Picking far-offset arrival times on common-source gathers

Joe Wong

ABSTRACT

A processing flow that uses a modified energy ratio (MER) attribute is effective for automatically picking first-arrival times near the hyperbolic apexes of common-source gathers. At the flanks of the hyperbolas, where the source-receiver offsets are large and the signal-to noise ratios of the first arrivals are small, a signal-enhancement technique employing three-trace summation along a range of time-space slopes must be used to increase signal-to-noise ratios before MER picking. Even after signal enhancement, many of the MER picked times at far source-receiver offsets will be outliers. However, when source-receiver offsets are large, arrivals on common-source gathers tend to have time moveouts that are almost linear with receiver position. A Radon transform helps to identify slowness and time-intercept values that most closely match the observed linear trends. Outlier MER picked times will have large deviations from the linear trends identified by the Radon transform. These outliers can be eliminated, and the remaining times and their receiver positions can be fitted with least-squares straight lines to provide interpolated times for receivers with missing or rejected MER times.

INTRODUCTION

First-arrival time-picking on seismic data has been well-studied for many years (Willis and Toksöz, 1983; Coppens, 1985; Boschetti et al., 1996; Withers et al., 1998; Baziw, 2001; Munro, 2004; Qiao and Bancroft, 2010; Chen and Stewart, 2005). This fundamental task is ubiquitous in seismic processing, and in particular it provides the arrival times required by refraction analysis for determining the velocities and thicknesses of near-surface low-velocity zones. Similarly, first-arrival times from full-waveform sonic well logs and VSP seismograms are needed to calculate interval velocities in a well. Arrival times are also required for locating focal points of global earthquakes and microseisms caused by hydraulic fracturing.

High-resolution 3D surveys may produce hundreds of millions or even billions of individual seismograms. For such large datasets, it is impossible to pick all the first-arrival times manually. Picking must be done automatically by computer software. Because speed is a prime consideration when picking hundreds of millions of first-arrival times, it is imperative that the computer-based decisions used in automatic picking be made as fast as possible. This means that simple calculations with few floating point operations must be given precedence over more involved calculations. In addition, quality control (QC) should be applied to time picks to identify which of the very many time picks are reliable, regardless of whether they are made by humans or by machines.

A standard method widely used for identifying valid seismic events and picking arrival times is based on short-term average to long-term average (STA/LTA) ratios of seismogram energies (Withers, 1991; Munro, 2005). Han et al. (2009) proposed the modified energy ratio (MER) attribute as an alternative to the STA/LTA ratio. Both types of energy ratios are simple and straightforward to calculate, and so are suitable for rapid automated time-picking. Han (2010) compared the effectiveness of techniques based on the two types of ratios for arrival-time picking on high-noise microseismic data. She found that, for both field and synthetic microseismograms, STA/LTA ratios and MER attributes were equally effective in picking first-break times on traces with good signal-to-noise (SNR) levels. However, for noisy seismograms, MER time-picking yielded more consistent results, and was significantly faster.

Quantitative estimates of the signal-to-noise ratios (SNRs) associated with the picked arrivals are an important by-product of MER time-picking. These estimates can be used for QC: the higher the SNR value, the more reliable the picked first-arrival time will be. For example, we might retain MER picked times if their SNRs are greater than 10dB, and reject those with SNRs less than 10dB. Formulas and other details about the MER attribute are given in Appendix A.

MER PICKING ON COMMON-SOURCE GATHERS

The common-source gathers (CSGs) of seismograms plotted on Figure 1 were acquired as part of a large 3D land survey. Reflections from subsurface geological boundaries are visible, but this article focuses on the first or early arrivals. The MER-based procedure employed to obtain the initial estimates of first-break times shown on Figure 1 is outlined in Appendix B. The initial picks appear to be very good estimates of first-arrival times for receivers near the source positions of the CSGs. However, they become erratic and less reliable for the noisy traces at large source-receiver offsets. The results of Figure 1 suggest that automatic picking of first-arrival times using the MER attribute is straightforward for most traces, but that extra effort is needed to improve the automatic picks for the high-noise traces at far source-receiver offsets.

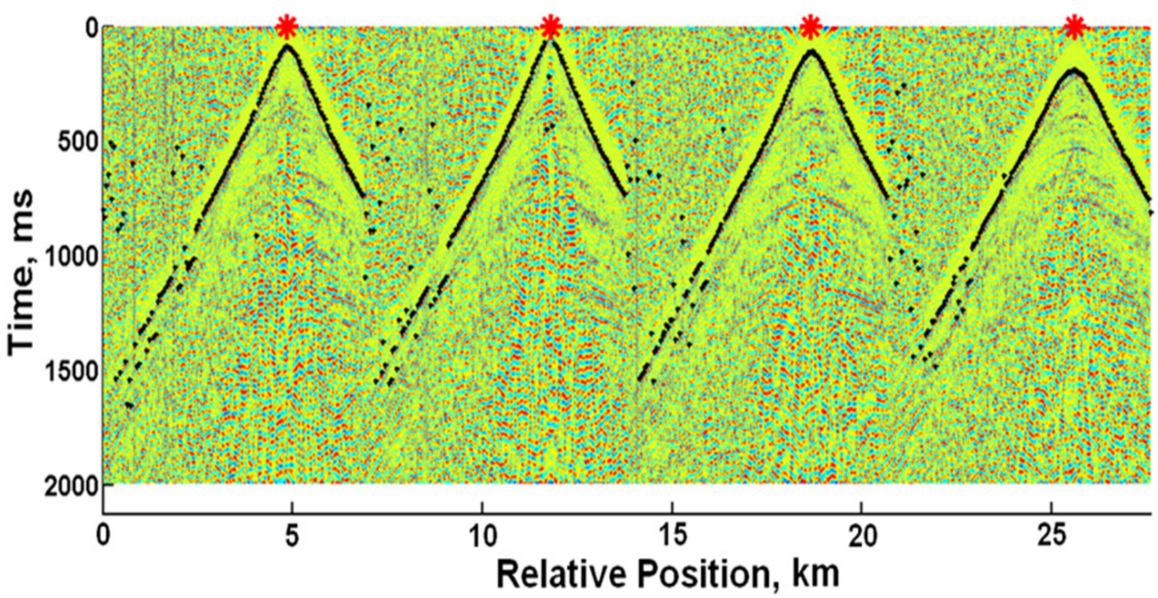

FIG. 1: AGC plot of four common-source gathers. Black crosses are initial MER time picks. Receiver spacing for the gathers is 50m. The red stars indicate the positions of the sources. Although the seismograms are plotted with AGC, the time picking was done on normalized traces.

Figure 2 is an expanded plot of one of the CSGs from Figure 1 showing each trace normalized by its maximum value. The first-break times obtained by an initial application of the MER-based picking procedure are indicated by the red crosses.

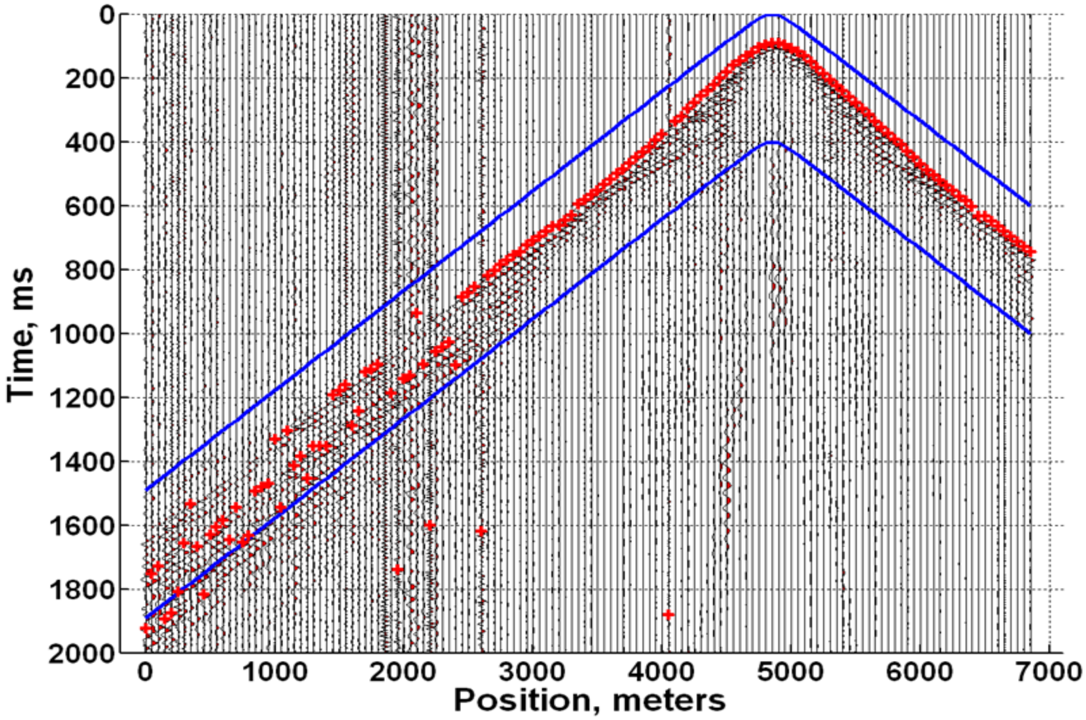

FIG. 2: Trace-normalized common source gather. Red crosses are first-break times t_{B1} from an initial application of MER first-break time-picking. Blue lines are the limits of a retention corridor (width = 400ms, or about 10 to 15 dominant periods).

On the time scale at which Figure 2 is displayed, we cannot see any great detail about how closely the picked times approach the first breaks on the seismograms. In order to see more detail, we generated a reduced-seismogram gather by windowing the original traces. The first arrivals on Figure 2 roughly follow a normal moveout (NMO) or hyperbolic trajectory. We estimated a velocity v_0 and a time t_0 to find a suitable hyperbolic trajectory to fit the picked times with good SNR values using the standard equation for calculating NMO times:

$$t_r = \sqrt{(t_0^2 + [(x_r - x_s)^2 + (y_r - y_s)^2]/v_0^2} \ , \tag{1}$$

where t_0 and v_0 are time and velocity parameters determined from the initial time picks, and t_r are the hyperbolic times of receivers at position (x_r, y_r) separated from the source at position (x_s, y_s). One way of finding good estimates for t_0 and v_0 is to fit the initial picked times with high SNR values to Equation 1 by nonlinear optimization. For the example of Figure 2, the parameters t_0 and v_0 were found to be 25ms and 3200m/s, respectively.

Once a suitable hyperbola is found, we shift it up by 200ms so that it lies entirely above all the first arrivals on the CSG. We shift it down by 200ms to create a second hyperbola that lies entirely below all the first arrivals on the CSG. The two hyperbolas define a retention corridor. We obtain reduced seismograms by muting all trace values outside the retention corridor, and resetting zero time to be at its leading edge. We perform a second pass of MER first-break picking on the reduced seismograms. We keep track of the shifted hyperbolic times of the leading edge so we can add them back to the second pass of MER picks.

The common-source gather of reduced seismograms and the second-pass MER estimates of first-break times are plotted on Figure 3. Also shown are the SNR values associated with the second-pass picked times. Picked times with SNRs greater than a cut-off value of 10dB are considered to be reliable first-arrival times. Increased scatter and low SNR values at receiver positions less than 2500m show a clear divide in the quality and reliability if the picked times..

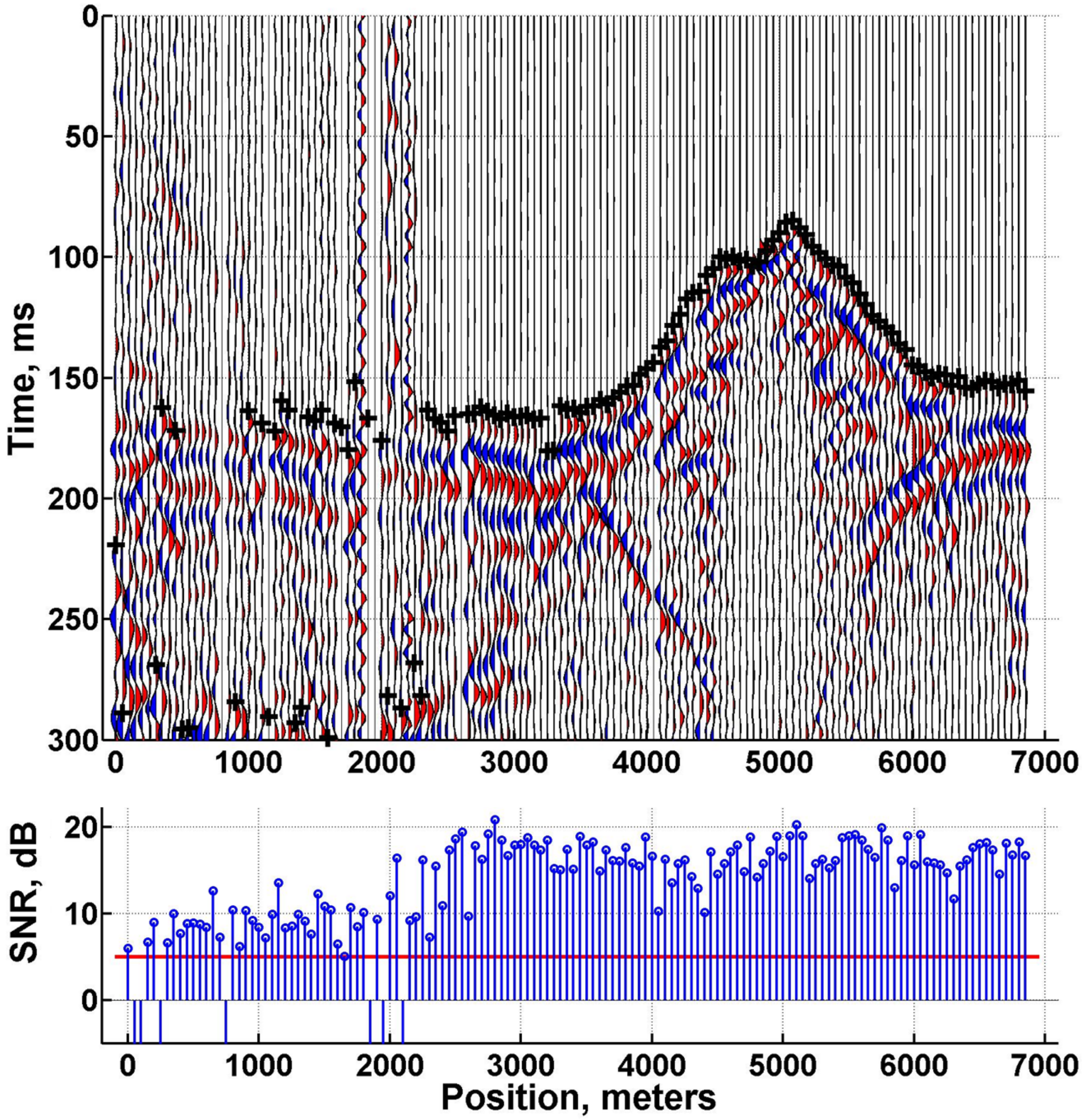

FIG. 3: Reduced seismograms with MER time picks displayed as black crosses. SNR values less than 0dB are associated with killed traces. MER picks associated with SNR values of less than 10dB are erratic and show little spatial coherence. The red line indicates an SNR value of 5dB.

TIME PICKING ON NOISY FAR-OFFSET ARRIVALS

The unreliable time picks at receiver positions 0m to 2500m are partly due to decreased SNRs and partly due to relatively strong arrivals below the reduced time of 250ms. Some of the erratic picks for these far-offset seismograms can be eliminated by muting trace values for times later than 250ms. Figure 4 displays expanded plots of the muted far-offset seismograms. We can see two events with (nearly) linear moveout: a relatively strong, almost flat event at about 170ms, and a weaker earlier event with steep slope. Our problem is to modify our picking procedure so we can obtain reasonably reliable arrival times for both events.

Prior to MER re-picking, we enhanced the SNRs for the signals on Figure 4(a) by performing a three-trace summation. To the value at each digital point on each seismogram, we added values from digital points on the preceding and following traces. The times for the added points were determined by a given linear slope on the gather. We calculated such sloped sums for a range of

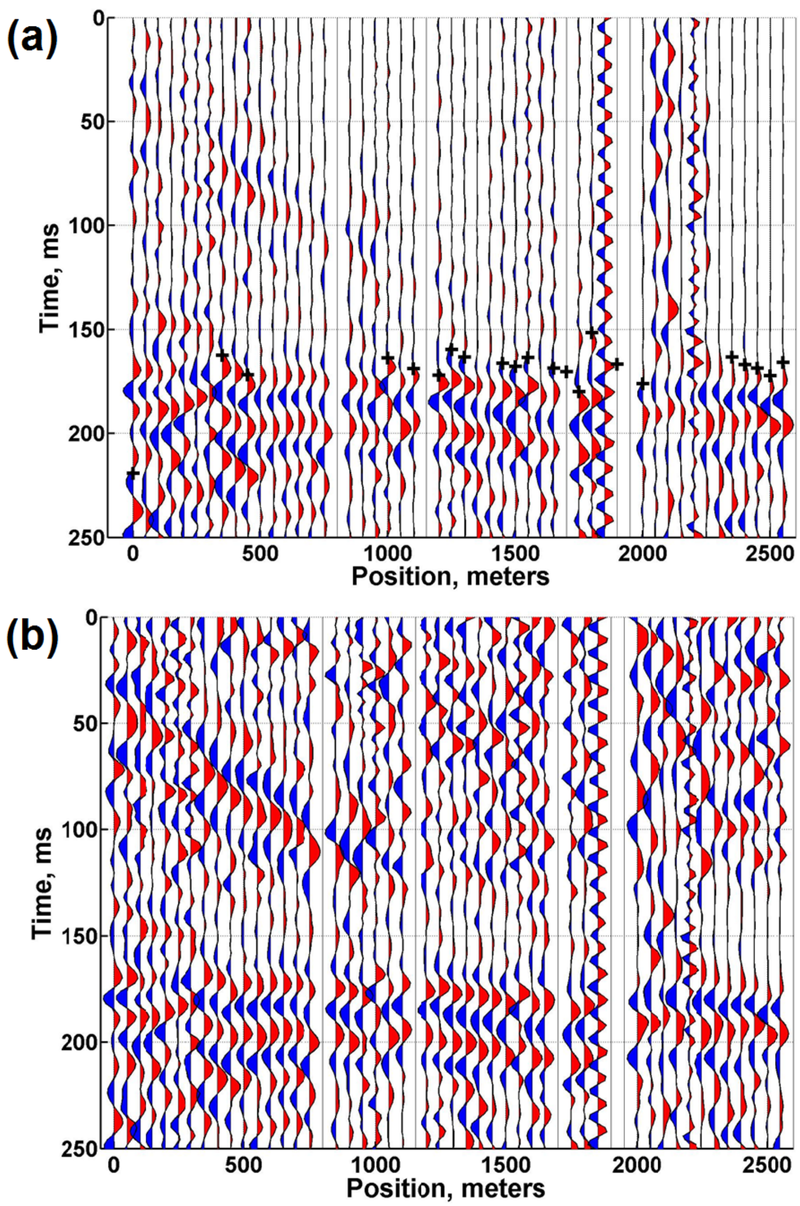

FIG. 4: Reduced traces at the far left flank of the CSG; (a) normalized plot, together with some MER picks (black crosses) from Figure 3; (b) AGC plot with AGC window length = 100ms. Traces are cut off to eliminate any strong arrivals later than 250ms.

slopes, and over this range we retained the sum with the maximum absolute value (along with its sign or polarity). Signals with enhanced SNRs result when we replace the original value at the digital point by the retained sum (with its polarity). The summation is equivalent to localized slant stacking. Figure 5(a) displays the normalized far-offset seismograms after signal enhancement. Most of the subsequent MER picks, shown as yellow dots, are quite good estimates of the first break times on the enhanced traces.

To deal with the earlier weak event with steep slope, we muted the enhanced trace values on Figure 5(a) following the times indicated by the yellow dots. The muted traces are plotted on Figure 5(b) after re-normalizing. MER time-picking on these traces gave reasonably consistent

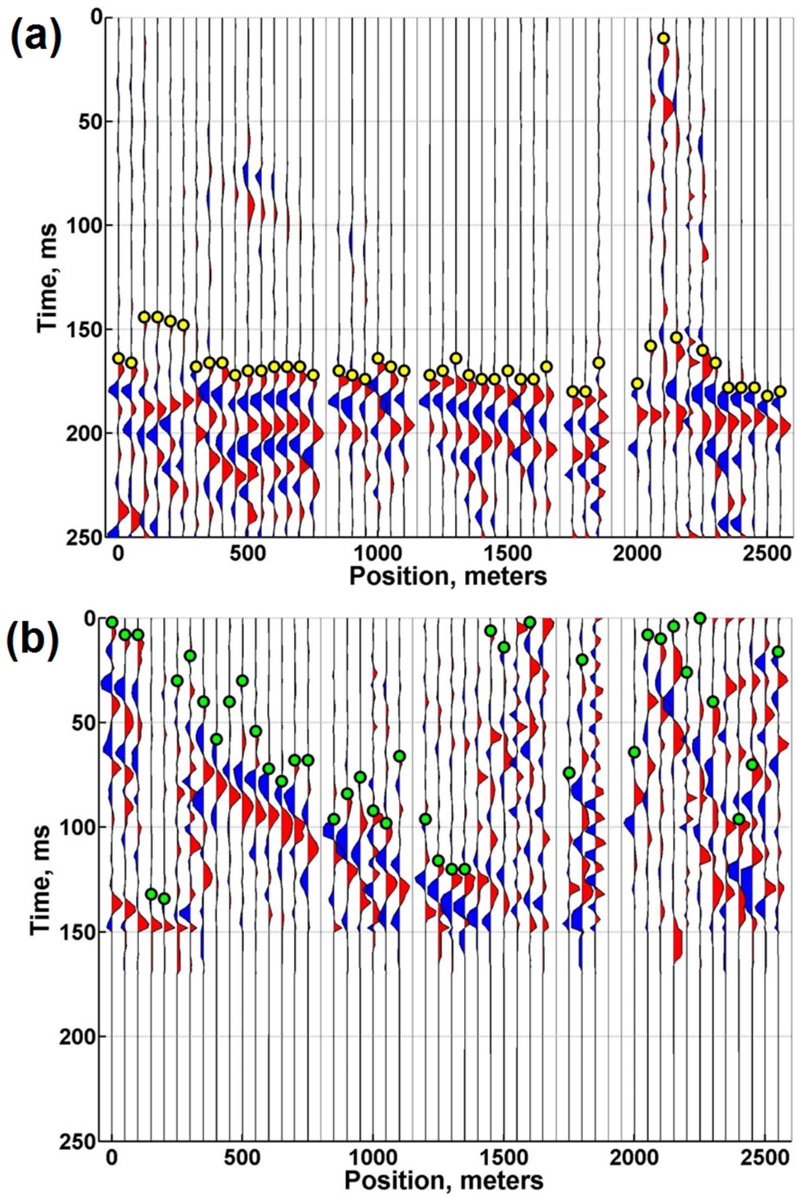

FIG. 5: Normalized plots of the enhanced far-offset traces, together with MER picks (yellow and green dots). (a) For the lower, flat-lying strong event. (b): For the upper, weaker event, after muting of the strong event below times of 150ms, and trace re-normalization.

arrival times for the steeply sloping event, as is indicated by the green dots.

Visually, there appears to be fairly good spatial coherence for both events on Figure 5. However, the MER picks themselves are much more erratic than the visual appearance of the events. This is especially true for the earlier sloping event, for which many of the picks are clearly outliers. We need a way to identify the outliers automatically, and to replace them with times which are more consistent with the visual appearance of the traces. We can do this by exploiting the apparently linear coherence of both events. The linear coherence of arrivals on a seismic gather is most easily brought out by a Radon transform of the data.

Figure 6 displays the data combined from Figures 5(a) and 5(b), together with several typical trajectories for calculating the Radon transform. The origin position for the summation trajectories was chosen to be 1250m (the Radon transform is somewhat dependent on the origin position chosen). The intercept times of the trajectories go from 0ms to 250ms in 2ms steps. Figure 7(a) is a colour-coded display of the complete Radon transform of the Figure 6 data.

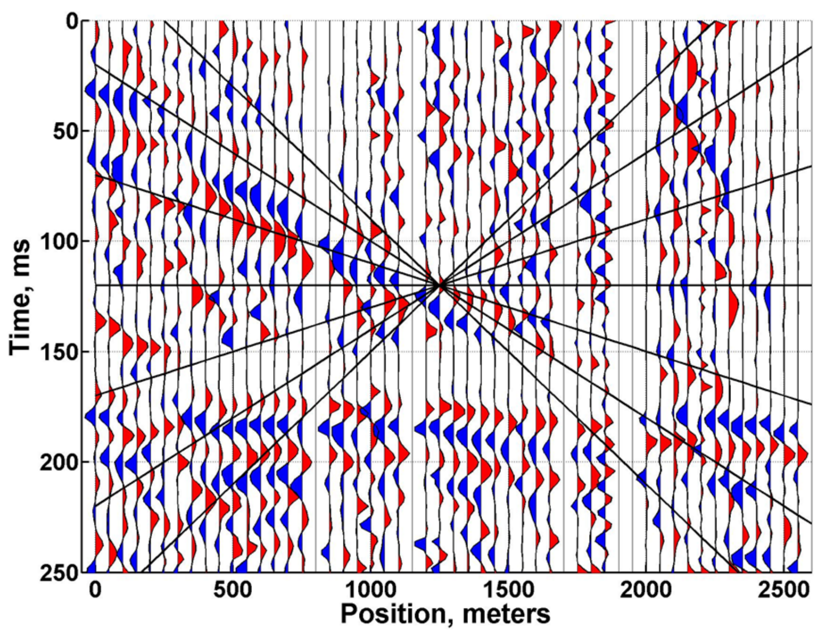

FIG. 6: Radon transform trajectories drawn on combined data from Figure 5 for various slopes (slownesses) with intercept times of 120ms at the origin position of 1250m,.

The strongest maxima and minima on the Radon transform on Figure 7(a) are associated with the possible linear events observable on the far-offset seismograms. Figure 7(b) plots the Radon transform in wiggle trace format. The red and black crosses indicate the maximum and minimum values for the three strongest of the possible events. For each of the three picked slowness value, the small green circle is located at a time equal to the least time of the red and black crosses, minus an eighth of the dominant observed period. For matching the first-break times of an event with linear moveout, the green circle is assumed to give the best estimate for the intercept time associated with that slowness.

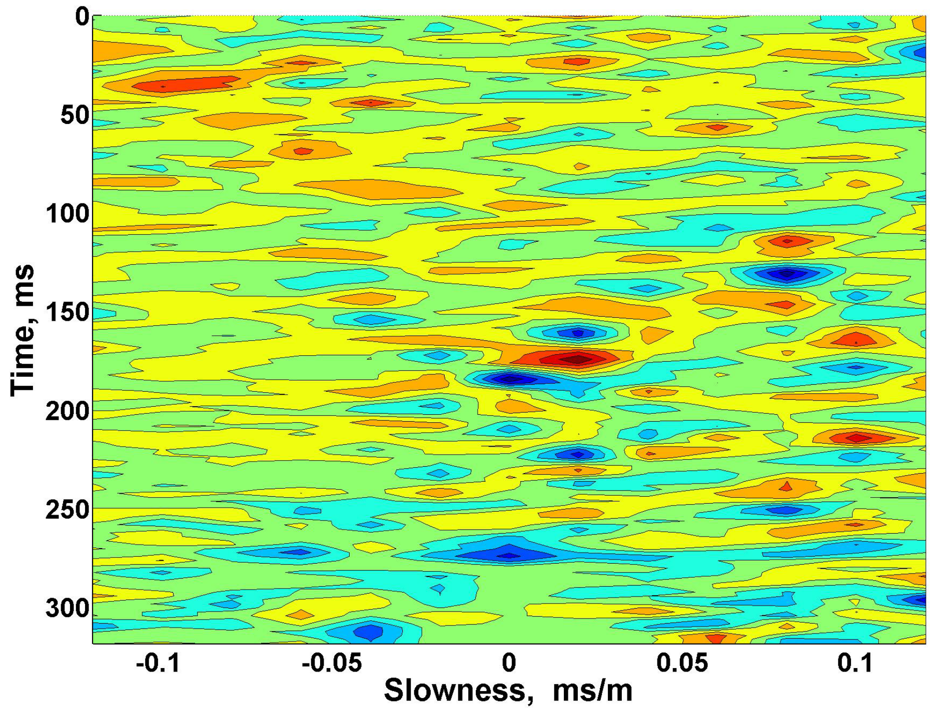

FIG. 7(a): Radon transform of the data in Figure 6, in coloured-coded format.

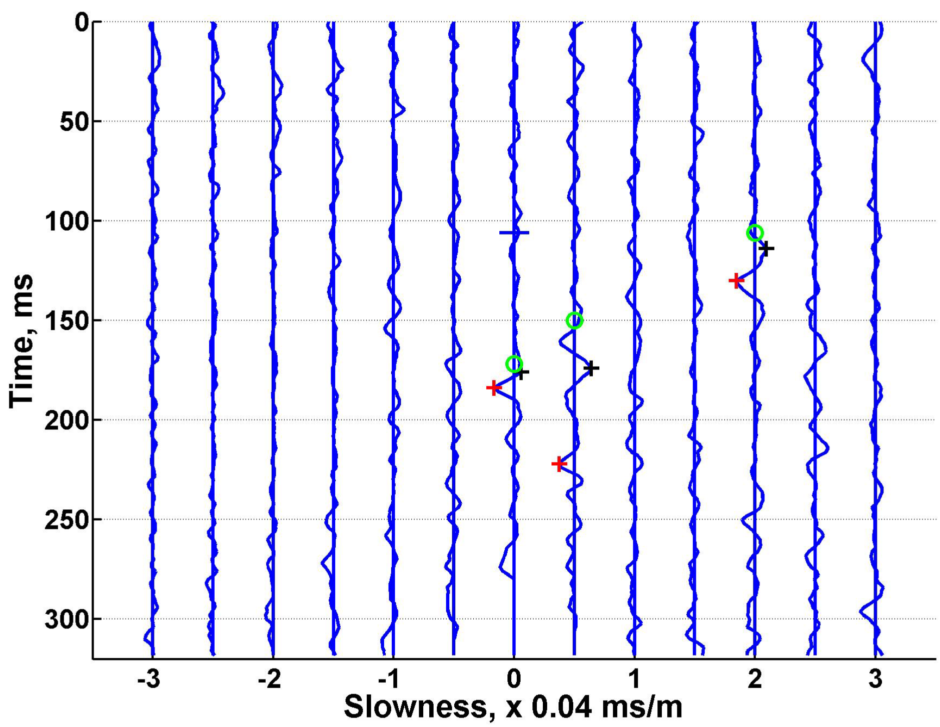

FIG. 7(b): Slownesses and time-intercepts for the three strongest Radon events. Green circles give the best possibilities for estimating first-break times for traces on Figure 6.

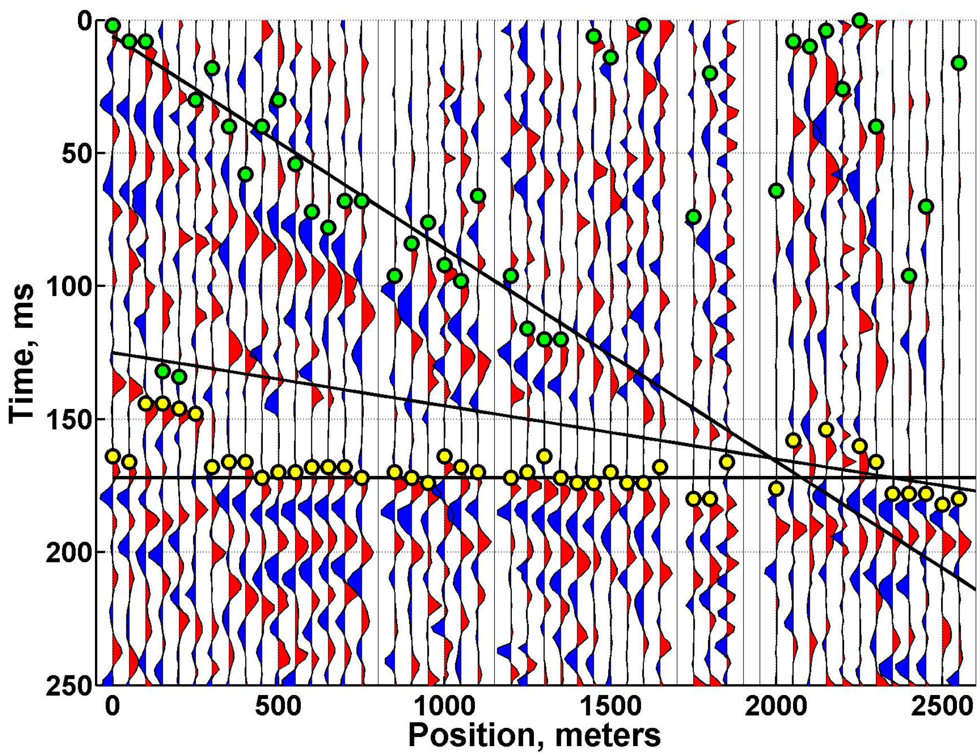

FIG. 8: Enhanced traces combined from Figures 5(a) and 5(b). Black lines are based on the parameters from the three strongest linear events on the Radon transform of Figure 7(b).

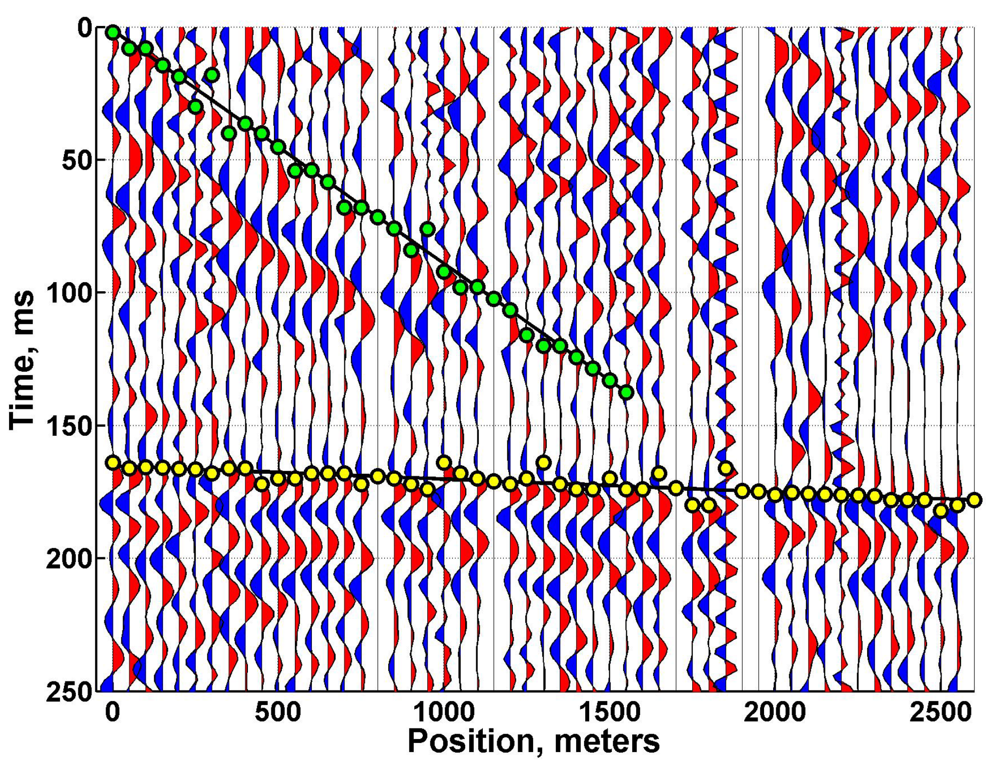

FIG. 9: Best estimates for the arrival times of the two linear events at far offsets, plotted on the AGC seismograms of Figure 4. Black lines ore the best-fit straight lines based on retained MER picks. Colored dots are a combination of retained MER picks and interpolated values.

The slowness and time intercept values obtained from Figure 7(b) were used to predict the arrival times plotted as three straight lines on Figure 8. We compared each set of MER picked times (yellow and green dots) with the time values on each of the three straight lines. We rejected MER picks deviating from the linear predictions by more than 10ms (about a quarter of the dominant observed period). For each set of MER picked times (yellow or green dots), we chose the straight line that gave the most retained MER picks. We then performed a least-squares linear fit to the chosen set of retained MER picks. Receiver positions with rejected MER picks, or with no observed seismic trace, are given values interpolated from the least-squares fitted line. Figure 9 displays the retained MER picks and the interpolated arrival times for both linear events. We consider these retained and interpolated values to be reliable estimates of the arrival times for the two earliest events observed on the far-offset seismograms of Figure 2.

CONCLUSIONS

We have described and tested a scheme for picking first-arrival times on common source gathers based on the modified energy ratio (MER), an attribute that is simple and fast to calculate. The simplicity makes the MER attribute attractive to use in an automated flow for first-arrival time picking on large datasets consisting of many common-source gathers with many seismograms. At the far flanks of such gathers, where the source-receiver offsets are large, first and second arrivals are noisy. A signal enhancement step equivalent to local slant stacking is required to boost signal-to-noise ratios before the MER-base picking method can give reasonably consistent arrival times. Even after signal enhancement, the automatically picked times at far offsets may exhibit many outliers. However, arrivals at far offsets have space-time moveouts that are almost linear. A Radon transform can be used to estimate best fitting linear parameters for the moveouts. These linear parameters help to eliminate outlier MER picks, and can predict interpolated values to replace the rejected picks. The scheme has been tested on a single common-source gather with 138 traces, but it can be applied repeatedly for as many CSGs as are available in a high-volume 3D seismic dataset.

ACKNOWLEDGEMENTS

We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and the industrial sponsors of CREWES for supporting this research.

REFERENCES

Baziw, E., 2002, Application of Kalman filtering techniques for microseismic event detection: Pure and Applied Geophysics, **159**, 449-471.

Boschetti, F., Dentith, M.D., and List, R.D., 1996, a fractal-based algorithm for detecting first arrivals on seismic traces: Geophysics, **61**, 1095-1102.

Coppens, F., 1985, First arrival picking on common-offset trace collection for automatic estimation of static corrections: Geophysical Prospecting, **33**, 1212-1231.

Chen, Z. and Stewart, R., 2005, Multi-window algorithm for detecting seismic first arrivals: Abstracts, CSEG 2005 National Convention, 355-358.

Han, L., Wong, J., and Bancroft, J., 2009, Time picking and random noise reduction on microseismic data: CREWES Research Reports, **20**, 30.1-30.13.

Han, L., 2010, Microseismic monitoring and hypocenter location: M.Sc. Thesis, University of Calgary.

Munro, K.A., 2005, Analysis of microseismic event picking with application to landslide and oil-field monitoring settings, M.Sc. Thesis, University of Calgary..

Qiao, B., and Bancroft, J., 2010, Picking microseismic first arrival times by Kalman filters and wavelet transforms: CREWES Research Report, **22**, 68.1-68.10.

Withers, M, Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., and Trujillo, J., 1998, A comparison of selected trigger algorithms for automated global seismic phase and event detection, Bulletin of the Seismological Society of America, **88**, 95-106.

Willis, M.E., and Toksöz, M.N., 1983, Automatic P and S velocity determination from full waveform digital acoustic logs: Geophysics, **48**, 1631-1644.

APPENDIX A: THE MODIFIED ENERGY RATIO (MER) ATTRIBUTE

Figure A1(a) summarizes the definition and use of the energy ratio and the MER attribute for a digitized seismic trace sampled with time *Δt*.

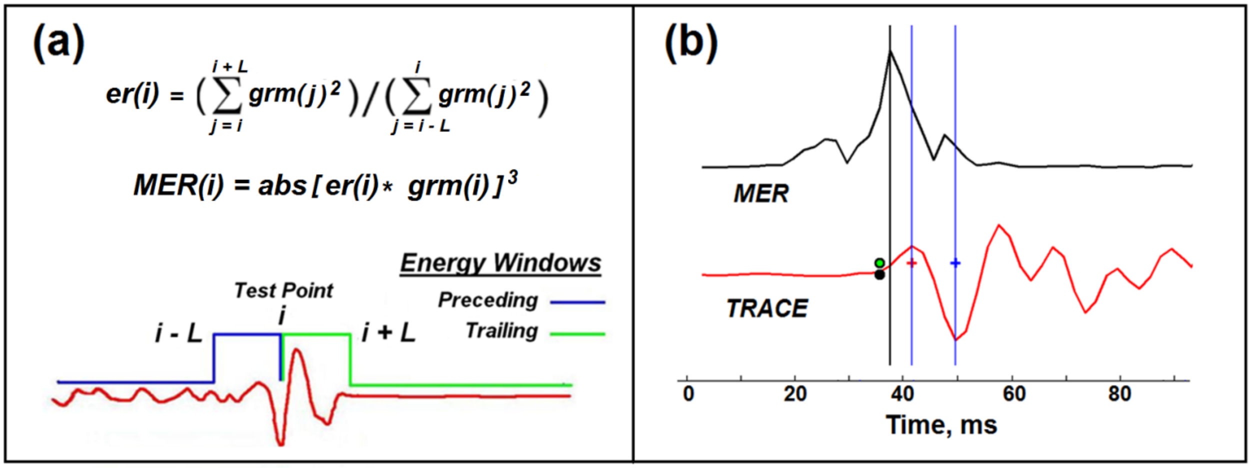

FIG. A1: (a) Energy and modified energy ratios at a test point with time index *i*. *L* is the length of the energy collection windows. (b) Vertical black line is the time t_{MER} of the peak MER value; the black dot indicates the estimated first-break time $t_{B1} = t_{MER} - \delta t$, with $\delta t = T_D/8$. Red and blue crosses indicate the first peak and trough times t_{PK} and t_{TR} following time t_{B1}. The green dot is located at t_{B2}, a second estimate for the first-break time (see Equation A3).

The energy ratio at a test point with index i on the digitized seismogram is

$$er(i) = (\sum_{j=i}^{i+L} grm(j)^2 / \sum_{j=i-L}^{i} grm(j)^2 , \tag{A1}$$

where $grm(j)$ are trace amplitudes, and L is the length of the energy collection windows. In the summations, if index j is less than 1, $grm(j) = [grm(1) + grm(2)]/2$; if j is greater than N the length of the seismic trace, $grm(j) = [grm(N-1) + grm(N)]/2$. The ratio is between energy in a short trailing window to energy in an equally short preceding window. The MER attribute for the seismogram is

$$mer(i) = [\, abs\big(grm(i)\big) * er(i) \,]^3 , \tag{A2}$$

which is independent of signal polarity. Figure A1(b) shows how the first-break time t_{B1} of a seismic trace (plotted in red) is estimated from the peak value of the MER attribute (plotted in black). The peak value of the MER attribute occurs somewhat later than the perceived first-break time. On common-source-gathers with noisy traces, the peak and trough times immediately following t_{B1} may show more trace-to-trace consistency. In such cases, an auxiliary estimate t_{B2} for the first-break time is useful:

$$T_{MAX} = max(t_{PK}, t_{TR}) \text{ , } T_{MIN} = min(t_{PK}, t_{TR}) \text{ ,}$$

$$T_D = 2 * (T_{MAX} - T_{MIN}) \text{ ,}$$

$$t_{B2} = [(T_{MIN} - 0.25 * T_D) + (T_{MAX} - 0.75 * T_D)]/2 \text{ .} \tag{A3}$$

The times t_{PK} and t_{TR} are the trace peak and trough times from Figure A1(b); T_D is an estimate of the dominant period of the arrival wavelet following the estimated first-break time t_{B1}. The difference between the MER pick t_{B1} and the auxiliary pick t_{B2} may be used as a figure of merit for quality control (if the difference is less than some small fraction of the dominant period, the picked first-break time for a particular seismograms is considered to be “good”).

When picking arrival times on field data, the signal-to-noise ratios (SNRs) of the first arrivals is an important consideration. For a given digitized seismogram, the SNR (expressed in decibels) of the first arrival is

$$SNR(i) = 10 \cdot \log_{10}[\ \textstyle\sum_{j=i}^{i+L} grm(j)^2 \ / \ \sum_{j=i-L}^{i} grm(j)^2)\ ] \text{ ,} \tag{A4}$$

where i is the time index of the first-break time pick. The term within the square brackets is the energy ratio in Equation A1. As indicated by the example on Figure 3, SNR values are very useful in assessing the reliability and quality of automatically picked arrival times.

The MER attribute in Equation A2 is independent of the polarity of the trace values. Strictly speaking, the exponent 3 is not required for first-arrival picking. It is included to give plots of the MER attribute a sharper, spike-like appearance at the first-break time; see Figure A1(b).

The window lengths L in the formulas for accumulating energy may have a significant effect on time-picking. For clean seismograms, L can be set to as low as one half to two times of the observed dominant period in the arrival. For seismograms with significant noise, the window lengths may need to be set to longer times. The tradeoff is that longer window lengths give more stable values of the attributes in the presence of noise, but increases the error in locating the first-break time. This error is always towards later times.

The example seismograms used for this article exhibit dominant periods of about 30ms to 50ms; rather arbitrarily, we have set the length of MER energy-collection window L at 32ms. However, testing the procedure with L over several values ranging from 24ms to 72ms gave very similar results for traces with SNR values greater than 10dB.

APPENDIX B: AUTOMATIC PICKING OF FIRST-ARRIVAL TIMES ON COMMON SOURCE GATHERS

Trace pre-conditioning

To minimize the effect of high-amplitude surface wave arrivals on the picking of first arrivals, we applied a bandpass filter with appropriate low-frequency cutoff to each trace. The trace was then normalized by its maximum peak-to-trough amplitude, and the normalized amplitudes (optionally) were divided by a suitable increasing function of time to reduce late-arriving amplitudes. MER attributes are then calculated for every point the trace, and we estimate the first-break time on the trace by picking the time of the maximum MER value and applying the correction $\delta t = T_D/8$. The uncertainty in the time-pick for a low-noise first arrival is in the range

$\pm T_D/8$; the uncertainty for high-noise arrivals is greater, or the pick could be totally false. The SNR values can be used as a figure of merit; low SNR values (less than 10dB) identify those picked first-break times that are potentially unreliable.

Summary of the MER-based time-picking procedure

1. Kill bad traces in the CSG; apply bandpass filtering (10-20-100-200) Hz to reduce surface wave amplitudes; normalize each trace by its maximum amplitude.

2. Optionally, divide trace amplitudes by the time indices, and taper the beginning of each trace.

3. Calculate MER attributes for each trace; make an initial pick of first-break times for all traces in the CSG by finding the maximum MER value for each trace.

4. Use the initial time picks to find parameters for a best-fit hyperbola in space-time coordinates; use the hyperbolic trajectory to define a retention corridor for the common source gather.

5. Mute all signals outside the retention corridor. Reject all initial time picks outside the corridor; retain initial time picks inside the corridor that have high SNRs.

6. Shift and flatten the muted traces using the times on the upper boundary of the retention corridor. This creates a CSG of reduced seismograms with a new time origin. Retain the shift times used to flatten.

7. Re-calculate MER attributes for the reduced seismograms, and re-pick the first-arrival times. Determine SNR values associated with the picks.

8. For the high-noise traces at the CSG flanks with large source-receiver offsets, apply the Radon-transform-assisted technique to estimate the first break times of events with nearly linear moveouts.