

Data-Driven Design of Wave-Propagation Models for Shale-Oil Reservoirs Based on Machine Learning

Fansheng Xiong^{1,2}, Jing Ba¹ , Davide Gei³, and José M. Carcione^{1,3} 

¹School of Earth Sciences and Engineering, Hohai University, Nanjing, China, ²Institute of Applied Physics and Computational Mathematics, Beijing, China, ³National Institute of Oceanography and Applied Geophysics—OGS, Trieste, Italy

Special Section:

Machine learning for Solid Earth observation, modeling and understanding

Key Points:

- Machine-learning methods are proposed to drive wave propagation based on seismic data
- Optimal prediction results of seismic attributes are obtained by feeding rock-physics properties into the trained models
- The proposed method has the potential to obtain additional physical quantities and shows a broad application prospect

Correspondence to:

J. Ba,
jingba@188.com

Citation:

Xiong, F., Ba, J., Gei, D., & Carcione, J. M. (2021). Data-driven design of wave-propagation models for shale-oil reservoirs based on machine learning. *Journal of Geophysical Research: Solid Earth*, 126, e2021JB022665. <https://doi.org/10.1029/2021JB022665>

Received 24 JUN 2021
Accepted 23 NOV 2021

Author Contributions:

Conceptualization: Jing Ba, José M. Carcione
Data curation: Jing Ba
Formal analysis: José M. Carcione
Funding acquisition: Jing Ba
Investigation: Fansheng Xiong, Davide Gei
Methodology: Fansheng Xiong, Jing Ba
Project Administration: Jing Ba
Resources: Jing Ba
Supervision: Jing Ba, José M. Carcione
Validation: Fansheng Xiong, Davide Gei
Writing – original draft: Fansheng Xiong, Jing Ba
Writing – review & editing: Jing Ba, Davide Gei, José M. Carcione

Abstract The exploration and exploitation of shale oil is an important aspect in the oil industry. Seismic properties and well-log data are essential to establish wave-propagation models. Specifically, the description of wave dispersion and attenuation under complex geological conditions needs proper lithological and petrophysical information. This complex physical mechanism has to be considered if a traditional modeling approach is adopted. In this sense, machine learning (ML) techniques provide new possibilities for this purpose. We compare two deep-neural-network (DNN)-based wave propagation models. In the first (pure data-driven), a DNN is trained to connect seismic attributes, such as wave velocities, to multivariate functions of rock-physics properties. By training DNNs with different initial parameters, the uncertainty of the proposed method can be quantified. The second method assumes the form of the wave equations. Then, the elastic constants of the constitutive relations are predicted by DNNs. The resulting dynamical equations describe the dispersion and attenuation and wavefield simulations can be performed to obtain more information. On the basis of a test, the two kinds of wave-propagation models yield acceptable estimations of the seismic properties, with the second approach showing a broader application because the DNN is trained without S wave data. The methodologies illustrate that the new wave-propagation model based on ML has high precision and can be general in terms of rheological description.

Plain Language Summary We develop a model to estimate the underground conditions, based on two machine-learning techniques by using measurement data as input. The first uses pure data-driven surrogate models, while the second is based on the classical wave propagation model in fluid-saturated porous medium proposed by Biot, where deep neural networks predict the elastic coefficients so that the theoretical seismic properties match the actual measurements. The two models are tested on the basis of data from five wells located in a shale-oil reservoir area. Compared with the pure data-driven model, the wave-equation method may be used to predict more reservoir properties.

1. Introduction

The exploration of unconventional oil/gas resources under complex geological conditions is an important aspect of applied geophysical studies in the petroleum industries (e.g., Avseth & Carcione, 2015). The interpretation based on seismic data is helpful for understanding the reservoir properties, but an optimal model and wave-propagation approach are required to honor the in situ geological conditions. Classical methods use a priori lithological and petrophysical information and partial differential equations (PDEs), while real data is used for verification and calibration. Different dynamical equations have been proposed. Biot studied wave propagation in fully saturated porous media (Biot, 1956a, 1956b, 1962), but could not explain the level of attenuation at seismic frequencies, while Dvorkin and Nur (1993) and Dvorkin et al. (1995) introduced the Biot-squirt model and made some progress in this sense. Pride and Berryman (2003a, 2003b) established the governing equations for propagation in double-porosity dual-permeability (DPDP) media, where the mesoscopic-loss mechanism properly describes the frequency dependency of the seismic properties. A similar double-porosity theory was proposed by Ba et al. (2011), known as the Biot-Rayleigh (BR) model, assuming porous spherical inclusions embedded in a host skeleton. Ba et al. (2017) and Zhang et al. (2021) presented a double double-porosity model, where the heterogeneities include the pore structure and patchy saturation, and W. Sun et al. (2018) proposed a three-layer ellipsoidal fluid patch based on the BR model. Berryman (1988) assume two immiscible fluids and derived wave equations based on a variational method. The theory by Santos, Corberó, & Douglas (1990) and Santos, Douglas, et al. (1990) generalizes that of Biot to partial saturation of immiscible fluids by considering capillary effects. Other similar models include Lo et al. (2005, 2006) and Tuncay and Corapcioglu (1996).

The presence of fabric heterogeneity and multi-phase fluids complicate the interpretation (Liu et al., 2018). Moreover, the number of differential equations increases the difficulty and computation cost of seismic inversion and wavefield simulation. Some studies focus on effective models (e.g., Pride et al., 2004), and the DPDP equations can be expressed as Biot's equations, but with complex and frequency-dependent coefficients. Liu et al. (2016, 2018) and Zhao et al. (2015, 2016, 2020) considered these effective poro-elastic and poro-viscoelastic equations, respectively.

Machine learning (ML) techniques, such as deep neural networks (DNN), can be applied to solve geophysical problems. For instance, the generative adversarial network has been used for digital rock reconstruction (You et al., 2021) and seismic waveform synthesis (N. Wang et al., 2021; T. Wang et al., 2021). Qadrouh et al. (2019) give a tutorial on ML, including petrophysical log prediction. Specifically, they applied a neural network (NN) to relate the density, sonic, gamma ray, and neutron-porosity logs with permeability. Synthetic data are generated according to a theoretical model developed by Carcione et al. (2000) for training and test. You et al. (2020) invert for shale-anisotropy by using a DNN model, based on the Hudson-Cheng forward model (Cheng, 1993) to generate a large amount of training data. Compared with the traditional time-consuming and computationally expensive methods (e.g., Barajas-Solano et al., 2015; Li et al., 2019; Wiese et al., 2018), this new approach is faster. In fact, this direct DNN-based method can also be used for efficient wavefield simulations (Moseley et al., 2020). However, this method is fully data-driven both for the forward and inverse problems.

Other techniques train the governing equations from data (e.g., Brunton et al., 2016; Champion et al., 2019; Karimpouli & Tahmasebi, 2020; Raissi et al., 2019; Rudy et al., 2019). There are also some model-building methods in high-dimension parameter space based on an optimization principle (Han et al., 2013, 2017). However, the governing equations can be well established, but their coefficients are not. Tartakovsky et al. (2020) studied subsurface flow assuming a known dynamical equation, but an unknown hydraulic conductivity. They applied a physics-informed neural network (PINN) proposed by Raissi et al. (2017a, 2017b, 2019) to simulate the flow and learn the unknown field from data. The DNN model was used to predict the hydraulic conductivity so that the calculated state quantity could match the data. Similarly, Wang et al. (2020), N. Wang et al. (2021) and T. Wang et al. (2021) proposed a theoretical guidance neural network to simulate subsurface flow. More realistic results can be obtained by adding more penalty terms to the loss function. PINN can also be used as a solver of the acoustic wave equations to speed up the simulation (Pettit & Wilson, 2020; Song et al., 2021).

One potential problem in the aforementioned studies is the source of training data, coming from the solution of theoretical models rather than from actual measurements. The purpose of this study is to learn the wave-propagation model from measured data by means of ML, which differs from the forward modeling methods. We develop a new approach to establish models based on DNN when log data and rock-physics data are available. Two DNN-based wave propagation model building methods are proposed and compared. In the first method, described in Section 2, DNN is trained to obtain a surrogate model describing wave propagation. Such a method is easy to implement, and the accuracy is highly dependent on the amount of training data. The physics is not present in this method since it produces a pure data-driven model. The second DNN-based method is introduced in Section 3. It assumes that the wave equation describing propagation in shale-oil reservoirs is known, while their coefficients need to be learned from data. DNN models are designed to predict these coefficients so that the predicted seismic properties computed by a plane-wave analysis method can match the data. As a result, wave equations in the form of PDEs along with surrogate models of elastic coefficients are obtained. The two kinds of DNN models are trained and tested with the same data set. Examples are given in Section 4, and the advantages and disadvantages of the two methods are discussed.

2. DNN-Based Surrogate Model

2.1. Data, Input, and Output

The data include the well-log measurements of 16 wells of shale-oil reservoirs from Ordos basin, China, which is an inland depression lake in Late Triassic. A set of continental clastic rock sedimentary formations developed in the Yanchang Formation. The Chang 7 member contains organic rich shale with an extensive distribution area, which is the main target of unconventional oil exploration (Bai & Ma, 2019; Fu et al., 2020). The oil in tight sandstones, siltstones and shales of the Chang 7 member are considered typical shale-oil resources. The data were measured by the instruments of Schlumberger's SonicScope and provided by the Changqing Oilfield of

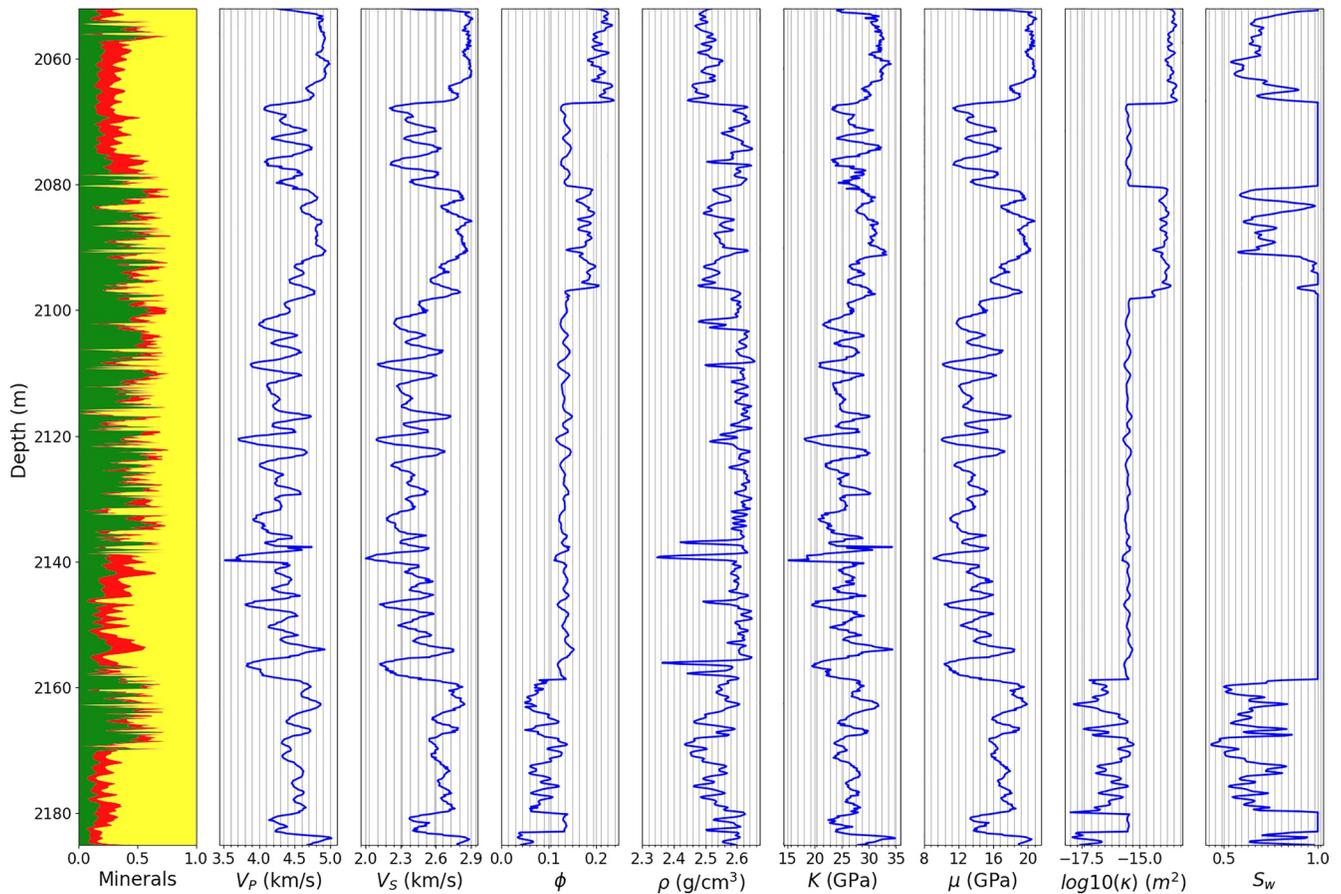


Figure 1. Logging data from a well intersection shale-oil reservoirs. The subplot on the left end gives the volume fractions of the mineral constituents, where green, red, and yellow represent clay, shale, and quartz, respectively. The other subplots from left to right are P wave velocity, S wave velocity, porosity, bulk density, bulk modulus, shear modulus, permeability in log10 scale and water saturation.

PetroChina. After pre-processing, there are 21,215 effective data points from these wells for training and tests, where all quantities are available. The amount of data is different for each well. Five wells with relatively abundant data points are selected as a test set and the rest, with 10,004 data points, constitute the training set. Figure 1 shows the data samples from one of the wells, including porosity, lithology, and the seismic attributes.

By analogy with the classical wave-propagation model, six parameters are selected, namely, porosity ϕ , bulk modulus of fully saturated rock K , shear modulus of fully saturated rock μ , density ρ , permeability κ , and water saturation S_w . Saturation is selected since there is oil and water in the pore space. The properties of water and shale oil can be obtained from the literature (e.g., Fu et al., 2020). Permeability κ can be estimated by the Kozeny-Carman equations (Carman, 1961) or calculated based on pore network model (Bernabé et al., 2011, 2016) in the absence of actual measurements. The stiffnesses of the rock cannot be measured directly in logging data, but can be estimated from the P and S slownesses and density. The bulk and shear moduli of the mineral, K_s and μ_s , can be inferred from K and μ based on Gassmann's equation (Gassmann, 1951), where the bulk and shear moduli of the skeleton, K_b and μ_b , are estimated from the Krief model (Krief et al., 1990), and it is assumed that μ equals μ_b .

Note that K , μ , and κ are dependent on ϕ in general, thus by adding these three quantities, the influence of porosity is enhanced. Other quantities, such as depth, gamma ray, and resistivity are not selected, since they are not used in classical elastic wave-propagation models. Moreover, in DNN training, more quantities may lead to overfitting. The seismic properties include the velocities of the P - (V_p) and S - (V_s) waves (labels), commonly available data from downhole measurements. In addition, the corresponding quality factors can be obtained by using the median

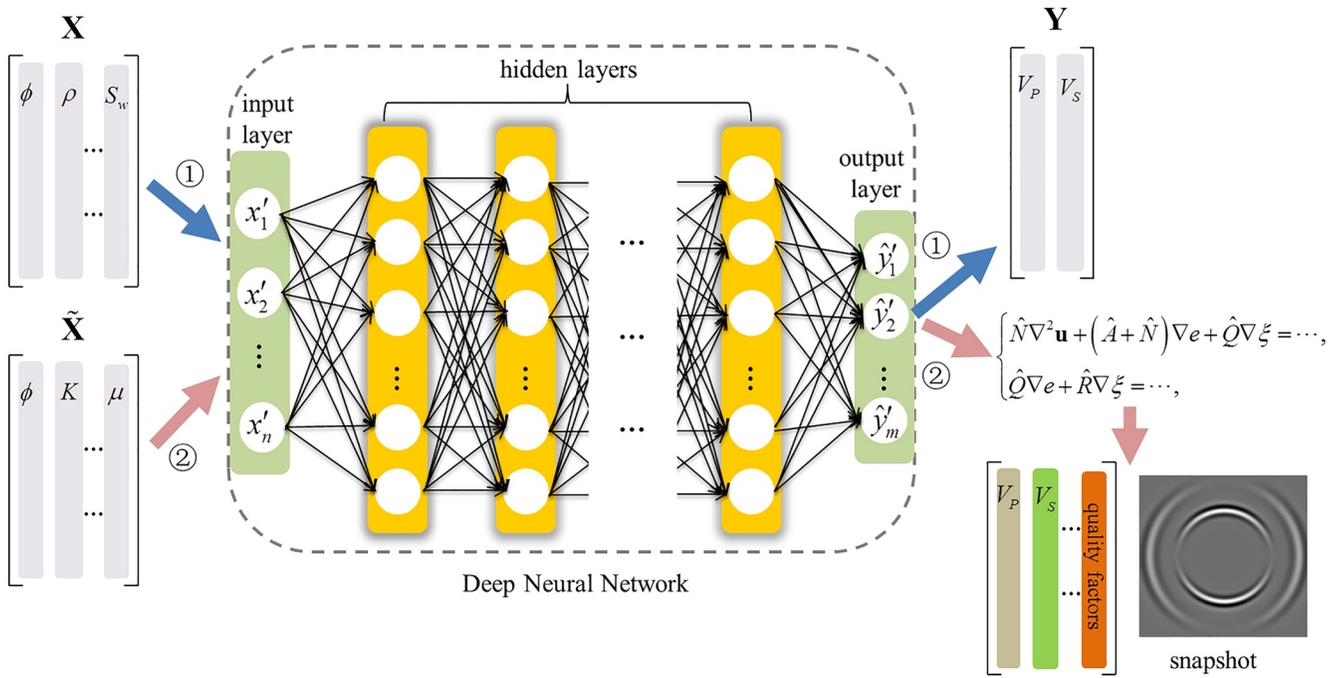


Figure 2. Diagram of the deep neural network as a surrogate model, where ① and ② correspond to the method described in Sections 2 and 3, respectively.

frequency-shift method (Frazer et al., 1997; Qi et al., 2021; X. Sun et al., 2000) if the conditions of the data allow it. To summarize, the inputs to DNN, \mathbf{X} , are ϕ , K , μ , ρ , κ , and S_w , and the outputs, \mathbf{Y} , are V_p and V_s .

2.2. DNN-Based Surrogate Model

DNN can process nonlinear dependencies between different data sets. Its performance is severely affected by hyperparameters (variables that need to be determined a priori), which determine the structure of the network and how this is trained. How to determine these optimal parameters remains an active area of research but falls outside the scope of this study. Here, we do it by trial and error, requiring the highest test accuracy. It is found the optimal DNN has five hidden layers, and each layer 30 units. As shown in Figure 2, in addition to hidden layers, DNNs have input and output layers. The NNs in this study are built based on PyTorch (Paszke et al., 2017). The activation function is the hyperbolic tangent (tanh). The adaptive moment estimation (Adam) optimizer is used (Kingma & Ba, 2014), and the initial learning rate is 0.001. The training process updates the network parameters, that is, weights and bias, to minimize the loss function, which is the mean square error (MSE) between the DNN predictions and the given labels (e.g., Qadrouh et al., 2019).

A normalization of the input quantities is needed before training because the magnitudes differ, affecting the convergence and stability of the training process, and the problem of a vanishing gradient can be improved after data normalization by using tanh. The input of the DNN is the data after processing, where each input, denoted as x_i , is normalized to $[-1, 1]$ as:

$$x'_i = \frac{2}{\max\{x_i\} - \min\{x_i\}}(x_i - \min\{x_i\}) - 1, \quad (1)$$

where $i = 1, 2, \dots, 6$ corresponds to the six quantities, respectively.

There is the same problem with \mathbf{Y} and normalization is required. Each label (V_p or V_s) is normalized by dividing by 1,000. This is due to the fact that \mathbf{Y} is unknown when the DNN is used to predict the seismic properties. The real input and output of the DNN model are \mathbf{X}' and $\hat{\mathbf{Y}}'$, respectively, and $\hat{\mathbf{Y}}'$ is re-scaled to the physical range via denormalization. Alternatively, one can set weights for each \mathbf{Y} component when defining the loss function.

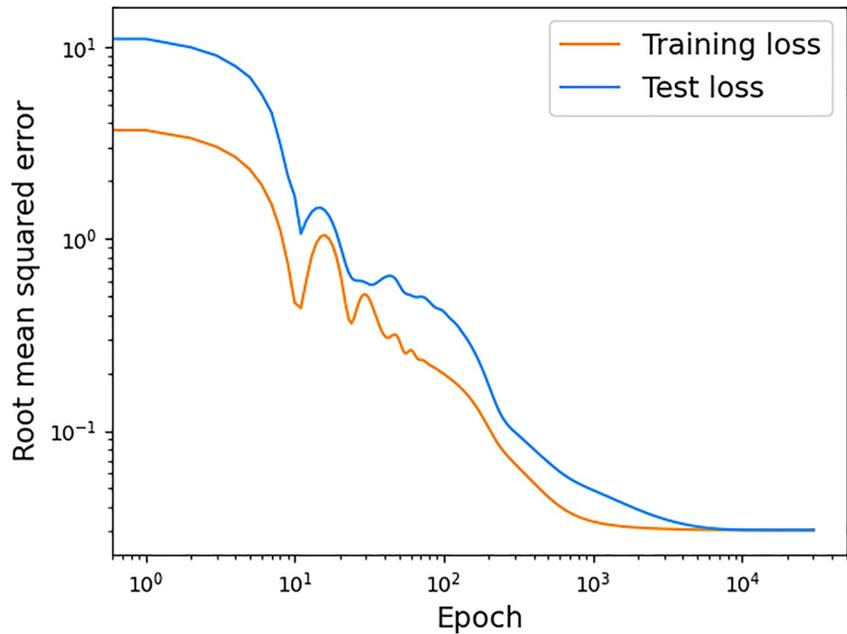


Figure 3. Loss curve of a DNN model. Training and test loss are denoted with the orange and blue solid curves, respectively.

The six input quantities are selected in this section, each having a different impact on V_p and V_s . The sensitivity of the seismic properties to each rock-physics parameter can be analyzed once the DNN-based surrogate model is obtained. By taking V_p as an example, the sensitivity function χ_i vs. x_i can be defined as (Hamby, 1994):

$$\chi_i = \left| \frac{\partial \hat{V}_p}{\partial x_i} \cdot \frac{x_i}{\hat{V}_p} \right|, \quad (2)$$

where $i = 1, 2, \dots, 6$. Equation 2 can be computed by using the automatic differential mechanism (Paszke et al., 2017).

3. Data-Driven Design of Wave Equations

To build the wave equations, their form and related coefficients are to be determined. A similar DNN-based method can be used to learn the new model from data.

3.1. DNN Approach of Unknown Constants

To simplify the problem, it is assumed that the form of the wave equations is consistent with that of Biot's theory (Carcione, 2014), while the elastic coefficients are unknown. The wave equations to be learned can be written as:

$$\begin{cases} N \nabla^2 \mathbf{u} + (A + N) \nabla e + Q \nabla \xi = \rho_{11} \ddot{\mathbf{u}} + \rho_{12} \ddot{\mathbf{U}} + b \cdot (\dot{\mathbf{u}} - \dot{\mathbf{U}}), \\ Q \nabla e + R \nabla \xi = \rho_{12} \ddot{\mathbf{u}} + \rho_{22} \ddot{\mathbf{U}} - b \cdot (\dot{\mathbf{u}} - \dot{\mathbf{U}}), \end{cases} \quad (3)$$

where \mathbf{u} and \mathbf{U} are the basic unknowns that represent the displacements of the solid and fluid phases, respectively, $e = \nabla \cdot \mathbf{u}$ and $\xi = \nabla \cdot \mathbf{U}$ are the corresponding dilatations, ∇ is the gradient operator and a dot above each quantity denotes the time derivative. ρ_{11} , ρ_{12} , and ρ_{22} are density constants, and b is the dissipation coefficient. Biot's theory is applied to determine ρ_{11} , ρ_{12} , ρ_{22} , and b (See Appendix A). It is assumed that the elastic constants A , N , Q , and R in Equation 3 are unknown, and DNNs are used to estimate them from data. In this case, the coefficients are dependent on ϕ , K_s , μ_s , and K_f according to Biot's theory, where K_f is the fluid bulk modulus. The input of

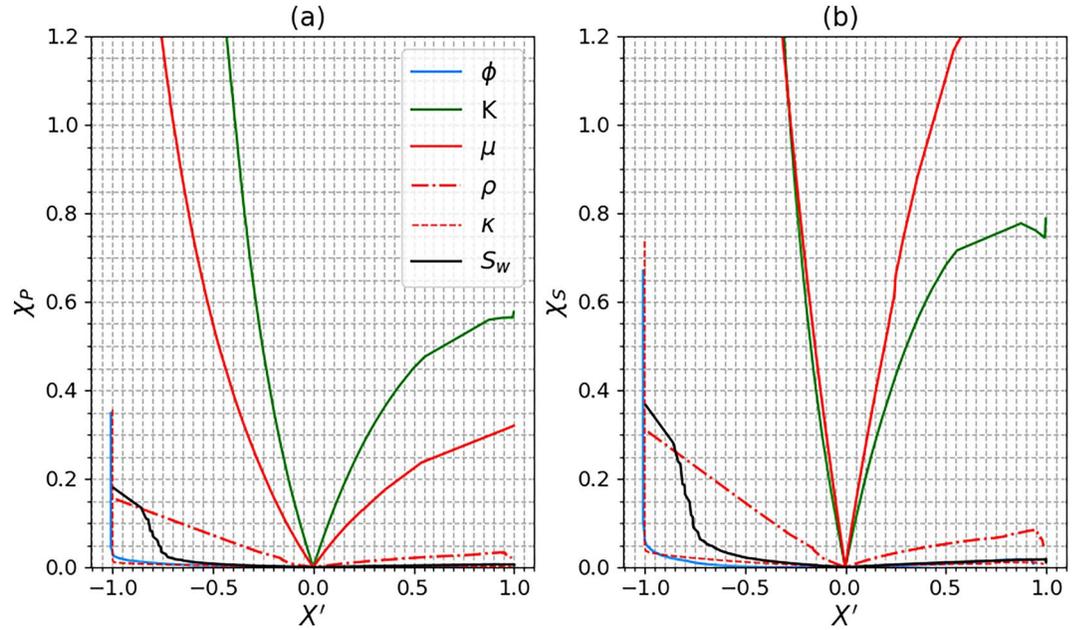


Figure 4. Sensitivity curves by a well-trained DNN model: (a and b) respectively show the sensitivity of V_p and V_s to each input quantity. The abscissa X' represents each normalized quantity.

the DNN model $\tilde{\mathbf{X}}$ consists of ϕ , K , μ , and S_w , where S_w is used instead of the bulk modulus of the fluid mixture. DNN models are introduced to obtain the elastic constants as:

$$\begin{aligned} A(\tilde{\mathbf{X}}) &\approx \hat{A}(\tilde{\mathbf{X}}; \theta_1) = \text{DNN}(\tilde{\mathbf{X}}; \theta_1), & N(\tilde{\mathbf{X}}) &\approx \hat{N}(\tilde{\mathbf{X}}; \theta_2) = \text{DNN}(\tilde{\mathbf{X}}; \theta_2), \\ Q(\tilde{\mathbf{X}}) &\approx \hat{Q}(\tilde{\mathbf{X}}; \theta_3) = \text{DNN}(\tilde{\mathbf{X}}; \theta_3), & R(\tilde{\mathbf{X}}) &\approx \hat{R}(\tilde{\mathbf{X}}; \theta_4) = \text{DNN}(\tilde{\mathbf{X}}; \theta_4), \end{aligned} \quad (4)$$

where θ_1 , θ_2 , θ_3 , and θ_4 are neural-network parameters, obtained by minimizing the loss function. The predictions of V_p and V_s have to be computed by using Equation 3 to build the loss function. A plane-wave analysis (Carrione, 2014) is used to obtain V_p and V_s (See Appendix A). Then we have:

$$\begin{aligned} V_p &\approx \hat{V}_p(\mathbf{X}; \theta), & V_s &\approx \hat{V}_s(\mathbf{X}; \theta), \\ Q_p &\approx \hat{Q}_p(\mathbf{X}; \theta), & Q_s &\approx \hat{Q}_s(\mathbf{X}; \theta), \end{aligned} \quad (5)$$

where $\theta = [\theta_1, \theta_2, \theta_3, \theta_4]$.

With these equations, the loss function can be defined.

3.2. Loss Function and DNN Training

The loss function can be defined as MSE:

$$L_1 = \sum_{i=1}^M |V_{p,i} - \hat{V}_{p,i}(\mathbf{X}; \theta)|^2 + \sum_{i=1}^M |V_{s,i} - \hat{V}_{s,i}(\mathbf{X}; \theta)|^2, \quad (6)$$

where $V_{p,i}$ and $V_{s,i}$ are the i th data point of wave velocities from logs as the ground truth, and M is the number of data points used for training. The quality factor is not incorporated in Equation 6 since there is no available real data. In most cases, the more information, that is, constraints, can be added to the loss function, the more likely the learned model will approach the reality. However, most ML algorithms are sensitive to the errors and noises of the data.

Table 1
The Precision of Predictions by the Two Methods Validated on the Test Set

Well	Attribute	Method in Section 2		Method in Section 3	
		Relative RMSE	r^2	Relative RMSE	r^2
W1	V_p	0.4831%	99.64%	1.649%	95.77%
	V_s	0.4498%	99.78%	1.910%	96.04%
W2	V_p	0.2432%	99.88%	1.713%	94.34%
	V_s	0.2540%	99.92%	1.288%	97.92%
W3	V_p	0.3840%	99.79%	1.836%	95.22%
	V_s	0.3737%	99.85%	2.298%	92.28%
W4	V_p	0.2656%	99.78%	1.273%	94.87%
	V_s	0.2599%	99.86%	1.212%	97.07%
W5	V_p	0.4992%	99.33%	1.805%	91.24%
	V_s	0.4611%	99.66%	1.911%	94.20%

Meanwhile, physical constraints need to be added into the training process to assure that the learned model is consistent with the real data and physics. According to Biot's theory, the following constraints are proposed: $|\hat{A}| > |\hat{Q}| > |\hat{R}| > 0$, based on the definitions of the coefficients. The loss function with physical constraints can be written as:

$$L_2 = \sum_{i=1}^M |\text{ReLU}(|\hat{Q}| - |\hat{A}|)|^2 + \sum_{i=1}^M |\text{ReLU}(|\hat{R}| - |\hat{A}|)|^2 + \sum_{i=1}^M |\text{ReLU}(|\hat{R}| - |\hat{Q}|)|^2, \quad (7)$$

where ReLU is the activation function with the full name of rectified linear unit. The Softplus function is used in the output layer to ensure its positivity.

Consequently, the total loss function is defined as:

$$\text{Loss} = \lambda_1 L_1 + \lambda_2 L_2, \quad (8)$$

where λ_1 and λ_2 are weight factors, which we set to 1.

The design of the DNN is similar to that of Section 2. A fully connected NN is used here, with five hidden layers, each layer with 50 units. The Adam optimizer is used and ReLU is selected as the activation function. Input normalization and label denormalization are not needed. The DNN models are trained simultaneously by minimizing the total loss function. The training and test sets are the same as in Section 2. Only after testing can the practicability be proved.

The workflow of the method presented in this section is summarized as follows. First, the density constants (ρ_{11} , ρ_{12} , and ρ_{22}) and dissipation coefficient (b) of Equation 3 are computed according to Biot's theory. Second, the DNN are fed with $\tilde{\mathbf{X}}$ to predict the elastic constants (A , N , Q , and R). Once a preliminary prediction is achieved, wave velocities (V_p and V_s) are computed by Equations A3 and A4. Then the loss function of data mismatch between predictions and the given log data along with the constraints can be formulated. Then follows the train of the DNN models and the updating to minimize the loss function.

4. Results and Discussions

We consider examples to test the methods presented in Sections 2 and 3. Two indicators for evaluating the precision of the DNN models are given. By taking V_p as an example, the first one is the root MSE:

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^M (V_{p,i} - \hat{V}_{p,i})^2}, \quad (9)$$

where $V_{p,i}$ is the ground truth of data point i , $\hat{V}_{p,i}$ is the prediction by the DNN model, and M is the number of data points used to test here. In an actual application, it is usually divided by the mean value of V_p to get the relative RMSE.

The other indicator is the so-called coefficient of determination:

$$r^2 = 1 - \frac{\sum_{i=1}^M (V_{p,i} - \hat{V}_{p,i})^2}{\sum_{i=1}^M (V_{p,i} - \bar{V}_{p,i})^2}, \quad (10)$$

where $\bar{V}_{p,i} = \sum_{i=1}^M V_{p,i} / M$. The closer the prediction from the DNN model is to the ground truth, the closer r^2 is to one.

4.1. Results With the DNN-Based Surrogate Model

A DNN model that serves as a surrogate one to match the relationship between \mathbf{X} and \mathbf{Y} is trained. Data normalization is performed before training and then the data is scaled back to the normal range after training. Data of five wells are selected as test set. The loss values in the process of training are shown in Figure 3. The whole test set is taken as the validation set. As can be seen, the training and test loss decrease and converge to a low value after

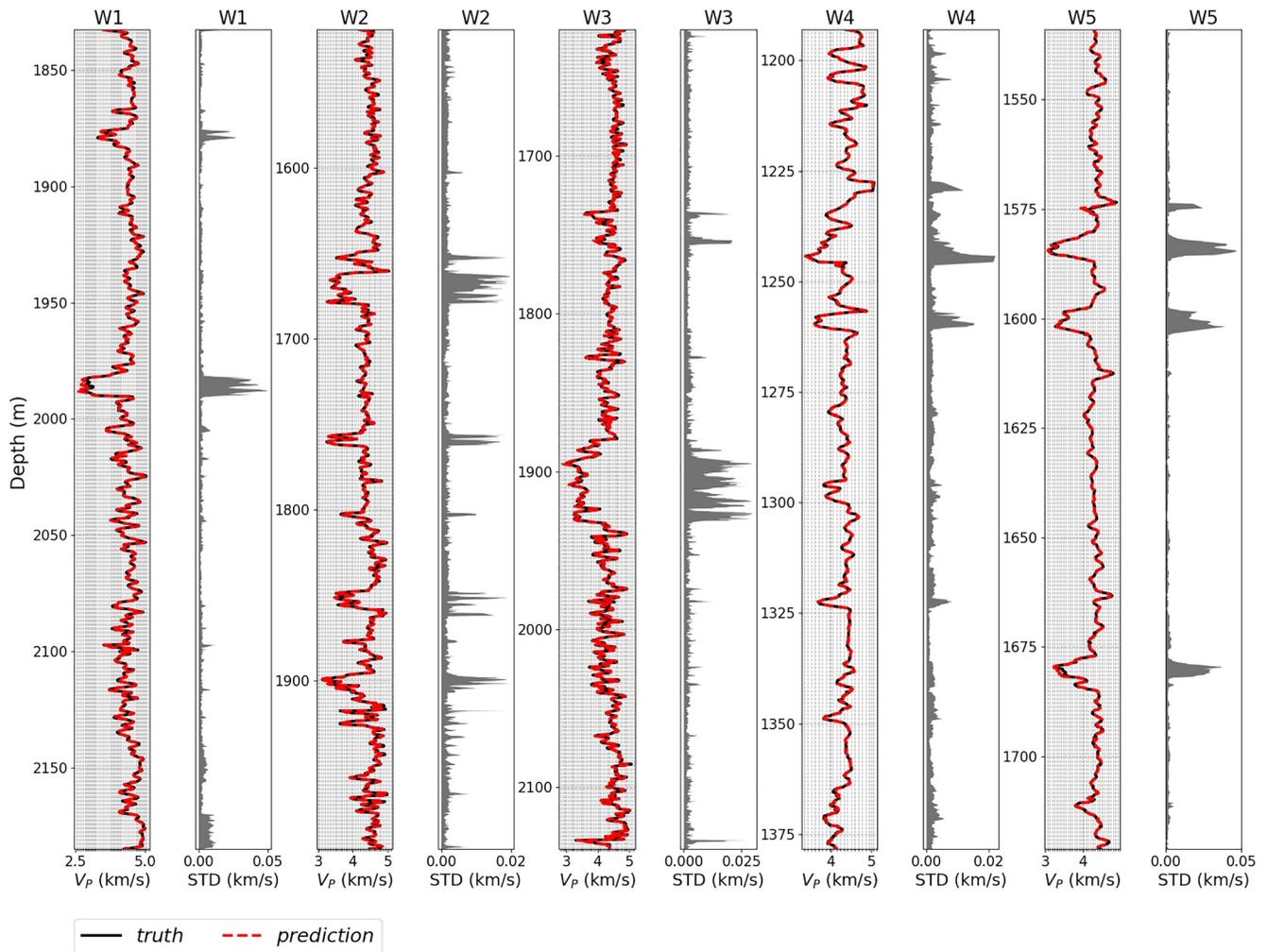


Figure 5. Mean and standard deviation of V_p from 15 DNN-based surrogate models. Predictions are validated on test data from W1 to W5. The black solid curves represent the ground truth, the red dashed curves are the mean predictions from 15 different DNN initializations, and the gray shadow represents the standard deviations as a degree of uncertainty.

6,000 epochs. The final test loss is the same as the convergence value of the training loss, which indicates that no overfitting occurs and the generalization capability of the trained DNN model can be guaranteed. High precision of predictions on the test set is obtained with relative RMSEs for predictions of V_p and V_s equal to 0.4325% and 0.4407%, and the corresponding r^2 values are 99.12% and 99.51%, respectively.

The sensitivity of V_p and V_s to each component of \mathbf{X} is investigated. The sensitivity curves are shown in Figure 4, where X' denotes each normalized component of the training set, and χ_p and χ_s are computed by using the well-trained DNN model according to Equation 2. Figure 4a shows that the bulk modulus K (green curve) is the most sensitive property to V_p , that is, a small variation of K may cause a significant change in V_p . The shear modulus μ also affects V_p .

The influence of porosity ϕ and permeability κ on V_p is significant only when it is very small. The trends of the two curves are consistent, indicating that there is a correlation. The effect of density ρ is more significant than κ and S_w in general. It should be noted that ϕ may be the most important parameter since K , μ , and κ are dependent on ϕ . The effect of ϕ is also included in the three quantities. By comparison, the water saturation S_w is the least sensitive.

Unlike the P wave, as can be seen in Figure 4b, μ affects more V_s as expected. K is also affected, due to the effect of porosity ϕ . The sensitivity curves of the other quantities are similar to Figure 4a. From the perspective of

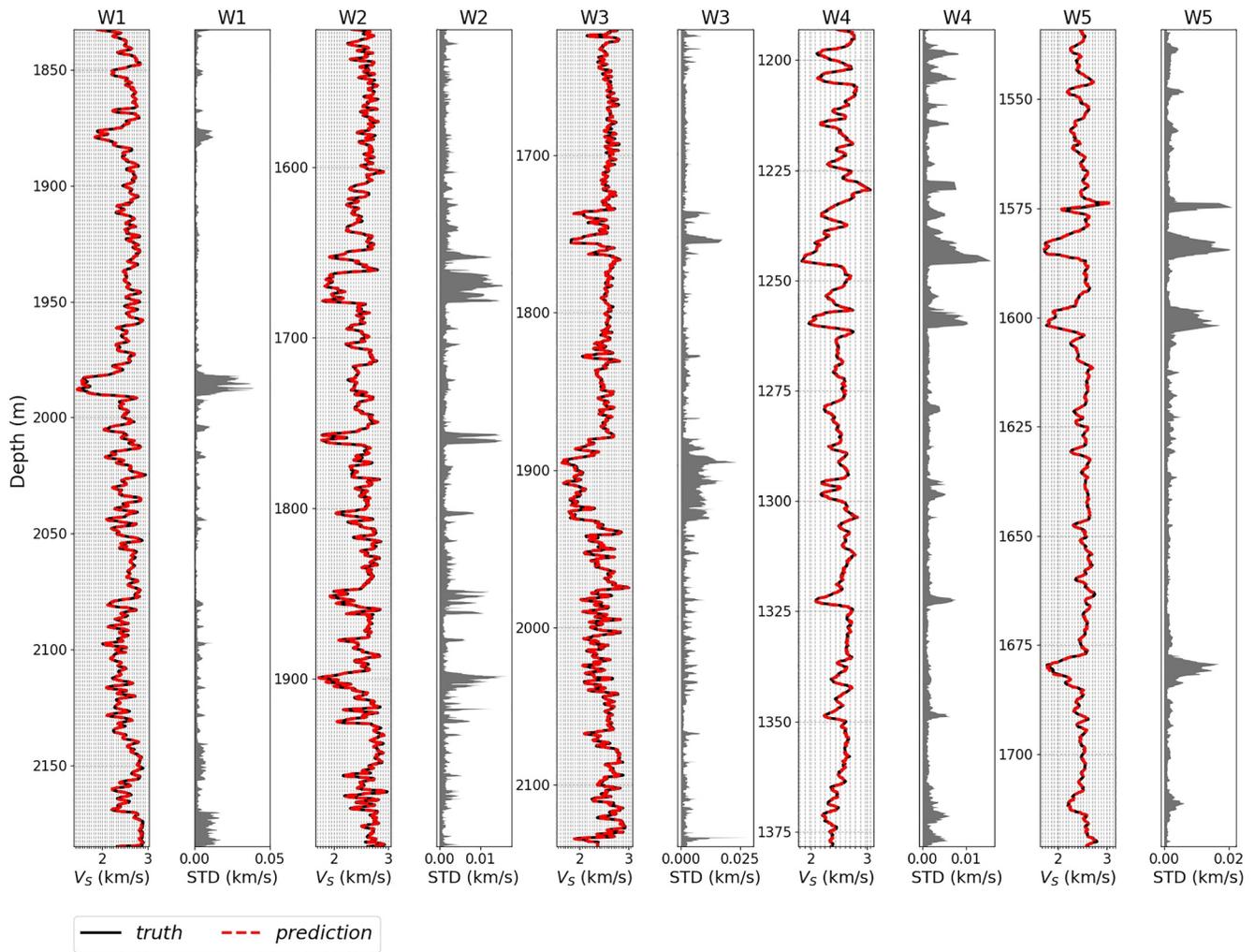


Figure 6. Mean and standard deviation of V_s from 15 DNN-based surrogate models. Predictions are validated on test data from W1 to W5. The black solid curves represent the ground truth, the red dashed curves are the mean predictions from 15 different DNN initializations, and the gray shadow represents the standard deviations as a degree of uncertainty.

multivariate functions, the influence of each variable on the dependent variables is obtained according to the sensitivity analysis. It suggests that the insensitive quantities can be neglected to achieve a dimensionality reduction in establishing the wave-propagation model. Regarding the log data, the sensitivity analysis shows that S_w and κ have a weak influence on V_p and V_s . Therefore, data points that lack these two parameters will still be used for testing, where a mean value can be taken. In this way, the number of test data points increases.

Next, we analyze the effect of the DNN initialization, as an uncertainty quantification of the DNN-based surrogate method, on the precisions of V_p and V_s . In this example, the 15 DNN models with different initializations but identical hyper parameters as described in Section 2 are trained. Specifically, the weights of the DNN models are initialized by feeding random values from the standard normal distribution. As a result, 15 well-trained DNN models are obtained and then validated by the five wells (denoted as W1–W5). For each input, mean values of outputs from these DNN models are calculated as predictions, and the standard deviation (STD) is obtained simultaneously.

Table 1 gives the precision of the prediction for each well. As can be seen the DNN results match well the ground truth. Figures 5 and 6 show the predictions with STDs for V_p and V_s from W1 to W5, respectively. The gray shadow shows degrees of prediction uncertainty, where the larger the STD, the higher the uncertainty. Overall, the trained DNN models perform well on the five wells, suggesting that the surrogate model is determined

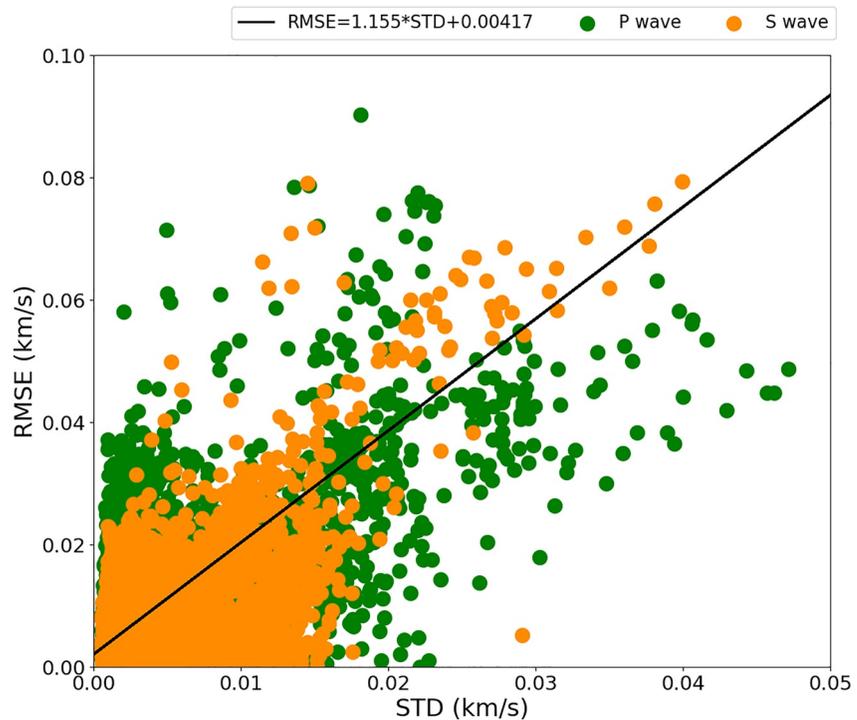


Figure 7. Scatter plot of the root mean squared error (RMSE) vs. the standard deviation (STD) of data from wells W1 to W5 based on DNN-based surrogate models. The black solid line represents the fitting according to the linear regression method.

successfully. It means that wave velocity prediction with a high precision can be achieved when the rock-physics parameters are obtained.

Note that the STDs of data points with a high precision are small, while the uncertainty is high for the test data such as the data points from depths of 1,980 to 1,990 m in W1 (where a relatively poor prediction is given as shown in Figure 5). There is a correlation between the prediction errors and the STDs. Furthermore, Figure 7 gives the scatter plot of RMSE vs. STD of all the test data. It is shown that most of the V_p - and V_s -related points overlap, and the prediction precision of V_p is slightly lower than that of V_s . There is some correlation between them, which can be approximated with a straight line. On the other hand, for both V_p and V_s , the mean STD is 20 times smaller than their mean value, indicating that the effect of the DNN initialization on the prediction of V_p and V_s is small, and each DNN model can output high-precision predictions.

4.2. Results With Data-Driven Design of Wave Equations

The aforementioned examples show the DNN-based surrogate model established with log data. Predictions of V_p and V_s with a high precision can be obtained once six input quantities are fed into the well-trained DNN models. However, such a model can only be used to link the rock-physics parameters with V_p and V_s , but further information other than these quantities cannot be obtained. For instance, the quality factor cannot be inferred and wavefield simulation cannot be performed. Here, DNN models are trained to predict the four elastic constants to match V_p and V_s . In this way, the wave equations for shale-oil reservoirs are obtained by feeding $\tilde{\mathbf{X}}$ into the well-trained DNN models to get these four constants, along with others coefficients which are determined according to Biot's theory.

Similar to the previous examples, 15 DNN models with different initializations are trained here, and the mean value of predictions of the learned wave equations based on the plane-wave analysis is taken as the result. See Table 1 for the precision. Although it is slightly lower than the previous method, the DNN-predicted results match well with the ground truth in general. Figures 8 and 9 show the predicted results and the STDs for V_p and V_s , respectively. The learned wave equations achieve high precision predictions on the data from wells W1 to W5.

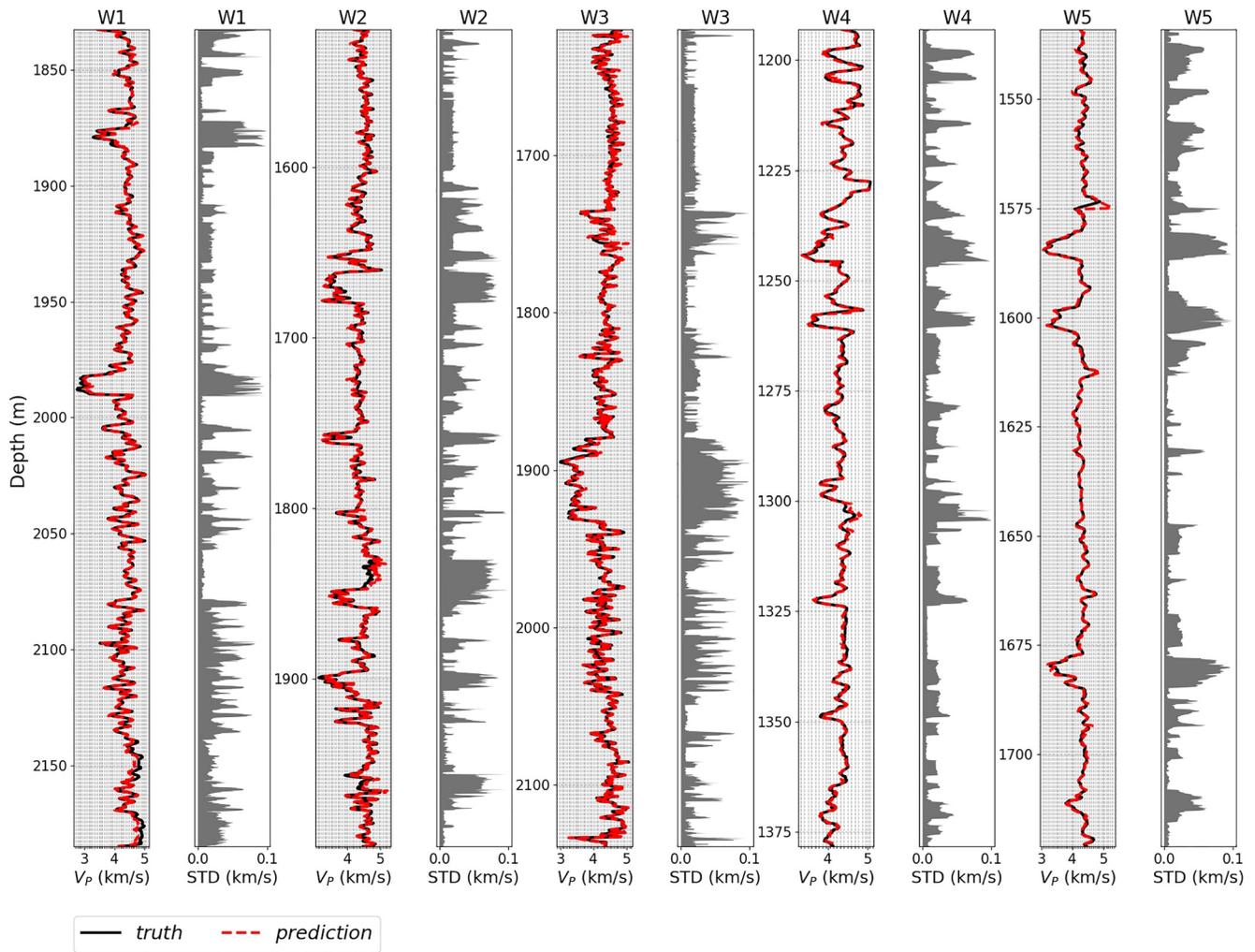


Figure 8. Mean and standard deviation of V_p from 15 learned wave equations. Predictions are validated on test data from wells W1 to W5. The black solid curves represent the ground truth, the red dashed curves are the mean predicted results from 15 learned wave equations, where the DNN training includes 15 initializations, and the gray shadow represents the standard deviation as a degree of uncertainty.

In addition, it is shown that there is a certain correlation between the prediction error and the uncertainty. For example, as shown in Figure 8, the STD value corresponding to the data of well W1 at depths around 1,850 m, is relatively large where the prediction is poor. Figure 10 gives the scatter plot of the RMSE vs. STD of all the test data. Most of the data points are concentrated in an area, indicating that there is a correlation between RMSE and STD. The prediction precision of V_p is almost the same as that of V_s but with a relatively higher uncertainty. Similar to Figure 7, the mean STD is 20 times smaller than the mean value for V_p and V_s , which indicates that the effect of the DNN initialization on the prediction of V_p and V_s is small.

In the examples, the wave velocities are used for training, while the quality factors of P and S waves can be predicted from Equation A4 without any other further information since the wave equations are obtained. However, the prediction cannot be tested since there is no measurement available. In order to explore if the learned wave equations can predict the unknown quantities under the condition of limited and zero S wave labels are available in training, a special example is given here to investigate whether V_s can be predicted correctly. Here, the data of V_p is used entirely in the process of DNN training, while different V_s values are used to train the DNN models whose outputs are the elastic coefficients of the wave equations.

Table 2 lists the precisions of predictions of V_p and V_s by the learned wave equations in the cases that five and no values of V_s are used. Compared with Table 1, the precisions of V_p and V_s decreases, which is not surprising

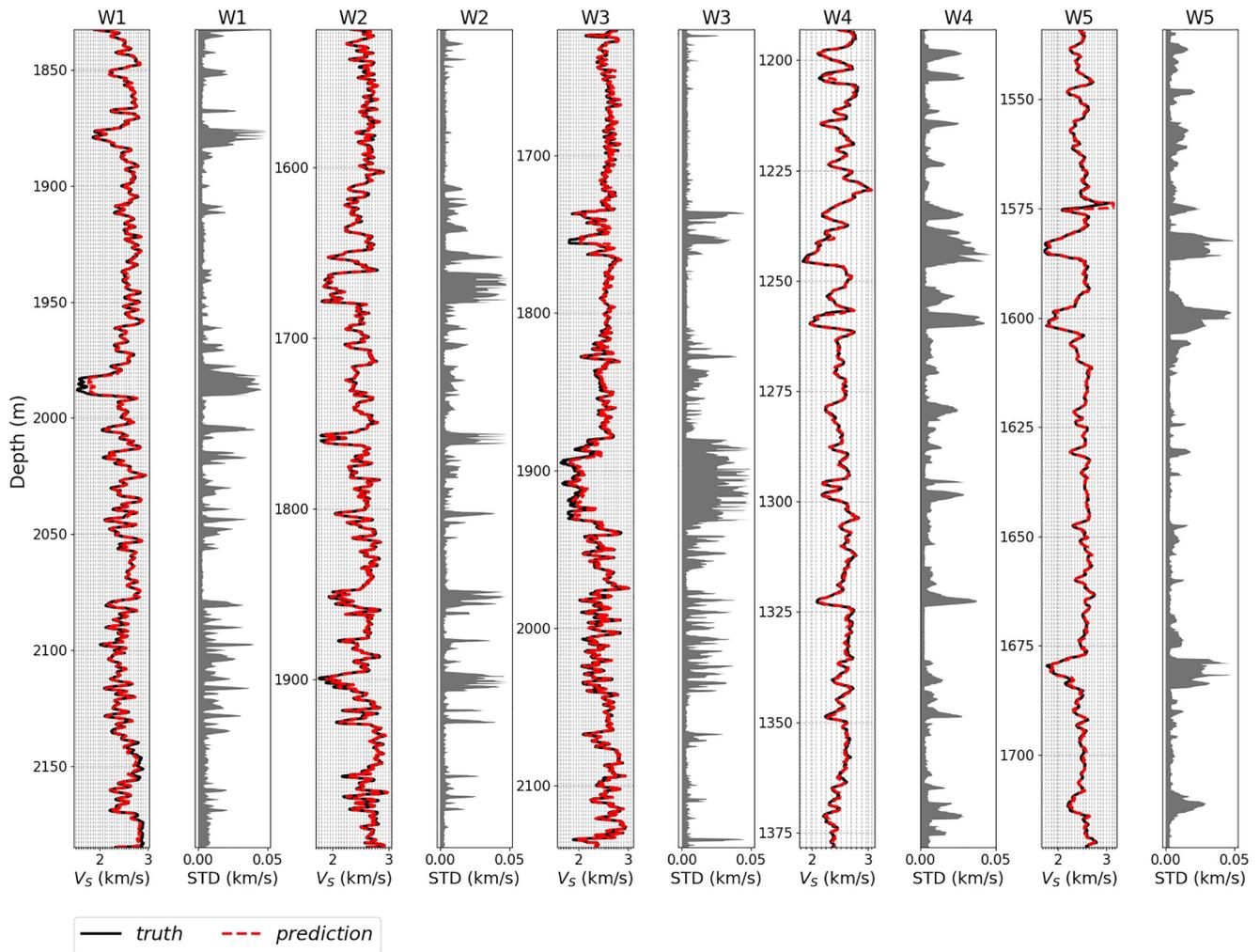


Figure 9. Mean and standard deviation of V_s from 15 learned wave equations. Predictions are validated on test data from wells W1 to W5. The black solid curves represent the ground truth, the red dashed curves are the mean predicted results from 15 learned wave equations where the DNN training include 15 initializations, and the gray shadow represents the standard deviation as a degree of uncertainty.

since the constraints on the DNN model are less. The precision of V_p is lower than that of the previous examples. Note that V_s is only related to the constant μ as can be seen in Equation A3, while V_p is affected by the four elastic constants. The fewer labels are given, the lower the prediction precisions of the four constants. Figure 11 shows the predictions for V_p and V_s for two special cases, in which data from wells W1 to W5 are used, respectively. It can be seen that the learned wave equations still successfully predict V_p and V_s , indicated as prediction 1 (blue dashed curve) in the case that five values of V_s are used for training. In the case of zero values (prediction 2, red dashed curve), although the precision is greatly reduced, the errors of the relative RMSE are less than 5%, which may be considered acceptable.

This example shows the advantages of the second method, that is, more physical quantities can be obtained once the wave equations are learned from data. Specifically, a relatively ideal prediction can be obtained even in case of zero values (labels). On the other hand, once the label of a quantity is available, even for a few points, the prediction precision can be improved.

Finally, the results are compared with those predicted by the classical rock physics model. The Xu and White (1995) one is used to predict V_s . This model assumes that the geometry of the pores of the sandy part is significantly different from that of clays and shales. The Kuster and Toksöz (1974) and differential-effective-medium theories

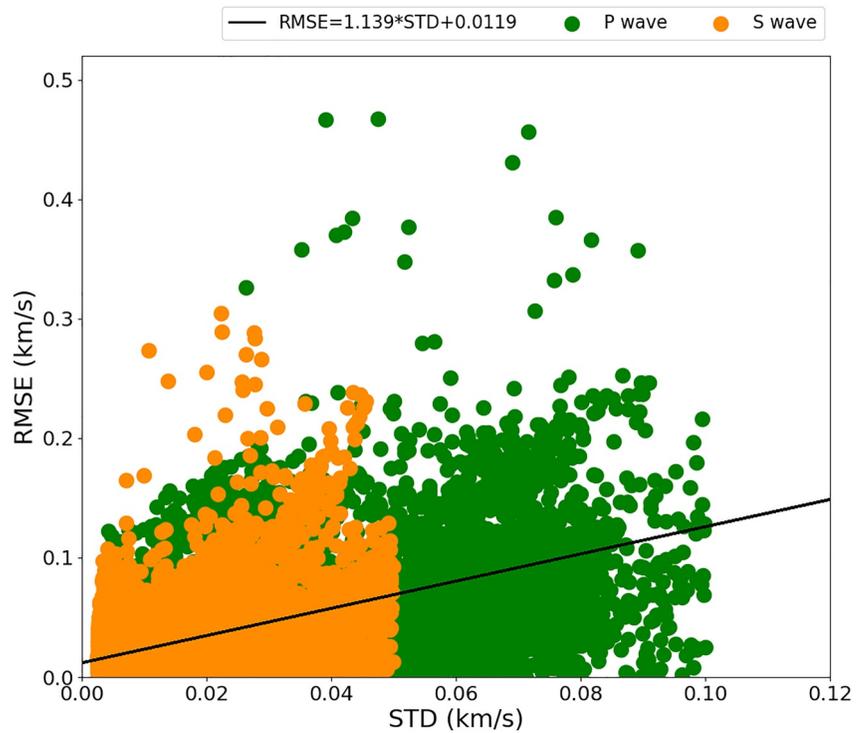


Figure 10. Scatter plot of RMSE vs. STD of data from wells W1 to W5 based on data-driven building of wave equations. The black solid line represents the fitting according to the linear regression method.

are employed to compute the dry-rock bulk and shear moduli. Then, the wet-rock bulk modulus K_{sat} is determined by using the Gassmann equation (Gassmann, 1951). The wave velocities are determined as:

$$V_P = \sqrt{\left(K_{sat} + \frac{4}{3}\mu_{sat}\right) / \rho}, \tag{11}$$

$$V_S = \sqrt{\mu_d / \rho}, \tag{12}$$

where μ_d is the shear modulus of the dry rock.

The porosity, clay content, and the density of rock needed in modeling are taken from well-log data. The bulk moduli of sand and clay's matrix are taken as 37 and 21 GPa, respectively, and the corresponding shear moduli are 44 and 7 GPa, respectively. The bulk moduli of water and oil are taken as 2.25 and 1.02 GPa, respectively. Other parameters of fluids are taken as the same values as the aforementioned analysis. The aspect ratio of sand-related pores is taken as 0.12. V_S can be predicted by using Equation 12.

Figure 12 shows the results for V_S , compared to those of the Xu-White model. For wells W1 and W3, there is a relatively good agreement between the curves, while the Xu-White model predicts lower velocities for wells W4 and W5. The relative error of these predictions for each well is higher than 5%. It shows that this classical rock-physics model may find it difficult to estimate the wave velocities of the shale reservoirs. Moreover, the rock-physics model mainly predicts the wave velocities, and ML can obtain more information. It is shown in Figure 9 that high precision results can be obtained according to the second ML-based method.

Table 2
Precisions of Wave Equations Learned in the Case of Five and Zero Values of V_S

Well	Attribute	Training with five values of V_S		Training without V_S	
		Relative RMSE	r^2	Relative RMSE	r^2
W1	V_P	2.179%	92.60%	1.462%	96.67%
	V_S	1.672%	96.97%	3.284%	88.30%
W2	V_P	1.530%	95.49%	2.001%	92.22%
	V_S	1.208%	98.17%	2.640%	91.28%
W3	V_P	2.613%	90.32%	1.538%	96.65%
	V_S	2.044%	95.41%	4.220%	80.46%
W4	V_P	1.360%	94.14%	1.942%	88.05%
	V_S	1.220%	97.04%	2.828%	84.08%
W5	V_P	2.220%	86.75%	1.909%	90.20%
	V_S	1.602%	95.92%	3.355%	82.12%

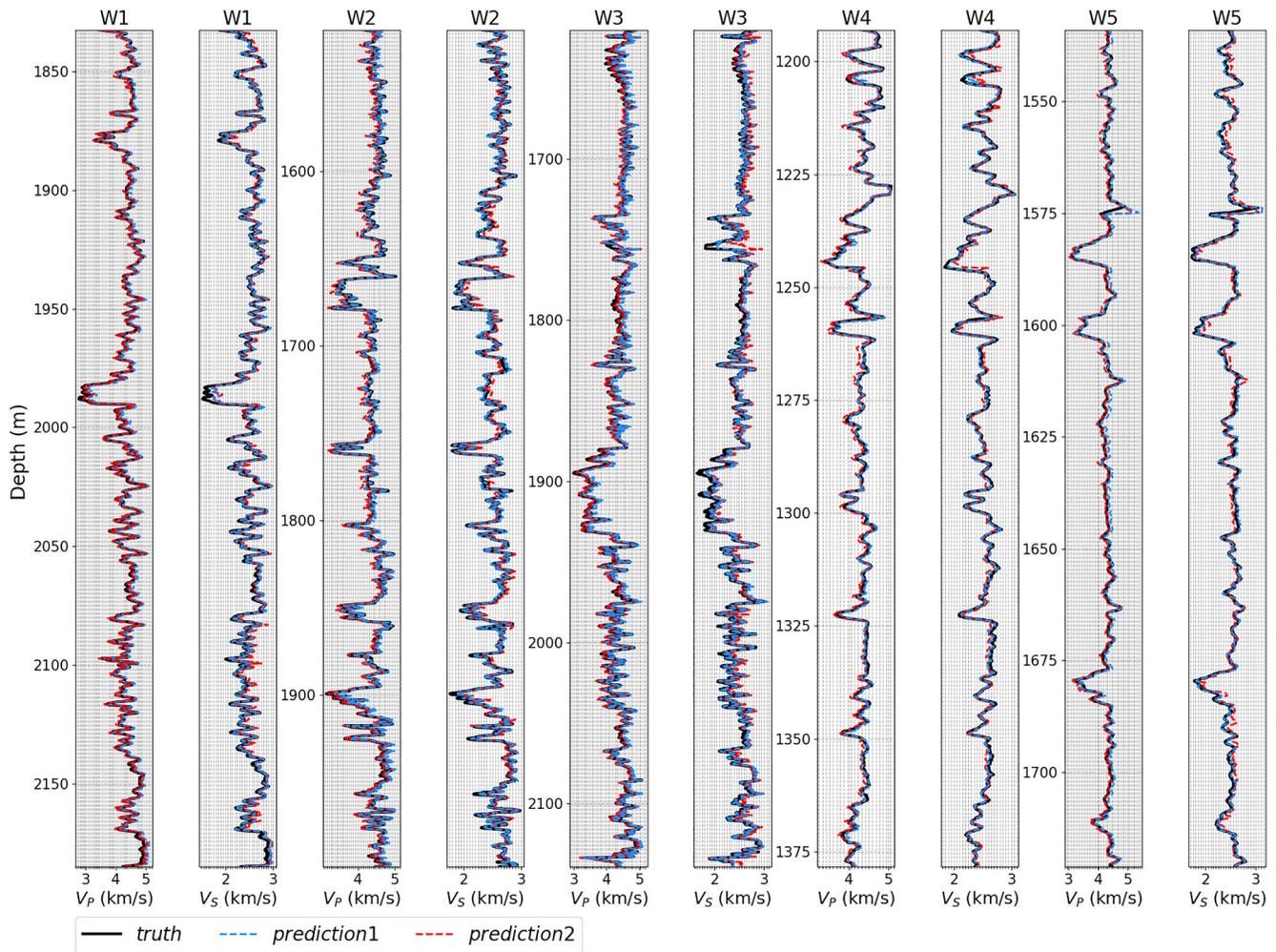


Figure 11. Predictions of V_p and V_s by ML, validated on data from wells W1 to W5. The black solid curves are the ground truth, and the blue (prediction 1) and red (prediction 2) dashed curves represent the case of five and zero values of V_s used in the process of DNN training.

4.3. Generalizability of the Proposed Method

To further show the general aspects of the proposed method, log data from a carbonate reservoir is selected to train the networks. The data is from the Shunbei Oilfield, Tarim Basin, China. Among them, 2,534 data points include porosity ϕ , rock density ρ , P wave velocity V_p and S wave velocity V_s . The bulk and shear moduli K and μ can be obtained based on the wave velocities. In addition, the density, bulk modulus, and viscosity of the fluid are determined with Batzle-Wang equations (Batzle & Wang, 1992) according to the in situ temperature and pressure. The average permeability is 1.703 mD. In fact, the parameters of fluid do not affect the accuracy of the results since constant values are used.

Figure 13 compares the true values with the ML predictions, where 25% of data with 634 points of V_p and V_s are selected as training set and the whole data as test set. It can be seen that the match is satisfactory. The overall accuracy is high with relative RMSE for V_p and V_s of 1.90% and 1.66%, and corresponding r^2 values of 96.66% and 95.87%, respectively, indicating that a good accuracy can be obtained with relatively fewer training data. This example also shows that the wave-propagation phenomenon corresponding to these log data can be described with Equation 3 as long as the elastic constants are determined properly.

Similar to Figure 11, an example is given to investigate whether ML can predict V_s accurately under the condition of limited and zero labels. Figure 14 shows the results by using different training sets. In Figures 14a and 14b, 25% of V_p with 634 points are used, but only five points of V_s are selected randomly. The overall accuracy is high

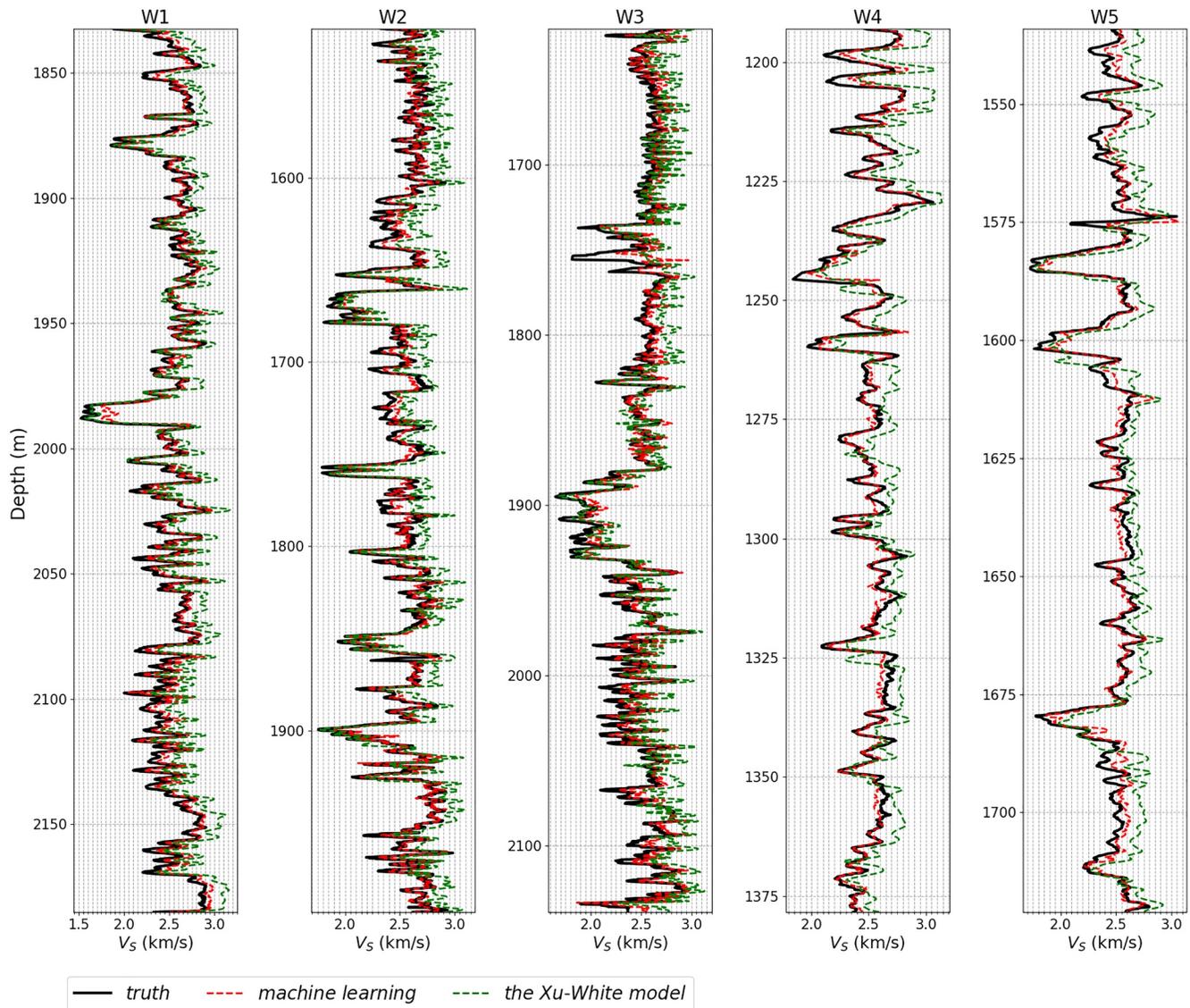


Figure 12. Comparison of V_s obtained with different methods, corresponding to wells W1–W5. The black solid curves represent the ground truth, and the red and green dashed curves the ML predictions and Xu-White model, respectively.

enough with relative RMSE of V_p and V_s of 2.28% and 1.58% and r^2 values of 95.54% and 94.96%, respectively. On the other hand, in Figures 14c and 14d, only 80% of V_p with 2,028 points are used, without V_s . As can be seen that the precision is smaller, although it may be considered acceptable with relative RMSE of V_p and V_s of 1.81% and 3.46% and r^2 values of 96.98% and 82.13%, respectively. Another case, trained only with 25% of V_p without V_s , is omitted since the precision is not good.

The previous examples are mainly focused on a shale reservoir, while the examples in this subsection show the proposed method can accurately be used for carbonate reservoirs. These examples show that wave equations based on ML can yield high precision and general applicability.

5. Conclusions

Two DNN-based methods for designing wave propagation models are proposed. Unlike the classical models, well-log data is required for training. In the first method, the trained DNNs serve as surrogate models to approximate the relation between the rock-physics properties and the seismic attributes. The second method makes use

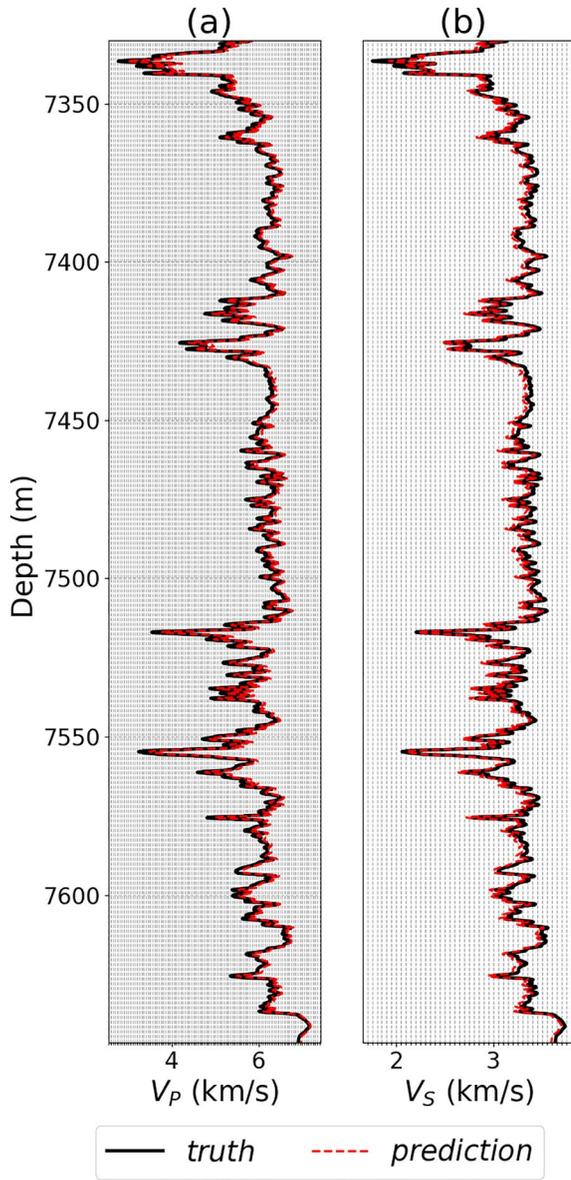


Figure 13. Predictions of V_p and V_s by ML, validated with the whole data set. The black solid curves are the ground truth, and the red dashed curves represent the case of 25% of V_p and V_s used in the process of DNN training.

of the wave equation (PDEs) and measurements to train the DNNs, to obtain the unknown elastic constants of the stress-strain relation. Large training data are needed to ensure the applicability of the methods. The two methods are tested on the data of shale reservoirs and both methods achieve good results. From the point of view of precision, models based on ML have higher accuracy than classical models. The second method has the advantage that additional physical information such as wavefield can be obtained even if some information is not used in the DNN training, which makes full use of the learned model. In addition, tests are performed on the carbonate rocks by using the second method, showing a good performance. Future research based on the second method needs to be performed to investigate if the form of the wave equations can also be obtained from the machine-learning algorithms other than the coefficients of the PDEs. Moreover, including attenuation other properties such as viscosity and permeability could be predicted. This can be done by assuming complex and frequency-dependent the coefficients in Biot equation to effectively describe squirt-flow and mesoscopic dissipation.

Appendix A: Biot's Theory

According to Biot's theory (Biot, 1956a, 1962), the density constants of Equation 3 are $\rho_{12} = [1 - \alpha(1 + 1/\phi)]\phi\rho_f$, $\rho_{11} = (1 - \phi)\rho_s - \rho_{12}$ and $\rho_{22} = \phi\rho_f - \rho_{12}$, where ρ_s and ρ_f are the densities of the solid and fluid, respectively, and $\alpha = 0.5$. ρ_s is inferred from the bulk density $\rho = (1 - \phi)\rho_s + \phi\rho_f = (1 - \phi)\rho_s + \phi\rho_f$. The dissipation coefficient $b = \eta\phi^2/\kappa$, where η is the viscosity of the fluid.

For the case of porous media saturated with two fluids, the effective-fluid theory (Carcione et al., 2006) is used to estimate the properties of the mixture as:

$$\begin{cases} \rho_f = S_w\rho_{f1} + (1 - S_w)\rho_{f2}, \\ \eta = \eta_1(\eta_2/\eta_1)^{S_w}, \end{cases} \quad (\text{A1})$$

where ρ_{f1} and ρ_{f2} are 1 g/cm³ (water) and 0.75 g/m³ (oil), respectively, and η_1 and η_2 are 1 mPa·s and 1.44 mPa·s for shale oil reservoir, respectively (Fu et al., 2020).

In Biot's theory, the expressions of the four elastic constants are given as follows (Biot, 1956a, 1962):

$$\begin{cases} N = \mu, \\ A = \frac{(1 - \phi)(1 - \phi - K_b/K_s)K_s + \phi K_b K_s / K_f}{1 - \phi - K_b/K_s + \phi K_s / K_f} - \frac{2}{3}N, \\ Q = \frac{(1 - \phi - K_b/K_s)\phi K_s}{1 - \phi - K_b/K_s + \phi K_s / K_f}, \\ R = \frac{\phi^2 K_s}{1 - \phi - K_b/K_s + \phi K_s / K_f}, \end{cases} \quad (\text{A2})$$

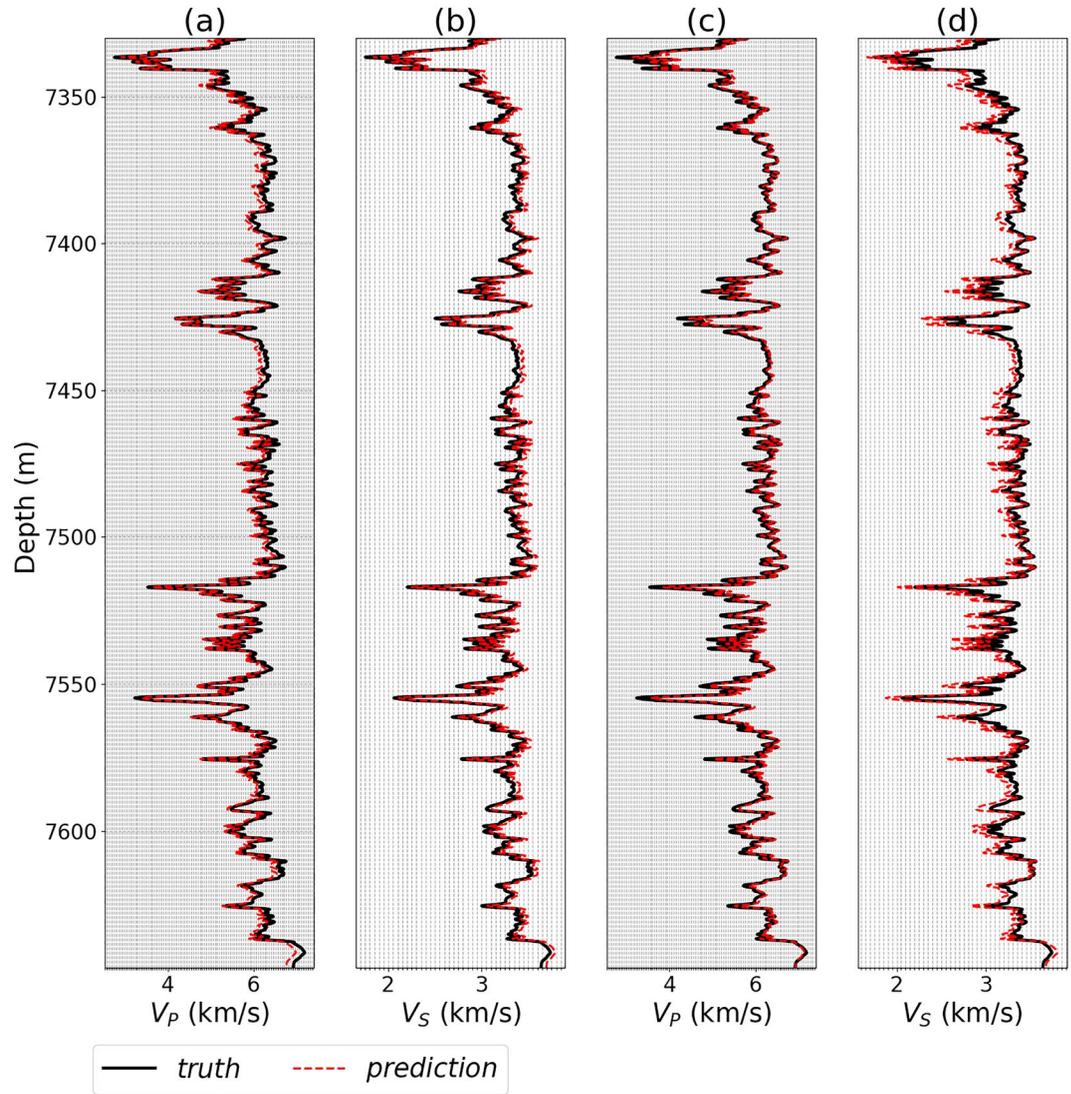


Figure 14. V_p and V_s predicted by ML, validated with the whole data set. The black solid curves are the ground truth, and the red dashed curves represent ML predictions for the two cases: trained with 25% of V_p and five points of V_s in (a and b), and trained with 80% of V_p and no V_s in (c and d).

where K_f is the bulk modulus of the fluid. The meaning of other symbols is the same as above. Instead of using Equation A2, as can be seen in Section 3, this study proposes to predict them by deep neural networks in reality. Once the required basic rock-physics parameters are given, a plane-wave analysis (Carcione, 2014) is used to solve the wave Equation 3, giving:

$$Z_1 \left(\frac{\omega}{k} \right)^4 + Z_2 \left(\frac{\omega}{k} \right)^2 + Z_3 = 0, \quad (\text{A3})$$

$$Z_1 \left(\frac{\omega}{k} \right)^2 - Z_4 = 0,$$

with

$$c_{11} = \rho_{11} + \sqrt{-1} \cdot b/\omega, \quad d_{11} = A + 2N, \quad c_{12} = c_{21} = \rho_{12} - \sqrt{-1} \cdot b/\omega,$$

$$d_{12} = d_{21} = Q, \quad c_{22} = \rho_{22} + \sqrt{-1} \cdot b/\omega, \quad d_{22} = R,$$

$$Z_1 = c_{11}c_{22} - c_{12}c_{21}, \quad Z_2 = c_{12}d_{21} + c_{21}d_{12} - c_{11}d_{22} - c_{22}d_{11}, \quad Z_3 = d_{11}d_{22} - d_{12}d_{21}, \quad Z_4 = Nc_{22}.$$

where ω is the angular frequency and k is the complex wave number. P and S solutions are obtained from the first and second equations, respectively. The phase velocity and quality factor are defined as (Carcione, 2014):

$$V = \left[\operatorname{Re} \left(\frac{k}{\omega} \right) \right]^{-1}, \quad Q = \frac{\operatorname{Re}(k)}{2\operatorname{Im}(k)}. \quad (\text{A4})$$

Data Availability Statement

The training and test data from log data used in this study are available online (at <https://zenodo.org/deposit/5003019>). The open source DNN python package PyTorch is used in this study.

Acknowledgments

The authors are grateful to the support of the Jiangsu Innovation and Entrepreneurship Plan, Jiangsu Province Science Fund for Distinguished Young Scholars (BK20200021), National Natural Science Foundation of China (41974123) and the research funds from SINOPEC Key Laboratory of Geophysics.

References

- Avseth, P., & Carcione, J. M. (2015). Rock-physics analysis of clay-rich source rocks on the Norwegian shelf. *The Leading Edge*, 34(11), 1340–1348. <https://doi.org/10.1190/tle34111340.1>
- Ba, J., Carcione, J. M., & Nie, J. X. (2011). Biot-Rayleigh theory of wave propagation in double-porosity media. *Journal of Geophysical Research: Solid Earth*, 116, B06202. <https://doi.org/10.1029/2010JB008185>
- Ba, J., Xu, W., Fu, L. Y., Carcione, J. M., & Zhang, L. (2017). Rock an elasticity due to patchy saturation and fabric heterogeneity: A double double-porosity model of wave propagation. *Journal of Geophysical Research: Solid Earth*, 122, 1949. <https://doi.org/10.1002/2016jb013882>
- Bai, Y., & Ma, Y. (2019). Geology of the Chang 7 Member oil shale of the Yanchang Formation of the Ordos Basin in central north China. *Petroleum Geoscience*, 26(2), 355–371. <https://doi.org/10.1144/petgeo2018-091>
- Barajas-Solano, D. A., Wohlberg, B. E., Vesselinov, V. V., & Tartakovsky, D. M. (2015). Linear functional minimization for inverse modeling. *Water Resources Research*, 51, 4516–4531. <https://doi.org/10.1002/2014WR016179>
- Batzle, M. L., & Wang, Z. (1992). Seismic properties of pore fluids. *Geophysics*, 57(11), 1396–1408. <https://doi.org/10.1190/1.1443207>
- Bernabé, Y., Wang, Y., Qi, T., & Li, M. (2016). Passive advection-dispersion in networks of pipes: Effect of connectivity and relationship to permeability. *Journal of Geophysical Research: Solid Earth*, 121, 713–728. <https://doi.org/10.1002/2015JB012487>
- Bernabé, Y., Zamora, M., Li, M., Maini, A., & Tang, Y. B. (2011). Pore connectivity, permeability, and electrical formation factor: A new model and comparison to experimental data. *Journal of Geophysical Research: Solid Earth*, 116, B11204. <https://doi.org/10.1029/2011JB008543>
- Berryman, J. G. (1988). Seismic wave attenuation in fluid-saturated porous media. *Pure and Applied Geophysics*, 128, 423–432.
- Biot, M. A. (1956a). Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range. *Journal of the Acoustical Society of America*, 28(2), 168–178. <https://doi.org/10.1121/1.1908239>
- Biot, M. A. (1956b). Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range. *Journal of the Acoustical Society of America*, 28(2), 179–191. <https://doi.org/10.1121/1.1908241>
- Biot, M. A. (1962). Mechanics of deformation and acoustic propagation in porous media. *Journal of Applied Physics*, 33(4), 1482–1498. <https://doi.org/10.1063/1.1728759>
- Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937. <https://doi.org/10.1073/pnas.1517384113>
- Carcione, J. M. (2014). *Wave fields in real media. Theory and numerical simulation of wave propagation in anisotropic, anelastic, porous and electromagnetic media* (3rd ed.). Elsevier.
- Carcione, J. M., Gurevich, B., & Cavallini, F. (2000). A generalized Biot-Gassmann model for the acoustic properties of shaley sandstones. *Geophysical Prospecting*, 48(3), 539–557. <https://doi.org/10.1046/j.1365-2478.2000.00198.x>
- Carcione, J. M., Picotti, S., Gei, D., & Rossi, G. (2006). Physics and seismic modeling for monitoring CO₂ storage. *Pure and Applied Geophysics*, 163, 175–207. <https://doi.org/10.1007/s00024-005-0002-1>
- Carman, P. C. (1961). *L'écoulement des gaz à Travers les milieux poreux, Bibliothèque des Sciences et Techniques Nucléaires*. Presses Universitaires de France.
- Champion, K., Lusch, B., Kutz, J. N., & Brunton, S. L. (2019). Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, 116(45), 22445–22451. <https://doi.org/10.1073/pnas.1906995116>
- Cheng, C. H. (1993). Crack models for a transversely isotropic medium. *Journal of Geophysical Research: Solid Earth*, 98, 675–684. <https://doi.org/10.1029/92JB02118>
- Dvorkin, J., Mavko, G., & Nur, A. (1995). Squirt flow in fully saturated rocks. *Geophysics*, 60(1), 97–107. <https://doi.org/10.1190/1.1443767>
- Dvorkin, J., & Nur, A. (1993). Dynamic poroelasticity: A unified model with the squirt and the Biot mechanisms. *Geophysics*, 58(4), 524–533. <https://doi.org/10.1190/1.1443435>
- Frazer, L. N., Sun, X., & Wilkens, R. H. (1997). Changes in attenuation with depth in an ocean carbonate section: Ocean Drilling Program sites 806 and 807, Ontong Java Plateau. *Journal of Geophysical Research: Solid Earth*, 102, 2983–2997. <https://doi.org/10.1029/96JB03189>
- Fu, J., Li, S., Niu, X., Deng, X., & Zhou, X. (2020). Geological characteristics and exploration of shale oil in Chang 7 Member of Triassic Yanchang Formation, Ordos Basin, NW China. *Petroleum Exploration and Development*, 47(5), 931–945. [https://doi.org/10.1016/s1876-3804\(20\)60107-0](https://doi.org/10.1016/s1876-3804(20)60107-0)
- Gassmann, F. (1951). Über die Elastizität poröser Medien. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 96, 1–23.
- Hamby, D. M. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, 32, 135–154. <https://doi.org/10.1007/bf00547132>
- Han, Z., Görtz, S., & Zimmermann, R. (2013). Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. *Aerospace Science and Technology*, 25(1), 177–189. <https://doi.org/10.1016/j.ast.2012.01.006>
- Han, Z., Zhang, Y., Song, C., & Zhang, K. (2017). Weighted gradient-enhanced kriging for high-dimensional surrogate modeling and design optimization. *AIAA Journal*, 55(12), 4330–4346. <https://doi.org/10.2514/1.J055842>
- Karimpouli, S., & Tahmasebi, P. (2020). Physics informed machine learning: Seismic wave equation. *Geoscience Frontiers*, 11(6), 1993–2001. <https://doi.org/10.1016/j.gsf.2020.07.007>
- Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*.

- Krief, M., Garat, J., Stellingwerff, J., & Ventre, J. (1990). A petrophysical interpretation using the velocities of P and S waves (full waveform sonic). *The Log Analyst*, 31, 355–369.
- Kuster, G. T., & Toksöz, M. N. (1974). Velocity and attenuation of seismic waves in two phase media: Part I. Theoretical formulation. *Geophysics*, 39, 587–606.
- Li, Y. E., Cheng, A. C., & You, N. (2019). Shale anisotropy estimation from logs in vertical wells. *Journal of Geophysical Research: Solid Earth*, 124, 6602–6611. <https://doi.org/10.1029/2019jb017572>
- Liu, X., Greenhalgh, S., Zhou, B., & Greenhalgh, M. (2018). Effective Biot theory and its generalization to poroviscoelastic models. *Geophysical Journal International*, 212(2), 1255–1273. <https://doi.org/10.1093/gji/ggx460>
- Liu, X., Greenhalgh, S., Zhou, B., & Heinson, G. (2016). Generalized poroviscoelastic model based on effective Biot theory and its application to borehole guided wave analysis. *Geophysical Journal International*, 207(3), 1472–1483. <https://doi.org/10.1093/gji/ggw345>
- Lo, W., Sposito, G., & Majer, E. (2005). Wave propagation through elastic porous media containing two immiscible fluids. *Water Resources Research*, 41(2), W02025. <https://doi.org/10.1029/2004WR003162>
- Lo, W., Sposito, G., & Majer, E. (2006). Low-frequency dilatational wave propagation through fully-saturated poroelastic media. *Advances in Water Resources*, 29(3), 408–416. <https://doi.org/10.1016/j.advwatres.2005.05.012>
- Moseley, B., Nissen-Meyer, T., & Markham, A. (2020). Deep learning for fast simulation of seismic waves in complex media. *Solid Earth*, 11, 1527–1549. <https://doi.org/10.5194/se-11-1527-2020>
- Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al. (2017). Automatic differentiation in pyTorch. *Paper presented at the 31st Conference on Neural Information Processing Systems (NIPS 2017)*.
- Pettit, C. L., & Wilson, D. K. (2020). A physics-informed neural network for sound propagation in the atmospheric boundary layer. *The Journal of the Acoustical Society of America*, 148, 2739. <https://doi.org/10.1121/1.5147609>
- Pride, S. R., & Berryman, J. G. (2003a). Linear dynamics of double-porosity dual-permeability materials. I. Governing equations and acoustic attenuation. *Physical Review E*, 68(3), 036603. <https://doi.org/10.1103/PhysRevE.68.036603>
- Pride, S. R., & Berryman, J. G. (2003b). Linear dynamics of double-porosity dual-permeability materials. II. Fluid transport equations. *Physical Review E*, 68(3), 036604. <https://doi.org/10.1103/PhysRevE.68.036604>
- Pride, S. R., Berryman, J. G., & Harris, J. M. (2004). Seismic attenuation due to wave-induced flow. *Journal of Geophysical Research: Solid Earth*, 109, 59–70. <https://doi.org/10.1029/2003jb002639>
- Qadrouh, A. N., Carcione, J. M., Alajmi, M., & Alyousif, M. M. (2019). A tutorial on machine learning with geophysical applications. *Bollettino di Geofisica Teorica ed Applicata*, 60(3), 375–402. <https://doi.org/10.4430/bgta0274>
- Qi, Q., Fu, L.-Y., Deng, J., & Cao, J. (2021). Attenuation methods for quantifying gas saturation in organic-rich shale and tight gas formations. *Geophysics*, 86(2), D65–D75. <https://doi.org/10.1190/geo2020-0291.1>
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2017a). *Physics informed deep learning (part I): Data-driven solutions of nonlinear partial differential equations*.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2017b). *Physics informed deep learning (part II): Data-driven discovery of nonlinear partial differential equations*.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Rudy, S. H., Alla, A., Brunton, S. L., & Kutz, J. N. (2019). Data-driven identification of parametric partial differential equations. *SIAM Journal on Applied Dynamical Systems*, 18(2), 643–660. <https://doi.org/10.1137/18M1191944>
- Santos, J. E., Corberó, J. M., & Douglas, J. (1990). Static and dynamic behavior of a porous solid saturated by a two-phase fluid. *The Journal of the Acoustical Society of America*, 87(4), 1428–1438. <https://doi.org/10.1121/1.399439>
- Santos, J. E., Douglas, J., Corberó, J., & Lovera, O. M. (1990). A model for wave propagation in a porous medium saturated by a two-phase fluid. *The Journal of the Acoustical Society of America*, 87(4), 1439–1448. <https://doi.org/10.1121/1.399440>
- Song, C., Alkhalifah, T., & Waheed, U. B. (2021). Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks. *Geophysical Journal International*, 225(2), 846–859. <https://doi.org/10.1093/gji/ggab010>
- Sun, W., Xiong, F., Ba, J., & Carcione, J. M. (2018). Effects of ellipsoidal heterogeneities on wave propagation in partially saturated double-porosity rocks. *Geophysics*, 83(3), WC71–WC81. <https://doi.org/10.1190/geo2017-0549.1>
- Sun, X., Tang, X.-M., Cheng, C. H. A., & Frazer, L. N. (2000). P and S wave attenuation logs from monopole sonic data. *Geophysics*, 65, 755–765. <https://doi.org/10.1190/1.1444774>
- Tartakovsky, A. M., Marrero, C. O., Perdikaris, P., Tartakovsky, G. D., & Barajas-Solano, D. (2020). Physics-informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. *Water Resources Research*, 56(5), e2019WR026731. <https://doi.org/10.1029/2019WR026731>
- Tuncay, K., & Corapcioglu, M. Y. (1996). Body waves in poroelastic media saturated by two immiscible fluids. *Journal of Geophysical Research: Solid Earth*, 101, 25149–25159. <https://doi.org/10.1029/96JB02297>
- Wang, N., Chang, H., & Zhang, D. (2021). Deep-learning-based inverse modeling approaches: A subsurface flow example. *Journal of Geophysical Research: Solid Earth*, 126, e2020JB020549. <https://doi.org/10.1029/2020JB020549>
- Wang, N., Zhang, D., Chang, H., & Li, H. (2020). Deep learning of subsurface flow via theory-guided neural network. *Journal of Hydrology*, 584, 124700. <https://doi.org/10.1016/j.jhydrol.2020.124700>
- Wang, T., Trugman, D., & Lin, Y. (2021). SeismoGen: Seismic waveform synthesis using GAN with application to seismic data augmentation. *Journal of Geophysical Research: Solid Earth*, 126, e2020JB020077. <https://doi.org/10.1029/2020JB020077>
- Wiese, B. U., Wagner, F. M., Norden, B., Maurer, H., & Schmidt-Hattenberger, C. (2018). Fully coupled inversion on a multi-physical reservoir model—Part I: Theory and concept. *International Journal of Greenhouse Gas Control*, 75, 262–272. <https://doi.org/10.1016/j.ijggc.2018.05.013>
- Xu, S., & White, R. E. (1995). A new velocity model for clay-sand mixtures. *Geophysical Prospecting*, 43(1), 91–118. <https://doi.org/10.1111/j.1365-2478.1995.tb00126.x>
- You, N., Li, Y. E., & Cheng, A. (2020). Shale anisotropy model building based on deep neural networks. *Journal of Geophysical Research: Solid Earth*, 125, e2019JB019042. <https://doi.org/10.1029/2019JB019042>
- You, N., Li, Y. E., & Cheng, A. (2021). 3D carbonate digital rock reconstruction using progressive growing GAN. *Journal of Geophysical Research: Solid Earth*, 126, e2021JB021687. <https://doi.org/10.1029/2021JB021687>
- Zhang, L., Ba, J., & Carcione, J. M. (2021). Wave propagation in infinituple-porosity media. *Journal of Geophysical Research: Solid Earth*, 126, e2020JB021266. <https://doi.org/10.1029/2020JB021266>

- Zhao, L., Cao, C., Yao, Q., Wang, Y., Li, H., Yuan, H., et al. (2020). Gassmann consistency for different inclusion-based effective medium theories: Implications for elastic interactions and poroelasticity. *Journal of Geophysical Research: Solid Earth*, *125*, e2019JB018328. <https://doi.org/10.1029/2019JB018328>
- Zhao, L., Han, D., Yao, Q., Zhou, R., & Yan, F. (2015). Seismic reflection dispersion due to wave-induced fluid flow in heterogeneous reservoir rocks. *Geophysics*, *80*(3), D221–D235. <https://doi.org/10.1190/geo2014-0307.1>
- Zhao, L., Yao, Q., Han, D., Yan, F., & Nasser, M. (2016). Characterizing the effect of elastic interactions on the effective elastic properties of porous, cracked rocks. *Geophysical Prospecting*, *64*(1), 157–169. <https://doi.org/10.1111/1365-2478.12243>