislocated by between 4 – 9 km for this range of values of ΔU (Figure 5). This observed increase in mislocation of 5 km between ΔU = 0.1 and ΔU = 0.6 is comparable to the increase in mislocation found by Bondár and McLaughlin (2009), which was ~6 km in the same range.

The lower panel of Figure 4 shows the comparison of absolute mislocation with CPQ. The linear groupings of iterations each relate to an individual explosion event, demonstrating the correlation between high mislocation and low CPQ on an event by event basis. The majority of these linear structures demonstrate that increased CPQ values are correlated with decreasing mislocation, consistent with the general trend observed. In a minority of linear groupings, the opposite is observed, with increasing CPQ correlated with greater mislocation. This can be explained by an increasing proportion of stations with poor quality picks contributing to the constrained hypocentre, or by the greater chance of the location procedure being influenced by unmodelled local velocity perturbations.

These features suggest that for some of the explosions it is impossible to achieve a low mislocation or high CPQ, most likely due to the limitations from the available station geometry. Additionally small differences between this study and Bondár and McLaughlin (2009) may result from the different global velocity models used in the relocation procedures. Bondár and McLaughlin (2009) used iasp91 (Kennett and Engdahl, 1991), whereas this study uses the ak135 (Kennett et al., 1995). In both cases, however, there are many events with a ΔU greater than 0.35 and a mislocation less than 5 km, suggesting that ΔU is not a primary control on the mislocation of the event.

To investigate this further we compare the absolute mislocation of these seismically recorded anthropogenic explosions to the semi-major axis of the error ellipse calculated by ISCloc (Figure 5). For all permutations of the explosion data set ~8.8% (11,923 out of 124,811) are mislocated by greater than the semi-major axis. Focusing on the 5,393 iterations where the semi-major axis is less than 5 km, there are 151 iterations where the mislocation is greater than 5 km. This represents ~2.8% of iterations with a semi-major axis of less than 5km.

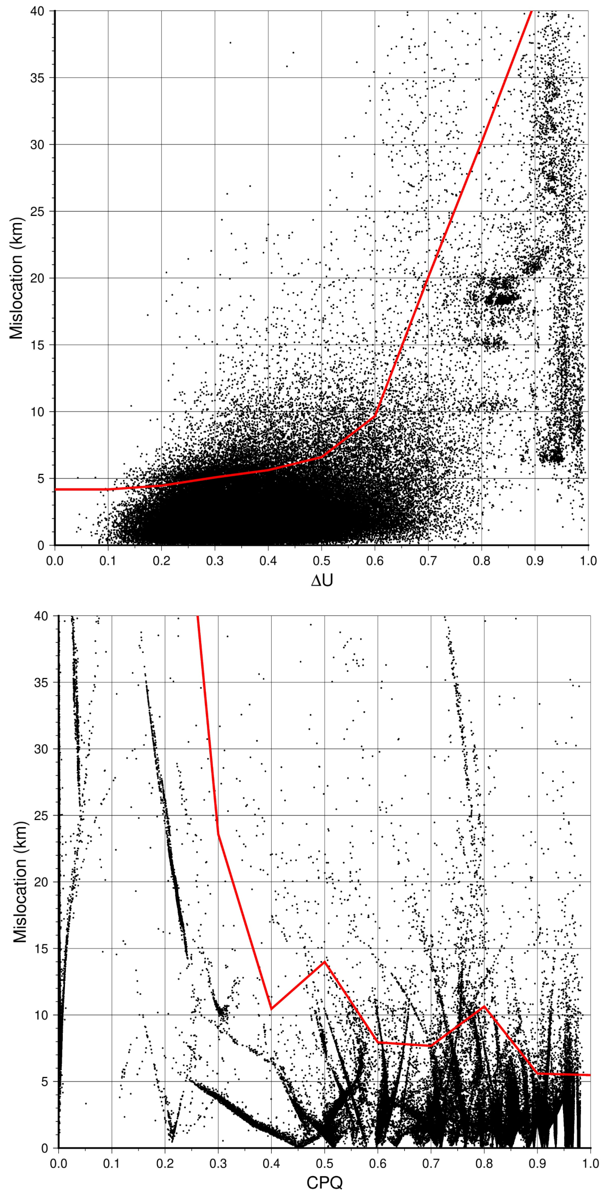


**Figure 4** Mislocation with respect to ΔU (top) and CPQ (bottom), for 124,811 random station selections on 133 explosions. The black dots are for each individual relocated permutation. For each bin of width 0.1, 95% of permutations are below the red line. The linear structures in the CPQ plot represent each of the individual 133 anthropogenic explosions, with the distribution reflecting the randomly sampled combinations of stations for each explosion.

From the top panel of Figure 6 we can see that events with a semi-major axis of less than 5 km have a limited range of values of CPQ (> 0.9) and ΔU (0.2 – 0.5), suggesting that the relationship between semi-major axis and mislocation is untested for cases where the station distribution is unbalanced or poor (i.e. high ΔU or low

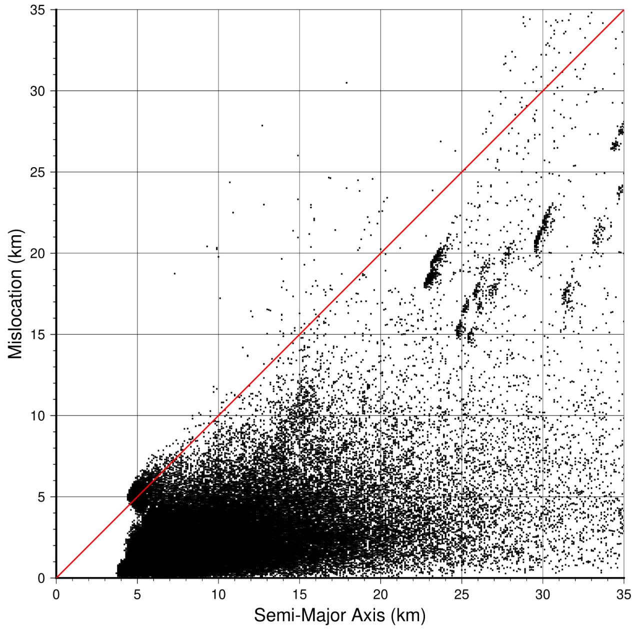


**Figure 5** Absolute mislocation of anthropogenic explosions with respect to semi-major axis of the error ellipse for 124,811 random station iterations. Red line shows where mislocation equals semi-major axis. 93.2% of the events are mislocated by less than the 90% semi-major axis.

CPQ). We therefore consider a wider subset of samples to include cases where the semi-major axis of the error ellipse is less than 10 km. As shown in the top panel of Figure 6, this allows for a wider range of values of CPQ (> 0.5) and ΔU (< 0.7) to be included. In this subset there are 4,565 iterations where the mislocation is greater than the semi-major axis, corresponding to ~7 % of iterations with a semi-major axis less than 10 km.

As a further verification of the validity of the relocation error ellipse to quantify location uncertainty, we calculate the number of iterations for which the known location of the anthropogenic explosion is within the error ellipse calculated during the relocation. We find that for ~87% of the explosion data set the known location is within the calculated (90% confidence interval) error ellipse. This corroborates the results from Bondár and Storchak (2011) who defined the 90% confidence interval when using ISCloc with global data, however, in our case we have replicated this result using only local data (e.g. seismic phases reported less than 150 km from the event).

This indicates (from Figure 5) that whereas the semi-major axis generally overestimates mislocation it rarely underestimates mislocation. This holds true even for poor and unbalanced station distributions. Based on this result we propose that the semi-major axis of the error ellipse is the best control on whether an event has a mislocation of 5 km or less and that other criteria are secondary. This demonstrates that the effect of unbalanced station distributions is accounted for effectively and successfully when the location procedure accounts for station-station correlations, as is the case for ISCloc (Bondár and Storchak, 2011). This removes the requirement for earthquakes with unbalanced reported station distributions to be excluded from the GT candidate list.

This combined with CPQs related feature of only improving with the addition of more stations, leads us to propose replacing ΔU with CPQ in GT candidate event selection.

The anthropogenic explosions considered in the testing described above are a specialised data set that are limited in number and geographical distribution. The number of anthropogenic explosions that can be used to calibrate to is limited as there are few events where the location is accurately known from non-seismological constraints. Additionally, these events are geographically limited to a small number of sites. This data set is further specialised as they primarily occur at the surface and in remote areas. Therefore, in order to conduct a broader test of GT criteria it is necessary to move away from testing absolute mislocation and instead use the semi-major axis to evaluate how accurately the event has been located. Given the relationship between mislocation and the semi-major axis of the error ellipse established above (i.e. ~93% of events with a semi-major axis of the error ellipse < 10 km are mislocated by less than the semi-major axis of the error ellipse) we propose using the semi-major axis of the error ellipse as a proxy for absolute mislocation for these events.

### 3.2 Testing Event Relocation Accuracy for ISC Reported Events

To investigate the GT criteria with a broader dataset we consider all reviewed events in the ISC Bulletin during the years 2018 – 2020, which have five or more stations within 150 km, resulting in a test dataset of 47,161 events. The bottom panel of Figure 6 shows the semi-major axis of the error ellipse with respect to ΔU and CPQ for this test data set of locally well recorded earthquakes.

In both cases there is a value of ΔU or CPQ at which it is highly unlikely that an event will have a semi-major axis of the error ellipse less than 5 km. For CPQ this value is 0.4 (0.02% of events with CPQ < 0.4 have a semi-major axis < 5 km) and for ΔU the value is 0.7 (0.05% of events with ΔU > 0.7 have a semi-major axis < 5 km). A minimum criteria of 0.4 CPQ excludes 18,089 events (~38.4%) whereas ΔU with a value of 0.7 excludes 7,857 events (~16.7%). Therefore, by using CPQ as a candidate criteria we will evaluate fewer events without a resulting reduction in the number of GT events identified.

Given the extended test data set we propose here, we also test other criteria for selecting GT candidates, including secondary azimuthal gap, stations within 10 km of the event and finally a potential new GT criteria considering the number of stations reporting both P & S phases. Figure 7 shows each of these constraints with respect to the semi-major axis of the error ellipse. As with CPQ and ΔU there is a value for secondary gap (~210°) that can be chosen above which it is highly unlikely that an event can have a semi-major axis of less than 5 km (0.31% of events with secondary gap < 210° have a semi-major axis < 5 km). Using this value a total of 21,611 (~46 % of initial candidates) events can be excluded. For stations reporting P & S phases or stations within 10 km, it is clear that even with no stations meet-

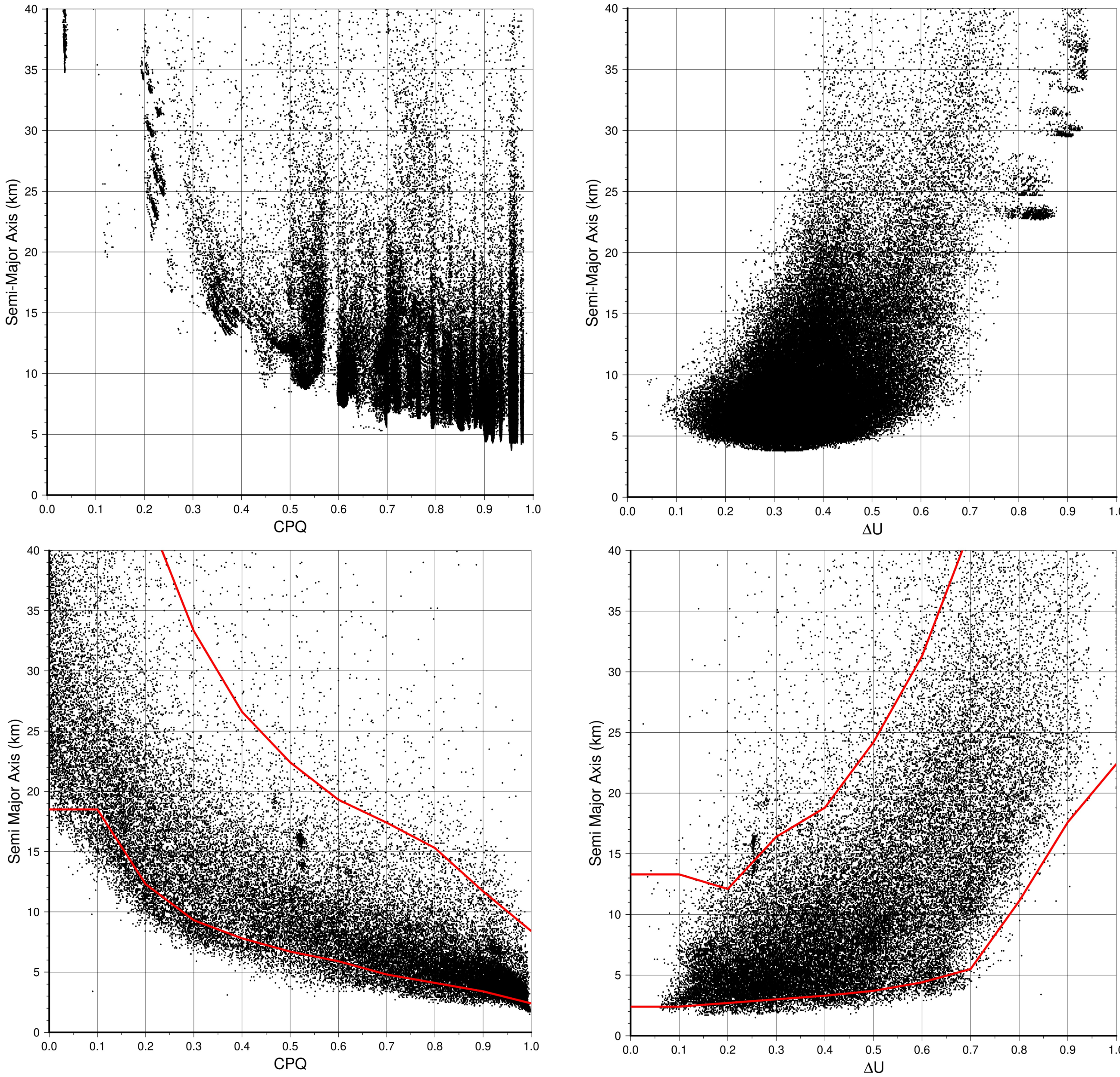


**Figure 6** Top: Semi-major axis of the error ellipse with respect to CPQ (left) and ΔU (right) for the resampled anthropogenic explosion relocation data set. Bottom: Semi-major axis of error ellipse with respect to CPQ (left) and ΔU (right) for the recent earthquake test data set. Red lines bound the area containing the central 90 % of samples, bins of size 0.1.

ing either of these criteria it is still possible to have a semi-major axis of less than 5 km, suggesting that these factors have limited control over the location accuracy. These criteria are however still considered, due to their potential control on earthquake depth, as discussed below.

Based on both the explosions and ISC Bulletin test datasets we propose the following criteria to identify GT candidates that are likely to have adequate location resolution: secondary azimuthal gap of less than or equal to 210° and a CPQ greater than or equal to 0.4. This would result in 24,811 events in the ISC Bulletin test dataset (covering the period 2018-2020) being considered as GT candidates, corresponding to ~53% of all events that were assessed. Additionally applying these criteria to the explosion test dataset results in 10,019 of all samples meeting GT criteria, 93.2% of which have an absolute mislocation of less than or equal to 5 km. This suggests that CPQ and secondary azimuthal gap are useful as both a candidate selection criteria and as certification criteria.

## 3.3 Event Depth Constraints for ISC Reported Events

An additional constraint for selecting GT events after relocation is a requirement for the locator to be able to resolve the event depth. In this study, we continue to require a free or resolved depth for an earthquake to be considered a GT event. At this stage, we do not discriminate GT events based on the size of the depth error, as the error in depth may be as much controlled by the deviation of the unknown local velocity structure from ak135, as by the station distribution geometries that are

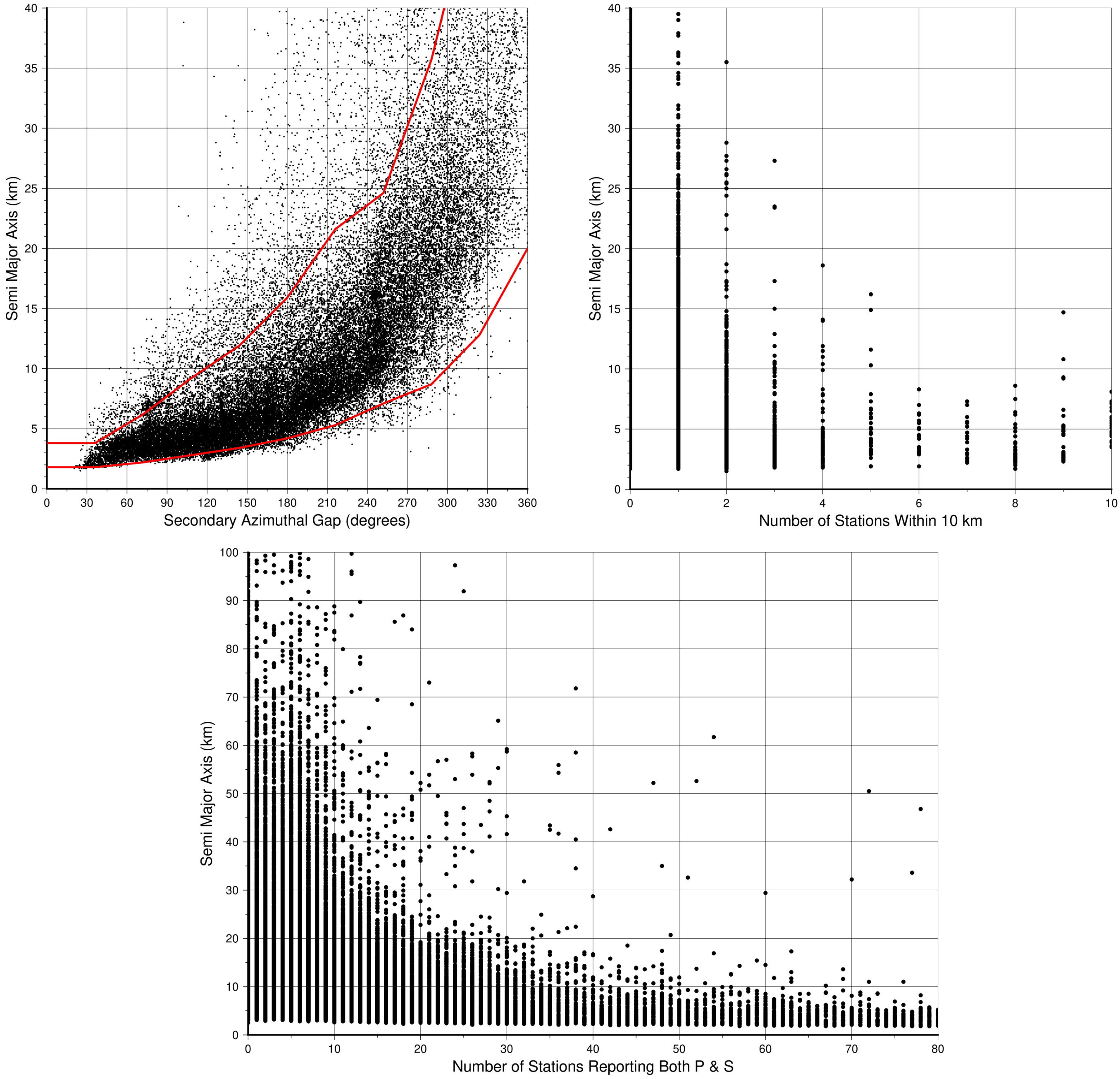


**Figure 7** Top left: Semi-major axis of the error ellipse with respect to secondary azimuthal gap. Top right: Semi-major axis of the error ellipse with respect to number of stations reporting P & S phase pairs. Bottom: Semi-major axis of the error ellipse with respect to number of stations within 10 km of the event.

primarily considered in this study. In addition, we currently have no reliable benchmark data set that is appropriate for testing depth resolution. The assertions made in this work concerning horizontal location errors are based on the well constrained locations of explosions, with sources located very close to the surface, making them ill-suited for testing depth resolution. We therefore consider only if the event depth can be resolved given the available phase data.

Bondár and McLaughlin (2009) introduced the requirement that GT candidate events have at least one defining station within 10 km of the event in order to add some constraint to the event depth. This requirement of at least one very close station is designed to limit potential trade-offs between the event depth and origin time, that is an established limiting factor in earthquake depth resolution. Another way of addressing the trade-off between event depth and origin time is by requiring S and P phases recorded at a local station. This has been shown to significantly reduce the trade-off between depth and origin time (Gomberg et al., 1990). Here we test whether using a defined number of pairs of P & S observations (i.e. where P and S phases are reported at the same station) can provide a comparable reduction in the trade-offs between earthquake origin time and depth to that from having a defining station within 10 km of the event.

To test the potential depth resolution of these constraints, we use the 47,161 natural event test dataset based on events from the 2018 – 2020 Reviewed ISC Bul-

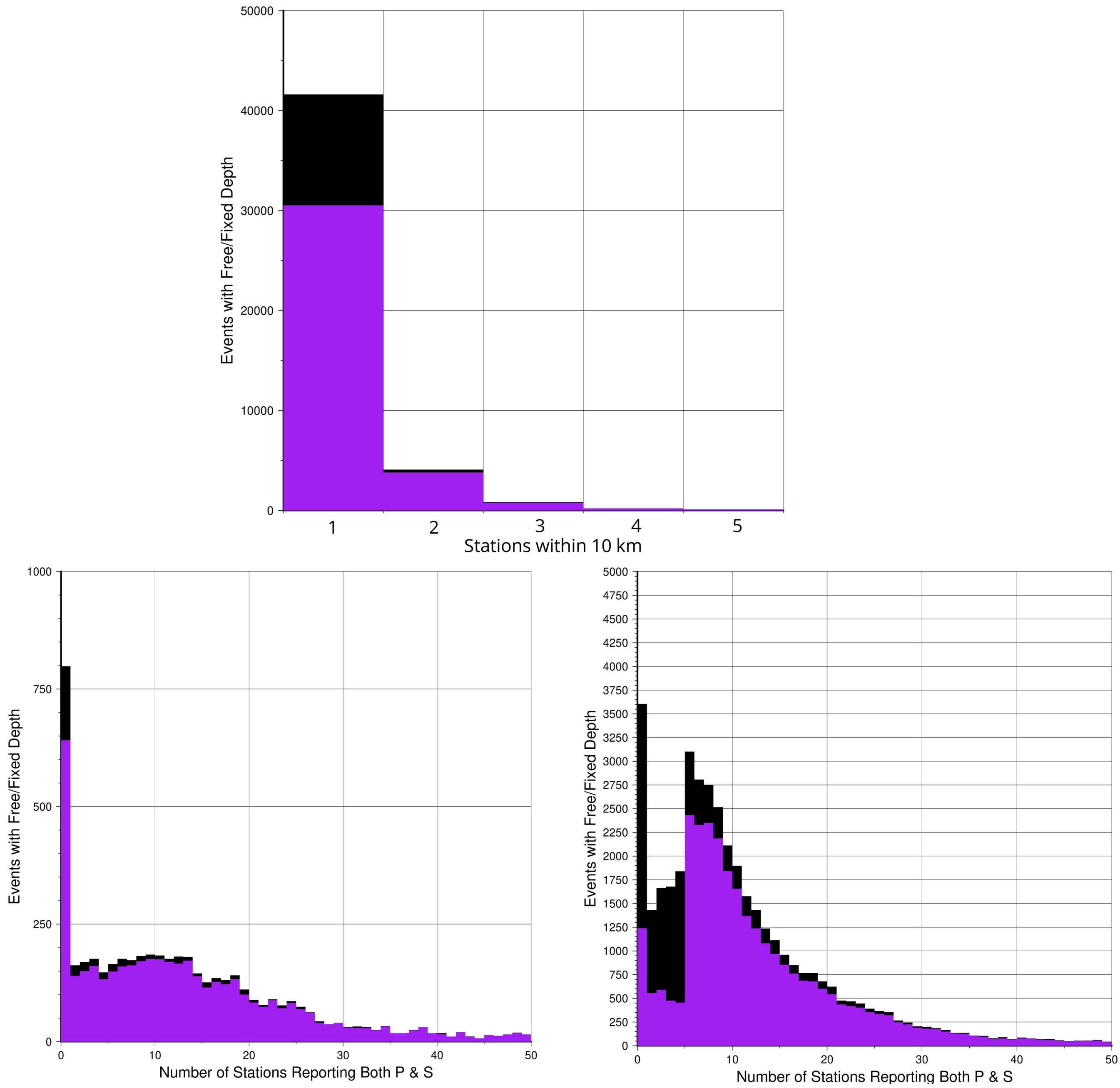


**Figure 8** Top: Histogram quantifying the variation in the number of resolved (purple) and unresolved event depths (black) for number of stations within 10 km of the event. Bottom left: Histogram as above, quantifying the number of resolved and unresolved depths varying with the number of stations reporting P & S phase pairs, where there is also one or more station within 10 km of the event. Bottom right: As above, but with a varying number of P & S phase pairs, and no stations within 10 km of the event.

letin that are well recorded locally. The top panel of Figure 8 illustrates the number of events where the depth is resolved for all events in the test dataset with respect to number of stations within 10 km. It is clear from this figure that having one or more stations within 10 km of the event significantly improves the potential depth is resolution. Without a station within 10 km of the event the depth is resolved for 73.3% of events, and with one or more stations within 10 km of the event the depth resolved in 91.7% of events. This confirms the proposal of Bondár and McLaughlin (2009) that having a station within 10 km of the event is a strong predictor of depth resolution.

The second potential criteria controlling depth resolution considered here is the number of stations within 150 km of the event reporting P & S phase pairs at the same station. The bottom left panel of Figure 8 demonstrates that events with one or more stations within 10 km, and one or more stations reporting P & S, 93.6% of events have a resolved depth. Without at least one station reporting P & S, 80.3% of events have a resolved depth. This reduction indicates that having at least one station reporting both P & S phases improves the ability to resolve the event depth.

The depth resolution provided by stations reporting P & S phases (S-P times) is also evident when looking

| Current Ground Truth Criteria | New Ground Truth Criteria |
|---|---|
| Event Magnitude <6.1 | Event Magnitude <6.1 |
| At least one station within 10 km of event | At least one station within 10 km of event **or** five or more stations reporting both P and S phases |
| Secondary Azimuthal Gap <160° | Secondary Azimuthal Gap $\leq$210° |
| Balanced station distribution, ΔU <0.36 | CPQ $\geq$0.4 |
| Event recorded at a station at a distance of at least 2° | Event recorded at a station at a distance of at least 2° |
| *Semi-major axis of the error ellipse $\leq$5 km* | *Semi-major axis of the error ellipse $\leq$5 km* |
| *Earthquake depth is not fixed to a set value* | *Earthquake depth is not fixed to a set value* |
| | *Earthquake depth is $\leq$35 km* |

**Table 2** Table comparing the current GT criteria (Bondár and McLaughlin, 2009) with the newly proposed GT criteria. Criteria in italics are applied after relocation, other criteria are applied before and after relocation.

at events where there are no stations within 10 km of the event. For these events, when there are no stations reporting P & S phases only and no stations within 10 km only 34.4% of events have depth resolution. Where there are no stations within 10 km and five or more stations reporting both P & S phases, 86.6% of events have a resolved depth.

Together these results show that with either a station within 10 km (80.3% solved for depth) or five or more stations reporting both P & S phases (86.6% solved for depth) there is a high likelihood of resolving the depth of an event. We therefore propose amending the GT criteria to allow depth resolution to be added by either approach, requiring either one or more stations within 10 km of an event or five or more stations within 150 km reporting both P & S phases or both. For the test dataset this would result in 36,948 events (78.3%) passing the depth criteria. These criteria, along with the requirement that the event depth is resolved by the relocation, would be necessary for adding an event to the GT list. In total 32,288 of 47,161 events (68.5%) from the ISC Bulletin test dataset would pass both the revised depth criteria, and provide a free depth solution after relocation.

## 4 GT Candidate and Certification Criteria

Based on the above work we propose the following criteria for identifying GT candidate events:

- Five or more stations within 150 km of the event
- CPQ greater than or equal to 0.4
- Secondary Azimuthal Gap less than or equal to 210°
- One or more stations within 10 km of the event OR five or more stations reporting both P & S phases
- The event is recorded at distances greater than or equal to 2°

The final criteria outlined above has not been discussed in detail, but effectively acts as a magnitude filter, removing smaller events that meet the other GT criteria, but are of limited use as they are not recorded at adequate distance from the candidate event. This criteria is chosen to select events of the most use for travel-time calibration while maintaining an acceptable workload for event processing.

From the ISC Bulletin test dataset (2018 – 2020) the above criteria would result in 20,921 events (corresponding to 44.3% of events with five or more stations within 150 km) being identified as GT candidates and being relocated to test for GT event certification. After relocation we propose the following criteria for certification of GT events:

- Five or more stations within 150 km of the event
- CPQ greater than or equal to 0.4
- Secondary Azimuthal Gap less than or equal to 210°
- One or more stations within 10 km of the event OR five or more stations reporting both P & S phases OR both conditions are met
- Recorded at distances greater than or equal to 2°
- Semi-major axis of the error ellipse less than or equal to 5 km
- The event depth is resolved during relocation, and is not fixed to a default reference depth
- The resolved event depth is 35 km or shallower

The final criteria which has been added excludes earthquakes occurring deeper than 35 km, due to the inherent uncertainty in the velocity structure of regions where deeper earthquakes occur. A comparison between the new set of GT criteria and the existing GT criteria is shown in Table 2.

From the test dataset this would result in 8,742 events (i.e. 18.5% of the initial 47,161 test data set) qualifying for the GT dataset. Figure 9 shows the geographic distribution of these events along with the 1,799 GT events that would have been identified with the previous criteria of Bondár and McLaughlin (2009). This demonstrates that the proposed revisions to the GT criteria outlined here significantly increase the overall number of GT events, as well as extending the coverage of the GT catalogue to geographic regions that were previously unsampled.

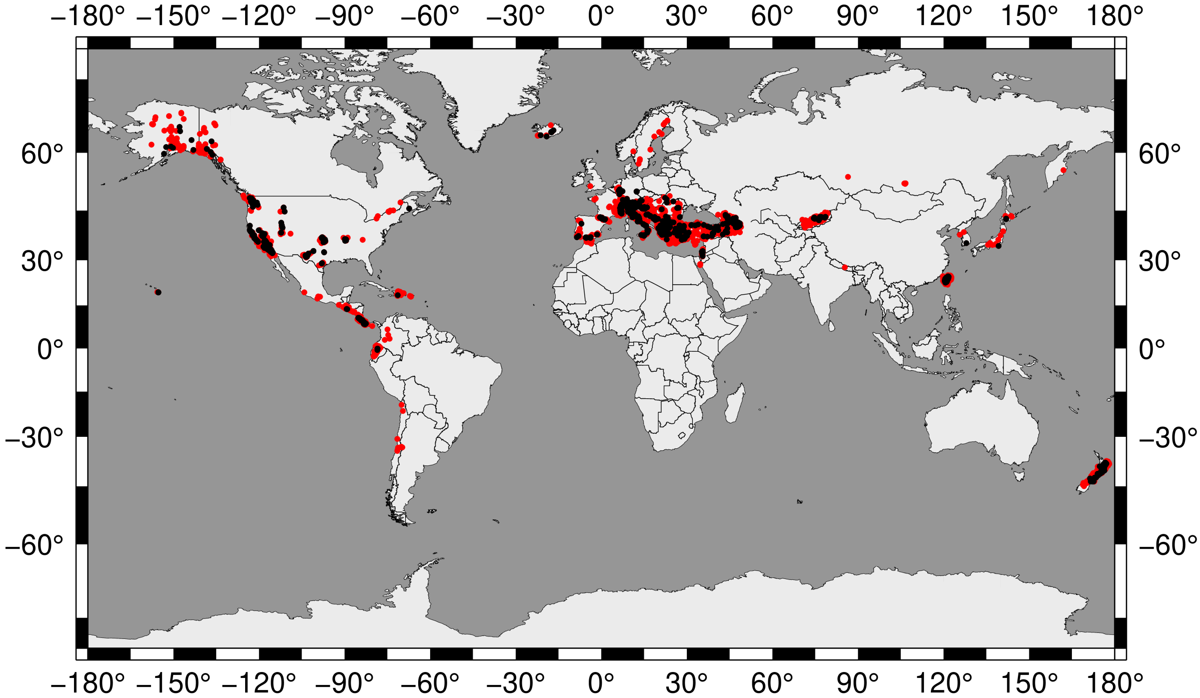


**Figure 9** Map of events meeting current GT criteria of Bondár and McLaughlin (2009) (black circles) and the new GT criteria proposed here (red circles) for events occurring in the years 2018–2020.

## 5 Conclusion

We have shown that with a modern location algorithm, such as ISCloc, unbalanced station distributions no longer significantly bias event location. This removes a key rationale for the implementation of the ΔU criteria. We therefore propose to replace ΔU with a new metric for azimuthal coverage, CPQ, which is a more appropriate measure of station azimuthal coverage and will always increase when stations are added. For location algorithms that are not able to account for unbalanced station distributions, the ΔU parameter remains an important constraint that can still be considered when identifying well located events.

To identify GT candidate events, we use a minimum criteria of CPQ (> 0.4), secondary azimuthal gap (< 210°) and require five or more stations within 150 km, to remove events where it is unreasonable to expect a GT event after relocation. Similarly, we use a minimum criteria of one or more stations within 10 km of the event or five or more stations reporting both P & S phases, to remove events where it is unreasonable to expect a depth solution after relocation.

Using 136,011 iterations of random station selections for 133 explosion events, we have demonstrated that the semi-major axis of the error ellipse (90% confidence) after relocation is normally (~93% of iterations) greater than the mislocation. We therefore propose using the semi-major axis of the error ellipse and that event depth is resolved (not fixed), as the two main discriminants for selecting GT events after relocation, in addition to reapplying the candidate event selection criteria to the relocated events. We have demonstrated that this new set of criteria will allow for more GT events to be identified across a greater geographic extent without compromising the integrity of the event locations within the GT database. The new criteria have been tested and developed using the Reviewed ISC Bulletin, using ISCloc, a modern location algorithm that accounts for uneven station distributions.

Given this specificity, we note that these criteria may not directly apply to all seismic event bulletins and location techniques. The proposed update to GT identification criteria are reliant on measures of horizontal precision calculated by ISCloc during the relocation procedure. This reliance is justified, when investigating epicentral location, by the testing using the anthropogenic explosion dataset that allows benchmarking with the absolute mislocation of the event. We therefore have evaluated the true performance of ISCloc in relation to horizontal location accuracy, in the context of reasonable local seismic station coverage. For depth there is no similar dataset to perform this benchmarking, and the criteria proposed here are therefore designed to optimise the proportion of events where the depth can be resolved. Further work may focus on developing testable depth constraints which quantify the uncertainty in the calculated depth.

## Acknowledgements

The ISC is able to fulfil its mission thanks to the support of its 78 Member-Institutions in 49 countries (www.isc.ac.uk/members/, last access: February 2025), including the National Science Foundation (NSF) award EAR 2414178 and the Royal Society Grant INT004. The ISC is also grateful to the following organizations that help to fund the ISC development projects: CTBTO, FM Global, the Lighthill Risk Network, AXA XL, LLNL, NTR, USGS, SARA, GeoSIG, GaiaCode and MS&AD. We thank Johannes Schweitzer for helpful discussions. We also thank Anthony Lomax and Yifan Yu for their helpful comments and contributions that improved this manuscript.

## Data and code availability

All data used in this study are available through the ISC Bulletin which can be accessed through the ISC website: https://doi.org/10.31905/D808B830. The original IASPEI Reference Event Dataset can be accessed through the ISC website: https://doi.org/10.31905/32NSJF7V. The 2018-2020 dataset used to highlight the impact of the updated criteria is query-able on the ISC website at the following link: https://doi.org/10.31905/A1CMX9W7. A code for calculating the value of CPQ from a list of event station azimuths is provided online at the following link: https://github.com/Ryan-isc/cpq. The locator used in this study, ISCloc, is available for download from the ISC website: https://www.isc.ac.uk/iscbulletin/iscloc/.

## Competing interests

The authors have no competing interests.

## References

Begnaud, M. L., Myers, S. C., Young, B., Hipp, J. R., Dodge, D., and Phillips, W. S. Updates to the Regional Seismic Travel Time (RSTT) Model: 1. Tomography. *Pure and Applied Geophysics*, 178 (7):2475–2498, Nov. 2020. doi: 10.1007/s00024-020-02619-5.

Belinić, T. and Markušić, S. Empirical criteria for the accuracy of earthquake locations on the Croatian territory. *Geofizika*, 34(1): 1–17, 2017. doi: 10.15233/gfz.2017.34.5.

Bennett, T. J., Oancea, V., Barker, B. W., Kung, Y.-L., Bahavar, M., Kohl, B. C., Murphy, J. R., and Bondar, I. K. The Nuclear Explosion Database (NEDB): A New Database and Web Site for Accessing Nuclear Explosion Source Information and Waveforms. *Seismological Research Letters*, 81(1):12–25, Jan. 2010. doi: 10.1785/gssrl.81.1.12.

Bergman, E. A., Benz, H. M., Yeck, W. L., Karasözen, E., Engdahl, E. R., Ghods, A., Hayes, G. P., and Earle, P. S. A Global Catalog of Calibrated Earthquake Locations. *Seismological Research Letters*, 94(1):485–495, Oct. 2022. doi: 10.1785/0220220217.

Bittner, P., Le Bras, R., Mialle, P., and Nielsen, P. International Data Centre Bulletin Events Triggered by Controlled Underwater Explosions of World War 2 Ordnances. *Pure and Applied Geophysics*, 180(4):1303–1315, Dec. 2022. doi: 10.1007/s00024-022-03146-1.

Bondar, I. and McLaughlin, K. Seismic Location Bias and Uncertainty in the Presence of Correlated and Non-Gaussian Travel-Time Errors. *Bulletin of the Seismological Society of America*, 99 (1):172–193, Feb. 2009. doi: 10.1785/0120080922.

Bondár, I. and McLaughlin, K. L. A New Ground Truth Data Set For Seismic Studies. *Seismological Research Letters*, 80(3):465–472, May 2009. doi: 10.1785/gssrl.80.3.465.

Bondár, I. and Storchak, D. Improved location procedures at the International Seismological Centre. *Geophysical Journal International*, 186(3):1220–1244, July 2011. doi: 10.1111/j.1365-246x.2011.05107.x.

Bondár, I., Yang, X., North, R., and Romney, C. Location Calibration Data for CTBT Monitoring at the Prototype International Data Center. *Pure and Applied Geophysics*, 158(1):19–34, Feb. 2001. doi: 10.1007/pl00001155.

Bondár, I., Engdahl, E., Yang, X., Ghalib, H., Hofstetter, A., Kirichenko, V., Wagner, R., Gupta, I., Ekström, G., Bergman, E., Israelsson, H., and McLaughlin, K. Collection of a Reference Event Set for Regional and Teleseismic Location Calibration. *Bulletin of the Seismological Society of America*, 94(4):1528–1545, Aug. 2004a. doi: 10.1785/012003128.

Bondár, I., Myers, S. C., Engdahl, E. R., and Bergman, E. A. Epicentre accuracy based on seismic network criteria. *Geophysical Journal International*, 156(3):483–496, Mar. 2004b. doi: 10.1111/j.1365-246x.2004.02070.x.

Bondár, I., Bergman, E., Engdahl, E. R., Kohl, B., Kung, Y.-L., and McLaughlin, K. A hybrid multiple event location technique to obtain ground truth event locations. *Geophysical Journal International*, 175(1):185–201, Oct. 2008. doi: 10.1111/j.1365-246x.2008.03867.x.

Bondár, I., Godoladze, T., Cowgill, E., Yetirmishli, G., Myers, S. C., Gunia, I., Buzaladze, A., Czecze, B., Onur, T., Gök, R., and Chiang, A. Relocation of the Seismicity of the Caucasus Region. *Bulletin of the Seismological Society of America*, 114(2):857–872, Nov. 2023. doi: 10.1785/0120230155.

Boomer, K. B., Brazier, R. A., and Nyblade, A. A. Empirically Based Ground Truth Criteria for Seismic Events Recorded at Local Distances on Regional Networks with Application to Southern Africa. *Bulletin of the Seismological Society of America*, 100 (4):1785–1791, July 2010. doi: 10.1785/0120090237.

Braden, B. The Surveyor's Area Formula. *The College Mathematics Journal*, 17(4):326–337, Sept. 1986. doi: 10.1080/07468342.1986.11972974.

Gomberg, J. S., Shedlock, K. M., and Roecker, S. W. The effect of S-wave arrival times on the accuracy of hypocenter estimation. *Bulletin of the Seismological Society of America*, 80(6A): 1605–1628, Dec. 1990. doi: 10.1785/bssa08006a1605.

Kennett, B. L. N. and Engdahl, E. R. Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2):429–465, May 1991. doi: 10.1111/j.1365-246x.1991.tb06724.x.

Kennett, B. L. N., Engdahl, E. R., and Buland, R. Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1):108–124, July 1995. doi: 10.1111/j.1365-246x.1995.tb03540.x.

Sweeney, J. *Accuracy of teleseismic event locations in the Middle East and North Africa*. Dec. 1996. doi: 10.2172/514441.

Sweeney, J. *Criteria for selecting accurate event locations from NEIC and ISC bulletins*. Lawrence Livermore National Laboratory, 1998.

Zeng, X., Xie, J., and Ni, S. Ground Truth Location of Earthquakes by Use of Ambient Seismic Noise From a Sparse Seismic Network: A Case Study in Western Australia. *Pure and Applied Geophysics*, 172(6):1397–1407, Dec. 2014. doi: 10.1007/s00024-014-0993-6.

Zhu, C., Wang, C., Zhang, B., Qin, X., and Shan, X. Differential Interferometric Synthetic Aperture Radar data for more accurate earthquake catalogs. *Remote Sensing of Environment*, 266: 112690, Dec. 2021. doi: 10.1016/j.rse.2021.112690.