Comparing the performance of stacking-based methods for microearthquake location: A case study from the Burträsk fault, northern Sweden

Ruth A. Beckel\textsuperscript{1\!*}, Björn Lund\textsuperscript{1}, Gunnar A. Eggertsson\textsuperscript{1}, Christopher Juhlin\textsuperscript{1}

\textsuperscript{1} Department of Earth Sciences, Uppsala University, SE-752 36 Uppsala, Sweden

\!* corresponding author, e-mail: ruth.beckel@geo.uu.se

28 September 2021

SUMMARY

Traditional earthquake location relying on first arrival picking is challenging for microseismic events with low signal-to-noise ratio. Over the past years, alternative procedures have been explored based on the idea of migrating the energy of an earthquake back into its source position by stacking along theoretical traveltime curves. To avoid destructive interference of signals with opposite polarity, it is common to transform the input signals into positive timeseries. Stacking-based source location has been successfully applied at various scales, but existing studies differ considerably in the choice of characteristic function, the amount of pre-processing and the phases used in the analysis. We use a dataset of 62 natural microearthquakes recorded on a 2D seismic array of 145 vertical geophones across the glacially-triggered Burträsk fault to compare the performance of five commonly used characteristic functions: the noise filtered seismograms and the semblance, the envelope, the short term average/long term average ratio and the kurtosis gradient of the seismograms. We obtain the best results for a combined P- and S-wave location using a polarity-sensitive characteristic function, i.e., the filtered seismograms or the semblance. In contrast, the absolute functions often fail to align the signals properly, yielding biased location estimates. Moreover, we observe that the success of the procedure is very sensitive to noise suppression and signal shaping prior
to stacking. Our study demonstrates the usefulness of including lower quality S-wave data to improve the location estimates. Furthermore, our results illustrate the benefits of retaining the phase information for location accuracy and noise suppression. To ensure optimal location results, we recommend carefully pre-processing the data and test different characteristic functions for each new dataset. In spite of the sub-optimal array geometry, we obtain good locations for most events within ~30-40 km of the survey and the locations are consistent with an image of the fault trace from an earlier reflection seismic survey.

**Key words:** Earthquake source observations; Computational seismology; Time-series analysis; Body waves; Fractures, faults, and high strain deformation zones

1 **INTRODUCTION**

Traditionally, earthquake location is based on the inversion of traveltimes which has the disadvantage that these have to be picked from the seismic records. If automatic pickers are not working properly, this is a time-consuming and, to a certain extent, subjective procedure that can introduce errors into the analysis. Automatic or manual, a prerequisite for picking-based source location is that the events need to be clearly detectable on the individual records. To address these problems, alternative methods for source location using waveforms instead of the traveltime information have been explored over the past years (see Li et al. 2020, for a review). The most common method is based on the idea of focusing the energy of earthquakes into the source location by reverse time migration through a pre-defined velocity model. The migration can be realized either by back-propagating the whole wavefield by reverse modelling (e.g. Gajewski et al. 2007; Artman et al. 2010) or, analogue to Kirchhoff migration, by stacking along diffraction hyperbolas defined by the traveltimes from potential source points to the receivers (e.g. Kao & Shan 2004; Gharti et al. 2010). The latter approach is more common since it is less sensitive to the velocity model and computationally less expensive (Li et al. 2020). Since the method is still relatively new, there are several different names in use, such as source scanning (Kao & Shan 2004, 2007), backprojection imaging (Beskardes et al. 2017), diffraction stacking (Gajewski et al. 2007; Li et al. 2019), (partial) waveform stacking (Grigoli et al. 2013; Cesca & Grigoli 2015; Li et al. 2020) and migration-based
Comparing stacking-based location methods for microearthquake location

location (Pesicek et al. 2014; Staněk et al. 2015; Gharti et al. 2010; Trojanowski & Eisner 2017). In the following, we will use the term stacking-based location since it is the most descriptive term for the methods presented in this paper.

The main advantage of stacking-based location methods is that phase picking is not required so that the first arrivals do not necessarily need to be detectable on individual records. Depending on the number of stations and the noise characteristics, this can facilitate the detection of very weak events. Moreover, the detection and location process can be carried out fully automatically (after initial parameter tuning). However, there is one major challenge: Polarity reversals across the stations lead to destructive interference during stacking which can significantly bias the locations or, in extreme cases, cause the method to fail. Therefore, it is common to pre-process the signals and/or to transform them into a characteristic function before stacking. Previous studies of stacking-based source location methods differ mainly in the choice of the characteristic function. To circumvent the polarity problem, it is common practice to convert the input traces into a positive time series using the envelope (Baker et al. 2005; Gharti et al. 2010; Kao & Shan 2007; Pesicek et al. 2014), the STA/LTA ratio (Drew et al. 2013; Grigoli et al. 2013; Li et al. 2017) or the kurtosis gradient (Langet et al. 2014; Poiata et al. 2016; Wagner et al. 2017) of the signal. However, dismissing the phase information can reduce the signal-to-noise ratio significantly since incoherent noise is no longer eliminated by destructive interference. To make use of the noise suppression potential in the phase, while avoiding stack degradation due to phase changes across the array, some investigators suggest retaining the phase information but including a polarity correction in the analysis. As characteristic function, the noise-filtered seismic traces (Anikiev et al. 2014; Beskardes et al. 2017; Trojanowski & Eisner 2017), the semblance of these traces (Chambers et al. 2010; Staněk et al. 2015) or a correlation based characteristic function (Shi et al. 2018a; Li et al. 2017) have been suggested. All three have the advantage of being less affected by high amplitude noise but are more sensitive to small unmodeled time shifts, e.g. due to near-surface heterogeneities or inaccuracies of the velocity model.

Another difference between existing studies is the amount of data pre-processing performed. Some investigators do not apply any pre-processing (Beskardes et al. 2017; Shi et al. 2018b),
some normalize (Kao & Shan 2004; Cesca & Grigoli 2015) or bandpass filter (Langet et al. 2014; Wagner et al. 2017) the data and some do both (e.g. Chambers et al. 2010; Pesicek et al. 2014; Zeng et al. 2014). Recently, new noise suppression techniques have been suggested that are specially designed for microseismic monitoring (Birnie et al. 2017; Trojanowski 2019; Alexandrov et al. 2020).

Finally, not all existing studies use both P- and S-waves for the location procedure. If three-component data are available, it is common to use the vertical component for the P-wave analysis and the horizontal components for the S-wave analysis (Cesca & Grigoli 2015; Drew et al. 2013; Gharti et al. 2010; Grigoli et al. 2013, 2016; Pesicek et al. 2014; Wagner et al. 2017; Shi et al. 2018b). If only vertical component data are available, most researchers choose to exclude the S-waves (Anikiev et al. 2014; Beskardes et al. 2017; Chambers et al. 2010; Stanek et al. 2015).

There are a few studies comparing the performance of different characteristic functions for stacking based migration but most of them are mainly based on synthetic data (Cesca & Grigoli 2015; Trojanowski & Eisner 2017; Shi et al. 2018a) and/or a comparison of selected, usually high-quality, events (Cesca & Grigoli 2015; Shi et al. 2018b; Beskardes et al. 2017).

In this paper we analyse a local microseismicity dataset registered on 145 vertical geophones arranged (mostly) along a crooked, 20 km long profile across the glacially-triggered Burträsk fault in northern Sweden. The array follows an existing reflection profile and was primarily designed for structural imaging with earthquake sources. Part of this process was to obtain precise location estimations for the earthquakes. The array is certainly not optimal for this task, but the simple geometry and the close spacing of the stations facilitate visual identification of weak events and an assessment of the alignment across the stations. We use this dataset to test how well different characteristic functions perform for stacking-based source location on a noisy dataset. In the analysis we focus on signal alignment and the influence of pre-processing. Moreover, we examine the potential of including lower quality S-wave data in order to improve the location results.
Comparing stacking-based location methods for microearthquake location

Figure 1. Glacially triggered faults (GTFs) in northern Scandinavia (Munier et al. 2020) and seismicity in northern Sweden recorded by the Swedish National Seismic Network (SNSN) since 2006 (Lund et al. 2021). The virtual Burträsk network is shown in blue.

2 SURVEY AREA AND DATA

The recording array is situated directly above the glacially-triggered Burträsk fault in northern Sweden (Lagerbäck & Sundh 2008). Glacially-triggered faults (GTFs, also called post- or end-glacial faults) are conspicuous geological features forming up to 15 m high fault scarps that can extend for tens of kilometres (Fig. 1) (e.g. Lagerbäck & Sundh 2008; Olesen et al. 2013; Kuivamäki et al. 1998). They are predominantly found in northern Fennoscandia and are thought to have formed by reactivation of pre-existing weakness zones in response to sudden stress changes in connection with the last glacial retreat (e.g. Lund 2015).

The Burträsk fault comprises a 45 km long series of mostly southwest-northeast oriented segments. The fault scarp is on average 5-10 m high and mostly covered by Quaternary sediments (Lagerbäck & Sundh 2008; Mikko et al. 2015). As many other GTFs, it follows a regional weakness zone, the Burträsk Shear Zone (BSZ) (Fig. 2). The BSZ is a wide, dextral shear zone that...
is suggested to have formed by lateral escape during the Svecokarelian orogeny (Romer & Nisca 1995). Regional metamorphism terminated around \( \sim 1.8 \) Ga (Weihed et al. 2002) and no major re-activation of the BSZ after that is documented until the rupture of the Burträsk fault.

Based on the widespread occurrence of sediment liquefaction structures and landslides in the vicinity of the GTFs, they are believed to have formed in single events, major earthquakes with magnitudes up to 7-8 (Arvidsson 1996; Lagerbäck & Sundh 2008; Lindblom et al. 2015). Present seismicity in northern Scandinavia is much smaller in magnitude, but the majority of events north of 66° occur in direct vicinity of the GTFs (Lindblom et al. 2015) (Fig. 1), indicating that the GTFs are still active and thus locations of current stress release. The Burträsk area is the seismically most active area in Sweden (Lund et al. 2021) (Fig. 1). The highest magnitude event recorded in this area was a ML 3.6 earthquake registered in June 2010 (Juhlin and Lund 2011), but there are historic indications of M4 events (Renqvist 1930).

In order to study the seismicity in the area in greater detail, the Swedish National Seismic Network (SNSN, Lund et al. 2021) installed a temporary network of six stations in the Burträsk area in December 2012, complementing the four stations in the permanent network that surrounds the Burträsk fault. The temporary network was extended by six additional stations in August 2015 and the network operated until October 2017, when the temporary stations were decommissioned. In the following, we refer to the combination of temporary and permanent stations as the virtual Burträsk network (blue stations in Fig. 1). The data were processed by the automatic South Iceland Lowland (SIL) detection and location system using the standard SNSN velocity model (Lund et al. 2021). The events were manually revised and relocated, first in the SNSN velocity model and then VELEST (Kissling et al. 1994) was used to derive a local 1D velocity model for the area and to relocate the events. Finally many of the events were relatively relocated using HypoDD (Waldhauser & Ellsworth 2000). Work on the full dataset is ongoing and will be reported elsewhere, but we will refer here to the re-located events as the Burträsk catalogue. Preliminary results show that the majority of events concentrate in the area southeast of the fault scarp, in agreement with a southeast dipping reverse fault, but also occur in a band of seismicity extending to the northeast.
Comparing stacking-based location methods for microearthquake location

Figure 2. Left) Seismicity of the survey area and selected events from the Burtræsk catalogue. The Burtræsk fault follows the general trend of the Burtræsk Shear Zone (BSZ). The seismicity recorded by the permanent SNSN stations is shown in green. The events from the Burtræsk catalogue included in this paper are shown in red. Right) Close-up of the study area showing the Burtræsk fault scarp (blue lines) and the location of the BUR array (cyan triangles). Events discussed in this paper are highlighted in yellow. Note that not all stations of the virtual Burtræsk network are shown here. The complete network is shown in Fig. 1.

(Lund et al. 2016) (Fig. 2). The seismically active zone is associated with two reflectors imaged with a shallow reflection seismic survey (Juhlin & Lund 2011; Beckel & Juhlin 2019).

In addition to the virtual Burtræsk network, we installed an array of 145 stations that followed the reflection profile of Juhlin & Lund (2011), along with a few additional stations to increase the cross-profile coverage (Fig. 2). We will refer to this as the BUR array. The profile is oriented nearly perpendicular to the Burtræsk fault and follows existing roads and forest paths. We used 4.5 Hz vertical component geophones and recorded the data on SERCEL RAU-X wireless units at a 4 ms sampling interval. The units were deployed mostly in the soil cover at 150-200 m spacing and left to record for 25 days from 28 July to 22 August in 2017. During this period, the Burtræsk catalogue comprises 38 microearthquakes with magnitudes ranging from -0.6 to 2.1. After manual inspection of waveforms, 24 additional, low quality microearthquakes were identified that had not been detected by the automatic SIL system so that the total number of events for our analysis increased to 62. The events occurred mostly southeast of the fault (Fig. 2). For the 24 later added events there is no magnitude information available. We used the pick-based locations from the
Burträsk catalogue as a starting point for our analysis. Otherwise, the data from the virtual Burträsk network are not included in the stacking-based location process.

3 METHOD

3.1 Calculation of the image function

The principle of stacking-based location is derived from Kirchhoff or diffraction stack migration. The most likely source parameters for an earthquake are identified by a grid search across a 4D image function that is calculated by stacking along diffraction hyperbolas defined by the traveltimes from the potential sources to the receivers (Fig. 3). The value of the image function quantifies how consistent a potential source location - origin time pair is with the recorded data for a given velocity model.

Around each preliminary location from the Burträsk catalogue, we constructed a 3D volume of potential source points \( r_i \) and computed the traveltimes \( \tau_{ij} \) between each potential source point and each station \( r_j \) of the BUR array using a traveltime calculator based on the Earth flattening approximation (Müller 1971) together with a 1D velocity model obtained from the earthquake data in the virtual Burträsk network using VELEST. For most events we used a 4 km x 4 km x 4 km volume with 50 m grid spacing, but for some events we increased this to a 6 km x 6 km x 6 km volume. We also tried estimating static corrections for the stations from five shot points along the profile, but did not obtain any consistent results.

We use a 2 s window and a 10 ms increment to construct a vector of potential origin time \( t_0 \) around the origin time from the Burträsk catalogue. Together with the traveltime, this defines the predicted arrival times \( t_p \):

\[
t_p(t_0, r_i, r_j) = t_0 + \tau_{ij}(r_i, r_j)
\]

(1)

The image function, also called brightness function, is obtained by stacking the signals of the stations along the predicted arrival time (Fig. 3):

\[
I_i(t_0) = \sum_{j=1}^{N} A_j(t_p) = \sum_{j=1}^{N} A_j(\tau_{ij} + t_0)
\]

(2)

where \( I_i \) is the image function at the \( i^{th} \) source grid point, \( N \) is the number of stations and \( A_j \) is
Comparing stacking-based location methods for microearthquake location

Figure 3. Simplified sketch illustrating the principle of stacking-based source location. The source volume is used to calculate the traveltimes to the receivers (black triangles). The pre-processed data are stacked along the traveltimes for each potential source point and each origin time. This comprises a 4D image function whose maximum corresponds to the most likely source location and origin time. A detection function can be obtained by extracting the maximum of the image function for each origin time. Maxima in the detection function exceeding a certain threshold are considered as possible events.

the characteristic function used for stacking at the $j^{th}$ station. This is done for each potential source point and each potential arrival time, resulting in a 4D data volume. The most likely location for an event corresponds to the maximum of the image function (Fig. 3).

We tested five different characteristic functions. In the simplest case, $A_j$ is just the (pre-processed) input waveform $u_j$ itself:

$$I_i^{raw}(t_0) = \sum_{j=1}^{N} u_j(t_p)$$  \hspace{1cm} (3)

To reduce the influence of high amplitude noise, we also tested using the semblance of the signal as characteristic function:

$$I_i^{semblance}(t_0) = \frac{\sum_{k=-w}^{+w} \left( \sum_{j=1}^{N} u_j(t_p + k) \right)^2}{N \sum_{k=-w}^{+w} \sum_{j=1}^{N} (u_j(t_p + k))^2}$$  \hspace{1cm} (4)

Where $w$ is the half-width of the time window in which the semblance is calculated.

As an alternative to these two coherency-sensitive methods, we also tried three characteristic functions that convert the signal to a positive time series. The most direct way to eliminate the polarity is to calculate the envelope:
\[ I_{i}^{\text{envelope}}(t_0) = \sum_{j=1}^{N} |H(u_j(t_p))| \]  

where \( H(u) \) is the Hilbert transform of the signal.

Two other commonly used functions are the short term average/long term average (STA/LTA) ratio and the positive gradient of the kurtosis that both act as first-arrival detection functions:

\[ I_{i}^{\text{STA/LTA}}(t_0) = \sum_{j=1}^{N} \frac{\text{STA}}{\text{LTA}}(u_j(t_p)) \]  

\[ I_{i}^{\text{kurtosis}}(t_0) = \sum_{j=1}^{N} K^+(u_j(t_p)) \]

where \( \frac{\text{STA}}{\text{LTA}}(u) \) is the STA/LTA ratio of the signal and \( K^+ \) is the positive gradient of the kurtosis of the signal (Langet et al. 2014).

In the following, we will refer to the different image functions \( I^{\text{raw}}, I^{\text{semblance}}, I^{\text{envelope}}, I^{\text{STA/LTA}} \) and \( I^{\text{kurtosis}} \) as the raw stack, semblance stack, envelope stack, STA/LTA stack and kurtosis gradient stack, respectively. Furthermore, we distinguish between the polarity-sensitive methods, i.e. the raw and semblance stack and the absolute methods, i.e. the envelope, STA/LTA and kurtosis gradient stack.

We tested the different image functions using both only the P-wave arrival and a combination of P- and S-wave arrivals. Since we only recorded the vertical component of the signal, the S-wave data have lower quality than the P-wave data. Therefore, we down-weighted the contribution from the S-wave by:

\[ I^{P+S} = \frac{I^P}{\max(I^P)} + 0.5 \frac{I^S}{\max(I^S)} \]  

when including the S-wave arrival. Here \( I_P \) is the image function calculated using the P-wave signal and \( I_S \) is the image function calculated using the S-wave signal.

As an event detection function, it is common to extract the maximum value of the image function for each origin time, often called the maximum stack amplitude (Fig. 3). Both detection level and the background noise value depend on the characteristic function used for stacking. To compensate for these differences and make the events more comparable, we use a scaled detection function \( s(t_0) \):
Comparing stacking-based location methods for microearthquake location

\begin{equation}
  s(t_0) = \max_{r_i} (I(t_0, r_i)) \cdot \frac{1}{S_{\text{noise}}}
\end{equation}

where \( I \) is the image function, \( r_i \) is the location, \( t_0 \) is the origin time and \( S_{\text{noise}} \) is the background noise value estimated from the maximum stack amplitude before or after an event (Fig. 3).

Here, \( s(t_0) \) comprises the relative stack amplitude, i.e. it estimates how much the maximum stack amplitude for an event increased relative to the typical value for background noise before or after the analysed event. This is similar to setting a time varying detection threshold to compensate for varying noise levels (Wagner et al. 2017). Since we only analyse known events, we do not actually use the function for detection but instead for comparing the results of the different methods as it visualizes how well an event can be distinguished from the background noise and how well its origin time is constrained. Moreover, we set a relative stack amplitude of 2.5 as reliability threshold. Events with a maximum below this threshold are considered undetectable. The value of 2.5 was chosen after visual inspection of all events and might not be an optimal threshold. If a study aims at comparing event detection, it might be a better choice to test the probabilistic scheme of Trojanowski (2018) for estimating comparable thresholds for different methods.

3.2 Pre-processing

Pre-processing and the choice of a characteristic function should be viewed as a combined step with the aim of optimizing the signal for stacking. The efficiency of stacking depends mainly on the signal shape, the signal amplitudes and the suppression of unwanted signals. Ideally, the signals should have a clear peak and a repeatable waveform to ensure that first arrivals align properly. The signal amplitude defines the weight of a station in the stack. Therefore, the amplitudes should be balanced to prevent single stations from dominating the image function. Unwanted parts of the signal, like coda waves, should be suppressed as far as possible before stacking. This is particularly important for the absolute value characteristic functions since noise suppression during stacking is much less efficient without the phase information. Furthermore, the signal width should be as narrow as possible since it controls the lateral resolution. Note that these are not requirements and that stacking-based locations can even be obtained for low quality signals if enough stations are used for stacking.
Most events we recorded featured an extended coda with amplitudes comparable to the first arrival, but higher frequencies. To reduce the coda energy, regularize the signal shape and prevent spatial aliasing, we bandpass filtered the signals before calculating the characteristic functions. We also tested more advanced methods like deconvolution but simple bandpass filtering worked best for our dataset. We used a frequency band of 3.5 Hz – 17 Hz for enhancing the P-wave and 3.5 Hz – 10 Hz for enhancing the S-wave before calculating the raw, semblance and envelope stack. The kurtosis gradient signal was enhanced best with a frequency band of 3.5 Hz – 30 Hz for the P-wave and 3.5 Hz – 10 Hz for the S-wave. However, a minor number of events contained higher dominant frequencies and required broader bandpass windows (Table 1). The STA/LTA signal did not improve consistently after filtering, so we did not apply any filter for this function. The choice of the calculation windows for the STA/LTA ratio, the kurtosis gradient and the semblance was a trade-off between getting a well-defined peak and limiting the width of the signal.

Another common problem of surface data with stations deployed in soil are unbalanced energy levels due to differences in coupling and transient noise bursts caused by passing vehicles. Stations 7 and 18-23 in Fig. 4a are dominated by high amplitude noise bursts that nearly drown out the earthquake signal and would dominate the stacking result. To ensure all stations contribute to the image function, we balanced the data traces before stacking by dividing each trace by its average absolute amplitude. After balancing, both P- and S-wave arrivals become clearly visible in Fig. 4a. The pre-processing was carried out mostly using ObsPy (Beyreuther et al. 2010) and is summarized in Table 1.

3.3 Uncertainty estimation

Since the image function is calculated for the whole 4D volume, it can be directly used to extract the source parameters that have the highest consistency with the data. The distribution of these source parameters can be regarded as a preliminary estimation of their uncertainty. We extract the distribution of all source location – origin time pairs with a stack value of 95% of the maximum of the image function and measure the uncertainty as the maximum distance and time between the most likely location and a point from this distribution. The preliminary uncertainty values are
Comparing stacking-based location methods for microearthquake location

Table 1. Summary of the pre-processing for the different characteristic functions (CF). Note that the semblance is calculated during stacking.

<table>
<thead>
<tr>
<th>CF</th>
<th>Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>• Butterworth bandpass filter (P: 3.5 Hz – 17/22.5 Hz, S: 3.5 Hz – 10/13 Hz)</td>
</tr>
<tr>
<td></td>
<td>• Trace balance</td>
</tr>
<tr>
<td>Semblance</td>
<td>• Butterworth bandpass filter (P: 3.5 Hz – 17/22.5 Hz, S: 3.5 Hz – 10/13 Hz)</td>
</tr>
<tr>
<td></td>
<td>• Trace balance</td>
</tr>
<tr>
<td></td>
<td>• (Semblance window: 84 ms)</td>
</tr>
<tr>
<td>Envelope</td>
<td>• Butterworth bandpass filter (P: 3.5 Hz – 17/22.5 Hz, S: 3.5 Hz – 10/13 Hz)</td>
</tr>
<tr>
<td></td>
<td>• Trace balance</td>
</tr>
<tr>
<td></td>
<td>• Envelope</td>
</tr>
<tr>
<td>STA/LTA</td>
<td>• Classic STA/LTA ratio (P: 20 ms/400 ms, S: 80 ms/800 ms)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>• Butterworth bandpass filter (P: 3.5 Hz – 30/40 Hz, S: 3.5 Hz – 10/15/20 Hz)</td>
</tr>
<tr>
<td></td>
<td>• Kurtosis (P: 1.0 s, S: 4.0 s)</td>
</tr>
<tr>
<td></td>
<td>• Positive gradient</td>
</tr>
<tr>
<td></td>
<td>• Trace balance</td>
</tr>
</tbody>
</table>

confined to multiples of the grid spacing and the maximum value is limited by the size of the search volume. Note that this is not a statistical but a data-driven approach and, thus, the choice of the percentage level is based on the variability in the image function and does not have a solid statistical justification. Therefore, the preliminary uncertainties should not be directly compared to uncertainties from picking-based locations.

4 RESULTS

4.1 Pre-processing results

Careful pre-processing proved to be essential for obtaining good locations. Without filtering and balancing the traces, strong coda arrivals and varying noise levels corrupted the image function so that all methods except for the STA/LTA stack failed completely. An example of this can be found in the supplements (Fig. S1).

The signal properties of the characteristic functions differ significantly. Fig. 4 (b) shows ex-
Figure 4. Pre-processing of the seismograms. (a) Comparison between the raw and the filtered and balanced signal for the event E2 in Fig. 2. The blue arrows mark the arrival of the P-wave. (b) Example data traces showing the results of pre-processing with the different characteristic functions for a strong event and a weak event. The blue arrows mark the arrival of the P- and S-waves.

ample traces for a strong and a weak event and Fig. 5 illustrates the waveforms for a typical event with a relatively good P-wave signal and a clearly visible but much more diffuse S-wave signal. The traces are shifted by the predicted arrival time of the best location for this event so that the first arrivals should align at 0 s but residual time shifts cause some deviation from this value. The bandpass filtered traces feature medium wide P-wave signals where the coda is often not completely suppressed, but weaker than the first arrival. The first arrival typically consists of a main peak with two major to minor side-lobes (Fig. 4b) and the signal shape is mostly consistent across the stations in spite of the variable ground conditions (Fig. 5). The S-wave signal has a much more complicated shape (Fig. 4b) that differs significantly between events (Fig. 5 and second column in Figs 6-9), but it is often relatively strong. The consistency across the stations varies.
Comparing stacking-based location methods for microearthquake location

Figure 5. Waveforms of the different characteristic functions for the P-wave (left column) and the S-wave (right column) of an event (E1 in Fig. 2) with intermediate quality. For better visibility, only every fourth station is plotted. The traces are shifted by the predicted arrival time of the best location for this event so that the first arrivals should align at 0 s. The stack columns show the stacks using the different characteristic functions for P- and S-waves, respectively. Additionally, the orange trace in the stack of the bandpass filtered waveforms shows the semblance of the bandpass filtered signals.

The semblance is a measure of the coherency between different traces and is calculated during stacking. Therefore, it does not have a waveform as such but Fig. 5 shows the semblance trace calculated for the bandpass filtered waveforms. Compared to the simple stack of the bandpass filtered traces, the semblance trace has a wider and less sharp maximum since it is calculated in a sliding window. Another feature of the semblance is that it is not necessarily highest for the peak of the input signals because the semblance is not very sensitive to amplitudes. In our dataset, the maximum of the semblance frequently corresponded to the onset of the first arrival.

The envelope is an amplitude property. Therefore, it is very sensitive to pre-processing and the signals are strongly influenced by remaining coda energy, background noise and local amplitude.

Figure 5. Waveforms of the different characteristic functions for the P-wave (left column) and the S-wave (right column) of an event (E1 in Fig. 2) with intermediate quality. For better visibility, only every fourth station is plotted. The traces are shifted by the predicted arrival time of the best location for this event so that the first arrivals should align at 0 s. The stack columns show the stacks using the different characteristic functions for P- and S-waves, respectively. Additionally, the orange trace in the stack of the bandpass filtered waveforms shows the semblance of the bandpass filtered signals.

The semblance is a measure of the coherency between different traces and is calculated during stacking. Therefore, it does not have a waveform as such but Fig. 5 shows the semblance trace calculated for the bandpass filtered waveforms. Compared to the simple stack of the bandpass filtered traces, the semblance trace has a wider and less sharp maximum since it is calculated in a sliding window. Another feature of the semblance is that it is not necessarily highest for the peak of the input signals because the semblance is not very sensitive to amplitudes. In our dataset, the maximum of the semblance frequently corresponded to the onset of the first arrival.

The envelope is an amplitude property. Therefore, it is very sensitive to pre-processing and the signals are strongly influenced by remaining coda energy, background noise and local amplitude.
focusing. As a result, both P- and S-wave signal are very wide and the peak is often shifted to later
times by coda waves (Figs 4b and 5).

The STA/LTA ratio provides relatively wide P-wave signals that often lack a clear peak (Fig.
4b) and are less consistent across the stations (Fig. 5). Since it is sensitive to amplitudes as well,
it is affected by strong coda arrivals so that the maximum in several cases does not correspond to
the first arrival (Fig. 7). The S-wave response is often very weak and scattered and rarely matches
the first arrival (Figs 4b and 5).

Ideally, the kurtosis waveform should feature a distinct peak with a very sharp increase at
the onset of a seismic event, followed by a more gradually decreasing tail so that the kurtosis
gradient has one sharp peak corresponding to the first arrival (e.g. Langet et al. 2014). In our
dataset, however, the peak in the kurtosis signal is less sharp and often has a staircase structure
that results in several very narrow peaks in the kurtosis gradient function instead of just one main
peak (Fig. 4b). Even though this increases the effective width, the kurtosis gradient still provides
the narrowest signal with the best noise suppression in most cases. However, the signal is often not
very repeatable and picks up later arrivals in several cases (Figs 7 and 9). Similar to the STA/LTA
ratio, the S-wave response is often scattered and incoherent, but better in matching the first S-wave
arrival (Figs 4b and 5).

4.2 Location results

In earthquake location the use of both P- and S-waves is an advantage. As the S-wave signal in our
vertical component data is in many cases of lower quality than the P-wave signal, we tried locating
the events both with and without the S-waves. Using only the P-wave worked reasonably well for
events that occurred below or very close to the BUR array. Fig. 6 summarizes the results of a typical
example. Fig. 6(a) shows the detection function and the location deviation, i.e. the difference
between the most likely location for each origin time and the initial location from the Burtrask
catalogue (not to be confused with the preliminary location uncertainty). Fig. 6(b) depicts how well
the first arrivals of P- and S-waves are aligned for the most likely location and Fig. 6(c) shows slices
through the maximum of the image function. All methods show a distinct peak in the detection
function that corresponds to the first arrival of the P-wave but there are significant differences in the shape of the peak and the alignment of the first arrivals. For the raw and semblance stack, the peak has a clear maximum indicating that the origin time is reasonably well constrained. In contrast to this, the envelope stack and the kurtosis gradient stack feature a broad peak with an almost flat top indicating that they are almost insensitive to changes in origin time in this part. Similarly, the first arrivals are aligned best for the raw and semblance stack and worst for the STA/LTA stack.

The shape of the maximum in the image function illustrates the spatial resolution of a location estimate. For a fixed origin time, it depends mainly on the geometry of the recording array, the sensitivity of the location method and the width of the input signal. The influence of the recording array can be observed best in the XY-slice of the image function (Fig. 6c). The maximum is smeared in the SW-NE direction for all methods because the BUR array has the smallest spread in the SW-NE direction and is therefore not very sensitive to changes along this direction. The amount of smearing in the NW-SE direction depends a lot on the type of location method. The raw and semblance stack are well focused in this direction (Fig. 6c) because destructive interference penalizes even small misalignments. In contrast, the absolute methods are relatively tolerant towards time shifts that are smaller than the effective wavelength of the input signal. This leads to a strong lateral smearing for the envelope and the STA/LTA stacks that have relatively wide input signals whereas the kurtosis gradient stack exhibits much less smearing because of its very narrow input signal. Overall, the kurtosis gradient stack features the sharpest, but also the most complicated, maximum in the image function whereas the envelope stack produces a very broad but also very smooth maximum. However, it is important to keep in mind that the slices through the image function shown in Fig. 6c illustrate just the resolution for a specific timeslice. As the example of the kurtosis gradient stack with a high spatial but low temporal resolution demonstrates, it is not enough to evaluate just the detection function or just a timeslice of the image function to assess the quality of an event location.
**Figure 6.** Summary of location results for an event (event E1 in Fig. 2) close to the BUR array located using only the P-wave. The rows show different location methods from top to bottom: raw stack, semblance stack, envelope stack, STA/LTA stack and kurtosis gradient stack. The columns show: (a) Detection function (red line) and deviation from the initial location (blue bars). The detection function depicts the relative stack amplitude of the most likely location for each potential origin time (as described in section 3.1). The maximum of the detection function (marked with a red star) corresponds to the most likely origin time of the event. The blue bars show the deviation of the most likely location from the initial location, i.e. the location from the Burtrask catalogue, in x, y and z for each origin time. (b) Data alignment for the P-wave (upper sub-panel) and S-wave (lower sub-panel): The traces are plotted in a 400 ms window centred around the predicted arrival time for the most likely location – origin time pair of the event (red star in (a)) to illustrate how well the first arrivals are aligned. The dashed blue lines mark the stacking lines and the dotted blue lines the semblance calculation window. For the envelope, STA/LTA and the kurtosis gradient stacks the seismic data are overlain in red by their respective signal amplitudes. (c) XY, XZ and YZ slices through the global maximum of the image function (i.e., the origin time corresponds to the red star in (a)).
Comparing stacking-based location methods for microearthquake location

For events that are located further away from the array, the sensitivity to changes in origin time drops quickly when only the P-wave signal is used. This effect occurs both for events located in the strike of the array and perpendicular to it. It is caused by the trade-off between an event’s origin time and its distance to the centroid of the array (e.g., Beskardes et al. 2017). Fig. 7(a) shows an extreme example where all methods feature a broad, almost flat peak in the detection function caused by the trade-off. Even though all methods pick a (more or less) similar location for a given origin time, the results of the different methods completely disagree because subtle changes in the stack amplitude push the maximum of the imaging function to completely different origin times. Moreover, none of the methods pick a time-location pair that is consistent with the S-wave arrival. Including the S-wave signal makes the image function much more sensitive to changes in origin time so that the detection functions obtain a well-defined maximum (Fig. 7b). The results from the different methods agree much better and the deviation from the location by the virtual Burträsk network decreases significantly. The exception is the STA/LTA stack where a very poor S-wave signal corrupts the image function, so that neither the P-wave arrival nor the S-wave arrival are properly matched.

In general, the quality of the S-wave signal varies greatly from strong and coherent (Fig. 7) to weak and diffuse (Figs 6 and 8). The raw and semblance stack are best in detecting and aligning the S-wave arrival. In our case, the absolute methods often misalign (Fig. 9) or, especially in the case of the STA/LTA stack, completely miss (Figs 7 and 8) the S-wave arrival because of the weaker and more scattered signal character. The performance of the absolute methods still improves when including the S-waves because the time – distance trade-off is reduced, but their locations are generally less reliable than the locations from the polarity-sensitive methods.

There are several weak events in the dataset that would be very hard to pick on individual traces, but which can still be visually identified when the whole array is plotted. The kurtosis gradient stack, the semblance stack and the raw stack are best suited to detect weak events because of their high relative stack amplitudes. Among these three, the kurtosis gradient stack has often the highest relative stack amplitudes, but it does not align the signals as well as the raw and semblance stack (Fig. 8).
In summary, the raw and semblance stacks provide high resolution and the best and most stable results for almost all events. The kurtosis gradient stack has the highest resolution, but is much less robust. The STA/LTA and envelope stacks feature the lowest resolution and the least stable results.

4.3 Polarity reversals

We observed hardly any events exhibiting signs of a polarity change over the length of the array. Fig. 9 shows the only unambiguous example with a reversal of the P-wave polarity around station 55. Towards lower stations number, the first arrival is a negative (white) phase and towards higher number it is a positive (black) phase. Close to the reversal point the signal amplitudes are significantly decreased so that it is difficult to identify the exact station at which the reversal occurs.
Comparing stacking-based location methods for microearthquake location

Figure 8. Summary of location results for a weak event (E3 in Fig. 2) located using both the P- and the S-wave. (a) Detection function (red line) and location deviation (blue bars), (b) data alignment for the P- and S-wave: shifted seismic traces and stacking lines (dashed blue) overlain by the signal amplitudes for the absolute methods, (c) slices through the maximum of the image function. For more details see caption of Fig. 6.

The polarity of the S-wave does not change across the stations. When using only the P-wave, the raw stack is clearly stacking along two different phases whereas the semblance stack seems to be less affected (Fig. 9a). However, the envelope and STA/LTA stacks do not align the first arrivals properly, either. If both P- and S-wave are included, the S-wave signal prevents the raw stack from cycle skipping and the P-wave gets aligned reasonably well (Fig. 9b). In this case, the raw and semblance stacks are consistent with the kurtosis gradient stack and align the P-wave significantly better than the envelope and STA/LTA stacks. Reversing the polarity for the stations with a negative phase and re-analysing the event does not change the results significantly (Figs 9c and d). However, this example is biased by residual time shifts that are largest in the area where the polarity is reversed.
Figure 9. Comparison of the P-wave only locations (a) and combined P- and S-wave locations (b) for an event with a polarity reversal around station 55 (E4 in Fig. 2). (c) and (d) re-analysis of the event with swapped polarity for part of the array (only polarity-sensitive methods). Left panels: detection function (red line) and location deviation (blue bars), right panels: data alignment for the P- and S-wave (i.e. shifted seismic traces and stacking lines (dashed blue) overlain by the signal amplitudes for the absolute methods). For more details see caption of Fig. 6.

We tried adding artificial polarity reversals to the P-wave signal on a couple of well aligned events and tested the performance of the polarity-sensitive methods on these events. We did not add any polarity reversal to the S-wave signal because down-weighting the S-wave contribution pushes...
Comparing stacking-based location methods for microearthquake location

the solution towards aligning the P-wave rather than the S-wave, anyways. A more comprehensive test including a proper modelling of both P- and S-wave polarities is beyond the scope of this study. Fig. 10 shows a selection of results from the polarity change tests. The remaining results can be found in the supplements (Figs S2-S5). Not surprisingly, the effect of a polarity reversal depends mostly on the number of stations that are affected. If only a minor number of the stations have reversed polarities, the locations of the raw and semblance stacks are in most cases close to the original locations (Fig. 10 lower panels). If the polarity reversal occurs approximately in the centre of the array, the locations are all clearly biased, but the more distant events seem to be more severely affected than the closer events (Fig. 10 top). Nevertheless, all events remain detectable and the raw and semblance stacks still feature higher relative stack amplitudes than the envelope and STA/LTA stacks. The origin times are in almost all cases very similar to the origin time without polarity reversal because they are primarily constrained by the S-wave signal. Except for the weakest event, the raw stack provides slightly better locations than the semblance stack.

4.4 Uncertainty factors

Fig. 11 illustrates the influence of the origin – time distance trade-off on the preliminary location uncertainty for the event in Fig. 7. If only the P-wave is used, the uncertainty is mostly caused by the time – distance trade-off while the spatial extent of the maximum in the image function for a given origin time is a minor factor (Fig. 11a). Using both P- and S-waves reduces the trade-off significantly so that in Fig. 11b the preliminary uncertainty is mainly due to the spatial extent of the maximum in the image function. Compared to the first case, the extent is larger since the combination of the two signals leads to lateral smearing. Note that the event in Figs 7 and 11 is an extreme case with both good P- and S-wave signal. Events with poor S-wave signal often exhibit a residual trade-off.

To identify factors influencing the uncertainty of the combined P- and S-wave locations, we plotted our preliminary uncertainty estimates extracted from the image function (Fig. 11) against distance from the centroid of the array, event depth, estimated signal strength and azimuth (Fig. 12). In this analysis, we only included ‘detectable’ events. To estimate the signal strength, we used
Figure 10. Location results for the raw stack of two events with artificially reversed polarities. (a) close event (E5 in Fig. 2) and (b) distant event (E2 in Fig. 2). Left panels: detection function (red line) and location deviation (blue bars), right panels: data alignment for the P- and S-wave (i.e. shifted seismic traces and stacking lines (dashed blue) overlain by the signal amplitudes for the absolute methods). For more details see caption of Fig. 6. The polarity distribution scenarios are shown in (c) with normal polarities in red and reversed polarities in blue.

the maximum STA/LTA ratio value in a 200 ms window around the P-wave arrival and averaged it over all stations. There is still a relatively clear correlation between the preliminary uncertainty and the distance to the event whereas neither depth nor estimated signal strength exhibit a pronounced correlation with the uncertainty. Fig. 12 indicates a slight azimuthal dependency, but the azimuth coverage is not sufficient to draw any conclusions.

4.5 Comparison of Burträsk locations

For many events, the difference between our best location for a specific origin time and the location in the Burträsk catalogue shows a consistent behaviour with origin time. The difference is high for times earlier than the preliminary origin time, becomes (almost) zero at the origin time from the Burträsk catalogue and increases significantly again afterwards (Figs 6a and 7a). This illustrates
Comparing stacking-based location methods for microearthquake location

Figure 11. Distribution of all potential source points with a stack amplitude $\geq 95\%$ of the maximum stack amplitude for the distant event (E2 in Fig. 2); (a) for P-wave only, (b) for combined P- and S-wave locations. The red dots show the distribution of locations for a fixed origin time, i.e. they illustrate the lateral extent of the maximum in the image function. The dark blue cross shows the preliminary uncertainty in $x$, $y$ and $z$.

that the locations from our analysis are principally consistent with the locations from the virtual Burträsk network even though they are not necessarily the same because the origin times might still differ. However, for some events there is always a residual deviation and/or the minimum deviation does not correspond to the preliminary origin time (Fig. 8a) indicating that our locations for these events are not completely compatible with the locations from the virtual Burträsk network. Fig 13 shows a comparison of the locations from the virtual Burträsk network and the best locations from our analysis. Close to the BUR array the locations match well but the difference increases with increasing distance to the array. The depth section along the array direction in Fig. 13c shows the locations in relation to the projection of two reflectors (B1 and B2) imaged by Beckel & Juhlin (2019). The locations exhibit good consistency with reflector B1 if only events occurring within a 10 km window around the depth section are considered. The pick-based and the stacking-based locations match the reflector equally well since the difference between them is small in this area.
Figure 12. Factors influencing the location uncertainty of the events: preliminary uncertainty estimates vs. (a) distance from the centroid of the array, (b) event depth, (c) signal strength (see text for details) and (d) azimuth. The total uncertainty is calculated as the Euclidean distance of the uncertainties in x, y and z. Only events with a relative stack amplitude $\geq 2.5$ are included.

5 DISCUSSION

5.1 The importance of pre-processing

The idea of stacking-based source location comes from active seismic studies where considerable effort is spent on signal enhancement because stacking is very efficient in suppressing random noise, but can easily enhance other (semi-)coherent signals such as coda arrivals. Our study, where almost all methods completely failed without noise suppression and balancing, might be an extreme example, but it illustrates that careful pre-processing can be nearly as important for stacking-based location as first arrival picking for traditional, picking-based location.

Recently, the topic has gained more attention and a couple of specialized techniques for noise suppression in microseismic monitoring have been developed. The suggested methods include
Comparing stacking-based location methods for microearthquake location

Figure 13. Comparison between the pick-based locations of the Burträsk catalogue and the stacking-based locations from this analysis: (a) map view, (b) distance between the pick- and stacking-based locations and (c) depth section across the Burträsk fault. Only events within a distance of 55 km from the centroid of the BUR array shown. The error bars in (c) depict the preliminary uncertainty in depth and along the section. Note that this display does not show the uncertainty perpendicular to the section which is often significantly larger than the uncertainty along the section. The empty circles with dotted error bars in (c) correspond to events that have a perpendicular distance of >10 km to the depth section. The dashed grey lines are the projections of two reflectors (B1 and B2) imaged by Beckel & Juhlin (2019).

down-weighting or removal of stations based on their noise level (Trojanowski 2019; Alexandrov et al. 2020), coherent noise removal using multi-channel convolution filters (Trojanowski 2019), curvlet-based noise filtering (Trojanowski 2019) and statistical noise whitening (Birnie et al. 2017). These methods look promising but it is important to consider that filtering can potentially damage the signal. Therefore, the pre-processing should always be targeted to the noise characteristics of the data. Our noise suppression scheme was relatively simplistic since our dataset is mostly dominated by transient noise and lacks stationary noise sources but other settings will likely require different processing approaches.

The same is true for the choice of a characteristic function. The envelope and STA/LTA ratio gave the poorest enhancement of the first arrivals in our study but they might perform completely differently for another dataset with different stations, sources and background noise characteris-
tics. Cesca & Grigoli (2015), for example, mention coda suppression and a narrow peak as the main benefits of the STA/LTA ratio whereas Shi et al. (2018a,b) attribute the failure of the envelope and STA/LTA stack to their failure to accurately identify the first arrivals of P- and S-waves. Therefore, we emphasize here the importance of careful data pre-processing and testing of different characteristic functions to ensure optimal location results.

5.2 Combined P- and S-wave locations

Our results demonstrate that it is problematic to use only P-waves when locating events that did not occur directly below the array since the resolution drops quickly with distance from the array. This enhances the effect of residual time shifts and small amplitude variations and can lead to severely biased locations. As already noted by Grigoli et al. (2013), including the S-waves reduces the time-distance trade-off and improves the locations significantly. Moreover, the S-waves provide additional constraints that can help to reduce the influence of polarity reversals in the P-wave by stabilizing the results for the origin time and decreasing the location bias. The potential success of combining P- and S-waves depends on the quality of the S-wave signal and velocity model, but our results show that it is worthwhile to include even low quality S-wave signals as long as they are not dominated by coda waves. If they are of considerably lower quality than the P-wave signals, it is important to down-weight their contribution in the image function so that they mainly reduce the time-distance trade-off and the position of the event is primarily determined by the higher quality P-waves signal. Otherwise, strong but incoherent S-wave signals can dominate the image function and lead to erroneous locations. Another prerequisite for including S-waves is a sufficiently well-constrained velocity model. A very poor S-wave velocity model can significantly decrease the quality of the locations. Moreover, tests with reducing the station number by a factor of four indicate that having a large number of stations can be beneficial, as well. For events with diffuse S-wave signals (but high overall signal-to-noise ratio), removing stations effectively increases the influence of strong coda arrivals and can result in seriously biased locations because the first arrival of the S-wave is not matched any longer.
5.3 Features and performance of the characteristic functions

Picking a characteristic function for a specific survey is not an easy task and depends on a number of factors. Many investigators prefer the absolute methods to circumvent the polarity problem. However, for our dataset, the polarity-sensitive methods performed much better, both for P-wave only and combined P- and S-wave location procedures. This is partly due to that the bandpass filtered traces provide the most repeatable input signal for both P- and S-waves, making the estimates more stable, and partly due to that they are coherence-sensitive. Destructive interference of signals with opposite polarity suppresses noise very efficiently and increases the resolution. Furthermore, it penalizes misalignment which is especially relevant in cases where local focusing effects cause later arrivals to have higher amplitudes than the first arrivals (Fig. 7 envelope) or where locally unbalanced amplitudes in combination with residual time shifts favour misalignment (Fig. 6 STA/LTA, 9 envelope). We did not attempt any automatic detection, but since the polarity-sensitive methods have high relative stack amplitudes, we anticipate that they would be well suited for automatic detection, as well.

The results of the polarity-sensitive methods are often very similar but in comparison, the semblance stack is more tolerant to unbalanced amplitudes and amplitude outliers. However, it is more severely affected by residual time shifts (Trojanowski & Eisner 2017) and less consistent in timing. The raw stack always picks the peak amplitude of the first arrival whereas the semblance stack sometimes picks the onset of the first arrival and sometimes the peak amplitude. Furthermore, the raw stack requires approximately 2.7 times less computation time (27 minutes vs. 74 minutes for 512000 grid points and 200 origin times on a 6 core, 3.7 GHz processor). Therefore, the raw stack is more suited for applications that aim for a (near) real-time detection whereas the semblance stack might be better for smaller, very noisy datasets.

Our results for the polarity-sensitive methods agree well with with results of a numerical study by Trojanowski & Eisner (2017) and a case study by Beskardes et al. (2017). Both studies conclude that either the raw stack (Beskardes et al. 2017) or the semblance stack (Trojanowski & Eisner 2017) provide the best locations if no polarity reversals are affecting the data. In contrast to Beskardes et al. (2017), we do not observe a pronounced sensitivity to high amplitude noise bursts.
for the raw stack. This might be explained by the fact that we balance the signals and thus reduce the influence of such noise bursts.

However, the results of our dataset cannot be directly transferred to other surveys since it includes very few events with polarity reversals. As discussed before, polarity reversals can lead to severely biased locations. Even though our tests demonstrated that including the S-wave signal can reduce the harmful influence of polarity reversals, the location results remain unstable. Therefore, we would not recommend using a polarity sensitive method if a significant number of events exhibit signs of polarity reversals. Our survey did not have sufficient azimuthal coverage to attempt simultaneous location and moment tensor inversion as suggested by e.g. Anikiev et al. (2014) or Chambers et al. (2014) but for smaller surveys with better azimuthal coverage this is probably the best solution.

If this is not possible, the absolute methods might be better suited since they are insensitive to polarity. However, the lack of destructive interference has the disadvantage of making them more prone to misalign the first arrivals which also results in biased locations. Furthermore, they are much more sensitive to the signal shape, and the signal shape, again, depends a lot on the properties of the characteristic function. The envelope just extracts amplitude information from the seismic trace and has therefore no intrinsic noise suppression. Due to the loss of the phase information, it has a much larger effective wavelength than the original signal/trace resulting in very broad and smooth maxima in the image function. The STA/LTA ratio and the kurtosis gradient extract statistical properties of the seismic trace to detect the P- and/or S-wave arrival. Therefore, they can suppress noise internally and the signal can be narrower than the input waveform but they can also fail to detect the first arrivals (Figs 7 and 8) or falsely detect later arrivals or noise (Fig. 9).

In general, shorter input signals lead to a better focusing of the image function which makes the locations potentially more accurate but also more susceptible to small time shifts or inconsistencies in the signal. In our survey, this is observed for the kurtosis gradient stack. Among the absolute methods, it provided by far the best results due to its sharp signal and efficient noise suppression. However, it picked up later arrivals for several events which led to severely biased lo-
Comparing stacking-based location methods for microearthquake location

Comparing stacking-based location methods for microearthquake location

As expected from the pre-processing results, the envelope and the STA/LTA stack yielded the lowest resolution and, due to insufficient noise suppression, the least stable results. In addition, the envelope stack was most affected by insufficiently balanced stations and featured the lowest relative stack amplitudes which made it the worst overall choice.

The results of the absolute methods are more difficult to compare because other comparative studies come to quite different conclusions. In a study of Cesca & Grigoli (2015) using the energy traces, the envelope, the STA/LTA ratio and the kurtosis (not the gradient) as characteristic functions, the STA/LTA stack provided the best results and the kurtosis gradient stack the worst. In contrast, Beskardes et al. (2017) compared the results from stacking the raw traces, the envelope, the STA/LTA ratio and the kurtosis gradient and concluded that from the absolute methods, the kurtosis gradient provided the best and most stable locations and the STA/LTA ratio the worst. Our results agree best with the results from Beskardes et al. (2017) even though the kurtosis gradient stack was much less robust in our case.

These contradictions can be explained by the differences in survey design, recording instruments, target events and pre-processing applied, resulting in completely different waveforms and noise characteristics to which the characteristic functions respond differently. This highlights again the importance of testing which characteristic function enhances the signals best for each new dataset. Since this is often difficult to assess from the signal shape alone and depends on the array geometry, as well, we would strongly recommend selecting a couple of representative events and testing the performance of different characteristic functions on them.

5.4 Comparison to picking-based locations

A direct comparison between the picking-based locations from the virtual Burträsk network and the stacking-based locations from our survey is difficult because the two arrays have completely different sensitivities. Due to the large number of stations in our survey, we speculate that events occurring close to the BUR array might be better constrained with our method. However, this is difficult to assess without detailed modelling since the better azimuthal coverage, instrument sensitivity and sensor coupling of the virtual Burträsk network might also compensate for the low
number of stations. In any case, the picking-based locations are most likely more precise for events occurring farther to the NE. Nevertheless, the stacking-based location method worked surprisingly well for events occurring within ∼30-40 km of the BUR array considering the suboptimal survey setup and the challenging environment. Our study also illustrates the potential of smaller, dense arrays to locate microseismic events far outside of the area covered by the stations.

5.5 Geometry of the Burträsk fault

Results from the shallow reflection seismic survey showed two southeast dipping reflectors with a potential link to the Burträsk fault (Juhlin & Lund 2011; Beckel & Juhlin 2019). Reflector B1 is relatively short and weak and since it can be connected to a segment of the fault scarp, it has been interpreted as the fault trace. Reflector B2 is much more prominent and continues to greater depth but does not have any direct connection to the fault scarp. So far, it has been interpreted as an old shear zone belonging to the BSZ (Juhlin & Lund 2011; Beckel & Juhlin 2019) but since branching has been observed for other GTFs (Talbot 1986; Juhlin et al. 2010; Ahmadi et al. 2015), another possible explanation is that it represents a branch of the Burträsk fault. Unfortunately, the number of events in our study is too small to draw any real conclusions on the geometry of the seismogenic zone but the fact that the events clearly concentrate around the projection of reflector B1 seems to support the interpretation of reflector B2 as an inactive fault zone. Analysis of the full Burträsk catalogue will reveal more details about the structure of the seismogenic zone and hopefully allow for further conclusions about the nature of the Burträsk fault.

5.6 Uncertainties

One advantage of stacking-based location is that the image function itself provides some information on the quality of the location result. Of course, this is not a statistical uncertainty, but rather an estimation of the effects of data quality, the physical resolution and systematic uncertainties due to the time-distance trade-off, the sensitivity of the recording array and – to some extent – residual time shifts. All of these factors lead to a smearing of the maximum in the image function. However, the preliminary uncertainty does not include factors that lead to a systematic bias of the location
Comparing stacking-based location methods for microearthquake location

like inaccuracies in the velocity model, static time shifts or stacking along different phases. One implication is that the detection and image functions alone are not sufficient to assess the quality of an event location. It is important to take into account how well the signals are aligned, as well. This can be evaluated visually or by calculating a coherence measure, but quantifying the effects this has on the location is, unfortunately, much more difficult. Another implication is that our uncertainty estimates are always relative to the velocity model employed in the analysis.

Some investigators have attempted to estimate the uncertainties with statistical tests, such as Jackknife re-sampling or bootstrapping on sample events (Grigoli et al. 2013; Pesicek et al. 2014; Li et al. 2017, 2019) or studying the effect of random or real seismic noise on the data (Chambers et al. 2010; Beskardes et al. 2017), but these approaches examine mostly the effects of data noise and neither quantify the effects of the velocity model nor the influence of the recording array properly. Our results show clearly that the signal strength has only a minor effect on the estimated uncertainties compared to systematic factors like the distance of the event, the array geometry and the signal resolution. Similarly, Chambers et al. (2010) and Beskardes et al. (2017) noted that the signal-to-noise ratio mainly affects the detectability of the events, but not so much the location. Therefore, we believe that simple statistical tests focusing on data quality are not sufficient and that estimating preliminary uncertainties from the image function is more meaningful at the current stage. However, more work is required to establish a relationship between both methods and develop a comprehensive approach that includes the effects of the velocity and residual (e.g. static) time shifts as well, since both are major contributors to the location uncertainty.

6 CONCLUSIONS

We analysed a local microseismicity dataset acquired on a 2D deployment of vertical component geophones near the glacially triggered Burtrask fault in order to evaluate the location potential of the array. We compared the performance of five commonly used stacking-based location methods with a special focus on proper signal alignment, since misalignment leads to biased locations. The dataset comprises 62 events that occurred mostly to the southeast of the fault scarp. Even though the quality of the S-wave signal was often poor, the locations benefited significantly from including
the S-waves since these provide additional constraints, reducing the origin time – distance trade-off. However, it was important to down-weight the S-wave contribution in the image function so that it primarily influences the origin time estimation and the location estimation is mainly based on the higher quality P-wave signal.

Moreover, our results highlight the importance of noise suppression and signal shaping for the success of stacking-based location methods. Especially the absolute methods, i.e. the envelope, STA/LTA and kurtosis gradient stacks, are very sensitive to signal shape and consistency across the stations. In contrast to the common assumption that the absolute methods provide more robust locations, we found that they performed worse than the polarity-sensitive methods since they frequently misaligned the signals. We attribute this mainly to unfavourable signal shapes, weak S-wave responses and leaking of coda energy in our dataset. A comparison of our results with other case studies indicates that the performance of the different methods depends on the characteristics of the dataset. Therefore, we suggest to always try different pre-processing and stacking methods on a set of manually identified test events for each new dataset to assess which methods isolate the first arrivals best.

We estimated uncertainties using the image function since it encompasses both the effects of data quality and the effects of systematic factors like the physical resolution, the geometry of the recording array and the time-distance trade-off. An analysis of the estimated uncertainties revealed that the sensitivity of the array has a much larger influence than the signal strength, indicating that simple statistical tests focusing on the data quality are not sufficient to describe the uncertainties.

Even though our survey design was not optimal for the task, we obtained good locations for most events within ~ 30–40 km of the survey. The event locations are consistent with a reflection seismic image of the trace of the Burträsk fault obtained in an earlier survey along the same profile. This supports the interpretation of the fault trace as the currently active fault zone and of a second, more prominent and continuous reflector as an inactive fault zone.
ACKNOWLEDGEMENTS

C.J., B.L. and R.B. conceived and planned the fieldwork, R.B. and B.L. carried out the fieldwork. G.E. re-located the events in the Burträsk catalogue. R.B. implemented the stacking-based location method, carried out the analysis and wrote the manuscript. All authors discussed the results and revised the manuscript critically. C.J. supervised the project.

We would like to thank Theo Berthet and Karin Berglund for taking part in planning and conducting the fieldwork and Peter Schmidt for providing the traveltime calculation module. Furthermore, we would like to thank Simone Cesca, Lei Li, Michal Malinowski and an anonymous reviewer for providing constructive reviews and editorial comments that helped to improve the quality of the manuscript.

The Swedish Nuclear Fuel and Waste Management Co. (SKB) provided funding for the initial six station temporary network. The ObsPy toolbox was used for pre-processing the data and calculating the characteristic functions. The figures were prepared using the Matplotlib toolbox and the Generic Mapping Tools (GMT) package.

Data Availability

The data underlying this article are subject to an embargo of 12 months from the publication date of the article. Once the embargo expires the data will be shared on reasonable request to the corresponding author.

REFERENCES


Juhlin, C. & Lund, B., 2011. Reflection seismic studies over the end-glacial Burträsk fault, Skellefteå,
Comparing stacking-based location methods for microearthquake location

Sweden, Solid Earth, 2, 9–16.


Trojanowski, J., 2019. *Advanced migration and filtration techniques for microseismic data*, Doctoral
Comparing stacking-based location methods for microearthquake location

Thesis, Institute of Geophysics - Polish Academy of Sciences, Warsaw, Poland.


