White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/127030/

Version: Accepted Version

Article:
Yuan, S, Liu, J, Wang, S et al. (2018) Seismic waveform classification and first-break picking using convolution neural networks. IEEE Geoscience and Remote Sensing Letters, 15 (2). pp. 272-276. ISSN: 1545-598X

https://doi.org/10.1109/LGRS.2017.2785834

© 2018 IEEE. This is an author produced version of a paper published in IEEE Geoscience and Remote Sensing Letters. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Seismic waveform classification and first-break picking using convolution neural networks

Sanyi Yuan, Jiwei Liu, Shangxu Wang*, Tieyi Wang, and Peidong Shi

***Abstract*—The convolutional neural networks (CNNs) approach is rarely applied to seismic waveform classification and first-break picking, although it has recently been applied to a series of fields and has exceeded human performance in some visual tasks. This letter investigates how CNNs is adopted to classify time-space waveforms from seismic shot gathers and further pick first breaks of both direct wave and refracted wave. We use representative sub-image samples with two types of labeled waveform classification to supervise CNNs training. The goal is to obtain the optimal weights and biases in CNNs, which are solved by minimizing the error between predicted and target label classification. The trained CNNs can be utilized to automatically extract a set of time-space attributes or features from any sub-image in shot gathers. These attributes are subsequently input to the trained fully-connected layer of CNNs to output two values between 0 and 1. Based on the two-element outputs, we define a discriminant score function to provide a single indication for classifying input waveforms. The first break is then located from the calculated score maps by sequentially using a threshold, the first local minimum rule of every trace and a median filter. Finally, we adopt synthetic and real shot data examples to demonstrate the effectiveness of CNNs-based waveform classification and first-break picking. The results illustrate that CNNs is an efficient automatic data-driven classifier and picker.**

***Index Terms*—Waveform classification, first-break picking, seismic data, convolutional neural networks (CNNs).**

For Peer Review

I. Introduction

THE earth is increasingly understood through active or passive seismic data, which are recorded by sensors at the surface or in boreholes to interpret the subsurface structure, prospect mineral resources and predict natural hazards (see [1]). The first break or the traveltime of the first arrival is a key piece of

*Corresponding author Email: wangsx@cup.edu.cn

Sanyi Yuan, Jiwei Liu, Shangxu Wang, and Tieyi Wang are with the China University of Petroleum, State Key Laboratory of Petroleum Resources and Prospecting, CNPC Key Lab of Geophysical Exploration, Beijing 102249, China.

Peidong Shi is with the School of Earth & Environment, University of Leeds, Leeds LS2 9JT, United Kingdom.

information of seismic data; it has been widely applied to statics correction processing, traveltime tomography, velocity inversion, source location determination, source mechanism characterization and hazard assessment. Fundamentally, the waveform features of seismic sub-images centred by the first break and non-first break are discrepant in the time domain, space domain or time-space domain. Consequently, this provides interpreters with the chance to manually or automatically pick the first break and meanwhile classify seismic waveforms.

Manual first-break picking of the P- and/or S-wave is a simple and straightforward method that implicitly leverages waveform classification. However, manual picking is tedious and time consuming when large amounts of data are processed, which is very common in seismic exploration. In addition, picking accuracy depends on the experience of the interpreter. A large number of (semi-)automatic methods (see [2], [3]), such as the short- and long-term average (STA/LTA) ratio, autoregressive techniques, time-frequency transform and higher-order statistics, have been proposed to pick the first break of the P- or S-wave. Nevertheless, these methods are usually not adaptive, only work well under certain conditions and are often restricted to identifying a single type of first break [4]. Furthermore, these methods commonly employ a single-trace process [5], thereby ignoring the feature of spatial coherence among traces. There are also methods using artificial neural networks (ANNs) for picking the first break from (micro)seismic data (see [6], [7]). These methods take a window from a trace and calculate sensitive attributes or features (e.g., the STA/LTA ratio and autoregressive coefficients; the variance, skewness and kurtosis; the amplitude, phase and frequency) to the first break (see [8], [9]). These attributes are considered as ANNs input and the network has to decide whether the corresponding classification output is first break or non-first break. ANNs-based methods can adaptively pick different types of first breaks, but the extraction of sensitive attributes has large uncertainty. In addition, these methods seldom employ the spatial coherence features of waveforms, which probably affects the accuracy of first-break picking.

Convolutional Neural Networks (CNNs) generally including the convolution, pooling and fully-connected layers is a well-known deep learning architecture inspired by the natural visual perception mechanism of living creatures (see [10], [11]). In recent years, the deep CNNs has been widely developed and applied to a variety of fields (see [11]–[16]), such as speech recognition, natural language processing, genetic determinants of disease, playing Atari games, remote sensing image classification and the game of Go, due to

For Peer Review

the features of local connectivity and parameter sharing. The CNNs can extract different features or attributes directly from images or signals by using its multiple convolution layers, and subsequently classify them via the fully-connected layer. Therefore, CNNs has the advantage of combining attribute extraction and classification in one network. Moreover, CNNs has a strong classification function for very large datasets that has been demonstrated to exceed human performance in some visual tasks (see [13], [15], [17]).

However, the CNNs is rarely applied to seismic waveform classification and first-break picking despite the above advantages or characteristics. In this letter, we investigate how CNNs can be adopted to classify time-space waveforms from seismic shot gathers and further pick first breaks. Apart from the introduction of CNNs architectures and some training details, we propose three quality factors (QCs) to qualitatively evaluate the quality of the chosen CNNs input samples and the corresponding labeled output classification. We also define a discriminant score function to visually classify seismic waveforms and introduce a workflow with three operations to pick the first break in the theory section. The synthetic and real data examples are then adopted to illustrate the performances of the CNNs-based seismic classifier and picker. Finally, the conclusions of this investigation and future work are discussed.

II. THEORY

Three separate sections are considered to introduce the basic theory of CNNs-based automated time-space waveform classification and first-break picking. The first section is CNNs architectures including the design of CNNs input and output patterns as well as the introduction of three types of layers in the network. The second is CNNs training involving how CNNs obtains the optimal weights and biases. In the final section, we will describe CNNs validation and generalization involving three QCs, a discriminant score function for classifying waveforms and a three-step workflow for picking the first break.

A. CNNs architectures

A shot gather typically includes a variety of wave types, such as direct wave, reflected wave, multiples, refracted wave, diffracted wave, surface wave and incoherent noise. However, we can simply classify them into first-break waves and non-first-break waves according to the arrival time of waves. Fundamentally, these two types of waves are discrepant in both time and space directions. Consequently, we choose a series of

For Peer Review

time-space sub-images centred by first-break or non-first-break points as the input samples and adopt two-element vectors to quantify their classification outputs. The ideal two-element outputs (1 0) or (0 1) correspond to the presence of the first break or non-first break, respectively.

In addition to the input image layer and the output classification layer, three main types of layers including convolutional, pooling and fully-connected layers are stacked between the input and output layers to construct the CNNs architectures. The convolutional layer is a key component of CNNs, involving a series of data-driven kernels or filters, where each kernel can extract a time-space attribute or feature map from seismic data. The pooling layer can help reduce the time-space dimension of the extracted attributes. The fully-connected layer can translate a set of attributes corresponding to each input sub-image into a classification output vector with two values between 0 and 1.

B. CNNs training

The process of CNNs training can be regarded as solving a complex nonlinear inverse problem using interactive forward propagation and back propagation. The aim of forward propagation is to calculate the classification output according to the designed network and the updated parameters (weights and biases), while the goal of back propagation is to update these parameters. Detailed descriptions of CNNs training have already been presented in the vast literature (see [11], [16]). Here, we review several key formulas with slight modifications to clarify CNNs training of seismic data.

The input sub-image or feature map in the convolutional layer is first convolved with learned kernels, and then the convolved results are input into a nonlinear activation function to calculate a series of (new) feature maps. For each input feature map, the k-th output feature map at the l-th layer, $\mathbf{X}_k^l$, is expressed as

$$\mathbf{X}_k^l = \left[1 + \exp\left(\mathbf{X}^{l-1} * \mathbf{W}_k^l + b_k^l \mathbf{E}\right)\right]^{-1}, \tag{1}$$

where matrix $\mathbf{X}^{l-1}$ represents a certain output feature map of the $(l-1)$-th layer or input feature map of the l-th layer, matrix $\mathbf{W}_k^l$ represents the k-th kernel or filter at the l-th layer consisted of several unknown weights, symbol * represents the convolution operator, scalar b_k^l represents the bias corresponding to the k-th kernel of the l-th layer, and $\mathbf{E}$ is a matrix with all entries of 1; the exponential operator $\exp(\cdot)$ in the sigmoid activation function introduces nonlinearities to the network. Note that the kernel $\mathbf{W}_k^l$ can be shared by all input feature

For Peer Review

maps to automatically extract a type of time-space attribute. Such a weight-sharing mechanism has several advantages; for instance, it can reduce network complexity and make CNNs easier to train. The pooling layer aims to achieve shift-invariance by reducing the resolution of the feature maps. It is usually placed after a convolutional layer. A typical average pooling is implemented by taking the average of every 3×3 neighborhood in the output feature map of the preceding convolution layer to output a low resolution and low dimension feature map. The generated feature maps are input into a fully-connected layer to calculate a two-element classification output vector $\mathbf{o}^{cal}$, which is given as

$$\mathbf{o}^{cal} = \left[1 + \exp\left(\mathbf{W}^{FCL}\mathbf{x} + \mathbf{b}^{FCL}\right)\right]^{-1}, \tag{2}$$

where $\mathbf{W}^{FCL}$ is an unknown weight matrix, $\mathbf{x}$ is a column vector generated by arranging all final abstract feature maps, and $\mathbf{b}^{FCL}$ is a column vector including two biases.

The main task of CNNs training is to update the above weights and biases to minimize the error between the forward calculated classification and the target label classification for N training samples, which is defined as the following loss function

$$O\left(\mathbf{W}^l, \mathbf{b}^l\right) = \sum_{i=1}^{N} || \mathbf{o}^{cal}\left(\mathbf{W}^l, \mathbf{b}^l\right) - \mathbf{o}^{known} ||_2^2, \tag{3}$$

where $\mathbf{W}^l$ (l=1,2,…,L) represents all weights at the l-th layer, $\mathbf{b}^l$ represents all biases at the l-th layer, the L-th layer represents the final fully-connected layer, and $\mathbf{o}^{known}$ is the target label classification quantified as (1 0) or (0 1). The loss function is distinctly differentiable, since the l_2 norm and exponential function are both differentiable. The differentiable nonlinear function, therefore, is readily solved by using a conventional back propagation algorithm with the following parameter update expression

$$\mathbf{m}^l \leftarrow \mathbf{m}^l - \lambda \frac{\partial O}{\partial \mathbf{m}^l}, \tag{4}$$

where $\mathbf{m}^l$ are either the weights or biases, λ is the learning rate, and the derivatives $\partial O / \partial \mathbf{m}^l$ are obtained by using the chain rule from the L-th layer to the l-th layer.

C. CNNs validation and generalization

A shot or several shot gathers with carefully manually picked first break can be chosen to validate the trained CNNs, and in turn, it can probably help modify CNNs architectures or optimize the weights and biases

For Peer Review

of the network. We can calculate the two-element classification output $\mathbf{o}^{cal}$ of each sub-image from shot gathers by using forward-propagation Equations (1) and (2). When $\mathbf{o}^{cal}=(o_1(t,x)\ o_2(t,x))$ is closer to (1 0), the centre point (t,x) of the corresponding sub-image can be classified as first break. Otherwise, when $\mathbf{o}^{cal}$ is closer to (0 1), point (t,x) can be interpreted as non-first break. To provide a single indication for classifying waveforms, we define a discriminant score function as

$$F(t,x)=|o_1(t,x)-1|+|o_2(t,x)|. \tag{5}$$

The trough in $F(t,x)$ corresponds to a characteristic change in waveform, and its minimum indicates the first break. If the change is similar to a training time-space waveform centred by the labeled first break, the trough value should be close to 0. In essence, the value size of $F(t,x)$ decides the similarity between a time-space waveform change from tested (validated or generalized) data and the training time-space waveform samples labeled as first break. Therefore, the tested time-space sub-images corresponding to small $F(t,x)$, usually less than 1, can be roughly classified into first break, whereas those more than 1 can be interpreted as non-first break.

We subsequently pick the first break from the calculated discriminant image $F(t,x)$ by sequentially using a threshold, the first local minimum rule of every trace and a median filter. The role of a threshold, usually set to 1, is to detect first breaks including false first breaks. The sub-images corresponding to these false first breaks are usually similar to some training time-space waveform samples labeled as first break more or less. The first local minimum rule of each trace is then employed to limit the detection of some false first breaks, essentially taking advantage of the early arrival property of real first-break waves. Finally, a median filter operation is utilized to take the spatial coherence property of real first-break waves into account, thereby further improving the accuracy of first-break picking.

During the validation phase, three QC rules including (1) the separability of the referenced first-break classification appearance represented by the careful manual-picking first break and the other classification appearances (false first-break and non-first-break classification), (2) the match degree between the CNNs-based automatic-picking first break and the manual-picking first break, and (3) the quantity and randomness of false first breaks, are considered to evaluate the quality of the chosen CNNs input samples and output classification or the trained CNNs structure. Consequently, we can purposefully adjust CNNs training

For Peer Review

input samples along with the corresponding label outputs to optimize the trained CNNs architectures until the QC rules meet by testing several shot gathers. The finally trained optimal CNNs can be generalized to all other shot gathers.

III. Examples

A synthetic data example and a real data example are adopted to illustrate the performances of CNNs in classifying seismic waveforms and picking first break. For these two data examples, the CNNs input is 2D time-space amplitude data from shot gathers with 47 time samples and 11 space traces, and the CNNs output is the classification result with a size of 2×1 corresponding to the centre point of the CNNs input. Two convolution layers, each including 6 and 12 kernels with a size of 3×3, an average pooling layer with 6 panels of size 3×3, and a fully-connected layer with 156 neurons, are orderly connected between the input and output layers. For CNNs training, the initial weights of the network are randomly assigned, and all biases are initialized to zero. For first-break picking, the threshold value is set to 1. For the sake of simplification, two shot gathers for each example are chosen, where one is used to both train the CNNs and validate the trained network, and the other is used to illustrate the generalization performance of the trained CNNs.

A synthetic shot gather with 2455 time samples and 330 space traces [Fig. 1(a), 1(c) or 1(e)] is first employed to illustrate the influence of the chosen input and output patterns during training on waveform classification and first-break picking. We discuss three patterns here, as denoted in Fig. 1(a), 1(c) and 1(e). Red and blue lines are chosen as the centre points of CNNs input sub-image samples, and labeled as first-break and non-first-break classification outputs, which are mathematically expressed as (1 0) and (0 1), respectively. Fig. 1(a) involves 1420 sub-image samples as the CNNs input, where 330 images are labeled as first break, which is carefully manually picked from both the direct wave and the refracted wave of the shot gather. The manual-picking first break (red line) is also considered as a reference to assess the effectiveness of CNNs-based automatic waveform classification and first-break picking. In this case, we utilize all accurate first-break points and some non-first-break points associated with different representative time-space waveforms as correct labels to train CNNs. Fig. 1(c) involves 1320 sub-images as the CNNs input, where 330 images are labeled first break corresponding to the result of the referenced first break moving down 30 time samples. In this case, the given labeled first-break classification output is inaccurate. Fig. 1(e) involves 1350

For Peer Review

images as the input, where 260 images are labeled as first break mainly corresponding to that of the direct wave. Consequently, the input samples lack representative time-space waveforms related to the refracted waves for this case.

After CNNs input samples and output classification are devised to train the CNNs structure, the current built optimal network can be applied to the tested shot gather to classify all its time-space sub-images, pick the first break and further evaluate the quality of the trained CNNs. Fig. 1(b), 1(d) and 1(f) show waveform classification and first-break picking results, which are predicted via CNNs trained from the three input and output patterns of Fig. 1(a), 1(c) and 1(e), respectively. Comparing Fig. 1(b), 1(d) and 1(f), the following can be observed:

1) Fig. 1(b) presents the best first-break picking result (blue dashed line), which is consistent with the reference (red line). Although there is a slight false appearance of first-break classification below the referenced first break and above about 4 s, there is a good separation feature between these false appearances and those first-break classification appearances near the reference.
2) Fig. 1(d) presents the worst waveform classification with the most false first-break classification appearances, and the worst first-break picking result (blue dashed line). Note that the first break picked from CDP 26 to CDP 340 is consistent with the given incorrectly labeled first break.
3) Fig. 1(f) presents classification and first-break results from CDP 1 to CDP 271 comparable to Fig. 1(b), but waveform classification from CDP 272 to 340 is easily confused, and the first break picked within this CDP range shows a great deviation from the corresponding referenced first break.

Based on the comparisons of these results, we choose a CNNs structure trained from the input and output pattern of Fig. 1(a) to further test another shot gather [Fig. 2(a)], and conclusively validate its generalization performance. As Fig. 2(b) shows, we can see that there is an obvious separation among false first-break, non-first-break and those first-break classification appearances approximatively consistent with the reference (red line); in addition, there is a good match between CNNs-based picking first break (blue dashed line) and the reference.

Next, a real land shot gather data example [Fig. 3] is used to test the application potential of the CNNs-based method for waveform classification and first-break picking. Fig. 3(a) is a gather with 4955 time samples and 239 space traces chosen for training and validating the CNNs structure, where red and blue lines

For Peer Review

are designed as a set of the centre points of CNNs input sub-images and classified into first-break and non-first-break CNNs outputs, respectively. We adopt 1434 sub-image samples as the input, where 239 images are labeled as first break that are carefully manually picked from both the direct and refracted waves of the gather. The manual-picking first break (red line) is considered as a reference to evaluate the trained CNNs. Fig. 3(b) is CNNs-based waveform classification and first-break picking result predicted from all time-space sub-images in Fig. 3(a). Although there is some false random appearances of first-break classification, there is a clear separation among the false first-break, non-first-break and those first-break classification appearances near the reference, and thus it gives rise to an approximate match between the CNNs-based picking first break (blue dashed line) and the reference (red line). The trained CNNs is then generalized to another shot gather [Fig. 3(c)]. Fig. 3(d) shows a CNNs-based waveform classification map along with the first-break picking result. As expected, there is also good separation among the false first-break, non-first-break and first-break classification appearances near the manual-picking first break (also defined as the reference), in addition to a good match between the CNNs-based picking first break (blue dashed line) and the reference (red line).

IV. Conclusion

The CNNs can be trained to build an optimal nonlinear mapping model between seismic time-space sub-image inputs and the labeled first-break and non-first-break classification outputs. The trained model is dependent on the quality of the chosen inputs and the corresponding labeled classification outputs, but it can be evaluated and further adjusted via three QCs rules, which are (1) the separability between the referenced first-break classification appearance represented by the careful manual-picking first break and the other classification appearances, (2) the match degree between CNNs-based automatic picking first break and manual-picking first break, and (3) the quantity and randomness of false first break. When the input sub-image samples are chosen representatively and sufficiently, the corresponding labeled classification outputs are accurately given to train CNNs, and all time-space sub-images corresponding to the first-break type are not too similar to those corresponding to the non-first-break type, the trained CNNs is generally effective for classifying seismic waveform and picking first break. As the synthetic and real shot data examples illustrate, CNNs is a well-performing automatic classifier and picker without the pre-processing step of attribute extraction.

For Peer Review

The CNNs-based waveform classification and first-break picking method can be readily extended to the other time-space waveform datasets, such as micro-seismic, earthquake or ground penetrating radar datasets. As future work we plan to extend the method to process massive and higher-dimensional seismic datasets, and further investigate CNNs architectures. We also plan to test the robustness of the method to strong noise near the first-break waves.

ACKNOWLEDGMENT

This work was financially supported by the National Natural Science Foundation of China (41674127), the National Key Basic Research Development Program (2013CB228600), and the Science Foundation of China University of Petroleum, Beijing (2462015BJB04).

REFERENCES

[1] M. Shirzaei, W. L. Ellsworth, K. F. Tiampo, P. J. González, and M. Manga, "Surface uplift and time-dependent seismic hazard due to uid injection in eastern Texas," *Science*, vol. 353, no. 6306, pp. 1416–1419, Sep. 2016.

[2] J. Akram, and D. W. Eaton, "A review and appraisal of arrival-time picking methods for downhole microseismic data," *Geophys.*, vol. 81, no. 2, pp. KS71–KS91, Mar. 2016.

[3] S. M. Mousavi, C. A. Langston, and S. P. Horton, "Automatic microseismic denoising and onset detection using the synchrosqueezed-continuous wavelet transform," *Geophys.*, vol. 81, no. 4, pp. V341–V355, Jul. 2016.

[4] H. Dai, and C. MacBeth, "The application of back-propagation neural network to automatic picking seismic arrivals from single-component recordings," *J. Geophys. Res.*, vol. 102, no. B7, pp. 15105–15113, Jul. 1997.

[5] R. Di Stefano, F. Aldersons, E. Kissling, P. Baccheschi, C. Chiarabba, and D. Giardini, "Automatic seismic phase picking and consistent observation error assessment: application to the Italian seismicity," *Geophys. J. Int.*, vol. 165, no. 1, pp. 121–134, Apr. 2006.

[6] M. D. McCormack, D. E. Zaucha, and D. W. Dushek, "First-break refraction event picking and seismic data trace editing using neural networks," *Geophys.*, vol. 58, no. 1, pp. 67–78, Jan. 1993.

[7] D. Maity, F. Aminzadeh, and M. Karrenbach, "Novel hybrid artificial neural network based autopicking workflow for passive seismic data," *Geophys. Prospecting*, vol. 62, no. 4, pp. 834–847, May 2014.

[8] C. Castellazzi, M. K. Savage, E. Walsh, and R. Arnold, "Shear wave automatic picking and splitting measurements at Ruapehu volcano, New Zealand," *J. Geophys. Res. Solid Earth*, vol. 120, no. 5, pp. 3363–3384, May 2015.

[9] S. M. Mousavi, S. P. Horton, C. A. Langston, and B. Samei, "Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression," *Geophys. J. Int.*, vol. 207, no. 1, pp. 29–46, Jul. 2016.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

For Peer Review

[11] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, and G. Wang, "Recent advances in convolutional neural networks," *varXiv preprint arXiv*, 1512.07108, Jan. 2017.

[12] G. Hinton, et al., "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Process Mag.*, vol. 29, no. 6, pp. 82–97, Oct. 2012.

[13] V. Mnih, et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[14] H. Y. Xiong, et al., "The human splicing code reveals new insights into the genetic determinants of disease," *Science*, vol. 347, no. 6218, pp. 1254806, Jan. 2015.

[15] D. Silver, et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[16] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, Oct. 2017.

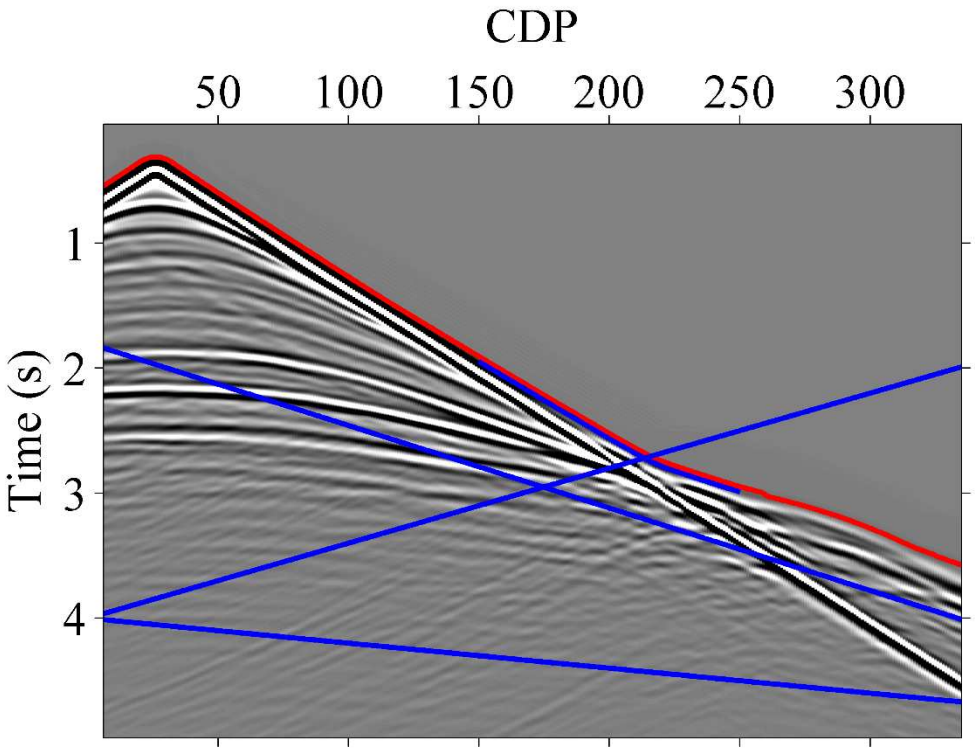
[17] O. Russakovsky, et al., "Imagenet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

Fig. 1. The influence of different seismic sub-image samples chosen in the synthetic shot gather (a, c and e) on CNNs-based waveform classification and first-break picking (b, d and f). Red and blue lines in (a), (c) and (e) represent three different sets of the centre points of all chosen training samples, which are adopted to construct the input of CNNs. The samples corresponding to red lines in (a), (c) and (e) are classified as first break and are labeled as a two-element output of (1 0), whereas those corresponding to blue lines are classified as non-first break and labeled as an output of (0 1). The red lines in (b), (d) and (f) represent the manual-picking first break, which is the same as the red line of (a). The blue dashed lines in (b), (d) and (f) are the CNNs-based picking first break. Different samples show different waveform classification effects and first-break picking effects. When the sub-images centred by the wrong labeled first break [(c)] or those involving too little first-break classification of the refracted wave [(e)] are adopted as the training samples, waveform classification and first-break picking are relatively poor [(d) and (f)].

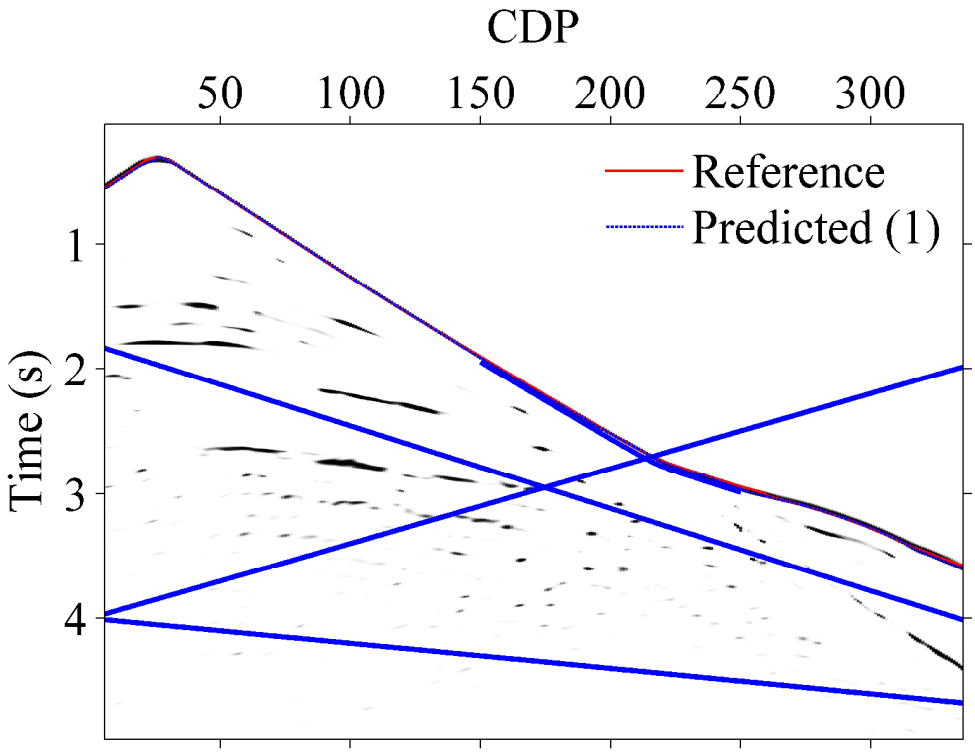
Fig. 2. The generalization of another synthetic seismic shot gather (a) for classifying waveform and picking first break (b) by using the CNNs model trained from the chosen input and output pattern in (a). The CNNs-based predicted first break (blue dashed line) matches with the careful manual-picking first-break reference (red line) well.

Fig. 3. The real seismic shot gather example for waveform classification and first-break picking. (a) A gather along with the labeled first-break (red line) and non-first-break (blue lines) classification used to train CNNs, (b) the CNNs-based waveform classification and first-break result of (a), (c) another gather for a generalization test, and (d) the CNNs-based waveform classification and first-break result of (c). The CNNs-based automatic picking first breaks [blue dashed lines in (b) and (d)] are consistent with the careful manual-picking first-break references [red lines in (b) and (d)].

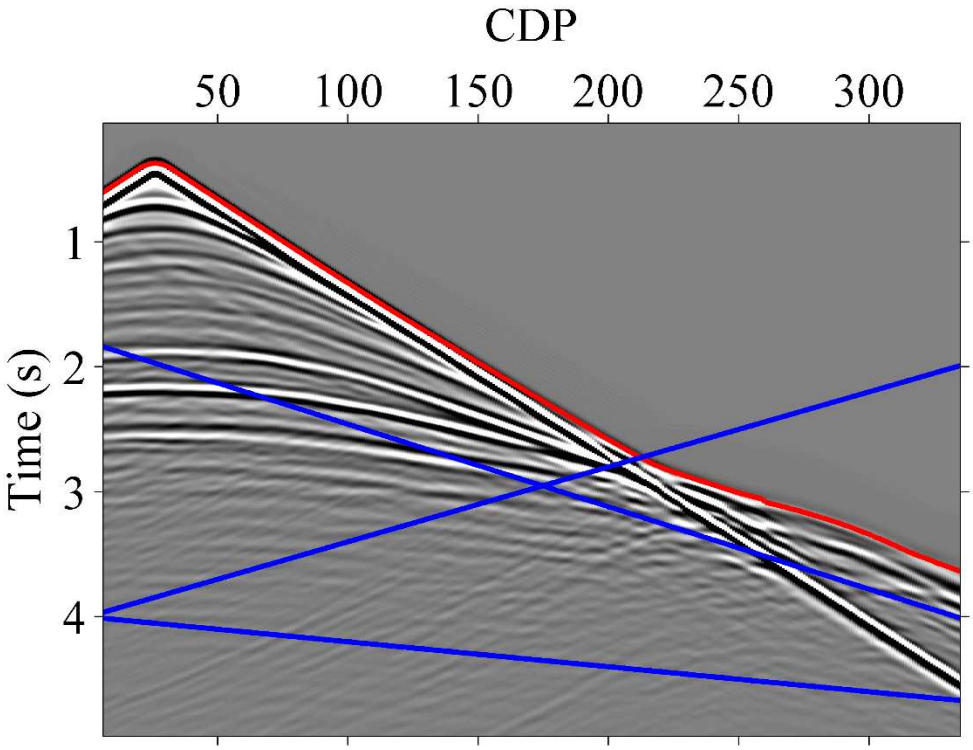
For Peer Review


CDP
50
100
150
200
250
300
Time (s)
1
2
3
4

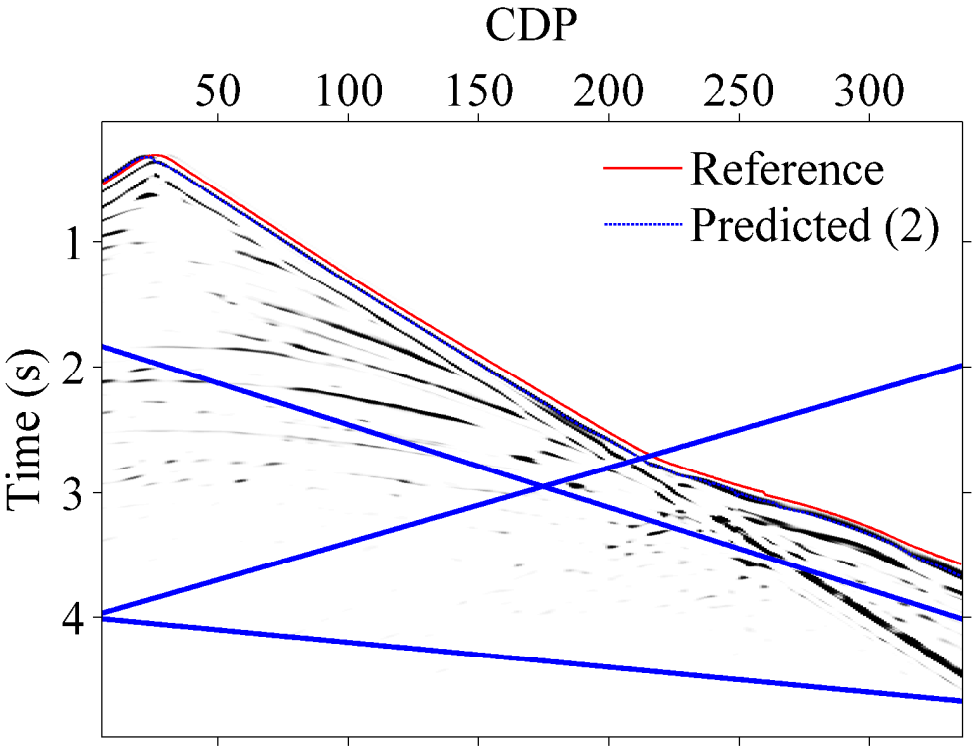
(a)


CDP
50
100
150
200
250
300
Reference
Predicted (1)
Time (s)
1
2
3
4

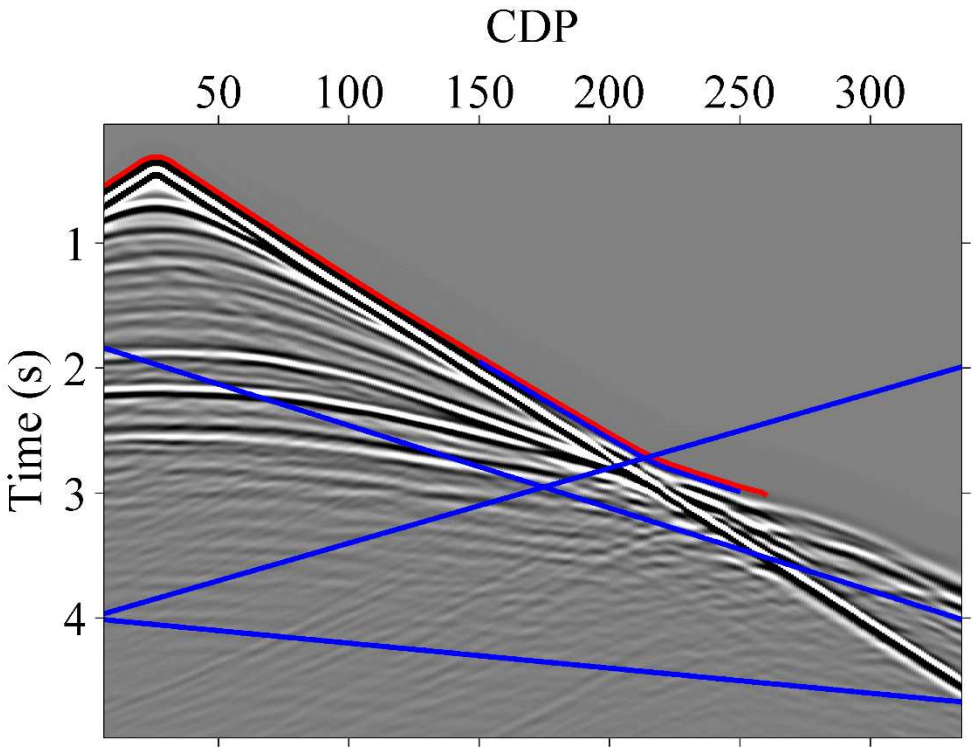
(b)


CDP
50
100
150
200
250
300
Time (s)
1
2
3
4

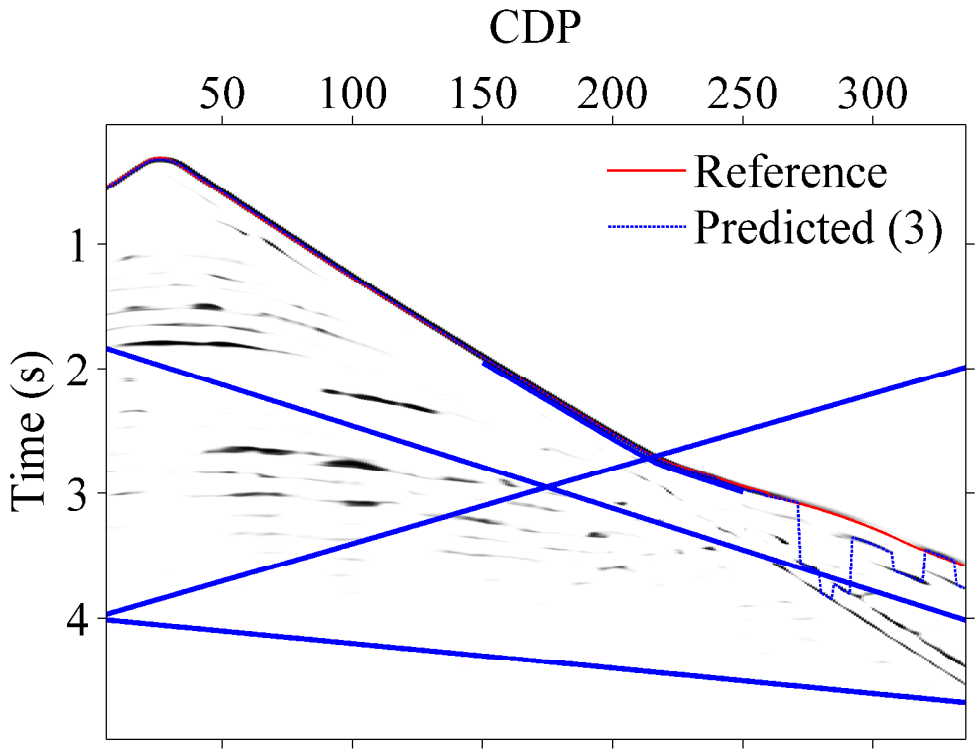
(c)


CDP
50
100
150
200
250
300
Reference
Predicted (2)
Time (s)
1
2
3
4

(d)

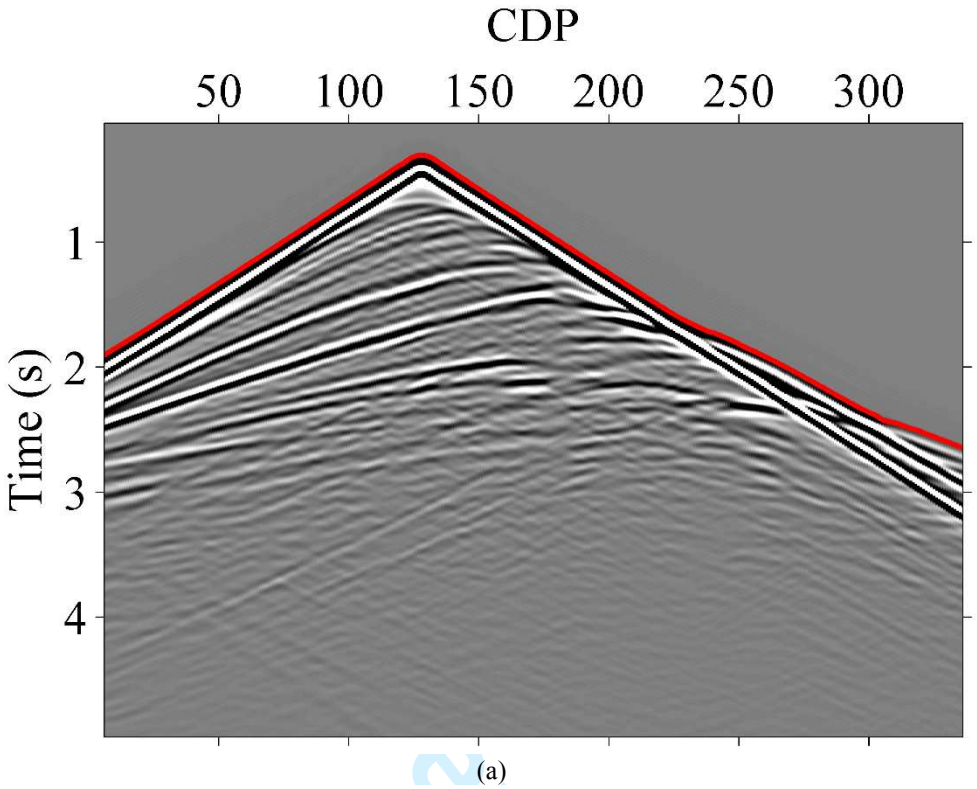

CDP
50
100
150
200
250
300
Time (s)
1
2
3
4

(e)

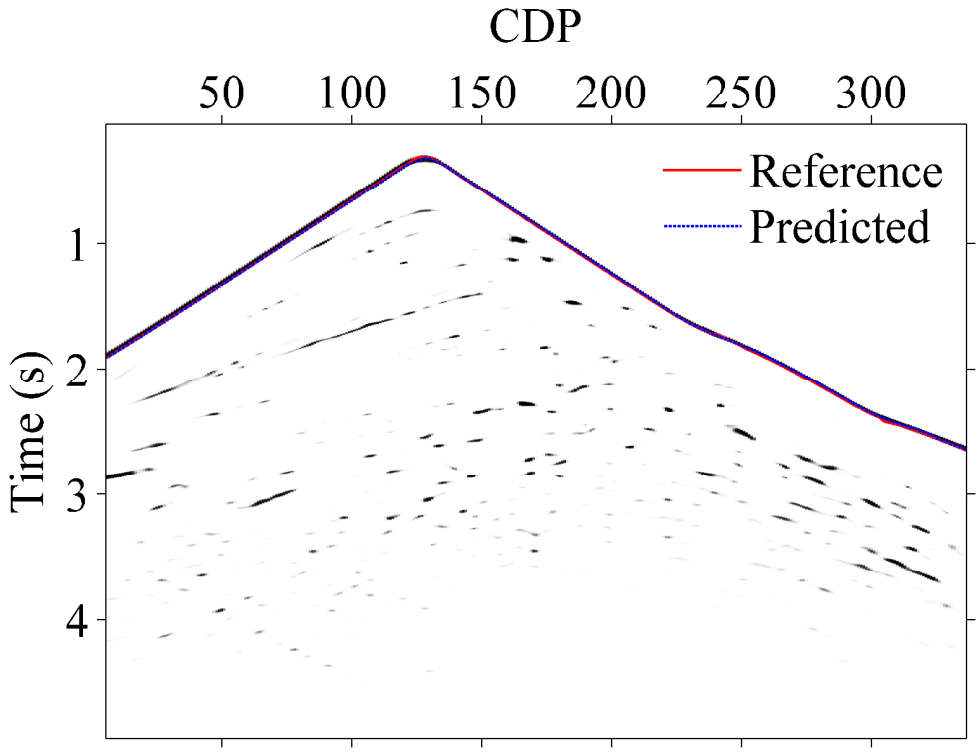

CDP
50
100
150
200
250
300
Reference
Predicted (3)
Time (s)
1
2
3
4

(f)

Figure 1

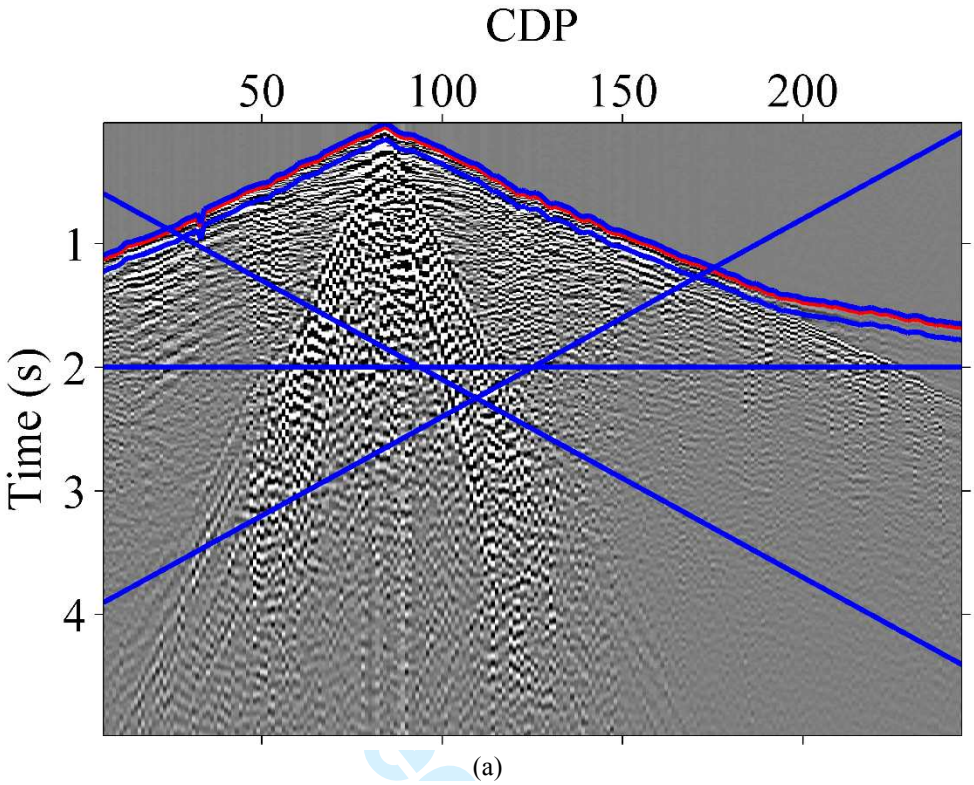

CDP
50
100
150
200
250
300
Time (s)
1
2
3
4

(a)

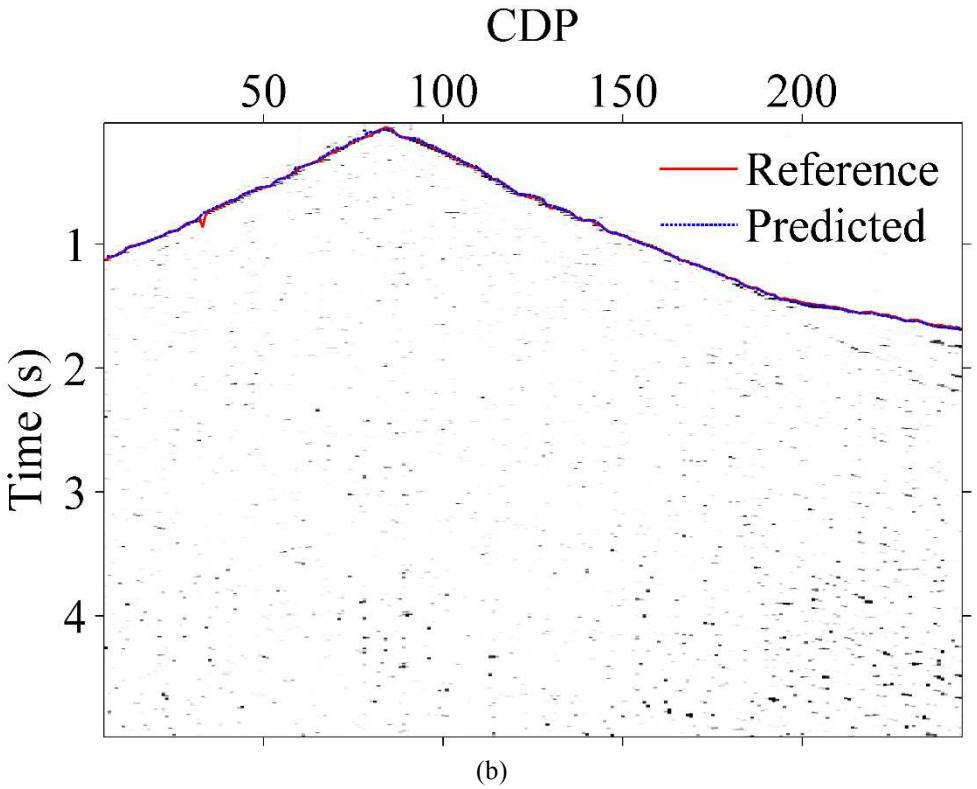

CDP
50
100
150
200
250
300
Time (s)
1
2
3
4
Reference
Predicted

(b)

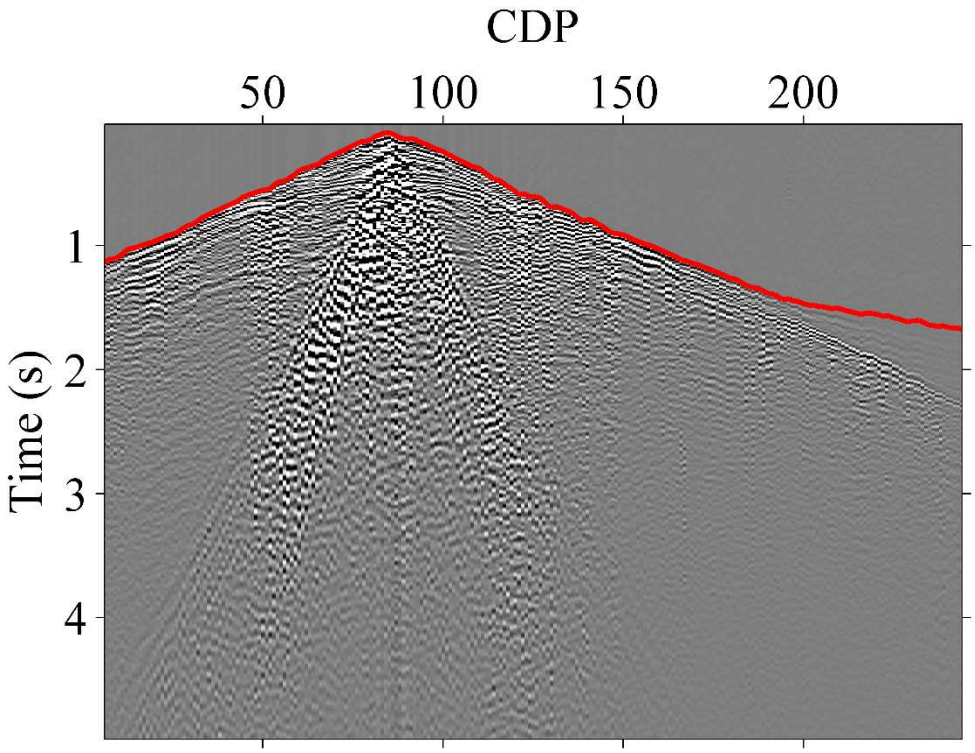
Figure 2


CDP
50
100
150
200
Time (s)
1
2
3
4

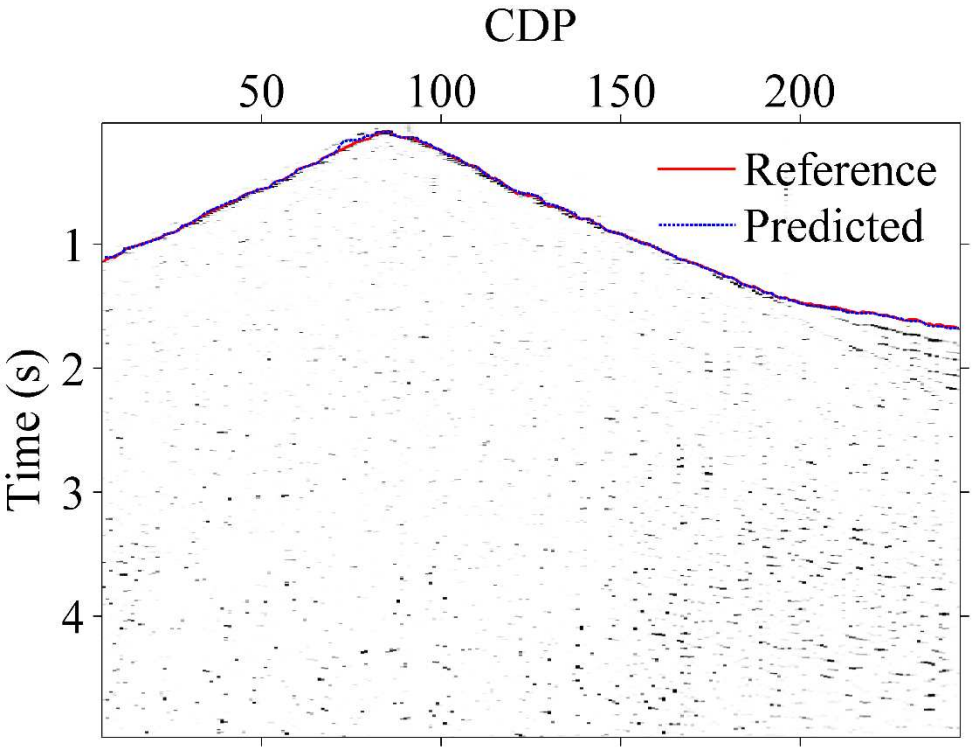
(a)


CDP
50
100
150
200
Reference
Predicted
Time (s)
1
2
3
4

(b)


CDP
50
100
150
200
Time (s)
1
2
3
4

(c)


CDP
50
100
150
200
Time (s)
1
2
3
4
Reference
Predicted

(d)

Figure 3