ACCELERATING THE U-NET BASED ADAPTIVE SUBTRACTION WITH TRANSFER LEARNING FOR REMOVING SEISMIC MULTIPLES IN FIELD DATA

ZHONGXIAO LI [1,2], JIAHUI MA[3], ZHEN QI [3], LINWEI ZHANG [3], FEI XIE [1,2] and JINLONG YANG [4]

[1] *State Key Laboratory of Shale Oil and Gas Enrichment Mechanisms and Effective Development, Beijing 100083, P.R. China. thulzx@163.com*
[2] *Sinopec Key Laboratory of Seismic Elastic Wave Technology, Beijing 100083, P.R. China.*
[3] *Department of Electronic Engineering, School of Electronic Information, Qingdao University, Qingdao 266071, P.R. China.*
[4] *Sinopec Geophysical Research Institute, Nanjing 211103, P.R. China.*

(Received May 3, 2023; accepted July 6, 2023)

ABSTRACT

Li, Z.X., Ma, J.H., Qi, Z., Zhang, L.W., Xie, F. and Yang, J.L., 2023. Accelerating the U-net based adaptive subtraction with transfer learning for removing seismic multiples in field data. *Journal of Seismic Exploration, 32: 373-384.*

Adaptive subtraction is essential for removing multiples effectively after the modeling of seismic multiples. The U-net based method has been proposed to conduct adaptive subtraction under the frame of non-linear regression. Compared with the linear regression method the U-net based method can remove more residual multiples and better protect primaries at the cost of higher computational cost. To accelerate the U-net based method for processing the field data efficiently we introduce transfer learning into the U-net based method in this paper. The adjacent seismic gathers have similar mismatches between the modeled multiples and true multiples. On the basis of the transfer learning theory the network parameters of U-net estimated in the previous data part can be used as the initial network parameters of U-net for the next data part. In this way the U-net based method can be accelerated with decreased epoch numbers. While achieving similar accuracy the accelerated U-net based method can decrease the computational time by 40% compared with the non-accelerated U-net based method without transfer learning in the field data example.

KEY WORDS: adaptive subtraction, U-net, transfer learning, acceleration.

0963-0651/23/$5.00

INTRODUCTION

Generally, we use seismic primaries for imaging and inversion in seismic exploration (Verschuur, 2013). Before recorded by the geophones, seismic primaries only reflect once under the surface while multiples reflect more than once (Berkhout and Verschuur, 1997; Verschuur and Berkhout, 1997). Seismic multiple removal is an important step before imaging or inversion. We can use the method of Wave Equation Modeling (Wiggins, 1988), Inverse Scattering Series (Weglein et al., 1997) and Surface Related Multiple Elimination (SRME) (Dragoset et al., 2010) to model multiples, which have temporal and spatial shift, amplitude and wavelet mismatch, phase discrepancy and so on compared with the true multiples (Li et al., 2021). Therefore, we conduct adaptive subtraction rather than direct subtraction of the modeled multiples from the original recorded data.

Adaptive subtraction is usually conducted under the frame of linear regression (LR) and 2D matching filter is solved with the assumption that primaries satisfy the L2 or L1 norm minimization constraint (Guitton and Verschuur, 2004; Li and Li, 2018). Compared with L2 norm, L1 norm can better express the super-Gaussian distribution of primaries. If strong primaries exist or multiples and primaries have overlap, the 2D matching filter estimated with L1 norm can better remove residual multiples and reduce distorted primaries than the 2D matching filter estimated with L2 norm (Li and Li, 2018). However, the 2D matching filter solved in the frame of LR has limited ability in removing the complicated discrepancies between the modeled multiples and true multiples effectively. It may harm primaries or leave residual multiples in complex geology condition (Jiang and Lu, 2020). Moreover, the 2D matching filter is solved independently in every 2D data window by the traditional LR method.

The U-net based method subtracts the modeled multiples adaptively from the original recorded data in the frame of non-LR (Li et al., 2021; Liu et al., 2022). In Li et al. (2021) the U-net based method outperforms the LR method in terms of better primary protection and residual multiple removal. The field data usually have a lot of gathers. In Li et al. (2021) how to process the field data efficiently is not discussed. For training U-net the seismic gathers are usually divided into overlapping 2D data windows. To express the complex discrepancies between the modeled multiples and true multiples of all 2D data windows simultaneously, U-net should have large network parameter capacity. In this case U-net has many layers of down-sampling and up-sampling computation. This may cause huge computational cost in training U-net, especially for processing field data.

In this paper we partition all the seismic gathers into many data parts. We train U-net and then conduct adaptive subtraction in every data part. In this case the U-net has small parameter capacity and this is benefit for training U-net efficiently. The adjacent parts have similar discrepancies between the modeled multiples and true multiples. On the basis of the transfer learning theory (Siahkoohi et al., 2019) the network parameters of U-net estimated in the previous data part are used as the initial network parameters of U-net for the next data part. In this way we do not need to train U-net in every data part with random initial network parameters. The U-net based adaptive subtraction is accelerated with decreased epoch numbers by using transfer learning.

The paper is organized as follows. We review the U-net based adaptive subtraction and then give the flow chart of using transfer learning for acceleration. After that the effectiveness of the proposed U-net based method with transfer learning is verified in the field data example. Finally the conclusion is given.

METHOD

For the U-net based adaptive subtraction its mathematical model is described with the following equation (Li et al., 2021):

$$\mathbf{P} = \mathbf{S} - N(\tilde{\mathbf{M}}, \boldsymbol{\Theta}) \quad , \tag{1}$$

where the modeled multiples, original recorded data and estimated primaries in a 2D data window are represented with the matrixes $\tilde{\mathbf{M}}$, $\mathbf{S}$ and $\mathbf{P}$ with size of $n \times n$. $\boldsymbol{\Theta}$ represents the network parameters of U-net $N(g)$, which are estimated during network training. The U-net architecture is shown in Fig. 1 and is described in detail in Li et al. (2021).

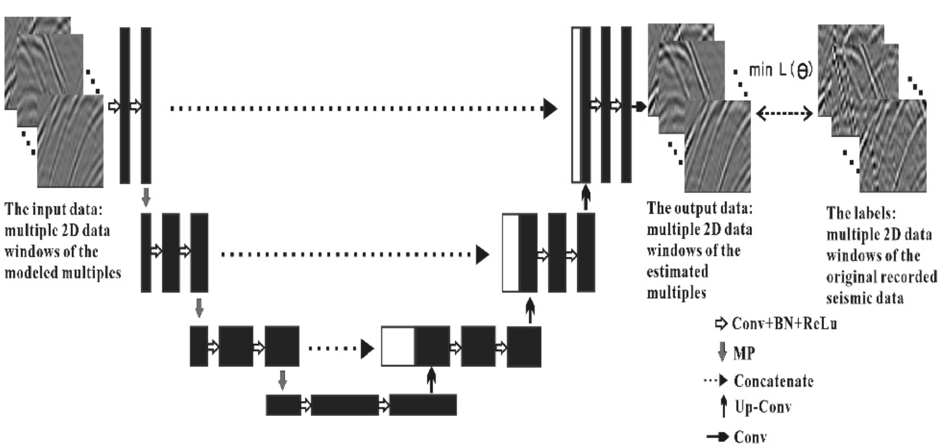

Fig. 1. U-net architecture (Li et al., 2021). The input data, labels and output data are the modeled multiples, original recorded data and estimated multiples in 2D data windows, respectively.

The modeled multiples and original recorded data are used as the input data and labels, respectively. We define the following loss function (Li et al., 2021)

$$L(\boldsymbol{\Theta}) = \sum_{i=1}^{N} \left[\left\| \mathbf{S} - N\left(\tilde{\mathbf{M}}, \boldsymbol{\Theta}\right) \right\|_1 + \lambda \left\| \boldsymbol{\Theta} \right\|_2^2 \right], \tag{2}$$

where λ is the regularization factor. The L1 norm is defined as $\|\mathbf{P}\|_1 = \sum_{j=1}^{n} \sum_{k=1}^{n} \|P_{j,k}\|_1$, where $\mathbf{P} = \{P_{j,k}\}$, $j = 1,2,\ldots,n$, $k = 1,2,\ldots,n$. To minimize the loss function $L(\boldsymbol{\Theta})$ and estimate the network parameters $\boldsymbol{\Theta}$ the Adam algorithm (Kingma and Ba, 2015) is introduced during training. We choose the number of 2D data windows N as the batch size to update $\boldsymbol{\Theta}$ in one training iteration. The total iteration number in one epoch of the Adam algorithm is computed as $\lceil A/N \rceil$, where the operator $\lceil \cdot \rceil$ rounds the value to the nearest integer and A is the total number of 2D data windows.

The flow chart for the U-net based method (Li et al., 2021) is shown in Fig. 2. The seismic gathers are divided into overlapping 2D data windows. The input data and labels are chosen as multiple 2D data windows of the modeled multiples and original recorded data for U-net training. A non-LR model is constructed by the trained U-net for multiple removal. Then we estimate multiples by using the modeled multiples in 2D data windows as the input data of the trained U-net. After that the estimated multiples are subtracted directly from the original recorded data to obtain the estimated primaries. In Li et al. (2021) how to process the field data efficiently for removing seismic multiples is not described.

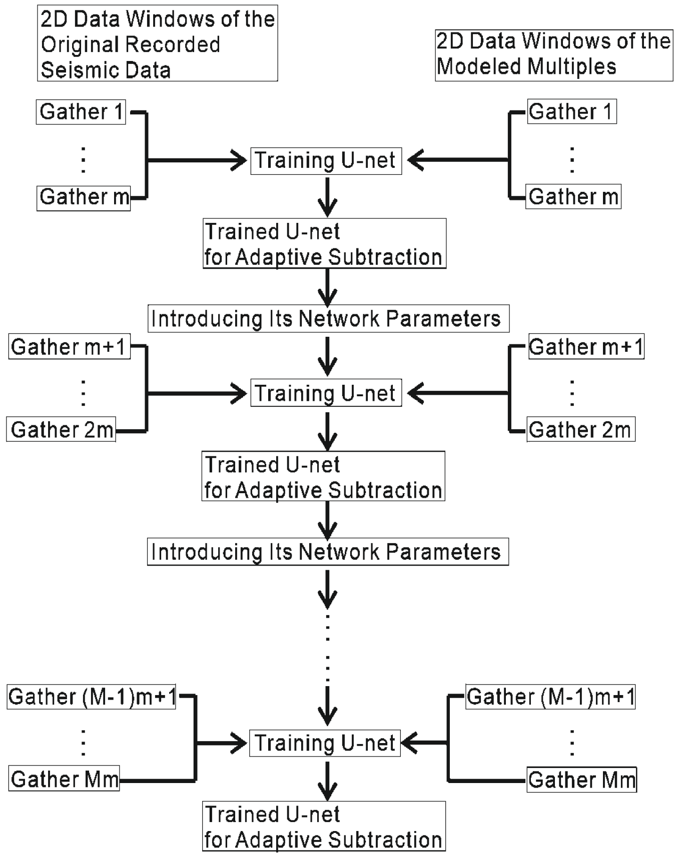

Fig. 2. Flow chart for the U-net based method (Li et al., 2021).

In this paper the seismic gathers are partitioned into M data parts sequentially. In every data part there are m gathers for training U-net. The adjacent parts have similar discrepancies between true multiples and modeled multiples. We introduce transfer learning (Siahkoohi et al., 2019) to use the network parameters of U-net estimated in the previous part as the initial parameters of U-net for the next part. In this way we do not need to train U-net with random initial network parameters in every data part. We denote the U-net based method with transfer learning as the accelerated method. Its flowchart is shown in Fig. 3. The accelerated method uses the epoch number as K in the first part and R in the other parts. Since transfer learning is used, R is much less than K. For U-net training the non-accelerated method uses the random initial network parameters. The accelerated U-net based method can reduce the iterative epoch numbers and improve the computational efficiency effectively compared with the non-accelerated U-net based method.

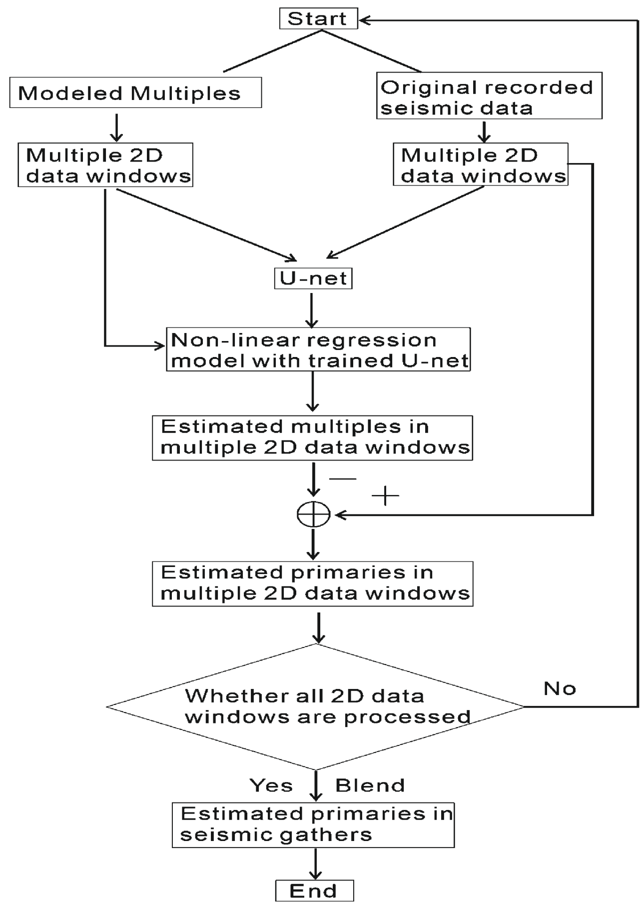

Fig. 3. Flow chart for the accelerated U-net based method with transfer learning.

EXAMPLES

In this section the field data is used to verify the validity of the proposed accelerated U-net based method.

There are 2241 common shot gathers in the field data, which has 240 traces in every gather and 3000 time samples in every trace. The temporal sampling interval is 2 ms. Fig. 4 shows the common offset gather of the original recorded data. Eight gathers with 500 traces in every gather from 930 ms to 1598 ms are shown in Fig. 4. Fig. 4a shows the first and second gathers, Fig. 4b shows the seventh and eighth gathers, Fig. 4c shows the eleventh and twelfth gathers and Fig. 4d shows the seventeenth and eighteenth gathers. 2D SRME gives the corresponding modeled multiples in Fig. 5. It is seen that the adjacent gathers have similar amplitude, phase, waveform and so on.

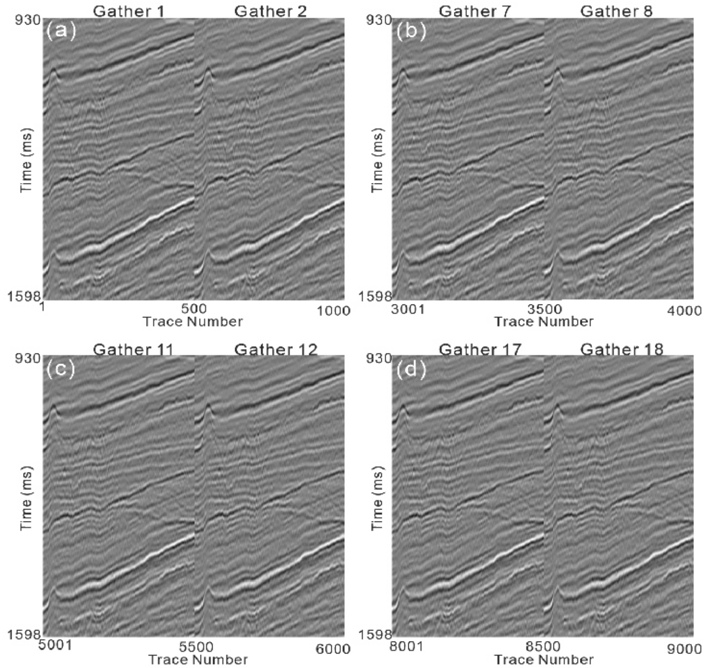

Fig. 4. Eight common offset gathers of the original recorded data.

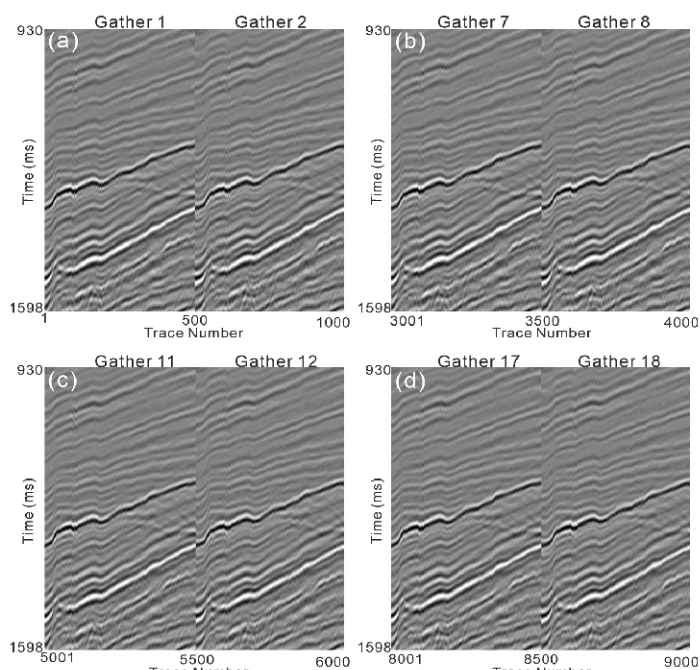

Fig. 5. Eight common offset gathers of the modeled multiples.

We choose $m = 2$, meaning that two gathers are used for U-net training in every data part. In the first data part the proposed accelerated U-net based method uses 45 epochs. The adjacent gathers have similar mismatches between the modeled multiples and true multiples. From the second data part the proposed accelerated U-net based method only uses 8 epochs to train U-net for seismic multiple removal. Since the non-accelerated U-net based method uses the random initial network parameters for U-net training of every data part, it needs 45 epochs to train U-net in every data part. Figs. 6a, 6b, 6c and 6d show the curves of the training loss value versus the epoch number in the first, fourth, sixth and nineth data part for the accelerated and non-accelerated U-net based methods. Both the accelerated and non-accelerated U-net based methods use the Gaussian random initial network parameters in the first data part. From Fig. 6 we can see that they obtain similar training loss value after the curve convergence in the four data parts. From the second data part the accelerated U-net based method uses very few epochs (8 epochs) compared with the non-accelerated U-net based method.

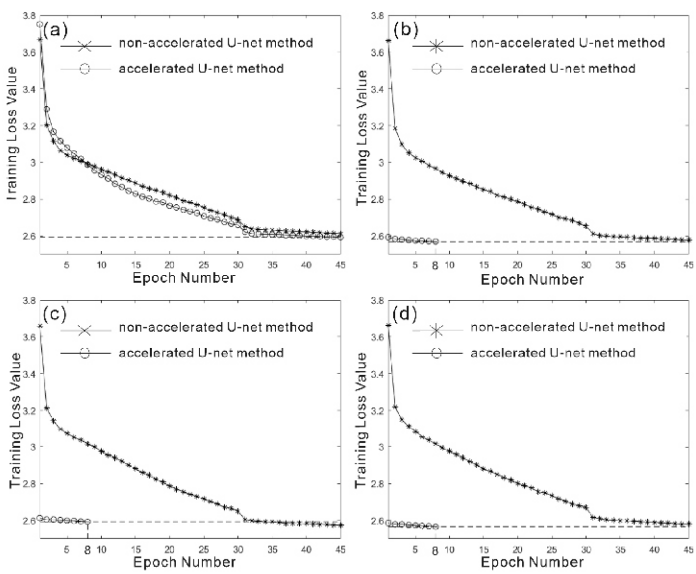

Fig. 6. The training loss curves in the first, fourth, sixth and nineth data part for the accelerated and non-accelerated U-net based methods.

The accelerated U-net based method estimates primaries in Fig. 7 and gives the removed multiples in Fig. 8. The non-accelerated U-net based method gives the estimated primaries in Fig. 9 and removed multiples in Fig. 10. The difference gathers in Fig. 11 are obtained with the subtraction of the gathers in Fig. 9 from the gathers in Fig. 7. Similar results are achieved by the accelerated and non-accelerated U-net based methods. The computational time of the accelerated and non-accelerated U-net based methods is 349 s and 1212 s, respectively. The accelerated U-net based method improves the computational efficiency by about 40% compared with the non-accelerated U-net based method.

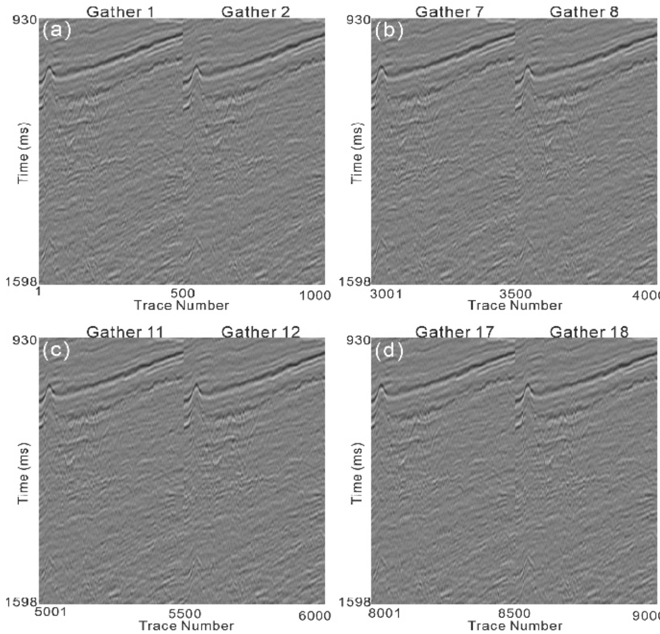

Fig. 7. Estimated primaries in eight common offset gathers of the accelerated U-net based method.

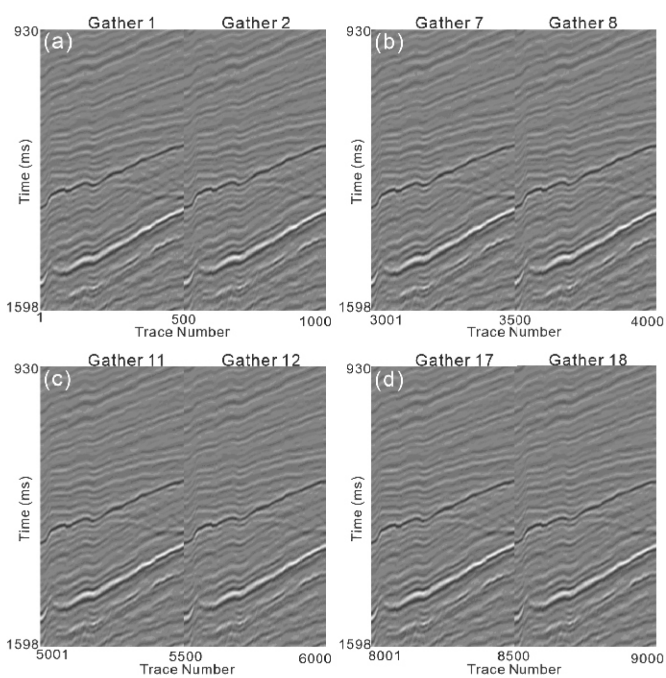

Fig. 8. Removed multiples in eight common offset gathers of the accelerated U-net based method.

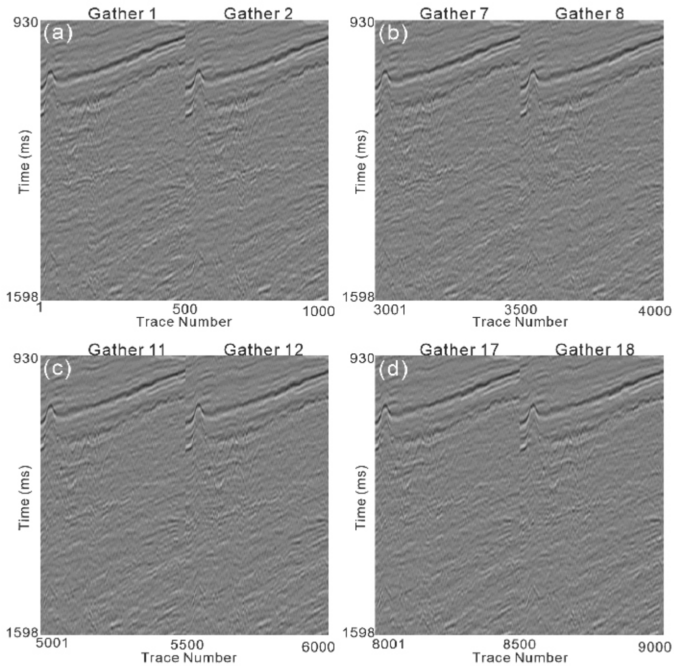

Fig. 9. Estimated primaries in eight common offset gathers of the non-accelerated U-net based method.

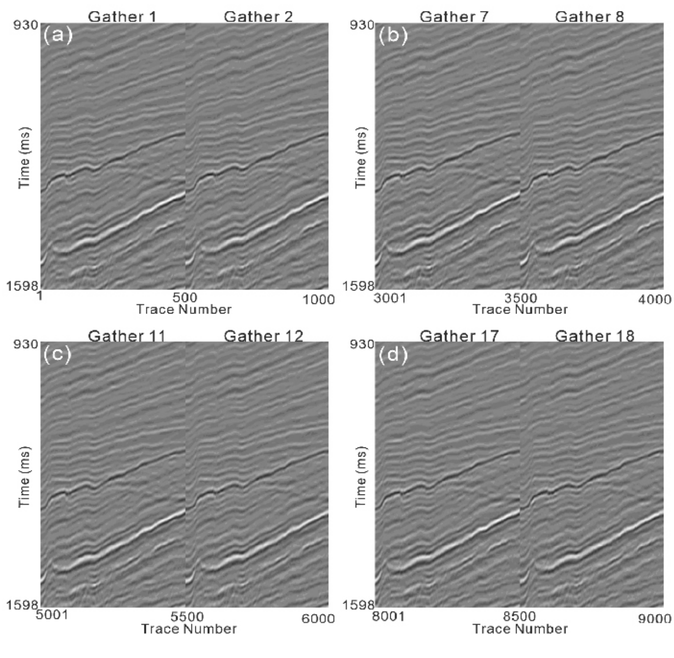

Fig. 10. Removed multiples in eight common offset gathers of the non-accelerated U-net based method.

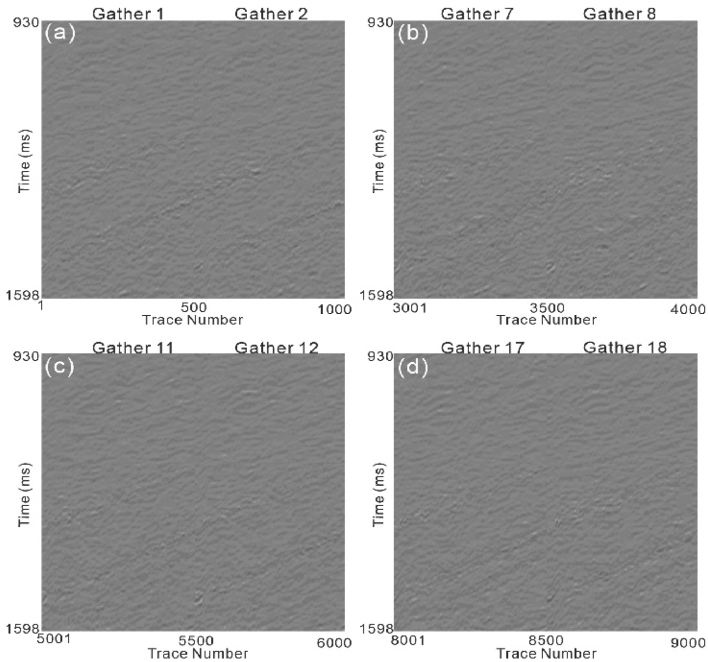

Fig. 11. Eight difference gathers by subtracting the gathers in Fig. 9 from the gathers in Fig. 7.

CONCLUSIONS

For efficient seismic multiple removal in field data transfer learning is introduced to accelerate the U-net based adaptive subtraction in this paper. We use the network parameters of U-net estimated in the previous data part as the initial parameters of U-net for the next data part. From the second data part the accelerated U-net based method uses very few epochs and is computationally efficient. The field data verifies the validity of the proposed accelerated U-net based method.

ACKNOWLEDGMENTS

The authors appreciate the support of Open Fund of Sinopec Key Laboratory of Seismic Elastic Wave Technology (No. 33550000-22-ZC0613-0286), which focuses on the seismic multiple separation method based on the interpretable closed-loop optimization algorithm driven neural network. Moreover, the authors appreciate the National Natural Science Foundation of China (No. 41804110), China Postdoctoral Science Foundation (No. 2022M723127) and Open Research Foundation of SINOPEC Key Laboratory of Geophysics (No. 33550006-22-FW0399-0020).

REFERENCES

Berkhout, A.J. and Verschuur, D.J., 1997. Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations. Geophysics, 62: 1586-1595.

Dragoset, B., Verschuur, E., Moore, I. and Bisley, R., 2010. A perspective on 3D surface-related multiple elimination. Geophysics, 75(5): A245-A261.

Guitton, A. and Verschuur, D.J., 2004. Adaptive subtraction of multiples using the L1-norm. Geophys. Prosp., 52: 27-38.

Jiang, B. and Lu, W., 2020. Adaptive multiple subtraction based on an accelerating iterative curvelet thresholding method. IEEE Transact. Image Process., 30: 806-821.

Kingma, D.P. and Ba, J.L., 2015. Adam: A Method for stochastic optimization. Internat. Conf. Learning Represent., San Diego, 1-15.

Liu, L., Hu, T., Huang, J. and Wang, S., 2022. Adaptive surface-related multiple subtraction based on convolutional neural network. IEEE Geosci. Remote Sens. Lett., https://ieeexplore.ieee.org/document/9584854.

Li, Z. and Li, Z., 2018. Accelerated 3D blind separation of convolved mixtures based on the fast iterative shrinkage thresholding algorithm for adaptive multiple subtraction. Geophysics, 83(2): V99-V113.

Li, Z., Sun, N., Gao, H., Qin, N. and Li, Z., 2021. Adaptive subtraction based on U-net for removing seismic multiples. IEEE Transact. Geosci. Remote Sens., 59: 9796-9812.

Siahkoohi, A., Louboutin, M. and Herrmann, F.J., 2019. The importance of transfer learning in seismic modeling and imaging. Geophysics, 84(6): A47-A52.

Verschuur, D.J. and Berkhout, A.J., 1997. Estimation of multiple scattering by iterative inversion, part II: Practical aspects and examples. Geophysics, 62: 1596–1611.

Verschuur, E., 2013. Seismic Multiple Removal Techniques: Past, Present and Future. EAGE, Houten.

Weglein, A.B., Gasparotto, F.A., Carvalho, P.M. and Stolt, R.H., 1997. An inverse-scattering series method for attenuating multiples in seismic reflection data. Geophysics, 62: 1975-1989.

Wiggins, J.W., 1988. Attenuation of complex water-bottom multiples by wave-equation-based prediction and subtraction. Geophysics, 53: 1527-1539.