PHYSICS-GUIDED UNSUPERVISED DEEP-LEARNING SEISMIC INVERSION WITH UNCERTAINTY QUANTIFICATION

YU ZHANG[1], SAGAR SINGH[2], DAVID THANOON[1], PANDU DEVARAKOTA[2], LONG JIN[1] and ILYA TSVANKIN[3]

[1] *Shell International Exploration and Production Inc., Houston, TX 77082, U.S.A.*
[2] *Shell Global Solutions (US), Inc., Houston, TX 77082, U.S.A.*
  Present address: NVIDIA, Boulder, CO 80301, U.S.A.
[3] *Colorado School of Mines, Golden, CO 80401, U.S.A.*

ABSTRACT

Zhang, Y., Singh, S., Thanoon, D., Devarakota, P., Jin, L. and Tsvankin, I., 2023. Physics-guided unsupervised deep-learning seismic inversion with uncertainty quantification. *Journal of Seismic Exploration*, 32: 257-270.

Data-driven seismic inversion techniques are often used for estimation of subsurface properties. Employing the acoustic or elastic wave equation, inversion starts with approximate initial values of subsurface parameters, which are typically updated in iterative fashion. Here, we propose a two-stage unsupervised machine-learning (ML) methodology for efficient and accurate seismic impedance inversion. The first stage utilizes the generalization capability of convolutional neural networks (CNN) to produce realistic estimates of the acoustic impedance (AI), whereas the second stage incorporates physics information to generate synthetic data from the subsurface AI distribution. We also add Bayesian layers to the first stage of the network to evaluate the model errors. The proposed probabilistic approach to deep learning allows one to estimate the uncertainty of the inverted parameters, which enhances the interpretability of the model. We apply the algorithm to a poststack data set generated using the CGG Hampson-Russell software. After conducting network training with a sufficient number of data points, the network is applied to the rest of the data to estimate the model parameters. The developed approach has a significant advantage over more conventional ML strategies because it produces statistically justified uncertainty maps and eliminates the need to use labeled data for training.

KEY WORDS: seismic inversion, physics guided machine learning, unsupervised learning, uncertainty quantification, model evaluation.

INTRODUCTION

Seismic inversion has been widely used to estimate the elastic properties of the subsurface and generate high-resolution reservoir models (e.g., Sen, 2006; Singh et al., 2021). For example, elastic parameters or reservoir properties can be obtained by deterministic or stochastic inversion of the amplitude-variation-with-offset (AVO) response (e.g., Russell, 1988; Sen and Stoffa, 2013). Deterministic algorithms, which use the gradients of the objective function to update the model parameters, are computationally efficient but often get trapped in local minima of the objective function depending on the accuracy of the initial model and inversion nonlinearity. Regularization terms are often added to improve the convergence of parameter updating. However, regularization may not be effective if the initial model is inaccurate, and the regularized objective function has to operate with smoothed input data (Tikhonov and Arsenin, 1977; Loris et al., 2010; Guitton, 2011; Sen and Biswas, 2014). Stochastic algorithms typically use a Bayesian framework and random sample the model space to minimize the objective function. This approach may help avoid local minima during parameter updating (e.g., Sen and Biswas, 2017).

Recent progress in computational resources (including the advent of graphic processing units or GPUs) has made it feasible to use deep-learning (DL) algorithms for seismic interpretation (Guitton, 2018; Shi et al., 2019; Singh et al., 2022). In particular, DL networks can provide highly efficient segmentation and regression results because they are capable of extracting essential features from large-scale data sets in high-dimensional spaces (Ronneberger et al., 2015). DL algorithms have also been successfully applied to other geophysical problems including classification of salt bodies and fault picking (e.g., Di et al., 2018; Wu et al., 2018).

However, current deep-learning applications, such as those based on CNNs (convolutional neural networks), are mostly limited to classification problems rather than inversion for subsurface properties. Das et al. (2018) applied CNNs to a supervised regression problem of estimating acoustic impedance from seismic data. Their inversion algorithm is based on convolutional operators because seismic reflection data can be represented as the result of convolving the source wavelet with the reflectivity series. Here, we assume the same model but solve the regression problem using a modified network trained in unsupervised fashion. Forward simulation is employed to generate the data for the predicted model and compare them to the observed data. The error of the predicted data then allows us to update the CNN weights (Biswas et al., 2019). This machine-learning algorithm referred to as DLI (deep-learning inversion).

Most convolutional networks produce pointwise estimates of their weights, but that information does not fully capture the uncertainty in the values of the weights. In our architecture, the encoder half of the network is supplemented with Bayesian layers to evaluate the uncertainties in the weights' space. The decoder half uses a physics-based approach to produce estimates of poststack data; we call this approach "Bayesian DLI." Then we introduce "Dropout DLI," which evaluates the uncertainty in the weights by employing dropout layers in the encoder part of the network. Finally, we compare the uncertainty quantification results of the two proposed algorithms.

UNSUPERVISED TRAINING MODEL

Unsupervised ML is designed to learn distinct distributions from the input data. Fig. 1 shows our proposed CNN architecture, whose main components are described below.

Encoder model: Stage 1

The input is a 1D tensor or vector of a fixed time length (NT), which represents a stacked seismic trace. The output is the 1D acoustic impedance (AI) of the same size (NT). The CNN consists of a feed-forward stack of two convolutional layers (Conv1 and Conv2). The size of the first layer (Conv1) is (time length of the wavelet $\times$ 1), with 60 filters or output channels and a stride of one. The filter dimension in time is the length of the wavelet, which is sufficient for capturing enough features or spatial dependencies in the region where that filter operates. Because this is a trace-by-trace operation, the second dimension is unity. The "stride" represents the number of rows/columns in the filter shifts over the input matrix while applying the convolution operation, and "1" in the second dimension implies no shifting.

The second convolutional layer (Conv2) has the same size as the first one, but only a single output channel with a stride of one. We conduct hyperparameter tuning to determine the filter size (i.e., 60) in the first convolutional layer. The second layer employs a filter of unit size, which yields a single time-series output. Setting the stride to one maintains the output dimension equal to that of the input. This network produces a time series, which is treated as our desired output (AI). After each convolutional layer, activation is performed by employing a nonlinear rectified linear-unit (ReLU) function.
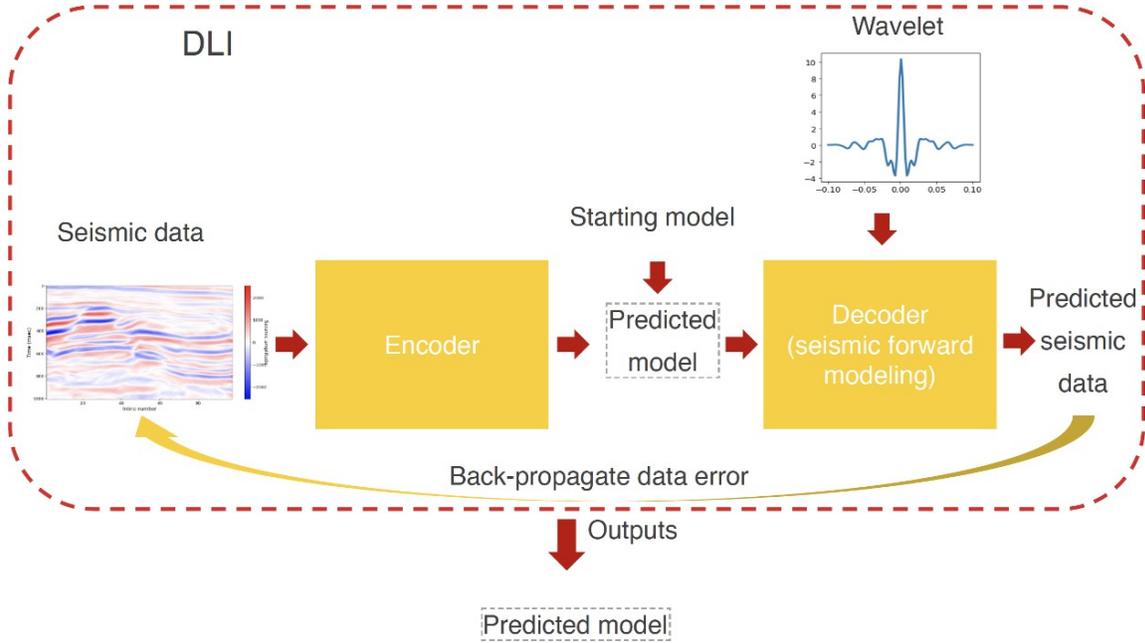
Fig. 1. Outline of the proposed CNN architecture for physics-based inversion. The acoustic impedance is estimated from stacked seismic gathers, and the obtained impedance model is used to generate seismic data by convolving the wavelet with the reflectivity series. Then the simulated data are compared with the observed data set.

Decoder model: Stage 2

Because our objective is to estimate the AI from stacked seismic traces, a forward-modeling operator is included in the decoder. We extend the encoder architecture, use the output model generated by the encoder, add a low-frequency component of the acoustic impedance (assumed to be known), and simulate seismic data using the forward-modeling operator. Finally, we compare the modeled and recorded seismic data to calculate the data misfit. Fig. 1 shows the modified network designed for poststack impedance inversion. We add a low-frequency model to the encoder network output, calculate the reflectivity series $r\,(t)$, and convolve it with the wavelet $w(t)$ to generate the poststack seismic trace d_{cal}:

$$r\,(t) = \frac{\Delta AI}{AI} \quad , \qquad (1)$$

$$d_{\mathrm{cal}} = r\,(t) \, * \, w\,(t) \quad , \qquad (2)$$

where $\overline{AI} = (\rho_1 V_{P\,1} + \rho_2 V_{P\,2})/2$ is the mean acoustic impedance for two consecutive layers and ΔAI is the difference between those impedances. The mean-square difference between the observed data d_{obs} and simulated data d_{cal} is:

$$E = \sum_{i=1}^{NT} \frac{(d_{\text{cal}} - d_{\text{obs}})^2}{NT} \qquad . \tag{3}$$

We also add to eq. (3) a regularization term, $\lambda||AI_{\text{Pred}} - AI_{\text{Initial}}||$, to constrain the predicted AI model; λ is a tunable parameter.

Bayesian inference

Probabilistic approach to deep learning allows one to enhance the interpretability of the model by estimating the uncertainty in the output (Singh et al., 2022). Each Bayesian convolutional layer is initialized with the standard normal prior $P(\mathbf{w}) = N(0, 1)$ (the vector $\mathbf{w}$ represents the network weights) and employs a flipout estimator (Wen et al., 2018) to approximate the posterior distribution during forward passes. The flipout estimator provides the Monte Carlo approximation of the posterior distribution by integrating over the Bayesian layer's kernel and bias and significantly lowering the network variance (Wen et al., 2018). We employ the variational free-energy loss function F to approximate the posterior distribution over the network's weights:

$$\mathcal{F}_i(\mathcal{D}_i, \theta) = \frac{1}{M} KL\Big[q(\mathbf{w}|\theta)||P(\mathbf{w})\Big] - E_{q(\mathbf{w}|\theta)}\Big[\log P(\mathcal{D}_i|\mathbf{w})\Big]$$

$$\tag{4}$$

where M is the total number of training examples, i is the minibatch, and $q(\mathbf{w}|\theta)$ represents the variational inference (θ is a variational parameter). The first and second terms on the right-hand side of eq. (4) are the so-called KL (Kullback-Leibler) divergence and crossentropy (D_i represents labels), respectively. The first term is divided by M to optimize the minibatch $I \in \{1,2,...M\}$, as proposed by Graves (2011). This helps distribute the KL divergence penalty evenly over the minibatches. Eq. (4) can be interpreted as a tradeoff between fitting the data (the crossentropy term) and satisfying the simplicity prior (the KL term).

To stabilize the network, we employ the Adam optimizer (Bowman et al., 2016). Bayesian layers are included only in the encoder half of the network to maximize the information transfer between the input and the latent space (LaBonte et al., 2020).

VALIDATION OF DECODER MODEL (FORWARD MODEL)

To validate the self-developed decoder model (forward-modeling operator), we compare it with a widely used geophysics software RokDoc (by Ikon Science). We pick a trace of synthetic acoustic impedance from RokDoc, use it as an input to our self-developed decoder model to generate a trace of synthetic seismic data, and compare that trace with the seismic calculated by RokDoc (Fig. 2).
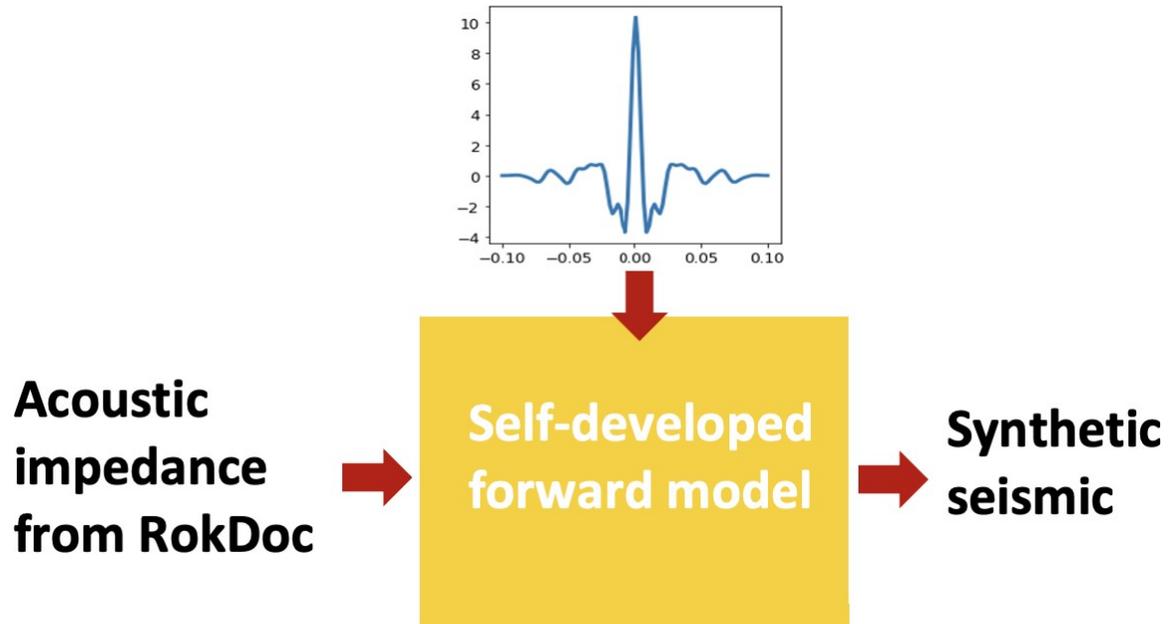

Fig. 2. Validation of the decoder model with RokDoc: Workflow.

As shown in Fig. 3, the synthetic seismic generated by our self-developed (forward) decoder model matches very well with the output of Rokdoc. (The convention of convolu in Tensorflow (Abadi et al., 2016) is different from that in RokDoc, so a factor of -1 needs to be applied to make the comparison.)

NUMERICAL EXAMPLES

Three-layer model

We first test our workflow on a synthetic layer-cake model composed of three homogeneous layers. The actual acoustic impedance of this model is shown in Fig. 4, and this synthetic impedance is used to generate seismic data. The initial acoustic impedance model is similar to the actual model, but

with a smaller AI amplitude in the second layer. The generated seismic and the initial AI model represent two inputs to our proposed unsupervised physics-guided deep-learning algorithm. The obtained prediction of the acoustic impedance is displayed in Fig. 5. The predicted seismic data accurately matches the actual seismic data, while the predicted acoustic impedance is close to the initial acoustic impedance model, although it can add fluctuations to the initial model.
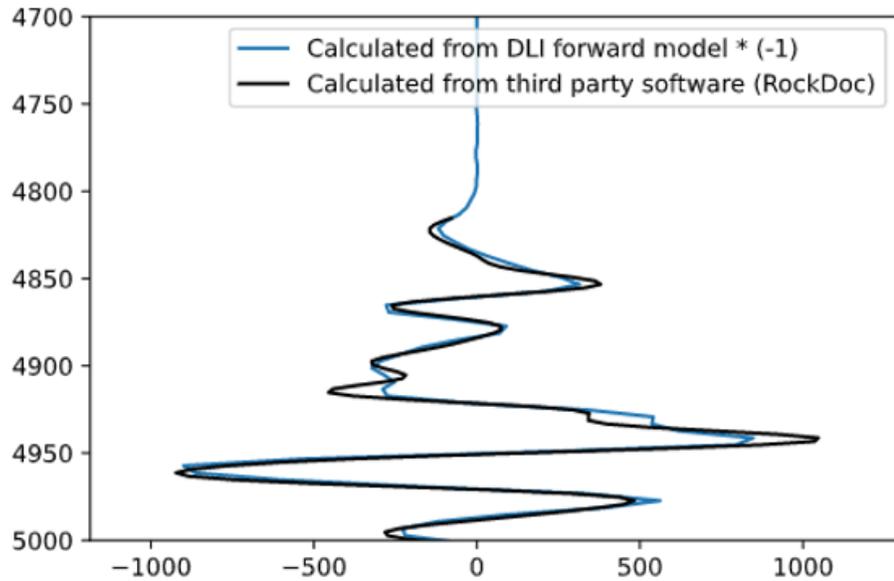

Fig. 3. Validation of the decoder model with RokDoc: Result.
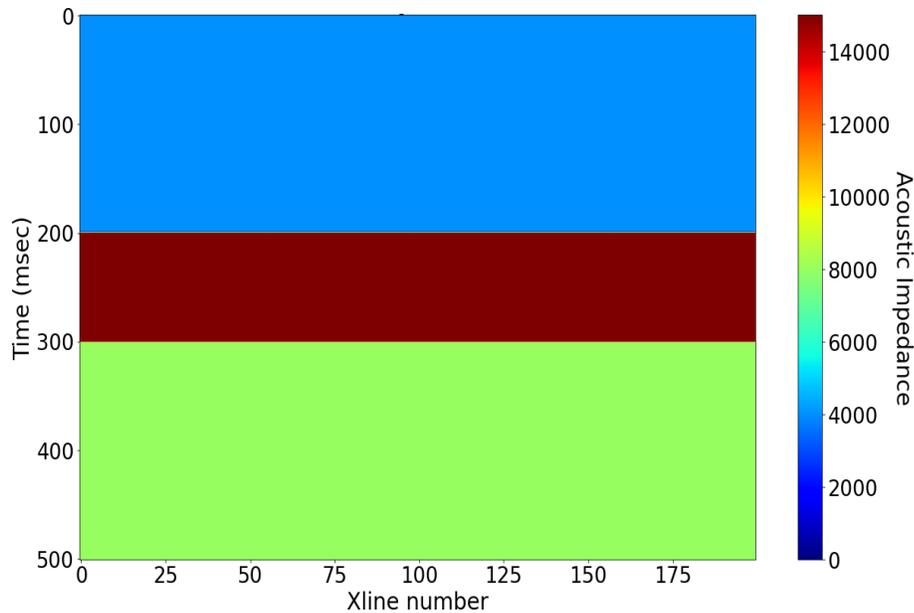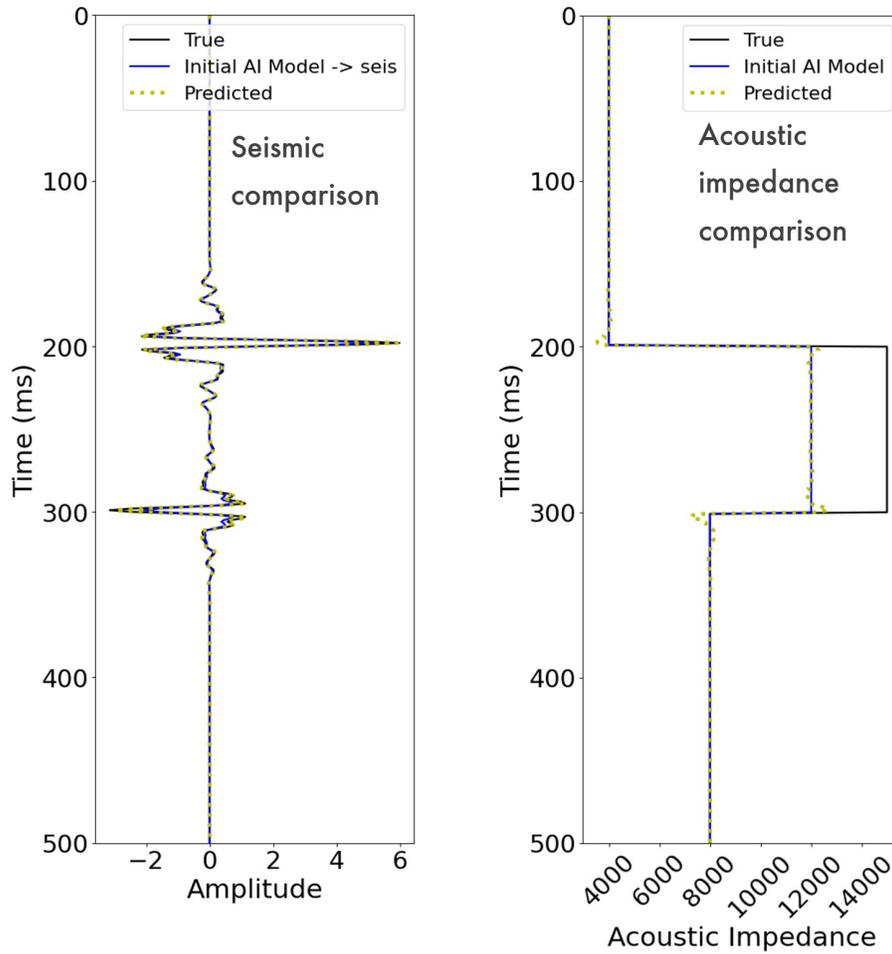

Fig. 4. Synthetic three-layer AI model.

Fig. 5. Performance of the developed DLI algorithm on the model from Fig. 4.

This demonstrates the nonuniqueness of the inverted impedance and the importance of the accurate initial model. Although this test confirms that our DLI algorithm has the ability to predict the high-frequency component of acoustic impedance, it still requires a reasonably accurate initial model for predicting the low-frequency AI component.

CGG Hampson-Russell model

To demonstrate the network's performance, we use poststack seismic data generated by the Hampson-Russell CGG software (Fig. 6a). This is a relatively small 3D data set, which includes the actual acoustic model and the source wavelet. We randomly select 80% of the traces for training, whereas the remaining data are set aside for network testing purposes. Based on the

training history, we stop training at approximately 200 epochs to minimize possible overfitting. The batch size is set to 16 traces, with batch shuffling at the end of each epoch. The input includes stacked seismic gathers and the low-frequency acoustic impedance model constructed by smoothing the actual model. We utilize a dynamic learning-rate methodology that has a starting value of 0.001.
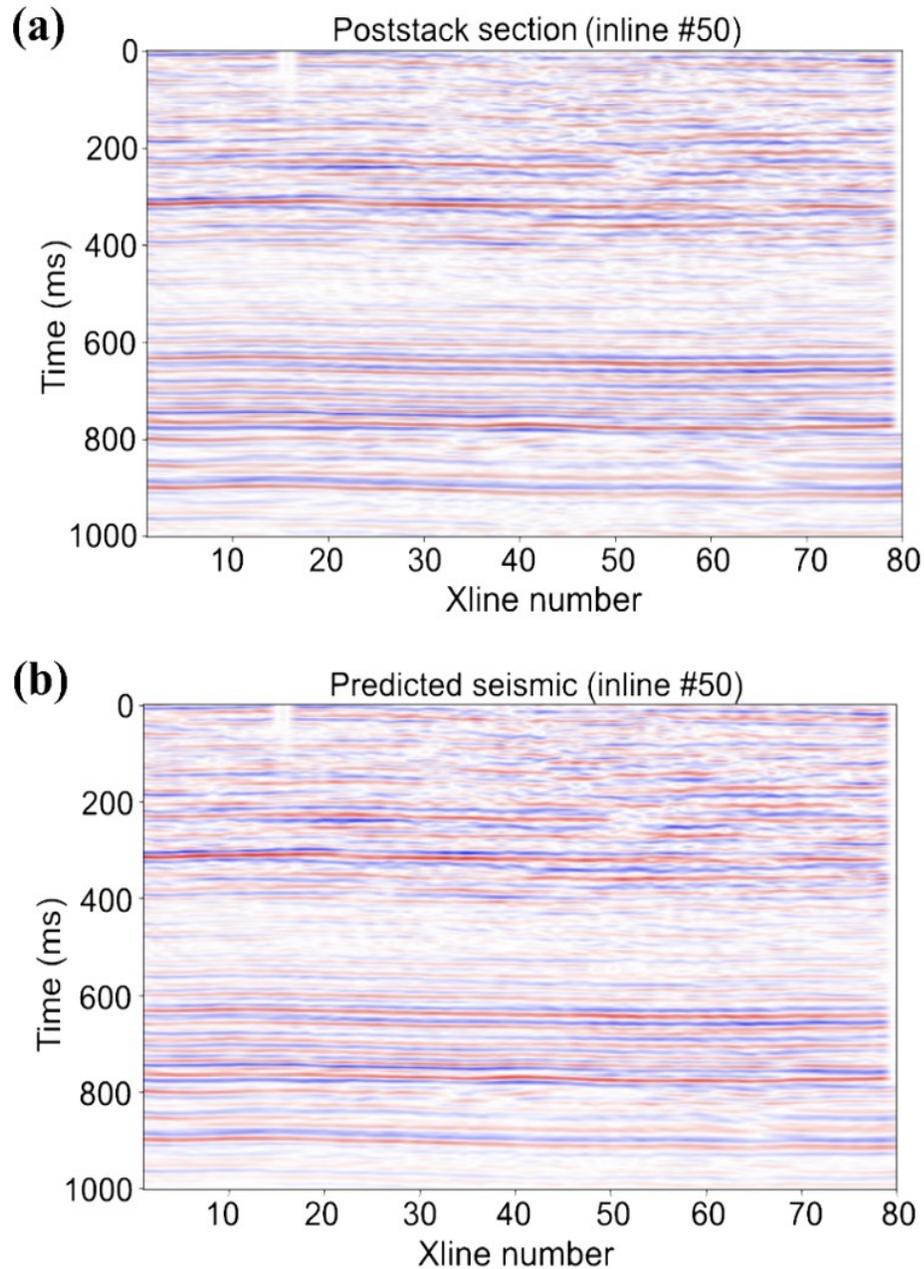

Fig. 6. (a) Inline #50 of the Hampson-Russell data set. (b) The mean poststack seismic data estimated by DLI.

To validate the network after training, we display the mean of both the poststack data and the acoustic impedance obtained from the testing data set. Figs. 6-8 illustrate the accuracy of the seismic data and of the impedance generated by the proposed networks. The parameters estimated by both networks are close to the actual values, with an R^2 score of 0.73.
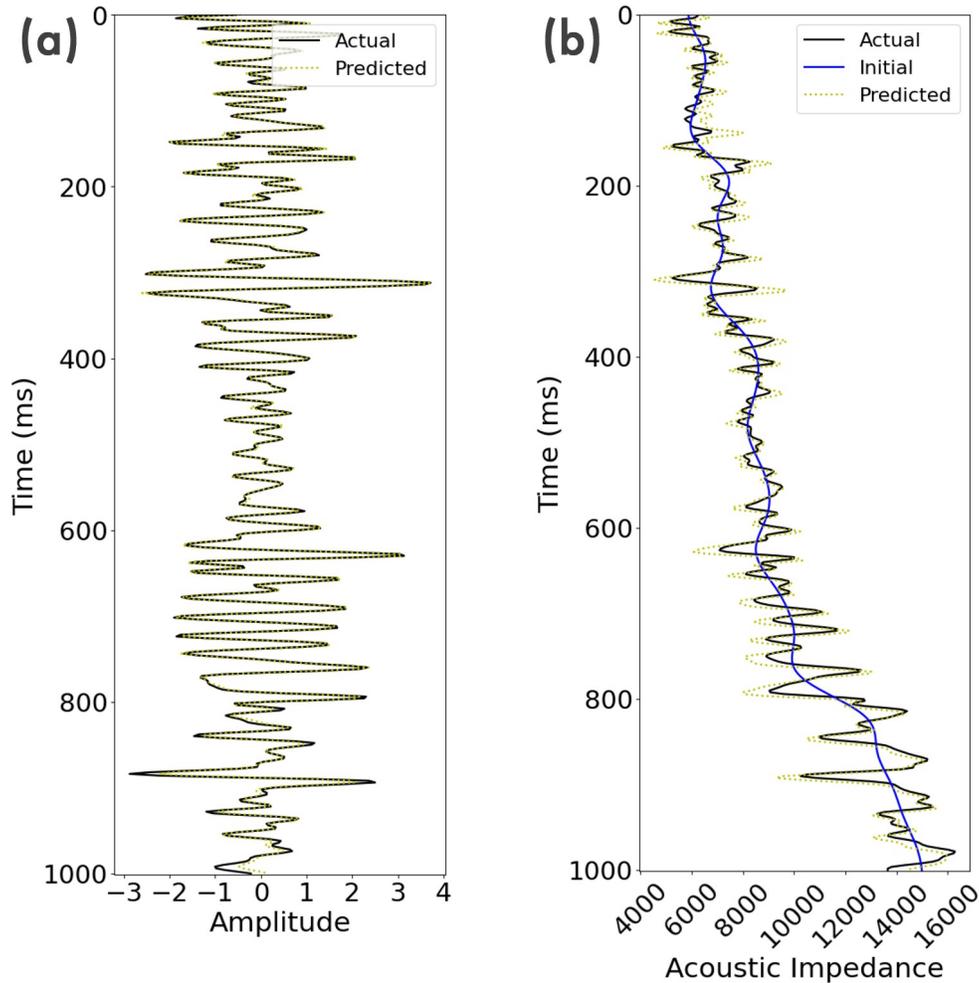

Fig. 7. (a) Comparison of the actual (black) and predicted (yellow) seismic traces for one of the CDPs. (b) Comparison of the actual (black), initial (blue), and predicted (yellow) acoustic impedance (AI) at the same CDP location

To obtain the posterior distribution for each grid point, 100 Monte Carlo samples are computed from both the Bayesian DLI and the Dropout DLI (the dropout probability is set to 0.2). We define the confidence intervals for the inverted AI as the 30th and 70th percentiles of the softmax values and the uncertainty as the difference between these percentile values (Fig. 9). It is clear that the Bayesian DLI has a lower standard deviation and provides a better uncertainty quantification compared with the Dropout DLI.
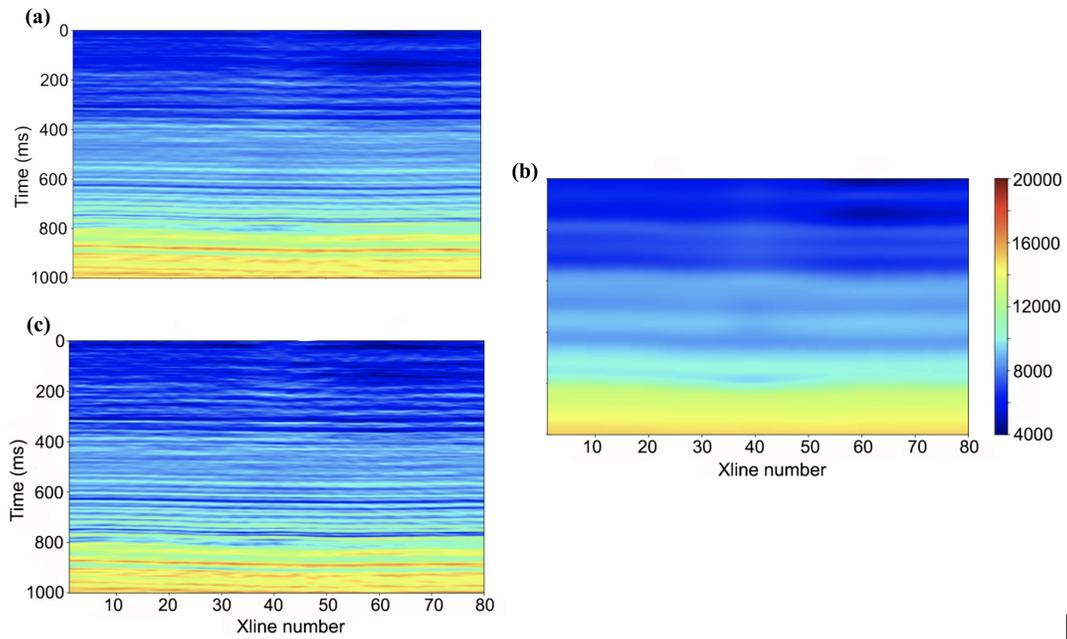
Fig. 8. (a) Actual, (b) initial, and (c) inverted AI models for inline #50.
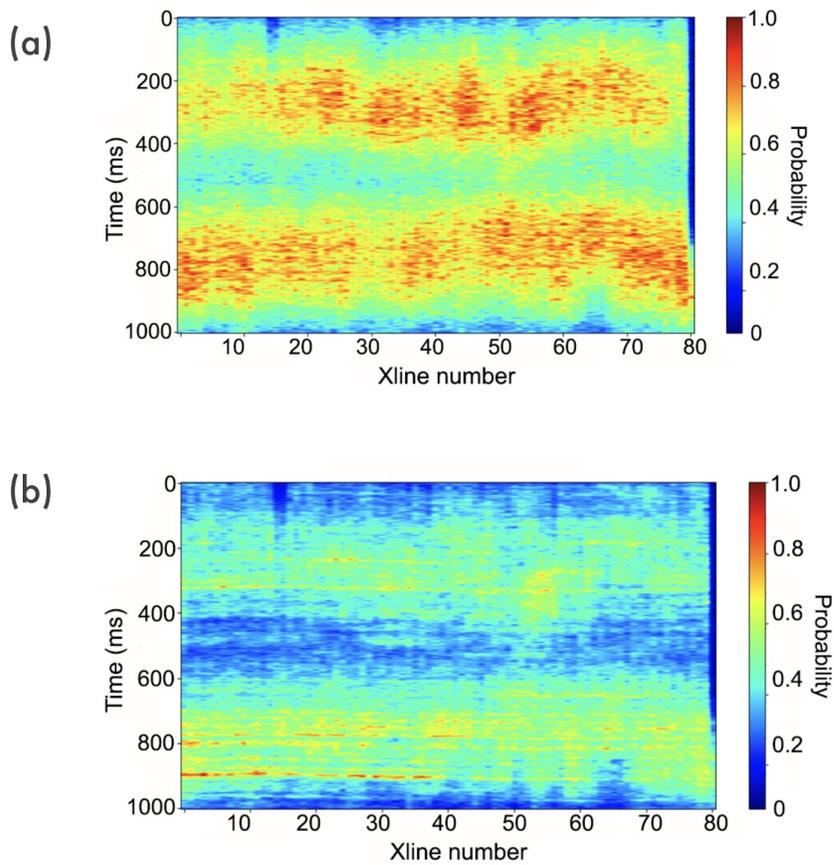

Fig. 9. Uncertainty maps of the posterior prediction of the acoustic impedance corresponding to the confidence interval between 30-70%: (a) Bayesian DLI; (b) Dropout DLI.

ANALYSIS OF SENSITIVITY TO INITIAL MODEL

Because one of the most important inputs to our algorithm is the initial model of the inverted properties (i.e., the acoustic impedance), it is essential to test the influence of the initial AI distribution on the inversion results.

The impedances predicted using homogeneous (Fig. 10a) and Hampson-Russell's (Fig. 10b) initial models look generally similar. However, the AI obtained with the Hampson-Russell model (which is more accurate) has a higher vertical resolution and better separation between layers.
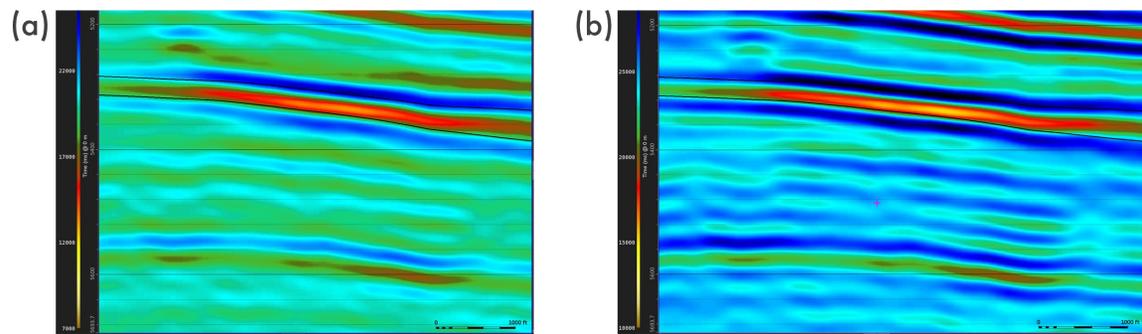

Fig. 10. Acoustic impedance predicted using (a) a homogeneous (constant AI) initial model and (b) the Hampson-Russell initial model.

CONCLUSIONS

We presented an application of Convolutional Neural Networks (CNNs) to the practically important problem of seismic impedance inversion. We made two significant modifications in the CNNs by estimating the geophysical parameters instead of employing classification and by making the network training unsupervised via incorporating physics information.

The results show that physics-based methods applied to poststack seismic inversion can predict the spatial distribution of the acoustic impedance with acceptable accuracy. The unsupervised physics-based network eliminates the need to have a known response (i.e., the actual label parameters) for the training, which is often difficult to obtain for field data. By estimating the uncertainty in the weight space, the employed Bayesian scheme provides interpretable quantification of the inversion errors. The method was successfully applied to realistic synthetic data, which confirms the high potential of deep-learning techniques in seismic inversion and

quantitative interpretation. We conclude that deep learning provides a robust framework for combining seismic data and low-frequency information about elastic properties in automated estimation of attributes for reservoir characterization. The developed strategy is currently being tested on a field data set from the Gulf of Mexico.

ACKNOWLEDGMENTS

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Schuster, M., Monga, R., Moore, S., Murray, D., Olah, C., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y. and Zheng, X., 2016. TensorFlow: Large-scale machine learning on heterogeneous systems. 12th *USENIX* Symposium Operating Systems, Design and Implementation: 265-283.
Biswas, R., Sen, M.K. and Das, V. and Mukerji, T., 2019. Prestack and poststack inversion using a physics-guided convolutional neural network. Interpretation, 7(3): SE161-SE174.
Das, V., Pollack, A., Wollner, U. and Mukerji, T., 2018. Convolutional neural network for seismic impedance inversion. Expanded Abstr., 88th Ann. Internat. SEG Mtg., Anaheim: 2071–2075.
Di, H., Wang, Z. and AlRegib, G., 2018. Deep convolutional neural networks for seismic salt-body delineation. AAPG Ann. Conv., Salt Lake City.
Graves, A., 2011. Practical variational inference for neural networks. Conf. Advances Neural Informat. Process. Syst.: 2348-2356.
Guitton, A., 2011. A blocky regularization scheme for full waveform inversion. Expanded Abstr., 81st Ann. Internat. SEG Mtg., San Antonio: 2418-2422.
Guitton, A., 2018. 3D convolutional neural networks for fault interpretation. Extended Abstr., 80th EAGE Conf., Copenhagen: 1–5.
LaBonte, T., Martinez, C. and Roberts, S.A., 2020. We know where we don't know: 3D Bayesian CNNs for credible geometric uncertainty.
arXiv Preprint, arXiv: 1910.10793.
Loris, I., Douma, H., Nolet, G., Daubechies, I. and Regone, C., 2010. Nonlinear regularization techniques for seismic tomography. J. Computat. Phys., 229: 890-905.
Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. arXiv Preprint, arXiv: 1505.04597.
Russell, B.H., 1988. Introduction to Seismic Inversion Methods. SEG, Tulsa, OK.

Sen, M.K., 2006. Seismic Inversion. SPE, Houston, TX.

Sen, M.K. and Stoffa, P.L., 2013. Global Optimization Methods in Geophysical Inversion. Cambridge University Press, Cambridge.

Sen, M.K. and Biswas, R., 2014. Choice of regularization weight in basis pursuit reflectivity inversion. J. Geophys. Engineer., 12: 70-79.

Sen, M.K. and Biswas, R., 2017. Transdimensional seismic inversion using the reversible jump Hamiltonian Monte Carlo algorithm. Geophysics, 82(3): R119-R134.

Shi, Y., Wu, X. and Fomel, S., 2019. SaltSeg: Automatic 3D salt segmentation using a deep convolutional neural network. Interpretation, 7(3): SE113-SE122.

Singh, S., Tsvankin, I. and Naeini, E.Z., 2021. Elastic FWI for orthorhombic media with lithologic constraints applied via machine learning: Geophysics, 86(4): R589-R602.

Singh, S., Tsvankin, I. and Zabihi Naeini, E., 2022. Facies prediction with Bayesian inference: Application of supervised and semisupervised deep learning. Interpretation, 10(2): 279-290. doi:10.1190/int-2021-0104.1.

Tikhonov, A. and Arsenin, V., 1977. Solution of Ill-posed Problems. Winston and Sons, London: 1-258.

Wen, Y., Vicol, P., Ba, J., Train, D. and Grosse, R., 2018. Flipout: Efficient pseudo-independent weight perturbations on mini-batches. arXiv Preprint, arXiv:1803.04386.

Wu, X., Shi, Y., Fomel, S. and Liang, L., 2018. Convolutional neural networks for fault interpretation in seismic images. Expanded Abstr., 88th Ann. Internat. SEG Mtg., Anaheim: 1946-1950.