Journal of Applied Geophysics 167 (2019) 172–179

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

Subsurface velocity inversion from deep learning-based data assimilation

Bo Mao *, Li-Guo Han, Qiang Feng, Yu-Chen Yin

College of Geo-Exploration Science and Technology, Changchun 130026, China

ARTICLE INFO

Article history:
Received 16 October 2018
Received in revised form 28 January 2019
Accepted 1 April 2019
Available online 4 June 2019

Keywords:
Deep learning
Data assimilation
Full waveform inversion
Prior velocity information

ABSTRACT

Data assimilation has widespread applications in meteorology and oceanography. This paper applies data assimilation for seismic exploration. Comparing with full waveform inversion, the distrust of the prior velocity information is preserved, to avoid the local minima caused by its inaccuracy. Prior velocity information is scarce in seismic exploration, but the introduction of deep learning methods makes it possible, to correct the major defects in the application of data assimilation in seismic velocity inversion. Finally, we design a simple salt body model, and we introduce deep learning to obtain prior velocity information and drive data assimilation for high-precision inversion of subsurface velocity. It is compared with the traditional full waveform inversion, to prove the effectiveness and advantages of this new process.

1. Introduction

The data assimilation theory is derived from meteorological (Anthes, 1974; Stauffer, 1990; Kalnay et al., 1996)and ocean (Derber and Rosati, 1989; Smedstad and O'Brien, 1991; Carton, 2000) monitoring systems, but few researches in seismic exploration (Dong et al., 2006; Liu et al., 2009). British mathematician Richardson first used the idea of assimilation in a weather forecast in 1922, laying a solid foundation (Richardson, 2007). Then, in 1950, the American meteorologist Charney used the subjective analysis method to determine the initial value, and successfully completed the first numerical weather forecast in human history on the world's first electronic digital computer (Charney et al., 2010). Until the mid-1980s, the variational data assimilation technique based on variational theory was raised (Barker et al., 1984; Dimet and Talagrand, 1986; Courtier et al., 1998, Courtier and Talagrand, 2010). Daley (1991) pointed out that variational assimilation is the development direction of data assimilation. The core idea of the variational data assimilation method is to use the minimum quadratic fitting and statistical theories for prior information such as background field error and observation field, and then use the global optimization algorithm to find the minimum value (Sasaki, 2007). This is similar to the full waveform inversion (FWI) theory in high precision seismic exploration (Tarantola, 1984; Pratt, 1999). Chen (2011) pointed out that the seismic inversion problem has much in common with the data assimilation method. They all compare simulation-based prediction with actual observations to make effective scientific inference, and it is necessary to continue the iterative cycle of the model.

In the process of applying data assimilation theory to seismic exploration, unlike the prior background error problem that traditional data assimilation needs to overcome (Bergthórsson and Bo, 1955; Cressman, 1959; Lorenc, 2010), the biggest difficulty is the lack of prior information (Asnaashari et al., 2013; Datta, 2016). Reliable prior information is generally derived from logging data, but this does not exist in the early stages of finding resources. The development of deep learning theory makes it possible to build reliable prior velocity models using only seismic data.

The application of artificial intelligence and deep learning to oil and gas exploration has been a hot topic in recent years (Guillen et al., 1949; Adler et al., 2017; AlRegib et al., 2018; Di et al., 2018), and has achieved certain results. Lewis and Vigh (2017) takes migrated seismic images as input, and the trained deep learning model outputs a probability of each image patch containing structural features, but no specific velocity distribution is obtained. Araya-Polo et al. (2017) first extracts features from of the gather data and uses these features as input to train the deep learning model. Although velocity model is finally output, the feature extraction in the first step contains too many human factors. Richardson (2018) embeds full-waveform inversion into the recurrent neural network (RNN) in deep learning and uses the powerful automatic gradient computing power of the deep learning system to speed up efficiency. Li et al. (2018) successfully uses deep learning for ground-roll noise attenuation. Wu et al. (2018) propose an automatic fault interpretation method by using convolutional neural networks. Sun and Demanet (2018) proposes a deep-learning-based bandwidth extension method and has good application in FWI.

In this paper, a convolutional neural network is used, and a zero-offset gather is directly used as the input data, whereas the output is the velocity model, and effectively avoids excessive

* Corresponding author.
E-mail address: maobo17@mails.jlu.edu.cn (B. Mao).

https://doi.org/10.1016/j.jappgeo.2019.04.002
0926-9851/

human intervention. Then, based on the deep learning model after training, the data assimilation process is driven to perform a high-precision inversion of the subsurface velocity. Finally, this process is tested on the designed high-velocity salt body model to prove, that deep learning-based data assimilation has significant advantages over the traditional full waveform inversion in high-precision velocity inversion. The prior velocity model obtained by deep learning completes the velocity inversion under the high velocity body and successfully jumps out of the local minima caused by its inaccuracy.

2. Methods

2.1. Variational data assimilation theory and adjustment

The basic principle of data assimilation is to use background information and observation data with information of prior error to give a global optimal solution for both the background information and the observation data through an optimization method. In general, the data assimilation problem solution requires three pieces of information: 1) background information and observation data that already contains certain errors; 2) error information, including background information and observation data error. For different errors it is estimated that a completely different solution will be obtained; 3) covariance, the physical spatial correlation between various observations (Yue et al., 2010). The covariance information is not needed in this paper because we only have seismic gather information. For joint inversion of gravity, magnetism, electricity, and earthquakes, the covariance information will be indispensable.

The mainstream data assimilation algorithms include variational algorithms, Kalman filter series algorithms, and the emerging Bayesian-based particle filter algorithm and layered Bayesian model in recent years (Ma and Qin, 2012). The variational algorithm is closest to the full waveform inversion in mathematical principle. The theory and application of this algorithm with respect to the changes brought about by seismic exploration are as follows.

Traditionally, the objective function is given by (Nichols, 2010)

$$J=\frac{1}{2}\left(x_0-x_0^b\right)^T B_0^{-1}\left(x_0-x_0^b\right)+\frac{1}{2}\sum_{k=0}^{N}\left(H_k(x_k)-y_k\right)^T R_k^{-1}\left(H_k(x_k)-y_k\right) \quad (1)$$

where x_k, y_k, H_k, and R_k ($k = 0, \ldots, N$) are n-dimensional model state vectors, observation vectors, simulation operators, and observation error matrices at time t_k, respectively. B_0 is the error matrix of the prior background state vector x_0^b. Applying it to seismic exploration, x means velocity v and y means gather u, consider each iteration as a prediction, the above formula is rewritten as follows

$$J=\frac{1}{2}\sum_{k=0}^{N}\left(\left(v_k-v^b\right)^T B_k^{-1}\left(v_k-v^b\right)+\left(d_k-u\right)^T R_k^{-1}\left(d_k-u\right)\right) \quad (2)$$

where

$$d_k=H(v_k)+\epsilon_k^0, k=0,\ldots,N \quad (3)$$

v_k represents the velocity vector at time t_k, the initial velocity v_0 is assumed to be known at time t_0, v^b is a prior velocity estimate, usually provided by previous predictions, but this is difficult to obtain in seismic exploration, so it will be provided by deep learning which introduced below. d_k is the simulated observation of v_k, u is the observation gather. H is forward simulation and it is a nonlinear operator. ϵ_k^0 represents observation error consisting of instrument and representative error. The weight matrices B_k and R_k are positive definite, as the uncertainty of prior estimates and observations, respectively.

Eq. (2) can be abbreviated as follows

$$J(v)=\frac{1}{2}\|f(v)\|_2^2 \quad (4)$$

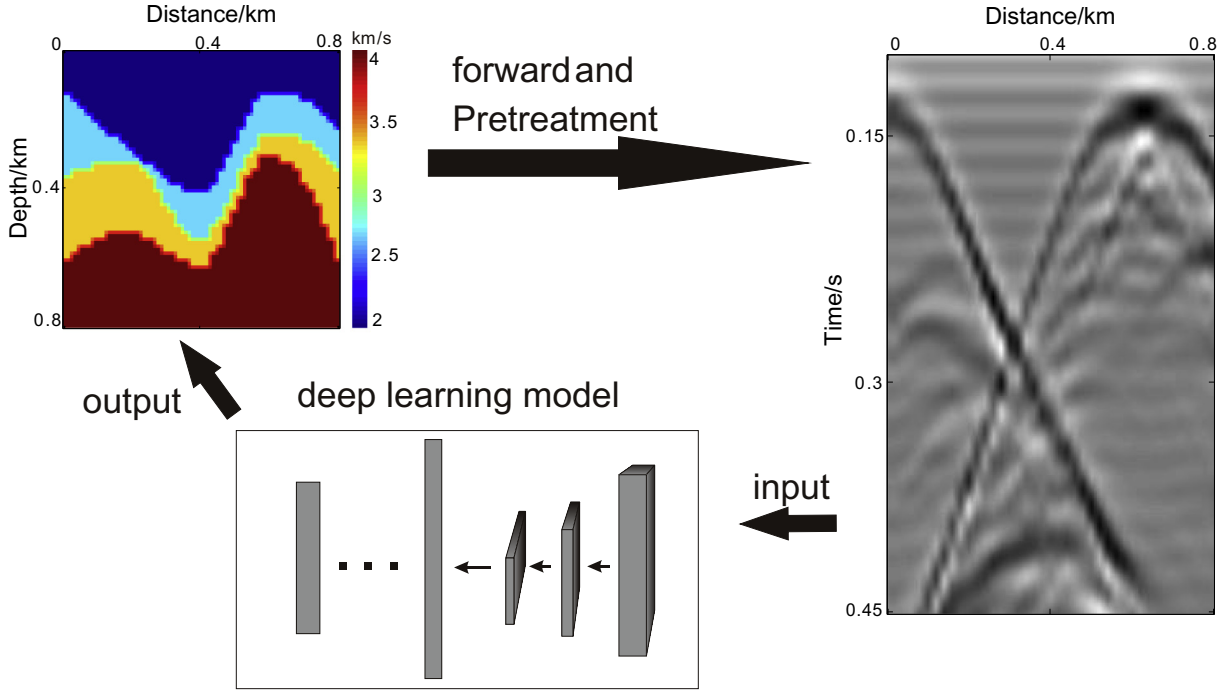

Fig. 1. Input, output and structure of the convolutional neural network (CNN).

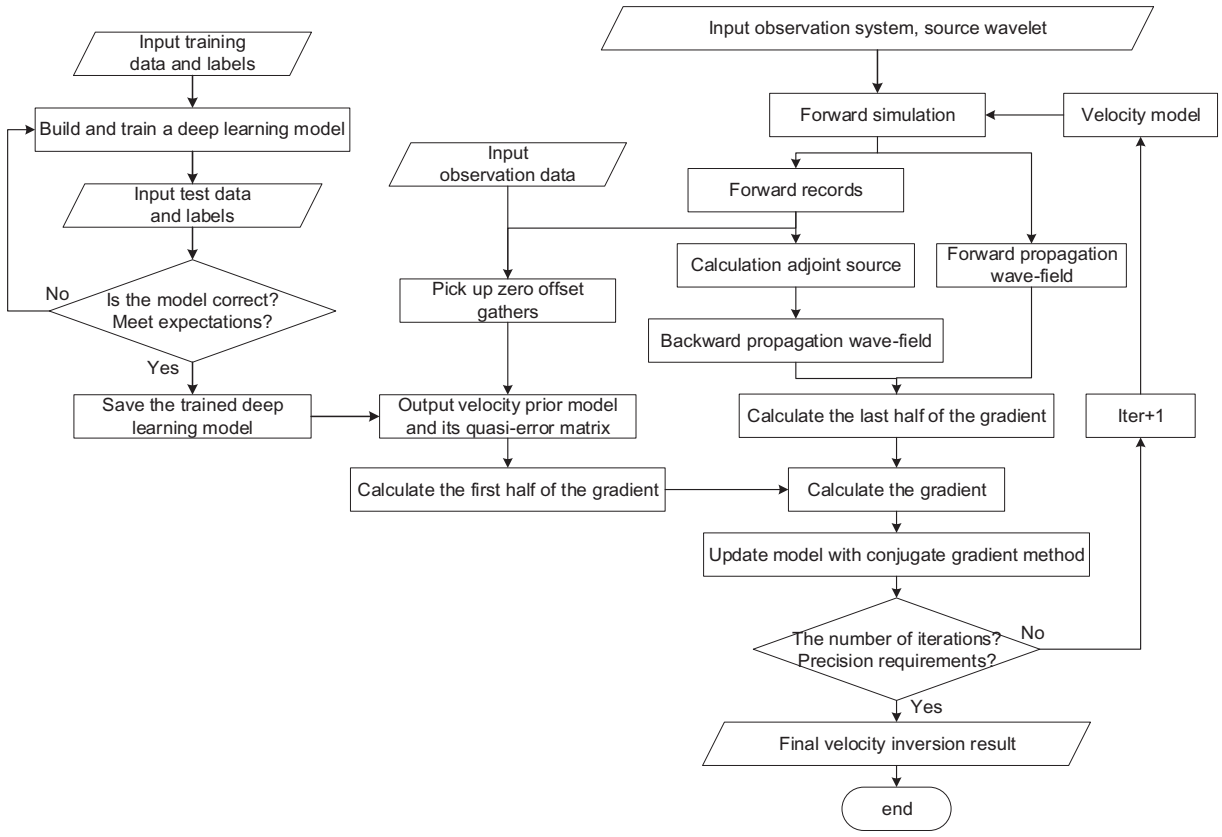

Fig. 2. Subsurface velocity inversion flow chart for deep learning driven data assimilation.

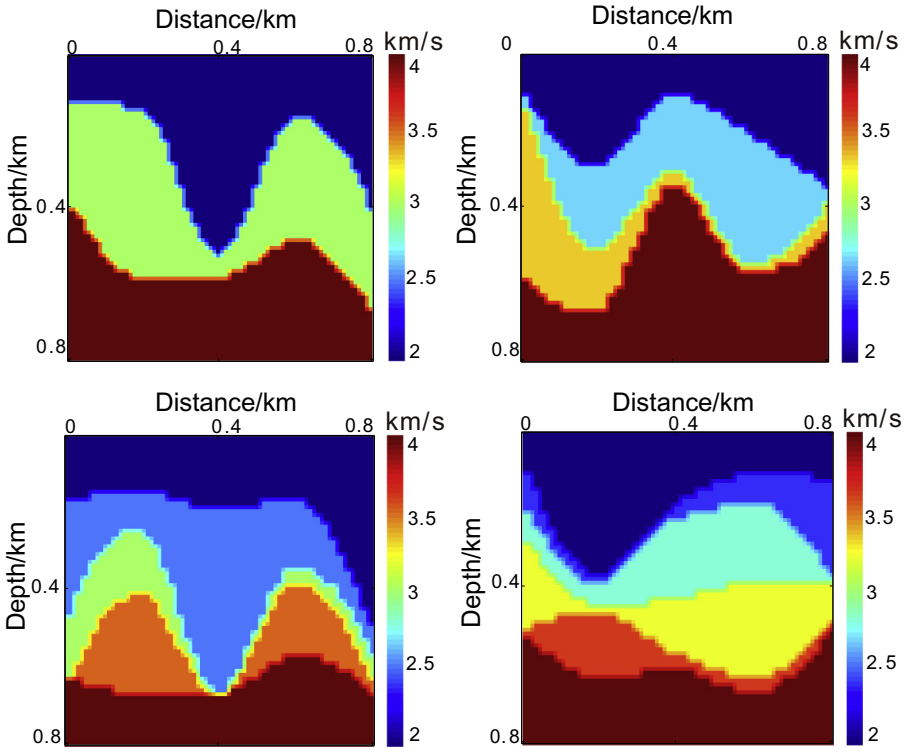

Fig. 3. Training data for random velocity layer structure from simple to complex.

where

$$f(v) = \begin{pmatrix} B_0^{-1/2}\left(v_0 - v^b\right) \\ \vdots \\ B_N^{-1/2}\left(v_N - v^b\right) \\ R_0^{-1/2}(d_0 - u) \\ \vdots \\ R_N^{-1/2}(d_N - u) \end{pmatrix} \tag{5}$$

The minimum point of Eq. (4) satisfies the gradient equation and is

$$\nabla_v J = Jc^T f(v) = 0 \tag{6}$$

where Jc represents the Jacobian matrix of the vector function f(v) defined by Eq. (5). Combining with the adjoint method (Tromp et al., 2010), we obtain

$$Jc = \begin{pmatrix} B_0^{-1/2} \\ \vdots \\ B_N^{-1/2} \\ R_0^{-1/2} S^{-1}\left[\frac{-\partial S}{\partial v} d_0\right] \\ \vdots \\ R_N^{-1/2} S^{-1}\left[\frac{-\partial S}{\partial v} d_N\right] \end{pmatrix} \tag{7}$$

where S is a large sparse matrix called the impedance matrix. The meanings of the weight matrices B_k and R_k are different in seismic exploration. R_k represents observation uncertainty; however, the observation data after denoising should be considered completely reliable (in the seismic exploration process, this is the only actual data). Therefore, we choose to place the energy compensation in the weight matrix R_k to improve the deep velocity disturbance during each iteration. For B_k, the inaccuracy of prior velocity estimation is not available without our understanding of the subsurface real velocity structure. Observing the velocity term on both sides of B_k in Eq. (2), we see that the effects of v_k and v^b are the same in each iteration. Therefore, we forward the velocity model v_k to obtain a zero-offset gather, and we input the gather into the trained deep learning model to obtain the output $\tilde{v_k}$. The $\widetilde{B_k}$ in each iteration is obtained by the difference between v_k and $\tilde{v_k}$, and we use $\widetilde{B_k}$ to approximate B_k.

2.2. Deep learning structure construction

To obtain reliable prior information with error matrices for driving data assimilation problems, deep learning is a novel choice. Convolutional neural network is one of the deep learning method. It can directly take the image as input and automatically extract features, avoiding the interference of human processing (Lecun et al., 2015), and its weight sharing attribute significantly reduces the complexity of the model (Lecun and Bengio, 1998; Dong et al., 2016).

After selecting the convolutional neural network as framework, we must first define the input and output of the deep learning model. As shown in Fig. 1, the input we selected is a zero-offset seismic gather that removes the direct wave, and the output is the subsurface velocity structure that we want. This neural network should have the following scheme,

$$V = \mathrm{Cnn}(U) \tag{8}$$

where V is the subsurface prior velocity information that we want to get, U is the observation data we processed (the zero-offset seismic gather of the direct wave is removed), and Cnn represents the neural network system we obtained through deep learning. It is just a symbol and cannot be described by simple mathematical expressions.

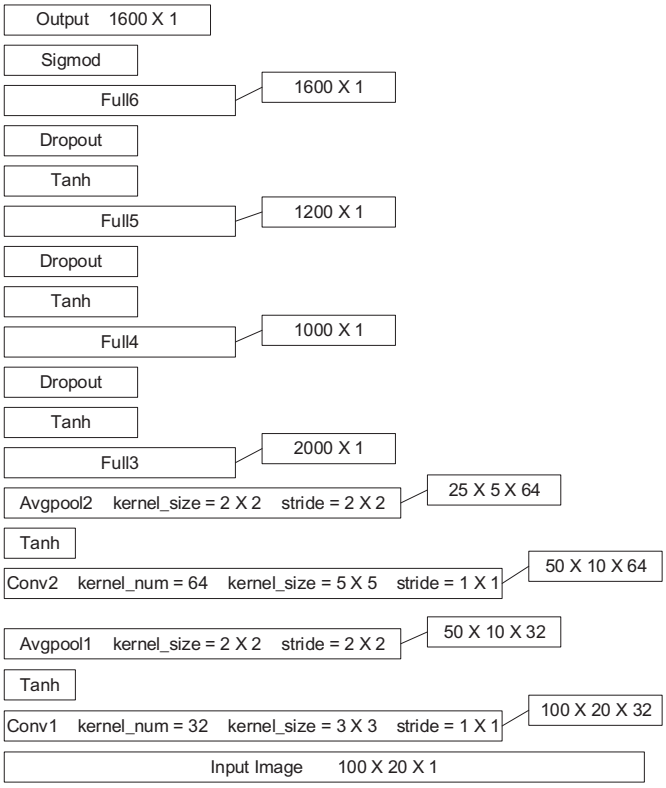

Fig. 4. Architecture of deep learning model.

After building the convolutional neural network, we input the training data prepared in advance to start training the deep learning model. The whole training process can be simply regarded as a least squares optimization problem with the following objective function

$$\mathrm{Loss} = \frac{1}{2}\sum_{i=1}^{N}\left(V - \tilde{V}\right)^2 \tag{9}$$

where $\tilde{V}$ is the label of input data, the true velocity of the subsurface structure. In the training process, data is divided into dozens or hundreds of batches, and then a batch of data is input for overall training, instead of inputting data one by one. This not only accelerates the efficiency but also reduces the overfitting phenomenon of the deep learning model to the training data. The N in the formula represents the amount of training data for each batch. To train the convolutional neural network Cnn to minimize the value of the objective function Loss, a Newton-like gradient algorithm can be selected for optimization.

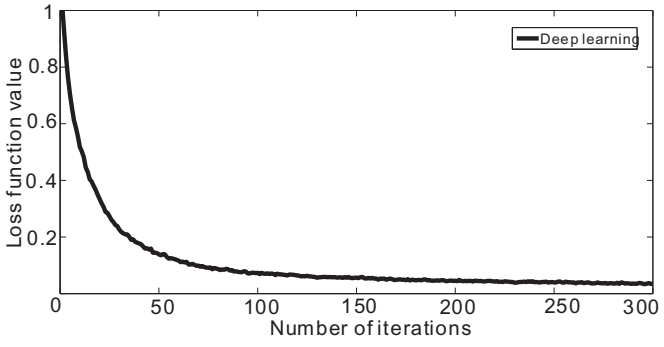

Fig. 5. Decline of cost function curve.

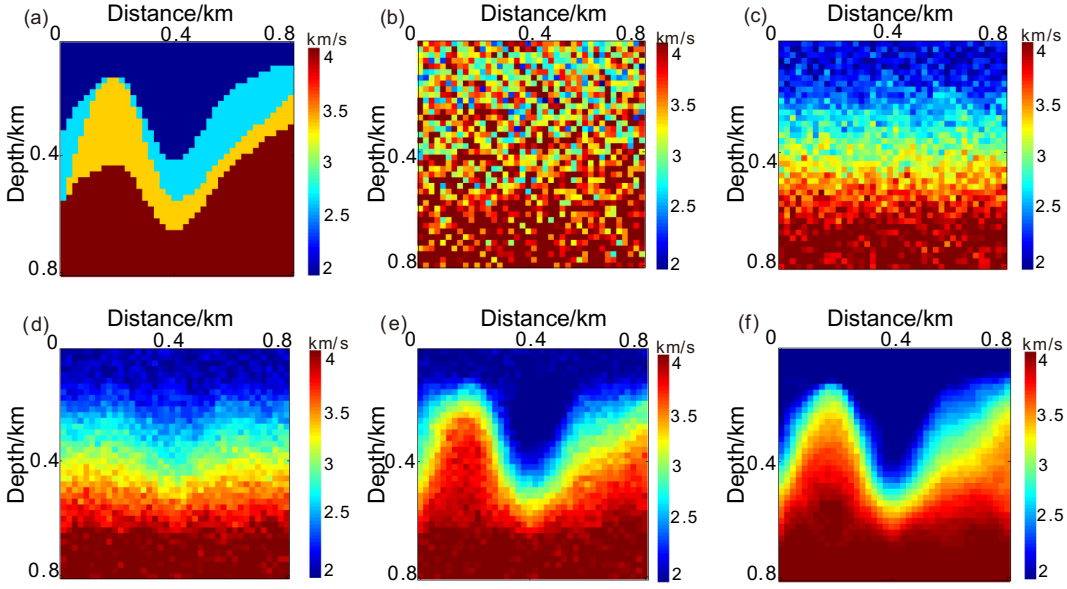

Fig. 6. Output of the deep learning model during training. (a) Text label; (b) After training 2 times; (c) After training 10 times; (d) After training 50 times; (e) After training 100 times; (f) After training 200 times.

In convolutional neural networks, the selection of convolution kernels affects the entire neural network, and the features extracted by them are the actual input data.

$$X_i = U * j_i \tag{10}$$

where j_i represents a total of i convolution kernels. It can be seen that the input data U will become complicated and huge after extracting features through i convolution kernels, so a subsequent pooling layer (down-sampling the input) is required to compress the data to ensure sufficient memory for the computer. The convolution part of the convolutional neural network can be regarded as the input preprocessing of the neural network part. The same convolution kernel sequence will extract the input data into completely different feature data, and the input data with less difference will become easier to distinguish.

As shown in Fig. 2, the left side is the deep learning module, which finally outputs our desired deep learning model and uses the trained model to obtain the prior velocity model and its quasi-error matrix in each iteration. The right side is the data assimilation module. Comparing with traditional full-waveform inversion, the prior velocity model and its quasi-error matrix are considered in the calculation of the gradient (its left module). In this way, deep learning drives data assimilation to achieve inversion results of required accuracy.

3. Numerical examples

3.1. Deep learning model training

After building the deep learning model structure, we need enough data to train the model, and we need a batch of data to determine whether the trained model meets our expectations. First, we generate 5000 velocity models with 3 to 6 layers of random velocity layers. These velocity models are the same size, and the velocity range is 2.0–4.0 km/s, as shown in Fig. 3. Then, through seismic forward simulation

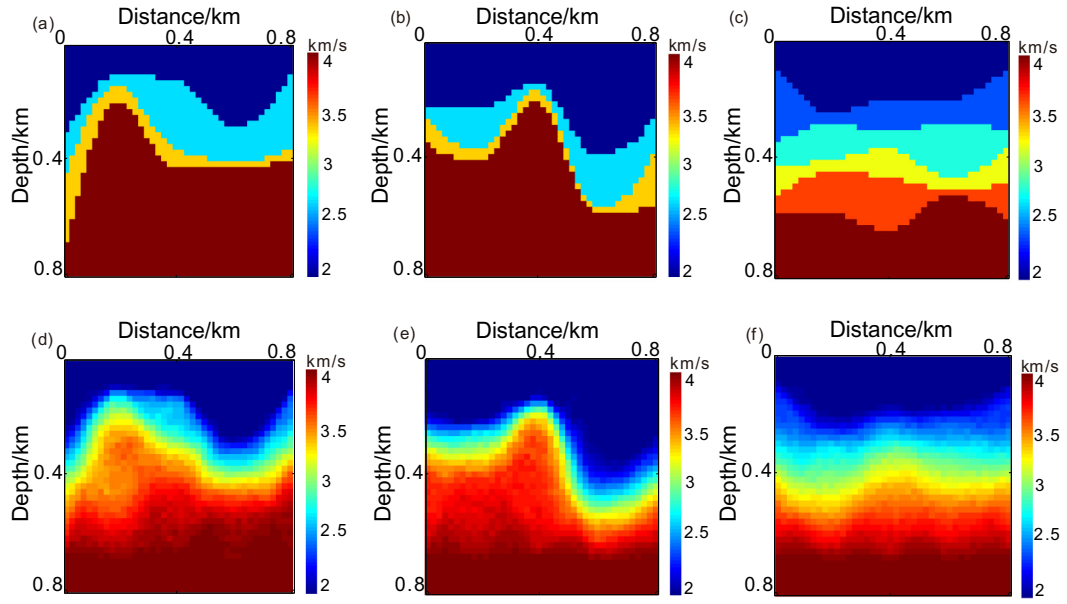

Fig. 7. Test of deep learning model after training. Test label from simple to complex (a), (b), (c)and their corresponding model output(d), (e), (f).

Table 1
Hardware specification and time cost (training and loading).

CPU: Xeon E5–2620 v3		Graphics card: RTX 2080 Ti	
Training			
Environmental preparation (s)	Iter once (s) and number of Iter	Save model (s)	Total time cost (s)
11.81	1.28 * 200	3.13	270.94
Loading			
Environmental preparation (s)	Prediction once (s)	Total time cost (s)	
8.24	0.02	8.26	

and preprocessing, the corresponding zero-offset gather for 5000 velocity models are obtained. In all the experiments, in order to more realistically simulate the real situation, the data of the forward simulation artificially cut off the effective information below the frequency of 10 Hz. We used the same process to get 50 sets of test data again and to ensure that the test data are different from the training data.

We enter the data and its corresponding labels and start training our deep learning model. Figs. 4 and 5 are schematic diagrams of the deep learning model structure and the drop of cost function (Eq. (9)) of the learning process, respectively. The convolutional neural network consists of 2 convolutional layers and 4 fully connected layers. In particular, the activation function uses "Tanh" because it fits the gather. And it can be seen from Fig. 5 that the objective function continues to decrease steadily. To display the whole training process more intuitively, we track and output a set of test data during training, and display the difference between the output of the test data and the corresponding label at any time, as shown in Fig. 6. Fig. 6a is the label of the test data, that is, the real velocity structure, and Fig. 6b, c, d, e, and f are the outputs when the deep learning model is trained for 2, 10, 50, 100, and 200 times, respectively. It can be seen that the trained deep learning model output is more and more similar to the label, and this can be used as the prior velocity information we are eager for, which provides a basis for a smooth solution to the data assimilation problem.

We save the deep learning model after 200 training iterations, input 50 sets of test data prepared in advance, calculate the cost function of the output and the corresponding label, and test the generality of the deep learning model, that is, whether there is an overfitting phenomenon. To avoid overfitting, dropout technology was used (Srivastava et al., 2014). This technology uses the artificial probability p ($0 < p < 1$). Every neuron in the network has a probability p that is suppressed (output is 0), and at the same time, the output of each suppressed neuron will be amplified to 1/p.

$$\mathrm{op} = \begin{cases} 0, & \mathrm{p} \\ \mathrm{ip} * \dfrac{1}{\mathrm{p}}, & 1-\mathrm{p} \end{cases} \tag{11}$$

where ip and op are the neuron input and output before and after the dropout unit, respectively. Fig. 7 is a comparison of the 3 label sets and their corresponding model outputs, which are a random selection of 50 test data sets according to label complexity from low to high. It can be seen that the output of the deep learning model has higher fit to simple than complex labels, but fortunately, the overall trend is correct. Although it cannot fit the details more accurately, we consider the output of the deep learning model as a prior velocity information and use it to drive the data assimilation problem. Table 1 showing the hardware specification and the computational cost of the method.

3.2. Data assimilation to solve for a simple salt body structure

The area containing salt domes is particularly exciting because it usually has hydrocarbon reservoirs on its sides or below (Kadu et al., 2017). However, the effect of full waveform inversion is not satisfactory in this respect (Lewis et al., 2012). The more popular improvement method is the level set method, but it also needs the support of prior information (such as the velocity of the salt dome). Therefore, it is great significance to verify the new proposed method for the salt body structure.

Firstly, we build a simple salt model, as shown in Fig. 8. The velocity model is changed on the test data of Fig. 6, with a high velocity block embedded in the shallow part. To make the results more intuitive, we use the L2 norm ($\mathrm{D} = \frac{1}{2}(v_p - v_r)^2$) to quantify the distance between the various results v_p and the real velocity model v_r. Fig. 8a shows the constructed real velocity model (label), and Fig. 8b shows the output (prediction) of the trained deep learning model. It can be seen that 8b is a good example of the general trend in 8a. The position of the salt body is reflected, and the deeper part of the cover also has fuzzy information, which is an effective prior velocity information. And the L2 norm D = 130.3790.

Then, we perform full waveform inversion on model 8a without prior velocity information and with deep learning model's output prior velocity information (Fig. 8b), respectively. The purpose is to illustrate the validity of the prior velocity information and to show the defect of traditional full waveform inversion. Finally, we use this prior velocity information to drive the data assimilation process and compare the obtained final output with the full waveform inversion result to prove the superiority of the method.

Fig. 9a is the initial model (its L2 norm D = 213.4281) used for full waveform inversion without prior velocity information, and Fig. 9b is its final output. We can see that the traditional full waveform inversion can only invert the upper part of the salt dome and the shallow subsurface structure without reliable prior information and misses the effective low-frequency information of the seismic gather. The inversion is caught in a local minimum that is far from the global minimum. And its L2 norm D = 162.6941. Fig. 9c is the final output of the full waveform inversion using prior velocity information 8b as the initial velocity

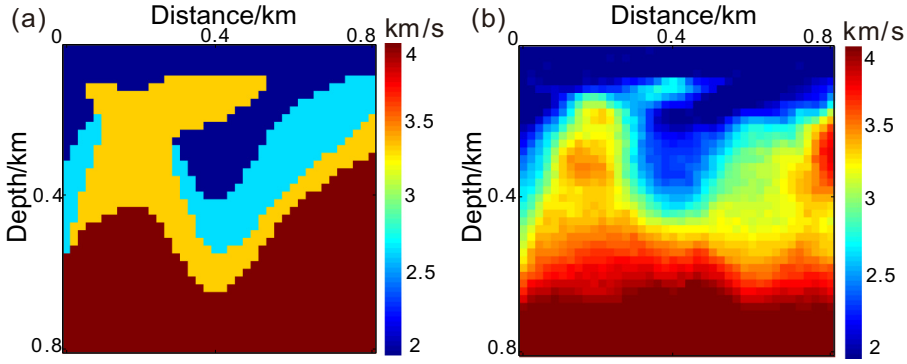

Fig. 8. Preparation of simple salt body test. (a) Real salt body model; (b) Output of the deep learning model.

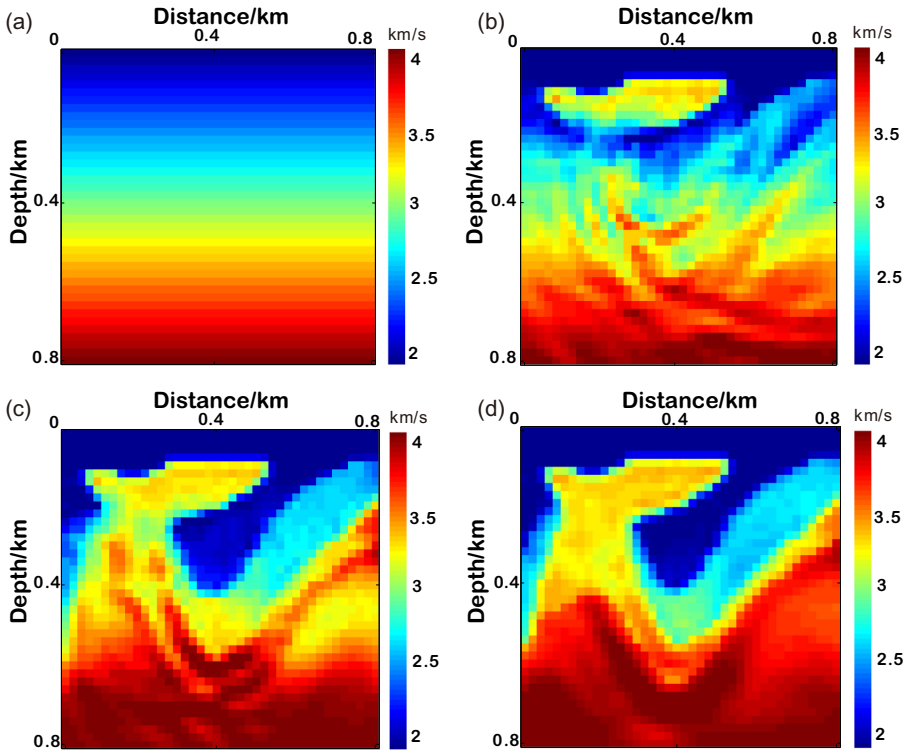

Fig. 9. Results of simple salt body test. (a) Linear initial velocity model; (b) Full waveform inversion results without prior information; (c) Full waveform inversion results with prior information; (d) Deep learning using priori information drives data assimilation results.

model. Comparing with the full waveform inversion result 9b without reliable prior velocity information, the result of the full waveform inversion has been greatly improved, and it is close to the real velocity model Fig. 8a. However, directly using the prior velocity information indicates that it is fully trusted, and in fact, the prior velocity information is not completely correct, with large errors in the deep and right sides of the model. It can be seen that the inversion results of these two places in Fig. 9c are inaccurate, indicating that the inversion is caught in a local minimum close to the global minimum. And its L2 norm D = 67.2214.

Fig. 9d is the final output of the deep learning driven data assimilation process using the linear gradient initial velocity model Fig. 9a as the initial combined with prior velocity information 8b. In the data assimilation process, we use the deep learning model to continually correct the prior velocity information error estimate, and quantify the degree of prior velocity information distrust in each iteration. It can be seen from Fig. 9d that not only has the salt body been fully embodied but also the large error in the prior velocity estimation is improved in the data assimilation process. And the L2 norm D = 21.4784. This proves the success of the proposed deep learning driven data assimilation process in the application of subsurface velocity inversion.

4. Conclusion

In this paper, a convolutional neural network is developed to predict the "prior velocity information" (the subsurface velocity information), then it is used to improve the full waveform inversion process with some error weighted on the generated prior velocity information. To solve the problem of full waveform inversion being unable to accurately converge to a global minimum in the absence of prior velocity information, we use deep learning method to train a reliable deep learning model. This model can directly take the processed zero-offset gather as input and the velocity of the corresponding region as output. Considering the accuracy of prior velocity information, data assimilation theory is applied to seismic subsurface velocity inversion. In the experiment of the simple salt body model, good results are obtained.

The numerical examples show that the following. (1) The deep learning model proposed and trained in this paper is successful, and the prior velocity information is basically correct and has practical value (the L2 norm D of initial model dropped from 213.4281 to 130.3790). (2) Using prior velocity information from the deep learning model improves the inversion quality and can solve the problem of salt body inversion (the L2 norm D of the full waveform inversion results dropped from 162.6941 to 67.2214). (3) The proposed deep learning driven data assimilation method effectively avoids the local convergence caused by inaccurate prior information (the L2 norm D dropped from 67.2214 to 21.4784).

Future work will mainly take three directions. The first is to extend the method to 3D data, the main difficulty is not formulation, but the memory demand. The second is Improve the accuracy of the deep learning model's output. We hope that can directly output high-precision predictions. Finally, how to apply this method to field data. Maybe this experiment is successful only because the same forward simulation is used.

Acknowledgements

Thanks to the sponsor of National Natural Science Foundation of China (No. 41674124).

References

Adler, A., Boublil, D., Zibulevsky, M., 2017. Block-based compressed sensing of images via deep learning. IEEE International Workshop on Multimedia Signal Processing , pp. 1–6. https://ieeexplore.ieee.org/document/8122281.

AlRegib, G., Deriche, M., Long, Z., et al., 2018. Subsurface structure analysis using computational interpretation and learning: a visual signal processing perspective. IEEE Signal Process. Mag. 35 (2), 82–98. https://ieeexplore.ieee.org/document/8312469.

Anthes, R.A., 1974. Data assimilation and initialization of hurricane prediction models. J. Atmos. Sci. 31 (3), 702–719. https://journals.ametsoc.org/doi/pdf/10.1175/1520-0469(1974)031%3C0702%3ADAAIOH%3E2.0.CO%3B2.

Araya-Polo, M., Jennings, J., Adler, A., et al., 2017. Deep-learning tomography. Lead. Edge 37 (1), 58–66. https://www.researchgate.net/publication/322192763_Deep-learning_tomography.

Asnaashari, A., Brossier, R., Garambois, S., et al., 2013. Regularized seismic full waveform inversion with prior model information. Geophysics 78 (2), R25–R36. https://library.seg.org/doi/abs/10.1190/geo2012-0104.1.

Barker, D.M., Huang, W., Guo, Y.R., et al., 1984. A three-dimensional variational data assimilation system for MM5: implementation and initial results. Wea. Rev, 897–914. https://journals.ametsoc.org/doi/pdf/10.1175/1520-0493(2004)132%3C0897%3AATVDAS%3E2.0.CO%3B2.

Bergthórsson, P., Bo, R.D., 1955. Numerical weather map analysis. Tellus 7 (3), 329–340. https://www.mendeley.com/research-papers/numerical-weather-map-analysis/.

Carton, J.A., 2000. A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part I : methodology. J. Phys. Oceanogr. 30 (2), 294–309. http://mitgcm.org/~gforget/pdfs/soda1.pdf.

Charney, J.G., Fjörtoft, R., Neumann, J.V., 2010. Numerical integration of the barotropic vorticity equation. Tellus 2 (4), 237–254. https://onlinelibrary.wiley.com/doi/full/10.1111/j.2153-3490.1950.tb00336.x.

Chen, P., 2011. Full-wave seismic data assimilation: theoretical background and recent advances. Pure Appl. Geophys. 168 (10), 1527–1552. https://link.springer.com/article/10.1007/s00024-010-0240-8.

Courtier, P., Talagrand, O., 2010. Variational assimilation of meteorological observations with the adjoint vorticity equation. Ii: numerical results. Q.J.R.Meteorol.Soc 113 (478), 1329–1347. https://www.tandfonline.com/doi/abs/10.3402/tellusa.v42i5.11896.

Courtier, Andersson P., Heckley, E., et al., 1998. The ECMWF Implementation of Three Dimensional Variational Assimilation. Part 1: Formulation. Ecmwf. http://www.atmosp.physics.utoronto.ca/~dbj/PHY2506/rabier_etal2000.pdf.

Cressman, G.P., 1959. An operational objective analysis system. Mon. Weather Rev. 87 (10), 367–374. https://core.ac.uk/display/23640123.

Daley, R., 1991. Atmospheric Data Analysis. Cambirdge Atmopsheric and Space Science Series 2. , p. 457. http://www.cambridge.org/catalogue/catalogue.asp?isbn=9780521458252.

Datta, D., 2016. Estimating starting models for full waveform inversion using a global optimization method. Eage Conference and Exhibition. Geophysics 81 (4), R211–R223. https://library.seg.org/doi/abs/10.1190/geo2015-0339.1.

Derber, J., Rosati, A., 1989. A Global Oceanic Data Assimilation System. J. Phys. Oceanogr. 19 (19), 1333–1347. https://www.researchgate.net/publication/243690312_A_Global_Oceanic_Data_Assimilation_System.

Di, H., Wang, Z., AlRegib, G., 2018. Real-time seismic-image interpretation via deconvolutional neural network. SEG Technical Program Expanded Abstracts 2018 , pp. 2051–2055. https://library.seg.org/doi/abs/10.1190/segam2018-2997303.1.

Dimet, F.L., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus Ser. A-Dyn. Meteorol. Oceanogr. 38A (2), 97–110. https://www.tandfonline.com/toc/zela20/current.

Dong, Y., Gu, Y., Dean, S., et al., 2006. Sequential assimilation of 4D seismic data for reservoir description using the ensemble Kalman filter. J. Pet. Sci. Eng. 53, 83–99. https://www.sciencedirect.com/science/article/pii/S0920410506000805.

Dong, C., Chen, C.L., He, K., et al., 2016. Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell. 38 (2), 295–307. https://ieeexplore.ieee.org/document/7115171.

Guillen, P., Larrazabal, G., Gonzalez, G., et al., 1949. Supervised learning to detect salt body. Seg's International Exposition and Meeting in New Orleans, Louisiana https://www.researchgate.net/publication/286937526_Supervised_learning_to_detect_salt_body.

Kadu, A., Leeuwen, T.V., Mulder, W.A., 2017. Salt Reconstruction in full-waveform inversion with a parametric level-set method. IEEE Trans. Comput. Imaging 3 (2), 305–315. https://ieeexplore.ieee.org/document/7784771.

Kalnay, E., Kanamitsu, M., Kistler, R., et al., 1996. The NCEP/NCAR 40-year reanalysis project. Bull. Am. Meteorol. Soc. 77 (3), 437–472. https://www.researchgate.net/publication/203931920_The_NMCNCAR_40-year_reanalysis_project_Bull_Am_Meteorol_Soc.

Lecun, Y., Bengio, Y., 1998. Convolutional Networks for Images, Speech, and Time Series. MIT Press http://www-labs.iro.umontreal.ca/~lisa/pointeurs/handbook-convo.pdf.

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436. https://www.nature.com/articles/nature14539.

Lewis, W., Vigh, D., 2017. Deep learning prior models from seismic images for full-waveform inversion. SEG Technical Program Expanded Abstracts 2017 , pp. 1512–1517. https://library.seg.org/doi/pdf/10.1190/segam2017-17627643.1.

Lewis, W., Starr, B., Vigh, D., 2012. A level set approach to salt geometry inversion in full-waveform inversion. Seg Technical Program Expanded: 4609 https://library.seg.org/doi/abs/10.1190/segam2012-0737.1.

Li, H., Yang, W., Yong, X., 2018. Deep learning for ground-roll noise attenuation. SEG Technical Program Expanded Abstracts 2018 , pp. 1981–1985. https://library.seg.org/doi/abs/10.1190/segam2018-2981295.1.

Liu, Q., Tape, C., Tromp, J., 2009. 3D seismic tomographic inversion as a data assimilation problem. American Geophysical Union, Spring Meeting 2009 http://adsabs.harvard.edu/abs/2009AGUSMDI74A..01L.

Lorenc, A.C., 2010. Analysis methods for numerical weather prediction. Q. J. R. Meteorol. Soc. 112 (474), 1177–1194. https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711247414.

Ma, J., Qin, S., 2012. Recent advances and development of data assimilation algorithms. Adv. Earth Science 27 (7), 747–757. http://www.adearth.ac.cn/EN/Y2012/V27/I7/747.

Nichols, N.K., 2010. Mathematical Concepts of Data Assimilation. Springer Berlin Heidelberg, pp. 13–39. http://centaur.reading.ac.uk/2001/.

Pratt, R.G., 1999. Seismic waveform inversion in frequency domain, Part 1 : Theory and verification in physical scale model. Geophysics 64 (3), 888. http://library.seg.org/doi/abs/10.1190/1.1444597.

Richardson, L.F., 2007. Weather prediction by numerical process. N. Y. vi. https://www.researchgate.net/publication/229061284_Weather_Prediction_by_Numerical_Process.

Richardson, A., 2018. Seismic Full-Waveform Inversion Using Deep Learning Tools and Techniques. http://cn.arxiv.org/abs/1801.07232.

Sasaki, Y.K., 2007. An objective analysis based on the variational method. J.Meteor.Soc. Japan 36 (3), 77–88. https://www.jstage.jst.go.jp/article/jmsj1923/36/3/36_3_77/_article.

Smedstad, O.M., O'Brien, J.J., 1991. Variational data assimilation and parameter estimation in an equatorial Pacific ocean model. Prog. Oceanogr. 26 (2), 179–241. https://doi.org/10.1016/0079-6611(91)90002-4.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15 (1), 1929–1958. http://jmlr.org/papers/volume15/srivastava14a.old/srivastava14a.pdf.

Stauffer, D.R., 1990. Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I : Experiments with synoptic-scale data. Mon. Weather Rev. 118, 1250–1277. https://journals.ametsoc.org/doi/pdf/10.1175/1520-0493(1991)119%3C0734:UOFDDA%3E2.0.CO.

Sun, H., Demanet, L., 2018. Low frequency extrapolation with deep learning. SEG Technical Program Expanded Abstracts , pp. 2011–2015. https://library.seg.org/doi/abs/10.1190/segam2018-2997928.1.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation. Geophysics 49 (8), 1259–1266. http://library.seg.org/doi/abs/10.1190/1.1441754.

Tromp, J., Tape, C., Liu, Q., 2010. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. Geophys. J. R. Astron. Soc. 160 (1), 195–216. https://academic.oup.com/gji/article/160/1/195/712020.

Wu, X., Shi, Y., Fomel, S., et al., 2018. Convolutional neural networks for fault interpretation in seismic images. SEG Technical Program Expanded Abstracts 2018 , pp. 1946–1950. https://library.seg.org/doi/abs/10.1190/segam2018-2995341.1.

Yue, X.A., Wan, W.X., Liu, L.B., et al., 2010. Development of an ionospheric numerical assimilation nowcast and forecast system based on Gauss-Markov Kalman Filter – an observation system simulation experiment taking example for China and its surrounding area. Chin. J. Geophys. 53 (4), 787–795. http://www.en.cnki.com.cn/Article_en/CJFDTOTAL-DQWX201004004.htm.