Geophysical Journal International

A benchmark case study for seismic event relative location

S.J. Gibbons [1,2] T. Kværna,[1] T. Tiira[3] and E. Kozlovskaya[4]

[1]*NORSAR, P.O. Box* 53, 2027 *Kjeller, Norway. E-mail:* steven.gibbons@ngi.no
[2]*NGI, Sognsveien* 72, 0855 *Oslo, Norway*
[3]*Institute of Seismology, Department of Geosciences and Geography, P.O. Box* 68, *FI*-00014 *University of Helsinki, Finland*
[4]*Oulu Mining School, Faculty of Technology, POB* 3000, *FIN*-90014 *University of Oulu, Finland*

SUMMARY

'Precision seismology' encompasses a set of methods which use differential measurements of time-delays to estimate the relative locations of earthquakes and explosions. Delay-times estimated from signal correlations often allow far more accurate estimates of one event location relative to another than is possible using classical hypocentre determination techniques. Many different algorithms and software implementations have been developed and different assumptions and procedures can often result in significant variability between different relative event location estimates. We present a Ground Truth (GT) dataset of 55 military surface explosions in northern Finland in 2007 that all took place within 300 m of each other. The explosions were recorded with a high signal-to-noise ratio to distances of about 2°, and the exceptional waveform similarity between the signals from the different explosions allows for accurate correlation-based time-delay measurements. With exact coordinates for the explosions, we are able to assess the fidelity of relative location estimates made using any location algorithm or implementation. Applying double-difference calculations using two different 1-D velocity models for the region results in hypocentre-to-hypocentre distances which are too short and it is clear that the wavefield leaving the source region is more complicated than predicted by the models. Using the GT event coordinates, we are able to measure the slowness vectors associated with each outgoing ray from the source region. We demonstrate that, had such corrections been available, a significant improvement in the relative location estimates would have resulted. In practice we would of course need to solve for event hypocentres and slowness corrections simultaneously, and significant work will be needed to upgrade relative location algorithms to accommodate uncertainty in the form of the outgoing wavefield. We present this data set, together with GT coordinates, raw waveforms for all events on six regional stations, and tables of time-delay measurements, as a reference benchmark by which relative location algorithms and software can be evaluated.

Key words: Time-series analysis; Controlled source seismology; Earthquake monitoring and test-ban treaty verification.

1 INTRODUCTION

Closely spaced or clustered seismic events can often be located with far greater accuracy than they could be in isolation. An important reason for this is that the bias in the modelled traveltime to a given station is essentially the same for the different events in the cluster. This can be exploited by multiple-event location algorithms to reduce bias in the resulting location estimates (e.g. Douglas 1967; Richards-Dinger & Shearer 2000; Myers *et al.* 2009; Anderson & Myers 2010). The so-called double-difference location algorithms can reduce another source of error: uncertainty in the signal arrival-time estimate. Very closely spaced events may show sufficiently

high waveform similarity for cross-correlation or similar methods to be able to measure relative arrival-times far more accurately than the absolute onsets can be measured (e.g. Waldhauser & Ellsworth 2000; Richards *et al.* 2006). The family of algorithms that exploit full-waveform methods to estimate accurate relative arrival-times are referred to collectively as *precision seismology*.

There are many situations in which it is desirable to increase the accuracy of seismic event location estimates. Even without the application of cross-correlation derived relative onset-times, the delineation of natural seismicity patterns and fault structures can be sharpened greatly using contextual location algorithms (Gibbons & Kværna 2017; Nooshiri *et al.* 2017; Jerkins *et al.* 2020). Even

greater improvement is possible using accurately measured relative times (e.g. Lin *et al.* 2007). The time-evolution and operational significance of induced seismicity can also be revealed using precision seismology (e.g. Goertz-Allmann *et al.* 2017; Yu *et al.* 2019). In forensic seismology, accurate relative location estimates of underground nuclear tests (e.g. Fisk 2002; Waldhauser *et al.* 2004) can place constraints on emplacement and depth with possible consequences for yield estimation (e.g. Pasyanos & Myers 2018; Voytan *et al.* 2019).

There are sources of error in relative event location which are often overlooked by practitioners. First is the assumption that the time of maximum correlation between two waveforms is an accurate indicator of the time separating two equivalent segments of the wave trains from different events. Simply displaying the signals from the 3 September 2007 North Korean underground nuclear test together with the signals from the significantly smaller previous explosions at the same site made it clear that a correlation-based alignment of waveform features following the signal onset could result in qualitatively misleading estimates of the relative onset-time (Gibbons *et al.* 2018b). The limitations of such estimates were discussed in greater depth by Bachura & Fischer (2019) for earthquakes in Czechia. Secondly is the assumed mapping between a relative time-delay and an interevent distance. Although a bias in the traveltime model may be cancelled out by the simultaneous location of two events, there is an underlying assumption of how far a *P* wave, *S* wave, or surface wave has travelled in a given time. Selby (2010) and Wen & Long (2010) had both calculated the location of the 25 May 2009 North Korean test relative to the 9 October 2006 test and found rather different locations from different sets of measurements: one teleseismic and one regional. Gibbons *et al.* (2017b) argued that the outgoing seismic wavefield from the test-site must be rather more complicated than could be explained using a 1-D velocity model and estimated an azimuthally dependent set of slowness corrections which made the regional and teleseismic estimates more internally consistent. The slowness corrections obtained appeared to vary approximately sinusoidally with azimuth, as would be observed given a dipping layer beneath the test-site. However, we cannot confirm a dipping layer hypothesis on the basis of these observations since the formulation of Gibbons *et al.* (2017b) only allowed the slowness, and not the backazimuth, to vary.

There are few situations in which the results of double-difference location estimates can be confirmed absolutely. Cross-validation studies can be performed by using different subsets of measurements (e.g. Murphy *et al.* 2013; Gibbons *et al.* 2017b) or, for example, body wave relative location estimates can be compared with surface wave relative location estimates (e.g. Ichinose *et al.* 2017). However, ultimately, assumptions about the propagation of seismic waves will have been made. In this paper we present a set of 55 surface explosions carried out at Hukkakero in Finland in August and September 2007. The explosions all take place within 300 m of one another and all were well recorded by many seismic stations within a distance of 2°. The exact coordinates of the explosions were provided by the Finnish military, allowing for an absolute check on quality of double-difference location estimates.

We select a set of six seismic stations: three permanent and three temporary. The temporary stations were part of the LAP-NET/POLENET deployment as part of the International Polar Year in 2007 and 2008 (e.g. Kozlovskaya *et al.* 2016; Usoltseva & Kozlovskaya 2016; Kozlovskaya 2007). Raw seismic waveforms from these stations are provided in miniSEED format for all events in the Supporting Information. In addition, cross-correlation calculations have been performed for both *P* and *S* arrivals on all six stations for every pair of events. The results of these calculations, with delay-time estimates and correlation coefficient values, are provided in a single ASCII file in the Supporting Information. Details of the Supporting Information are provided in the Appendix. With the raw waveforms, readers are encouraged to test relative event location algorithms and software through measurement of the delay times and subsequent location estimates. We describe the explosions and the corresponding set of seismic signals and discuss some of the issues involved in our attempts to locate the events relative to each other.

2 THE GROUND-TRUTH DATA SET

For many decades, the Finnish military have disposed of expired ammunition at a site in Central Lapland (northern Finland) in a sequence of controlled surface explosions between August and September every year. These events, referred to as the Hukkakero explosions, have received a lot of attention due to the fact that they generate strong infrasound signals recorded at distances up to several thousand kilometres (e.g. Gibbons *et al.* 2018a). It was demonstrated by Gibbons *et al.* (2007) that the seismic signals generated were essentially identical from explosion to explosion, meaning that the significant differences in the infrasound signals from event to event were due entirely to changes in the atmospheric conditions. Accurate measurements of backazimuth and phase velocity could be made of the infrasonic signals at distances of up to several hundred kilometres and Blixt *et al.* (2019) demonstrated that the azimuthal deviation in these measurements could be used to validate cross-wind predictions from atmospheric models.

Whereas the events have been of greatest interest for the infrasound signals, the repeatability of the seismic signals has been exploited for studies of the stability of direction estimates using *f–k* analysis on seismic arrays. Gibbons *et al.* (2010) measured the direction of arrival (DOA) of the initial *P*-wave arrival at the ARCES seismic array for many Hukkakero events in different frequency bands. It could be demonstrated that the frequency band in which the most stable backazimuth estimates were made was the 2–4 Hz band. In bands of increasing dominant frequency, the signal-to-noise ratio (SNR) improved while the stability of backazimuth estimates became worse. This was the exact opposite to what was observed for several other more distant sites, for which both the stability and the SNR were optimal at a higher frequency. It should be noted that at ARCES, the Pg, Pb and Pn phases (see Storchak *et al.* 2003) from Hukkakero events arrive essentially simultaneously. These phases are associated with slightly different slowness vectors and dominant frequencies and this may also have consequences for the relative event location procedure.

Most Hukkakero events have an explosive yield of between 15 and 30 tonnes TNT equivalent, although a number of far lower yield explosions have also been associated with the same site (Gibbons *et al.* 2018a). The explosives are raised above ground level and are not well-coupled to the bedrock. This limits damage to ground, such that the same sites can be reused repeatedly. This is the reason that so much of the energy in the explosions is converted to infrasound, with the seismic source resulting from the downward shock wave. Table 1 gives the coordinates of 55 explosions from 2007 as provided by the Finnish military. We do not have exact values for the elevation at each location but differences in elevation from site to site are small compared with the lateral distances. Entering the coordinates into the satellite map system https://www.bing.com/maps (last accessed 16 June 2020) allows the reader to identify visible ground

Table 1. Ground Truth coordinates for 55 explosions at Hukkakero between 15 August and 16 September 2007. Note that the sources have dimensions of the order of many metres such that there are only likely four significant figures in the latitude and longitude for the effective seismic source. mm-dd indicates the month and day-of-month in 2007, and hh.mm gives the hour and minute of the origin time, UT.

mm-dd	evID	hh.mm	lat.	lon.	evID	hh.mm	lat.	lon.
08-15	H01	08.00	67.93590	25.83491	H02	12.00	67.93354	25.83291
08-16	H03	08.00	67.93580	25.83511	H04	12.00	67.93352	25.83229
08-17	H05	07.30	67.93577	25.83428	H06	11.00	67.93351	25.83215
08-18	H07	07.30	67.93490	25.82991	H08	11.00	67.93352	25.83167
08-19	H09	08.00	67.93471	25.82915	H10	11.30	67.93344	25.83108
08-20	H11	07.30	67.93588	25.83496	H12	11.00	67.93362	25.83100
08-21	H13	07.30	67.93573	25.83528	H14	10.30	67.93354	25.83079
08-22	H15	07.30	67.93487	25.82938	H16	10.30	67.93355	25.83246
08-23	H17	07.30	67.93499	25.82933	H18	10.30	67.93370	25.83242
08-24	H19	07.30	67.93510	25.82961	H20	11.00	67.93344	25.83229
08-25	H21	07.30	67.93486	25.82948	H22	10.30	67.93367	25.83288
08-26	H23	08.45	67.93385	25.83475	H24	11.30	67.93597	25.83472
08-27	H25	07.30	67.93373	25.83438	H26	10.30	67.93607	25.83409
08-28	H27	07.30	67.93364	25.83402	H28	10.30	67.93590	25.83355
08-29	H29	07.30	67.93402	25.83551	H30	10.30	67.93363	25.83374
08-30	H31	07.30	67.93428	25.83612	H32	10.30	67.93367	25.83357
08-31	H33	07.30	67.93437	25.83597	H34	10.30	67.93359	25.83391
09-01	H35	07.30	67.93438	25.83606	H36	10.30	67.93363	25.83391
09-02	H37	08.15	67.93410	25.83538	H38	10.45	67.93337	25.83239
09-03	H39	07.15	67.93416	25.83586	H40	10.00	67.93336	25.83216
09-04	H41	07.30	67.93433	25.83564	H42	10.00	67.93340	25.83189
09-05	H43	07.30	67.93414	25.83507	H44	10.00	67.93351	25.83186
09-06	H45	10.30	67.93363	25.83200				
09-07	H46	10.00	67.93375	25.83185				
09-08	H47	09.30	67.93365	25.83159				
09-09	H48	09.30	67.93368	25.83214				
09-10	H49	09.00	67.93371	25.83271				
09-11	H50	09.00	67.93358	25.83155				
09-12	H51	10.00	67.93376	25.83295				
09-13	H52	09.30	67.93375	25.83156				
09-14	H53	09.30	67.93374	25.83280				
09-15	H54	09.00	67.93361	25.83202				
09-16	H55	09.30	67.93373	25.83233				

features associated with each of the explosions. The explosions are carried out within seconds of the full-hour, half-hour, or quarter-hour, and these approximate times are provided in Table 1. Between 15 August 2007 and 5 September 2007, two explosions were carried out on each day at different parts of the Hukkakero site. Between 6 September 2007 and 16 September 2007, only a single explosion was carried out each day.

Fig. 1 shows the location of the explosion site in relation to the six selected seismic stations. Although other stations were and are operating in this region, we deemed it desirable to present a realistically sparse regional network. The azimuthal coverage of the source region is good, although the distances range from about 0.5° (with the crustal Pg-type phases anticipated) to about 2.0° (with the Moho and uppermost mantle Pn-type phases anticipated). We saw little value in including stations at greater distances as these are often unlikely to record events of magnitudes of interest when monitoring induced seismicity or earthquake swarms. We also avoided adding more stations at short distances as this may give the impression of applicability only in the case of far denser networks than are likely to be found in regions of interest. LP34, LP53, and LP61 were among the first stations to be deployed in the temporary LAPNET deployment. All three components of each station are displayed in Fig. 2 for event H19, which took place at 07:30 UT on 24 August 2007. The stations are ordered by epicentral distance and

it is possible to see the waveform morphology at each of the stations. At the shortest distance (LP53 at 59 km), only 7 s of *P* arrival and coda are available before the *S*-wave arrival. At ARCES and KEV, over 20 s of *P*-wave energy is observed before the *S* arrival. The length of window to be used for time-delay calculations in double-difference locations is a theme which needs careful consideration. The desire for a long time-bandwidth product for an accurate delay-time measurement must be balanced against the understanding that the phase velocity of the coda waves decreases with time after the signal onset. Ideally, only a few cycles at the start of the signal would be used for the correlation calculation. In practice, a somewhat longer window is likely to make a clearer and less ambiguous peak in the cross-correlation trace. All calculations performed here used a template with length 6.0 s and a frequency band 2–5 Hz.

Fig. 3(a) displays the locations of the individual explosions. They take place in three or four clusters in a near ring with a diameter of between 250 and 300 m. Since the exact origin times of the explosions were not available in the ground truth (GT) information provided by the Finnish military, a correlation detector (using the formulation of Gibbons & Ringdal 2006) was run. A full-waveform signal for event H04 (2007-08-16T12.00.00) on the ARCES array was selected as a template and correlated against continuous ARCES array data for all of August and September 2007. This was done to confirm the presence of explosions at all the times
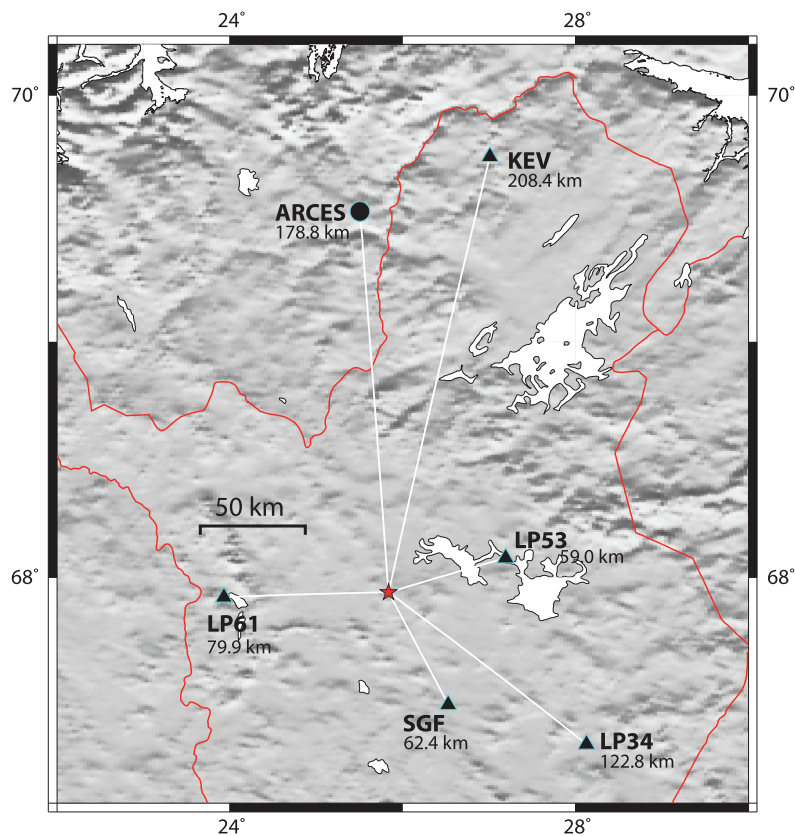
Figure 1. Location of the Hukkakero explosion site in nothern Finland in relation to the six seismic stations used in this relative location study. The coordinates of the stations are 69.5349N, 25.5058E, 403 m (ARCES), 69.7553N, 27.0067E, 81 m (KEV), 67.4421N, 26.5261E, 180 m (SGF), 67.2657N, 28.1253E, 208 m (LP34), 68.0843N, 27.1888E, 285 m (LP53), and 67.9141N, 23.9322E, 321 m (LP61).

indicated, at only the times indicated, and also to refine the origin time estimates. Note that the origin time estimates obtained from this detector assume that all events are exactly colocated and so will be inaccurate by the short differences in traveltime over the extent of the explosion site. In addition to the time of maximum correlation for each detection, the detector writes out the maximum value of the correlation coefficient at this time. The location symbols in Fig. 3 are coloured according to these values, providing a visual validation of the coordinates provided. The highest values of the correlation coefficients obtained are for the explosions set off closest to the Reference Event. The full waveform correlation values diminish with distance from this location. Fig. 3(b) displays waveforms at ARCES from four events as indicated from different clusters. It is clear how the waveform similarity is higher for events in the same cluster than for events in different clusters.

3 SEISMIC RELATIVE LOCATION ESTIMATES

As a first step to attempting a relative location calculation for these events, manual arrival-time estimates were made for both first *P* and first *S* arrivals for a single event (H04) on each of the available stations: 12 onset-time estimates in total. For each station, the *P*-arrival time was estimated on the vertical component trace filtered into many different frequency bands. The corresponding *S*-wave arrival time was made on the transverse rotation of the horizontal component traces (i.e. the ground motion perpendicular to the great-circle path connecting source and station), also observed in many different

bands simultaneously. For the ARCES array, a beam was formed for both *P* and *S* arrivals using steering-parameters chosen to optimize the SNR for each arrival. These manually picked times for event H04 were used as the starting times for provisional waveform templates, and a set of reference times was obtained for the same phases for all other events using a simple time-domain correlation detector on the appropriate waveform segments. All waveforms were upsampled to 1000 Hz in order to perform the time-domain cross-correlation calculation; this is not necessary for signal detection but mitigates a source of uncertainty in the time-delay estimates. The upsampling was performed using a time-domain sinc-interpolation (Schanze 1995). With provisional reference *P* and *S* times for all events on all stations, an exhaustive all-versus-all correlation calculation was performed in order to measure all necessary relative time delays. All of these correlation results are provided in the file `absolute_CC_times.txt` in the Supporting Information.

We follow the procedure described by Gibbons *et al.* (2017b) for calculating relative location estimates. This algorithm considers only the differences between pairs of arrival-time differences, eliminating the need to solve for the origin time. In addition to the time-delay measurement for each phase, we need a model prediction for the traveltime for that phase from any given source location to the station. For the initial calculation, we use two different 1-D models: the AK135 global model (Kennett *et al.* 1995) and the Fescan, or Fennoscandian, model (Mykkeltveit & Ringdal 1981).

The depth is held fixed in all calculations performed here. While there is no difficulty in scanning hypocentres at different depths, the traveltimes to stations at regional distances provide very little vertical resolution (see also Pasyanos & Myers 2018). Moving an
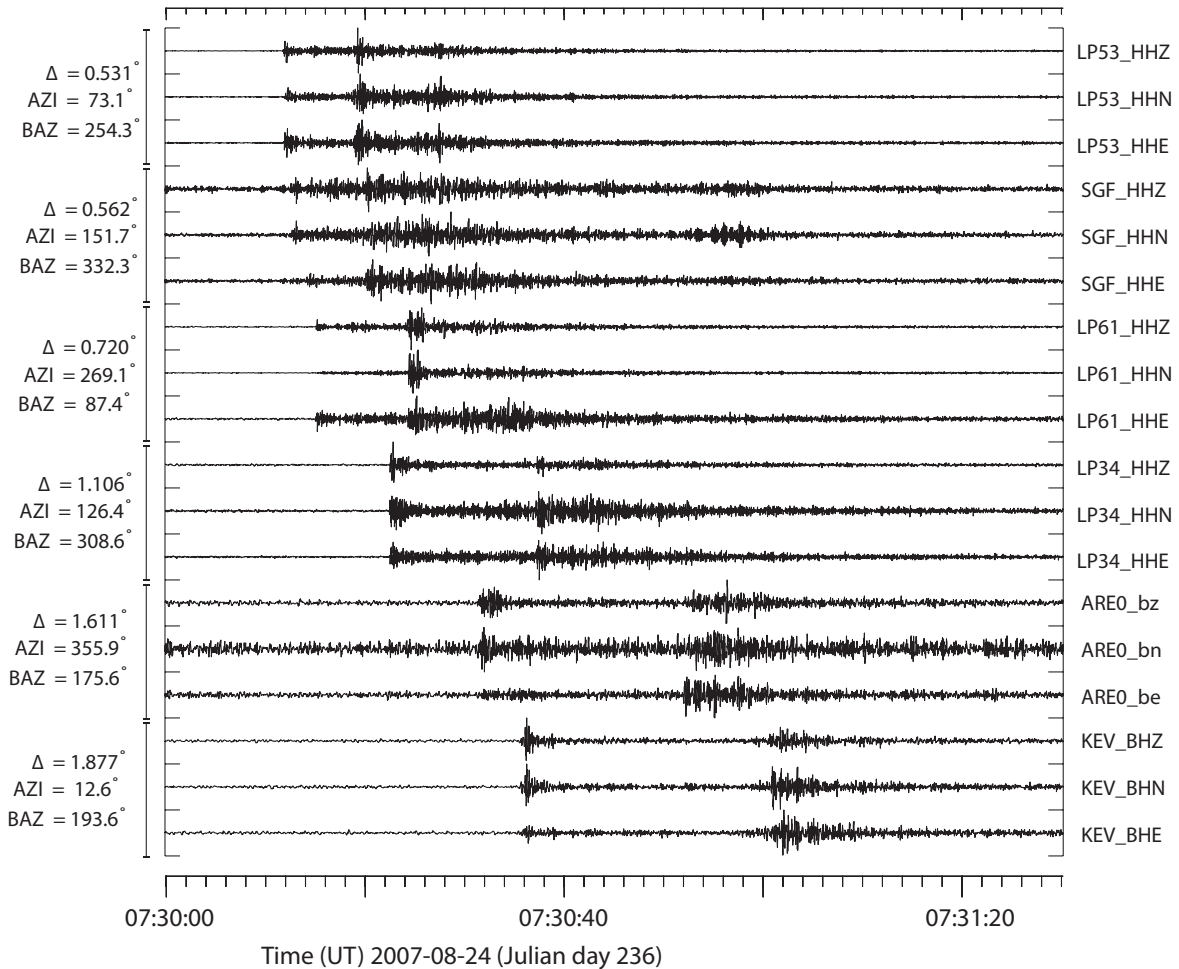
Figure 2. Seismic waveforms, bandpass filtered 2–8 Hz, from the six stations displayed in Fig. 1 for the 24 August 2007, 07:30 Hukkakero blast. The station ARE0 is at the reference site for the ARCES seismic array.

event hypocentre laterally in one direction decreases the travel time to stations in that direction and increases the travel times to stations in the opposite direction; moving an event hypocentre vertically increases or decreases the travel times to all stations approximately equally. In this study, where all explosions are carried out on the surface, the fixed depth assumption is approximately correct. This may not be the case for natural or induced seismicity, although a fixed depth assumption is still recommended unless data is available which actually provides resolution in depth differences.

Fig. 4 displays contour plots of the double-difference misfit residual norms for six different calculations with different velocity models and sets of measurements used. In each panel, the fixed location of the reference event (H03) is displayed and the (assumed unknown) location of event H02 is assumed to be at the location for which the double-difference misfit residual norm is at a minimum. In each of the calculations, the residual minimum (indicated by the red square) coincides quite well with the GT location (indicated by the white circle). The results for the two different models, AK135 in panels (a), (c) and (e) and fescan in panels (b), (d) and (f), are very similar. While the crustal and Moho velocities in fescan are rather faster than in AK135, the favourable azimuthal coverage means that the effect of a shorter or longer distance indicated in one direction is approximately balanced by a similar effect in the opposite direction. One of the most important visual lessons to be learned from Fig. 4 is the spatial resolution for different combinations of phases.

The residual contour lines for the (slower) *S*-phases (panels e and f) are denser than for the (faster) *P*-phases (panels c and d). The calculations with both *P*- and *S*-phases (panels a and b) show a density of the contour lines somewhere between those for the *P*-only and *S*-only calculations. The temptation is to claim that it would be best to use *S*-phases in isolation for improved spatial resolution. In practice, the *S*-phase may be poorly observed or competing with significant *P*-wave coda energy. While the location of the residual minimum is approximately in the correct location, the uncertainty is significant relative to the interevent distance.

We can perform the same calculation displayed in Fig. 4 with all of the events in the GT dataset. That is to say that we hold the location of event H03 fixed and attempt to locate the other 54 events relative using the misfit norm minimization technique. Fig. 5 displays the relative location estimates together with the events' GT locations and the corresponding mislocation vectors. For all phase combinations and both 1-D velocity models, the length of the mislocation vectors varies from close to zero to over 50 m (over 20 per cent of the true interevent distances). The mislocation vectors are encouragingly consistent for neighbouring events, providing confidence that the mislocations are due to fundamental model assumptions about the outgoing wavefield rather than, for example errors in measurement of the time-delays. The mislocation vectors for the northernmost cluster (in which the reference event H03 is located) are essentially zero, give or take noise within the region of
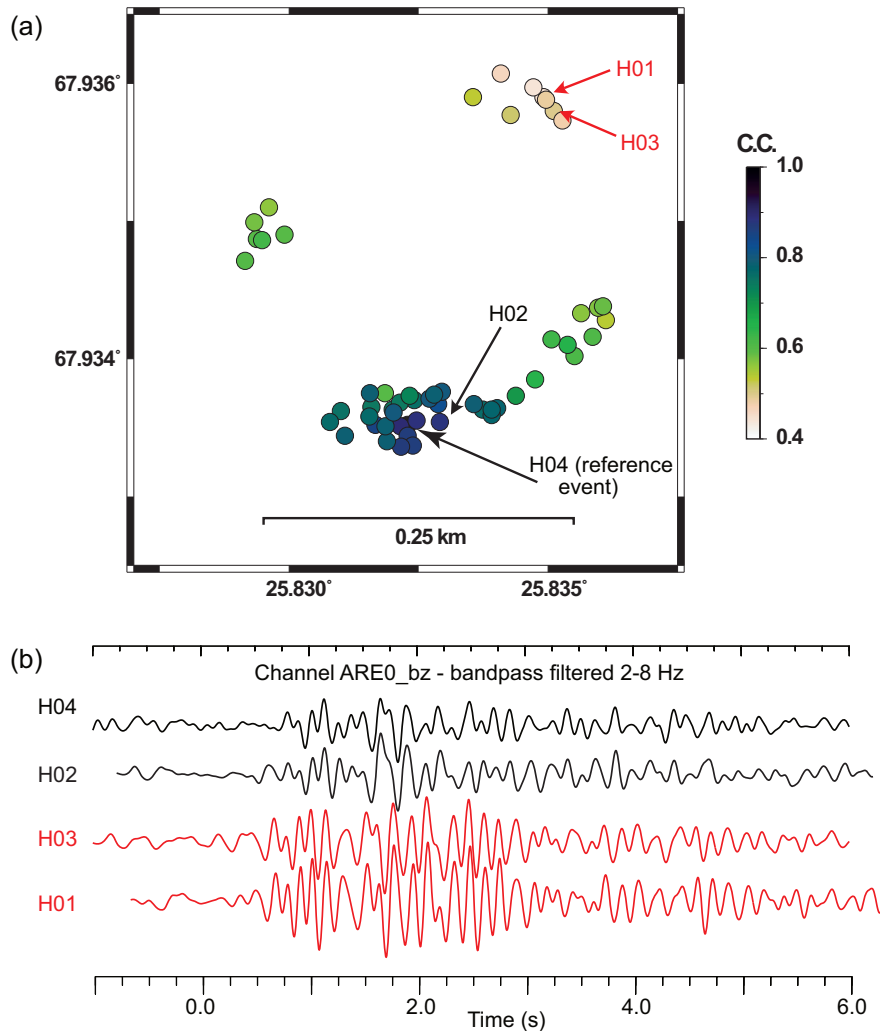
Figure 3. (a) Absolute locations of the 55 explosions listed in Table 1 coloured according to the fully normalized array correlation coefficient obtained using a correlation detector on the ARCES array with a signal template from event H04. The given locations of the explosions are consistent with visual features seen in satellite imagery using https://www.bing.com/maps with search coordinates 67.934°N, 25.830°E (last accessed 16 June 2020). (b) Bandpass filtered waveforms for events H01, H02, H03 and H04 recorded on sensor ARE0 bz. The locations of these four events are labelled on the map.

uncertainty. The spread here is likely due to measurement errors, possibly from low SNR signals. For the other clusters of explosions, the mislocation vectors generally point towards the reference event. This means that the interevent distances are underestimated, and that the velocity with which (at least some) phases have left the explosion site have been underestimated.

In this section we have demonstrated that, using the available waveforms and simple time-domain cross-correlation time-delay estimates, we have largely managed to reconstruct the large-scale structure of the event cluster using a simple residual minimization algorithm. While the spatial resolution appears coarse from the contour plots, the locations of the minimum misfit are quite stable. The relationship between the absolute levels of the misfit residual norms and size of the regions enclosed by each contour is a clue as to how stable the calculation would be to a given time-measurement error, for example due to a low SNR signal or a significant waveform dissimilarity introduced by a difference in the source. The consistency within clusters of the mislocation vectors is an indicator of robustness of the procedure itself and identifies assumptions about the velocity of the phases leaving the explosion site as the most probable source of error.

4 IMPROVING RELATIVE EVENT LOCATION ESTIMATES USING WAVEFIELD MEASUREMENTS

Gibbons *et al.* (2017b) sought the relative locations of the declared underground nuclear tests carried out under Mount Mantap in North Korea. They recognized that errors in the assumed velocities of the outgoing wave fronts were the likely source of discrepancy between earlier relative location estimates. An empirical algorithm was devised for estimating correction terms for each of the predicted traveltime differences. The correction terms obtained were subject to considerable uncertainty; the time-differences themselves are tiny fractions of a second and the correction terms estimate marginal adjustments to these delays. In addition, a dipping layer below the test-site would result in a deviation in both the apparent velocity and the direction of the outgoing waves; any azimuthal deviation in an outgoing wavefront was neglected. In spite of these difficulties, a set of corrections was obtained which greatly improved the consistency between estimates made using different sets of seismic phases. The obtained corrections showed a near sinusoidal azimuthal variability resembling that which would be observed given a dipping layer beneath the source.
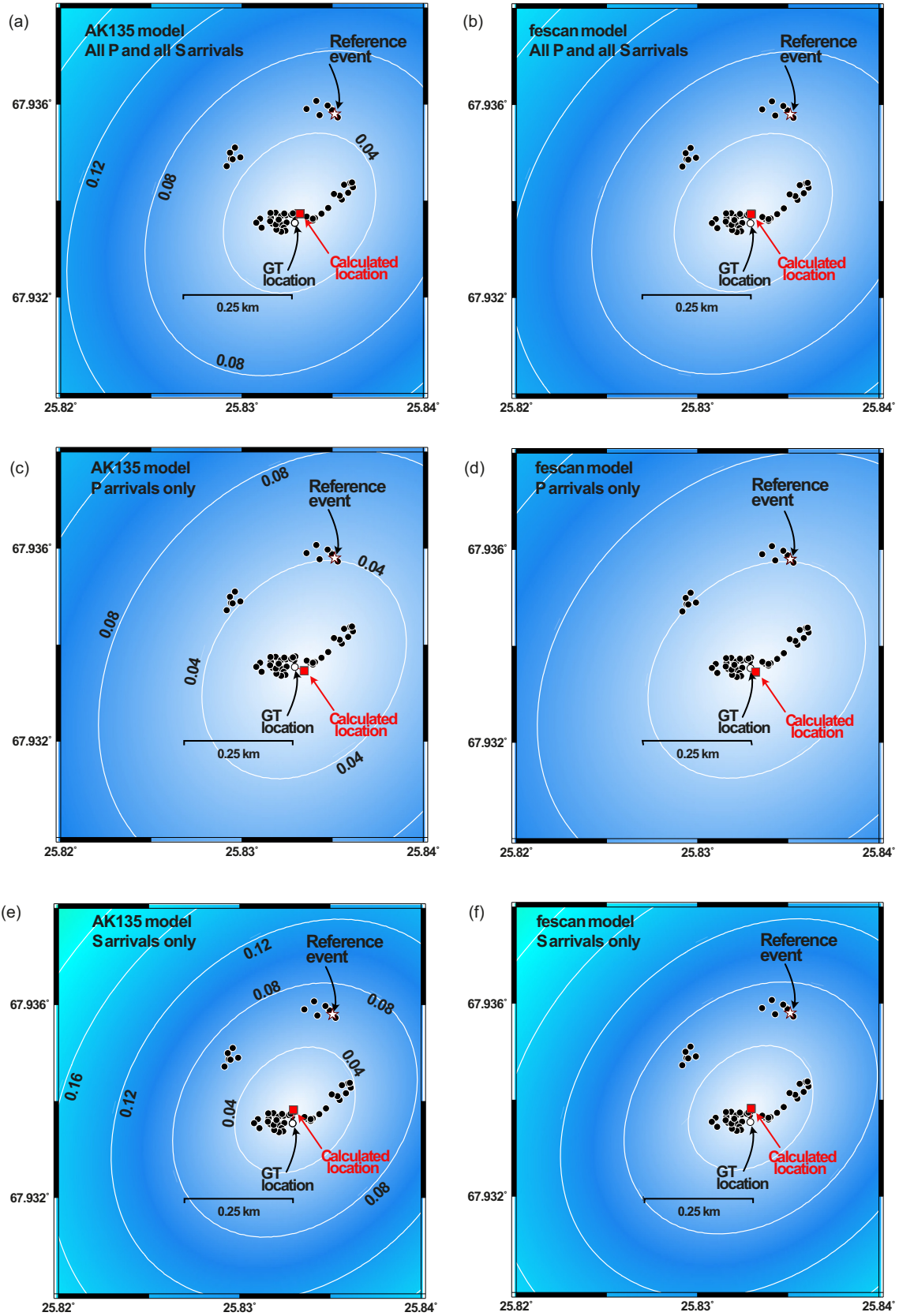
Figure 4. Double-difference location estimates for event H02 relative to event H03, which has fixed location denoted by a white star. The Ground Truth location of event H02 is indicated with a white circle and, for each panel, the red square indicates the location for which the double-difference misfit time-residual norm is a minimum. The panels differ in the velocity model applied and the subset of phases used as indicated. The contour lines are all at multiples of 0.04 s and are only labelled in the left-hand panels.
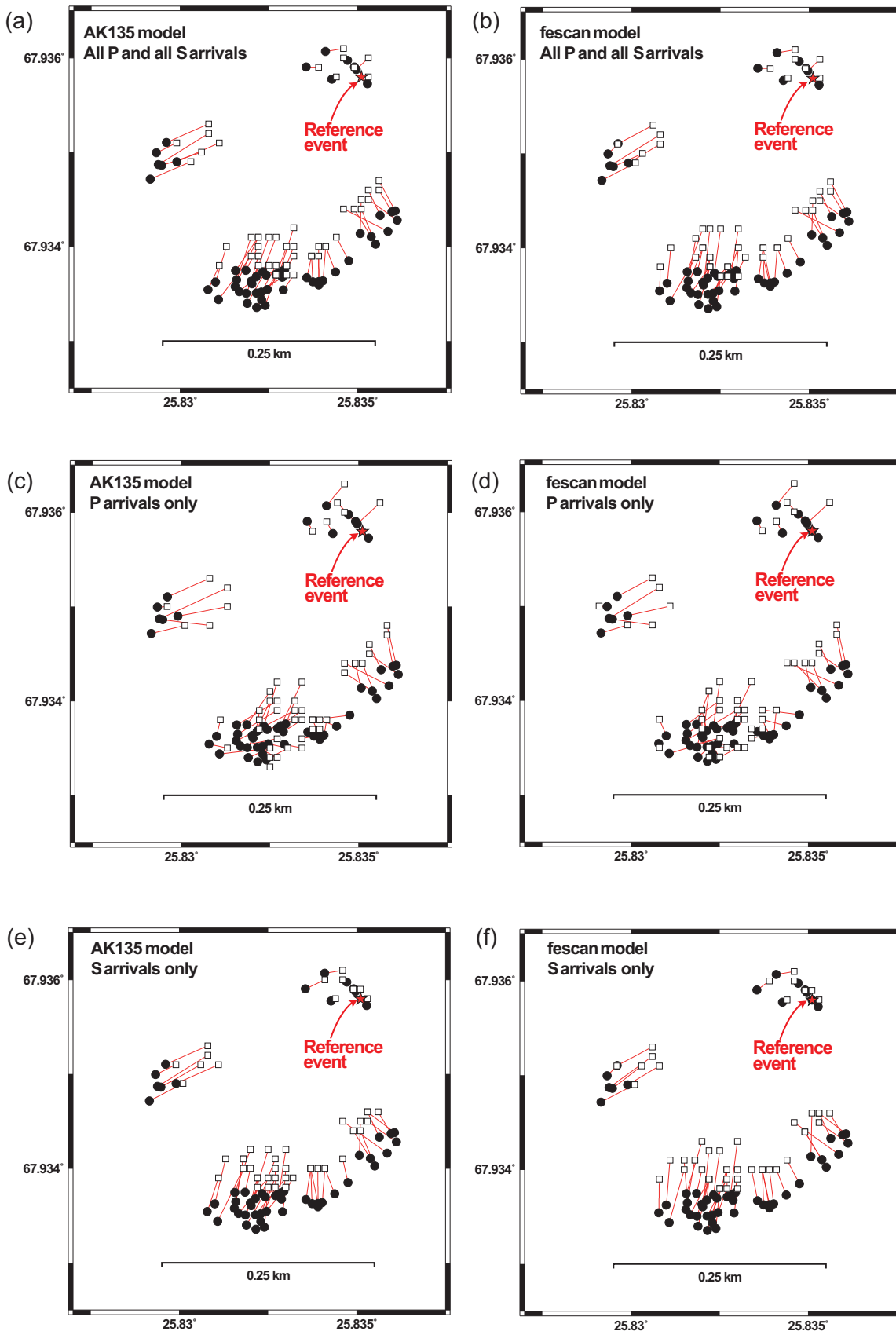
Figure 5. Location estimates for each of the explosions in Table 1 relative to the fixed event H03 (red star) using velocity models and phase subsets as indicated. For each event, the black circle indicates the Ground Truth location the white square at the end of the red line indicates the location for which the double-difference time residual norm was a minimum. Note that the scale of the maps is larger than in Fig. 4 in order to better display the mislocation vectors.

With this benchmark data set, unlike for the North Korean events, we have the exact coordinates of the sources. This may allow us to make direct measurements of the slowness vectors of the outgoing wavefield using so-called source-array analysis. Source-array analysis can be viewed as treating each event as a sensor observation at the coordinate of the event, where the data first are time-shifted to a common origin time. In this way, standard array processing methods can be applied. However, we have the disadvantage of not having exact origin times in the GT data and hence will have to solve for them. The initial origin time estimates we have from the correlation detectors are quite accurate, especially compared with say those obtained from classical event hypocentre estimation. However, they are based only on measurements on the ARCES array and will need to be adjusted for the short differences in the traveltime to ARCES based upon the locations within the explosion site.

The Bayesloc program (Myers *et al.* 2007) estimates joint probability distributions of hypocentre locations, origin times, arrival time uncertainty and corrections to modelled traveltimes for multiple events. If prior information is available, this can be applied to constrain the solution. In particular, we know the exact geographical hypocentres of these events and these can therefore be fixed in the input to Bayesloc. We have very accurate relative arrival time estimates from the cross-correlation calculations and the set of input files to run Bayesloc to solve for the event origin times is included in the Supporting Information. The Bayesloc solution uses traveltime tables as input and it could be argued that this introduces circularity, since the slowness vectors for the outgoing wavefield are implicit in the gradients of the traveltime tables. However, the ability of Bayesloc to compensate for deficiencies in traveltime predictions is exploited (e.g. Gibbons *et al.* 2017a) and we anticipate that this will mitigate any complications of circularity. To address the sensitivity of the origin time estimates to the traveltimes used, we ran the Bayesloc program using both the global AK135 velocity model and the regionally calibrated Fennoscandian model. The origin times solved for using the Fennoscandian model are given in Table 2. The AK135 model, with slightly slower crustal velocities, returned origin time estimates that were earlier than the estimates in Table 2 by an average of 0.0941 s (94.1 ms). However, the values of $(t_{orig}^{fescan} - t_{orig}^{AK135})$ were highly consistent from event to event with a standard deviation of 0.00145 s (1.45 ms). Since this variability is of the same order as the precision at which the traveltime differences are recorded, we are confident that the relative origin time estimates obtained are as accurate as we can demand given the accuracy of the differential traveltime data. The user is encouraged to investigate the stability of these origin time estimates by experimenting with parameters for the Bayesloc program and repeating the calculations as required.

Any error in the origin times given in Table 2 which is constant for each event is of no consequence to the source-array analysis. This is because the origin times are used only as a reference point for cutting data segments for each of the stations. Fig. 6 displays the waveforms for all 55 events on station SGF aligned according to these origin time estimates (panel a); five signals are singled out for closer inspection (panel b). The waveforms were upsampled to a sampling interval of 0.001 s in order to mitigate any inaccuracy due to the temporal discretization when the traces are cut according to the origin time estimates. The direction of the outgoing *P*- and *S*-wave fronts at SGF are estimated using broadband frequency–wavenumber (*f–k*) analysis in the windows as indicated. Panels (c) and (d) show the results of *f–k* analysis in the two different time-windows together with the AK135-predicted slowness vectors. For both *P*- and *S*-wave arrivals, the *f–k* maximum falls slightly closer to

Table 2. Origin time estimates for the Hukkakero explosion listed in Table 1 using Bayesloc using the Fennoscandian velocity model (Mykkeltveit & Ringdal 1981).

mm-dd	evID	hh.mm.ss.sss	evID	hh.mm.ss.sss
08-15	H01	07.59.59.936	H02	12.00.00.150
08-16	H03	08.00.00.275	H04	12.00.00.374
08-17	H05	07.30.00.774	H06	11.00.00.380
08-18	H07	07.30.00.369	H08	11.00.00.667
08-19	H09	08.00.00.684	H10	11.30.00.937
08-20	H11	07.30.01.185	H12	11.00.01.234
08-21	H13	07.30.01.571	H14	10.30.01.187
08-22	H15	07.30.01.742	H16	10.30.01.481
08-23	H17	07.30.01.885	H18	10.30.01.830
08-24	H19	07.30.02.157	H20	11.00.02.229
08-25	H21	07.30.00.357	H22	10.30.00.682
08-26	H23	08.44.59.193	H24	11.29.58.999
08-27	H25	07.29.59.073	H26	10.29.58.938
08-28	H27	07.29.59.148	H28	10.29.59.313
08-29	H29	07.29.59.183	H30	10.29.59.170
08-30	H31	07.29.58.879	H32	10.29.59.131
08-31	H33	07.29.58.802	H34	10.29.58.900
09-01	H35	07.29.58.890	H36	10.29.59.133
09-02	H37	08.14.59.515	H38	10.44.59.132
09-03	H39	07.14.59.264	H40	09.59.59.430
09-04	H41	07.29.59.399	H42	09.59.59.594
09-05	H43	07.29.59.402	H44	09.59.59.292
09-06	H45	10.29.59.639		
09-07	H46	10.00.00.372		
09-08	H47	09.29.59.721		
09-09	H48	09.30.00.422		
09-10	H49	09.00.01.169		
09-11	H50	09.00.00.644		
09-12	H51	10.00.01.260		
09-13	H52	09.30.00.390		
09-14	H53	09.30.00.455		
09-15	H54	09.00.01.577		
09-16	H55	09.30.01.265		

the centre of the grid than the model-predicted slowness. This is to say that the outgoing wavefield has a slightly faster apparent velocity than is predicted by the velocity model—and it is the apparent velocity (the speed at which the wavefront appears to cross the array) which is critical for distance determination.

This procedure was repeated for each of the six stations, for both *P* and *S* arrivals. Panel (a) of Fig. 7 displays the AK135-predicted slowness vectors for the outgoing *P* and *S* waves to each of the stations, while panel (b) shows the corresponding slowness vectors estimated empirically using the array analysis displayed in Fig. 6. The differences, analogous to the slowness corrections used to calibrate array measurements of incoming wavefronts (e.g. Schweitzer 2001), are displayed in panels (c) and (d) for *P*- and *S*-wave fronts, respectively. In panel (a), we see a clear difference between the slowness vectors for the the more distant stations (ARCES and KEV) and those for the closer stations (LP53, LP61, SGF and LP34). For ARCES and KEV, the predicted apparent velocities of *P* and *S* arrivals are about 8.0 km s^{-1} and 4.5 km s^{-1}, respectively. For the other stations, the predicted apparent velocities for *P* and *S* arrivals are about 5.9 and 3.5 km s^{-1}, respectively. The empirical results for ARCES and KEV (panel b) are very similar to the model-predicted slowness vectors (panel a). For the closer stations, the differences are greater. (This becomes clear in the slowness corrections displayed in panels c and d.) Consider station LP53; the apparent velocities associated with tohe measured *P*- and *S*-arrival slowness vectors are almost identical to those for ARCES and KEV. While the layered
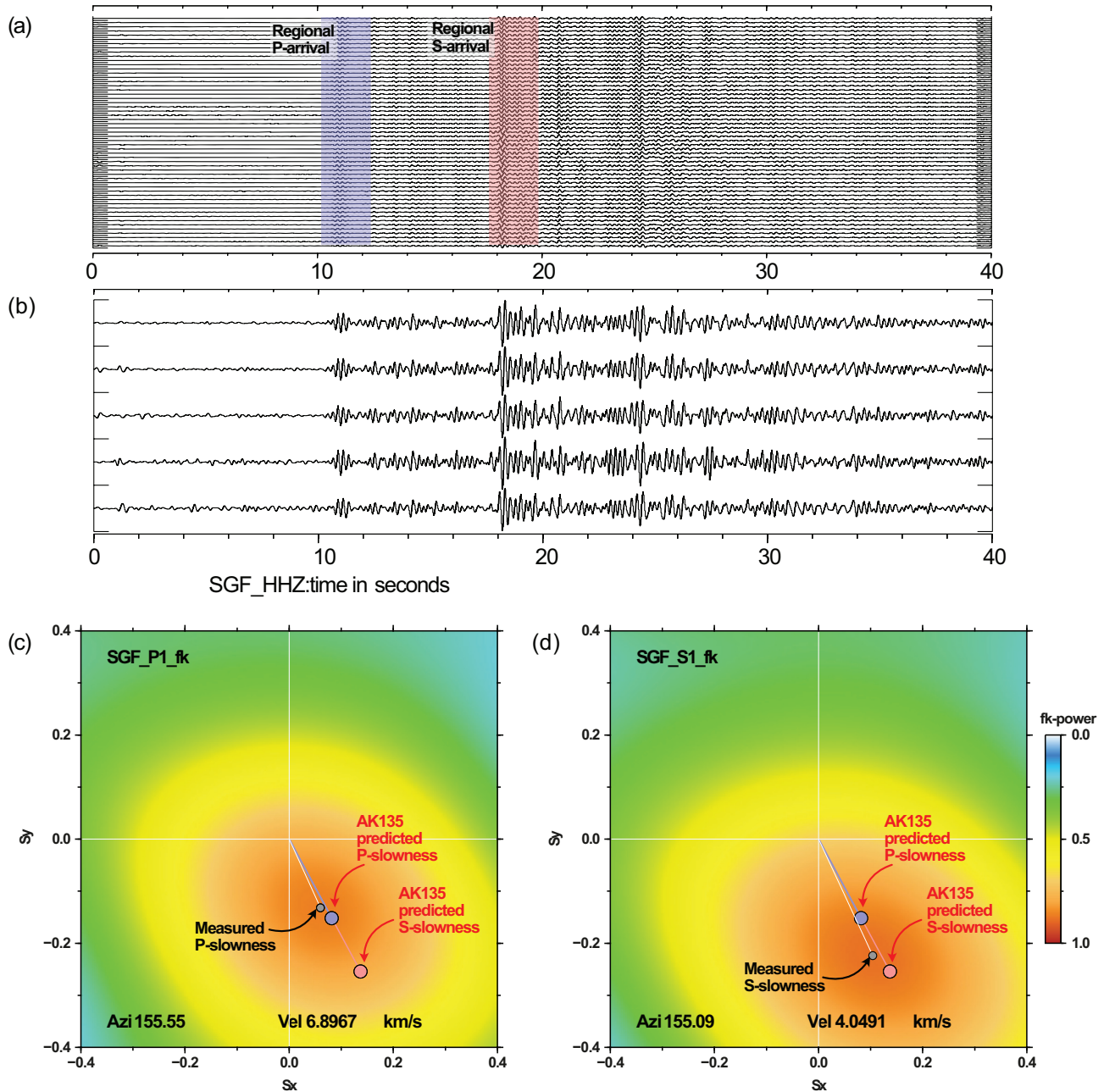
Figure 6. (a) Vertical seismograms from station SGF for the 55 events, bandpass filtered 2–4 Hz, interpolated to 1000 Hz sampling frequency, and aligned according to the origin times listed in Table 2. The *P*-wave and *S*-wave time intervals are shaded blue and red, respectively. (b) Zoom of five signals from panel (a). *f*–*k* analysis performed on virtual source array, with each trace ascribed the Ground Truth coordinates of the corresponding explosion, for *P* window (c) and for *S* window (d). The small grey circles in panels (c) and (d) indicate the slowness vector with the optimal relative power and the blue and red circles indicate the theoretical *P*- and *S*-slowness vectors for the wavefield leaving the source under the AK135 velocity model and great-circle path assumption. Both measured *P* and *S* arrivals appear closer to the centre of the plots than the AK135 predictions indicating faster apparent velocity.

models predict a Pg-type phase to this closest station, the apparent velocities are higher. The apparent velocities for station LP61, to the West, are significantly higher than predicted. However, we urge some caution regarding the accuracy of the empirical slowness vectors; some uncertainty will result from the relatively low resolution of the calculations. (The peaks in the relative power functions in Fig. 6 are relatively broad since our virtual source array has aperture 300 m, considerably smaller than most operational seismic arrays.) We recall also the results of Gibbons *et al.* (2010) in which significantly different slowness vectors for an incoming wavefield

were obtained when using different frequency bands in the *f*–*k* analysis.

In several cases, a small azimuthal deviation is observed in the measured slowness vector. This was ignored in the subsequent relative event location calculation since the procedure described by Gibbons *et al.* (2017b) assumes great-circle propagation; a factor α multiplies the traveltime gradient. The α were calculated using the apparent velocities displayed in Fig. 7; $\alpha > 1$ implies a slower-than-predicted wave front while $\alpha < 1$ indicates a faster-than-predicted wave front.
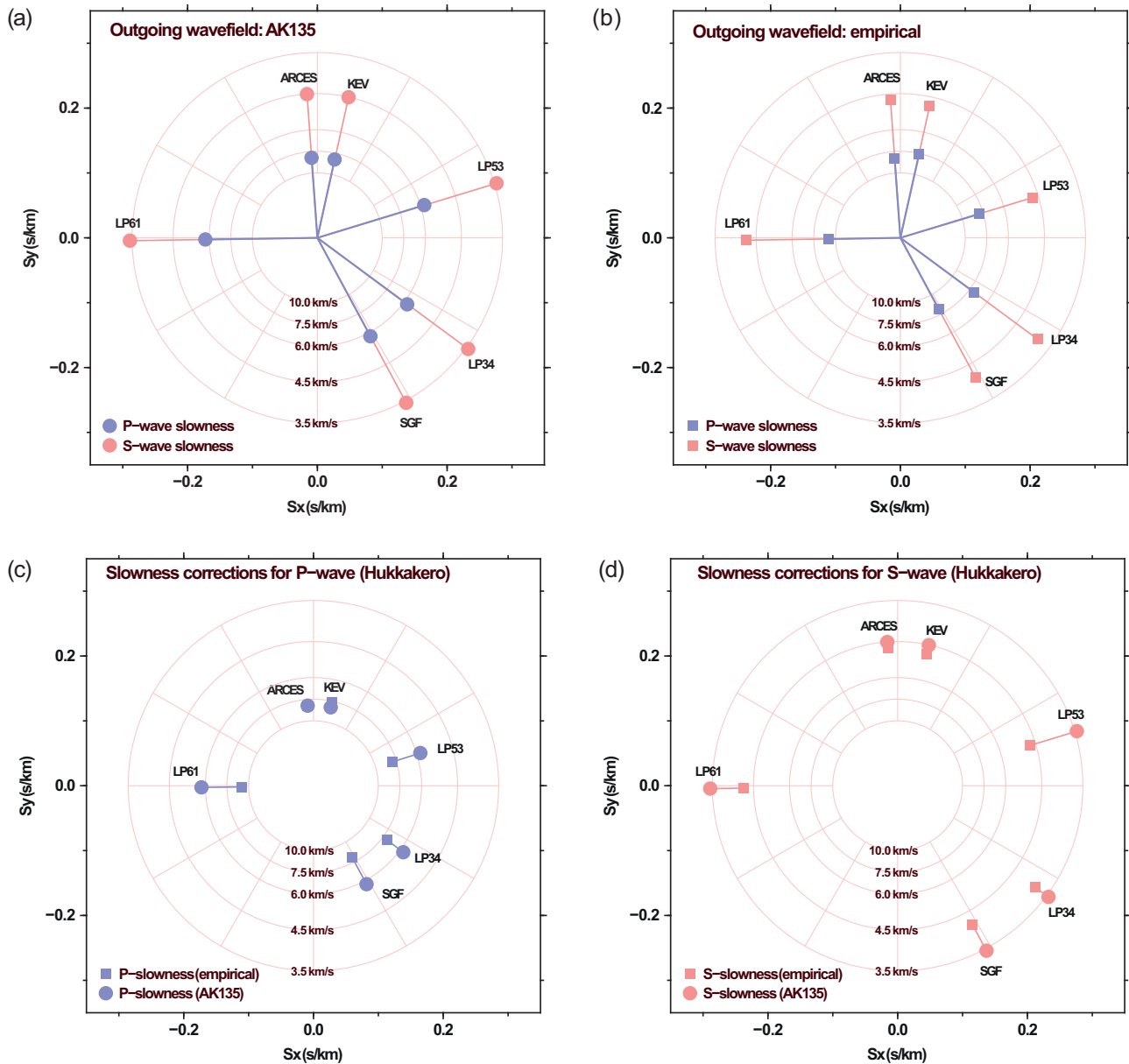
Figure 7. Slowness maps for the outgoing wavefield at Hukkakero using (a) the AK135 velocity model and great-circle path assumption and (b) the source-array analysis as displayed in Fig. 6. Panels (c) and (d) display the corrections for *P* and *S* arrivals, respectively, needed to move from the AK135-predicted slowness vector to the empirically determined slowness vector.

The relative event location results, with and without the empirical slowness corrections, are displayed in Fig. 8. The relative event locations using the corrections (displayed in panel b) have much smaller mislocation vectors than those without (displayed in panel a). The empirical slowness corrections have increased the apparent velocities of the outgoing wavefield to South, East, and West. The resulting location estimates are far closer to the GT locations than the uncalibrated estimates. Significantly, whereas the mislocation vectors displayed in panel (a) of Fig. 8 are relatively uniform within the clusters of explosions, the mislocation vectors in panel (b) are almost random with no clear pattern. We conclude that the event mislocations after the application of slowness corrections are primarily due to accuracy of time-delay measurements rather than the

fundamental assumptions regarding the propagation of the outgoing seismic wavefield.

While we have obtained provisionally satisfactory relative location estimates for the 2007 Hukkakero explosions, we sit with the fundamental caveat that we used the GT coordinates to calculate the necessary slowness corrections. In other words, there is a circularity inherent in these calculations and we would not be able to follow an identical procedure for a set of earthquakes or explosions for which we had no *a priori* knowledge. What we have demonstrated is that systematic mislocation is most likely dominated by incorrect assumptions about the speed of the outgoing wavefield to each of the stations used. All approaches to accurate relative event location need to take this into account.
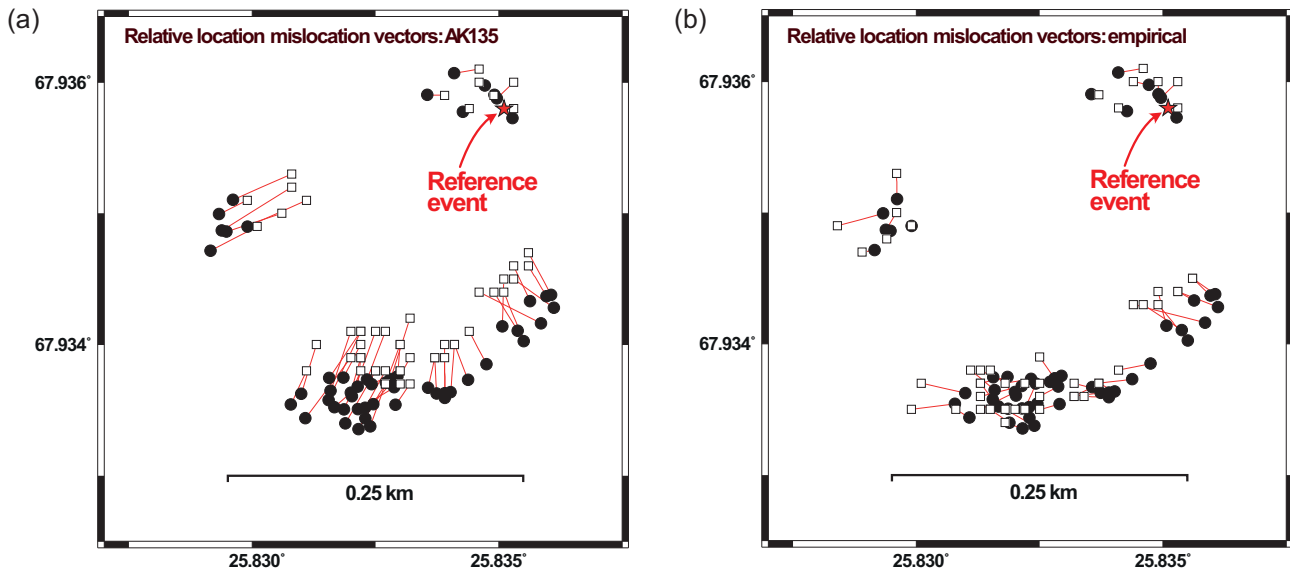
Figure 8. Mislocation vectors for location estimates for all explosions relative to event H03 (fixed), denoted by a red star, using AK135 (a) and empirical slowness corrections (b) for all *P*- and *S*-phase arrivals. Note that panel (a) is identical to panel (a) of Fig. 5.

5 CONCLUSIONS

We present a set of GT explosions, well-recorded seismically up to distances of over 200 km, for which the source coordinates are known exactly. We provide a set of waveforms for all events and accurate time-delay measurements for all event and phase pairs. The distances between events are relatively short and are of relevance to interevent distances of small earthquakes and induced events. Assessing both the accuracy and precision of an algorithm's ability to locate the events presented here using the available data will be of great benefit to assessing the likely fidelity of maps of relocated seismicity in case studies of natural and induced seismicity.

We demonstrate that, under the assumptions of two 1-D velocity models, the event-to-event distances are underestimated. From a source-array analysis, we obtain empirical slowness corrections which result in a significant improvement of the relative location estimates. In subsequent practical situations, we will not have enough information to perform the source-array analysis described here. Event mislocations are likely to be dominated by assumptions about the outgoing wavefield and we advocate theoretical advancements and implementations which address this issue. The GT data set includes 55 events, providing the possibility of working on many different subsets of events. We recommend that future studies address how much information about the relative location and the outgoing wavefield can be obtained in situations in which far fewer events and far less data is available. The GT will allow assessment and validation of any subsequent algorithms and software implementations. We hope that this study will motivate the publication of other GT data sets, with potentially very different network configurations and types of seismic source.

ACKNOWLEDGEMENTS

Most maps and figures in this paper are created using GMT software (Wessel & Smith 1995).

REFERENCES

Anderson, M.L. & Myers, S.C., 2010. Assessment of regional-distance location calibration using a multiple-event location algorithm, *Bull. seism. Soc. Am.,* **100**(2), 868–875.

Bachura, M. & Fischer, T., 2019. Waveform crosscorrelation for differential time measurement: bias and limitation, *Seismol. Res. Lett.,* **90**(5), 2005–2014.

Blixt, E.M., Näsholm, S.P., Gibbons, S.J., Evers, L.G., Charlton-Perez, A.J., Orsolini, Y.J. & Kværna, T., 2019. Estimating tropospheric and stratospheric winds using infrasound from explosions, *J. acoust. Soc. Am.,* **146**, 973–982.

Douglas, A., 1967. Joint epicentre determination, *Nature,* **215**(5096), 47–48.

Fisk, M.D., 2002. Accurate locations of nuclear explosions at the lop nor test site using alignment of seismograms and IKONOS satellite imagery, *Bull. seism. Soc. Am.,* **92**(8), 2911–2925.

Gibbons, S., Kværna, T. & Näsholm, P., 2018. Characterization of the infrasonic wavefield from repeating seismo-acoustic events, in *Infrasound Monitoring for Atmospheric Studies,* pp. 387–407, eds Le Pichon, A., Blanc, E. & Hauchecorne, A., Springer, Cham.

Gibbons, S.J. & Kværna, T., 2017. Illuminating the seismicity pattern of the October 8, 2005, M = 7.6 Kashmir earthquake aftershocks, *Phys. Earth planet. Inter.,* **270,** 1–8.

Gibbons, S.J. & Ringdal, F., 2006. The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.,* **165,** 149–166.

Gibbons, S.J., Ringdal, F. & Kværna, T., 2007. Joint seismic-infrasonic processing of recordings from a repeating source of atmospheric explosions, *J. acoust. Soc. Am.,* **122**(5), EL158.

Gibbons, S.J., Kværna, T. & Ringdal, F., 2010. Considerations in phase estimation and event location using small-aperture regional seismic arrays, *Pure appl. Geophys.,* **167,** 381–399.

Gibbons, S.J., Harris, D.B., Dahl-Jensen, T., Kværna, T., Larsen, T.B., Paulsen, B. & Voss, P.H., 2017. Locating seismicity on the Arctic plate boundary using multiple-event techniques and empirical signal processing, *Geophys. J. Int.,* **211**(3), 1613–1627.

Gibbons, S.J., Pabian, F., Näsholm, S.P., Kværna, T. & Mykkeltveit, S., 2017. Accurate relative location estimates for the North Korean nuclear tests using empirical slowness corrections, *Geophys. J. Int.,* **208**(1), 101–117.

Gibbons, S.J., Kværna, T., Näsholm, S.P. & Mykkeltveit, S., 2018. Probing the DPRK nuclear test site down to low seismic magnitude, *Seismol. Res. Lett.,* **89**(6), 2034–2041.

Goertz-Allmann, B.P., Gibbons, S.J., Oye, V., Bauer, R. & Will, R., 2017. Characterization of induced seismicity patterns derived from internal structure in event clusters, *J. geophys. Res.,* **122**(5), 3875–3894.

Ichinose, G.A., Myers, S.C., Ford, S.R., Pasyanos, M.E. & Walter, W.R., 2017. Relative surface wave amplitude and phase anomalies from the Democratic People's Republic of Korea announced nuclear tests, *Geophys. Res. Lett.,* **44**(17), 8857–8864.

Jerkins, A.E., Shiddiqi, H.A., Kværna, T., Gibbons, S.J., Schweitzer, J., Ottemöller, L. & Bungum, H., 2020. The 30 June 2017 North Sea earthquake: location, characteristics, and context, *Bull. seism. Soc. Am.,* **110**(2), 937–952.

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

Kozlovskaya, E., 2007. Seismic network XK:LAPNET/POLENET seismic temporary array (RESIF-SISMOB). RESIF - Réseau Sismologique et géodésique Français, https://doi.org/10.15778/RESIF.XK2007.

Kozlovskaya, E., Narkilahti, J., Nevalainen, J., Hurskainen, R. & Silvennoinen, H., 2016. Seismic observations at the Sodankylä Geophysical Observatory: history, present, and the future, *Geosci. Instrum. Method. Data Syst.,* **5,** 365–382.

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

Murphy, J.R., Stevens, J.L., Kohl, B.C. & Bennett, T.J., 2013. Advanced seismic analyses of the source characteristics of the 2006 and 2009 North Korean nuclear tests, *Bull. seism. Soc. Am.,* **103**(3), 1640–1661.

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

Myers, S.C., Johannesson, G. & Hanley, W., 2009. Incorporation of probabilistic seismic phase labels into a Bayesian multiple-event seismic locator, *Geophys. J. Int.,* **177**(1), 193–204.

Mykkeltveit, S. & Ringdal, F., 1981. Phase identification and event location at regional distances using small-aperture array data, in *Identification of Seismic Sources—Earthquake or Underground Explosions,* pp. 467–481, eds Husebye, E.S. & Mykkeltveit, S., Reidel Publishing Company.

Nooshiri, N., Saul, J., Heimann, S., Tilmann, F. & Dahm, T., 2017. Revision of earthquake hypocentre locations in global bulletin data sets using source-specific station terms, *Geophys. J. Int.,* **208**(2), 589–602.

Pasyanos, M.E. & Myers, S.C., 2018. The coupled location/depth/yield problem for North Korea's declared nuclear tests, *Seismol. Res. Lett.,* **89**(6), 2059–2067.

Richards, P., Waldhauser, F., Schaff, D. & Kim, W.-Y., 2006. The applicability of modern methods of earthquake location, *Pure appl. Geophys.,* **163**(2), 351–372.

Richards-Dinger, K.B. & Shearer, P.M., 2000. Earthquake locations in southern California obtained using source-specific station terms, *J. geophys. Res.,* **105**(B5), 10939–10960.

Schanze, T., 1995. Sinc interpolation of discrete periodic signals, *IEEE Trans. Signal Process.,* **43**(6), 1502–1503.

Schweitzer, J., 2001. Slowness corrections one way to improve IDC products, *Pure appl. Geophys.,* **158,** 375–396.

Selby, N.D., 2010. Relative locations of the October 2006 and May 2009 DPRK announced nuclear tests using international monitoring system seismometer arrays, *Bull. seism. Soc. Am.,* **100,** 1779–1784.

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

Usoltseva, O. & Kozlovskaya, E., 2016. Studying local earthquakes in the area Baltic-Bothnia Megashear using the data of the POLENET/LAPNET temporary array, *Solid Earth,* **7**(4), 1095–1108.

Voytan, D.P., Lay, T., Chaves, E.J. & Ohman, J.T., 2019. Yield estimates for the six North Korean nuclear tests from teleseismic p wave modeling and intercorrelation of p and pn recordings, *J. geophys. Res.,* **124**(5), 4916–4939.

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

Waldhauser, F., Schaff, D., Richards, P.G. & Kim, W.-Y., 2004. Lop nor revisited: underground nuclear explosion locations, 1976-1996, from double-difference analysis of regional and teleseismic data, *Bull. seism. Soc. Am.,* **94**(5), 1879–1889.

Wen, L. & Long, H., 2010. High-precision location of North Korea's 2009 nuclear test, *Seismol. Res. Lett.,* **81,** 26–29.

Wessel, P. & Smith, W.H.F., 1995. New version of the generic mapping tools, *EOS, Trans. Am. Geophys. Un.,* **76,** 329.

Yu, H., Harrington, R.M., Liu, Y. & Wang, B., 2019. Induced seismicity driven by fluid diffusion revealed by a near-field hydraulic stimulation monitoring array in the montney basin, british columbia, *J. geophys. Res.,* **124**(5), 4694–4709.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.
relloc_supplement
Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX

The raw waveforms for all events are contained within five tar files. Each file contains the waveforms for 11 events. Unpacking each file will generate 11 directories. For example, typing

`tar xvf events01_11.tar.gz` will generate directories H01, H02 and up to H11. Each of the available channels are contained within a single miniseed file with the naming convention `event_station_channel.mseed`, for example `H01_ARE0_bz.mseed`. Note that there is no data from station KEV for the final few events in the sequence.

The file `Relloc_Bayesloc_files.tar` contains all the files required to run the Bayesloc program to solve for approximate origin times for the GT events. If you have successfully installed the Bayesloc program on your system, the command

```
bayesloc bayesloc.cfg
```

will run the program with constraints on the events stored in the files `arrival.dat`, `station.dat` and `origin_prior.dat`, and traveltime tables stored in the files starting with the prefix `fescan`. The bayesloc program is obtained from

https://www-gs.llnl.gov/nuclear-threat-reduction /nuclear-explosion-monitoring/bayesloc

(last accessed 16 June 2020). Please read the Bayesloc documentation available at the above site for details on file formats and program output.

The file `Benchmark.site` contains the station coordinates in the so-called CSS3.0 format.

The file `absolute_CC_times.txt` contains the results of all cross-correlation calculations carried out. There are seven columns in this file:

```
H01  H02  2007-08-15T08:00:27.857  2007-08-15T12:
00:28.110  ARCES  P1  0.828
H01  H02  2007-08-15T08:00:48.693  2007-08-15T12:
00:48.956  ARCES  S1  0.755
H01  H02  2007-08-15T08:00:32.148  2007-08-15T12:
00:32.407  KEV   P1  0.808
H01  H02  2007-08-15T08:00:55.371  2007-08-15T12:
00:55.641  KEV   S1  0.796
H01  H02  2007-08-15T08:00:19.109  2007-08-15T12:
00:19.299  LP34  P1  0.850
H01  H02  2007-08-15T08:00:33.100  2007-08-15T12:
00:33.286  LP34  S1  0.798
H01  H02  2007-08-15T08:00:08.467  2007-08-15T12:
00:08.709  LP53  P1  0.926
H01  H02  2007-08-15T08:00:15.228  2007-08-15T12:
00:15.476  LP53  S1  0.813
H01  H02  2007-08-15T08:00:11.736  2007-08-15T12:
00:11.945  LP61  P1  0.678
H01  H02  2007-08-15T08:00:21.048  2007-08-15T12:
00:21.232  LP61  S1  0.731
H01  H02  2007-08-15T08:00:09.053  2007-08-15T12:
00:09.243  SGF   P1  0.658
H01  H02  2007-08-15T08:00:16.574  2007-08-15T12:
00:16.744  SGF   S1  0.702
```

Column (1) provides the event ID string for the reference event and column (2) provides the event ID string for the detected event. Column (5) gives the station name. Note that it states ARCES and not ARE0 since the full array correlation stack was used to measure the time-delay as accurately as possible. For the three-component stations, the three-channel correlation stack was used. Column (6) indicates the phase (either first P, P1 or first S, S1). Column (3) provides the starting time (in ISO 8601 format) of the reference event waveform template, and column (4) provides the time at which the maximum correlation coefficient is obtained. This maximum value is provided in column (7). Note that this list is generated using a fully automated correlation procedure and has not been checked manually for potential spurious correlation maxima, for example resulting from data errors. As a sanity check, all autocorrelation calculations have also been included. That is to say that if columns (1) and (2) are identical, columns (3) and (4) should also be identical and column (7) should read unity. In practice, these lines will always be ignored in relative event location calculations.