

Physics-guided neural network for velocity calibration using downhole microseismic data

Hongliang Zhang, Jubran Akram, and Kristopher A. Innanen

ABSTRACT

We present an unsupervised physics-guided neural network to calibrate the simplified 1D layered velocity model based on downhole recordings of perforation shots. This novel neural network incorporates five fully connected layers, a Scaling & Shifting layer as well as a forward modeling layer that generates theoretical travel times of P- and S-waves. Due to the inclusion of the forward modeling layer, our network eliminates the need for labeled data which is unavailable or limited in many cases. In addition, compared with conventional theory-based inversion, the neural network can solve the velocity optimization problem without explicit programming. To yield better constraint for both velocity-calibration and event-location problems, a hybrid objective function is used, which combines misfits of both arrival times and arrival-time difference between P- and S-waves. We apply the proposed neural network to a numerical example with six simulated perforation shots, yielding robust inversion results for layer velocities in the presence of noise. This neural network will be further examined with field data in the future research.

INTRODUCTION

A good knowledge of the velocity information is crucial to obtain accurate microseismic event locations. A detailed 3D velocity model can provide more constraints in event locations by accounting for spatial velocity variations due to geological structures, facies changes, strata dipping, etc. While constructing 3D velocity models normally requires 3D seismic data and/or tomography techniques (Kissling, 1988; Cameron et al., 2007), which are not available in many cases. Instead, 1D flat-layered velocity models are often used to locate microseismic events in practical microseismic data processing. The initial rough 1-D velocity is often built based on well log data. Then various calibrations events (e.g., perforation shots, string shots or ball-dropping events) can be used to update the initial velocity model by minimizing arrival time/arrival-time difference residuals for P- and/or S-waves. Many approaches to calibrate the 1-D velocity model have been suggested in previous studies; for example, Pei et al. (2008, 2009) optimized the velocity model using Occam's and very fast simulated annealing algorithms, in which the arrival-time residual was minimized. Tan et al. (2013) and Jiang et al. (2016) applied time difference-based velocity calibration algorithms using downhole and surface microseismic data, respectively. Tan et al. (2018) adopted the neighborhood algorithm, in which a hybrid objective function that combines misfits for both arrival times and arrival-time difference between P- and S-waves was introduced. All the above velocity-calibration algorithms are theory-based and need explicit programming for both forward modeling and inversion process.

Recently, machine learning has become unprecedentedly popular due to the improvements of the computational capacity of computers and rapid developments within the big data revolution. Various machine learning techniques have been applied to different microseismic data processing steps, such as event detection (Qu et al., 2019; Stork et al., 2020),

arrival-time picking (Chen et al., 2019; Ma et al., 2020) and event location (Huang et al., 2018; Zhang et al., 2020), exhibiting promising results. While most of these applications are data driven and may fail to yield robust results in cases of limited or unavailable training data.

In this study, we propose an unsupervised physics-guided neural network to calibrate the 1D layered velocity model based on P- and S- wave arrival-time measurements of multiple perforation shots. A forward modeling layer (i.e., ray-tracing layer) is incorporated into the fully connected (FC) neural work to guide the training process, which thus eliminates the need for training data. In addition, neural network weights and layer velocities are updated through the training process without explicitly programming for inversion algorithms.

METHOD

Figure 1 shows the architecture of the proposed physics-guided neural network, in which the input is the picked arrival times for both P- and S-waves of multiple perforation shots. Five FC layers and a Scaling & Shifting layer are used to generate layer velocities for P- and S-waves. Differing from the conventional supervised learning which is normally data driven and requires examples of input-label pairs to train the network, we implement the network in an unsupervised fashion, in which a forward modeling layer, i.e., the Ray Tracing layer, is included to model theoretical arrival times of P- and S-waves based on the layer velocities produced from the previous Scaling & Shifting layer. Then a loss function that evaluates the error between observed arrival times and modeled travel times is used to update the weights of neural networks. Our designed network follows a similar architecture to an encoder-decoder type of neural network, in which the combination of FC layers and the Scaling & Shifting layer acts as an encoder to map the input arrival times into layer velocities, then a decoder, the Ray Tracing layer, is used to model travel times. The implementation details of the proposed neural networks are presented in the followings.

In the physics-guided neural network, five fully connected layers are used. The mathematical operation for the FC layers can be written as

$$\mathbf{y}^i = f(\mathbf{w}^i \mathbf{y}^{i-1} + \mathbf{b}^i), (i = 1, 2, \dots, 5), \quad (1)$$

where \mathbf{y}^{i-1} and \mathbf{y}^i are the output vectors for layers $i-1$ and i , and \mathbf{w}^i and \mathbf{b}^i represent the weight matrix and bias vector for the i th layer. Since velocities for both P- and S-waves are considered in the inversion. The number of neurons for the first and fifth FC layers are equal to twice of the numbers of geophones and velocity layers, respectively. For the three intermediate FC layers, each layer has 32 neurons in this study. To introduce nonlinearity into the output of each FC layer, the activation function f is used. For the first four FC layers, the rectified linear unit (ReLU) is used, and the sigmoid activation is adopted for the last FC layer. Mathematical representations of ReLU and sigmoid activations are

$$f(x) = \max(0, x), \quad (2)$$

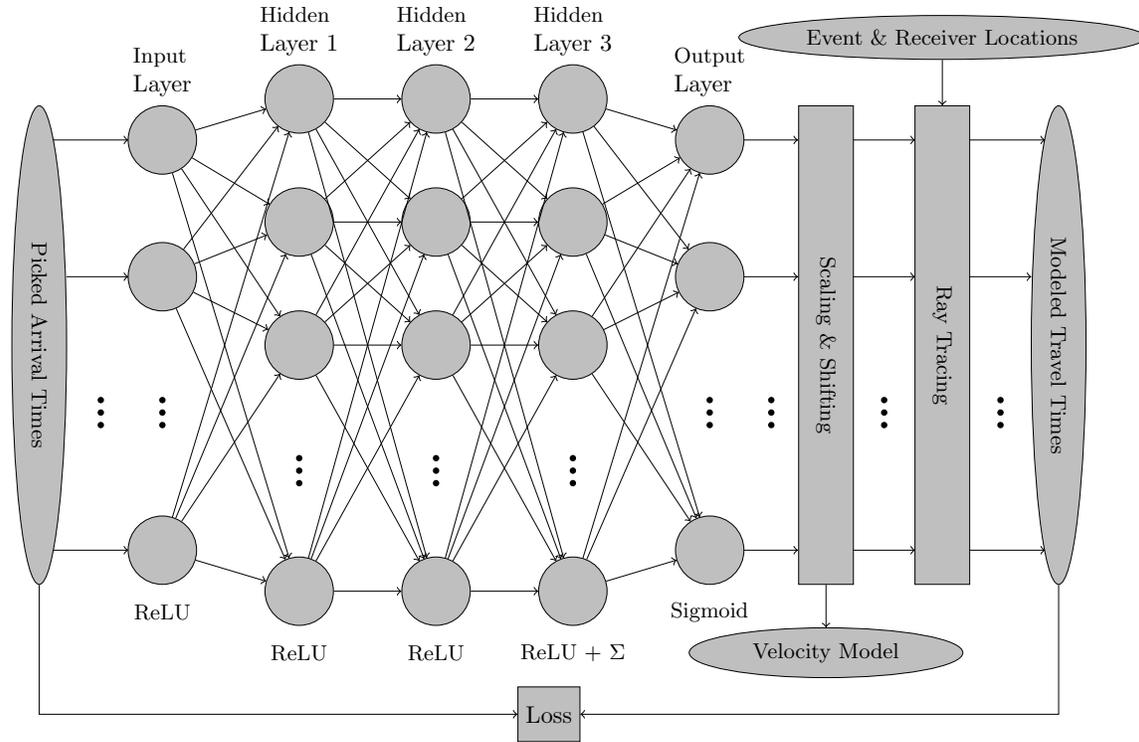


FIG. 1. Architecture of the physics-guided neural network. Activation functions and extra operations are indicated at the bottom of FC layers. Inputs and outputs of the network are represented by ellipses.

and

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (3)$$

For the regression problem, the numbers of input vectors and output vectors are normally the same. While for this velocity-calibration case, arrival times of multiple perforation shots are used to optimize a single layered velocity model, i.e., the numbers of input and output vectors are different. Thus, after the ReLU activation at the third hidden layer, a summation operation over all the vectors is implemented to output only a single vector.

Due to the sigmoid activation function, elements of output vector from the last FC layer have values within the range of 0 – 1. To make the generated layer velocities in desired ranges, a Scaling & Shifting layer is used. The output from this layer can be denoted as

$$\mathbf{v} = \mathbf{a}_0 + \mathbf{a}_1 \odot \mathbf{y}, \quad (4)$$

where \mathbf{y} is the output of the previous layer, and \mathbf{v} is the resulting velocity vector. The symbol \odot denotes the element-wise multiplication. Vectors \mathbf{a}_0 and \mathbf{a}_1 are the shifting and scaling vectors, respectively. The shifting vector \mathbf{a}_0 consists of the lower layer-velocity boundaries, and each element of the scaling vector \mathbf{a}_1 represents the velocity difference between upper and lower boundaries for the corresponding layer.

After the Scaling and Shifting layer, the Ray Tracing layer is built to calculate theoretical travel times of P- and S-waves. In addition to the velocity model generated from the previous layer, this forward modeling layer also requires extra inputs, i.e., locations of receivers and perforation shots that are prepared before the training phase. For down-hole microseismic data, the classic ray-shooting algorithms may be unstable since large velocity perturbations between two adjacent layers may lead to shadow zones related to the post-critical incidence (Eaton et al., 2011). Thus, this study uses an efficient ray-bending algorithm (e.g., Cerveny, 2005; Eaton et al., 2011, which can yield an approximate, but robust solutions under the downhole monitoring geometries. In the ray-bending algorithm adopted here, the initial ray path is guessed, then updated iteratively to optimize the ray parameter p . Compared with ray-shooting algorithms, the ray-bending method can significantly reduce the computation time, which further improves the efficiency of the neural network.

In the physics-guided neural network, we use a similar loss function developed by Tan et al. (2018), which is defined as

$$\phi = \sqrt{\alpha_1 \sum_{i=1}^M \sum_{j=1}^N (T_P^{ij} - t_P^{ij} - T_P^{i0})^2 + \alpha_2 \sum_{i=1}^M \sum_{j=1}^N (T_S^{ij} - t_S^{ij} - T_S^{i0})^2 + \alpha_3 \sum_{i=1}^M \sum_{j=1}^N [(T_P^{ij} - T_S^{ij}) - (t_P^{ij} - t_S^{ij})]^2}, \quad (5)$$

where T_P^{ij} and T_S^{ij} represent the observed arrival times for P- and S-waves at the j th receiver for the i th perforation shot, and t_P^{ij} and t_S^{ij} denote the theoretical travel times. M and N are the total numbers of perforation shots and receivers, respectively. According to the work by Nelson and Vidale (1990), the best fitting origin time T^{i0} can be estimated by the formula

$$T^{i0} = \frac{1}{N} \sum_{j=1}^N (T^{ij} - t^{ij}). \quad (6)$$

In Equation (5), three terms are included within the square root, i.e., residuals of arrival time difference for P-wave (the first term) and S-wave (the second term), and the residual for S-P arrival/travel time difference (the third term). Three weighting factors, α_1 , α_2 and α_3 , are used to balance the contributions from the three terms. As presented by Tan et al. (2018), the combination of the three terms in the loss function can yield better constraints than conventional ones not only for the velocity-calibration problem but also for the subsequent event-location problem.

For the physics-guided neural network presented here, no training data is needed, and the 1-D layered velocity model is updated during the training phase of model parameters (network weights and biases) based on the perforation-shot data. In the training process, the adaptive momentum (Adam) algorithm (Kingma and Ba, 2014) is used to search the optimal model parameters. This algorithm can calculate learning rates adaptively for individual model parameters from estimates of first and second moments of the gradients. In

addition, prior to the training, the weights and biases of the proposed network are assigned randomly with Xavier initialization (Glorot and Bengio, 2010).

SYNTHETIC EXAMPLE

Numerical simulation setup

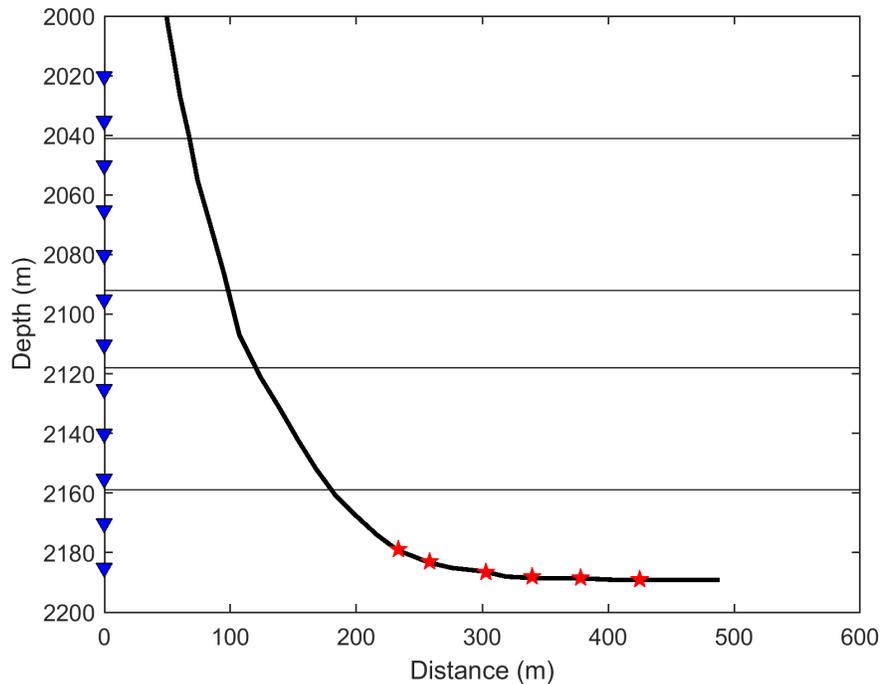


FIG. 2. Acquisition geometry of the synthetic study. Blue triangles and red stars denote geophones and perforation shots, respectively. The thick black line represents the well trajectory. Layer interfaces of the 1D velocity model are marked by horizontal lines.

A synthetic dataset is used to examine the performance of the proposed neural network. As shown in Figure 2, six simulated perforation shots along the horizontal well are used for calibrating the initial 1D layered velocity model derived from well log data. The simulated downhole monitoring system consists of 12 geophones with a constant depth interval of 15 m. Synthetic P- and S-wave travel times are calculated based on the velocity model in Figure 3. To simulate the picking error, zero-mean Gaussian noise with standard deviation of 0.5 ms is added to the P- and S-wave travel times at each individual geophone. Due to the generally high signal-to-noise ratio of calibration shot events recorded by the downhole array, selecting this level of noise is reasonable (Tan et al., 2013; Akram and Eaton, 2016). The velocity model in Figure 3 consists of five layers, and search areas for P-wave and S-wave in the network training are specified by the light blue and light red regions. This example uses a vertical downhole array and a layered isotropic velocity model, thus theoretical travel times for any event along a given circle on the horizontal plane are the same due to the symmetry property. In practical microseismic data processing, the back azimuth (orientation with respect to the downhole array) of the microseismic events is first determined through P-wave polarization, then a grid search can be implemented on the 2D space along the pre-determined vertical plane. For the sake of simplicity, this study

only investigates the velocity model calibration in 2D space.

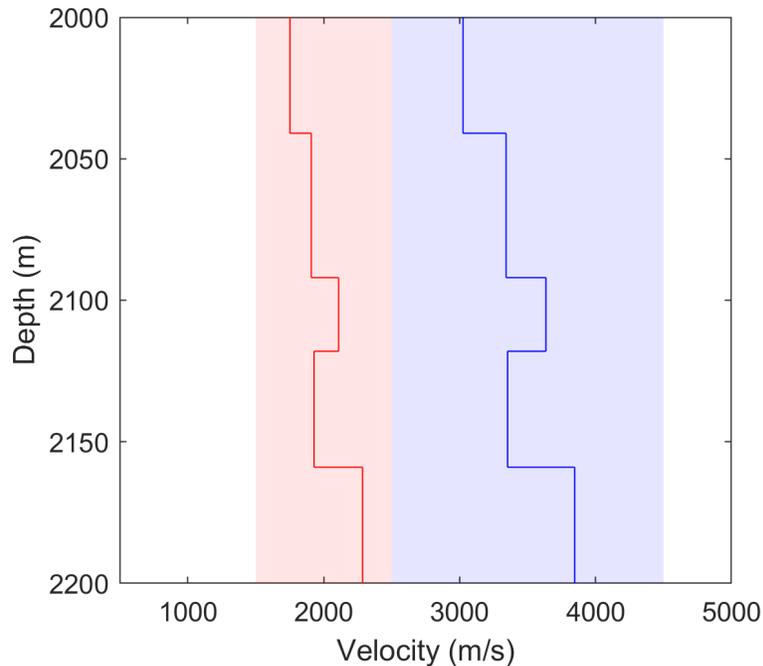


FIG. 3. P-wave (blue line) and S-wave (red line) velocity models used to generate synthetic data. The light blue and red areas denote search regions for P- and S-wave velocities, respectively.

Results

For the loss function in Equation (5), the values for α_1 , α_2 and α_3 are set to be 5, 0.5 and 0.001. With an initial value of 0.001, the learning rate is reduced by 25% after every 100 epochs, and 1,000 epochs are used in total in the training. Figure 4 shows the loss curve, in which the loss value greatly decreases at the beginning. In addition, significant fluctuation can also be observed within the first 200 training epochs which is mainly due to that the scheduled learning rate is too large for some certain epochs leading to the overshooting problem. Then, the loss value decreases gradually despite some minor fluctuations after the first 200 epochs, and it converges to a relatively low level after 800 epochs.

Figure 5 shows the P- and S-wave velocity models estimated at the final training epoch with each layer exhibiting minor deviations from the true velocity values, e.g., the average deviations for P- and S-waves are 17.3 *m/s* and 19.2 *m/s*, respectively. This minor deviation is mainly caused by the random noise within the synthetic data. With the retrieved velocity model, we also estimate the P- and S-wave arrival times (blue dots in Figures 6 and 7) at each geophone for all perforation shots, and the mean RMS errors between observed and estimated arrival times are 0.45 *ms* and 0.51 *ms* for P- and S-waves, respectively.

To evaluate the inverted velocity model through the proposed neural network, we use the grid search algorithm to relocate the six perforation shots, in which the same loss function as in Equation (5) is used. We use a grid size of 1 *m* for both depth and distance, and the grid point with the minimum loss value is then taken as the relocated position. Figure

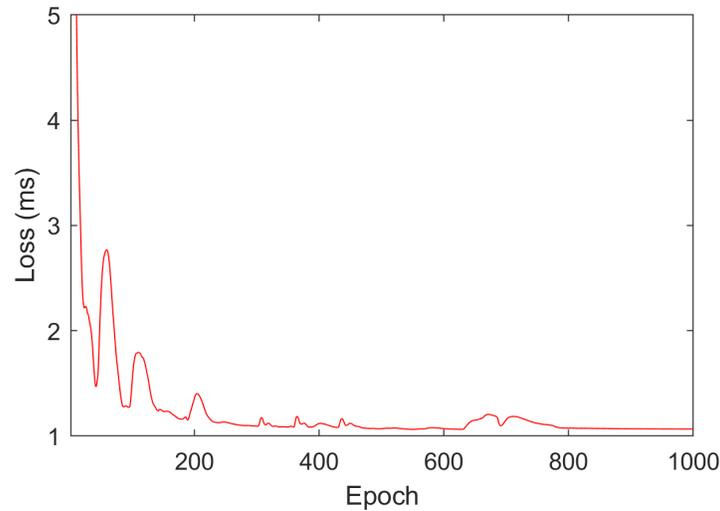


FIG. 4. Loss curve of the training process.

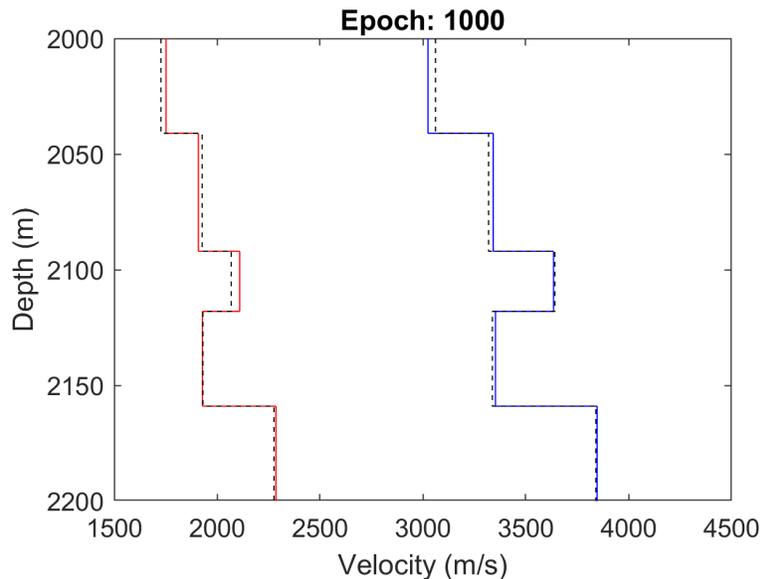


FIG. 5. Inversion result for the P- (blue dashed line) and S-wave (red dashed line) velocity models. Solid lines represent true velocity models.

8 shows contour maps of the loss value for the six perforation shots, and the relocated positions exhibit relatively small errors from the true positions with mean deviations of 2.2 *m* and 3.6 *m* for depth and distance, respectively. As comparisons, we also relocate these six perforation shots using the other two loss functions, i.e., RMS residuals in arrival times (loss functions in Equation (5) with only the first two terms within the square root) and in arrival-time difference (loss functions in Equation (5) with only the third term within the square root). The results are shown in Figures 9 and 10. It can be observed that the use of residuals of either arrival times or arrival-time difference yields much larger location errors than the results with hybrid loss function (i.e., the use of the three terms within the square root in Equation (5)). In Figure 9, the contour lines of arrival-time residuals greatly elongate along a certain direction where both the depth and distance increase or decrease

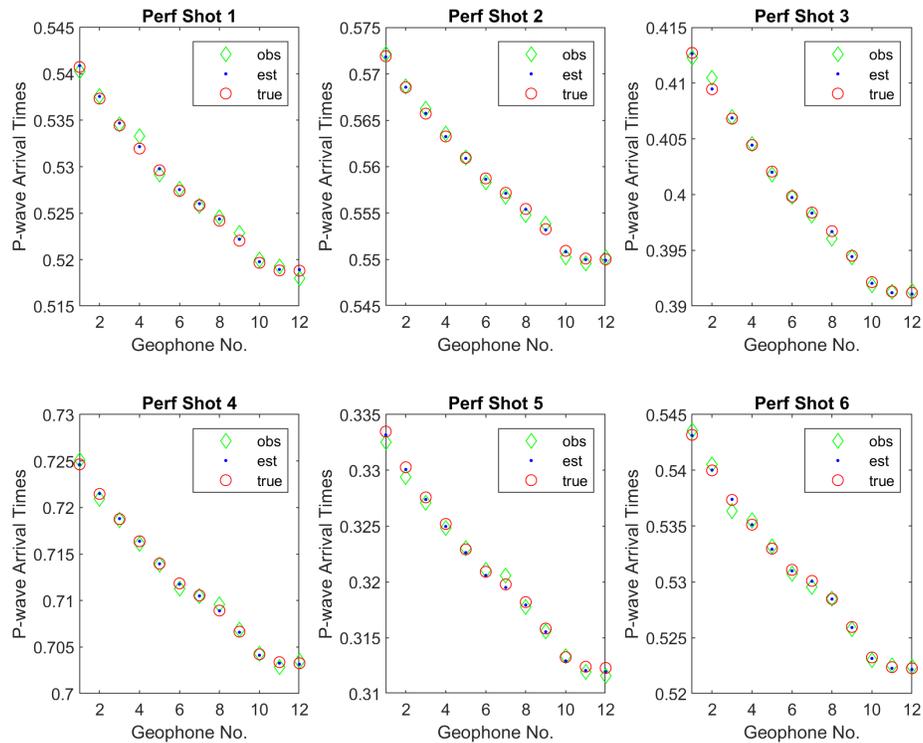


FIG. 6. Comparisons of observed (green diamond), estimated (blue dot) and true (red circles) P-wave arrival times for the six perforation shots.

concurrently. While, the use of the time-difference residual can yield better lateral positions but poor depth solutions (shown in Figure 10). In comparison, the use of the hybrid loss function could provide better constraints for both lateral and vertical locations than the other two loss functions, which is in good agreement with the study by Tan et al. (2018).

For the event-location problem, the effectiveness of the hybrid loss function has been demonstrated through the above analysis. While for the velocity-calibration problem, due to the multilayers of the velocity model, it is not feasible to draw such a contour map as Figure 8. Instead, similar to the study by Tan et al. (2018), we use a constant velocity model to illustrate the impacts of different loss functions on the velocity inversion results. For the constant velocity model, P- and S-wave velocity values are set to be 3000 m/s and 1800 m/s , respectively. The acquisition geometry is the same as in Figure 2. The depth and distance (to the vertical well) of the source are set to be 2100 m and 350 m , respectively. Figure 11 shows contour maps of RMS residual of the arrival times (Figure (a)), RMS residual of arrival-time difference (Figure (b)) and the loss value calculated using Equation (5) (Figure (c)). It can be observed that contour maps for both arrival-time residual and loss value obtained with Equation (5) are comparable, and both of them can provide better constraints for V_P and V_S than the contour map of residual for arrival-time difference. While for the event-location problem, similar to Figures 8–10, the contour map of loss value obtained with Equation (5) can provide the best constraint for the event location (shown in Figure 12). Thus, we can conclude that the adoption of the hybrid loss function can provide better

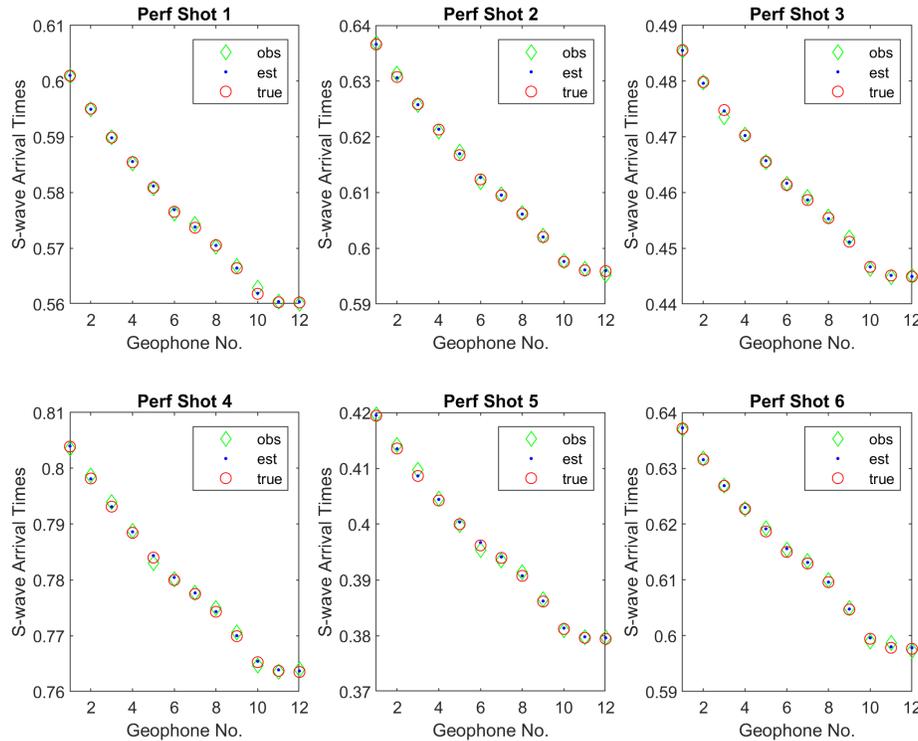


FIG. 7. Comparisons of observed (green diamond), estimated (blue dot) and true (red circles) S-wave arrival times for the six perforation shots.

constrains for both the velocity-calibration and event-location problems.

Uncertainty analysis

In order to investigate the uncertainty of the inverted velocity model caused by noise, we run the velocity-model inversion 100 times using the proposed neural network, and the 100 sets of inversion results are used for the uncertainty analysis. Considering that neural network weights are initialized randomly, the initial velocity values for both P- and S-waves are also randomly generated, which will further affect the final inversion results. Figure 13 shows the density map for estimated layer velocities, in which true values, mean values as well as standard deviations of layer velocities are also indicated. The mean deviations from true velocities are 76.0 m/s and 32.6 m/s for P- and S-waves, respectively. The mean standard deviations are 96.9 m/s and 47.0 m/s , respectively. For the sake of comparison, we also carry out the inversion with only one perforation shots, and the uncertainty-analysis results are shown in Figure 14. The mean deviations from the true values are 120.4 m/s and 46.7 m/s , respectively, and the standard deviations of layer velocities are 119.5 m/s and 48.3 m/s , respectively. Both deviations from true values and standard deviations are larger than the results using six perforation shots, indicating that the inclusion of more perforation shots can provide more accurate solutions with less uncertainty.

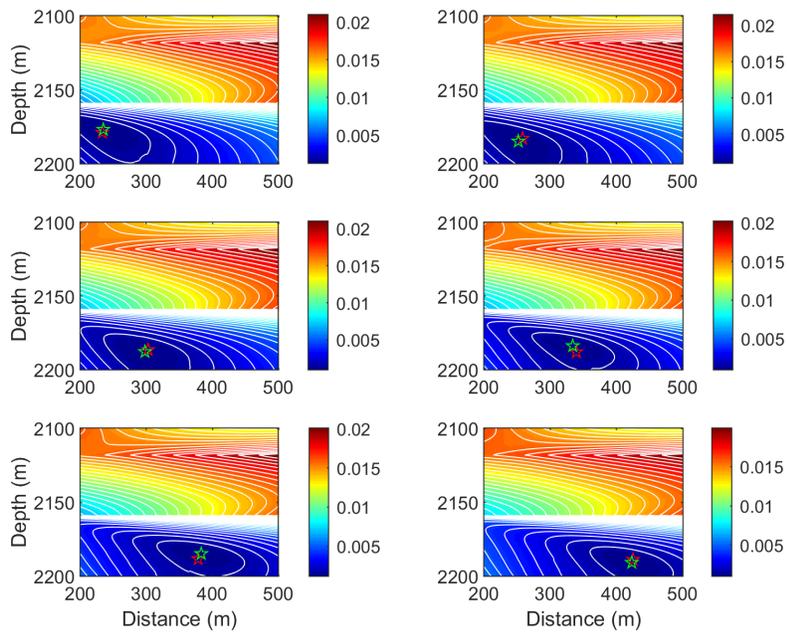


FIG. 8. Contour maps the loss function for the six perforation shots. The loss function in Equation (5) is used. The relocated and true perforation-shot positions are marked by green and red stars, respectively.

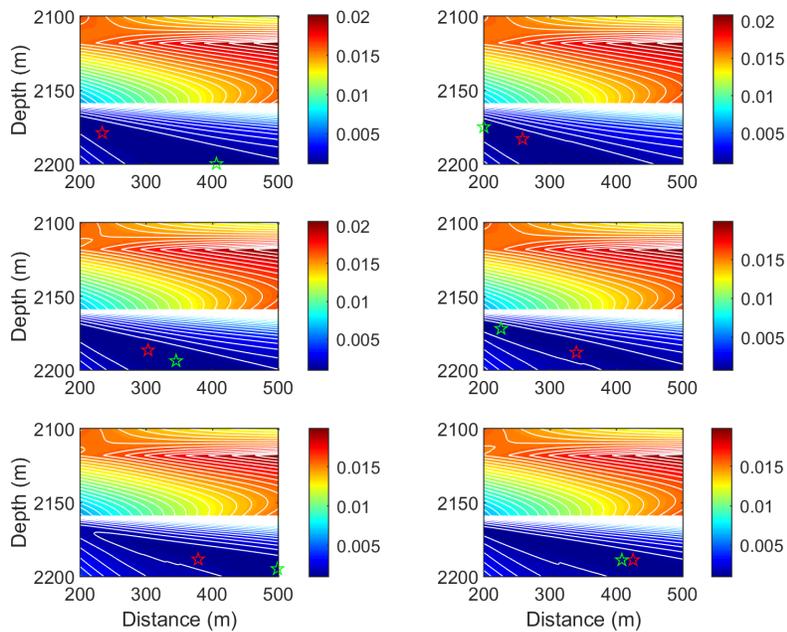


FIG. 9. Contour maps the loss function for the six perforation shots. The RMS residual in arrival times is used as the loss function. The relocated and true perforation-shot positions are marked by green and red stars, respectively.

CONCLUSIONS

We have developed an unsupervised physics-guided neural network to calibrate the 1D layered velocity model with perforation shots recorded by the downhole array. In the

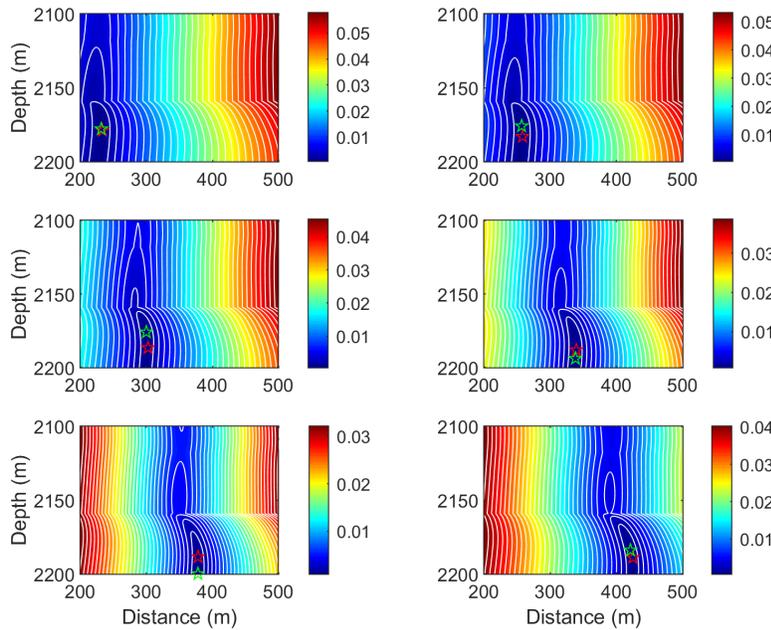


FIG. 10. Contour maps the loss function for the six perforation shots. The RMS residual in arrival-time difference is used as the loss function. The relocated and true perforation-shot positions are marked by green and red stars, respectively.

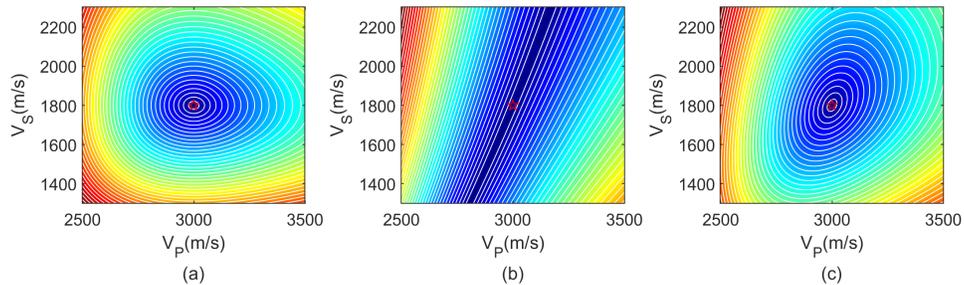


FIG. 11. The contour maps of (a) RMS residual of arrival times, (b) RMS residual of arrival-time difference and (c) the loss value calculated using Equation (5). These contour maps are used for the velocity-calibration problem. The red star denotes the position of true model parameters.

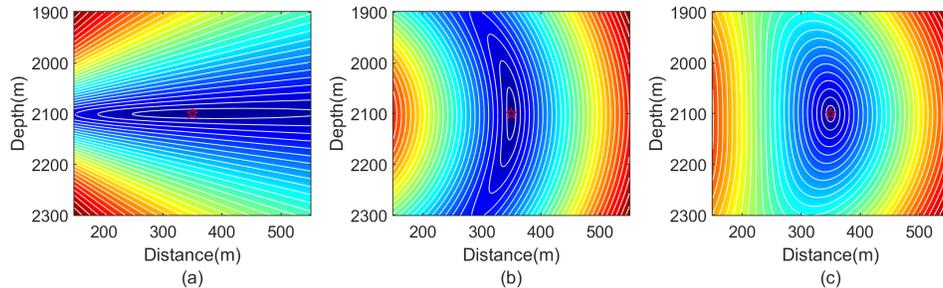


FIG. 12. The contour maps of (a) RMS residual of arrival times, (b) RMS residual of arrival-time difference and (c) the loss value calculated using Equation (5). These contour maps are used for the event-location problem. The true event location is denoted by the red star.

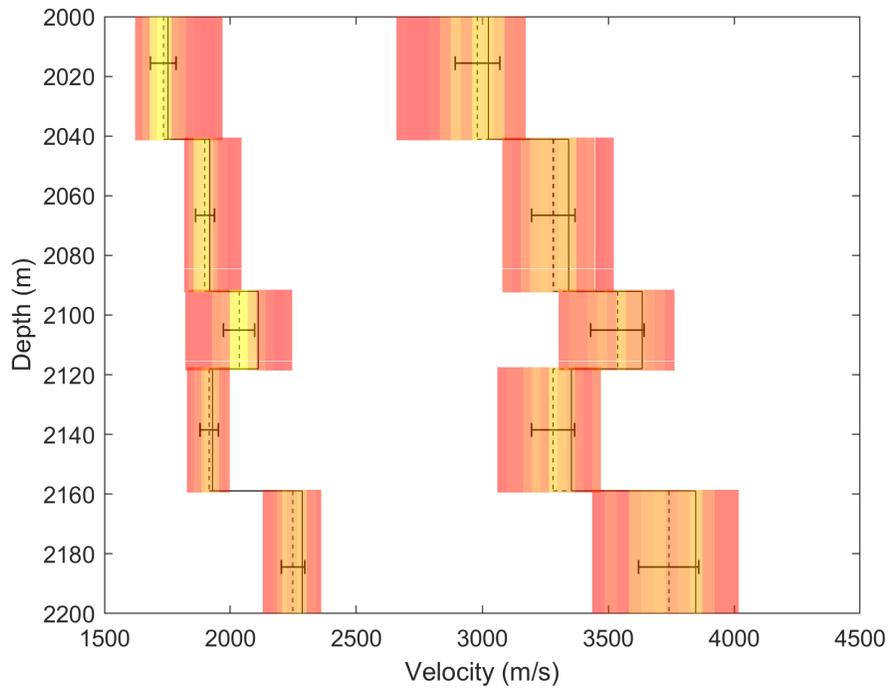


FIG. 13. Uncertainty analysis results using six perforation shots. This figure is color-coded based on the count frequency for layer velocities. Solid lines denote true velocities. Dashed lines denote mean values of inverted layer velocities. The standard deviations are also denoted by horizontal bars.

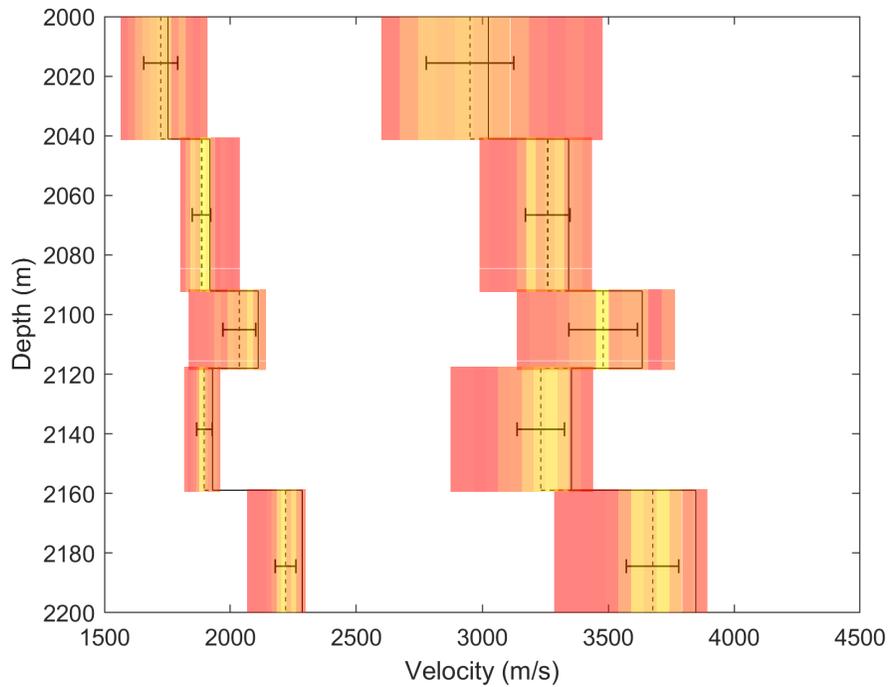


FIG. 14. Uncertainty analysis results using only one perforation shot. All the symbols have the same meaning as in Figure 13.

proposed neural network, five fully connected layers and a Scaling & Shifting layer are used to generate layer velocities. Due to the inclusion of the forward modeling layer for P- and S-wave travel times, the labeled data is not required in the training process. The network operates by minimizing a hybrid loss function that combines residuals of arrival times and time-difference between P- and S-waves. Numerical examples show that, both the updated velocity and relocated event locations exhibit relatively small errors from true values, demonstrating the effectiveness of proposed algorithm. In addition, the effects of noise and random initialization of neural network weights on inversion results have also been investigated through uncertainty analysis. In the future research, field data will be used to further validate the proposed neural network.

ACKNOWLEDGEMENTS

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Sciences and Engineering Research Council of Canada) through the grants CRDPJ 461179-13 and CRDPJ 543578-19. Partial funding also came from the Canada First Research Excellence Fund.

REFERENCES

- Akram, J., and Eaton, D. W., 2016, A review and appraisal of arrival-time picking methods for downhole microseismic data arrival-time picking methods: *Geophysics*, **81**, No. 2, KS71–KS91.
- Cameron, M. K., Fomel, S. B., and Sethian, J. A., 2007, Seismic velocity estimation from time migration: *Inverse Problems*, **23**, No. 4, 1329.
- Cerveny, V., 2005, *Seismic ray theory*: Cambridge university press.
- Chen, Y., Zhang, G., Bai, M., Zu, S., Guan, Z., and Zhang, M., 2019, Automatic waveform classification and arrival picking based on convolutional neural network: *Earth and Space Science*, **6**, No. 7, 1244–1261.
- Eaton, D. W., Akram, J., St-Onge, A., and Forouhideh, F., 2011, Determining microseismic event locations by semblance-weighted stacking, *in* *Proceedings of the CSPG CSEG CWLS Convention*.
- Glorot, X., and Bengio, Y., 2010, Understanding the difficulty of training deep feedforward neural networks, *in* *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 249–256.
- Huang, L., Li, J., Hao, H., and Li, X., 2018, Micro-seismic event detection and location in underground mines by using convolutional neural networks (cnn) and deep learning: *Tunnelling and Underground Space Technology*, **81**, 265–276.
- Jiang, H.-Y., Chen, Z.-B., Zeng, X.-X., Lv, H., and Liu, X., 2016, Velocity calibration for microseismic event location using surface data: *Petroleum Science*, **13**, No. 2, 225–236.
- Kingma, D. P., and Ba, J., 2014, Adam: A method for stochastic optimization: *arXiv preprint arXiv:1412.6980*.
- Kissling, E., 1988, Geotomography with local earthquake data: *Reviews of Geophysics*, **26**, No. 4, 659–698.
- Ma, Y., Cao, S., Rector, J. W., and Zhang, Z., 2020, Automated arrival time picking using a pixel-level network: *Geophysics*, **85**, No. 5, 1–40.
- Nelson, G. D., and Vidale, J. E., 1990, Earthquake locations by 3-d finite-difference travel times: *Bulletin of the Seismological Society of America*, **80**, No. 2, 395–410.

- Pei, D., Quirein, J. A., Cornish, B. E., Quinn, D., and Warpinski, N. R., 2009, Velocity calibration for microseismic monitoring: A very fast simulated annealing (vfsa) approach for joint-objective optimization: *Geophysics*, **74**, No. 6, WCB47–WCB55.
- Pei, D., Quirein, J. A., Cornish, B. E., Zannoni, S., and Ay, E., 2008, Velocity calibration using microseismic hydraulic fracturing perforation and string shot data, *in* SPWLA 49th annual logging symposium, Society of Petrophysicists and Well-Log Analysts.
- Qu, S., Guan, Z., Verschuur, E., and Chen, Y., 2019, Automatic high-resolution microseismic event detection via supervised machine learning: *Geophysical Journal International*, **218**, No. 3, 2106–2121.
- Stork, A. L., Baird, A. F., Horne, S. A., Naldrett, G., Lapins, S., Kendall, J.-M., Wookey, J., Verdon, J. P., Clarke, A., and Williams, A., 2020, Application of machine learning to microseismic event detection in distributed acoustic sensing data: *Geophysics*, **85**, No. 5, KS149–KS160.
- Tan, Y., He, C., and Mao, Z., 2018, Microseismic velocity model inversion and source location: The use of neighborhood algorithm and master station method: *Geophysics*, **83**, No. 4, KS49–KS63.
- Tan, Y., He, C., and Zhang, H., 2013, Time difference-based velocity model inversion for microseismic event location, *in* SEG Technical Program Expanded Abstracts 2013, Society of Exploration Geophysicists, 2223–2227.
- Zhang, X., Zhang, J., Yuan, C., Liu, S., Chen, Z., and Li, W., 2020, Locating induced earthquakes with a network of seismic stations in oklahoma via a deep learning method: *Scientific reports*, **10**, No. 1, 1–12.