GEOPHYSICS, VOL. 79, NO. 6 (NOVEMBER-DECEMBER 2014); P. KS61–KS68, 4 FIGS., 1 TABLE.
10.1190/GEO2013-0404.1

Picking versus stacking in a modern microearthquake location: Comparison of results from a surface passive seismic monitoring array in Oklahoma

Jeremy D. Pesicek[1], Duncan Child[1], Brad Artman[1], and Konrad Cieślik[1]

ABSTRACT

We present location results for a group of ~200 microearthquakes that occurred in 2012 in a region of Oklahoma hosting ongoing exploration activities. Using a local passive surface seismic monitoring network of 15 broadband stations, we applied two modern location techniques that use fundamentally different approaches. The first is a pick-based double-difference relocation method with waveform crosscorrelation. Multiple-event location techniques such as these are generally regarded as the best approach for obtaining high-precision locations from pick data. The second approach is an automated waveform migration stacking method. These types of methods are becoming increasingly common due to increasing network station density and computer power. The results from the two methods show excellent agreement and provide similar results for the interpreter. Both methods reveal spatial and temporal patterns in the locations that are not visible in results obtained using a more traditional pick-based approach. We performed two statistical uncertainty tests to assess the effects of data quality and quantity on the two methods. We show that the uncertainties for both methods are comparable, but that the stack-based locations are less sensitive to station geometry, likely due to the different treatment of outliers and the beneficial inclusion of noisier data. Finally, we discuss the favorable conditions in which to apply each method and argue that for small aperture surface arrays where accurate velocity information exists, such as in this study, the stack-based method is preferable due to the higher degree of automation. Under these conditions, stack-based methods better allow for rapid and precise determination of microearthquake locations, facilitating improved interpretations of seismogenic processes.

INTRODUCTION

Determining the locations of earthquakes beneath a local seismic network is a fundamental problem in seismology and remains an active area of research. Significant advances have been made in recent years using two fundamentally different approaches. The traditional and still most widely used class of location techniques is based on Geiger's method (Geiger, 1910, 1912) and relies on picks of the arrival times of seismic phases as input data. The picks are compared with those predicted by an available earth model to determine best-fitting hypocentral parameters in a least-squares sense (Thurber, 2011). Almost all automated location algorithms in use today fall into this class, including those that are used to routinely monitor global and regional seismicity (e.g., Buland et al., 2009; Olivieri and Clinton, 2012). Herein, we refer to these methods and their variants as "pick-based" methods.

With pick-based methods, the precision of the picks is a fundamental control on the quality of the resulting event location catalog. Therefore, any effort to increase the quality of the picks, via human interaction or algorithmic efforts in pre- and post-processing, is immediately rewarded. For example, the precision of pick data can be greatly improved by application of advanced automatic pickers (e.g., Nippress et al., 2010) and by the use of waveform crosscorrelation (WCC) techniques (e.g., Rowe, 2002; Du et al., 2004; Schaff, 2004). In addition, the development and widespread use of multiple-event location (MEL) techniques (e.g., Douglas, 1967; Got et al., 1994; Waldhauser and Ellsworth, 2000), usually combined with WCC, has greatly improved the precision of location catalogs,

Manuscript received by the Editor 31 October 2013; revised manuscript received 4 June 2014; published online 11 November 2014.
[1]Spectraseis, Inc., Denver Colorado, USA. E-mail: jeremy.pesicek@spectraseis.com; duncan.child@gmail.com; brad.artman@spectraseis.com; konrad.cieslik@spectraseis.com.

even when using sparse networks (e.g., Rowe et al., 2004; Pesicek et al., 2008; Statz-Boyer et al., 2009). These techniques are the vanguard of modern location methods and can improve location precision by up to two orders of magnitude compared with more traditional single event location (SEL) methods (Richards et al., 2006).

Alternative methods that do not require time picks of seismic phases are becoming increasingly popular for locating earthquakes using local surface seismic networks. These methods exploit full or partial waveforms to rebroadcast the seismic signals to the location in which the energy stack is maximized. We refer to all varieties of these non-pick-based methods as "stack-based" methods. More specifically, we can subdivide them into two categories: (1) those that exploit full waveforms such as time-reversal (TR) methods (e.g., McMechan, 1982; Gajewski and Tessmer, 2005; Artman et al., 2010) and (2) those that are limited to primary seismic phases only, and use some sort of migration or "delay and sum" approach to search for the optimal location (e.g., Kao and Shan, 2004; Baker et al., 2005). All of these stack-based methods are more computationally intensive than pick-based methods. Nevertheless, they are now being routinely applied for many applications. Furthermore, the second class of stack-based methods, those that perform partial stacking of primary phases only, is gaining popularity for real-time monitoring (Baker et al., 2005; Rentsch et al., 2007; Gharti et al., 2010; Grigoli et al., 2013).

The increasing use of relatively dense local seismic arrays, coupled with modern computing capabilities has resulted in increased interest in stack-based methods in recent years. However, the use of such dense arrays makes it more difficult to apply pick-based methods; manually reviewing and improving automated picks for these arrays is an onerous task. These inherent differences in methodology and applicability between the two different classes make it difficult to directly compare their results. For example, using dense seismic networks, TR methods can locate microearthquakes in environments with low signal-to-noise ratios (S/N) where accurate picks cannot reliably be obtained. Although some small subset of the events located using TR methods could also be located using traditional pick-based SEL methods, these few events would likely not be an ideal target for application of more precise MEL methods. Conversely, large event catalogs resulting from pick-based MEL studies may not be amenable to location using TR methods, due to network sparseness, inadequate velocity information, and/or computational demand. However, simplifications made in partial stacking methods (compared with TR methods) make them faster and thus more suitable for broader use. The result is that more overlap now exists between the applicability of pick-based location methods and stacking methods, presenting a choice between microearthquake location techniques that has not existed until recently.

Developers of these new partial stacking methods have conducted exhaustive synthetic location experiments to validate their methods and results (Kao and Shan, 2004; Baker et al., 2005; Gharti et al., 2010; Liao et al., 2012; Grigoli et al., 2013). In addition, limited comparisons of locations produced by stacking methods versus those produced by pick-based SEL methods have also been performed. For example, Grigoli et al. (2013) compare their stack-based locations using P- and S-arrivals with those of a pick-based "manual procedure" using P only, whereas Liao et al. (2012) locate one $M \sim 4.5$ using a stack- and pick-based SEL method. However, no one has yet systematically compared stack-based location results with those obtained using MEL methods with WCC, which is generally regarded as the best approach available for obtaining high-precision relative event locations from pick data. In this paper, we use these two different modern location methods to locate a group of microearthquakes that occurred in a region of Oklahoma hosting exploration activities. We describe the two different methods, compare, and contrast their results, and discuss the advantages, disadvantages, and applicability of each method and the implications for the future of microseismic location studies.

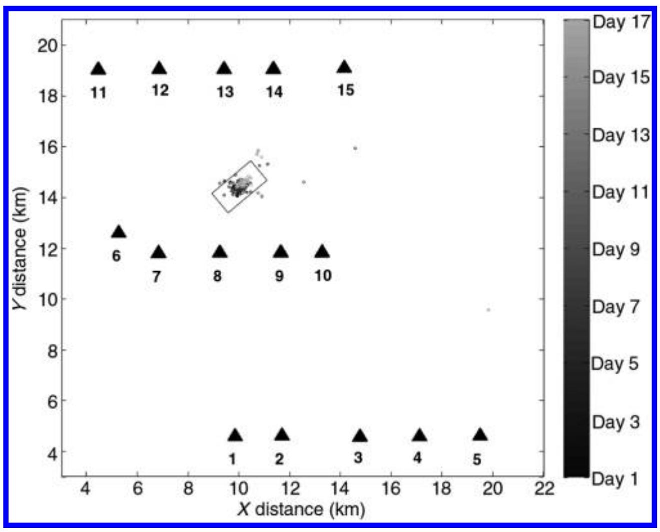

Figure 1. Station (triangles) and event location map for approximately 200 events ($M0$–$M1.7$) located using a pick-based SEL method. Event symbols are shaded by time. The black rectangle is the area shown in Figure 2.

DATA AND VELOCITY MODEL

The earthquakes discussed in this study (~200) comprise a subset of events recorded in the region shown in Figure 1, which hosts ongoing seismic monitoring operations. They occurred over the course of 17 days during the year 2012 and have magnitudes ranging from 0 to 1.7. Only one of the events (the eastern most event in Figure 1 with $M = 1.6$) was reported by the Oklahoma Geological Survey (OGS). However, the largest magnitude event occurring in the box in Figure 1 has a magnitude of 1.7, which was not reported by the OGS. In fact, none of the events shown in the box in Figure 1 were reportedly felt nor were they recorded by any monitoring agencies despite many similar magnitude events having been recorded elsewhere in the state by the OGS.

We recorded these events using 15 3C broadband seismometers deployed in three lines (Figure 1). Two types of instruments were used, both of which have approximately flat transfer functions in the frequency band of interest (120 s to 50 Hz and 40 s to 50 Hz). To compute traveltimes, we used a fast-marching Eikonal solver (Fomel and Alkhalifah, 2001) and a 3D P-wave velocity model from a prestack time migration rms volume (after conversion to interval velocity in depth), which covers the entire area. S-wave traveltimes were computed using a constant V_P/V_S ratio of 1.73. No station corrections were computed or applied.

PICK-BASED METHODS

The general work flow for automated pick-based SEL methods is as follows: (1) an event is detected when the short-term average to long-term average ratio (STA/LTA) exceeds a specified threshold on a subset of stations within a specified time window, (2) an automatic picking algorithm is applied to a window around the detection to improve the STA/LTA onset times, and (3) the difference between the observed and predicted traveltimes (computed using the time picks and available earth model) is minimized to find a preliminary best-fitting location and origin time. Although automated picking algorithms have improved significantly lately (e.g., Nippress et al., 2010), the accuracy and precision of automated picks remains relatively low. Therefore, the automatic picks are usually reviewed and improved a posteriori by a human analyst and the hypocenter is recomputed. For our data, we have implemented this procedure using the algorithm of Oye and Roth (2003). To improve upon these automated solutions, we have invested considerable effort manually reviewing and improving all the picks used in this study. The locations computed using this approach have a final residual rms of 0.50 s.

To improve the precision of the resulting locations, we next applied an MEL technique to relocate the event catalog shown in Figure 1. Of the established MEL algorithms in use today, the double-difference (DD) algorithm *hypoDD* (Waldhauser and Ellsworth, 2000; Waldhauser, 2001) is perhaps the most popular. The DD method minimizes the difference between observed and predicted traveltime differences (i.e., differential or DD times) for pairs of events observed at common stations. This effectively minimizes errors due to unmodeled velocity structure outside the source regions in which the travel paths are highly similar, providing more precise relative event locations. Differential data can be computed directly from the existing picks and/or by using WCC methods (Waldhauser, 2001). In this study, we use catalog-based and WCC-based DD data in the relocation process. To compute more precise differential times, we used the time-domain WCC package named GISMO (Reyes and West, 2011). We crosscorrelated arrivals within 0.6 and 1.0-s windows around the P- and S-picks, respectively, and accepted time delays with CC coefficients above 0.8. A hierarchical dynamic weighting scheme was applied that properly accounts for the different data types and quality (Waldhauser and Ellsworth, 2000; Waldhauser, 2001). We obtained final misfits of 0.004 and 0.019 s for the WCC and catalog-based differential data, respectively, corresponding to 97% and 78% misfit reductions. The relocation results are shown in Figure 2a and 2b and illustrate yet another example of the increased location precision that can be achieved by postprocessing of SEL locations using MEL and WCC methods.

PARTIAL STACKING METHOD

Many studies now exist that aptly describe the development and successful implementation of stack-based methods. For this study, we have developed our own automated partial stacking location algorithm, which we name EMMA (efficient multi-component migration algorithm). The method is based on those of other workers (Kao and Shan, 2004; Gharti et al., 2010; Liao et al., 2012; Grigoli et al., 2013) and is designed for real-time acquisition, detection, and location of microearthquakes in areas hosting exploration and mining activities. However, the focus of this study is to present the comparison of results obtained by pick- and stack-based methods rather than to present the details of another partial stacking method. Therefore, interested readers are referred to Appendix A for a description of EMMA and to the included references. The results of applying EMMA to the same events located using pick-based methods are shown in Figure 2c.

UNCERTAINTY ESTIMATES

Automated pick-based SEL methods commonly used for real-time monitoring (e.g., Oye and Roth, 2003; Buland et al., 2009; Olivieri and Clinton, 2012) provide formal uncertainty estimates for each event by applying standard error formulas based on the precision of the pick data. Such methods for estimating uncertainty are not as easily implemented for large MEL problems and are not applicable to stack-based optimization methods. Instead, we have conducted two commonly used statistical resampling tests, namely the jackknife and the bootstrap (e.g., Efron, 1982; Waldhauser and Ellsworth, 2000; Pesicek et al., 2010; Grigoli et al., 2013), which we have implemented for the DD method and our partial stacking method. For the jackknife test, we recomputed the location for each event N times, where N is the total number of stations used in the preferred location. For the nth location, we removed data from the

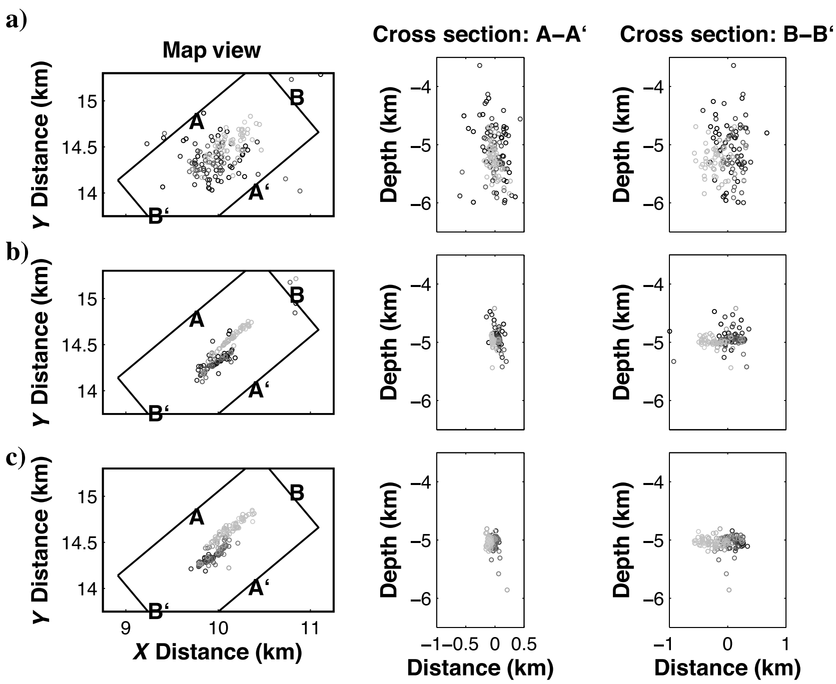

Figure 2. (a) Microearthquake locations using a traditional SEL method, (b) relocations of the SEL events using DD and WCC methods, and (c) locations produced using our migration-stacking method. Refer to Figure 1 for additional details.

*n*th station and recomputed the location. We then used these relocations from the data subsamples to estimate data covariance matrices and confidence regions (e.g., Aster et al., 2013) for both the location catalogs (Figure 3). The mean and standard deviations for the data subsample solutions are listed in Table 1.

Not surprisingly, the jackknife results confirm for both methods that the locations are better constrained in epicenter than focal depth. The results also reveal that the epicentral confidence ellipse for the EMMA solutions is elongated in the same direction (northeast–southwest) as the real locations (Figure 2c). Although there are multiple faults in the region that also trend northeast–southwest, this pattern might suggest that some of the elongation of the real locations is artificial and might be due to the network geometry. The same pattern is not apparent in the DD jackknife results, however, perhaps because it is obscured by the larger deviations obtained with this method.

The results in Figure 3 provide insight into the behavior of the two methods with respect to station geometry and data quantity. It is clear from the results that the DD locations are more sensitive to missing stations than the EMMA locations. This effect can be understood by the differing nature of the two methods and the reliance of the DD method on pick data. For microearthquakes, reliable picks are harder to obtain on noisy stations. However, waveform data from these noisy stations are still used by EMMA and directly contribute to maximizing the energy stack. The result is that pick-based methods generally have fewer stations contributing to the solution, making each observation more critical and increasing the negative effects of removing any one station. Said another way, the jackknife simulations more negatively impact the pick-based solutions by excluding a larger percentage of the original data. The result is that there are more observations used by EMMA, providing improved stability of the solutions with respect to station geometry.

For the same reason, this test may also be illustrating the differing effects of data and/or model errors between the two methods. Removing an observation during the jackknife test reduces the number of constraints on the solution, effectively upweighting outliers. In pick-based methods, large outliers can result from erroneous picks or from forwarding modeling bias due to systematic velocity model errors. The least-squares solutions assume Gaussian data distributions, but it is well known that traveltime residuals are typically not Gaussian, due to such outliers (e.g., Buland, 1986). The result is that the pick-based least-squares solutions can be significantly influenced by these outliers (e.g., Aster et al., 2013), and we observe outlier residuals up to 3.16 s. These issues manifest differently in stack-based methods and may partly explain the disparity in the jackknife results. Removing a station during stacking will change the strength of the stack maximum but is less likely to change its location. Similarly, large systematic velocity model errors that may not be accounted for in the EMMA stacking window (0.1 s; see Appendix A) will serve to reduce the stack strength. In these cases, such errors contribute more to event detectability rather than location uncertainty. Detectability (see also "Discussion" section) is a separate issue that has not influenced our jackknife results because we have limited our analysis to easily identifiable events whose stack maximum is well above our chosen detection threshold. In this regard, the pick-based jackknife uncertainty estimates are more affected by model errors than the stack-based estimates.

The jackknife test has been applied in essentially the same way to both the location techniques, providing a direct comparison of the sensitivity of the techniques with variations in data distribution. However, the inherent differences in the location methods have the effect that the two methods are affected unequally, as discussed above. To provide an alternative independent estimate of uncertainty, we have also conducted a bootstrap analysis, designed to assess the effects of noise in the data for both the methods in a more equitable manner. For pick-based methods, noise is often added based on the data misfit of the preferred location (e.g., Shearer, 1997; Waldhauser and Ellsworth, 2000). However, no misfit criterion is used in stack-based methods. Furthermore, data noise is a function of time for pick-based methods but is a function of amplitude for stack-based methods, complicating the comparison. Nevertheless, assessing the effects of noise in the data may provide valuable insight into the quality of the locations for each method. To simulate the effects of noise on the pick-based DD catalog, we added noise to the traveltimes from a random Gaussian distribution with standard deviation equal to 0.02 s, which is representative for our data. We recomputed the locations 100 times, each time drawing noise from the random distribution and again used the results to estimate location uncertainty. To assess the effects of noise on our stacking method, we computed the average noise level of all the traces prior to the event and added it to the amplitudes of the traces containing the event. We repeated this process 100 times to obtain 100 independent location estimates.

Figure 4 and Table 1 show the results of the bootstrap tests. The two methods are generally affected similarly by noise in the data. The confidence regions are larger for the EMMA results than the DD results, but this difference is relatively small. Interestingly, even though the DD confidence regions are smaller than those of EMMA, the mean deviations (Table 1) are

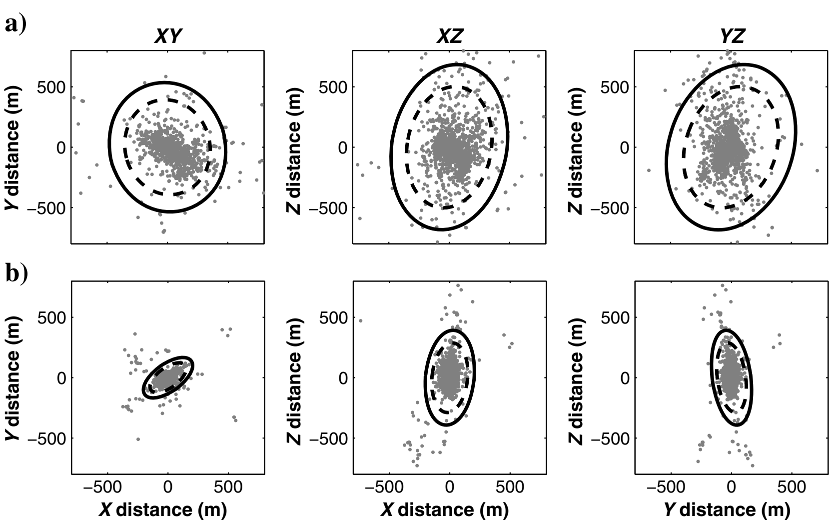

Figure 3. Results of the jackknife uncertainty analyses. Panel (a) shows results of the analyses applied to the DD relocation results, and panel (b) shows the results using our migration stacking method. The 95% and 80% error ellipses are determined using 3000 trials (1 trial for each of 15 stations× approximately 200 events).

slightly larger. This point illustrates the effect of the trial solution outliers (not shown in Figure 4), which are larger for EMMA. These outlier solutions are likely illustrating the rare cases in which the addition of noise produces a new global maximum in a different location rather than slightly altering the location of the preferred maximum. In general, the bootstrap tests illustrate the similar sensitivity of both the methods to the presence of random data noise.

DISCUSSION

Precision and accuracy of earthquake locations depends on many factors, including station geometry, velocity model errors, and data noise. In this study, we have applied two different methodologies to the same real data set with the goal of comparing and validating each method's results and estimating each event's location uncertainty. As is true for all real earthquakes, location error remains unknown and we are limited to only assessing location uncertainty due to data noise, which we can accurately estimate. We cannot accurately assess location errors, such as those due to velocity model error and array geometry because such effects cannot be known for these events. Furthermore, we have not considered uncertainty associated with event detectability and false positives (e.g., Thornton and Eisner, 2011); we have only discussed high-quality events that are easily identifiable to the analyst and far above our detection thresholds. In this regard, the uncertainties discussed in this study only inform us about real event precision and they provide no information about event accuracy. Although beyond the scope of this work, many studies have investigated such effects in various ways with either synthetic data or explosions with known locations, for pick- (e.g., Michelini and Lomax, 2004; Lin and Shearer, 2005; Bondar and McLaughlin, 2009) and stack-based methods (e.g., Kao and Shan, 2004; Gharti et al., 2010; Liao et al., 2012; Grigoli et al., 2013).

Despite these limitations, the comparison between location results that we have presented illustrates the capability of pick- and stack-based methods to precisely determine microearthquake locations (Figure 2) using a local surface seismic array (Figure 1). Our partial stacking method and the DD method reveal similar spatiotemporal structure in the locations that is not visible in the locations determined using the SEL method (Figure 2). However, there are significant differences inherent in the two approaches that warrant further discussion. First, assuming for a moment that the design of our network is ideal for locating earthquakes and that our velocity model is perfect, the most important difference between the two methods is the difference in the amount of human effort invested to obtain the data that are required for each algorithm. For our automated stacking algorithm, only minimal automated processing of the raw trace data is performed — no analyst grooming of the data is required. Although automated earthquake locations are routinely computed using pick-based methods, significant additional effort is required to achieve the higher level of precision obtained by the use of MEL and WCC methods. These methods are always applied as postprocessing after the initial efforts to obtain the picks and preliminary locations using an SEL method. This postprocessing usually involves considerable additional human and computational effort to test and optimize the many parameters involved. Although MEL methods can potentially be applied rapidly following an event ("near real-time"; Waldhauser, 2009) in some cases, such applications are inherently limited to regions in which large archives of historical seismicity are already available for rapid correlation and comparison to new events. In areas of new scientific or industrial exploration, no such data exist. Thus, in many cases in which rapid and precise hypocenters are required beneath a local array, a partial stacking method such as EMMA can outperform other methods.

Unfortunately, the benefits of the higher degree of automation of stacking methods can have drawbacks in some circumstances. In our comparison, we have used the same 3D velocity model for both the location methods and have so far neglected any discussion of the effects of unmodeled velocity heterogeneity. For stacking methods,

Table 1. Results of uncertainty tests: Mean and standard deviations of the trial locations from the preferred locations.

		Jackknife			Bootstrap		
	Direction	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
DD	Mean	106	93	122	40	32	84
EMMA		25	20	57	24	27	90
DD	Standard deviation	193	217	279	125	85	234
EMMA		83	68	161	79	165	271

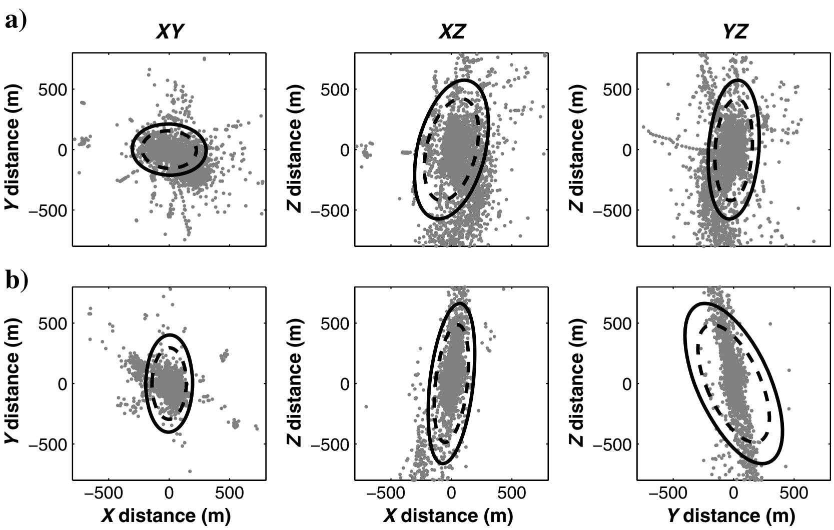

Figure 4. Results of the bootstrap uncertainty analyses. (a) Shows results of the analyses applied to the DD relocation results, and (b) shows the results using our migration stacking method. The 95% and 80% error ellipses are determined using approximately 20,000 trials (100 trials × approximately 200 events). See the text and Table 1 for additional details of the tests.

the reference velocity model is generally obtained independently of the seismic data used for location purposes, as we have done using a priori information. In contrast, for pick-based methods, the velocity model is often constructed and/or improved using the picks from the acquired data. In cases without sufficient preexisting velocity information, generating picks may be unavoidable. Furthermore, the accuracy of the velocity model is more critical for stacking methods than picking methods. Inaccurate traveltime prediction will negatively impact both location methods, but pick-based methods will always produce a best-fitting location for each event. In contrast, inaccurate traveltime prediction can prevent localization for stack-based methods altogether by inhibiting focusing of the stacked energy to a single clear, preferred solution.

In addition to the availability of accurate velocity information, another factor that will influence the choice of location method is the size of the search volume and the density of stations. For stack-based methods, increasing the area being monitored and/or the station density comes at the expense of increased computational burden. Although this is also true for automated pick-based SEL methods, the effect is much greater for stack-based optimization methods, due to the additional scanning in space and time required by larger aperture arrays. At some scale, the computational burden may become too great and pick-based methods may be preferred. In any case, future advances in computer power will surely facilitate expanded applicability of stack-based methods.

Another important difference between pick- and stack-based methodologies is their approach to event detection. In this paper, we have limited our analysis to events that were easily detected by both the methodologies, in an effort to make appropriate and direct comparisons of the location results from both methods. Thus, we focused on a very limited subset of the catalog of events recorded in the area. Here, we note that EMMA was able to detect and locate roughly twice as many events over the study period compared with the pick-based approach (see also Bardainne et al., 2009). Similar to the jackknife results, this can be attributed to the beneficial inclusion of noisier data and the power of stacking to increase S/N. However, this uplift in detectability depends on the accuracy of the velocity model, as previously discussed. In addition, detected events with lower S/N lie closer to the detection threshold, often have reduced precision, and require additional consideration to avoid false positives (e.g., Thornton and Eisner, 2011). Nonetheless, in our study area (Figure 1), the stack-based method (EMMA) proved faster, less user intensive, and detected many more events for the same time period. The final EMMA catalog for the entire data set is providing significantly more information for interpretation.

Finally, the two classes of location methods do not have to be viewed as completely independent. Liao et al. (2012) state that their stacking method is not designed as a replacement for pick-based MEL methods. Instead, they suggest a complementary role in which their method serves as an alternative or replacement for the detection and SEL steps prior to application of more precise MEL methods. Alternatively, Rentsch et al. (2007) require preliminary automated picks for appropriate time windowing, but then apply a stack-based method for location. More recently, Drew et al. (2013) develop another hybrid approach. Clearly, there is opportunity for synergy between these two disparate methods. For our data, the EMMA algorithm achieved essentially equivalent results to the DD method but with much less human interaction. Thus, we have shown that stacking methods can in some cases serve as full replacements for SEL and MEL methods. However, even in such cases, the use of both methods in concert with one another is beneficial. By applying both methods with similar results, we have achieved greater confidence in the results of each. This type of approach may help mitigate potential problems inherent in either methods. For example, it has been shown that MEL methods can produce spurious results in some circumstances (Michelini and Lomax, 2004). Presumably, this is true for migration methods as well, but due to their relative novelty, less is known about their limitations. Thus, a strategy employing both methods may be preferred, leaving little doubt in the results when agreement is achieved (Figure 2) or leading to better understanding of the shortcomings of the methods and/or the data when it is not.

CONCLUSIONS

We have conducted a direct comparison of microearthquake locations obtained using a pick-based MEL method and a partial stacking method from surface seismic data recorded at a site in Oklahoma hosting ongoing exploration activities. We find excellent agreement between the results and show that both methods reveal spatiotemporal structures in the locations that were not visible in the results obtained using a traditional SEL method. We have assessed location uncertainty due to station geometry and find the results of the stack-based method to be more stable, due in part to the different handling of outliers and to the additional constraints provided by noisy stations for which reliable picks cannot be obtained. We also assessed uncertainty due to data noise and find that both methods are affected similarly. We conclude that these methods provide similar results for the interpreter. However, we argue that for small aperture local surface arrays with detailed 3D velocity information available, such as in this application, stack-based methods are preferable to pick-based methods due to their comparable precision but higher automation, better facilitating rapid high-quality results for interpretation.

ACKNOWLEDGMENTS

We thank R. Habiger and C. Thurber for their comments on an earlier version of the manuscript. We also thank the editors and several anonymous reviewers for their comments, which helped improve the quality of the manuscript.

APPENDIX A

EFFICIENT MULTICOMPONENT MIGRATION ALGORITHM

We have developed an efficient multicomponent migration algorithm (EMMA). The method is designed for real-time acquisition, detection, and location of microseismicity in areas hosting exploration and mining activities. Following Kao and Shan (2004), we refer to our objective function as the "brightness" function and define it for a particular origin time τ and location r as

$$b(r,\tau) = \sum_{n=1}^{N} u_n(\tau + t_{r,n}), \qquad \text{(A-1)}$$

where N is the number of recording stations, u_n is the seismogram recorded at station n, and $t_{r,n}$ is the predicted traveltime of a given seismic phase from the assumed source location r to station n. We band-pass and normalize u_n and use the envelope to compensate for the radiation pattern (Gharti et al., 2010), and refer to the resulting seismogram as $\hat{u}_n$. To compensate for errors in the predicted traveltimes, equation A-1 is modified to incorporate contributions within a time window of size $2M$ centered on the predicted traveltime

$$b(r,\tau) = \sum_{n=1}^{N} \left\{ \sum_{m=-M}^{M} W_m \hat{u}_n(\tau + t_{r,n} + m\delta t) \right\}, \qquad \text{(A-2)}$$

where δt is the sampling rate and W_m is the weighting factor (between 0 and 1) for sample m that depends on the distance from the predicted arrival time in the center of the window. To include P and S arrivals and assuming that P and S energies are confined to the vertical and horizontal components, respectively, we modify equation A-3 as

$$\begin{aligned} b(r,\tau) = \sum_{n=1}^{N} \Bigg\{ \sum_{m=-M}^{M} W_m [& w_Z \hat{u}_n^Z(\tau + t_{r,n}^{\mathrm{P}} + m\delta t) \\ & + w_Y \hat{u}_n^Y(\tau + t_{r,n}^{\mathrm{S}} + m\delta t) \\ & + w_X \hat{u}_n^X(\tau + t_{r,n}^{\mathrm{S}} + m\delta t)] \Bigg\}, \end{aligned} \qquad \text{(A-3)}$$

where $t_{r,n}^{P,S}$ are now the predicted traveltimes for the P- and S-waves, respectively and $\hat{u}_n^{X,Y,Z}$ and $w_{X,Y,Z}$ are the preprocessed seismograms and weights for the east, north, and vertical components. We use equation A-3 for event detection and location. For detection, we decimate the seismic data and then search for and save the location corresponding to the maximum brightness for every sample. For each minute, we then compute the STA and LTA of the brightness function and detect an event when the STA/LTA ratio exceeds a chosen threshold. If an event is detected, the process is repeated on the undecimated data around the preliminary location to determine the final hypocenter. To make the algorithm fast enough for real-time monitoring, we forego the brute force full-grid search and instead use a global optimization algorithm called the covariance matrix adaptation evolution strategy (Hansen et al., 2003).

REFERENCES

Artman, B., I. Podladtchikov, and B. Witten, 2010, Source location using time-reverse imaging: Geophysical Prospecting, **58**, 861–873, doi: 10.1111/j.1365-2478.2010.00911.x.

Aster, R. C., B. Borchers, and C. H. Thurber, 2013, Parameter estimation and inverse problems, 2nd ed.: Academic Press.

Baker, T., R. Granat, and R. W. Clayton, 2005, Real-time earthquake location using Kirchhoff reconstruction: Bulletin of the Seismological Society of America, **95**, 699–707, doi: 10.1785/0120040123.

Bardainne, T., E. Gaucher, F. Cerda, and D. Drapeau, 2009, Comparison of picking-based and waveform-based location methods of microseismic events: Application to a fracturing job, 79th Annual International Meeting, SEG, Expanded Abstracts, 1547–1551.

Bondar, I., and K. McLaughlin, 2009, Seismic location bias and uncertainty in the presence of correlated and non-Gaussian travel-time errors: Bulletin of the Seismological Society of America, **99**, 172–193, doi: 10.1785/0120080922.

Buland, R., 1986, Uniform reduction error analysis: Bulletin of the Seismological Society of America, **76**, 217–230.

Buland, R. P., M. Guy, D. Kragness, J. Patton, B. Erickson, M. Morrison, C. Bryon, D. Ketchum, and H. Benz, 2009, Comprehensive seismic monitoring for emergency response and hazard assessment: Recent developments at the USGS National Earthquake Information Center, Eos — Transactions of the American Geophysical Union, **90**, S11B-1696, Fall Meeting Supplement, Abstract

Douglas, A., 1967, Joint epicentre determination: Nature, **215**, 47–48, doi: 10.1038/215047a0.

Drew, J., R. S. White, F. Tilmann, and J. Tarasewicz, 2013, Coalescence microseismic mapping: Geophysical Journal International, doi: 10.1093/gji/ggt331.

Du, W.-X., C. H. Thurber, and D. Eberhart-Phillips, 2004, Earthquake relocation using cross-correlation time delay estimates verified with the bispectrum method: Bulletin of the Seismological Society of America, **94**, 856–866, doi: 10.1785/0120030084.

Efron, B., 1982, The jackknife, the bootstrap and other resampling plans: SIAM.

Fomel, S., and T. Alkhalifah, 2001, Implementing the fast marching eikonal solver: Spherical versus Cartesian coordinates: Geophysical Prospecting, **49**, 165–178.

Gajewski, D., and E. Tessmer, 2005, Reverse modelling for seismic event characterization: Geophysical Journal International, **163**, 276–284, doi: 10.1111/j.1365-246X.2005.02732.x.

Geiger, L., 1910, Herdbestimmungen bei Erdbeden aus den Ankunftzeiten: Königlichen Gesellschaft der Wissenschaften Zu Göttingen, **4**, 331.

Geiger, L., 1912, Probability method for the determination of earthquake epicenters from the arrival time only: Bulletin of St. Louis University, **8**, 60.

Gharti, H., V. Oye, M. Roth, and D. Kühn, 2010, Automated microearthquake location using envelope stacking and robust global optimization: Geophysics, **75**, no. 4, MA27–MA46, doi: 10.1190/1.3432784.

Got, J.-L., J. Fréchet, and F. W. Klein, 1994, Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea: Journal of Geophysical Research, **99**, 15375–15386, doi: 10.1029/94JB00577.

Grigoli, F., S. Cesca, M. Vassallo, and T. Dahm, 2013, Automated seismic event location by travel-time stacking: An application to mining induced seismicity: Seismological Research Letters, **84**, 666–677, doi: 10.1785/0220120191.

Hansen, N., S. D. Müller, and P. Koumoutsakos, 2003, Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES): Evolutionary Computation, **11**, 1–18, doi:10.1162/106365603321828970.

Kao, H., and S.-J. Shan, 2004, The source-scanning algorithm: Mapping the distribution of seismic sources in time and space: Geophysical Journal International, **157**, 589–594, doi: 10.1111/j.1365-246X.2004.02276.x.

Liao, Y.-C., H. Kao, A. Rosenberger, S.-K. Hsu, and B.-S. Huang, 2012, Delineating complex spatiotemporal distribution of earthquake aftershocks: An improved source-scanning algorithm: Geophysical Journal International, **189**, 1753–1770, doi: 10.1111/j.1365-246X.2012.05457.x.

Lin, G., and P. Shearer, 2005, Tests of relative earthquake location techniques using synthetic data: Journal of Geophysical Research: Solid Earth, **110**, B04304, doi: 10.1029/2004JB003380.

McMechan, G. A., 1982, Determination of source parameters by wavefield extrapolation: Geophysical Journal International, **71**, 613–628, doi: 10.1111/j.1365-246X.1982.tb02788.x.

Michelini, A., and A. Lomax, 2004, The effect of velocity structure errors on double-difference earthquake location: Geophysical Research Letters, **31**, L09602, doi: 10.1029/2004GL019682.

Nippress, S. E. J., A. Rietbrock, and A. E. Heath, 2010, Optimized automatic pickers: Application to the ANCORP data set: Geophysical Journal International, **181**, 911–925, doi: 10.1111/j.1365-246X.2010.04531.x.

Olivieri, M., and J. Clinton, 2012, An almost fair comparison between Earthworm and SeisComp3: Seismological Research Letters, **83**, 720–727, doi: 10.1785/0220110111.

Oye, V., and M. Roth, 2003, Automated seismic event location for hydrocarbon reservoirs: Computers & Geosciences, **29**, 851–863, doi: 10.1016/S0098-3004(03)00088-8.

Pesicek, J. D., C. H. Thurber, H. R. DeShon, S. G. Prejean, and H. Zhang, 2008, Three-dimensional P-wave velocity structure and precise earthquake relocation at Great Sitkin Volcano, Alaska: Bulletin of the Seismological Society of America, **98**, 2428–2448, doi: 10.1785/0120070213.

Pesicek, J. D., C. H. Thurber, H. Zhang, H. R. DeShon, E. R. Engdahl, and S. Widiyantoro, 2010, Teleseismic double-difference relocation of earthquakes along the Sumatra-Andaman subduction zone using a 3D model: Journal of Geophysical Research, **115**, 115(B10), doi: 10.1029/2010JB007443.

Rentsch, S., S. Buske, S. Lüth, and S. Shapiro, 2007, Fast location of seismicity: A migration-type approach with application to hydraulic-fracturing data: Geophysics, **72**, no. 1, S33–S40, doi: 10.1190/1.2401139.

Reyes, C. G., and M. E. West, 2011, The waveform suite: A robust platform for manipulating waveforms in MATLAB: Seismological Research Letters, **82**, 104–110.

Richards, P., F. Waldhauser, D. Schaff, and W.-Y. Kim, 2006, The applicability of modern methods of earthquake location: Pure and Applied Geophysics, **163**, 351–372, doi: 10.1007/s00024-005-0019-5.

Rowe, C., C. Thurber, and R. White, 2004, Dome growth behavior at Soufriere Hills Volcano, Montserrat, revealed by relocation of volcanic event swarms, 1995–1996: Journal of Volcanology and Geothermal Research, **134**, 199–221, doi: 10.1016/j.jvolgeores.2004.01.008.

Rowe, C. A., 2002, An automatic, adaptive algorithm for refining phase picks in large seismic data sets: Bulletin of the Seismological Society of America, **92**, 1660–1674, doi:10.1785/0120010224.

Schaff, D. P., 2004, Optimizing correlation techniques for improved earthquake location: Bulletin of the Seismological Society of America, **94**, 705–721, doi: 10.1785/0120020238.

Shearer, P. M., 1997, Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence: Journal of Geophysical Research: Solid Earth, **102**, 8269–8283, doi: 10.1029/96JB03228.

Statz-Boyer, P., C. Thurber, J. Pesicek, and S. Prejean, 2009, High precision relocation of earthquakes at Iliamna volcano, Alaska: Journal of Volcanology and Geothermal Research, **184**, 323–332, doi: 10.1016/j.jvolgeores.2009.04.016.

Thornton, M., and L. Eisner, 2011, Uncertainty in surface microseismic monitoring: 81st Annual International Meeting, SEG, Expanded Abstracts, 1524–1528.

Thurber, C. H., 2011, Earthquake location techniques, *in* H. K. Gupta, ed., Encyclopedia of solid earth geophysics: Springer, 201–206.

Waldhauser, F., 2001, HypoDD: A computer program to compute double-difference earthquake locations, USGS Open File Rep., 01–113.

Waldhauser, F., 2009, Near-real-time double-difference event location using long-term seismic archives, with application to Northern California: Bulletin of the Seismological Society of America, **99**, 2736–2748, doi: 10.1785/0120080294.

Waldhauser, F., and W. L. Ellsworth, 2000, A double-difference earthquake location algorithm: Method and application to the Northern Hayward Fault, California: Bulletin of the Seismological Society of America, **90**, 1353–1368, doi: 10.1785/0120000006.