Accepted Manuscript

Automatic first-arrival picking method based on an image connectivity algorithm and multiple time windows

Shulin Pan, Ziyu Qin, Haiqiang Lan, José Badal

PII: S0098-3004(18)30496-5

DOI: https://doi.org/10.1016/j.cageo.2018.12.001

Reference: CAGEO 4201

To appear in: *Computers and Geosciences*

Received Date: 21 May 2018

Revised Date: 6 November 2018

Accepted Date: 3 December 2018

Please cite this article as: Shulin Pan, Ziyu Qin, Haiqiang Lan, José Badal, Automatic first-arrival picking method based on an image connectivity algorithm and multiple time windows, *Computers and Geosciences* (2018), doi: 10.1016/j.cageo.2018.12.001

Automatic first-arrival picking method based on an image connectivity algorithm and multiple time windows

Shulin Pan[1, 2], Ziyu Qin[1], Haiqiang Lan[3, 4*], José Badal[5]

1. School of Earth Science and Technology, Southwest Petroleum University, Chengdu 610500, China.
2. State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu 610500, China.
3. State Key Laboratory of Lithospheric Evolution, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China.
4. University of the Chinese Academy of Sciences, Beijing 100049, China.
5. Physics of the Earth, Sciences B, University of Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain.

* Corresponding author: Haiqiang Lan, lanhq@mail.iggcas.ac.cn.

Abstract

The use of a computer to automatically pick the first-arrival of a seismic signal is an operation that involves picking and screening the first arrival of the wave according to the criteria established in the manual picking process. To increase the picking accuracy for data with low-to-moderate signal-to-noise ratio (SNR), we propose a new single-trace boundary detection algorithm. This algorithm includes three steps: (1) calculate the first-arrival characteristic values through multi time windows; (2) take the times corresponding to the maximum characteristic values given by different time windows as intermediate results; (3) compare the intermediate results: if the

difference is too large, it is marked the time is abnormal, otherwise the average time of the intermediate results is taken as the first-arrival time. Using this energy boundary detection method, the characteristic values obtained are bi-directionally expanded to allow the use of the trace connectivity algorithm which is improved from the region growing method. Determining the connectivity between the first-arrival characteristic values is a way to simulate how the human eye discriminates true first arrivals. This method significantly improves the elimination of false or abnormal first-arrivals. Next, a small-step fitting algorithm is applied to the remaining first-arrival characteristic values to complete the calculation of the final characteristic values. Based on the retained first-arrival characteristic values, the missing values are assigned by interpolation. The characteristic values are mapped on the original record and finally the first-arrival picking is completed using a small time window. Theoretical results as well as the results obtained from real data demonstrate that the proposed automatic first-arrival picking method effectively improves the accuracy of the first-arrival picking. Finally, the new picking algorithm is presented more efficient than the energy ratio method, as well as cross-correlation method.

Keywords: first-arrival picking; image connectivity algorithm; multi time windows; interpolation.

1. Introduction

Manual first-arrival picking may be of quality but of low efficiency, by which the use of a computer instead of manual picking is an "ultimate goal" in the context of the

research of an automatic picking algorithm. The picking of a qualified first-arrival provides the basis for static correction and subsequent data processing. Many researchers have advanced different methods for the automatic first-arrival picking. In an early stage and based on the similarity of adjacent seismic traces, Peraldi and Clement (1972) proposed cross-correlation of adjacent traces to obtain the time-lag between the first-arrival onset and the peak of the signal. But this method is not very effective when there is a large difference in the waveforms of adjacent seismic traces or when some trace is missing. Hatherly (1980) proposed a method for determining the quality of the first-arrival based on the time difference between the initial onset and the peak of the wave. However, this method provides poor performance when the similarity between the seismic traces is low, apart from the fact that the calculations involved in this algorithm are quite complicated. Gelchinsky and Shtivelman (1983) combined association methods with statistical methods and used the first-arrival time-distance curve to constrain the results of picking. With this method, similar results can be achieved to those offered by manual first-arrival picking in the case of signals with a high signal-to-noise ratio (SNR), but also large errors when SNR is low. Coppens (1985) was the first to propose the "energy ratio method", which uses the ratio between the energy of the signal in one cycle and the energy of the entire time window as the criterion for determining the first arrival. This method is comparatively more noise-resistant and many scholars conducted more research based on it. Baer and Kradolfer (1987) used an envelope function and a non-linear amplifier for automatic phase picking. Murat and Rudman (1992) and McCormak et al. (1993) used

a neural network approach for the automatic first-arrival picking. The detection algorithms became more "intelligent" with the improvement of the calculation resources. Boschetti (1996) proposed a fractal-based algorithm for detecting first arrivals on seismic traces. Liao et al. (2011) proposed an automatic first-arrival picking method based on time-frequency analysis using minimum uncertainty wavelets. Mousa et al. (2012) proposed a first-arrival enhancement method using the τ-p transform on energy-ratio seismic shot records, thus obtaining an increase in SNR. Senkaya and Karsli (2014) used the cross-correlation technique for automatic first-arrival detection. Tan et al. (2014), based on the difference in amplitudes, polarizations and statistical characteristics between ambient noise and seismic signal, used the SLPEA algorithm for the automatic microseismic event detection and first-arrival picking.

An et al. (2015) have proposed the combination of ultra-virtual interferometry for ground-scattered waves with traditional ultra-virtual interferometry for refracted waves as a means of enhancing the energy of the first-arrivals. Maity and Salahi (2016) have developed a neuro-evolutionary event detection technique for downhole microseismic surveys of low SNR. Recently, Chi-Durán et al. (2017) have conducted new strategies based on the Fourier transform and the fractal method for the automatic detection of P- and S-wave arrival times. These methods offer certain advantages for the automatic first arrival picking, but still present some application problems. The main issue is that these methods do not meet expectations in cases of low energy level and low SNR.

In this paper, we use for the first time the connectivity between first arrivals as an important criterion to determine the first-arrival of the wave. Abnormal first arrivals are identified and removed when examining its connectivity with the first arrivals in the adjacent seismic traces. To use the connectivity feature effectively in the determination of the first arrival, we develop a search algorithm based on the expansion of the first-arrival characteristic values. Also, the multi time-window energy ratio method is modified to improve the accuracy of the procedure with low SNR signals.

2. Binary image connectivity algorithm

The image connectivity algorithm is widely used in areas including medical, transportation, surveying and mapping research. Martin-Herrero (2007), Wu et al. (2009) and He (2009) used connected-component labeling algorithms in digital images. Commonly used image labeling methods include pixel labeling, linear labeling and block scanning. Pixels that are adjacent to each other form what are called connected components.

There are two common adjacency modes of 4 and 8 connected components. Fig. 1 shows the matrices for 4 and 8 connected components: For the mode of 4 components the central component is 1, while any other outer component is marked 1 if it is adjacent to the previous one. In the following we use the 8 connected-component search algorithm. The connected-component labeling algorithm scans the data from left to right and from top to bottom, searches for points of interest and determines whether a component keeps connectivity with the neighbor

component. To reduce the number of scans and improve the labeling efficiency, we use the region growing method to label connected components (Pavlidis and Liow, 1990). The steps of this procedure are described below:

- Step 1: Input of the two-dimensional data set D(x,y) to be labeled and definition of the labeling matrix M(x,y) that has the same size as D(x,y), a queue L and label count N.
- Step 2. Scanning of D(x,y) from left to right and from top to bottom. When an unlabeled point is scanned, N increases by 1 and the current point is properly labeled in M(x,y). Scanning of the 8 points connected to the current point. If unlabeled points are found, then they are labeled in M(x,y) and put into L as seeds for growing.
- Step 3. If L is not empty, a point from L is taken as seed for growing and the 8 points that are connected with the seed point are scanned. Unlabeled points that are found will be labelled in M(x,y) and put into L.
- Step 4. Repetition of step 3 up to L is empty, so that the labeling of a connected component is completed.
- Step 5. Go back to step 2 up to the entire picture is scanned to obtain the label matrix M(x,y) and the number N of connected components.

To better understand the previous algorithm, Fig. 2 shows the flow chart illustrating the operations involved by the region growing method. All field data analyzed here come from the Zhungeer basin survey conducted in 2012 (data provided by courtesy of SINOPEC). Although the first arrivals may be visually

connected, this does not mean that they are well connected from the perspective of image processing. The characteristic values of visually well-connected first arrivals include many empty blocks in the pixel matrix. This makes difficult to distinguish the valid values from noise if the connectivity algorithm is used directly. For this reason, the first-arrival characteristic values must be preprocessed before applying the connectivity algorithm. Fig. 3a shows the raw seismic traces and Fig. 3b the input data, i.e. the first-arrival characteristic values calculated from the previous data. Fig. 3c shows the expansion operator matrix: the black circles represent the positions of the values involved in the calculation of the current value whose position is highlighted in red. The result of the expansion algorithm is shown in Fig. 3d.

Each seismic trace provides only one characteristic value and therefore the connectivity in conventional image processing needs to be modified and adapted for first-arrival picking. So, to better determine the connectivity between first arrivals, we refer to the concept of trace connectivity: if there is connected points between the current trace and the next one then we can say the current trace is connected to the next trace. The number of connected traces will be used to determine the connectivity of the first arrivals, instead of the number of connected points. The connectivity between valid first arrivals is much greater with real data than with noise.

3. Calculating first-arrival characteristic values

The energy ratio method allows improve the performance of the first-arrival picking. In particular, the single-trace boundary detection algorithm is more noise-resistant. It can be described as follows:

156 $$S_i = |(B/A)\times(B-A)| \quad (1)$$

157 $$A=\sum_{p=i}^{p=i-n} S_p \ , \ B=\sum_{p=i}^{p=i+n} S_p$$

158 where S_i is the boundary characteristic value obtained for the ith sample point; A is
159 the sum of the amplitudes of n points before the current point in the same trace and B
160 is the sum of the amplitudes of n points after the current point in the same trace. For
161 this type of algorithm, the choice of the time window plays a crucial role in the
162 picking result. Different time windows have significant different sensitivities to first
163 arrivals. A large time window embraces the general characteristics of the signal, while
164 a small time window provides a more accurate description of the details. A single time
165 window has strong limitations. Fig. 4a shows a trial seismic trace with a weak first
166 arrival. The first arrival occurs at about 400 ms. The first-arrival characteristic values
167 are calculated by equation (1) as the time window is opened more and more to times
168 of 5 ms, 25 ms, 45 ms and 65 ms. The results can be seen in Fig. 4b that the 25ms
169 time window is better than other time windows. The same initial trace after having
170 added random noise is shown in Fig. 4c. In this case the results obtained by following
171 the same procedure are shown in Fig. 4d. It can be observed that, the first arrival is
172 annihilated by noise, even with 25ms time window a much larger characteristic value
173 is observed much away from the true first arrival. In addition, the results calculated by
174 different windows are differ greatly. So the selection of time windows is a difficult
175 task in such data with weak first arrival signal or low SRN.

176 Repeating the process in the search for characteristic values from seismic data
177 with high SNR (Fig. 5a), using time windows of 25 ms, 45 ms, 65 ms and 85 ms, the

results reveal that the characteristic values appear basically in the same place (Fig. 5b), i.e. the results of the automatic first-arrival picking are highly reliable. On the contrary, when dealing with seismic data with low SNR (Fig. 6a) (the correct first arrival position is marked by red line.), the results are no longer satisfactory because the location and the magnitude of the characteristic values vary greatly (Fig. 6b), in spite of using time windows with similar opening. It is difficult to choose the first arrival accurately regardless of the size of the time window or the type of determination criterion that is used for processing. In fact, the first arrival is buried in noise and cannot be picked accurately. In this case the automatic picking offers poor reliability and should be discarded. Even so, the greater or lesser reliability in the determination of first arrivals in traces with high or low SNR can be easily discerned according to the distribution of the characteristic values calculated using different time windows.

After several tests, we propose to use a formula to detect the single-trace energy boundary with the help of multi time windows to obtain the first-arrival characteristic values, as well as to eliminate the noise interference. The calculation is as follows:

$$X_k = \max\left(\left|(B_k / A_k)\times(B_k - A_k)\right|\right) \tag{2}$$

$$I_k = pos(X_k) \tag{3}$$

$$pos(final) = \begin{cases} \text{ceil}\ (\dfrac{\sum I_1 + I_2 + \cdots I_K}{K}) & , \max(I_k) - \min(I_k) < w \\ 0 & , \max(I_k) - \min(I_k) \geq w \end{cases} \tag{4}$$

In these expressions X_k is the maximum characteristic value once calculated using the *k*th time window of size n_k (the index k is the sequential window number); the index *i* indicates the current point; $pos(X_k)$ is the position of X_k in the corresponding trace; $pos(final)$ is the final first-arrival position, given by the rounded average value of all previous positions if the difference between characteristic values obtained using different time windows does not exceed *w* ms; otherwise, if the difference is greater than *w* ms, the first arrival is marked as abnormal. After a large number of tests with real data, we concluded that the value *w* is given by half of the wavelet of the first-arrival wave. Take the peak frequency as the frequency of the first-arrival wave. If it were 50Hz, the length of the wavelet can be calculated to be 20ms, then the value of w would be 10 ms.

Fig. 7 shows an example of common-shot gather that is taken as a reference for the subsequent calculation. Fig. 8 shows the first-arrival characteristic values (highlighted in color in the illustration) determined by applying different time windows (10 ms, 20 ms, 30 ms, 40 ms and 50 ms) to all traces of the original seismic record. Most of their respective positions mapped directly on the record are overlapped to the first arrivals due to the relatively high SNR. This example reveals that for those traces affected by a high noise level, the differences between the positions of the first arrivals and those determined using different time windows are large. Fig. 9 shows the final result of mapping first-arrival characteristic values (highlighted in color in the illustration) obtained by applying the single-trace energy boundary detection (STEBD) method and increasingly large time windows (10 ms, 20

ms, 30 ms, 40 ms and 50 ms) to all traces of the record section shown in Fig. 7. The traces with first arrivals at time zero are abnormal traces. A simple comparison of the picking results (Fig. 8 and Fig. 9) indicates that the STEBD method is able to effectively overcome the biased information due to first arrivals with low SNR and eliminate abnormal first arrivals. The eliminated first arrivals can be re-picked again by interpolation and searching for local peaks to obtain the definitive picking result (Fig. 10a). The comparison with the results obtained by cross-correlation and energy ratio reveals the better performance of the STEBD method against the two previous ones (Fig. 10b) due to the use of multiple time windows. The improved STEBD method effectively avoids the wrong picking and, through interpolation and local optimization, provides more stable and reliable results.

4. Connectivity treatment and refinement of characteristic values

In practice the reality can be very different and we can meet with a seismic record with low SNR or some ambiguity or difficulty to distinguish the first arrivals of energy, if not all at least some of them (Fig. 11a). The first-arrival boundary is much weaker than the energy boundary of a refracted or reflected wave at a certain depth. In such cases, the STEBD algorithm may not accurately recognize the limits of the first-arrivals. This is clear in Fig. 11b where we show the result of applying STEBD and multi time windows. Some first arrivals provided by erroneous picking (highlighted in color and enclosed by an ellipse) are clearly below their correct positions. The use of linear adjustment methods to correct these abnormal first arrivals can easily cause the erroneous elimination of nearby correct first-arrivals, thus

affecting the overall picking and making more difficult the re-reading of the abnormal first arrivals in the subsequent processing.

Continuing with the example, Fig. 11c shows the result of applying the "binarization" procedure to the first-arrival characteristic values processed by the improved STEBD method (section 3). The workflow was as follows: the characteristic values were put in binary format, i.e. the amplitude at the point corresponding to a first arrival was set to 1, while the amplitude at any other point was set to 0. The values given by the expansion algorithm (section 2) applied to the signals are shown in Fig. 11d. The (labeled) values obtained by connectivity processing of the signals are shown in Fig. 11e, and the number of connectivity of each connected component was counted. Connected components with less than 50 traces were considered as abnormal components and zero amplitude was assigned to each of these components. The dot product of the result of this process (Fig. 11e) and the original binary record (Fig. 11c) resulted in the new record in binary format (Fig. 11f). If the difference between the times of the current and adjacent first-arrival is too large, the current first-arrival is removed (Fig. 11g). Lastly, we obtained the missing arrivals by time picking and interpolation (for this last operation, 10 points before and after the missing point are selected and the linear fitting is performed using the current first-arrival time and the relative offsets of these points). The final result is the one shown in Fig. 11h, where the best first-arrival positions at points (with amplitude 1) determined by local optimization appear mapped on the original seismic record. The final result is indeed satisfactory compared to the situation at the beginning.

In order to test the proposed algorithm, we considered an indeed complex seismic velocity structure (Fig. 12a). The spatial dimensions of this 2D model are 80 km (length) x 7 km (depth). The shot point is marked by a red arrow on top. Fig. 12b shows the shot data obtained by forward modeling, while Fig. 12c shows the first arrivals determined with the method proposed in this study. Adding random noise and coherent noise to the data completed the seismic experiment, so we could obtain new results in the same way from data both with moderate SNR ((Fig. 12d) and low SNR (Fig. 12e). It can be seen that the random noise has little effect on the picking results (Fig. 12d). However, the coherent noise has a somewhat more pronounced effect, even giving rise to a lack of first arrival (Fig. 12d).

To appreciate the advantage of the automatic first-arrival picking method that we proposed, next we present an application example based on the data collected on the occasion of the 2D survey line across the Wulungu depression in the northern margin of the Junggar basin. The maximum offset distance is 6650 m, and the number of traces reaches 660 with horizontal spacing between traces of 20 m. For comparison purposes, first arrivals were automatically picked using the method proposed in this paper and alternatively the energy ratio method using commercial software. The static correction was calculated using the tomographic static correction technique. After data processing, we obtained the results plotted in Fig. 13. In the two cases the same portion of record section is highlighted to easily see the improvement in the information achieved by the new method (Fig. 13c) versus the energy ratio method (Fig. 13a) and the cross-correlation method (Fig. 13b). A simple visual inspection

allows to see that the seismic markers found through the implementation of the new first-arrival picking method appear better reconstructed, which result in more accurate information for exploratory practice.

5. Conclusions

Starting from a single-channel boundary detection algorithm, we propose a new STEBD method based on the fact that time windows of different sizes reveal different boundary detection sensitivities. In this method, the connectivity algorithm in image processing is applied to first-arrival picking. To reduce the number of scans and thus improve the labeling efficiency, we use the region growing method to label connected components. The STEBD algorithm is implemented using the energy ratio method, which improves the picking performance, since it is more resistant to noise. Several time windows of different sizes are used for single-trace energy boundary detection and so obtain the first-arrival characteristic values. This allows us to avoid the instability caused by the manual setting of the time-window width and to greatly mitigate the noise influence as well. The improved algorithm is able to effectively eliminate abnormal first arrivals in signals with low SNR.

The joint application of the trace connectivity algorithm and multi time-window boundary detection comes to effectively solve problems associated to data with low SNR and to the automatic picking of low-energy first arrivals. After performing a variety of tests with real data, the results prove to be clearer and more reliable than those obtained with standard software, which is of great interest for exploratory practice. However, choosing the connected parameter is an important step in the

proposed algorithm, and the characteristics of the data should be considered. In some cases, this parameter is difficult to adjust by experience, which will inevitably lead to repeated tests. Consequently, it is worth further studying to find an algorithm that can automatically select the connected parameters.

Acknowledgements

We appreciate the helpful comments and suggestions from three anonymous reviewers that allowed us to substantially improve the presentation of this article. We are grateful for the financial support of the following institutions: Open Projects Fund of the State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation (PLN201733), Open Projects Fund of the Natural Gas and Geology Key Laboratory of Sichuan Province (2015trqdz03), National Natural Science Foundation of China (NSFC 41674095), State Key Laboratory of Lithospheric Evolution (grant SKL-YT201802), and Youth Innovation Promotion Association of the Chinese Academy of Sciences (grant 2015051).

References

Baer, M., Kradolfer U., 1987. An automatic phase picker for local and teleseismic events. Bulletin of the Seismological Society of America 77(4), 1437-1445 .

Boschetti, F., 1996. A fractal-based algorithm for detecting first arrivals on seismic traces. Geophysics 61(4), 1095-1102.

Chi-Durán, R., Comte, D., Díaz, M., Silva, J. F. , 2017. Automatic detection of P- and S-wave arrival times: new strategies based on the modified fractal method and basic matching pursuit. Journal of Seismology 21(4), 1-14.

Coppens, F., 1985. First arrival picking on common-offset trace collections for automatic estimation of static corrections. Geophysical Prospecting 33(8), 1212-1231.

Gelchinsky, B., Shtivelman, V., 1983. Automatic picking of first arrivals and parameterization of traveltime curves. Geophysical Prospecting 31(6), 915-928.

Hatherly, P.J., 1982. A computer method for determining seismic first arrival times. Geophysics 47(10), 1431-1436.

He, L., Chao, Y., Suzuki, K, et al., 2009. Fast connected-component labeling. Pattern Recognition 42(9), 1977-1987.

Liao, Q., Kouri D., Nanda, D., Castanga J., 2011. Automatic first break detection by spectral decomposition using minimum uncertainty wavelets. Soceity of Exploration Geophysics, 81st Annual Meeting (San Antonio), 1627-1631 .

Maity, D., Salehi, I., 2016. Neuro-evolutionary event detection technique for downhole microseismic surveys. Computers & Geosciences 86, 23-33.

Martín-Herrero, J., 2007. Hybrid object labeling in digital images. Machine Vision Application 18(1), 1-15.

McCormack, M.D., Zaucha, D.E., Dushek, D.W., 1993. First-break refraction event picking and seismic data trace editing using neural networks. Geophysics 58(1), 67-78.

Mousa, W., Al-Shuhail, A., Abdullatif, A., 2012. Enhancement of first arrivals using the τ-p transform on energy-ratio seismic shot records. Geophysics 77(3), 101-1111.

Murat, M., Rudman, A., 1992. Automated first arrivals picking: A neural network approach. Geophysical Prospecting 40(6), 587-604.

Pavlidis T., Liow, Y.T., 1990. Integrating Region Growing and Edge Detection. Pattern Analysis & Machine Intelligence, IEEE Transactions 12(3), 225-233.

Peradi, R., Clement, A., 1979. Digital processing of refraction data Study of first arrival. Geophysical Prospecting 20(3), 529-548.

Şenkaya, M., Karsli, H., 2014. A Semi-Automatic Approach to Identify First Arrival Time: the Cross-Correlation Technique (CCT). Earth Sciences Research Journal 18(2), 107–113.

Sheng-Pei, A.,Tian-Yue, Hu., Yong-Fu, Cui., W. S, Duan ., G. X, Peng., 2015. Auto-pick first break swith complex ray paths for undulate surface conditions. Applied Geophysics 12(1), 93-100.

Tan, Y., Yu, J., Gang, F., He, C., 2014. A combined method for automatic microseismic event detection and arrival picking. SEG Annual Meeting. Denver, Colorado, USA, pp. 2335-2340.

Wu, K., Otoo, E., Suzuki, K., 2009. Optimizing two-pass connected-component labeling algorithms. Pattern Analysis and Applications 12(2), 117-135

Figure captions:

Figure 1. Matrices for 4 (left) and 8 (right) connected components.

Figure 2. Flow-chart illustrating the operations involved by the region growing method.

Figure 3. (a) Raw seismic traces (data provided by courtesy of SINOPEC). (b) Input data: first-arrivals characteristic values calculated from the above traces. (c) Representative 3*3 matrix of the structure expansion operator: the black circles represent the positions of the values involved in the calculation of the current value whose position is highlighted in red. (d) Output data: result of the expansion algorithm.

Figure 4. (a) A trial seismic trace with a weak first arrival (data provided by courtesy of SINOPEC). (b) First-arrival characteristic values calculated from the above data using increasingly time windows of 5 ms, 25 ms, 45 ms and 65 ms. (c) Same initial trace after having added random noise. (d) First-arrival characteristic values calculated from the data contaminated by noise.

Figure 5. (a) A trial seismic trace with high SNR (data provided by courtesy of SINOPEC). (b) First-arrival characteristic values calculated from the above data using increasingly large time windows of 5 ms, 25 ms, 45 ms, 65 ms and 85 ms.

Figure 6. (a) A trial seismic trace with very low SNR (data provided by courtesy of SINOPEC). (b) First-arrival characteristic values calculated from the above data using increasingly large time windows of 5 ms, 25 ms, 45 ms, 65 ms and 85 ms.

Figure 7. Example of common-shot gather taken as reference for subsequent calculation (data provided by SINOPEC).

Figure 8. First-arrival characteristic values (highlighted in color) obtained by applying different single time-windows (10 ms, 20 ms, 30 ms, 40 ms and 50 ms) to the record section shown in Fig. 7.

Figure 9. First-arrival characteristic values (highlighted in color) obtained by applying increasingly time windows (10 ms, 20 ms, 30 ms, 40 ms and 50 ms) and the single-trace energy boundary detection method to the record section shown in Fig. 7.

Figure 10. (a) Final result obtained after re-picking the traces shown in Fig. 9. (b) Results picked by the cross-correlation and energy ratio methods.

Figure 11. (a) Record section with low SNR seismic traces showing unclear first arrivals (courtesy of SINOPEC). (b) Result obtained using single-trace energy boundary detection and multi time windows (see red dotted line). Some first arrivals provided by erroneous picking (all enclosed by an ellipse) are clearly below their correct positions. (c) First-arrival characteristic values in binary format. (d) Values given by the expansion algorithm applied to the signals. (e) Values obtained by the connectivity processing of the signals shown in (d). (f) New record in binary format obtained after the connectivity processing of the signals previously treated. (g) Values obtained by automatic picking after removing abnormal points. (h) Best first-arrival positions obtained by time picking and interpolation (see red dotted line).

Figure 12. (a) Seismic velocity model with a complex geometry (the shot point is marked by a red arrow on top). (b) Record section obtained from the previous model by forward modeling. (c) First arrivals determined with the method proposed in this study. (d) Results obtained in the same way from data contaminated by random noise (with moderate SNR). (e) Results obtained in the same way from data contaminated by coherent noise (with low SNR). In all cases the first arrivals are highlighted by red dotted lines.

Figure 13. Comparison of the static correction effect on the automatic first-arrival picking using real field data. The data come from the profile across the Wulungu

421 depression in the northern margin of the Jungaar basin (data provided by courtesy of
422 SINOPEC). (a) Result obtained by the energy ratio method using commercial
423 software. (b) Result obtained by cross-correlation method. (c) Result obtained using
424 the method proposed in this paper. The rectangles delimit portions of the seismic
425 record where the improvement in the information can be appreciated clearly.

426

427

428

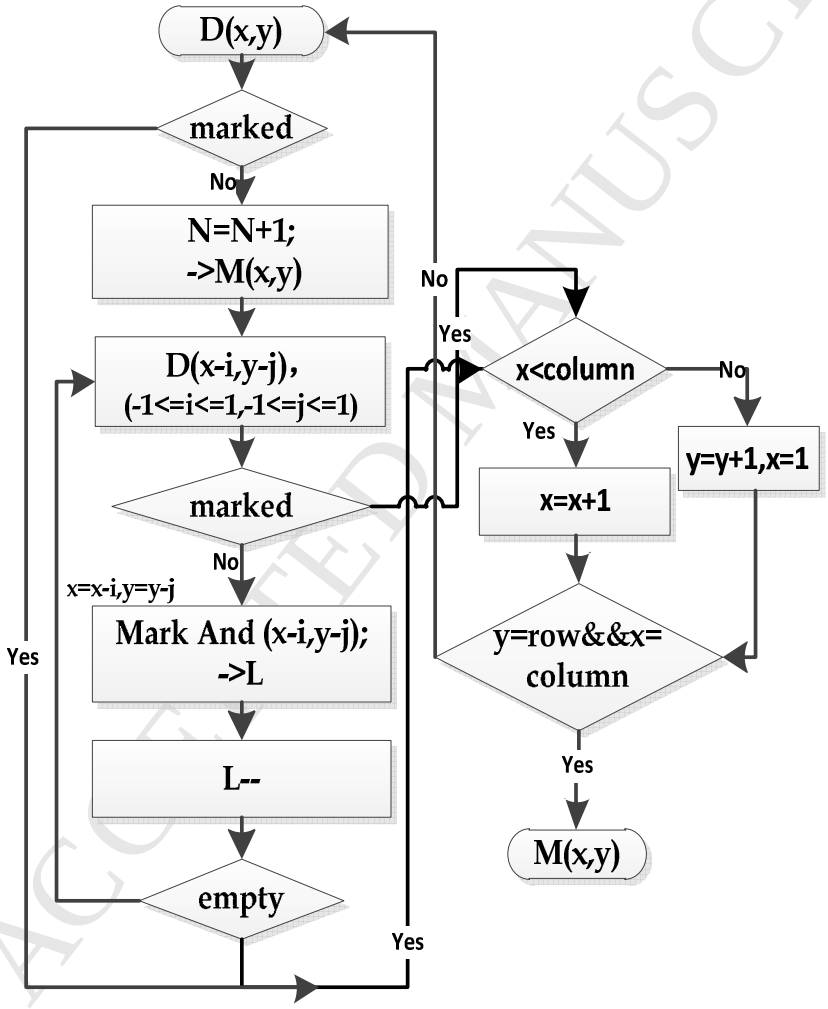
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

429

430 Figure 1. Matrices for 4 (left) and 8 (right) connected components.

431

432



433

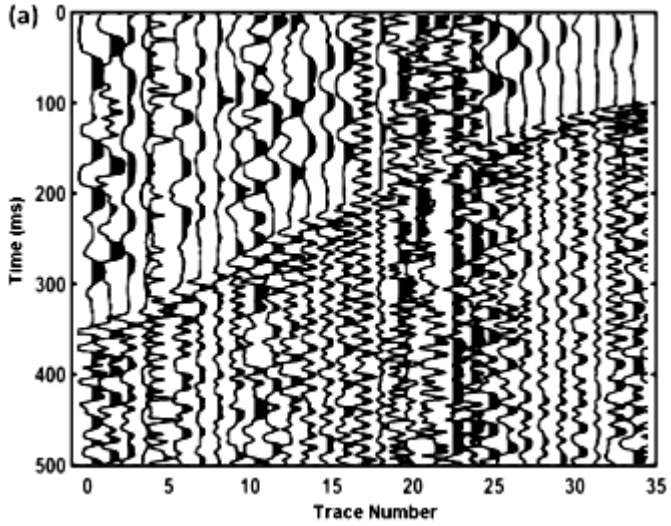
434 Figure 2. Flow-chart illustrating the operations involved by the region growing

435 method.

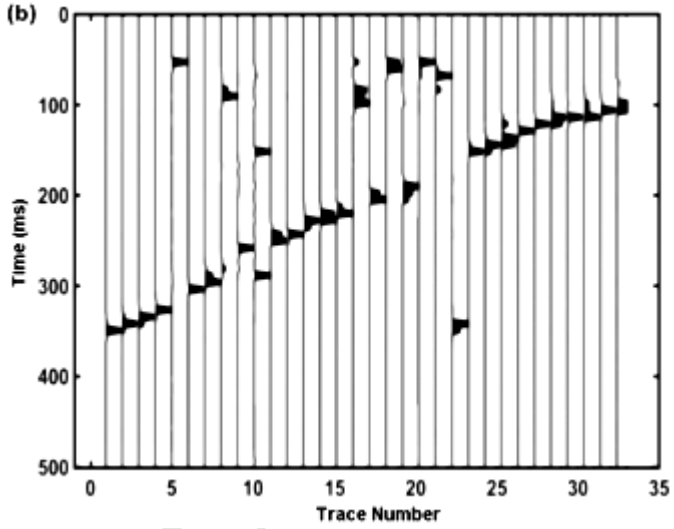
436

437

438

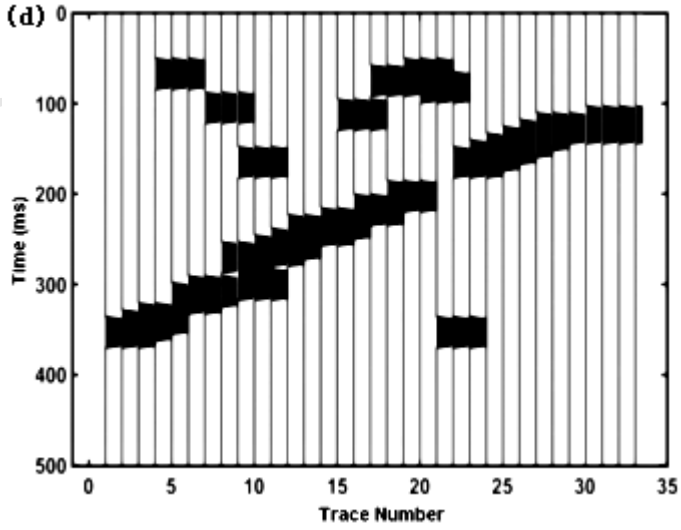


439



440

441



442

443 Figure 3. (a) Raw seismic traces (data provided by courtesy of SINOPEC). (b) Input
444 data: first-arrivals characteristic values calculated from the above traces. (c)
445 Representative 3*3 matrix of the structure expansion operator: the black circles
446 represent the positions of the values involved in the calculation of the current value
447 whose position is highlighted in red. (d) Output data: result of the expansion
448 algorithm.

449

450

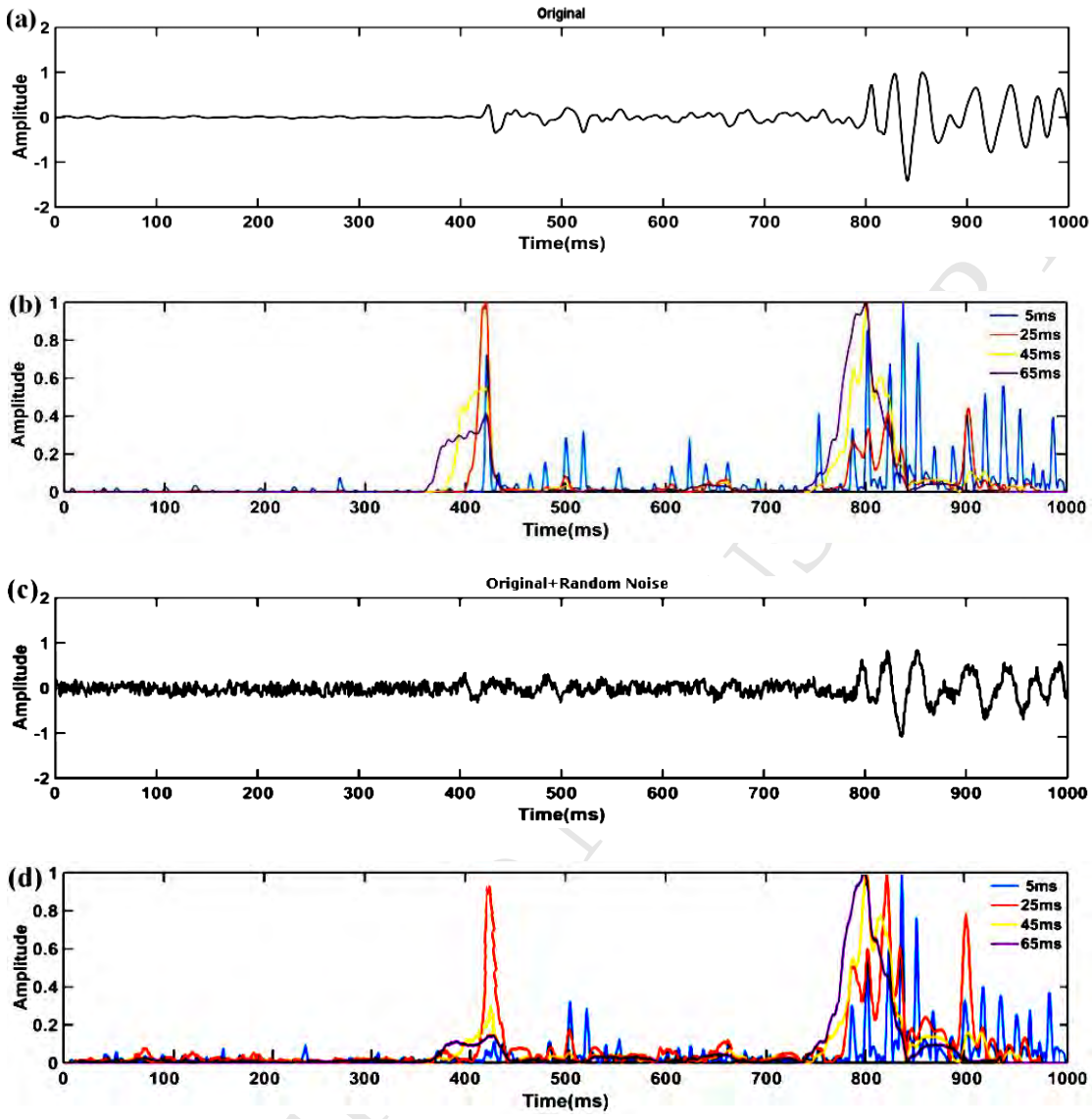


Figure 4. (a) A trial seismic trace with a weak first arrival (data provided by courtesy of SINOPEC). (b) First-arrival characteristic values calculated from the above data using increasingly time windows of 5 ms, 25 ms, 45 ms and 65 ms. (c) Same initial trace after having added random noise. (d) First-arrival characteristic values calculated from the data contaminated by noise.

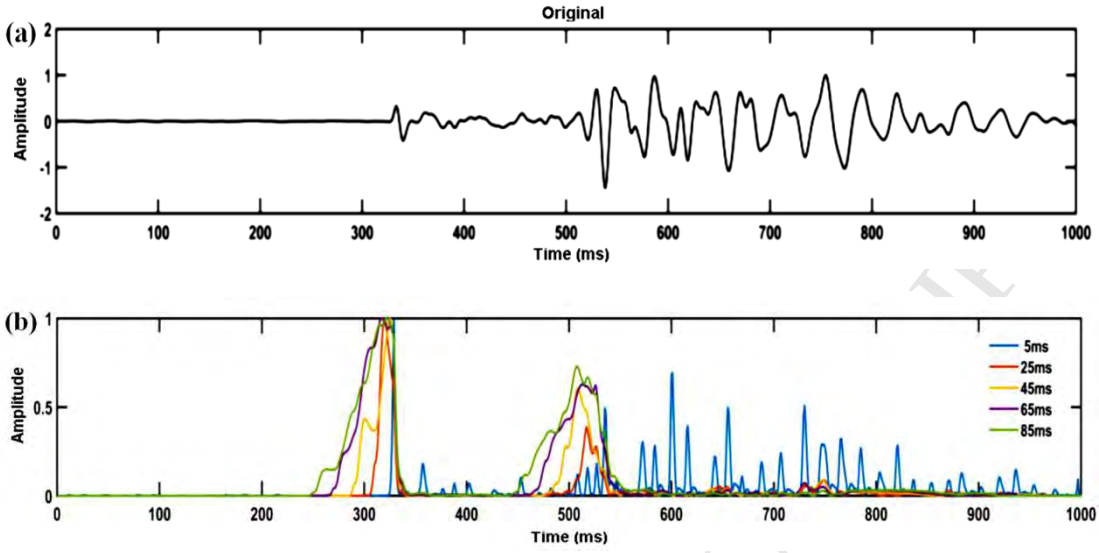


Figure 5. (a) A trial seismic trace with high SNR (data provided by courtesy of SINOPEC). (b) First-arrival characteristic values calculated from the above data using increasingly large time windows of 5 ms, 25 ms, 45 ms, 65 ms and 85 ms.

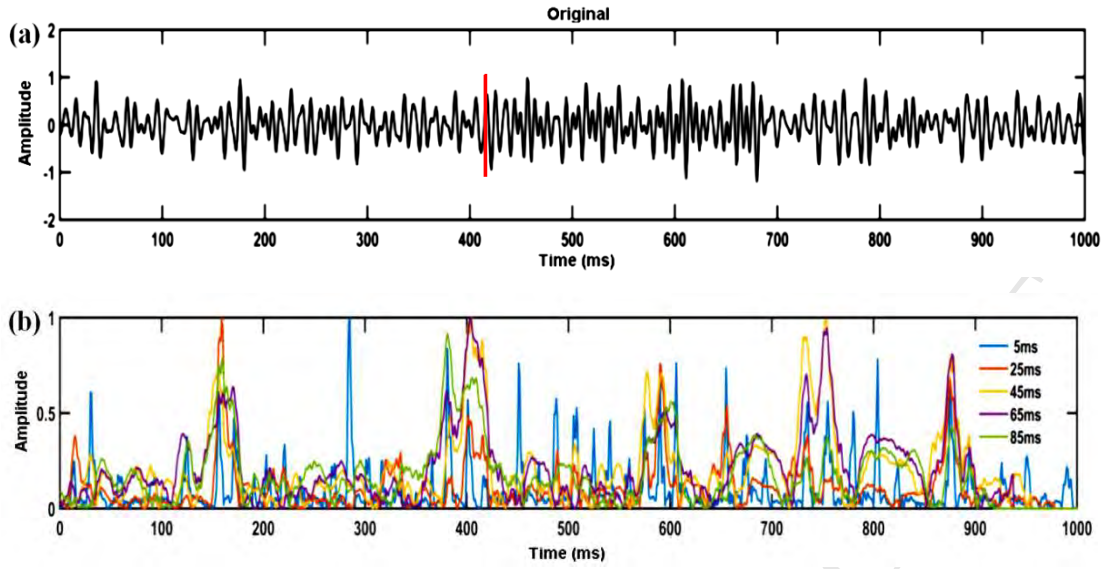


Figure 6. (a) A trial seismic trace with very low SNR (data provided by courtesy of SINOPEC). (b) First-arrival characteristic values calculated from the above data using increasingly large time windows of 5 ms, 25 ms, 45 ms, 65 ms and 85 ms.

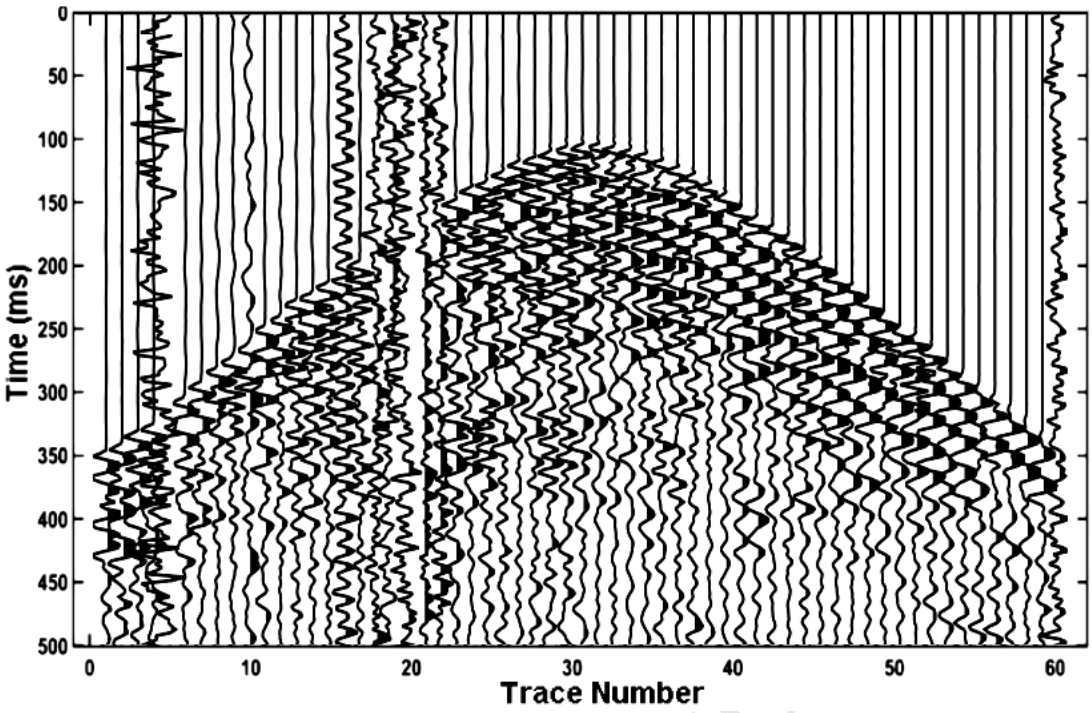


Figure 7. Example of common-shot gather taken as reference for subsequent calculation (data provided by SINOPEC).

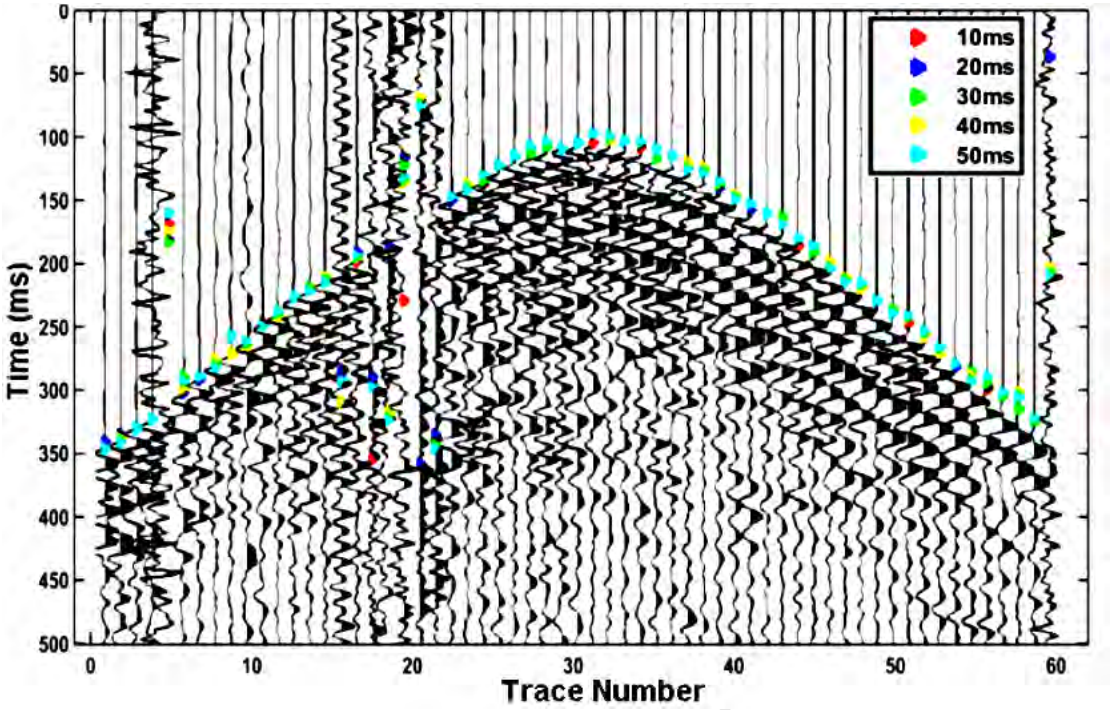


Figure 8. First-arrival characteristic values (highlighted in color) obtained by applying different single time-windows (10 ms, 20 ms, 30 ms, 40 ms and 50 ms) to the record section shown in Fig. 7.

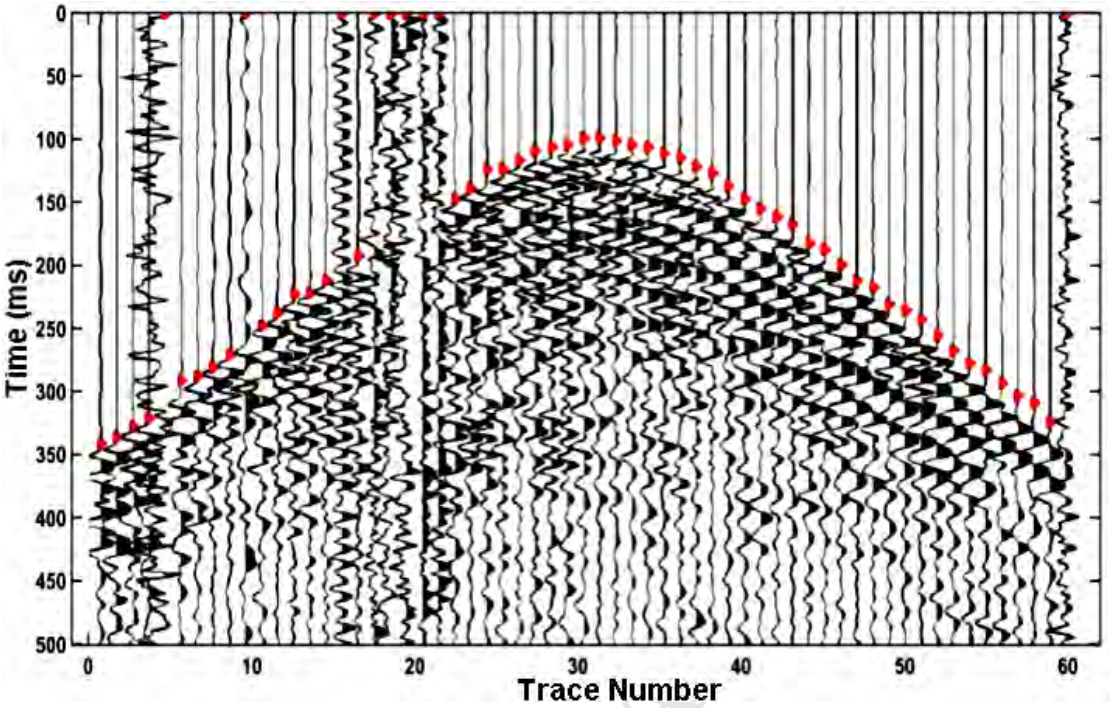
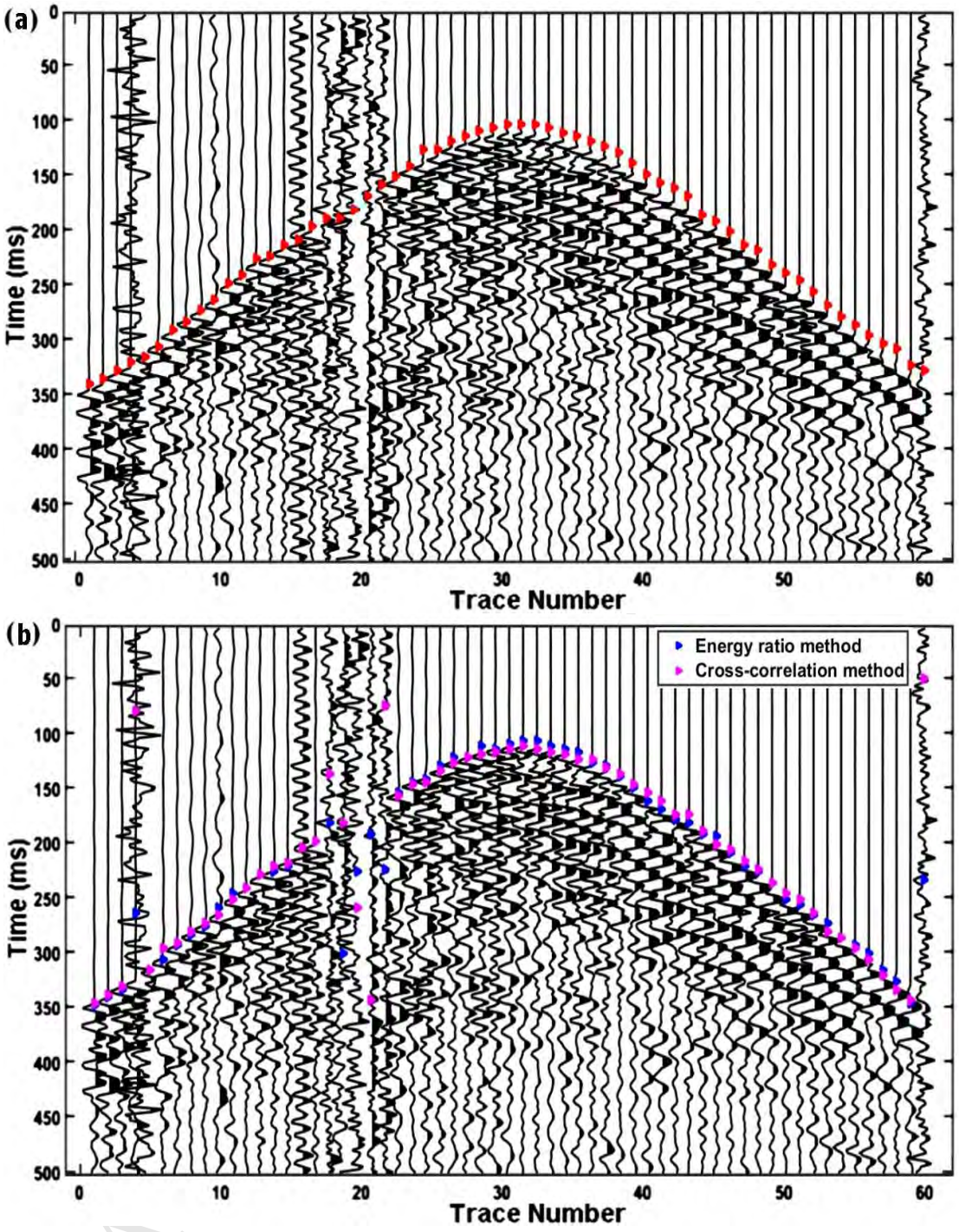


Figure 9. First-arrival characteristic values (highlighted in color) obtained by applying increasingly time windows (10 ms, 20 ms, 30 ms, 40 ms and 50 ms) and the single-trace energy boundary detection method to the record section shown in Fig. 7.

506



507

508

509 Figure 10. (a) Final result obtained after re-picking the traces shown in Fig. 9. (b)

510 Results picked by the cross-correlation and energy ratio methods.

511

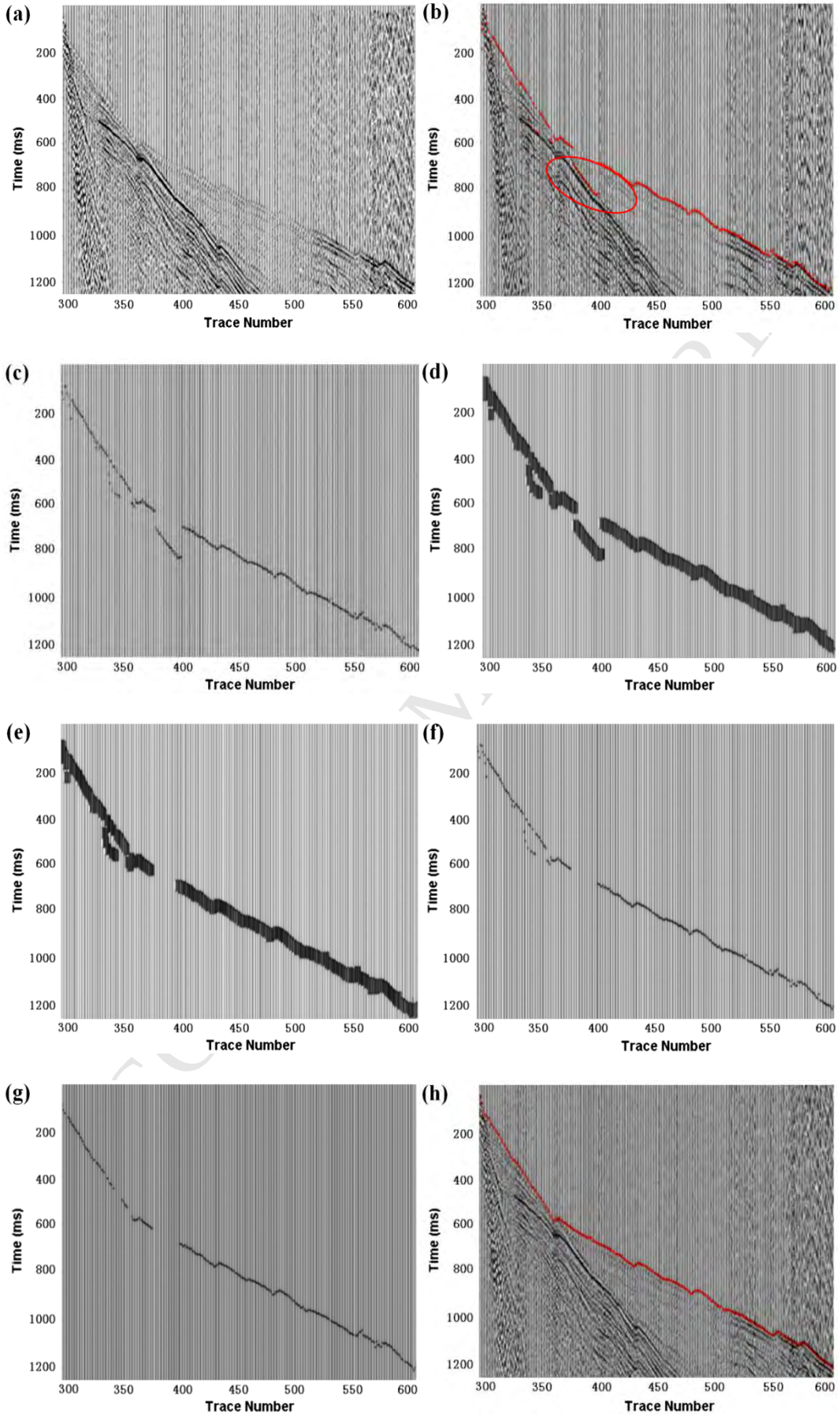

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
Time (ms)
Trace Number
200
400
600
800
1000
1200
300
350
400
450
500
550
600

Figure 11. (a) Record section with low SNR seismic traces showing unclear first arrivals (courtesy of SINOPEC). (b) Result obtained using single-trace energy boundary detection and multi time windows (see red dotted line). Some first arrivals provided by erroneous picking (all enclosed by an ellipse) are clearly below their correct positions. (c) First-arrival characteristic values in binary format. (d) Values given by the expansion algorithm applied to the signals. (e) Values obtained by the connectivity processing of the signals shown in (d). (f) New record in binary format obtained after the connectivity processing of the signals previously treated. (g) Values obtained by automatic picking after removing abnormal points. (h) Best first-arrival positions obtained by time picking and interpolation (see red dotted line).

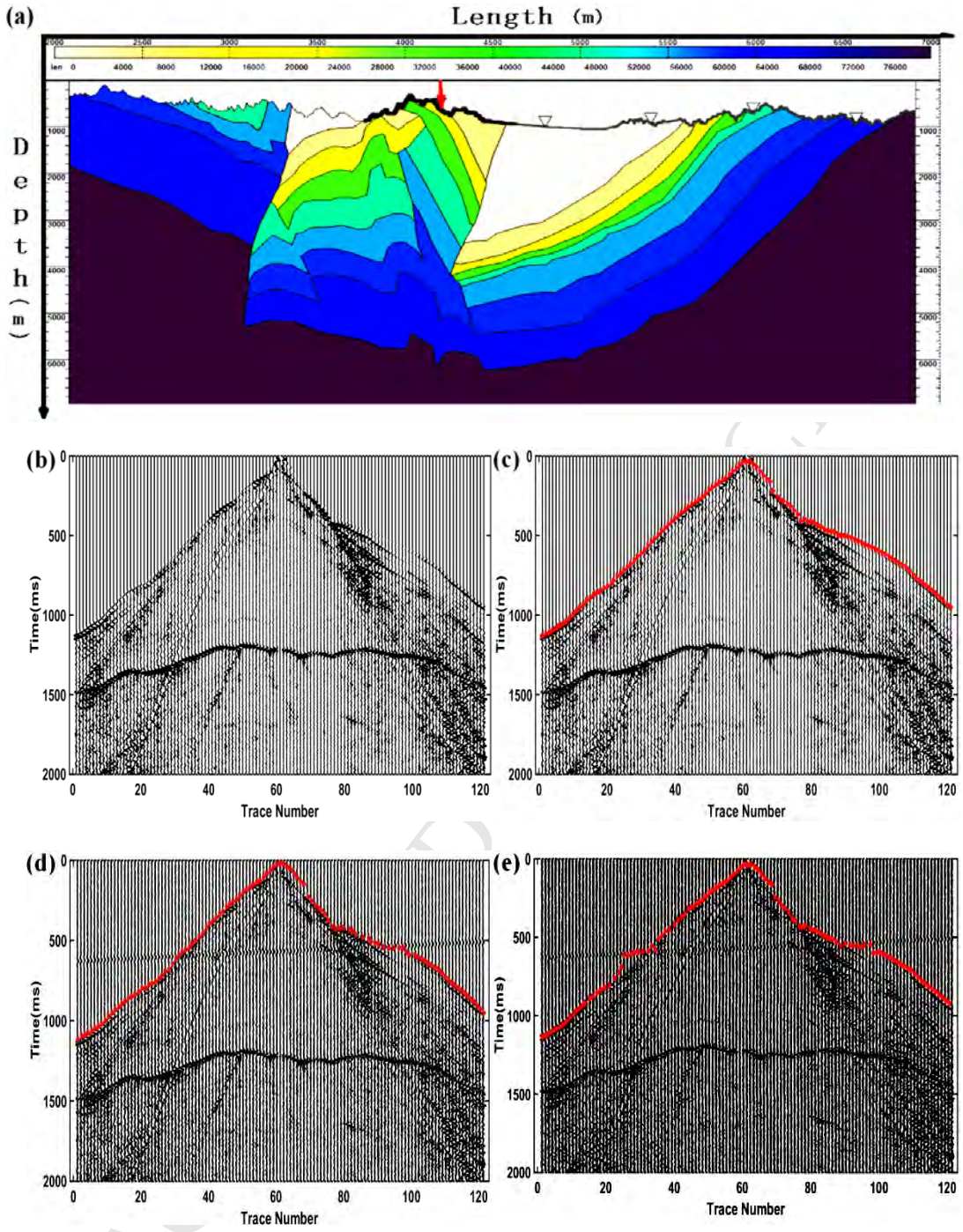
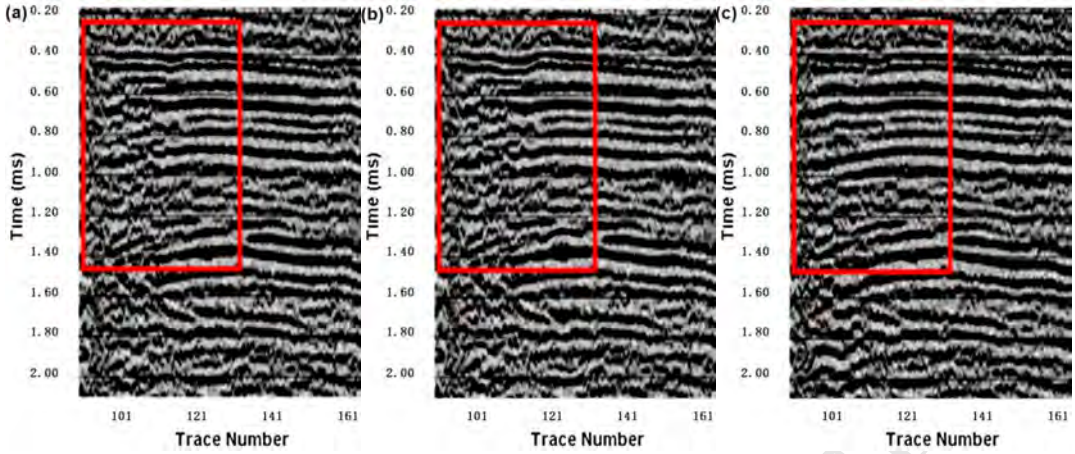


Figure 12. (a) Seismic velocity model with a complex geometry (the shot point is marked by a red arrow on top). (b) Record section obtained from the previous model by forward modeling. (c) First arrivals determined with the method proposed in this study. (d) Results obtained in the same way from data contaminated by random noise (with moderate SNR). (e) Results obtained in the same way from data contaminated by coherent noise (with low SNR). In all cases the first arrivals are highlighted by red

539 dotted lines.



540

541 Figure 13. Comparison of the static correction effect on the automatic first-arrival
542 picking using real field data. The data come from the profile across the Wulungu
543 depression in the northern margin of the Jungaar basin (data provided by courtesy of
544 SINOPEC). (a) Result obtained by the energy ratio method using commercial
545 software. (b) Result obtained by cross-correlation method. (c) Result obtained using
546 the method proposed in this paper. The rectangles delimit portions of the seismic
547 record where the improvement in the information can be appreciated clearly.

548

549

Highlights:

- The proposed method improves the elimination of false or abnormal first-arrivals.
- The method is more accurate as that of the single-trace boundary detection algorithm.
- The method has a great potential in applications automatic picking first-arrivals.