GEOPHYSICS, VOL. 84, NO. 6 (NOVEMBER-DECEMBER 2019); P. R869–R880, 10 FIGS., 1 TABLE.
10.1190/GEO2018-0838.1

Convolutional neural network for seismic impedance inversion

Vishal Das[1], Ahinoam Pollack[2], Uri Wollner[1], and Tapan Mukerji[3]

ABSTRACT

We have addressed the geophysical problem of obtaining an elastic model of the subsurface from recorded normal-incidence seismic data using convolutional neural networks (CNNs). We train the network on synthetic full-waveform seismograms generated using Kennett's reflectivity method on earth models that were created under rock-physics modeling constraints. We use an approximate Bayesian computation method to estimate the posterior distribution corresponding to the CNN prediction and to quantify the uncertainty related to the predictions. In addition, we test the robustness of the network in predicting impedances of previously unobserved earth models when the input to the network consisted of seismograms generated using: (1) earth models with different spatial correlations (i.e. variograms), (2) earth models with different facies proportions, (3) earth models with different underlying rock-physics relations, and (4) source-wavelet phase and frequency different than in the training data. Results indicate that the predictions of the trained network are susceptible to facies proportions, the rock-physics model, and source-wavelet parameters used in the training data set. Finally, we apply CNN inversion on the Volve field data set from offshore Norway. P-wave impedance I_P inverted for the Volve data set using CNN showed a strong correlation (82%) with the I_P log at a well.

INTRODUCTION

Artificial neural networks (ANNs) have been applied successfully to various geophysical problems to determine the nonlinear relationships from the data. Shallow neural networks have been used for inversion of seismic data to obtain subsurface elastic attributes (Röth and Tarantola, 1994; Calderón-Macías et al., 1998, 2000; Moya and Irikura, 2010; Shahraeeni et al., 2012), to estimate lithology and permeability from well logs (Wiener et al., 1991; Rogers et al., 1992; Huang et al., 1996) and to address many other problems. For a summary of the applications of ANNs in geophysics, see Poulton (2002).

In recent times, the availability of massive computational resources has resulted in novel data-driven approaches for addressing various research questions (Monajemi et al., 2016). This progress in data-driven approach is exemplified by the tremendous advances being made in deep-learning-related research and the numerous examples of their successful application in diverse fields. Deep-learning applications incorporate deep neural networks, which have evolved from shallow ANNs. Deep neural networks are comprised of multiple hidden layers, and the "depth" of the network is defined by the number of hidden layers. Training such deep networks in a supervised learning task includes mapping an input vector to an output vector and requires a sufficiently large training data set that includes labeled data (Goodfellow et al., 2016). The approximate form of the true mapping function established by the deep network is refined iteratively through a training process that updates the internal parameters of the network subject to minimizing a difference criterion between computed and labeled outputs. One form of deep neural networks called convolutional neural networks (CNNs) has recently gained attention as Krizhevsky et al. (2012) used them successfully in classifying images into 1000 different classes.

Deep neural networks are also gaining popularity in the geoscientific community. They have been successfully applied in solving geophysical problems such as seismic horizon and fault interpretation, seismic texture identification, and others. Araya-Polo et al. (2018) use deep CNNs to reconstruct subsurface velocity model from synthetically generated seismograms. Richardson (2018) uses another variation of deep neural network called the recurrent neural network (RNN) (when the output of a layer is fed back into the same layer as an input) for seismic full-waveform inversion. Wu and McMechan (2018) use a combination of conventional full-waveform inversion

Manuscript received by the Editor 4 December 2018; revised manuscript received 5 May 2019; published ahead of production 07 August 2019; published online 09 October 2019.
[1]Stanford University, Department of Geophysics, Stanford, California, USA. E-mail: vdas2@stanford.edu (corresponding author); wollner.uri@gmail.com.
[2]Stanford University, Department of Energy Resources Engineering, Stanford, California, USA. E-mail: ahinoamp@stanford.edu.
[3]Stanford University, Department of Geophysics, Stanford, California, USA and Stanford University, Department of Energy Resources Engineering, Stanford, California, USA. E-mail: mukerji@stanford.edu.
© 2019 Society of Exploration Geophysicists. All rights reserved.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

with CNNs to capture salient features in the geometry and obtain better inversion results. Zhang et al. (2018) use CNNs for seismic lithology prediction. Das et al. (2018) introduce CNN-based seismic impedance inversion that is extended using a physics-guided machine-learning approach by Biswas et al. (2019).

This paper is an extension of Das et al. (2018). In this paper, we address the quintessential geophysical problem of P-impedance I_P inversion from normal-incidence seismic data that has been solved previously using traditional deterministic and stochastic seismic-inversion methods (Tarantola, 2005; Sen, 2006). The research questions that we address are: Can modern deep neural networks solve the I_P inversion problem? If yes, are the results comparable or better than the traditional inversion results? To address these questions, we use a modern neural network architecture based on CNNs. CNNs include a convolution operator as part of their framework and form a natural fit to the I_P inversion problem as recorded seismic data are usually modeled as a wavelet convolved with reflectivity series.

In addition to testing the applicability of modern deep neural networks for I_P inversion, we also focus on other concerns related to successful application of CNNs in geophysical problems. Deep neural networks face the problem of hyperparameter tuning in the training stage to establish an optimum neural-network architecture. How can we effectively tune these parameters of the neural network? Can we use domain knowledge to systematically perform hyperparameter tuning? CNNs have been typically used for classification problems. Can we modify them for regression problems in which the output is a continuous variable rather than class labels? Additionally, CNNs require a large labeled training data set that has the input and the corresponding output known. How can we augment the available data, for example, data from the drilled well locations, and generate a large, geologically realistic training data set? CNNs can be considered as nonlinear regressors, which, after training, give us an *estimate*. However, the inherent uncertainties involved in geoscientific problems make it imperative to quantify the *uncertainty* associated with the estimate. How can we quantify the uncertainty in estimations from neural networks? The robustness of a neural network in its prediction for data sets with different statistical distributions as compared to training data set distribution is unknown. We focused on systematically testing the robustness of a trained network on samples that are outside the training set distribution, for example, different facies proportions, different source-wavelet parameters, and others. The question we address is whether the trained network is robust for some parameters as compared to others.

Compared to previous papers that used ANNs to solve the I_P inversion problem, this paper addresses the questions raised above and makes the following contributions:

- Proposes a convolutional-based neural network for the impedance-inversion problem.
- Combines CNN-based prediction with an approximate Bayesian computation (ABC) method to estimate the posterior distribution for uncertainty quantification related to the predicted I_P values.
- Studies the robustness of trained CNN when the input to the network consisted of data generated from a different distribution than that used for training. We test the predictions with seismograms generated with (1) earth models with different spatial correlations (i.e., variograms), (2) earth models with different facies proportions, (3) earth models with different underlying rock-physics relations, and (4) source-wavelet phase and frequency different than in the training data.

The paper is organized as follows: First, we discuss the CNN-based seismic-inversion workflow. Second, we describe the methodology of augmenting data available from a well location and generating a statistically similar labeled training data set. Third, we discuss the tuning process of the hyperparameters and describe the architecture of the neural network. Fourth, we describe the implementation of uncertainty quantification of the neural network's estimate using an ABC approach. Fifth, we systematically test the robustness of a trained network on samples outside the training data set distribution. Last, we apply the CNN-based inversion on field data from offshore Norway and discuss the results obtained.

SEISMIC INVERSION: WORKFLOW

The general workflow used in this work to invert I_P from seismic data can be described in the following steps:

1) Generate training data by (1) simulating multiple earth models with defined elastic properties (I_P) and (2) numerically computing the normal-incidence seismic signals.
2) Train CNN using generated training data, in which input to the network are seismic traces and output are I_P profiles.
3) Use trained CNN for prediction of I_P from seismic traces.
4) Calculate posterior uncertainty using CNN estimate as summary statistics for ABC.

METHOD: GENERATING A TRAINING DATA SET

The prediction accuracy of supervised ANNs depends heavily on the complexity of the architecture and the amount of data available for training. Unlike other areas where ANNs have found tremendous success, such as image classification (Krizhevsky et al., 2012), in which labeled data are relatively easy to obtain, the availability of real labeled data in exploration geophysics is often limited. An example of a labeled data in the I_P inversion problem is the impedance log at a well location and the corresponding seismic data at the same location. However, the number of wells drilled is often limited, thus limiting the number of available labeled training data. To overcome this problem, the labeled data at the well location must be augmented using physics and geology to generate a synthetic labeled data set for training the CNN model.

The geology of the earth is created by physical, nonrandom processes; thus, augmenting the well data and generating synthetic earth models must at least capture the spatial correlation imposed by depositional processes. In the first part of the work, we assume a clastic depositional environment with two predominant sand facies in the reservoir section — good-quality sand referred to as facies 1 and poor-quality sand referred to as facies 2. In this case, we had data from a well drilled in an offshore clastic turbiditic reservoir. We obtain the distributions of the facies and petrophysical properties: porosity φ and volume of clay V_{clay} from the well logs. We then simulate 2000 1D earth models each 200 m thick with geostatistical techniques using property distributions from the well. Sequential indicator simulation (Deutsch and Journel, 1992; Remy et al., 2009) is used to generate several realizations of facies with predefined facies proportions and variogram parameters (Figure 1a). Similarly, sequential Gaussian simulation (Deutsch and Journel, 1992; Journel, 1994; Remy et al., 2009) is used to populate each facies with petrophysical properties using predefined variogram parameters (Figure 1b). The variogram type assumed for both the simulations

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

is spherical. We use the constant-cement rock-physics model (Avseth et al., 2000; Mavko et al., 2009) to relate the generated petrophysical properties (φ and V_{clay}) to P-wave velocity and density and, finally, I_P (Figure 1c) as a product of density and P-wave velocity. The constant-cement model is chosen based on calibrating the model curves with I_P and φ data at the well location (Figure 1e). The same constant-cement model is used for all depths because the target window of inversion is narrow, with porosity variations mostly related to sorting or depositional trends as compared to diagenetic changes. Although the crossplot between I_P and φ (Figure 1e) indicates the presence of a mixed facies, only two facies were considered in this work for simplicity. The P-impedances, along with a wavelet, are used as inputs to a 1D full-waveform normal-incidence synthetic seismograms generator based on the invariant imbedding (reflectivity) method (Kennett, 1974, 1983) (Figure 1d). These generated normal-incidence synthetic seismograms were used as input training data to the CNN network. The I_P values, which are a product of the P-wave velocity and bulk density, serve as the labeled output data in the training process of the network. Because the generated seismic trace is a function of the two-way traveltime, we convert the impedances

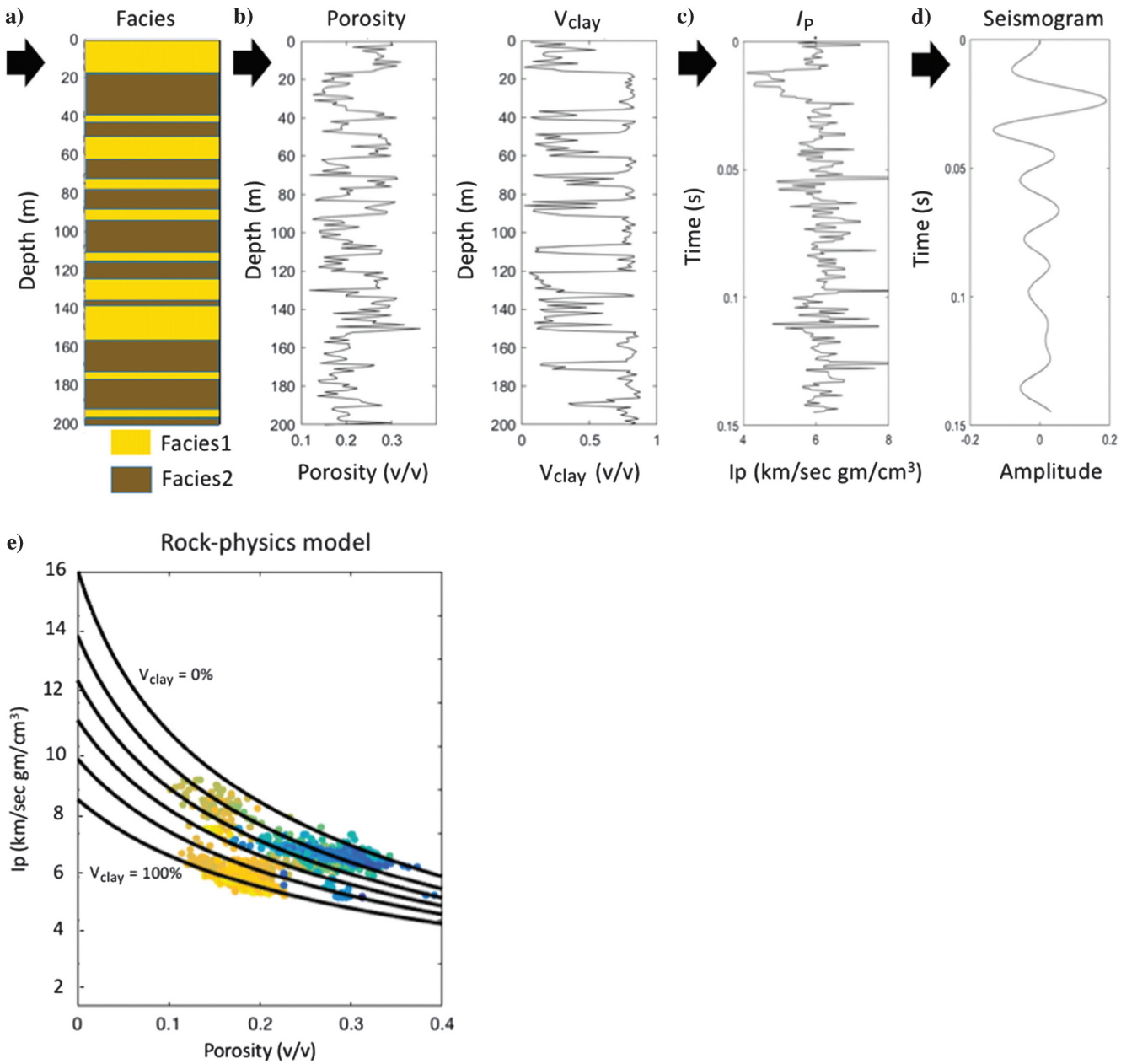

Figure 1. Workflow for generating seismic data. (a) Generating the facies and (b) petrophysical properties φ and V_{clay}, as a function of depth. (c) Calculating I_P using the petrophysical properties as input to the rock-physics model. (d) Generating synthetic seismic data using forward modeling. (e) The constant cement rock-physics model is fit to the I_P vs. Porosity crossplot. The colors represent V_{clay}, with warmer colors representing higher V_{clay} and cooler colors representing lower V_{clay}.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

from depth to time by using the corresponding velocity and the computed ray-theory traveltimes. Each pair of I_P seismic trace consists of 330 time samples.

The synthetic data set generation process has the following parameters: (1) two vertical variogram ranges defining the vertical correlation of facies and petrophysical properties, (2) facies proportions, (3) coordination number, an input parameter for the constant-cement rock-physics model referring to the average number of contacts per grain in the sediments, and (4) source-wavelet parameters — central frequency and phase. We generate three sets of synthetic data. The first base-case data set is generated using the base-case parameters described in Table 1. This data set was used to train and tune a neural-network architecture and its learning parameters. The second data set to train and test the robustness of the proposed neural network to the different parameters is generated using the regular testing (RT) parameters described in Table 1. The third data set is generated using the challenge testing (CT) parameters described in Table 1. The third data set is used to analyze the robustness of a network trained using the second data set to samples outside the training data set distribution.

An alternative approach to the data augmentation from well locations can also be performed without having to define facies, porosity, clay content, and a rock-physics model. This approach uses the I_P log and the time-depth relationship at the well locations along with a corresponding seismic trace. The source wavelet can be estimated by performing a seismic-to-well-tie. The I_P log can be used to estimate the vertical correlation. Sequential Gaussian simulation can then be used to generate several realizations of the I_P log with similar statistics, bypassing the generation of facies, and facies properties. Convolution of the reflectivity series estimated from the simulated I_P logs with the source wavelet estimated using the seismic-to-well-tie process can be used to obtain the labeled I_P seismic trace pairs. The synthetic seismograms generated using the reflectivity-series method and by the convolution of reflectivity series with a source wavelet are expected to give similar results when the impedance contrasts between the layers are moderate. This method of data augmentation has been used for the Volve field example described in the last section of this paper.

METHOD: NEURAL NETWORK ARCHITECTURE AND HYPERPARAMETER TUNING

To address the seismic-impedance inversion problem in a layered, linearly elastic earth model, we select a CNN-based neural network (Lee et al., 2009; Krizhevsky et al., 2012). We use the PyTorch implementation of 1D CNN (Paszke et al., 2017). The hyperparameters for the neural network include: (1) number of hidden layers and units, (2) kernel width and stride of the convolution operator that defines the length of the convolution operator and the sliding of the operator over the 1D input vector, respectively, (3) activation function, (4) weights initialization, (5) optimization method and the parameters for the chosen optimization scheme, and (6) number of epochs. We choose the simplest convolutional network architecture with two 1D convolutional layers (Figure 2). The 1D CNN layers perform convolution of the 1D input vector with a convolution kernel defined by the kernel width. The kernel width is fixed at 300 samples comparable to the size of the number of time samples in one wavelength corresponding to the central frequency of the source wavelet (30 Hz). The learning curve of the network converged better when the kernel size is chosen to be comparable to the number of samples in the source wavelet in the hyperparameter tuning stage. The stride of the convolution kernel is fixed at one sample. Zero padding is used at the end of the input so that the input and the output length are the same. At the end of each of the layers, we introduce nonlinearity using rectified linear unit (ReLU) defined as $f(x) = \max(0, x)$ (Nair and Hinton, 2010) as the activation function. The weights of the convolution layers are initialized using He initialization (He et al., 2015), which has been found to give better convergence of the weights in the training process. We randomly select 70% of the data from the 2000 I_P seismic trace pairs from the base-case data set for training the CNN, and 15%

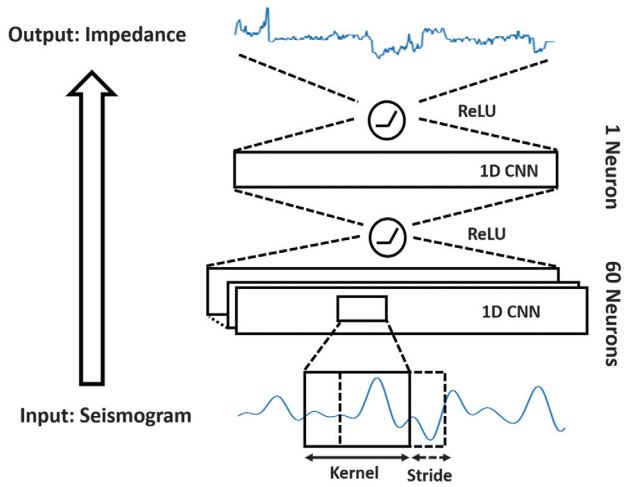

Figure 2. The CNN architecture used in this work. The architecture consists of two 1D CNN layers with a ReLU layer at the end of each layer. The training process on the known I_P and seismic traces was used to optimize the weights of the layers in the back-propagation step.

Table 1. Parameters used in generating different data set.

Parameters	Base case	Training and RT	CT
Spherical variogram ranges (m) of facies	20	Uniform ~ [10, 60]	Uniform ~ [(3, 8) ∪ (70, 100)]
Spherical variogram ranges (m) of φ and V_{clay}	20	Uniform ~ [10, 60]	Uniform ~ [(3, 8) ∪ (70, 100)]
Facies1 proportion (%)	60	Uniform ~ [30, 60]	Uniform ~ [(20, 25) ∪ (65, 80)]
Coordination number	14	Uniform ~ [9, 12]	Uniform ~ [(6, 8) ∪ (13, 14)]
Phases (°)	0	Uniform ~ [20, 70]	Uniform ~ [(0, 10) ∪ (80, 90)]
Central frequency (Hz)	30	Uniform ~ [30, 50]	Uniform ~ [(15, 25) ∪ (55, 65)]

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

each for the validation set, and the test set. The number of samples for training is selected based on the number of trainable parameters (approximately 36,000) for the current network architecture and also by evaluating the training and validation performance with different amounts of training data. We use mean-squared error as the loss function in the training process to measure the difference between the output from the network and the true output from the labeled data set. The optimization of the weights of the neural network in the back-propagation step of the network is done using Adam optimization (Kingma and Ba, 2014). We select the learning rate for the model update to be 0.001, and we used 500 epochs to train the network. These hyperparameters are tuned by assessing the performance of the network in the training and the validation data sets. The test set is untouched during the hyperparameter tuning. The final network architecture fixed after hyperparameter tuning is shown in Figure 2. The same neural network architecture is trained and validated using 10 realizations of random He initializations of the weights. The convergence plots of the training and the validation error for the fixed network architecture with 10 initializations are shown in Figure 3. The L2-norm of errors in the convergence plot is normalized by the mean of I_P values. The network with the lowest normalized training error of 6% among the 10 trained networks is selected for the rest of the analysis. The corresponding normalized validation error is approximately 10%. The training is performed using a GPU implementation of 1D CNN in PyTorch for faster computations and takes approximately 45 minutes for one training process. We maintain the same network architecture for the rest of the work.

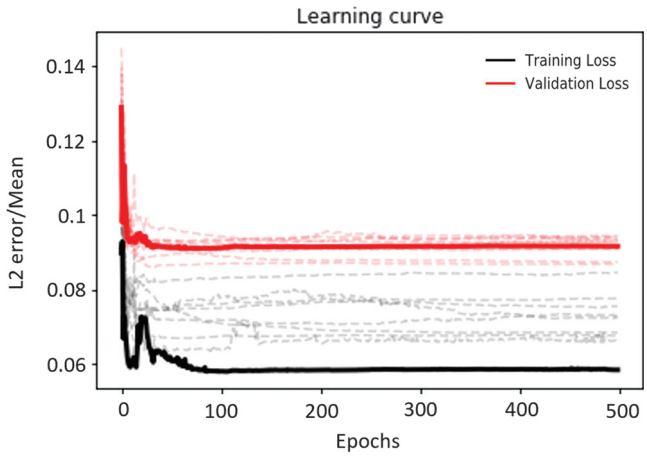

Figure 3. Convergence plot of L2-norm of error with respect to mean I_P values for training and validation set for 500 epochs with 10 different initializations. The training process was stopped at 500 epochs based on these curves to prevent the network from overfitting the training data set. The curves in bold correspond to the final network that was selected.

RESULTS AND COMPARISON WITH THE TRADITIONAL INVERSION METHOD

We first perform the training and validation of the CNN with the data set generated using the base-case values (Table 1) and compare the results of the CNNs predictions with a conventional model-based least-squares inversion technique that takes an initial earth model and perturbs the initial model until the derived synthetic seismic best fits the observed data. The comparison is performed to understand the differences and advantages of using CNN in comparison to a traditional inversion method. The trained and validated CNN is used to predict I_P-time profiles from seismic traces in the test set. We select a seismic trace from the test data set to perform model-based least-squares inversion (Russell and Hampson, 1991) using Hampson-Russell software. We used the source wavelet from the seismic forward-modeling step and a low-frequency I_P trace obtained by applying a moving-average window of 30 samples (approximately one-tenth of the seismic-source wavelength corresponding to the central frequency of the source wavelet) on the true I_P trace as inputs to the least-squares inversion. Figure 4 compares the true I_P, CNN-predicted I_P, and inverted I_P using the model-based inversion technique, along with its corresponding seismogram. The model-based inversion is performed such that the least-squares solution gives full weight to the data mismatch and no penalty is given to the norm of the model parameter perturbations. The model-based least-squares inversion and CNN inversion show a close match with the true I_P values. However, the correlation between the CNN inverted I_P and true I_P is 95%. This is higher than the

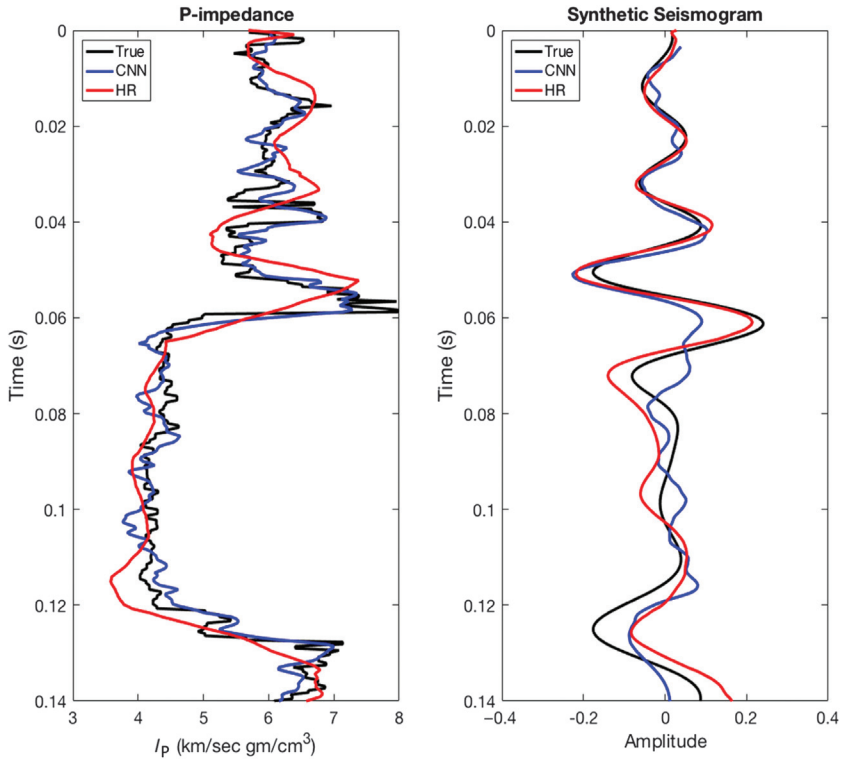

Figure 4. Comparison between CNN and model-based least-squares (LS) inversion (from the Hampson-Russell [HR] software suite). Synthetic seismograms generated from the impedance traces are plotted against the true seismogram. Correlation between the true and the CNN-inverted impedance is 95%, and for the LS impedance it is 89%. The data correlation for the CNN impedance is 66%, and the LS impedance is 85%.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

correlation of 89% between the model-based inverted I_P and true I_P. The correlation between the observed and synthetic seismogram generated from the CNN inverted I_P (66%) is lower than the correlation between observed and synthetic seismogram obtained using the model-based inverted I_P (85%). The objective function in the model-based inversion minimizes the data mismatch that leads to a higher data correlation in this case. The objective function in the CNN inversion only minimizes the model mismatch. A closer look at the inverted I_P profiles reveals that CNN-inverted I_P reproduces some of the high-frequency components of the impedance that are missing from the model-based inversion. This analysis shows that the CNN-based approach is at least as good as (or perhaps better than) conventional deterministic impedance inversion.

POSTERIOR UNCERTAINTY QUANTIFICATION USING APPROXIMATE BAYESIAN COMPUTATION

The training step in machine learning is a nonlinear regression to obtain a set of network parameters by minimizing the cost function. Thus, the regression gives a single estimate for each new input observation. However, quantitative geoscientific interpretations require some measure of uncertainty associated with the estimate. In an attempt to go beyond the neural network's estimate, we need to sample the posterior distribution of the response variables. To this end, we propose the use of ABC techniques (Rubin, 1984; Pritchard et al., 1999; Beaumont et al., 2002). The ABC method provides a framework for Bayesian inference in cases in which the likelihood distribution is intractable.

The approximation to the posterior distribution in ABC is made using various summary statistics or features extracted from observed data, as shown in equation 1,

$$f(m|d = d_{\text{obs}}) \sim f(m|[S(d) - S(d_{\text{obs}})] < \epsilon, \tag{1}$$

where $f(x)$ denotes the probability distribution of random variable x and m and d_{obs} are the model variable (i.e., I_P here) and observed seismic data, respectively. The term $S(d)$ denotes the data summary statistics, and ϵ is a threshold. Prior models whose data summary statistics match those of the observed data within a certain threshold are accepted as samples of the approximate posterior distribution. The use of data summary statistics in the ABC framework is what enables accounting for machine learning predictions during sampling of posterior distributions. This result was first established by Fearnhead and Prangle (2012), who use regression models for prediction variables as summary statistics. Pradhan and Mukerji (2018) demonstrate the use of this approach by using deep neural networks to estimate posterior uncertainties of average reservoir properties in subseismic reservoir intervals from seismic data.

In this problem, we can compute the posterior probability of I_P realizations using ABC. In the first step, we define the model for posterior probability density function (pdf) using Bayes' theorem (equation 2)

$$P(m|d) \propto P(m)P(d|m), \tag{2}$$

where $P(m|d)$ is the posterior pdf, $P(m)$ is the prior pdf, $P(d|m)$ is the likelihood, m represents the model (I_P), and d represents the data (seismic). The proportionality constant is not relevant for our purpose because we will only classify the realizations from our prior data set in quantiles without associating a probability value. The term $P(m)$ in our case is defined by the ABC data set that is kept separate from the training, validation, or testing process. The ABC data set consists of a large number of prior models and is kept separate for ABC computation. The bias in the prior data set can be estimated by prior falsification in which we check whether the prior is consistent with the data. In this case, as it is a synthetic example, by construction, the prior is consistent, and we do not need the prior falsification step. The likelihood pdf is estimated by considering adequate samples from the prior data set using the rejection ABC algorithm (Beaumont, 2010; Csilléry et al., 2010). The summary statistic in this case is considered as the sum of squares of residuals, given by

$$S^2(m) = \sum_{k=1}^{K} (d_k^{\text{calc}}(m) - d_k^{\text{obs}})^2 \leq \epsilon, \tag{3}$$

where $S^2(m)$ is the sum of squares of residuals corresponding to the model (m), $d_k^{\text{calc}}(m)$ is the seismic data corresponding to m calculated using seismic forward modeling, d_k^{obs} is the observed seismic trace, K is the total number of time samples in a seismic trace, and ϵ is the threshold for the ABC rejection algorithm. Assuming that the data residuals are characterized by an uncorrelated Gaussian distribution with mean 0 and standard deviation σ_d, $S^2(m)$ is distributed as a chi-square distribution with K degrees of freedom. For large values of K, such as the case in our problem (330 samples), the chi-squared distribution converges to a normal distribution with parameters given by

$$S^2(m) \sim N(K\sigma_d^2, 2K\sigma_d^4), \tag{4}$$

where K is the number of time samples in a seismic trace and σ_d is the standard deviation of the data residuals. The standard deviation σ_d involves two terms: the observation-error deviation and the calculation-error deviation. In our work, we assume the standard deviation of the residuals as 25% of the maximum amplitude of the observed seismic trace. The threshold in equation 3 for any specific percentile of the likelihood function can now be obtained using common tables for the pdf of the normal distribution (equation 4). The likelihood for different percentiles is given by the seismic traces from the ABC data set that are selected based on the defined threshold (equation 3). The posterior distributions for each of the likelihoods are defined by the I_P models corresponding to the selected seismic traces. We considered P1, P10, P30, P50, P70, P90, and P99 percentiles as thresholds in this work. The posterior data set is created by segmenting the models into different percentile bins and then weighting the bins by their appropriate probabilities. The weighting is performed because the thresholding in the ABC rejection algorithm takes all the selected realizations as "equiprobable," disregarding the seismic-data likelihood. To make further calculations easy, instead of using samples with unequal weights, we replicate the samples in each bin several times proportional to the weights to create a set of samples with effectively equal weights. The I_P models in the pth bin are weighted by replicating the I_P models in this bin w times. Here, weight w is obtained by subtracting maximum percentile value in the pth bin from 100. For example, in the 70th percentile bin, we replicate the traces in the bin 30 times.

In the synthetic example, we have the true and the CNN-predicted I_P traces corresponding to a seismic trace from the test data set. We follow the procedure described above twice — first for the true I_P trace and second for the CNN-predicted I_P trace. A comparison of the

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

posterior estimated for the two cases helps us to show that the CNN prediction honors the exact data likelihood with good approximation. Histograms of S^2 between the true seismic trace and seismic traces in the ABC data set (Figure 5a) and S^2 between the seismic trace obtained using seismic forward modeling of CNN I_P prediction and seismic traces in the ABC data set (Figure 5b) are plotted along with the P1, P50, and P99 thresholds. The plots show that we operate at the left tail of the S^2 distribution; hence, we need a large number of samples in the prior set to select at least a minimum number of traces (assumed 10 in our work) based on the threshold values for different percentiles. The posterior data sets (approximate ABC posterior) are created based on the procedure of binning and weighting described above for the true I_P and the CNN-predicted I_P cases. The range of the prior, the range of the estimated posterior, P10, P50, P90 along with true and CNN I_P traces are plotted for the true I_P case (Figure 6a) and CNN-predicted I_P case (Figure 6b). Both plots look reasonably similar to the posterior range for the true I_P case (Figure 6a) slightly narrower than the CNN I_P case (Figure 6b). This establishes that the CNN "inversion" honors the exact data likelihood with good approximation. The blue band in these figures can be related to the uncertainty associated with the CNNs prediction and can be especially useful in real problems in which the true I_P values are unavailable.

ANALYSIS OF ROBUSTNESS OF CNN

We set up several experiments to test the robustness of the seismic-inversion capability of the CNN under varying geologic conditions and seismic modeling parameters using the RT and CT data set described in the "Generating a training data set" section. We vary the variogram ranges of the facies and the petrophysical properties, the proportions of facies, the rock-physics model, and the source wavelet: phase and central frequency used in seismic forward modeling. The parameters of the experiments, along with the parameters used in the base case, are detailed in Table 1. In each experiment, only a single parameter is varied. For example, in the facies variogram data set, the training and testing data sets are created by keeping the source wavelet phase and central frequency, coordination number, facies proportion, and variogram ranges of the petrophysical properties constant with values used in the base case, and only varying the variogram range of facies. This makes it possible to isolate the CNN's capability to predict I_P profiles from seismic traces on models with varying variogram ranges of the facies.

Each experiment has two levels of testing: RT and CT. In RT, we train a network and tested on data from the RT data set that comes from the same distribution as used for training. In CT, we train the network on the RT data set but use the CT data generated using parameters outside of the training data set for testing. We train the CNN and calculated prediction errors for the experiments using parameters shown in Table 1. The training and RT are performed using the RT data set that consists of 2000 I_P-seismic trace pairs split randomly into 70% for training and 30% for RT. In this case, we did not have a validation data set because the network architecture and hyperparameters are kept the same as those obtained for the base case described earlier. The CT data set for each scenario consists of 600 I_P seismic trace pairs. The prediction performance of the CNN on different scenarios is assessed using the L2-norm between the observed and predicted I_P, normalized by the mean I_P for each 1D trace for RT and CT data set. Each of the testing scenarios and the corresponding results are discussed in detail in the following subsections.

Robustness of CNN to variogram parameters for facies and petrophysical properties

The vertical spatial correlation of facies and petrophysical properties (φ and V_{clay}) is governed by the range of the vertical variogram

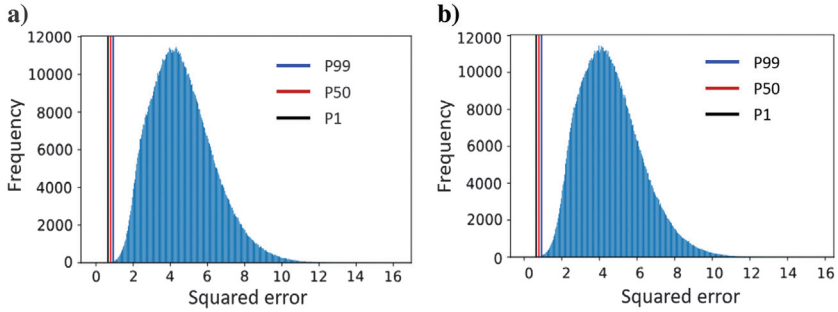

Figure 5. Histogram of squared-error residuals between (a) the true seismic trace and seismic traces in the ABC data set and (b) the seismic trace obtained using seismic forward modeling from CNN-predicted I_P trace and seismic traces in the ABC data set. The P1, P50, and P99 thresholds are plotted for both figures.

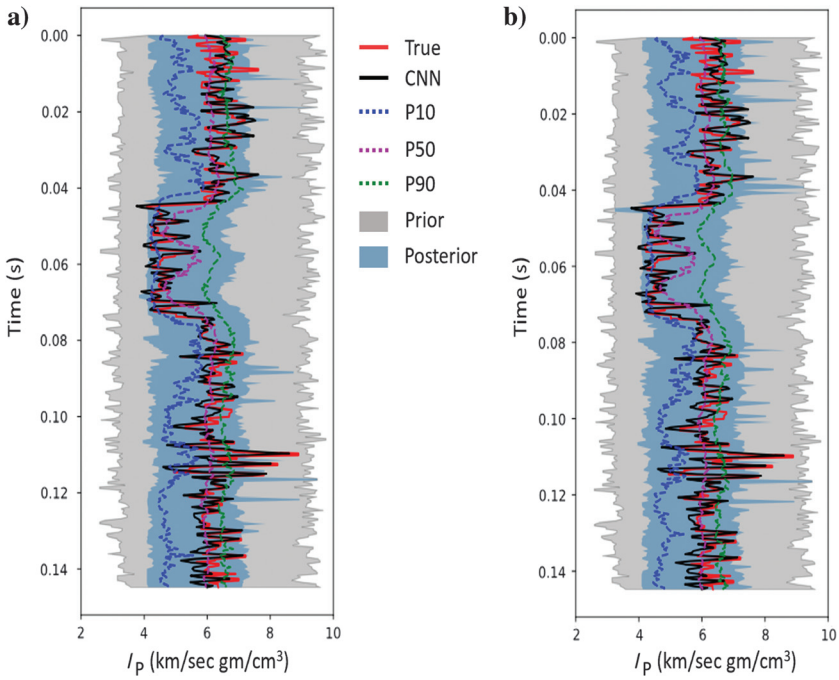

Figure 6. Prior range, estimated posterior range, true I_P, CNN-predicted I_P, along with P10, P50, and P90 from estimated posterior data set for (a) the true I_P case and (b) the CNN-predicted I_P case.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

used in the simulation. We assume that the porosity and volume of clay have the same vertical spatial distribution; hence, we simulated both properties using the same vertical variogram parameters. The variogram type for both cases is fixed as spherical. The variogram ranges in the training and the RT data set are drawn randomly from a uniform distribution between 10 and 60 m. The variogram range for the base case is 20 m. The CT data set consists of samples with higher or lower spatial correlation as compared to the training data set with variogram ranges drawn randomly from a uniform distribution between 3–8 and 70–100 m. CNNs I_P prediction matches the observed I_P traces closely for the RT and the CT data sets. The normalized (normalized by the mean I_P) L2-norm of error (ε_n) is calculated for the RT and CT data sets. For the facies variogram parameter tests, the ε_n for RT and CT data sets is approximately 10%. For the petrophysical variogram parameter tests, the ε_n for RT and CT data sets is approximately 7.5%. The error bars for the RT and CT data sets for both cases are small and have an equal spread around the mean (Figure 7). This demonstrates the robustness of the CNN predictions to subseismic spatial correlation of facies and petrophysical properties for this example.

Robustness of CNN to facies proportion

The facies simulation results also depend on the proportions of the two facies along with the variogram parameters. The base case proportions of facies1 and facies2 are fixed at 40% and 60% based on statistics obtained from a reference well. In the training and RT of the neural network for this case, the facies1 proportions are drawn randomly from a uniform distribution between 30% and 60%. The proportions of facies1 for the CT data set are drawn randomly from a uniform distribution between 20%–25% and 65%–80%. Thus, the CT data set has either a lower or higher proportion of facies1 than the training and RT data sets. CNN predictions in the training and RT data sets match closely with the observed I_P values. The ε_n for the RT data set is approximately 7.5%. However, the ε_n for the CT data set is approximately 12%. The error bar for the CT data set (Figure 7) has a much wider spread than that of the RT data set. This indicates that the CNN predictions depend on the facies proportion in the training data set distribution. When the facies proportion distributions in the training data are different from the test data, the network's I_P prediction in some of the samples from CT data set fails to capture the values and the trend of the true I_P trace.

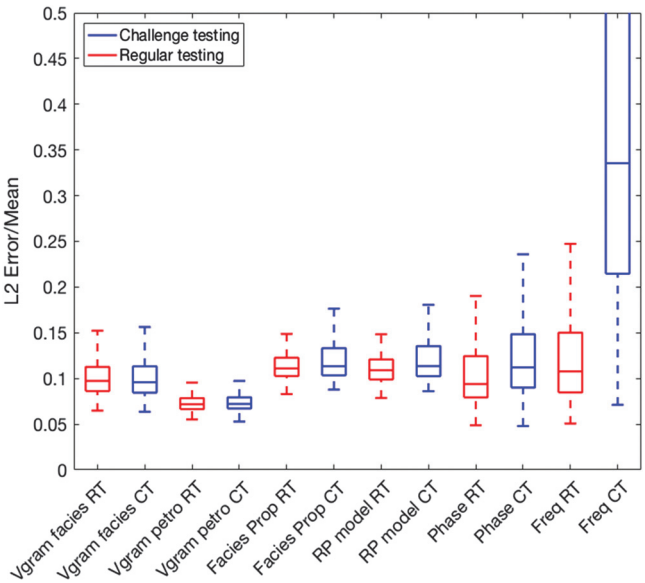

Figure 7. Box plot of the L2-norm of error with respect to mean I_P values. The error using the frequency CT data set is larger than other experiments and is outside of the range plotted.

Robustness of CNN to rock-physics model parameters

Rock-physics modeling is performed to compute the elastic properties from the facies and petrophysical properties. In this step, we made the choice of a rock-physics model and its parameters. To assess the robustness of the CNN, we train the CNN on the impedance generated by one set of rock-physics model parameters and tested the performance on the impedance generated by a different distribution of rock-physics model parameters. The different model parameters are selected by varying the coordination number in the contact-cement model (Avseth et al., 2000; Mavko et al., 2009). Changing the coordination number can be related to diagenetic changes normally happening over a large depth range. Change in sorting in sandstones can also be modeled by changing the coordination number in the constant-cement model. Although the impedance predictions in this work are target-oriented investigations (i.e., having a limited depth range) in which the coordination number can be assumed to be constant, the CT data set is generated with a wide range of coordination numbers for testing purposes and may not be a typically observed geologic scenario. The training and the RT data set are generated by randomly drawing coordination numbers from a uniform distribution between 10 and 12. The CT data set is generated by randomly drawing coordination numbers from a uniform distribution from 6 to 8 and 13 to 14. CNN's predicted I_P on the RT data set is similar to the results obtained on the training data set, indicating the good generalization power of the CNN. The ε_n for the RT data set is approximately 8%. The CNN predictions for the CT data set do not perform as well as the training or RT data set and have higher error values approximately 12%. The spread of the error around the mean in this case is much higher as compared to the RT data set (Figure 7). This indicates the importance of calibrating the parameters in the rock-physics model used for generating the training data set in the success of the I_P inversion using CNN.

Robustness of CNN to source-wavelet parameters

In addition to the above parameters, we also test the robustness of CNN predictions on the source-wavelet parameters — phase and central frequency used in seismic forward modeling. We systematically sample phases of the source wavelet from a uniform distribution between 20° and 70° to generate the training and RT data set. We generate the CT data set by randomly sampling phases from a uniform distribution between 0°–10° and 80°–90 . The ε_n for the RT data set is approximately 11%. The predictions in the CT data set are not as good as in the training or RT data sets. The ε_n for the CT data set is approximately 13%. The spread of the error in the CT data set around the mean is also larger than that in the RT data set (Figure 7). This shows that the CNN's prediction is moderately sensitive to the phase of the source wavelet used in the training process. In general, the training of the CNN with zero-phase seismograms performed better than seismograms generated using minimum or mixed phase wavelets. We also found that the CNN trained with

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

zero-phase seismograms did not give a good prediction when the test input is a minimum or mixed-phase wavelet.

We test the performance of CNN's robustness using different values of the central frequency of the source wavelet. The training and the RT data set are generated using central frequencies of the source wavelet sampled randomly from a uniform distribution between 30 and 50 Hz. The CT data set is generated by performing seismic forward modeling with source wavelets having central frequencies drawn from uniform distribution between 15–25 and 55–65 Hz. Overall, the CNN failed in predicting the I_P values for the CT data set. The ε_n for the RT data set is approximately 13% and higher than the other cases. The ε_n for the CT data set is much higher, i.e., approximately 41%. The spread of the errors in the CT data set around the mean is also very large, with the maximum error being outside the range of the box plot plotted in Figure 7. This gives a clear indication that CNN predictions for data with a different central frequency than the ones used in training the CNN are unreliable.

FIELD EXAMPLE: VOLVE DATA SET

We extend the above described synthetic example to a field data set recently made public by Equinor. The Volve field is located in offshore Norway and is a clastic reservoir (Figure 8). Our aim in this case is to use the network architecture obtained from the previous case to predict the impedance from the post-stack depth migrated (PSDM) seismic data in the time domain available as part of the Volve field data set. We also aimed to match the prediction of the I_P from CNN with the I_P computed from P-wave sonic velocity and density data at a well location. Statistics of the I_P values at the well location differed from the I_P examples in the synthetic data sets described in the previous section. This indicated the possibility of differences in the rock-physics model, facies proportion, and spatial correlation. The statistics and the wavelet extracted from the seismic trace (zero-phase Ricker wavelet with central frequency of 20 Hz) at the well location indicated that the source wavelet properties are similar to the synthetic data sets used in the previous sections. The CT described in the previous section indicated that the weights of the neural network trained in the previous case cannot be used directly if the rock-physics model or the facies proportions are different. We generated 2000 realizations of I_P traces based on the statistics of the I_P log from the well location using the alternate method of data augmentation described in the "Generating a training data set" section. The variogram for the vertical correlation is estimated using the I_P log at the well location and fitted using a spherical variogram. Check-shot data available for the well are used to establish the time-depth relationship and convert the well log from the depth domain to the time domain. We also generated synthetic seismograms using zero-phase Ricker wavelet with a central frequency of 20 Hz. We trained the weights of the neural network (Figure 2) using the generated I_P-synthetic seismogram pairs. We then used the trained neural network for

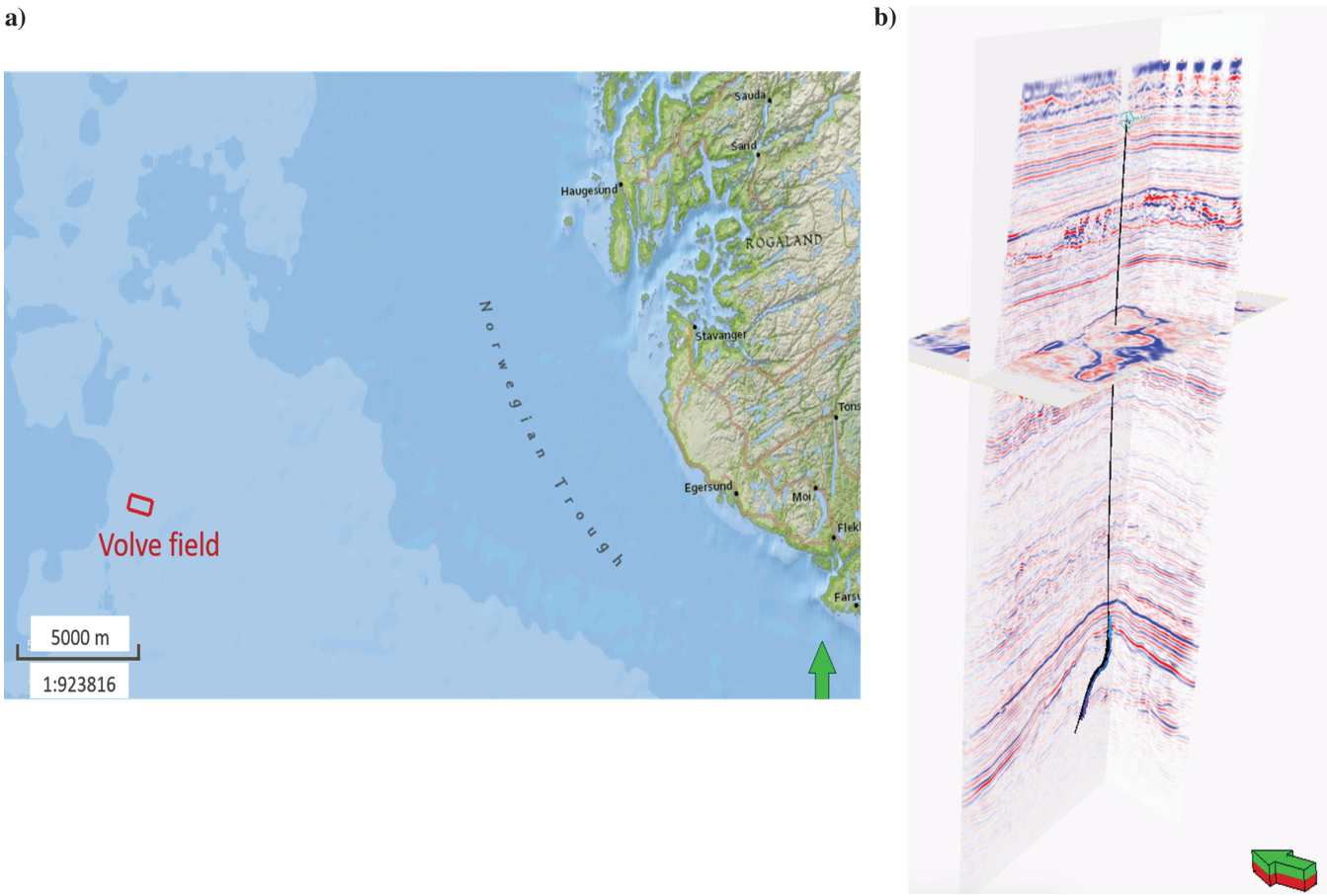

Figure 8. (a) Volve field shown in a map located in the offshore North Sea area and (b) seismic data along with a well trajectory from the Volve field.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

I_P inversion at the well location. Figure 9 shows a comparison between the predicted I_P and the true I_P at the well location. The predicted I_P shows a good match with the true I_P with a correlation coefficient of 82%. We also used the trained CNN to perform a trace-by-trace inversion of 3D PSDM (time domain) seismic cube available as part of Volve field data set in the zone of interest. We assume low impedance values to represent the facies of interest. Hence, as an example of using CNN-derived impedance for reservoir characterization, we extracted a geobody with low values of I_P from the inverted I_P cube obtained using CNN inversion. The validation of the extracted geobody is outside the scope of this paper. A cross-section of seismic data, along with a cross-section of inverted I_P and geobody extracted using low values of I_P are shown in Figure 10.

DISCUSSION

With the tremendous advances being made in machine learning-related research, training a network for different problems is becoming more feasible. Correspondingly, such advances create opportunities for enhancing our capabilities to apply these networks to more challenging geoscientific estimation problems. In this paper, we demonstrate the use of a deep neural network in solving a quintessential geophysical problem of impedance inversion. A limitation in using neural networks for geophysical problems is in the availability of a large training data set of labeled examples. In this paper, we propose a workflow using preexisting geoscientific techniques to generate large amounts of synthetic labeled data set by defining and sampling prior distributions on target variables. Although it might be possible to set up a neural network problem to directly predict reservoir properties from seismic data, the role of geostatistical simulations, rock-physics models, and seismic forward modeling is to augment the limited labeled data set (well logs) and generate a geologically realistic synthetic labeled data set that can be used for training the deep networks. We selected a CNN-based neural network architecture in this work because it forms a natural fit to the deconvolution problem of trace-by-trace seismic impedance inversion. Two-dimensional CNNs can be used to perform a similar workflow with pre-stack seismic gathers as input and elastic properties as output. An advantage of a neural network approach for impedance inversion is that once the weights of the network are fixed in the training process using a training data set (1500 I_P seismic traces for the synthetic example), the prediction of impedance is instantaneous and only requires a single-input seismic data trace.

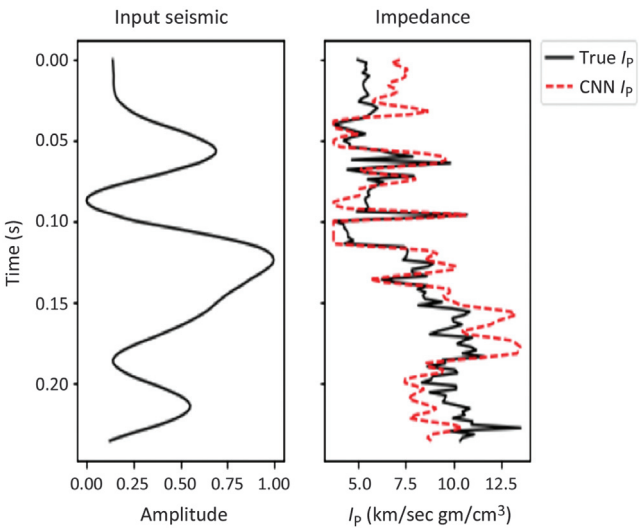

Figure 9. Input seismic data along the well and the corresponding true and CNN impedance. The correlation between the true impedance and the CNN impedance is 82%.

We found that systematically testing the trained network using a challenge data set can provide critical insights into the prediction of the robustness of the network. CT highlighted the importance of using data sets with similar facies proportions, rock-physics models, and source-wavelet parameters for training and application of the network for predictions. The variogram ranges for the facies and petrophysical properties have the least impact on the network's performance. CT showed how extensive tests need to be applied after training and validation to ensure the reliability of the neural network. Such tests can be devised for different geophysical problems using our preexisting knowledge of geology and physics and should be an indispensable part of the machine-learning-based approach for geophysical problems.

Our analysis during hyperparameter tuning in training the network also showed that the filter width of the kernel is comparable to the number of samples in one wavelength of the source wavelet. The CNN predictions also showed promise in predicting the high-frequency components of the I_P traces that are difficult to obtain in traditional deterministic seismic-inversion workflows unless prior geostatistical models are imposed in the inversion process. Our results also show that CNN-based seismic to impedance inversion has good generalization power. The computation time of the network is mainly attributed to the synthetic data set generation and training the network. In this case, we trained the network for 500 epochs with computation time of approximately 45 minutes. The learning curves for the same neural network with different random He initialization of weights showed a different convergence behavior, which is an inherent property of the stochastic gradient-descent-based (Adam) optimization algorithm. These effects, however, can be minimized by using a regularization constraint to the network's learning process, such as batch normalization. The neural network architecture and the hyperparameters are selected such that we obtain fairly consistent convergence behavior with random initializations of the weights. We selected the network with the least training error for all prediction purposes because it had the best training performance. The learning

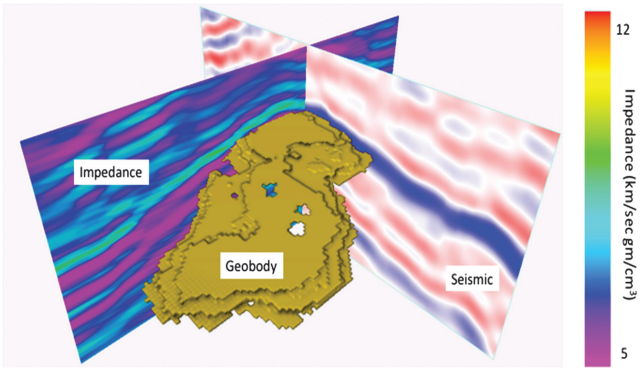

Figure 10. The PSDM (time) seismic data from the Volve field and the CNN impedance inversion are plotted along with a geobody extracted based on the low values of the inverted impedance. The blue color in the seismic section denotes positive amplitude, and the red color denotes negative amplitude.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

curves (Figure 3) plateau after 250 epochs. Hence, an early stopping of the training can be performed at 250 epochs that will reduce the training time for the network by approximately 50% if the computation time scales linearly.

In all our examples, we generated a training data set based on available information from one well location. However, if multiple wells are available for a field, we can either decide to have a single data set for training or multiple data sets based on the stationarity of the well data. A single data set can be used if the distance between the multiple wells is small such that the geology of the area for the multiple wells does not change. However, if the distance is large, we will need multiple data set and multiple CNNs corresponding to the different data sets. The final prediction of I_P can be obtained by a weighted average of proximity of the prediction location to the well log used in generating the data set. The decision of selecting one data set versus multiple data sets can be made using the results of challenge testing. For example, if the seismic source signatures corresponding to the seismic data available in the well locations are different, then the CT results show that multiple training data sets and multiple CNNs should be used. Overall, this study paves the way for a better and faster seismic reservoir characterization using modern machine-learning approaches. The findings from this study can also be used in applying deep networks for more complex geophysical inversion problems, such as inversion for facies models.

CONCLUSION

Various modern machine-learning techniques such as CNNs are being applied in solving different problems in geoscientific applications. Although the neural networks can be trained well on the available data set with sufficient accuracy, the predictions of the trained network on situations not encountered in the training data set may not be robust. Our study shows that in solving the problem of seismic impedance inversion from seismic traces, the underlying source wavelet characteristics, facies proportion, and rock-physics model play an important role in the prediction of the CNN. The CNN can predict meaningful results only when the training and the test data sets have statistically similar facies proportions, rock-physics relations (can be related to the depositional facies type), or source wavelet properties. An ABC-based approach was used in the quantification of the uncertainty in the CNN's prediction. The application of the CNN inversion on a real field data set also showed a good match with well logs. Thus, our study demonstrates that machine-learning-based approaches have great promise in performing better and faster seismic reservoir characterization. However, the true application of machine-learning algorithms to solve geoscientific problems can only be performed by carrying out systematic studies aimed at understanding the machine itself.

ACKNOWLEDGMENTS

This work is supported by the Stanford Rock Physics and Borehole Geophysics (SRB) project, the Stanford Center for Earth Resources Forecasting (SCERF), and Stanford University Geophysics Department. This material is based upon the work supported by the National Science Foundation (NSF) Graduate Research Fellowship Program under grant no. 1656518 and does not necessarily reflect NSF views. We would also like to thank the Dean of the School of Earth, Energy, and Environmental Sciences at Stanford University, Professor S. Graham, for funding, and CGG for the Hampson-Russell software used. We also acknowledge the helpful reviews from all of the reviewers especially from D. Draganov and M. Bosch. We would also like to acknowledge M. Bosch for his collaboration and comments that helped improve the uncertainty quantification part of the work.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

REFERENCES

Araya-Polo, M., J. Jennings, A. Adler, and T. Dahlke, 2018, Deep-learning tomography: The Leading Edge, **37**, 58–66, doi: 10.1190/tle37010058.1.

Avseth, P., J. Dvorkin, G. Mavko, and J. Rykkje, 2000, Rock physics diagnostic of North Sea sands: Link between microstructure and seismic properties: Geophysical Research Letters, **27**, 2761–2764, doi: 10.1029/1999GL008468.

Beaumont, M. A., 2010, Approximate Bayesian computation in evolution and ecology: Annual Review of Ecology, Evolution, and Systematics, **41**, 379–406, doi: 10.1146/annurev-ecolsys-102209-144621.

Beaumont, M. A., W. Zhang, and D. J. Balding, 2002, Approximate Bayesian computation in population genetics: Genetics, **162**, 2025–2035.

Biswas, R., M. K. Sen, V. Das, and T. Mukerji, 2019, Pre-stack and post-stack inversion using a physics-guided convolutional neural network: Interpretation, **7**, no. 3, SE161–SE174, doi: 10.1190/INT-2018-0236.1.

Calderón-Macías, C., M. Sen, and P. Stoffa, 1998, Automatic NMO correction and velocity estimation by a feed forward neural network: Geophysics, **63**, 1696–1707, doi: 10.1190/1.1444465.

Calderón-Macías, C., M. Sen, and P. Stoffa, 2000, Artificial neural networks for parameter estimation in geophysics: Geophysical Prospecting, **48**, 21–47, doi: 10.1046/j.1365-2478.2000.00171.x.

Csilléry, K., M. G. Blum, O. E. Gaggiotti, and O. François, 2010, Approximate Bayesian computation (ABC) in practice: Trends in Ecology and Evolution, **25**, 410–418, doi: 10.1016/j.tree.2010.04.001.

Das, V., A. Pollack, U. Wollner, and T. Mukerji, 2018, Convolutional neural network for seismic impedance inversion: 88th Annual International Meeting, SEG, Expanded Abstracts, 2071–2075, doi: 10.1190/segam2018-2994378.1.

Deutsch, C. V., and A. G. Journel, 1992, GSLIB: Geostatistical software library and user's guide: Oxford University Press.

Fearnhead, P., and D. Prangle, 2012, Constructing summary statistics for approximate Bayesian computation: Semi-automatic approximate Bayesian computation: Journal of the Royal Statistical Society: Series B (Statistical Methodology), **74**, 419–474, doi: 10.1111/j.1467-9868.2011.01010.x.

Goodfellow, I., Y. Bengio, and A. Courville, 2016, Deep learning: MIT Press.

He, K., X. Zhang, S. Ren, and J. Sun, 2015, Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification: Proceedings of the IEEE International Conference on Computer Vision, 1026–1034.

Huang, Z., J. Shimeld, M. Williamson, and J. Katsube, 1996, Permeability prediction with artificial neural network modeling in the Venture gas field, offshore eastern Canada: Geophysics, **61**, 422–436, doi: 10.1190/1.1443970.

Journel, A. G., 1994, Modeling uncertainty: Some conceptual thoughts, Geostatistics for the next century: Springer.

Kennett, B. L. N., 1974, Reflections, rays, and reverberations: Bulletin of the Seismological Society of America, **64**, 1685–1696.

Kennett, B. L. N., 1983, Seismic wave propagation in stratified media: Cambridge University Press.

Kingma, D. P., and J. Ba, 2014, Adam: A method for stochastic optimization: arXiv preprint arXiv:1412.6980.

Krizhevsky, A., I. Sutskever, and G. Hinton, 2012, ImageNet classification with deep convolutional neural networks: Advances in Neural Information Processing Systems, 1097–1105.

Lee, H., R. Grosse, R. Ranganath, and A. Y. Ng, 2009, Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations: Proceedings of the 26th Annual International Conference on Machine Learning, 609–616.

Mavko, G., T. Mukerji, and J. Dvorkin, 2009, The rock physics handbook: Tools for seismic analysis in porous media: Cambridge University Press.

Monajemi, H., D. L. Donoho, and V. Stodden, 2016, Making massive computational experiments painless: 2016 IEEE International Conference on Big Data, 2368–2373.

Moya, A., and K. Irikura, 2010, Inversion of a velocity model using artificial neural networks: Computers and Geosciences, **36**, 1474–1483, doi: 10.1016/j.cageo.2009.08.010.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user

Nair, V., and G. E. Hinton, 2010, Rectified linear units improve restricted Boltzmann machines: Proceedings of the 27th International Conference on Machine Learning, 807–814.
Paszke, A., S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, 2017, Automatic differentiation in PyTorch: NIPS Autodiff Workshop: The Future of Gradient-based Machine Learning Software and Techniques.
Poulton, M. M., 2002, Neural networks as an intelligence amplification tool: A review of applications: Geophysics, **67**, 979–993, doi: 10.1190/1.1484539.
Pradhan, A., and T. Mukerji, 2018, Seismic estimation of reservoir properties with Bayesian evidential analysis: 88th Annual International Meeting, SEG, Expanded Abstracts, 3166–3170, doi: 10.1190/segam2018-2998259.1.
Pritchard, J. K., M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman, 1999, Population growth of human Y chromosomes: A study of Y chromosome microsatellites: Molecular Biology and Evolution, **16**, 1791–1798, doi: 10.1093/oxfordjournals.molbev.a026091.
Remy, N., A. Boucher, and J. Wu, 2009, Applied geostatistics with SGeMS: A user's guide: Cambridge University Press.
Richardson, A., 2018, Seismic full-waveform inversion using deep learning tools and techniques, arXiv:1801.07232.
Rogers, S. J., J. H. Fang, C. L. Karr, and D. A. Stanley, 1992, Determination of lithology from well-logs using a neural network: AAPG Bulletin, **76**, 792–822.
Röth, G., and A. Tarantola, 1994, Neural networks and inversion of seismic data: Journal of Geophysical Research, **99**, 6753–6768, doi: 10.1029/93JB01563.
Rubin, D. B., 1984, Bayesianly justifiable and relevant frequency calculations for the applied statistician: The Annals of Statistics, **12**, 1151–1172, doi: 10.1214/aos/1176346785.
Russell, B., and D. Hampson, 1991, Comparison of poststack seismic inversion methods: 61st Annual International Meeting, SEG, Expanded Abstracts, 876–878, doi: 10.1190/1.1888870.
Sen, M. K., 2006, Seismic inversion: Society of Petroleum Engineers.
Shahraeeni, M. S., A. Curtis, and G. Chao, 2012, Fast probabilistic petrophysical mapping of reservoirs from 3D seismic data: Geophysics, **77**, no. 3, O1–O19, doi: 10.1190/geo2011-0340.1.
Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM.
Wiener, J., J. A. Rogers, J. R. Rogers, and R. Moll, 1991, Predicting carbonate permeabilities from wireline logs using a back-propagation network: 61st Annual International Meeting, SEG, Expanded Abstracts, 285–288, doi: 10.1190/1.1888943.
Wu, Y., and G. A. McMechan, 2018, Feature-capturing full-waveform inversion using a convolutional neural network: 88th Annual International Meeting, SEG, Expanded Abstracts, 2061–2065, doi: 10.1190/segam2018-2963265.1.
Zhang, G., Z. Wang, and Y. Chen, 2018, Deep learning for seismic lithology prediction: Geophysical Journal International, **215**, 1368–1387.

Downloaded from http://pubs.geoscienceworld.org/geophysics/article-pdf/84/6/R869/4934315/geo-2018-0838.1.pdf
by Univ Estadual de Campinas UNICAMP-IG user